

일반논문 (Regular Paper)

방송공학회논문지 제30권 제5호, 2025년 9월 (JBE Vol.30, No.5, September 2025)

<https://doi.org/10.5909/JBE.2025.30.5.772>

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

## 3D Gaussian Splatting의 실시간 렌더링 최적화를 위한 클러스터링 및 중요도 기반 프루닝 기법

조 수 빈<sup>a)</sup>, 박 성 준<sup>a)</sup>, 서 정 일<sup>a)†</sup>

### Efficient Real-Time Rendering of 3D Gaussian Splatting via Clustering and Pruning by Importance

Subin Cho<sup>a)</sup>, Seongjun Park<sup>a)</sup>, and Jeongil Seo<sup>a)†</sup>

#### 요 약

본 논문은 3D Gaussian Splatting의 실시간 렌더링 효율을 개선하기 위해, K-Means 클러스터링과 시각적 기여도에 따른 중요도 기반 프루닝(PPI)을 결합한 프레임워크를 제안한다. 본 기법은 각 클러스터 내 가우시안의 특성을 통합하여 시각적 중요도를 정량화하고, 중요도가 낮은 가우시안을 제거함으로써 연산량을 줄이고 메모리 효율을 개선한다. 실험 결과, 전체 가우시안 수를 약 43% 감소한 상태에서도 PSNR 9.01%, SSIM 3.12%, LPIPS 42.4%의 품질 향상에 도달하였다. 또한, FPS는 기존 대비 소폭 향상되었으며, GPU 메모리 사용량도 감소하여 실시간 렌더링 환경에서의 실용 가능성을 확인하였다. 본 연구는 별도의 신경망 기반 구조 없이 경량화와 품질 향상을 동시에 달성할 수 있는 효율적인 3DGS 구조를 제시함으로써, 다양한 응용 환경에 활용될 수 있는 기술적 기반을 제공한다.

#### Abstract

Real-time 3D Gaussian Splatting (3DGS) often requires a large number of Gaussians to maintain visual quality, resulting in significant computational overhead and memory consumption. To address this limitation, we propose a novel and lightweight framework that combines K-Means clustering with point-wise pruning importance (PPI). Our method first clusters Gaussians based on spatial and appearance features, then evaluates the perceptual importance of each Gaussian within its cluster. Gaussians with minimal visual contribution are selectively pruned, effectively reducing redundancy while preserving rendering quality. Unlike prior approaches that rely on neural network architectures, our framework employs purely analytical techniques, achieving efficiency without compromising simplicity. Experimental results show an average reduction of 43% in the number of Gaussians, along with perceptual quality improvements of 9.01% and 3.12%, and a 42.4% decrease in error-based metrics. Furthermore, our method yields modest gains in frame rate and substantial reductions in GPU memory usage, confirming its suitability for real-time applications. This work presents a scalable, high-quality rendering pipeline for 3DGS, with practical implications for virtual reality, gaming, and interactive visualization.

Keyword : 3D Gaussian Splatting, Pruning, Clustering, 3D Reconstruction, Real-Time

Copyright © 2025 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

3차원 렌더링 기술은 방송, 미디어, 가상현실(Virtual Reality, VR), 증강현실(Augmented Reality, AR), 메타버스 등 다양한 분야에서 몰입형 콘텐츠 제작의 핵심적인 역할을 담당하고 있다. 특히 Novel View Synthesis(NVS)는 제한된 시점 데이터를 활용해 기존에 없는 새로운 관점에서 고품질 3D 장면을 생성하는 기술이다<sup>[1]</sup>. 최근 각광받는 Neural Radiance Field(NeRF)는 신경망을 기반으로 정밀한 3D 장면을 재구성할 수 있지만, 연산 비용이 높아 방송이나 모바일처럼 연산 자원이 부족한 환경에서는 적용이 어렵다는 한계가 있다<sup>[2]</sup>.

이를 해결하기 위해 개발된 3D Gaussian Splatting(3DGS)은 3D 장면의 구성 요소를 가우시안 프리미티브로 표현하고, 미분 가능한 Splatting 기법을 적용하여 NeRF보다 빠른 렌더링 속도와 우수한 시각적 품질을 제공한다<sup>[3]</sup>. 3DGS는 실시간 3D 콘텐츠 렌더링, VR/AR 콘텐츠 제작, 360도 영상 스트리밍 등에 적합한 기술로 평가된다. 그러나 대규모 가우시안을 처리해야 하므로 메모리 사용량과 연산 복잡도가 높아지며, 특히 고해상도 장면은 수십만 개 이상의 가우시안을 요구해 렌더링 속도와 자원 효율성을 저하시킨다. 또한, 공간적으로 유사한 가우시안들이 서로 매우 유사한 정보를 가지며, 이는 높은 연산 복잡도와 불필요한 메모리 소비로 이어지는 한계점이 있다. 본 논문에서는 이러한 문제를 극복하기 위해, 공간적 클러스터링 및 중요도 기반 프루닝을 적용하여 불필요한 가우시안을 효과적으로 제거하는 프레임워크를 제안한다.

기존 연구에서는 가우시안의 수를 줄이는 프루닝(prun-

ing) 기법이나 속성을 압축하는 양자화(quantization) 방법을 제안했지만, 시각적 품질 저하와 제한된 압축률 문제가 있었다<sup>[4][5][6][7]</sup>. 또한, 신경망 기반의 최적화를 통해 가우시안 속성을 예측하거나, anchor를 활용한 압축 방식은 3DGS의 포인트 기반 구조와 호환되지 않거나 복잡한 학습 과정을 요구했다<sup>[8]</sup>. 이에 따라 3DGS의 기존 구조를 유지하면서 메모리와 계산 비용을 줄이고, 렌더링 품질을 보존하는 방법은 여전히 해결해야 할 과제로 남아 있다.

본 논문은 K-Means 클러스터링과 중요도 기반 프루닝(Point-wise Pruning Importance, PPI)을 결합한 새로운 프레임워크를 제안하여 3DGS의 효율성을 개선하고자 한다. 이 방법은 가우시안을 공간적으로 그룹화하고, 각 클러스터 내에서 위치, 스케일, 색상, 투명도 등의 속성을 바탕으로 중요도를 평가해 낮은 중요도의 가우시안을 제거한다. 이를 통해 가우시안 수를 줄이고, 계산 복잡도와 메모리 사용량을 최소화하며, 렌더링 품질 저하를 방지하기 위해 프루닝 후 속성을 재조정한다.

제안된 프레임워크는 세 가지 핵심 전략을 포함한다. 첫째, 공간적 클러스터링으로 가우시안 분포를 분석하여 유사한 특성을 가진 가우시안을 그룹화한다. 둘째, 중요도 기반 프루닝으로 각 가우시안의 렌더링 기여도를 평가해 시각적 품질에 미치는 영향이 적은 가우시안을 제거한다. 셋째, 프루닝 후 최적화를 통해 남은 가우시안의 속성을 조정하여 렌더링 품질을 유지한다. 이 접근법은 3DGS의 기존 구조를 유지하면서 실시간 렌더링 성능을 개선하며, 방송, VR/AR, 360도 영상, 실시간 스트리밍 등 다양한 미디어 플랫폼에서 3DGS의 활용 가능성을 높인다.

## II. 관련 연구

3DGS는 비등방성 가우시안을 사용해 3D 장면을 표현하는 기술로, 고해상도의 렌더링과 효율적인 시각화를 동시에 제공한다. 3DGS는 NeRF의 높은 계산 부담을 크게 줄이면서, NVS를 통해 다양한 미디어 활용 분야에서 실시간 렌더링을 가능하게 한다. 하지만 고해상도 장면을 정밀하게 표현하기 위해 수십만 개 이상의 가우시안 요소가 요구되며, 이는 메모리 사용량과 연산 비용의 증가를 초래한다.

a) 동아대학교 컴퓨터공학과(Dept. of Computer Engineering, Dong-A University)

‡ Corresponding Author : 서정일(Jeongil Seo)

E-mail: jeongilseo@dau.ac.kr

Tel: +82-51-200-7796

ORCID: <https://orcid.org/0000-0001-5131-0939>

※ This work was partly supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(No.2023-0-00076, National Program of Excellence in Software(Dong-A University)) and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(RS-2023-00273349\_3)

· Manuscript August 13, 2025; Revised September 12, 2025; Accepted September 12, 2025.

이러한 자원 소모 문제를 해결하기 위해 여러 경량화 전략들이 제안되었으며, 본 장에서는 기존 주요 경량화 방법과 함께 프루닝 및 적응형 밀도 제어에 기반한 최신 경량화 기법에 대해 설명한다.

## 1. Type Casting

초기 연구에서 가우시안의 핵심 속성인 위치, 색상, 스케일 등을 32비트 부동소수점(FP32) 대신 8비트 정수(INT8)로 변환하거나 엔트로피 부호화를 적용해 저장 효율을 높이는 방법이 주로 활용되었다<sup>[9]</sup>. 이러한 접근은 구현이 간단하지만, 압축률의 한계와 정밀도 손실 문제를 동반한다. 이후 유사한 가우시안들을 클러스터링한 후 각 클러스터의 속성을 단일 Multi-Layer Perceptron(MLP)으로 압축하는 신경망 기반 방식이 제안되었다<sup>[10]</sup>. 이 방식은 높은 압축률을 제공하지만, MLP 연산의 복잡성과 클러스터 간 속성 변동성이 큰 경우 학습이 어려운 단점이 있다.

## 2. 근접한 가우시안 속성 유사성 분석 기법

최근 연구에 따르면, 3DGS에서 공간적으로 가까운 가우시안들이 서로 유사한 속성을 가지는 것으로 나타났으며, 이는 중복된 정보를 가진 가우시안들로 인해 불필요한 연산 복잡도를 초래한다. 가우시안 간 거리와 속성 차이를 분석한 결과, 거리가 가까울수록 속성 유사도가 높은 밀도 분포 경향이 확인되었다. 이는 프루닝과 클러스터링 기법을 통해 중복된 시

각 정보를 효과적으로 제거할 수 있다는 가능성을 보여준다.

Shin은 이를 정량적으로 검증하기 위해, 가우시안 쌍의 유클리드 거리를 기준으로 속성 간 차이를 계산하였다<sup>[11]</sup>. 거리  $d$ 는  $\log_2(d) = -12$ 에서  $-8$ 까지 다섯 구간으로 나누어 속성 유사도 분포를 그림 1에 시각화하였다.

여기서 색상은 공간적 거리를 나타내고,  $x$  축은 속성의 유클리드 거리,  $y$  축은 밀도를 의미한다. 속성의 유클리드 거리는 값이 낮을수록 가우시안 쌍의 정보 유사성이 높다는 것을 나타내며, 예를 들어 그림 1의 A. Opacity 그래프에서는  $x = 0$  근처에서 밀도가 높아 공간적으로 가까운 가우시안의 Opacity값이 거의 동일함을 보여준다. 분석 결과, 거리 간격이 좁을수록 모든 속성에서 밀도가 좌측으로 집중되었으며, 이는 시각적으로 영향을 적은 중복 가우시안들이 다수 존재한다는 것을 시사한다.

이에 따라 프루닝 기반 접근법이 시각적 기여도가 낮은 중복 가우시안을 제거하여 연산량을 줄이는 방법으로 주목받고 있다<sup>[12]</sup>. 이는 구조적 단순성과 실시간 성능 개선 측면에서 강점을 가진다. 본 논문에서는 K-Means 클러스터링을 활용해 속성 유사성을 기반으로 가우시안을 공간적으로 그룹화하고, 각 가우시안의 중요도를 정량적으로 평가하여 선택적으로 제거하는 PPI 방식을 도입함으로써 기존 방식과 차별화된 접근을 시도한다.

## 3. 프루닝 기반 경량화 기법

프루닝은 모델의 복잡성을 낮추고 연산 효율성을 높이기

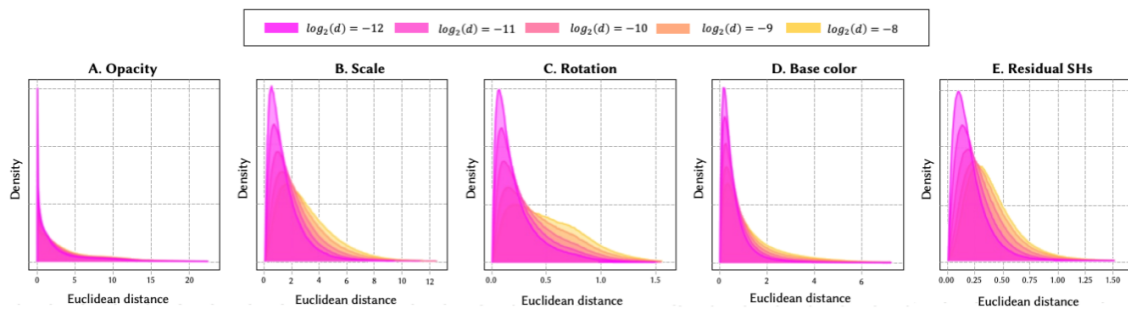


그림 1. 공간적 근접성에 따른 가우시안 속성 유사성 분석.  $\log_2(d) \in [-12, -8]$  범위의 거리 구간으로 나누어, 각 거리 구간 내에서 Gaussian 속성들 간의 유클리드 거리 분포를 시각화한 그림이다.

Fig. 1. Analysis of Gaussian property similarities based on spatial proximity. The range of  $\log_2(d) \in [-12, -8]$  is divided into intervals, and for each interval, the Euclidean distance distributions between Gaussian properties are visualized.

위한 경량화 기법으로, 딥러닝 신경망에서 중요도가 낮은 뉴런이나 가중치를 제거하는 방식으로 연구되어 왔다<sup>[13]</sup>. 이 기법은 연산량 감소, 메모리 절약, 추론 속도 향상 등의 이점을 제공하며, 최근에는 3차원 비전 분야로 확장 적용되고 있다. 3차원 비전에서는 포인트 클라우드나 볼륨 렌더링과 같은 데이터에서 공간적 중복 요소나 시야 밖의 요소를 제거하여 불필요한 점을 줄이는 방식이 주로 사용된다. 3DGS에서도 유사하게, 시각적 기여도가 낮은 가우시안을 제거하여 렌더링 효율성을 개선하려는 시도가 이어져 왔다. 기존 연구에서는 주로 가우시안의 가시성, 공간 밀도, 색상 대비 등을 기준으로 한 간단한 임계값 기반 프루닝을 적용하였다<sup>[14]</sup>. 그러나 이러한 방식은 프루닝으로 인한 시각적 품질 저하를 효과적으로 방지하지 못한다는 공통된 한계를 지닌다. 특히, 3DGS와 같이 직접적인 이미지 생성에 영향을 미치는 기술에서는 보다 정교한 프루닝 기법이 요구된다.

본 논문에서는 각 가우시안의 시각적 기여도를 정량적으로 평가해 프루닝 대상을 선정하는 PPI 기법을 제안한다. 이 방법은 위치, 스케일, 투명도 등 시각적 속성을 종합적으로 분석하여, 각 가우시안이 렌더링 품질에 미치는 영향을 계산한다. 이를 통해 단순한 임계값 기반이 아닌 실제 시각적 기여도에 따라 제거 대상을 선택함으로써 성능과 품질을 모두 개선하는 경량화를 가능하게 한다.

#### 4. 적응형 밀도 제어 기반 경량화 기법

최근 연구에서는 장면 특성에 따라 유연하게 가우시안 밀도를 조절하는 적응형 밀도 제어(Adaptive Density Control, ADC) 기법이 연구되었다. ADC는 각 가우시안의 시각적 기여도, 장면 내 위치, 복잡도 등을 고려하여 프루닝 비율을 조정함으로써, 높은 고품질 렌더링을 유지하면서도 연산 자원을 효율적으로 활용할 수 있게 한다.

Bulò는 픽셀 단위의 시각적 오차를 기반으로 각 가우시안의 누적 기여도를 정량화하여 밀도 조절의 기준으로 사용하는 방법을 제안하였다<sup>[15]</sup>. 구체적으로, 각 가우시안  $g_k$ 에 대해 다음과 같은 오차 기반 중요도 함수  $E_k$ 를 정의하였다:

$$E_k = \sum_{i \in \Omega_k} \alpha_{ik} \cdot e_i \quad (1)$$

식 (1)에서  $\Omega_k$ 는 가우시안  $g_k$ 가 영향을 미치는 픽셀 집합,  $\alpha_{ik}$ 는 픽셀  $i$ 에서 해당 가우시안의 알파 블렌딩 가중치, 그리고  $e_i$ 는 해당 픽셀의 렌더링 오차를 나타낸다. 이 값이 임계값  $\tau$ 를 초과할 경우, 해당 가우시안은 복제되거나 밀도 증가의 대상이 되며, 이는 텍스처 보존이 중요한 고주파 디테일 영역에서 렌더링 품질을 효과적으로 유지하는 데 기여한다. 고주파 영역은 물체 경계나 질감의 미세한 변화를 포함하고 있으며, 이는 저주파 영역보다 인간 시각에 더 큰 영향을 미친다. 따라서 ADC 기법은 저주파 영역에서 더 큰 프루닝을 적용하면서 고주파 영역에서는 밀도를 유지하거나 높여 디테일 손실을 최소화하며, 전체 렌더링 품질 저하 없이 연산 효율성을 높이는 데 초점을 맞춘다.

Grubert는 기존 ADC 방식의 훈련 안정성 문제를 개선하기 위해 학습 시간에 따라 임계값을 점진적으로 증가시키는 조절 방식을 제안하였다<sup>[16]</sup>. 이 방식에서 임계값은 다음과 같이 조정된다:

$$\tau(t) = \tau_0 \cdot \exp(\lambda t) \quad (2)$$

식 (2)에서  $\tau_0$ 은 초기 임계값,  $\lambda$ 는 상승률 계수, 그리고  $t$ 는 학습 반복(iteration) 수를 나타낸다. 이를 통해 학습 초기에는 더 많은 가우시안 생성을 유도하고, 학습 후반에는 불필요한 밀도 증가를 억제하여 자원 낭비를 줄인다. 또한, 가우시안의 시각적 기여도와 밀도를 통합적으로 고려한 중요도 기반 프루닝 기법을 적용해 단순 임계값 필터링보다 시각적 품질을 효과적으로 보존하였다.

이러한 적응형 밀도 제어 기법은 고정된 프루닝 비율로 처리하기 어려운 장면 복잡도나 시점 변화에 따른 장면 변화를 유연하게 대응할 수 있어, 실시간 3D 렌더링 응용에서 매우 유용하다고 평가된다. 본 논문에서는 정적 프루닝 비율에 기반한 클러스터링 및 PPI 기반 프루닝 기법을 적용했다.

### III. 제안 방법

본 논문은 3DGS의 메모리 사용량과 계산 복잡성을 줄이기 위해 K-Means 클러스터링과 PPI 기법을 통합한 새로운 프레임워크를 제안한다. 이 방법은 가우시안의 공간적 분

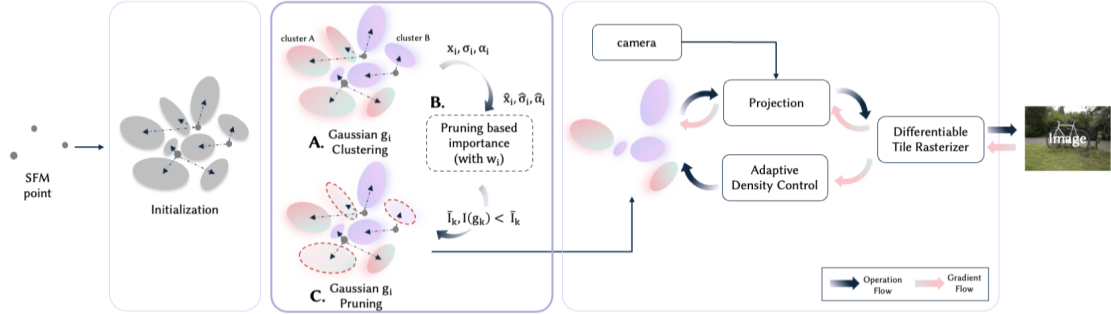


그림 2. 제안한 방법론의 전체 파이프라인. SfM 포인트로부터 초기 Gaussian을 생성한 후, 공간적 클러스터링(A), 중요도 계산(B), 프루닝 후 최적화(C)를 통해 재구성한다. 이후 Camera Projection 및 Differentiable Tile Rasterizer를 통해 최종 이미지를 생성하며, 전체 과정은 end-to-end 학습이 가능하다.

Fig. 2. Full pipeline of the proposed method. Starting from SfM points, initial Gaussians are generated and then refined through the Spatial Clustering(A), Importance estimation(B), and Pruning with optimization(C). The final image is rendered via Camera Projection and a Differentiable Tile Rasterizer, with the entire process being end-to-end trainable.

포를 분석하고 시각적 품질에 미치는 영향을 정량적으로 평가해 선택적으로 제거함으로써, 미디어 응용에 적합한 경량화된 3DGS를 구현한다. 그림 2는 제안된 방법의 전체 아키텍처를 보여준다. 본 장에서는 제안된 방법의 핵심 구성 요소인 (A)공간적 클러스터링, (B)중요도 계산, (C)프루닝 기법 및 최적화로 나누어 자세히 설명한다.

## 1. 공간적 클러스터링

3DGS는 장면을 수십만 개의 비등방성 가우시안 프리미티브로 표현하며, 각 프리미티브는 위치( $x$ ), 스케일( $s$ ), 회전( $r$ ), 색상( $c$ ), 투명도( $\alpha$ ) 등의 속성을 포함한다. 대규모 가우시안을 개별적으로 처리하면 연산 비용이 증가하므로, K-Means 클러스터링을 통해 가우시안의 공간적 분포를 그룹화한다. K-Means 클러스터링은 주어진 데이터를 중심점을 기준으로 가까운 데이터들을  $K$ 개의 그룹으로 묶는 기법이다<sup>[17]</sup>. 클러스터링은 다음과 같은 단계로 진행된다.

- 1) 각 가우시안의 위치  $x_i \in \mathbb{R}^3$ 를 기반으로 K-Means 알고리즘을 적용한다. 클러스터 수  $K$ 는 전체 가우시안 수  $N$ 에 비례하여 설정되며, 본 연구에서는  $K = N/100$ 으로 계산된다.
- 2) 각 클러스터  $k$ 의 중심  $c_k$ 와 해당 클러스터에 속한 가

우시안 집합  $G_k$ 를 구한다.

- 3) 클러스터링 결과는 인덱스 집합  $\{I_k\}_{k=1}^K$  형태로 저장되어 후속 중요도 평가 및 프루닝에 사용된다.

대규모 장면에서 클러스터 수  $K$ 를 과도하게 작게 설정하면 서로 다른 구조가 동일 군집 내에 함께 포함된다. 이 경우, 경계나 얇은 텍스처와 같은 세부 구조는 넓은 평탄 영역과 같은 집단에서 경쟁하게 되어 중요도가 과소평가되기 쉽다. 반대로  $K$ 를 크게 두면 세부 구조가 독립적으로 평가되어, 프루닝 시 보존될 가능성이 높아진다. 본 클러스터링 파이프라인에서는  $K = \lceil N/r \rceil$ 로 두고  $r \in \{25, 50, 100, 200, 400\}$ 을 비교 실험을 진행한 결과,  $r = 100$ 에서 PSNR과 SSIM이 최대, LPIPS는 최소로 세 지표 모두에서 우수한 성능을 나타냈다. 이 실험 결과는 표 1에 나타나 있다.

$r = 25$ 에서는 과도하게 큰 클러스터 수인  $K = 2170$ 로 인해 PSNR과 SSIM이 감소하였으며, 이는 세부 구조의 중요도 추정이 불안정해지고 경계 영역에서 과도한 프루닝이 발생할 위험이 있음을 시사한다. 반대로  $r = 400$ 에서는 클러스터 수인  $K = 134$ 가 지나치게 작아서 품질 성능이 떨어졌음을 확인할 수 있다. 결과적으로, 품질 지표는  $r$ 이 너무 크거나 작은 극단적 설정에서는 저하되는 반면, 중간 영역에서는 최적 성능을 달성하는 경향을 나타냈으며, 본 연구에서는 최적값인  $r = 100$ 으로 설정하여 실험을 진행하였다.

표 1.  $r \in \{25, 50, 100, 200, 400\}$ 에 대한 비교 결과.  $r = 100$ 가 세 지표 모두에서 최고 성능을 보였다.  $r = 100$ 은  $r = 50$  대비  $+0.1$  PSNR,  $+0.0028$  SSIM,  $-0.005$  LPIPS의 유의미한 개선을 보였으며, 메모리 자원 및 효율 지표는  $r$  변화에 미세한 차이로 민감하지 않았다.

Table 1. For  $r \in \{25, 50, 100, 200, 400\}$ ,  $r = 100$  achieved the best performance across all three metrics. Compared with  $r = 50$ ,  $r = 100$  showed significant gains:  $+0.1$  dB in PSNR,  $+0.0028$  in SSIM, and  $-0.005$  in LPIPS. Memory usage and efficiency metrics were largely insensitive to changes in  $r$ , showing only minor differences.

$r$ ( $K = \lceil N/r \rceil$ )	PSNR( $\uparrow$ )	SSIM( $\uparrow$ )	LPIPS( $\downarrow$ )	$K$ (Cluster count)
25	21.57	0.6841	0.1856	2170
50	21.48	0.6814	0.1894	1084
100	21.58	0.6842	0.1844	541
200	21.51	0.6812	0.1894	270
400	21.47	0.6802	0.1900	134

## 2. 중요도 계산

프루닝의 핵심 목표는 시각적 품질 손실을 최소화하면서 가우시안 수를 줄이는 것이다. 이를 위해 본 논문은 각 가우시안의 시각적 기여도를 정량화하는 중요도 지표를 정의한다. 이 중요도는 위치, 스케일, 색상, 투명도 네 가지 시각적 특성을 기반으로 계산된다. 각 특성은 Min-max 정규화를 통해 범위를 조정한 후 가중치 파라미터를 적용하여 통합된다. 가우시안  $i$ 에 대한 중요도  $I_i$ 는 식 (3)과 같이 계산된다:

$$I_i = w_{pos} \cdot \|\hat{p}_i\|_2 + w_{scale} \cdot \|\hat{s}_i\|_2 + w_{color} \cdot \|\hat{c}_i\|_2 + w_{opacity} \cdot \hat{o}_i \quad (3)$$

$\hat{p}_i$ ,  $\hat{s}_i$ ,  $\hat{c}_i$ ,  $\hat{o}_i$ 는 각각 위치, 스케일, 색상, 투명도의 정규화된 값을 나타내고, 각 특성의 상대적 중요도는 가중치  $w_{pos}$ ,  $w_{scale}$ ,  $w_{color}$ ,  $w_{opacity}$ 로 조정된다. 정규화는 다음과 같은 Min-max 방식으로 수행된다:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min} + \varepsilon} \quad (4)$$

식 (4)에서  $\varepsilon$ 은 분모가 0이 되는 것을 방지하기 위한 작은 값이다. 각 벡터 항목에는 L2 norm을 적용하여 특성의

시각적 영향력을 수치화한다. 예를 들어, 스케일 벡터의 L2 norm은 가우시안의 공간적 크기를 의미하고, 색상 벡터의 L2 norm은 밝기나 대비를 반영한다. 이 중요도 함수는 투명도가 낮고, 색상 대비가 크며, 화면 중심과의 거리가 가깝고, 스케일이 큰 가우시안이 최종 이미지에 더 큰 시각적 영향을 미친다는 전제를 기반으로 설계되었다. 이는 사람이 불투명하고 크기가 크며 주변과 대비되는 객체에 더 많은 주의를 기울인다는 시각 연구 결과를 반영한 것이다. 또한 위치 특성의 중요도는 사람의 시선이 보통 화면의 중심이나 관심 영역에 더 잘 머문다는 현상을 따른다. 따라서 이러한 시각적 주목성을 고려해 각 특성을 정규화하고 중요도를 계산하였다.

## 3. 프루닝 및 최적화 기법

### 3.1 Importance-based Pruning

Gaussian Splatting의 렌더링 연산 효율성을 높이기 위해, 계산된 중요도  $I_i$ 를 기준으로 중요도가 낮은 가우시안들을 제거하는 프루닝 과정을 진행한다. 프루닝 과정은 다음과 같다. 먼저 각 클러스터  $G_k$  내의 가우시안들을 중요도  $I_i$ 를 기준으로 정렬한다. 그 후, 프루닝 비율  $p \in [0, 1]$ 을 설정하여 클러스터  $G_k$ 에 속한 가우시안 수  $|G_k|$  중 하위  $(1-p) \cdot |G_k|$ 개의 가우시안을 제거한다. 본 연구에서는  $p = 0.7$ 의 비율을 사용하며, 이는 장면 복잡도나 렌더링 성능 요구에 따라 조정할 수 있는 파라미터이다. 이 비율을 통해 하위 30%의 상대적으로 중요도가 낮은 가우시안을 제거하면서도 의미 있는 가우시안의 손실을 최소화한다. 제거 비율이 40% 이상으로 증가할 경우 시각적으로 중요한 가우시안까지 제거되어 품질 저하가 발생할 우려가 있기 때문에  $p = 0.7$ 은 품질 균형을 고려한 최적의 선택으로 볼 수 있다. 줄제거할 가우시안을 선정할 때, 클러스터 평균 중요도  $\bar{I}_k$ 보다 낮은 가우시안  $g_k$ 는  $I_k(g_k) < \bar{I}_k$  조건을 만족하면 제거 대상으로 선정된다. 이 기준은 시각적 품질 손실을 최소화하면서 연산량을 효과적으로 줄인다. 제거된 가우시안의 인덱스를 마스크  $m \in \{0, 1\}^N$  형태로 저장되며, 남은 가우시안 집합은  $G' = \{g_i \mid m_i = 1\}$ 로 정의된다.

이 접근법은 3DGS의 포인트 기반 구조를 유지하면서 성

능을 개선하며, 추가적인 학습 네트워크나 복잡한 전처리 과정을 요구하지 않는다. 프루닝은 연산 속도를 크게 줄이며, 실시간 렌더링에서도 *PSNR*, *SSIM*, *LPIPS* 기준으로 품질과 성능을 향상시킨다.

### 3.2 Post-pruning Optimization

프루닝 후 남은 가우시안의 시각적 품질 손실을 최소화하기 위해, 가우시안 속성을 재조정하여 식 (5)를 통해 최적화를 수행한다. 이를 통해, 프루닝 후에도 시각적 품질을 유지할 수 있도록 하였다:

$$L = \lambda_1 L_{L1} + \lambda_2 (1 - L_{SSIM}) \quad (5)$$

이때,  $\lambda_1 = 0.7$ ,  $\lambda_2 = 0.3$ 으로 설정하였으며,  $L_{L1}$ 에 높은 가중치를 부여하여 최적화가 실제 데이터(Ground Truth, GT)에 근접하도록 유도하며,  $(1 - L_{SSIM})$ 에 낮은 가중치를 적용해 세부 디테일을 보존하였다. 프루닝 후 변경된 가우시안 분포를 반영하여 K-Means 클러스터링을 다시 수행하고, 클러스터 중심  $c_k$ 를 업데이트한다. 최적화된 가우시안 속성은 3DGS 렌더링 파이프라인에 통합되어, 기존 대비 적은 연산량으로 유사한 시각 품질을 유지한다. 최적화 과정은 과적합을 방지하기 위해 최대 1,000회 epoch으로 제한된다. 추가로,  $\lambda_2$  값을 0.1, 0.2, 0.3으로 변화시키며 성능을 비교한 결과,  $\lambda_2 = 0.3$ 일 때, *PSNR* 21.64, *SSIM* 0.6937, *LPIPS* 0.1571로 가장 좋은 성능을 나타냈다. 특히 *LPIPS*에서 뚜렷한 향상이 확인되어 시각적 품질 보존의 개선이 극대화되었다. 다만  $\lambda_2$  값을 0.4 이상으로 높이면 메모리 사용량이 증가하여 실험이 불가능했으므로, 자원 제약을 고려했을 때  $\lambda_2 = 0.3$ 이 최적값 설정으로 판단된다.

## IV. 실험

### 1. 실험 환경

본 장에서는 가우시안을 공간적 근접성을 기준으로 클러스터링한 후 중요도 기반 프루닝 적용한 실험을 통해 제안

모델과 기존 3DGS의 정확도 및 메모리 사용량, FPS 성능을 비교한다. 실험에는 Mip-NeRF 360 데이터셋의 Bicycle 장면이 사용되었으며, 이 데이터셋은 다양한 시점에서 촬영된 고해상도 RGB 이미지와 COLMAP 기반 카메라 포즈 정보를 포함한다<sup>[18]</sup>. 실험을 위해 원본 고해상도 이미지를 1080p 해상도로 수동 다운샘플링하였다. 학습에는 약 100개의 뷰를 사용하였으며, 테스트는 별도로 분리된 20개의 카메라 시점에서 진행되었다.

### 2. 구현 세부사항

본 논문에서 3DGS의 학습 파이프라인에 K-Means 클러스터링과 중요도 기반 프루닝을 통합하여 연산량을 줄이고 실시간 성능을 확보하였다. 클러스터링은 초기 가우시안의 위치 정보를 기반으로 수행되며, 클러스터 수  $K$ 는 전체 가우시안 수  $N$ 에 대해  $K = N/100$ 으로 설정하였다. 각 클러스터 내 가우시안의 공간 밀도는 이후 중요도 계산 및 프루닝에 활용된다.

프루닝은 투명도, 스케일, 위치 등의 속성을 종합하여 각 가우시안의 중요도를 평가한 후, 하위 30%의 중요도를 가진 가우시안을 제거하는 방식으로 진행된다. 이후 제거된 가우시안을 제외하고 최적화 파라미터를 재설정하여 효율적인 학습이 가능하게 하였다. 클러스터링 및 프루닝 효과를 정량적으로 분석하기 위해, 프루닝 전후의 클러스터별 가우시안 수 분포를 그림 3에 시각화하였다. 그림 3.a와 3.b는 각각 프루닝 전후의 클러스터별 가우시안 수를 막대그래프로 나타낸다. 가로축은 클러스터 ID, 세로축은 각 클러스터에 포함된 가우시안의 수를 나타낸다. 그림 3.a에서 볼 수 있듯이, 프루닝 전에는 클러스터 간 가우시안 수가 불균일하게 분포되어 있으며, 일부 클러스터에는 약 1,600개의 가우시안이 집중되는 등 밀도 집중 현상이 관찰되었다. 이는 특정 장면 영역에서 정보가 과도하게 표현되었음을 나타내며, 계산 자원의 비효율적 사용으로 이어질 수 있다.

그림 3.b는 프루닝 이후의 가우시안 분포를 보여주며, 전체 가우시안 수가 감소하고 특히 고밀도 클러스터에서 가우시안 수가 크게 줄어든 것을 확인할 수 있다. 중요도가 낮은 가우시안이 제거되면서 정보가 과도하게 집중된 클러스터에서 효과적인 경량화가 이루어졌으며, 시각적 기여도



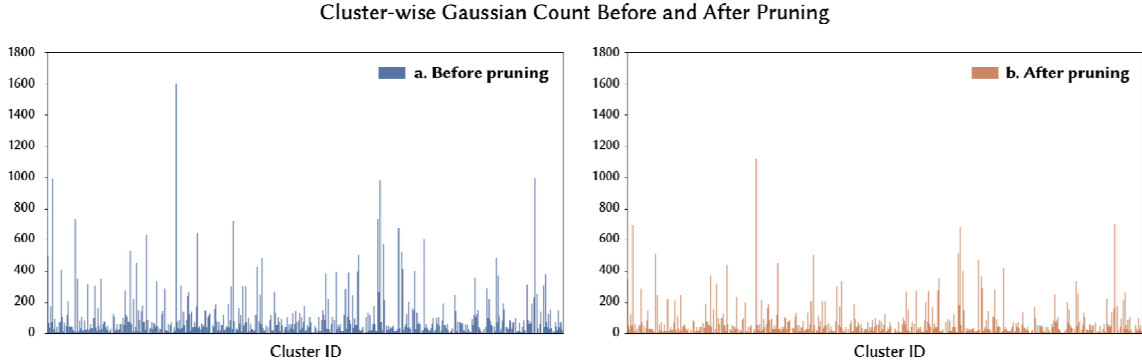


그림 3. 클러스터별 가우시안 수 분포 비교. (a)는 프루닝 전 각 클러스터에 포함된 가우시안 수를 나타내며, 일부 클러스터에 정보가 집중된 현상을 확인할 수 있다. (b)는 중요도 기반 프루닝 이후 분포로, 전체적으로 가우시안 수가 감소하고 클러스터 간 분포가 더욱 균일해진 형태를 보인다.

Fig. 3. Comparison of the number of Gaussian per cluster. (a) shows the number of Gaussian in each cluster before pruning, where some clusters exhibit a concentration of information. (b) presents the distribution after importance-based pruning, showing an overall reduction in the number of Gaussian and a more uniform distribution across clusters.

가 낮은 클러스터에서는 대부분의 가우시안이 제거되었다. 그 결과, 클러스터 간 밀도 균형이 개선되어 보다 균일한 가우시안 분포를 얻을 수 있었다.

이러한 결과는 제안한 PPI 기법이 단순한 임계값 기반이 아닌 시각적 기여도를 정량적으로 평가하여 클러스터 내 요소를 선택적으로 제거하는 데 효과적임을 보여준다. 또한, 클러스터 밀도 제조정은 렌더링 연산량 감소와 메모리 최적화에 기여하며, 실시간 렌더링 성능 향상에 중요한 역할을 한다.

### 3. 평가 방법

제안된 프레임워크의 렌더링 품질과 실시간 처리 성능을 정량적으로 평가하기 위해, 다음과 같은 지표를 사용해 성능을 측정하였다. 렌더링 품질은 GT와의 비교를 통해 세 가지 지표로 평가되었다. 첫째, Peak Signal-to-Noise Ratio (PSNR)은 평균 제곱 오차(Mean Squared Error, MSE)를 기반으로 하며, 식 (6)에 따라 계산된다:

$$PSNR = 10 \cdot \log_{10} \left( \frac{L^2}{MSE} \right) \quad (6)$$

여기서  $L$ 은 픽셀의 최대값(예:225),  $MSE$ 는 영상 간의 차이의 제곱 평균을 의미한다. PSNR값이 클수록 원본과

유사한 고품질 영상을 나타낸다. 둘째, Structural Similarity Index(SSIM)은 영상의 밝기, 대비, 구조 정보를 종합적으로 평가하는 지표로, 식 (7)로 정의된다<sup>[19]</sup>:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

여기서  $\mu$ 는 평균,  $\sigma$ 는 분산,  $\sigma_{xy}$ 는 공분산이며,  $C_1$ 과  $C_2$ 는 분모가 0이 되는 것을 방지하기 위한 작은 상수이다. SSIM값은 0에서 1 사이로, 1에 가까울수록 원본과 구조적으로 유사함을 나타낸다. 셋째, Learned Perceptual Image Patch Similarity(LPIPS)는 딥러닝 기반으로 사람의 시각적 유사성을 반영하는 지표로, 값이 낮을수록 원본과 시각적으로 유사함을 의미한다<sup>[20]</sup>.

실시간 렌더링 성능은 다음과 같은 지표로 측정하였다. Frames Per Second(FPS)는 초당 처리 가능한 프레임 수로, 실시간성의 주요 기준으로 사용된다. GPU 메모리 사용량은 CUDA 환경에서의 메모리 할당량(Allocated Memory)을 측정하여 시스템 자원 효율성을 분석한다. 각 지표는 학습 반복 횟수 7,000회와 30,000회에서 평가되었으며, 본 논문에서는 30,000회 시점의 최종 성능을 기준으로 주요 결과를 보고한다.

또한, 가우시안 파라미터의 최소 메모리 요구량인



Parameter Memory는 다음과 같이 계산되었다:

$$Memory_{params} = \sum_{t \in T} numel(t) \times element\_size(t) \quad (8)$$

여기서  $T$ 는 가우시안 모델의 파라미터 텐서 집합  $\{xyz, scaling, rotation, opacity, features\_dc, features\_rest\}$ 를 나타낸다.  $numel(t)$ 은 텐서  $t$ 의 원소 개수,  $element\_size(t)$ 는 각 원소의 크기(Byte)를 의미한다. 본 연구에서는 모든 파라미터가 32-bit 부동소수점(FP32) 형식을 사용하여 4 Byte이므로  $element\_size(t) = 4$ 이다. 따라서 가우시안 개수를  $N$ , 파라미터 총 개수를  $P$ 라 하면, 식 (9)와 같이 단순화된다:

$$Memory_{params} = N \times P \times 4 \quad (9)$$

이 식은 파라미터의 순수 메모리 사용량 나타내며, 이를 통해 메모리 효율성을 정량적으로 평가할 수 있다.

#### 4. 실험 결과

그림 4는 Bicycle 장면을 기준으로 GT, 기존 3DGS, 그리고 제안된 PPI 3DGS의 시각적 결과를 비교한 것이다. 세 결과는 전반적으로 유사한 품질을 유지하지만, 그림 4 하단의 확대된 영역에서 PPI 3DGS는 벤치와 나뭇가지 같은 고주파 영역에서 선명한 세부 묘사를 보여주며, 잔디와 같은 저주파 영역에서도 경계가 뚜렷하게 표현됨을 확인할 수 있다. 정량적 비교 결과는 표 2에 정리되었다. 제안된 PPI 3DGS는 가우시안 수를 54,272개에서 37,929개로 약 43% 줄였음에도 불구하고, PSNR은 19.85에서 21.64로 9.01% 향상되었고, SSIM은 0.672에서 0.693로 3.12% 개선되었다. 이는 프루닝으로 인해 연산량이 감소했음에도 픽셀 정확도와 품질이 향상되었음을 나타낸다. LPIPS는 0.273에서 0.157으로 42.4% 감소하여 이미지 유사성 측면에서도 매우 개선된 결과를 보였다. 결과적으로, 제안된 기법은 주요 품질 지표에서 향상된 성능을 달성했다.

표 3은 실시간 렌더링 성능의 정량적 비교 결과를 나타낸

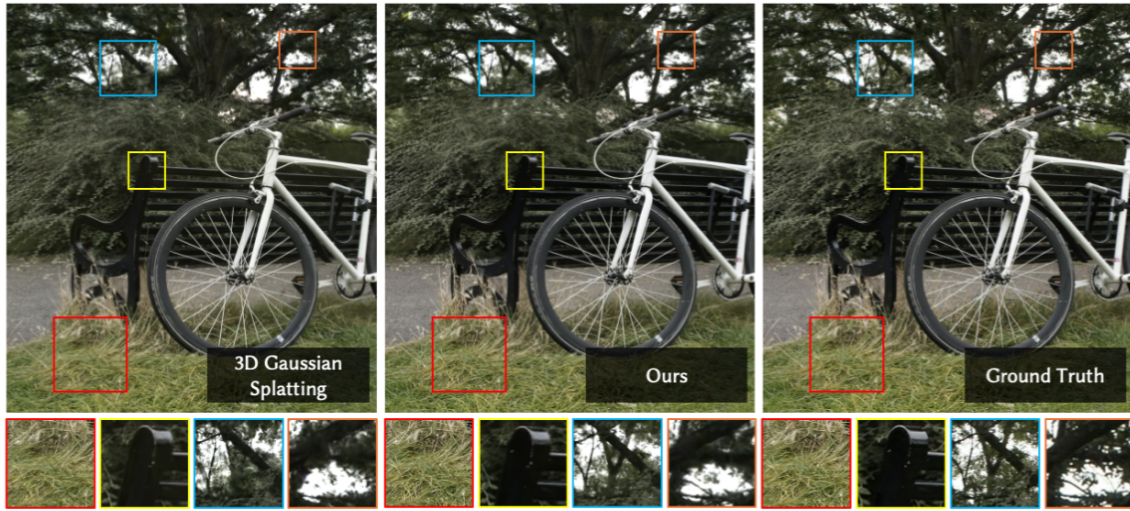


그림 4. Bicycle 장면에 대한 기존 3D Gaussian Splatting, 제안한 PPI 3DGS(Ours), 그리고 Ground Truth의 시각적 비교. 확대 영역을 통해 고주파 세부 묘사 및 경계 표현에서 제안 기법이 기존 대비 더 높은 품질을 나타낸 것을 확인할 수 있다. 특히, 잔디 및 나뭇가지와 같은 복잡한 영역에서도 제안 방법은 선명한 구조와 경계를 재현하고 있음을 보여준다.

Fig. 4. Visual comparison of the Bicycle scene between the conventional 3D Gaussian Splatting, the proposed PPI 3DGS (Ours), and the Ground Truth. The zoomed-in regions demonstrate that the proposed method achieves higher quality in terms of high-frequency detail preservation and boundary sharpness compared to the baseline. In particular, it effectively reconstructs clear structures and edges in complex areas such as grass and tree branches.

표 2. 제안한 PPI 3DGS와 기존 3DGS 간 렌더링 품질 및 가우시안 수 비교  
표는 동일한 반복 횟수(30K)를 기준으로, PPI 3DGS와 기존 3DGS 간의 *PSNR*, *SSIM*, *LPIPS*와 같은 렌더링 품질 지표 및 가우시안 수를 비교한 결과를 나타낸다. PPI 3DGS는 가우시안 수를 약 43% 감소하면서, 모든 품질 지표에서 기존 대비 향상된 성능을 보였다.

Table 2. Comparison of rendering quality and number of Gaussian between the proposed PPI 3DGS and the conventional 3DGS. The table presents the results based on the same number of training iterations (30K), comparing rendering quality metrics such as *PSNR*, *SSIM*, and *LPIPS*, as well as the number of Gaussians. PPI 3DGS achieves improved performance across all quality metrics while reducing the number of Gaussians by approximately 43% compared to the conventional approach.

Index	3DGS(30K)	PPI 3DGS(30K)
Gaussian count (↓)	54,272	37,929 (43.05%)
<i>PSNR</i> (↑)	19.85	21.64 (9.01%)
<i>SSIM</i> (↑)	0.672	0.693 (3.12%)
<i>LPIPS</i> (↓)	0.273	0.157 (42.4%)

표 3. 기존 3DGS와 제안한 PPI 3DGS의 실시간 렌더링 성능 정량적 비교 결과. 동일한 학습 조건(30K 반복)에서 제안 기법은 FPS가 2.35% 향상되었으며, 파라미터 메모리 사용량은 약 30.12% 절감되었다.

Table 3. Quantitative comparison of real-time rendering performance between the baseline 3DGS and the proposed PPI 3DGS. Under identical training conditions (30K iterations), the proposed method achieved a 2.35% improvement in FPS and reduced parameter memory usage by approximately 30.12%.

Evaluation index	3DGS(30K)	PPI 3DGS(30K)
FPS(↑)	85.02	87.02 (2.35%)
Parameter Memory(MB)(↓)	12.216	8.537 (30.12%)

다. PPI 3DGS는 평균 FPS에서 87.02를 기록하여 기존 3DGS의 85.02 대비 약 2.35% 향상된 성능 나타났다. 이는 동일 시간 내 더 많은 프레임을 처리함으로써 렌더링 효율성을 개선했음을 의미한다. 또한, 가우시안 파라미터 메모리 사용량 측면에서, PPI 3DGS는 8.537MB로 기존 3DGS의 12.216MB 대비 약 30.12% 절감된 결과를 보였다. 이러한 메모리 절감 효과는 제안된 방법이 가우시안 수 감소를 통해 파라미터 저장 효율성을 크게 향상시켰음을 시사한다. 특히, 파라미터 수의 감소는 메모리 사용량뿐 아니라 모델의 연산 비용을 줄여, 실시간 애플리케이션 환경에서 처리 지연 최소화에도 기여할 수 있다.

이러한 실험 결과는 클러스터링과 PPI 기반 프루닝이

3DGS의 실시간 렌더링 품질 및 성능을 개선함을 보여준다. 클러스터링을 통해 가우시안의 공간적 구조를 정리하고, 중요도 기반으로 불필요한 요소를 제거함으로써 모델 경량화와 연산 최적화가 동시에 달성했다. 프루닝 후에도 *PSNR*과 *SSIM*은 실용적으로 사용 가능한 성능을 유지했으며, *LPIPS* 또한 텍스처 품질의 실용적 기준을 충족하였다. 또한, 제한된 연산 자원 환경에서도 높은 품질의 실시간 3D 렌더링을 구현할 수 있는 가능성을 입증하였다.

## V. 결론

본 논문에서 3DGS의 렌더링 품질을 유지하면서 연산 효율성과 메모리 사용량을 개선하기 위해 K-Means 기반 공간 클러스터링과 시각적 중요도 기반 프루닝을 결합한 프레임워크를 제안했다. 이 기법은 각 클러스터 내 가우시안의 위치, 스케일, 투명도 등의 속성을 활용해 시각적 기여도를 정량화하고, 중요도가 낮은 가우시안을 선택적으로 제거하여 전체 연산량을 효과적으로 줄이는 구조를 갖는다. 실험 결과, PPI 3DGS는 가우시안 수를 약 43% 감소시켰음에도 불구하고 *PSNR*, *SSIM*, *LPIPS* 품질 지표에서 기존 3DGS보다 우수한 시각적 품질을 유지했다. 또한, FPS와 GPU 메모리 사용량에서 소폭의 성능 향상을 달성하였으며, 추가적인 신경망 없이 기존 포인트 기반 구조만으로 경량화와 품질 향상을 동시에 구현하였다.

특히, 품질 개선은 기존 3DGS가 시각적 기여도를 고려하지 않고 가우시안을 균일하게 분포시켜 고주파 영역에서 디테일 손실과 저주파 영역에서 자원 낭비를 초래하는 품질 분산 문제를 내포하고 있다. 반면, PPI 3DGS는 각 가우시안의 시각적 중요도를 기준으로 연산 대상을 선별하여 나뭇가지와 경계선 같은 고주파 영역에 자원을 집중하고, 저주파 영역은 중복되거나 기여도가 낮은 가우시안을 제거함으로써 품질 밀도를 효율적으로 재배치했다. 저주파 영역은 표현 복잡도가 낮아 소수의 가우시안만으로도 충분한 시각적 품질을 유지할 수 있다. 결과적으로 텍스처 선명도와 구조적 표현력이 모두 향상되었으며, 제한된 가우시안 자원을 중요한 영역에 효과적으로 분배함으로써 품질과 성능을 동시에 개선하였다.

제안 기법은 클러스터링과 프루닝을 통해 실시간 렌더링 성능을 2.35% 향상시켰으나, 정적인 프루닝 비율을 사용해 다양한 시점 변화에 적응하지 못하는 한계가 있다. 이러한 한계를 보완하기 위해 향후 연구에서는 장면 복잡도나 시점 변화에 따라 가우시안 밀도를 동적으로 조절하는 실시간 프루닝 기법을 도입할 계획이다. 또한, 클러스터링과 프루닝 연산의 병렬화를 통해 전체 렌더링 속도를 더욱 향상시키고자 한다. 이러한 연구 방향은 실시간 성능과 시각적 품질을 모두 만족시키는 정밀도 경량화 기술로 발전할 가능성이 있으며, VR/AR 및 실감형 콘텐츠와 같은 다양한 실시간 응용 분야에서의 활용성을 높일 것으로 기대한다.

#### 참 고 문 헌 (References)

- [1] A. Tewari, J. Thies, B. Mildenhall, P. Srinivasan, E. Tretschk, Y. Wang, C. Lassner, V. Sitzmann, R. Martin Brualla, S. Lombardi, T. Simon, C. Theobalt, M. Nießner, J. T. Barron, G. Wetzstein, and M. Zollhöfer, "Advances in Neural Rendering," *Computer Graphics Forum*, Vol. 41, No. 2, pp. 703-735, April 2022.  
doi: <https://doi.org/10.1111/cgf.14507>
- [2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis," *Proceedings of the European Conference on Computer Vision (ECCV)*, Lecture Notes in Computer Science, Vol. 12346, pp. I 405-I 421, September 2020.  
doi: [https://doi.org/10.1007/978-3-030-58452-8\\_24](https://doi.org/10.1007/978-3-030-58452-8_24)
- [3] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian Splatting for Real Time Radiance Field Rendering," *ACM Transactions on Graphics*, Vol. 42, No. 4, Article 1, pp. 1-14, July 2023.  
doi: <https://doi.org/10.1145/3592433>
- [4] X. Liu, X. Wu, P. Zhang, S. Wang, Z. Li, and S. Kwong, "CompGS: Efficient 3D Scene Representation via Compressed Gaussian Splatting," *Proceedings of the 32nd ACM International Conference on Multimedia (ACM MM '24)*, pp. 2936-2944, October 2024.  
doi: <https://doi.org/10.1145/3664647.3681468>
- [5] G. Kopanas, J. Philip, T. Leimkühler, and G. Drettakis, "Point Based Neural Rendering with Per View Optimization," *Computer Graphics Forum (Proceedings of the Eurographics Symposium on Rendering)*, Vol. 40, No. 4, pp. 1-12, June 2021.  
doi: <https://doi.org/10.1111/cgf.14339>
- [6] A. Hanson, A. Tu, V. Singla, M. Jayawardhana, M. Zwicker, and T. Goldstein, "PUP 3D GS: Principled Uncertainty Pruning for 3D Gaussian Splatting," *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*, pp. 5949-5958, June 2025.  
doi: <https://doi.org/10.48550/arXiv.2406.10219>
- [7] S. Xie, W. Zhang, C. Tang, Y. Bai, R. Lu, S. Ge, and Z. Wang, "MesonGS: Post training Compression of 3D Gaussians via Efficient Attribute Transformation," *Proceedings of the European Conference on Computer Vision (ECCV)*, Lecture Notes in Computer Science, Vol. 15091, pp. 434-452, October 2024.  
doi: [https://doi.org/10.1007/978-3-031-73414-4\\_25](https://doi.org/10.1007/978-3-031-73414-4_25)
- [8] Y. Wang et al., "Neural Compression for 3D Gaussian Splatting," *Proceedings of SIGGRAPH Asia*, Tokyo, Japan, pp.1-9, Dec. 2024.
- [9] Z. Sun, C. Ge, J. Wang, M. Lin, H. Chen, H. Li, and X. Sun, "Entropy Driven Mixed Precision Quantization for Deep Network Design," *Advances in Neural Information Processing Systems*, Vol. 35, pp. 21508-21520, December 2022.
- [10] Z. Tang, C. Feng, X. Cheng, W. Yu, J. Zhang, Y. Liu, X. Long, W. Wang, and L. Yuan, "NeuralGS: Bridging Neural Fields and 3D Gaussian Splatting for Compact 3D Representations," *arXiv preprint*, March 29, 2025.  
doi: <https://doi.org/10.48550/arXiv.2503.23162>
- [11] S. Shin, J. Park, and S. Cho, "Locality Aware Gaussian Compression for Fast and High Quality Rendering," *arXiv preprint*, January 10, 2025.  
doi: <https://doi.org/10.48550/arXiv.2501.05757>
- [12] Z. Liu, M. Sun, T. Zhou, G. Huang, and T. Darrell, "Rethinking the Value of Network Pruning," *arXiv preprint*, October 11, 2018.  
doi: <https://doi.org/10.48550/arXiv.1810.05270>
- [13] H. Cheng, M. Zhang, and J. Q. Shi, "A Survey on Deep Neural Network Pruning: Taxonomy, Comparison, Analysis, and Recommendations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 46, No. 12, pp. 10558-10578, December 2024.  
doi: <https://doi.org/10.1109/TPAMI.2024.3447085>
- [14] Z. Zhang, T. Song, Y. Lee, L. Yang, C. Peng, R. Chellappa, and D. Fan, "LP 3DGS: Learning to Prune 3D Gaussian Splatting," *Advances in Neural Information Processing Systems*, Vol. 37, pp. 122434-122457, NeurIPS 2024.  
doi: <https://doi.org/10.48550/arXiv.2405.18784>
- [15] S. R. Bulò, L. Porzi, and P. Kotschieder, "Revising Densification in Gaussian Splatting," *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, pp. 347-362, April 2024.  
doi: [https://doi.org/10.1007/978-3-031-73036-8\\_20](https://doi.org/10.1007/978-3-031-73036-8_20)
- [16] G. Grubert et al., "Improving Adaptive Density Control for 3D Gaussian Splatting," *arXiv preprint*, arXiv:2503.14274, Mar. 2025.  
doi: <https://doi.org/10.48550/arXiv.2503.14274>
- [17] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k Means Clustering Algorithm," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, Vol. 28, No. 1, pp. 100-108, 1979.  
doi: <https://doi.org/10.2307/2346830>
- [18] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "Mip NeRF 360: Unbounded Anti Aliased Neural Radiance Fields," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5460-5469, June 2022.  
doi: <https://doi.org/10.1109/CVPR52688.2022.00539>
- [19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity,"

IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600-612,  
April 2004.

doi: <https://doi.org/10.1109/TIP.2003.819861>

[20] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The

Unreasonable Effectiveness of Deep Features as a Perceptual Metric,"  
Proceedings of the IEEE/CVF Conference on Computer Vision and  
Pattern Recognition (CVPR), Vol. 2018, pp. 586-595, June 2018.

doi: <https://doi.org/10.1109/CVPR.2018.00068>

---

## 저 자 소 개

### 조 수 빈



- 2022년 ~ 현재 : 동아대학교 컴퓨터공학과 학사과정
- ORCID : <https://orcid.org/0009-0005-8833-3583>
- 주관심분야 : 영상 처리, 딥러닝, 컴퓨터비전

### 박 성 준



- 2025년 : 동아대학교 AI학과(공학사)
- 2025년 ~ 현재 : 동아대학교 컴퓨터공학과 석사과정
- ORCID : <https://orcid.org/0009-0006-3560-763X>
- 주관심분야 : 멀티미디어, 딥러닝, 머신비전, 3D

### 서 정 일



- 1994년 : 경북대학교 전자공학과(공학사)
- 1996년 : 경북대학교 대학원 전자공학과(공학석사)
- 2005년 : 경북대학교 대학원 전자공학과(공학박사)
- 1998년 ~ 2000년 : LG반도체 주임연구원
- 2010년 ~ 2011년 : 영국 Southampton University, ISVR 방문연구원
- 2000년 ~ 2023년 : 한국전자통신연구원 실감미디어연구실장
- 2023년 ~ 현재 : 동아대학교 컴퓨터AI공학부 부교수
- ORCID : <https://orcid.org/0000-0001-5131-0939>
- 주관심분야 : 멀티미디어 부호화, 컴퓨터비전, 실감 영상 및 음향, 멀티미디어 표준화