

## 실공간 객체 Gaussian Splat 생성 및 하이퍼파라미터 최적화 연구

최경호 한건희 이승미 정종범

국립군산대학교 컴퓨터정보공학과

{cgh3578, gunhee9369, qweasd030622, jongbeomjeong}@kunsan.ac.kr

## Research of Gaussian Splat Generation of Real-Space Object and Hyperparameter Optimization

GyeongHo Choi GunHee Han Seung-Mi Lee Jong-Beom Jeong

Department of Computer Science and Information Engineering,

Kunsan National University (KSNU)

## 요 약

본 논문은 3D Gaussian Splatting (3DGS) 기반의 gsplat 라이브러리를 활용하여 하이퍼파라미터 변화가 실공간 객체의 복원 품질에 미치는 영향을 체계적으로 분석한다. 직접 촬영한 실공간 객체 영상으로부터 COLMAP을 이용해 카메라 포즈를 추정하고, gsplat의 학습 파이프라인의 다양한 하이퍼파라미터를 조정하며 객체 복원 실험을 수행하였다. 복원 결과는 peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), learned perceptual image patch similarity (LPIPS) 지표를 통해 정량적으로 평가하였으며, 각 파라미터가 복원 품질과 학습 효율에 미치는 영향을 비교·분석하였다. 실험 결과, 하이퍼파라미터 설정에 따라 복원 품질 차이가 발생하였으며, 실공간 객체 복원에 효과적인 파라미터 조합을 도출하였다. 본 연구는 3DGS 기반 객체 복원 과정에서 성능 향상을 위한 하이퍼파라미터 설정 기준과 활용 방향을 제공할 것으로 기대된다.

## 1. 서론

최근 3차원 재구성 분야에서 Neural Radiance Fields (NeRF)[1]는 다각도에서 촬영된 2차원 이미지를 기반으로 볼륨 렌더링을 수행하여 3차원 객체와 장면을 정밀하게 복원할 수 있는 대표적인 기술로 주목받아 왔다. NeRF는 객체를 촬영한 영상 데이터만으로 실공간을 사실적으로 재구성할 수 있다는 장점을 가지지만, 신경망 기반의 복잡한 연산 구조로 인해 높은 계산 비용이 요구되며 실시간 렌더링에 한계가 존재한다. 2023년 소개된 3DGS는 이러한 한계를 극복하기 위해 제안되었다[2]. 3DGS는 Gaussian 기반 장면 표현과 래스터화 (rasterization) 기법을 활용하여 높은 복원 품질과 실시간 렌더링 성능을 달성하였다. 이후 객체 복원 및 실시간 렌더링 분야에서 3DGS 기반 연구가 활발히 진행되고 있으며, 이를 위한 최적화 기법과

라이브러리 또한 지속적으로 개발되고 있다.

최근에는 기존 3DGS의 학습 및 렌더링 성능을 더욱 향상시키기 위해 CUDA 커널 기반의 최적화 라이브러리인 gsplat[3]이 활용되고 있다. 최근 gsplat을 활용하여 렌더링 성능 향상, 동적 장면 처리, 대규모 장면 복원 등의 연구는 활발히 이루어지고 있다. 하지만 실제 환경에서 촬영된 실공간 객체를 대상으로 gsplat의 하이퍼파라미터 변화가 복원 품질에 미치는 영향을 다양한 실험을 통해 체계적으로 분석하는 연구 역시 필요하다.

본 논문은 직접 촬영한 실공간 객체 영상을 기반으로 COLMAP을 통해 카메라 포즈를 추정하고, gsplat 파이프라인의 하이퍼파라미터 변화에 따른 객체 복원 성능을 분석한다. 이를 위해 총 31종의 파라미터 실험을 수행하였으며, PSNR, SSIM,

LPIPS 등의 정량적 지표를 활용하여 각 하이퍼파라미터가 복원 품질과 학습 효율에 미치는 영향을 비교·분석하였다.

본 논문의 구성은 다음과 같다. 2절에서는 NeRF 및 3DGS 기술의 배경과 gsplat 라이브러리의 선행 연구를 소개한다. 3절에서는 실공간 객체 촬영 데이터 기반의 하이퍼파라미터 실험 방법을 설명한다. 4절에서는 31종 파라미터 중 기본값 대비 PSNR, SSIM, LPIPS 세 지표 모두에서 향상을 보인 11종의 주요 파라미터에 대한 복원 품질 분석 결과를 제시한다. 마지막으로 5절에서는 본 논문의 결론을 맺는다.

## 2. 선행 연구

NeRF는 여러 시점에서 촬영된 영상과 카메라 포즈 정보를 이용하여 장면을 연속적인 3차원 공간 정보로 표현하는 신경망 기반의 3차원 재구성 기법이다. 공간상의 위치와 시점 정보를 입력으로 사용하여 색상과 밀도를 추정하고, 볼륨 렌더링을 수행한다. 이를 통해 실공간 객체에서 새로운 시점의 이미지를 생성할 수 있다. 그러나 다층 퍼셉트론 기반의 반복적인 신경망 연산이 필요하기 때문에 학습 및 렌더링 속도가 느리며, 실시간 렌더링 환경에서는 처리 속도 측면의 제약이 존재한다.

이러한 문제를 개선하기 위해 제안된 3DGS는 장면을 수많은 3차원 Gaussian splats (GS) 분포로 표현하는 방식이다. 각 GS는 위치, 색상, 불투명도, 회전 및 공분산 정보를 가지며, 기존 NeRF와 달리 명시적 장면 표현 방식을 사용하여 장면을 구성한다. 이를 통해 병렬 연산 구조를 효율적으로 활용하여 고속 렌더링이 가능하다. 또한 adaptive density control (ADC) 기법을 통해 GS를 동적으로 추가하거나 제거하며, 복제 (cloning), 분할 (splitting) 및 가지치기 (pruning) 과정을 통해 세부 구조 표현과 연산 효율을 향상시킨다. 그러나 3DGS는 대규모 GS 처리 과정에서 메모리 사용량 증가와 연산 병목 문제가 발생할 수 있다. 특히 GS 개수가 증가할수록 렌더링 과정의 연산량과 GPU 메모리 사용량이 증가하며, 밀집화 (densification) 와 가지치기 과정이 반복적으로 수행되기 때문에 추가적인 최적화가 요구된다.

이에 따라 gsplat 같은 CUDA 커널 기반 라이브러리가 개발되었다. gsplat은  $16 \times 16$  tile 기반 GPU 병렬 처리 구조를 사용하여 래스터화를 수행하며, 투영, 정렬 및 알파 블렌딩 과정의 연산 효율을 향상시켜 학습 시간을 단축하였다. 3DGS의 밀집화 휴리스틱 (Densification heuristics)을 따르는 기본 학습 방법 외에도 Markov Chain Monte Carlo (MCMC) 기반 학습 전략과 다양한 하이퍼파라미터 설정을 지원한다. ADC, Absolute Gradient (AbsGrad) 및 MCMC 기반 밀집화 전략을 제공하며, GS 생성 및 제거 과정을 효율적으로 수행할 수 있도록 구성되어 있다. 특히 AbsGrad는 기울기 (gradient) 상쇄 현상으로 인해

세부 구조 표현이 감소하는 문제를 개선하며, MCMC 전략은 GS를 보다 효율적으로 배치하여 복원 품질 향상에 기여한다. 또한 안티앨리어싱 (anti-aliasing) 기반 렌더링을 지원하여 계단 현상을 감소시킬 수 있으며, 포즈 최적화 기능을 통해 카메라 포즈 보정 또한 가능하다.

최근에는 gsplat 기반 연구가 렌더링 성능 향상[4], 메모리 최적화[5], 동적 장면 처리[6] 및 대규모 장면 복원[7] 등을 중심으로 활발히 진행되고 있다. 그러나 실제 환경에서 촬영된 실공간 객체를 대상으로 다양한 하이퍼파라미터 변화가 복원 품질과 학습 효율에 미치는 영향을 체계적으로 분석한 연구[8] 역시 중요하다. 본 논문은 gsplat 파이프라인의 다양한 하이퍼파라미터를 대상으로 실공간 객체 복원 품질 변화를 분석하고, 효과적인 하이퍼파라미터 설정 방향을 제시하고자 한다.

## 3. 실험 방법

본 논문은 gsplat 파이프라인에서 제공하는 다양한 하이퍼파라미터가 실공간 객체 복원 품질에 미치는 영향을 분석하기 위해 실험을 수행하였다. 실험은 실제 환경에서 직접 촬영한 차량용 로봇 객체 데이터를 기반으로 진행하였으며, 동일한 데이터셋과 학습 환경에서 각 하이퍼파라미터를 개별적으로 변경하여 복원 품질 변화를 비교하였다.

실험 데이터셋은 연구실 환경에서 차량용 로봇 객체를 다양한 시점에서 촬영하여 구성하였다. 객체 중심 (object-centric) 방식으로 촬영을 수행하였으며, 객체의 형태와 표면 정보가 충분히 포함될 수 있도록 카메라 시점을 변화시키며 영상을 획득하였다. 획득한 영상 데이터는 ffmpeg을 사용하여 초당 2장의 프레임으로 이미지를 변환하였고, 총 155장의 이미지를 획득하였다. 이후 COLMAP을 이용하여 카메라 포즈와 sparse point clouds를 생성하였으며, 전체 이미지를 학습에 활용하였다.

본 실험은 Ubuntu 24.04 기반 환경에서 수행되었다. 실험 시스템의 GPU는 NVIDIA GeForce RTX 5090을 사용하였으며, CUDA Version 12.8 및 NVIDIA Driver Version 570.211.01 환경에서 실험을 진행하였다. 학습 환경은 Python 3.11.15 및 PyTorch 2.11.0+cu128 기반으로 구성하였으며, GPU 가속이 활성화된 상태에서 모든 학습 및 렌더링 실험을 수행하였다.

모든 실험은 gsplat의 기본 설정을 baseline으로 총 30,000 iteration 동안 학습을 수행하였다. 파라미터를 개별적으로 변경하여 baseline과의 성능을 비교하였으며, 평가 지표로는 PSNR, SSIM, LPIPS를 사용하였다. 또한, GS 개수와 실험 시간 비율을 비교하여 복원 품질뿐만 아니라 학습 효율 및 연산 비용 변화도 함께 분석하였다.

&lt;표 1&gt; 하이퍼파라미터 변경에 따른 정량적 품질 평가 비교

| Options                   | Value   | $\Delta$ PSNR | $\Delta$ SSIM | $\Delta$ LPIPS | Number of GS | Runtime Ratios |
|---------------------------|---------|---------------|---------------|----------------|--------------|----------------|
| baseline                  | -       | 27.89         | 0.9341        | 0.1142         | 999K         | 100.00%        |
| Init-scale                | 0.5     | + 0.25        | + 0.0002      | -0.0035        | 1,016K       | 98.69%         |
| sh-degree-interval        | 2,000   | + 0.33        | + 0.0007      | -0.0045        | 995K         | 39.73%         |
| scales-lr                 | 0.002   | + 0.38        | + 0.0017      | -0.0035        | 926K         | 36.69%         |
| quats-lr                  | 0.0005  | + 0.44        | + 0.0009      | -0.0037        | 988K         | 58.30%         |
| shN-lr                    | 0.00025 | + 0.68        | + 0.0015      | -0.0054        | 997K         | 48.96%         |
| antialiased               | TRUE    | + 0.53        | + 0.0006      | -0.0046        | 1,030K       | 52.00%         |
| strategy.grow-scale3d     | 0.02    | + 0.47        | + 0.0009      | -0.0046        | 993K         | 39.84%         |
| strategy.prune-scale3d    | 0.2     | + 0.58        | + 0.0021      | -0.0064        | 984K         | 39.08%         |
| strategy.refine-stop-iter | 10,000  | + 0.36        | + 0.0005      | -0.0040        | 845K         | 38.11%         |
|                           | 20,000  | + 0.25        | + 0.0003      | -0.0014        | 977K         | 40.17%         |
| strategy.refine-every     | 200     | + 0.40        | + 0.0009      | -0.0023        | 691K         | 33.76%         |
| strategy.revised-opacity  | TRUE    | + 0.54        | + 0.0018      | -0.0037        | 993K         | 39.63%         |

#### 4. 실험 결과

본 논문은 gsplat 파이프라인의 기본 설정을 기준으로 31종 하이퍼파라미터를 체계적으로 실험하였다. 각 파라미터를 기본값 대비 상향 및 하향 조정하여 실험을 진행하였으며, 기본 설정 대비 PSNR, SSIM, LPIPS 세 지표 모두에서 향상을 보인 경우를 선별하였다. 최종적으로 11종의 주요 파라미터와 그 결과를 표 1에 정리하였다.

quats-lr은 GS의 회전 파라미터에 대한 학습률이다. 기본값 대비 감소시킨 결과, 객체의 기하학적 형태가 더 정확하게 정렬되었고, 경계의 왜곡이 줄어들어 GS의 방향성이 보다 안정적으로 최적화되어 물체의 구조적 특징을 잘 반영한 결과로 판단된다.

shN-lr은 higher-order spherical harmonics (SH) 계수의 학습률을 제어하는 하이퍼파라미터이다. 본 실험에서는 기본값 대비 2배 증가시켜 0.00025로 설정하였을 때, PSNR 28.57 dB를 기록하며 전체 실험 중 최고 성능을 달성하였다. 이는 SH 계수가 시점에 따른 색상 변화를 주로 모델링하기 때문이다. 학습률을 높여 SH 계수를 적극적으로 업데이트함으로써, 장면 내 복잡한 광반사와 미세한 조명 효과를 더 정확하게 학습할 수 있었으며, 결과적으로 렌더링의 광학적 충실도가 향상되어 PSNR이 유의미하게 개선된 것으로 분석된다.

antialiased 옵션은 렌더링 과정에서 계단 현상을 완화하는 기능이다. 이를 활성화한 결과, 객체 가장자리의 선명도가 크게 향상되었으며, 특히 곡면과 세밀한 구조에서 계단 현상이 현저히 감소하였다. 이는 래스터화 과정에서 서브 픽셀 수준의 보정 처리가 적용되어 전체 화질의 정밀도가 높아졌기 때문이다.

보인다.

strategy 관련 파라미터들은 GS의 생성과 제거를 제어한다. 그중 strategy.prune-scale3d는 3차원 공간에서 일정 크기 이상으로 거대한 GS를 제거하는 파라미터이다. 이 값을 증가시킬 경우, 제거되어야 할 GS들이 잔존하게 되며, 이들이 배경과 경계면의 빈 공간을 채우면서 렌더링 오차가 감소하게 된다. 이로 인해 객관적 수치 평가에서 오차 요인이 상쇄되어 기본 설정보다 지표가 더 높게 측정된다. 다만, 이는 객관적 지표의 상승이 시각적 복원 품질의 향상으로 이어지지 않을 수 있음을 나타낸다.

또한, strategy.revised-opacity는 GS 복제 과정에서 투명도 (opacity) 임계값 기준을 재조정하는 기능이다. 이를 활성화함으로써 허공에 불필요하게 생기는 플로터 (floater) 현상이 제거되어 정량적 복원 품질이 뚜렷하게 개선되었다.

이러한 하이퍼파라미터 최적화가 실제 복원된 결과물에 미치는 영향을 시각적으로 검증하기 위해 하이퍼파라미터별 복원 품질을 비교하였다.



<그림 1> 객체 복원 결과



〈그림 2〉 하이퍼파라미터별 시각적 복원 품질 비교

그림 1은 본 논문의 실험 대상인 차량용 로봇 객체의 전체적인 3D 복원 결과이다. 그림 2에서 확인할 수 있듯이, 11종의 하이퍼파라미터는 전반적으로 불필요한 노이즈를 효과적으로 제거하고 렌더링 품질을 크게 향상시켰다. 하지만 strategy.prune-scale3d의 경우, 표 1의 객관적 수치에서 높은 지표 향상을 기록하였으나 실제 시각적 복원 품질에서는 오히려 기본 설정보다 좋지 않은 결과를 보여주었다. 이를 분석한 결과, 위에 언급한 바와 같이 거대한 가우시안들이 공간상에 제거되지 못하고 잔존하면서 시각적 품질을 낮춘 것으로 확인하였다. 특히 그림 2의 Region 1 영역에서 슬롯 구조 위의 빈공간을 확인하면, 경계면을 따라 미세한 가우시안 노이즈가 남아있는 것을 확인할 수 있었다. 반면, shN-lr은 수치적 지표의 향상과 더불어 시각적 복원 역시 우수하게 수행되었으며, antialiased의 경우 타이어 실루엣 영역이 매끄럽고 정밀하게 복원된 것을 확인하였다.

## 5. 결론

본 논문은 gsplat 파이프라인을 기반으로 3DGS의 하이퍼파라미터 31종에 대한 비교 실험과 성능 평가를 진행하였다. 그중 정량적 지표가 유의미하게 개선된 11종을 선별하여 체계적으로 분석하였다. 분석 결과, shN-lr, antialiased, quats-lr, strategy.revised-opacity 조정 시 정량적 지표와 시각적 품질이 모두 향상되었다. 반면, strategy.prune-scale3d는 정량적 수치의 상승에도 불구하고 실제 시각적 노이즈를 제거하지 못하는 것으로 확인되었다. 이를 통해 객체 중심 데이터셋 복원 시, shN-lr 학습률 상향 및 antialiased 옵션 활성화가 복원 품질 개선에 적합한 설정임을 확인하였다. 향후 연구에서는 대규모 장면 데이터셋을 대상으로 추가적인 하이퍼파라미터 최적화 연구를 진행할 예정이다.

## ACKNOWLEDGEMENT

본 논문은 2026년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065)

## 참고 문헌

- [1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, Nerf: Representing scenes as neural radiance fields for view synthesis, *Communications of the ACM*, Vol. 65, No. 1, pp. 99-106, 2021.
- [2] B. Kerbl, G. Kopanas, T. Leimkühler, G. Drettakis, 3D Gaussian Splatting for Real-Time Radiance Field Rendering, *ACM Transactions on Graphics (TOG)*, Vol. 42, No. 4, pp. 1-14, 2023.
- [3] V. Ye, A. Kanazawa, gsplat: An Open-Source Library for Gaussian Splatting, *Journal of Machine Learning Research (JMLR)*, Vol. 26, No. 34, pp. 1-7, 2025.
- [4] S. Niedermayr, J. Stumpfegger, R. Lin, Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10349-10358, 2024.
- [5] M. S. Ali, C. Zhang, M. Cagnazzo, G. Valenzise, E. Tartaglione, S. H. Bae, Compression in 3d gaussian splatting: A survey of methods, trends, and future directions, *IEEE Transactions on Circuits and Systems for Video Technology*, 2026.
- [6] J. Luiten, G. Kopanas, B. Leibe, T. Albrecht, Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis, *2024 International Conference on 3D Vision (3DV)*, pp. 800-809, 2024.
- [7] Y. Liu, C. Luo, L. Fan, N. Wang, J. Peng, Z. Zhang, Citygaussian: Real-time high-quality large-scale scene rendering with gaussians, *European Conference on Computer Vision (ECCV)*, pp. 265-282, 2024.
- [8] G. Chen, W. Wang, A survey on 3d gaussian splatting, *ACM Computing Surveys*, 2024.