

DBSCAN 기반의 3D Gaussian Splatting 성능 개선을 위한 모델 개발

1

우다연^{#1}, 서은서^{#2}, 한채현^{#3}, 이연경^{#4}, *진창균⁵

성균관대학교 데이터사이언스융합전공¹, 중앙대학교 응용통계학과²,

이화여자대학교 인공지능학과³, 세종대학교 수학과통계학과⁴,

서울과학기술대학교 인공지능융합학과⁵

e-mail : dyeonwu@gmail.com¹, jeo053472@gmail.com², chehunhan16@gmail.com³,

leeyeonkyung296@gmail.com⁴, jcg.brainai@gmail.com⁵

Developing a Model for Improving 3D Gaussian Splatting Performance Based on DBSCAN

Dayeon Woo^{#1}, Eunseo Seo^{#2}, Chehun Han^{#3}, Yeonkyung Lee^{#4}, * Changgyun Jin⁵

Department of Data Science, Sungkyunkwan University¹,

Department of Applied Statistics, Chungang University²,

Department of Artificial Intelligence, Ewha Womans University³,

Department of Mathematics and Statistics, Sejong University⁴,

Department of Applied Artificial Intelligence, Seoul National University of Science & Technology⁵

Abstract

Recently, 3D Gaussian Splatting (3DGS), which improves the slow rendering speed of conventional Neural Radiance Field (NeRF), has gained significant attention in the field of 3D scene reconstruction. However, 3DGS generates many Gaussians in complex scenes, which leads to unnecessary noise in areas where observations are scarce. In this study, we propose a new method that applies DBSCAN clustering to improve these problems and improves the generation performance by reducing the number of Gaussians by flexibly grouping them even in unstructured data-dense

areas. In the experiments, the performance of DBSCAN was evaluated according to the number of times it is used, and the results show that the proposed model is superior to the existing 3DGS by generating images with better quality while reducing memory usage by about 10%.

I. 서론

최근 3D 장면을 생성하고 렌더링하는 computer vision 기술인 3D scene reconstruction이 크게 발전하고 있다^{[1][2][3]}. 그 중 NeRF^[1]는 Multi-Layer Perceptron (MLP)을 기반으로 한 3D scene

[#]: Equally Contribution

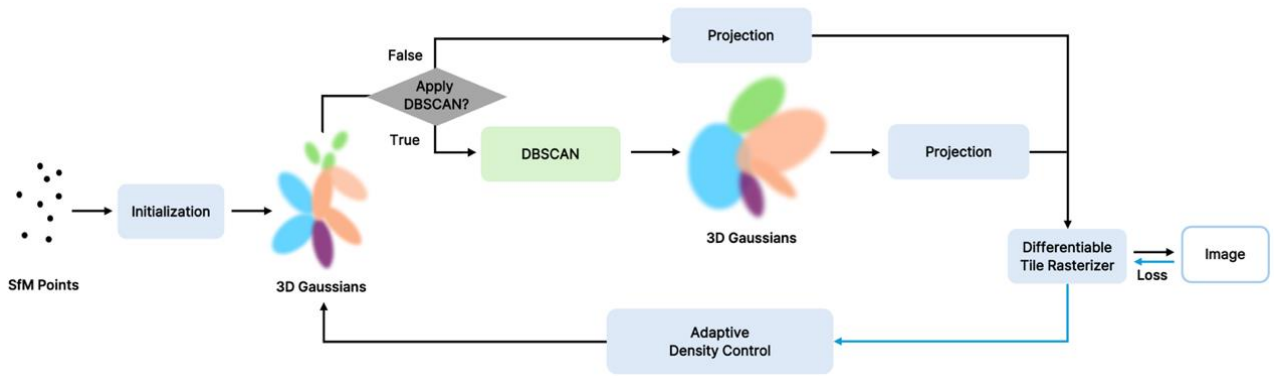


그림 1. DBSCAN을 적용한 3DGS모델의 구조도

reconstruction기법으로, MLP를 통해 각 ray의 위치와 방향을 입력 받아 해당 위치의 색상과 밀도를 예측함으로써 3D 장면을 복원한다. 이 방식은 2D 이미지에서 얻은 정보를 바탕으로 복잡한 3D 구조를 표현하는 데 매우 효과적이지만, 높은 연산 비용과 메모리 사용량을 요구하므로 대규모 데이터를 처리할 때 성능 상의 한계에 직면하게 된다. 이러한 문제를 해결하기 위해 제안된 3DGS^[2]는 NeRF와 달리 가우시안 분포를 활용하여 각 점을 splatting하는 방식을 취하며, 각 점의 색상과 밀도를 표현하여 장면을 재구성한다. 이를 통해 NeRF에 비해 처리 속도가 크게 향상되며, 복잡한 장면에서도 비교적 효율적으로 작동할 수 있다. 그러나 3DGS는 장면을 표현하기 위해 수많은 가우시안 분포를 생성하므로, 복잡한 장면에서는 가우시안 개수가 기하급수적으로 늘어나는 단점이 있다. 특히, 관측이 부족한 영역에서도 무분별하게 생성하게 되면서 불필요한 노이즈가 발생할 수 있다. 이러한 과도한 가우시안의 사용은 모델의 메모리 및 계산 효율성에 부정적인 영향을 미치며, 최적화를 통해 개선할 필요가 있다. 따라서 본 논문에서는 3DGS의 한계를 보완함과 동시에 성능을 개선하기 위해 DBSCAN^[4]을 적용한 새로운 알고리즘을 제안한다. 제안 모델은, 가우시안의 개수는 유지하면서 각 가우시안의 값 대신 클러스터의 대푯값으로 변경하여 사용하는 기존 방식과는 달리, 렌더링 도중 학습되고 있는 가우시안에 직접적으로 DBSCAN을 적용하여 수를 조절함으로써 생성되는 최종 이미지의 품질을 향상시키고 불규칙적인 데이터 분포에 유연하게 대응할 수 있도록 하였다. 실험 결과, 전체 메모리 사용량의 약 10%를 감소시켰으며 Structural Similarity Index Measure(SSIM), Peak Signal-to-noise ratio(PSNR)와 같은 이미지 품질 평가지표에서 3DGS에 비해 PSNR 기준 2.43 가량 향상된 성능을 보였다.

II. 제안 모델

2.1 3DGS

3DGS는 COLMAP과 같은 Structure-from-Motion(SfM) 기법을 통해 얻은 희소한 데이터 포인트로부터 3D 가우시안을 생성하는 과정으로 시작된다. 생성된 3D 가우시안은 카메라의 pose 값을 활용하여 world 좌표계에서 image 좌표계로 투영된다. 이후, 이미지를 16x16 크기의 타일 단위로 나누어 렌더링된 이미지를 생성하고, 이를 ground truth(GT) 이미지와 비교하여 손실 함수를 계산한다. 계산된 loss를 줄이기 위해 가우시안은 복제, 분할, 가지치기 및 수정을 수행하는 densification 과정을 거쳐 파라미터를 점진적으로 업데이트하며, 이를 통해 장면을 더 정밀하게 표현할 수 있다.

2.2 3DGS with DBSCAN

DBSCAN은 반경 Epsilon내 최소 Min Points 이상의 이웃 포인트가 있는 영역을 클러스터로 식별하는 기법으로, 데이터의 밀도를 기반으로 클러스터를 형성한다. 비교적 밀도가 낮은 영역의 값들을 노이즈로 분류하고 제거하여 메모리 사용을 효율적으로 줄이고 분포가 불규칙한 데이터에 유연하게 동작하는 장점이 있다.

DBSCAN 수행 시점까지 학습 되어있던 가우시안에 클러스터링을 수행해 파라미터가 유사한 가우시안을 하나의 클러스터로 묶는다. 이때, 각 클러스터에 속해있는 가우시안들을 하나의 큰 가우시안으로 변경하기 위해서는 대푯값을 정의하는 과정이 요구된다. 본 연구에서는 가우시안을 표현하는 데 필요한 위치, 색상, 투명도의 대푯값을 클러스터 내부 가우시안들의 평균값으로 설정하였다. 이는 한 클러스터 내의 가우시안들이 전체적으로 근사한 수치의 값을 가지고 있기 때문이다.



그림 2. 정량적평가를 위한 3DGS와 제안 모델이 생성한 이미지 비교

크기와 회전 대푯값은 변경된 위치에 대한 공분산 행렬을 계산한 후 분리하여 계산하였다. 이를 위해서 위치 대푯값과 클러스터 내부 가우시안의 위치값의 차이를 기반으로 공분산 행렬을 계산한다.

$$\Sigma = RSS^T R^T \quad (1)$$

$$\Sigma = Q\Lambda Q^T \quad (2)$$

이후 식 (1)과 같이 공분산을 R과 S로 분해한다. 이때 공분산 Σ 은 대칭 행렬이므로 식 (2)와 같이 고유값 분해를 수행할 수 있다. 이 때 R은 고유벡터 행렬 Q를 사용하고, S는 고유값 행렬 Λ 의 제곱근으로 정의된다. 이렇게 정의된 새로운 가우시안은 다시 3DGS의 방식을 통해 학습을 진행한다.

III. 실험

3.1 실험환경

본 논문에서는 성능평가를 위해 Tank&Temples^[5] 데이터셋의 Train과 Truck을 사용했다. 이미지 품질 비교를 위해 SSIM과 PSNR을 평가 기준으로 삼았다. SSIM은 두 이미지의 구조적 측면에서 GT 이미지와 렌더링 된 이미지의 유사도를 계산하며, 값이 1에 가까울수록 원본과의 일치도가 높음을 의미한다. PSNR은 픽셀 간 최대 신호 대 잡음비로, 값이 높을수록 이미지 품질이 우수함을 나타낸다. 모든

실험은 L4 GPU 환경에서 수행했으며, 기존 3DGS와 동일한 하이퍼 파라미터 설정을 사용하되, DBSCAN을 수행하는 시점을 별도로 정해두었다. DBSCAN을 1번 수행하는 경우엔 전체 3만 iteration 중 2만 번째에, 3번 수행하는 경우엔 3천 번째, 1만 번째, 3만 번째에 클러스터링이 이루어지도록 했다. 이와 함께 densification을 최종적으로 반복하는 iteration을 2만 5천 번째로 설정하여 DBSCAN 수행 이후 소실될 수 있는 학습된 디테일들을 최적화를 통해 다시 획득할 수 있도록 돕는다. 또한, DBSCAN의 epsilon은 0.1, min points은 5로 지정하여 실험을 진행했다.

표 1. 기존 알고리즘과 성능 및 메모리 사용량 비교

Method	Tanks&Temples			
	SSIM	PSNR	Mem	num
M-NeRF360 ^[1]	0.759	22.22	8.6MB	-
*3DGS ^[2]	0.841	23.14	411MB	1.78
Ours-1	0.887	25.57	387.6MB	1.63
Ours-3	0.885	25.21	384.4MB	1.62

* Reproduced

3.2 Results

가. Quantitative Results

Tanks & Temples 데이터셋에서 제안 모델(Ours-1, Ours-3)은 3DGS 대비 전반적으로 우수한 성능을 보였다. 여기서 1과 3은 DBSCAN을 수행한 횟수를 의미하며, SSIM

에서는 Ours-1과 Ours-3이 각각 0.887과 0.885를 기록하여 3DGS보다 높은 유사도를 보였다. 이는 제안 모델이 이미지의 구조적 유사성을 잘 보존했음을 나타낸다. PSNR 측면에서도 Ours-1과 Ours-3은 각각 25.57과 25.21로 3DGS보다 우수한 성능을 보였다. 또한, 메모리 사용량의 경우 제안 모델이 약 10% 감소된 메모리 사용량을 기록하면서 렌더링 성능을 증가시켜, 메모리 효율성에서도 경쟁력 있는 결과를 얻었다.

나. Qualitative Results

그림 2에서는 제안모델과 3DGS가 생성한 이미지를 GT와 비교하였다. 가우시안 개수를 의도적으로 DBSCAN을 통해 줄였음에도 큰 차이를 보이지 않는 것을 확인할 수 있으며, 이러한 결과는 메모리 감소를 위해 사용한 DBSCAN방법이 렌더링 학습을 크게 방해하지 않으며 메모리를 효율적으로 관리한다는 것을 의미한다.

IV. 결론 및 향후 연구 방향

본 연구는 3DGS의 주요 한계점인 관측이 부족한 영역에서 불필요한 노이즈가 발생하는 문제를 해결하기 위해 DBSCAN Clustering을 적용하여 3DGS를 개선하는 새로운 알고리즘을 제안했다. 제안 모델은 10%가량 메모리효율을 올렸을 뿐만 아니라 기존 3DGS에 비해 우수한 성능을 보였다. 하지만 densification 중단 시점, DBSCAN 적용 시점 등 여러가지의 파라미터들이 존재하기 때문에, 최적의 성능을 도출하기 위해서는 더 많은 실험이 필요하다. 또한 DBSCAN을 5회 이상 적용할 경우 렌더링 속도에 영향을 미칠 수 있다는 한계가 있다. 따라서 향후에는 추가 실험과 함께 DBSCAN을 더 적용시킬 수 있는 방법에 대해 연구할 예정이다.

참고문헌

- [1] 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," ECCV, 2020.
- [2] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," SIGGRAPH, 2023.
- [3] J. C. Lee, D. Rho, X. Sun, J. H. Ko, and E. Park, "Compact 3D Gaussian Representation for Radiance Field," CVPR, 2024.
- [4] Ester, Martin, et al, "A density-based algorithm for discovering clusters in large spatial databases with noise," kdd, vol.96, no. 34, pp.226-231, 1996.
- [5] A. Knapitsch, J. Park, Q.-Y. Zhou, and V. Koltun, "Tanks and Temples: Benchmarking Large-Scale Scene Reconstruction," ACM Transactions on Graphics (ToG), vol. 36, no. 4, pp. 1-13, 2017.