Journal of Next-generation Convergence Information Services Technology Vol.15, Suppl.1, November (2026) pp.15-23
ISSN 2384-101X(p) 2672-1163(e) *http://dx.doi.org/10.29056/jncist.2026.11.02*

Train-View Reliability Map을 이용한 투명 객체 Gaussian Splatting 손실 재가중

Train-View Reliability Map-Based Loss Reweighting for Transparent Object Gaussian Splatting

이은서[1], 황보택근[2]*

EunSeo Lee[1], Taegkeun Whangbo[2]*

요 약

3D Gaussian Splatting(3DGS)은 다중 시점 RGB 영상을 활용한 새로운 시점 합성에서 우수한 성능을 보이나, 투명 객체의 시점 의존적 외관으로 인해 재구성 잔차가 공간적으로 불균일하게 나타날 수 있다. 본 연구는 RGB 재구성 잔차, 렌더링 깊이 경계, 시점 간 렌더링 일관성을 결합한 학습 시점 신뢰도 맵의 영역 구분 특성과 손실 재가중 효과를 구분하여 분석하였다. 동일한 3DGS 체크포인트에서 1,000회 추가 학습한 세 장면을 대상으로 균일 가중치와 복합 단서 기반 가중치를 각각 세 개 시드에서 비교하고, 잔차 기반 가중치를 단일 시드의 보조 조건으로 평가하였다. 신뢰도 맵과 평가용 투명 객체 영역의 대응 정도는 투명 의자와 격자형 프레임에서 각각 AUROC 0.837과 0.813으로 나타나, 학습 영상만으로 구성한 복합 단서가 투명 객체 주변의 특징적 영역을 구분하는 데 활용될 가능성을 확인하였다. 반면 복합 단서 가중치 적용에 따른 전체 PSNR 변화는 각각 -0.00636 dB와 +0.00180 dB로, 영역 구분 성능이 미관측 시점 렌더링 개선으로 직접 연결되지는 않았다. 이러한 분석을 통해 신뢰도 맵의 진단적 유효성과 최적화 가중치로서의 유효성이 상이한 평가 대상임을 확인하였으며, 투명 객체 학습에서는 난이도 영역의 식별과 해당 영역의 손실 증폭을 구별하여 설계해야 함을 제시한다.

핵심어 : 3D Gaussian Splatting, 투명 객체, 새로운 시점 합성, 손실 재가중, 신뢰도 맵

Abstract

Three-dimensional Gaussian Splatting (3DGS) enables efficient novel-view synthesis from multiview RGB images, while transparent objects can produce spatially nonuniform reconstruction residuals. This study separately investigates the region-discrimination behavior and loss-reweighting effect of a train-view reliability map composed of RGB residuals, rendered depth discontinuities, and cross-view rendering consistency. Uniform and multi-cue weighting were compared across three scenes and three random seeds

1 Department of Computer Engineering, Gachon University, Seongnam, Korea [Ph.D. Student]
e-mail: dmstj5308@gachon.ac.kr
2 Department of Computer Engineering, Gachon University, Seongnam, Korea [Professor]
e-mail: tkwhangbo@gachon.ac.kr (Corresponding author)

*이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원을 받아 수행된 연구임(IITP-2026-RS-2020-II201846, 기여율 100%).

Received(October 10, 2026), Review Result(1st: October 27, 2026), Accepted(November 12, 2026), Published(November 30, 2026)

© 2026 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/4.0/.

using a matched 1,000-step continuation from the same 3DGS checkpoint; residual-only weighting was assessed as an auxiliary single-seed condition. The reliability map showed correspondence with evaluation-only transparent-object regions, yielding AUROC values of 0.837 and 0.813 for the transparent chair and slatted-frame scenes, respectively. However, the changes in overall held-out PSNR were -0.00636 dB and +0.00180 dB, respectively. These observations distinguish the diagnostic utility of a train-view map from its effectiveness as an optimization weight. The results identify the need to evaluate where training difficulty is localized separately from whether amplifying the corresponding loss improves novel-view rendering.

Keyword : 3D Gaussian Splatting, Transparent Object, Novel-View Synthesis, Loss Reweighting, Reliability Map

1. 서론

Neural Radiance Fields(NeRF)는 다중 시점 영상과 카메라 자세를 이용한 새로운 시점 합성의 기반을 제시하였으며 [1], 3D Gaussian Splatting(3DGS)은 3차원 Gaussian primitive와 미분 가능한 래스터화를 이용해 효율적인 장면 표현 및 렌더링을 수행한다 [2]. 그러나 투명 객체에서는 투과와 시점 의존적 외관 변화에 따라 관측 RGB와 실제 표면의 대응 관계가 불명확해지며, 일부 학습 픽셀에 큰 재구성 잔차가 지속될 수 있다.

기존 3DGS는 영상 복원 오차와 구조적 유사도에 기초한 손실을 이용해 최적화한다 [3]. 한편 영상 복원 연구에서는 오차가 크게 나타나는 성분을 선택적으로 강조하는 학습 전략이 제안되었으나 [4], 높은 오차에 더 큰 가중치를 부여하는 것이 새로운 시점의 성능 개선을 보장하지는 않는다. 투명 객체 영역은 영상 내 면적이 작거나 시점 간 외관 차이가 크므로, 픽셀별 잔차를 바탕으로 한 재가중 전략의 실제 효과를 분리하여 검증할 필요가 있다.

투명 객체의 Gaussian Splatting 연구에서는 표면 법선 및 조명 관련 사전 정보 [5], 환경 표현과 역렌더링 [6] 등을 활용하여 복원 문제에 접근하였다. 또한 Gaussian 표현의 주파수 특성과 시점 적응성을 다룬 연구에서는 렌더링 오차와 최적화 구조 간의 관계를 분석하였다 [7-9]. 이에 비해 본 연구는 별도의 깊이, 법선 또는 객체 마스크를 학습 입력으로 요구하지 않고, 학습 영상에서 계산되는 복합 단서를 픽셀별 손실 가중치로 사용했을 때의 효과를 분석한다.

본 연구에서는 복합 단서 기반 신뢰도 맵을 적용한 결과를 영역 구분 특성과 렌더링 최적화 효과의 두 측면에서 분석하였다. 첫째, 학습 영상에서 구성한 점수가 평가용 투명 객체 영역과 어느 정도 대응하는지 정량화하였다. 둘째, 동일 초기 체크포인트와 학습 예산을 적용하여 신뢰도 맵 기반 재가중에 따른 전체 영상 및 투명 객체 영역의 PSNR과 SSIM 변화를 비교하였다. 이를 통해 영상적 난이도 단서의 식별 가능성과 손실 가중치 조정의 효용 사이에 존재하는 차이를 실험한다.

2. 관련 연구

2.1 Gaussian Splatting과 투명 객체 렌더링

3DGS는 위치, 공분산, 불투명도 및 색상 계수를 갖는 Gaussian primitive를 최적화하며 [2], 2D Gaussian Splatting(2DGS)은 표면에 정렬된 2차원 primitive를 활용하여 표면 기하의 정확성을 고려한다 [10]. 이후 Mip-Splatting은 다중 해상도에서 나타나는 aliasing을 완화하였으며 [8], Scaffold-GS는 장면 구조와 시점에 따라 Gaussian 속성을 조정하는 방식을 제안하였다 [9]. 투명 객체에서는 관측 광선의 투과와 굴절로 기하·외관의 모호성이 증가하며, 이를 보완하기 위한 법선 및 조명 사전 정보 기반 접근 [5]과 역렌더링 기반 접근 [6]이 보고되었다.

2.2 영상 잔차 기반 가중치와 연구 범위

영상 재구성 과정에서 상대적으로 어려운 성분을 강조하는 접근으로는 Focal Frequency Loss가 있으며, 이 방법은 공간적 픽셀 재가중이 아니라 주파수 성분의 손실 기여를 조정한다 [4]. Gaussian Splatting에서도 FreGS는 주파수 정규화를 이용해 Gaussian 최적화 과정을 개선하였다 [7]. 따라서 높은 오차를 강조하는 전략 자체만으로 본 연구의 차별성을 주장하기는 어렵다. 본 연구는 RGB 재구성 잔차, 렌더링 깊이 경계, 시점 간 렌더링 일관성의 결합 효과를 동일한 초기 체크포인트와 추가 학습 예산에서 평가한다. 복합 단서 점수는 실제 기하 오차 또는 광학적 굴절량의 추정값으로 해석하지 않는다.

3. 연구 방법

3.1 입력 및 기본 렌더링

학습 데이터는 V개 시점의 RGB 영상 I_v, 카메라 내부 행렬 K_v, 회전 행렬 R_v 및 평행 이동 벡터 t_v로 구성된다.

$$D_{tr=}(I_v,K_v,R_v,t_v)_{v=1}^{V} \tag{1}$$

3DGS 모델의 매개변수를 Θ로 두며, 학습 시점 v에 대한 렌더링 RGB 영상, 깊이 및 알파 값을 각각 $\hat{I}_v$, $\hat{D}_v$, $\hat{A}_v$로 나타낸다. 신뢰도 맵은 학습 과정에서 계산되는 보조 점수이며, 추론 시 별도 센서 입력을 요구하지 않는다.

3.2 복합 단서 기반 신뢰도 맵

첫 번째 단서는 관측 RGB 영상과 렌더링 결과의 픽셀별 평균 절댓값 잔차로 정의한다.

$$r_v(p) = \frac{1}{3} \| |I_v(p) - \hat{I}_v(p) \|_1 \tag{2}$$

두 번째 단서는 렌더링 깊이의 국소 기울기 크기이다.

$$d_v(p) = \| \nabla \hat{D}_v(p) \|_2 \tag{3}$$

이 값은 참 깊이와 예측 깊이 사이의 오차가 아니라 렌더링 깊이가 급격하게 변화하는 영상 위치를 나타낸다. 세 번째 단서는 인접한 학습 시점의 렌더링 영상을 카메라 자세와 렌더링 깊이를 이용하여 현재 시점으로 변환한 후 계산하는 시점 간 일관성 오차이다.

$$c_v(p) = median_{u \in N(v)} \| \hat{I}_v(p) - W_{u \to v}(\hat{I}_u; \widehat{D}_v, K, R, t)(p) \|_1 \tag{4}$$

여기서 $N(v)$는 인접 학습 시점의 집합이며, $W_{u \to v}$는 카메라 보정 정보를 이용한 영상 변환 연산이다. 변환 결과가 영상 범위를 벗어나거나 알파가 낮은 픽셀은 일관성 계산에서 제외한다. 해당 연산에는 관측되지 않은 시험 영상이나 정답 기하가 사용되지 않는다.

세 단서는 학습 영상별 5백분위수와 95백분위수를 이용하여 정규화한다.

$$\tilde{x} = clip\left(\frac{x - Q_{0.05}(x)}{Q_{0.95}(x) - Q_{0.05}(x) + \epsilon}, 0, 1 \right) \tag{5}$$

$$U_v(p) = \frac{\tilde{r_v}(p) + \lambda_d \tilde{d_v}(p) + \lambda_c \tilde{c_v}(p)}{1 + \lambda_d + \lambda_c} \tag{6}$$

λ_d와 λ_c는 깊이 경계 및 시점 간 일관성 단서의 상대적 기여를 조정하는 계수이다. $U_v(p)$는 영상 공간에서의 학습 난이도를 나타내기 위한 점수이며, 물리적 굴절 특성이나 실제 기하 오차의 직접적인 추정값으로 해석하지 않는다.

3.3 손실 함수의 제한적 재가중

기본 3DGS 손실로 초기 학습을 수행한 뒤 신뢰도 맵을 주기적으로 갱신하며, 가중치 계산 시 기울기 전파를 차단한다. 정규화된 픽셀별 가중치는 다음과 같다.

$$w_v(p) = \frac{clip(1 + \beta sg(U_v(p)), 1, w_{\max})}{\frac{1}{|\Omega_v|} \sum_{q \in \Omega_v} clip(1 + \beta sg(U_v(q)), 1, w_{\max})} \tag{7}$$

Ω_v는 유효 픽셀 집합이며, sg는 stop-gradient 연산을 의미한다. 가중치 평균을 1로 정규화함으로써 가중치 적용에 따른 전체 손실 규모의 변화를 제한한다. 최종 목적함수는 다음과 같다.

$$L = \sum_{v} \left[\frac{1}{|\Omega_v|} \sum_{p \in \Omega_v} w_v(p) \| I_v(p) - \hat{I}_v(p) \|_1 + \lambda_{ssim} L_{D-SSIM}(I_v, \hat{I}_v) \right] \tag{8}$$

4. 실험 구성

4.1 비교 방법 및 학습 조건

실험에서는 [표 1]과 같이 균일 가중치 기반의 기존 3DGS(B0), RGB 잔차만 사용한 가중치(B1), RGB 잔차와 깊이 경계를 결합한 가중치(B3), 그리고 RGB 잔차·깊이 경계·시점 간 일관성을 결합한 가중치(B4)를 구분하였다. 무작위 가중치(B2)는 가중치 분산의 영향을 확인하기 위한 비교 조건으로 설계하였다. 다만 본 논문의 정량 예비 결과는 실제 측정값이 확보된 B0, B1 및 B4에 한정한다.

[표 1] 손실 재가중 방식별 비교 방법 및 목적
[Table 1] Comparison Methods and Objectives of Loss Reweighting Strategies

구분	가중치 방식	비교 목적
B0	균일 가중치(기본 3DGS)	기본 성능
B1	RGB 잔차 기반	단일 잔차 가중치의 영향
B2	무작위 가중치	가중치 분산에 대한 대조 조건
B3	RGB 잔차 + 깊이 경계	깊이 단서의 기여 분석 조건
B4	RGB 잔차 + 깊이 경계 + 시점 간 일관성	복합 단서 기반 재가중

기존 3DGS의 7,000회 학습 체크포인트에서 시작하여 각 조건에 동일한 1,000회 추가 학습을 적용하였다. 학습 영상, 카메라 분할, 해상도, 초기 체크포인트 및 추가 학습 예산을 동일하게 유지하였다. B0와 B4는 세 개 시드에서 반복 수행하였으며, B1은 하나의 시드에서 진단적 비교를 수행하였다. 따라서 B1과 B4의 우열에 대한 반복 실험 기반 결론은 제시하지 않는다.

4.2 평가 장면 및 지표

평가에는 투명 의자(잔디 배경), 투명 격자형 프레임 및 불투명 목재 의자 장면을 사용하였다. 전체 미관측 시점 영상의 PSNR과 SSIM을 측정하고, 두 투명 장면에 대해서는 평가용 투명 객체 마스크 내부의 PSNR을 추가 계산하였다. 마스크는 학습 가중치 계산이나 모델 입력에 사용하지

않았다. 신뢰도 맵과 투명 영역의 대응 정도는 AUROC와 평균 정밀도(AP)로 구분하여 평가하였다. 이들 분류 지표는 새로운 시점의 렌더링 성능이 아니라 학습 시점에서의 영역 식별 성향을 분석하는 보조 지표이다.

5. 실험 결과

5.1 새로운 시점 렌더링 성능

[표 2]는 1,000회 추가 학습 이후 B0, B1 및 B4의 새로운 시점 렌더링 성능을 나타낸다. B0와 B4의 값은 세 시드의 평균과 표준편차이며, B1은 단일 시드의 결과이다.

[표 2] 손실 재가중 방식에 따른 새로운 시점 렌더링 성능 비교

[Table 2] Comparison of Novel-View Rendering Performance across Loss Reweighting Methods

평가 장면	B0 PSNR (dB)	B4 PSNR (dB)	B0 SSIM	B4 SSIM	B1 PSNR / SSIM (seed 0)
투명 의자 (잔디 배경)	25.5470 ± 0.0035	25.5406 ± 0.0034	0.904992 ± 0.000030	0.904849 ± 0.000024	25.5347 / 0.904725
투명 격자형 프레임	24.9738 ± 0.0044	24.9756 ± 0.0035	0.854297 ± 0.000038	0.854130 ± 0.000040	24.9804 / 0.854026
불투명 목재 의자	25.5766 ± 0.0080	25.5904 ± 0.0079	0.922298 ± 0.000174	0.922347 ± 0.000183	25.6028 / 0.922550

투명 의자에서 B4의 PSNR은 B0보다 0.00636 dB 낮았고, SSIM 역시 0.000143 감소하였다. 투명 격자형 프레임에서는 PSNR이 0.00180 dB 증가했으나 SSIM은 0.000168 감소하였다. 불투명 목재 의자에서 B4의 PSNR은 0.01378 dB 증가하였다. 따라서 추가 학습 구간에서 관찰된 미세한 변화는 투명 장면에서 일관된 개선 경향을 형성하지 않았다.

5.2 투명 객체 영역의 렌더링 품질

평가용 마스크로 구분한 투명 객체 영역의 PSNR을 [표 3]에 제시하였다.

[표 3] 투명 객체 영역에서의 PSNR 성능 비교

[Table 3] Comparison of PSNR Performance in Transparent Object Regions

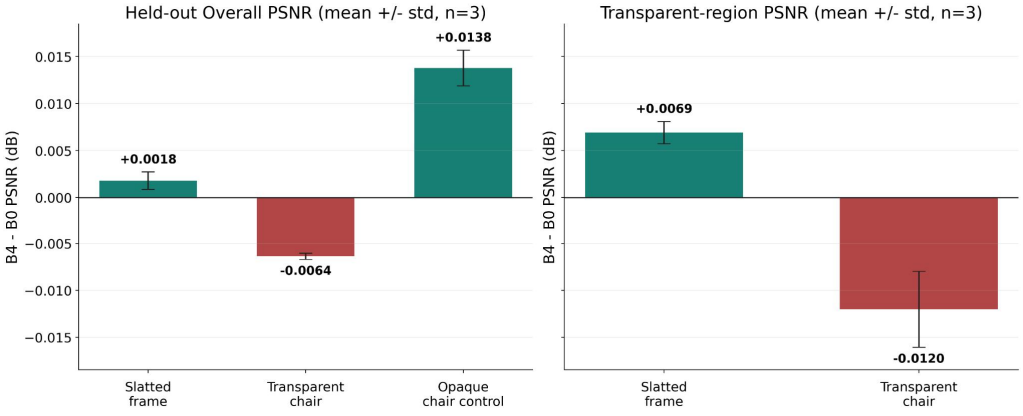
투명 장면	B0 영역 PSNR (dB)	B4 영역 PSNR (dB)	B4 − B0 (dB)
투명 의자(잔디 배경)	27.0484 ± 0.0014	27.0364 ± 0.0031	−0.01201
투명 격자형 프레임	22.9452 ± 0.0029	22.9521 ± 0.0026	+0.00693

투명 의자에서는 전체 영상과 투명 객체 영역의 PSNR이 모두 감소하였다. 격자형 프레임에서는 객체 영역 PSNR이 소폭 증가했으나, 변화 폭은 0.01 dB 미만이었다. 이러한 결과는 복합 단서 기반 가중치가 투명 객체 영역에서 안정적인 이득을 제공했다고 판단하기 어렵다는 점을 보여준다.

5.3 신뢰도 맵의 영역 식별 특성

신뢰도 맵과 평가용 투명 객체 마스크의 대응 관계를 분석한 결과, 투명 의자의 AUROC와 AP는 각각 0.837과 0.373, 투명 격자형 프레임에서는 각각 0.813과 0.275로 나타났다. 두 장면에서 AUROC가 0.8을 상회한다는 점은 복합 단서 점수가 투명 객체 영역과 비투명 영역을 구분하는 순위 정보를 일정 수준 포함함을 보여준다. 다만 AP는 양성 픽셀 비율에 영향을 받으며, 이러한 결과만으로 투명성의 물리적 원인이나 기하 오차를 식별했다고 볼 수는 없다. 따라서 본 결과는 복합 단서의 투명 영역 연관성을 나타내는 진단적 관찰로 해석한다.

신뢰도 맵의 진단 결과와 렌더링 최적화 결과를 함께 비교하면, 영역 구분 능력과 손실 재가중의 효과는 별개의 성질임을 확인할 수 있다. 투명 의자에서는 AUROC 0.837의 영역 대응이 관찰되었지만 B4의 전체 및 객체 영역 PSNR은 각각 감소하였으며, 격자형 프레임에서는 AUROC 0.813과 함께 객체 영역 PSNR이 소폭 증가하였으나 SSIM은 감소하였다. [그림 1]의 장면별 변화 역시 두 투명 장면에서 동일한 개선 경향이 나타나지 않음을 보여준다. 이러한 차이는 구분 점수의 크기를 곧바로 최적화 가중치에 연결하는 전략을 별도로 검증해야 함을 뒷받침한다.



[그림 1] 복합 단서 기반 손실 재가중(B4)과 균일 가중치(B0)의 PSNR 차이: (a) 전체 영상, (b) 투명 객체 영역

[Fig. 1] Differences in novel-view PSNR between multi-cue loss reweighting (B4) and uniform weighting (B0): (a) overall image, (b) transparent object region.

6. 결론

본 연구는 3DGS의 학습 영상에서 계산한 RGB 재구성 잔차, 렌더링 깊이 경계 및 시점 간 일관성을 결합하여 신뢰도 맵을 구성하고, 해당 맵의 영역 구분 특성과 손실 재가중 효과를 분석하였다. 두 투명 장면에서 신뢰도 맵은 평가용 투명 객체 영역과 AUROC 0.837 및 0.813의 대응 정도를 나타내어, 추가 정답 정보를 학습에 투입하지 않고도 영상적 난이도와 연관된 공간 정보를 추출할 가능성을 보여주었다. 그러나 동일 체크포인트와 학습 예산에서 재가중에 따른 전체 PSNR 변화는 각각 -0.00636 dB와 +0.00180 dB로 나타났으며, 렌더링 개선은 일관되지 않았다. 이 결과를 통해 투명 객체 주변의 특징적 픽셀을 식별하는 과정과 그 픽셀의 손실을 증폭하는 최적화 과정은 구분하여 평가해야 함을 확인하였다. 본 연구는 신뢰도 맵의 진단적 활용 가능성과 단순 가중치 증폭 방식의 한계를 동시에 제시하며, 향후 단서별 기여 분석과 가중치 갱신 전략 설계를 위한 실험적 근거를 제공한다.

References

[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," in *European Conference on Computer Vision*, Glasgow, UK, Aug. 23-28, 2020, pp. 405-421, doi: 10.1007/978-3-030-58452-8_24.

[2] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," *ACM Transactions on Graphics*, vol. 42, no. 4, Art. no. 139, pp. 1-14, Jul. 2023, doi: 10.1145/3592433.

[3] 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, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[4] L. Jiang, B. Dai, W. Wu, and C. C. Loy, "Focal Frequency Loss for Image Reconstruction and Synthesis," in *IEEE/CVF International Conference on Computer Vision*, Montreal, Canada, Oct. 11-17, 2021, pp. 13919-13929, doi: 10.1109/ICCV48922.2021.01366.

[5] M. Li, P. Pang, H. Fan, H. Huang, and Y. Yang, "TSGS: Improving Gaussian Splatting for Transparent Surface Reconstruction via Normal and De-lighting Priors," in *33rd ACM International Conference on Multimedia (MM '25)*, Dublin, Ireland, Oct. 27-31, 2025, pp. 7220-7229, doi: 10.1145/3746027.3754548.

[6] L. Huang et al., "TransparentGS: Fast Inverse Rendering of Transparent Objects with Gaussians," *ACM Transactions on Graphics*, vol. 44, no. 4, pp. 1-17, Aug. 2025, doi: 10.1145/3730892.

[7] J. Zhang, F. Zhan, M. Xu, S. Lu, and E. Xing, "FreGS: 3D Gaussian Splatting with Progressive Frequency Regularization," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 16-22, 2024, pp. 21424-21433, doi: 10.1109/CVPR52733.2024.02024.

[8] Z. Yu, A. Chen, B. Huang, T. Sattler, and A. Geiger, "Mip-Splatting: Alias-Free 3D Gaussian Splatting," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, Jun. 16-22, 2024, pp. 19447-19456, doi: 10.1109/CVPR52733.2024.01839.

[9] T. Lu et al., "Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, Jun. 16-22, 2024, pp. 20654-20664, doi: 10.1109/CVPR52733.2024.01952.

[10] B. Huang, Z. Yu, A. Chen, A. Geiger, and S. Gao, "2D Gaussian Splatting for Geometrically Accurate Radiance Fields," in *SIGGRAPH 2024 Conference Papers*, Denver, CO, USA, Jul. 27-Aug. 1, 2024, Art. no. 32, doi: 10.1145/3641519.3657428.