

Supplemental Material for GS³: Efficient Relighting with Triple Gaussian Splatting

ZOUBIN BI, YIXIN ZENG, CHONG ZENG, FAN PEI, XIANG FENG, KUN ZHOU, and HONGZHI WU, State Key Lab of CAD&CG, Zhejiang University, China

ACM Reference Format:

Zoubin Bi, Yixin Zeng, Chong Zeng, Fan Pei, Xiang Feng, Kun Zhou, and Hongzhi Wu. 2024. Supplemental Material for GS³: Efficient Relighting with Triple Gaussian Splatting. In *SIGGRAPH Asia 2024 Conference Papers (SA Conference Papers '24)*, December 3–6, 2024, Tokyo, Japan. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3680528.3687576>

1 ABLATIONS

1.1 Angular Gaussian vs. MLP

In Fig. 3, we replace the appearance function with an MLP of {30, 100, 600, 100, 30}. The input of the MLP are ω_i , ω_o and a 32D latent vector, and the output is a shading color. This MLP is pre-trained on 100 SGGX BRDFs with random parameters [Heitz et al. 2015]. As shown in the figure, Solely using a dedicated appearance MLP fails to model complex all-frequency appearance, unlike our representation.

1.2 Shadow Splatting vs. MLP

We train our representation without using the shadow splatting process (Sec. 4.2 in our paper), and directly adopt the refinement MLP alone to represent the shadow function. As shown in the figure, a pure MLP-based approach does not produce high-quality shadows as ours. Moreover, such an approach is prone to over-fitting (i.e., less generalizable than shadow splatting).

1.3 Shadow Refinement

We train our representation using the shadow value directly computed from splatting without the refinement MLP (Sec. 4.2 in our paper). As shown in Fig. 3, while the main features are preserved, the shadows are not as detailed as our pipeline with the refinement step.

1.4 MLP for Other Effects

We train our representation with the MLP for approximating other effects removed (Sec. 4.3 in our paper). In Fig. 3, we can observe noisy results where effects like interreflections should be presented. Other components in the representation tend to compensate for the

Authors' address: Zoubin Bi; Yixin Zeng; Chong Zeng; Fan Pei; Xiang Feng; Kun Zhou; Hongzhi Wu, hwu@acm.org, State Key Lab of CAD&CG, Zhejiang University, 866 Yuhangtang Rd., Hangzhou, China, 310058.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SA Conference Papers '24, December 3–6, 2024, Tokyo, Japan

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-1131-2/24/12

<https://doi.org/10.1145/3680528.3687576>

Table 1. Ablation studies of components in our representation. We list the average quantitative errors in SSIM, PSNR and LPIPS of all synthetic scenes to quantify the impact of ablated components. Our choices for individual components are shown in bold.

Ablation Variant	SSIM↑	PSNR↑	LPIPS↓
Full	0.9715	31.39	0.0355
w/o shadow splatting	0.9661	29.93	0.0391
w/o Φ (shadow refining)	0.9514	28.03	0.0556
w/o Ψ (other effects)	0.9707	31.30	0.0366
1 basis angular Gaussians	0.9655	29.70	0.0407
2 basis angular Gaussians	0.9694	30.93	0.0377
4 basis angular Gaussians	0.9709	31.25	0.0363
8 basis angular Gaussians	0.9715	31.39	0.0355
16 basis angular Gaussians	0.9721	31.50	0.0350
500 images	0.9670	30.61	0.0390
1,000 images	0.9698	31.09	0.0370
2,000 images	0.9715	31.39	0.0355

effects that would have been modeled by this missing MLP, making it more prone to over-fitting.

1.5 Number of Basis Angular Gaussians

In Fig. 1, we evaluate the impact of the number of basis angular Gaussians on a specific scene. And Tab. 1 lists the quantitative scores averaged over all synthetic objects/scenes. Note that the qualitative differences are more obvious in the figure, compared with the differences in average scores in the table. Our current selection of 8 is made after balancing under- and over-fitting of appearance.

1.6 Number of Input Images

Fig. 2 evaluates the impact of input image number. And the average errors are also reported in Tab. 1. More input images improve the reconstruction quality, due to more sampling in the view/lighting domain. We observe that it requires less input images to faithfully reconstruct scenes with simple geometry and appearance (e.g., HOTDOG) than more complex ones.

2 METRICS

We report detailed quantitative metrics here, due to the limited space in the main paper. First, Tab. 2 lists the errors of the comparison experiments with alternative approaches that take in environment-lit input images (corresponding to Fig. 12 in the main paper). Next, Tab. 3 reports the errors of the comparisons with [Zeng et al. 2023],

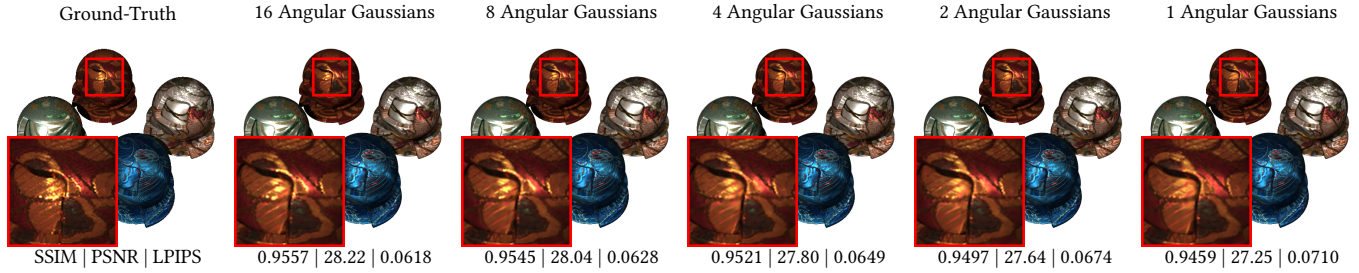


Fig. 1. Impact of the number of angular Gaussians. From the left image to right, the ground-truth, results from our representations trained with different numbers of basis angular Gaussians. Average errors in SSIM, PSNR and LPIPS are reported at the bottom of each related image.

Table 2. Quantitative metrics corresponding to Fig. 12 in the main paper. We list the quantitative errors in SSIM, PSNR and LPIPS.

Scene	Method	SSIM \uparrow	PSNR \uparrow	LPIPS \downarrow
Hotdog	GaussianShader	0.9439	29.25	0.0600
	GS-IR	0.9289	29.13	0.0816
	Relightable3DGaussian	0.9526	30.22	0.0555
	TensoIR	0.9590	31.68	0.0466
	Ours	0.9712	36.47	0.0342
	Lego	GaussianShader	0.8964	24.33
GS-IR		0.9230	27.66	0.0550
Relightable3DGaussian		0.9463	30.31	0.0446
TensoIR		0.9540	30.96	0.0356
Ours		0.9651	32.06	0.0302
MaterialBalls		GaussianShader	0.8900	24.31
	GS-IR	0.8756	23.41	0.0969
	Relightable3DGaussian	0.9045	25.34	0.0848
	TensoIR	0.9026	25.05	0.0765
	Ours	0.9459	28.01	0.0508

Table 3. Per-object/scene quantitative comparison with NRHints [Zeng et al. 2023] (corresponding to Fig. 9 in the main paper). We list the quantitative errors in SSIM, PSNR and LPIPS.

Scene	Method	SSIM \uparrow	PSNR \uparrow	LPIPS \downarrow
Drums	NRHints	0.9745	29.88	0.0294
	Ours	0.9714	30.45	0.0276
FurBall	NRHints	0.9522	35.06	0.0777
	Ours	0.9669	35.34	0.0534
Lego	NRHints	0.9583	29.90	0.0393
	Ours	0.9511	30.19	0.0468
Fish	NRHints	0.9140	31.28	0.1137
	Ours	0.9252	31.13	0.0866
Cluttered	NRHints	0.9280	29.61	0.0879
	Ours	0.9521	30.82	0.0696
Cat	NRHints	0.8560	27.34	0.1667
	Ours	0.8981	26.43	0.1381

Table 4. Per-object/scene quantitative errors of our approach for all synthetic and captured examples.

Scene	SSIM	PSNR	LPIPS
AnisoMetal	0.9494	27.25	0.0462
Drums	0.9714	30.45	0.0276
FurBall	0.9669	35.34	0.0534
Hotdog	0.9733	32.98	0.0295
Lego	0.9511	30.19	0.0468
Translucent	0.9740	32.34	0.0318
Cup	0.9905	32.74	0.0203
Egg	0.9833	31.84	0.0257
Fabric	0.9812	31.15	0.0317
MaterialBalls	0.9545	28.04	0.0628
Tower	0.9913	32.96	0.0150
Boot	0.8980	28.84	0.1013
Container	0.9745	36.65	0.016
Fox	0.9225	34.21	0.0745
Li'lOnes	0.9809	38.61	0.0182
Nefertiti	0.956	36.58	0.0434
Zhaojun	0.9321	32.08	0.1071

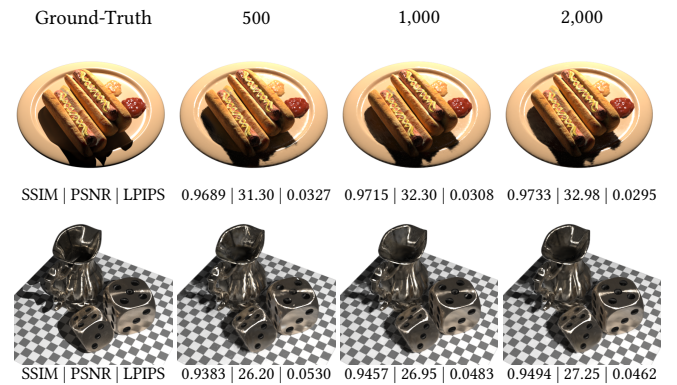


Fig. 2. Impact of the number of training images. From the left column to right, the ground-truths, and reconstruction results trained with 500, 1,000 and 2,000 images. Average errors in SSIM, PSNR and LPIPS are reported at the bottom of each related image.

