Disco-GS: Gaussian Splatting in Dynamic Color Lighting (Supplementary Material)

Ashish Kumar A. N. Rajagopalan
Indian Institute of Technology Madras, India
ashish.k.research@gmail.com, raju@ee.iitm.ac.in

All figures and tables in the supplementary are labeled with prefix S. The supplementary material is structured as follows:

1. Quantitative comparisons on additional videos
2. Optimization time
3. Qualitative Results on Intrinsic-NeRF and NeRF-OSR
4. Visualization of different components of Disco-GS
5. Discussion on Disco dataset acquisition
6. Hyperparameter sensitivity analysis
7. Visualization of Ablations

This supplementary material is also accompanied by a *'.mp4'* video file.

S1. Quantitative comparisons on additional videos

Table S1 shows the overall average PSNR, SSIM, and LPIPS values across the Disco dataset. Disco-GS outperforms all the competing methods. It should be noted that the metrics calculations have been performed with pseudo ground truth. Please refer to the accompanying *'.mp4'* file for qualitative comparisons.

Method	PSNR↑	SSIM↑	LPIPS↓
3DGS [2]	18.71	0.75	0.35
Wild-Gaussians [3]	16.17	0.66	0.49
Gaussians-Wild [6]	16.18	0.72	0.45
RNG [1]	5.11	0.03	0.87
IntrinsicNeRF [5]	10.45	0.36	0.62
NeRF-OSR [4]	16.29	0.79	0.45
Disco-GS	**20.68**	**0.85**	**0.21**

Table S1. Comparison of PSNR, SSIM, and LPIPS on Disco dataset.

S2. Optimization Time

Table S2 compares the per-scene average optimization time (on NVIDIA RTX 3090 GPU) of different methods on our Disco dataset.

Method	Training Time (hours)
3DGS [2]	0.28
Wild-Gaussians [3]	1.5
Gaussians-Wild [6]	2.5
RNG [1]	3
Disco-GS	1.1

Table S2. While 3DGS offers best optimization performance, it lacks a dedicated mechanism for handling color variations in the input images. RNG [1], on the other hand, follows a two-stage pipeline. Among the methods specifically designed to address color-tinted inputs [3, 6], Disco-GS stands out as the only approach that delivers both superior quantitative and qualitative results, while also achieving faster optimization.

S3. Qualitative Results on Intrinsic-NeRF and NeRF-OSR

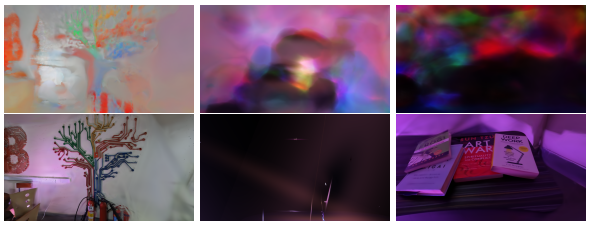

Figure 1. Additional comparisons with NeRF-based methods on scenes 3D Artwork-A, Casual and Books-A respectively. The first row shows results from Intrinsic NeRF [5], and the second row shows results from NeRF-OSR [4]. Both methods struggle with accurate scene reconstruction.

S4. Visualization of different components of Disco-GS

Fig. S2 visualizes $E(t)$, B, and $I(\underline{v})$ for five different complex scenes from Disco-GS outputs. The first two rows cor-

Figure S2. Visualization of Disco-GS outputs on the proposed Disco dataset (from left to right: input colored image, $E(t)$, B, and $I(\underline{v})$). Disco-GS remains robust under varying color light intensities, diverse scene contents, and different spatial distributions of colored illumination.

respond to the same scene captured at different times and from different viewpoints. This example involves *local colored lighting*: the first row is influenced by blue light, while the second row is affected by pink light (in the left region). Disco-GS accurately recovers $E(t)$ and successfully captures the transient reflection on the glass window as part of the transient component, while $I(\underline{v})$ remains free of this reflection. The third row shows a *global red colored lighting* scenario. Although the scene contains a red armchair, Disco-GS does not confuse object color with illumination; $E(t)$ cleanly captures the global red lighting. The fourth row presents a *global green lighting* case, where Disco-GS again correctly extracts the illumination in $E(t)$ and reconstructs a clean $I(\underline{v})$. Rows five and six depict *local colored lighting* with strong initial color influence that gradually decays (Note that in these two cases, the color strength of light is relatively light compared to the above rows examples.). Despite the complexity, Disco-GS reliably recovers both $E(t)$ and $I(\underline{v})$. These examples demonstrate the robustness and fidelity of Disco-GS in recovering the canonical scene appearance and transient colored illuminations, even under highly challenging artificial lighting conditions.

S5. Discussion on Disco dataset acquisition

The color of the light source was varied either manually or synchronized with music. For manual control, the color was changed using a mobile application, as shown in Fig. S3.

The application provides two tabs: the first offers a wide range of colors, while the second contains only a few options, including white light. During data capture, we varied the light color by randomly selecting positions on the color panel, either by tapping or dragging, resulting in non-systematic and natural color changes. For pseudo ground truth, we switched to the second tab and selected the white light option (top-right corner) at 100% brightness. We refer to this as *pseudo* ground truth because the actual brightness of the white light may vary slightly in practice.

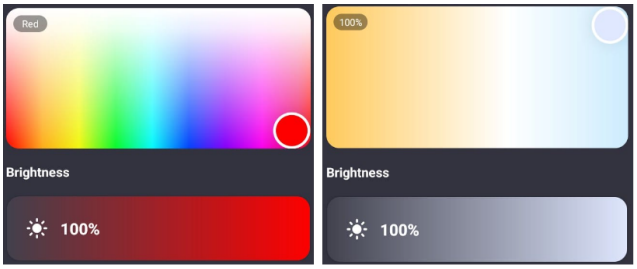

Figure S3. **Illumination Panel:** The left panel illustrates the color illumination variation interface. Illumination can be adjusted either continuously by sliding the circular cursor (shown in red) or discretely by tapping anywhere within the panel for an abrupt illumination change. The right panel shows the pseudo ground-truth illumination capture, obtained by fixing the circular cursor at the top-right corner to emulate a white lighting condition. Note that while capturing color illuminations (left), both hue and brightness were varied with brightness randomly sampled between 50% and 100%, whereas for the pseudo ground truth, brightness was fixed at 100%.

S6. Hyperparameter sensitivity analysis

Disco-GS struggles in the presence of moving objects, as illustrated in Fig. S4.

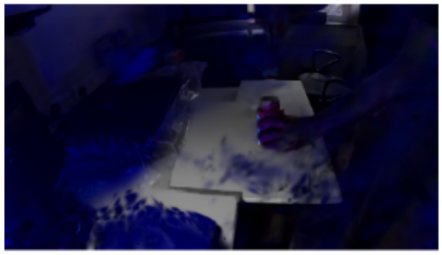

Figure S4. Failure case in presence of a dynamic object.

We additionally perform experiment using one at a time (OAT) method. The average PSNR/SSIM values under different settings are: (i) $\lambda_1 = \{0.7, 0.8, 0.9, 0.95\}$: 21.64/22.88/23.17/23.17; (ii) $\lambda_2 = \{0.3, 0.4, 0.6, 0.7\}$: 22.97/22.98/23.11/23.11; (iii) $\lambda_3 = \{0.7, 0.75, 0.85, 0.9\}$: 23.01/23.11/23.15/23.12; (iv) $\lambda_4 = \{0.0015, 0.002, 0.08, 0.09\}$: 23.18/23.17/23.01/23.03. Disco-GS is less sensitive to λ_4; mildly sensitive to λ_2 and λ_3. It is more sensitive to λ_1 as it aids in preserving structures in J_{can}.

S7. Visualization of Ablations

Figures S5–S13 visualize the novel-view renderings produced by different configurations used in our ablation study (Sec. 5 of the main paper). Except for the full loss $\mathcal{L}$, all other settings fail on at least one scene or introduce view inconsistencies. This highlights the necessity of each component in $\mathcal{L}$ (Eq. 10 of the main paper).

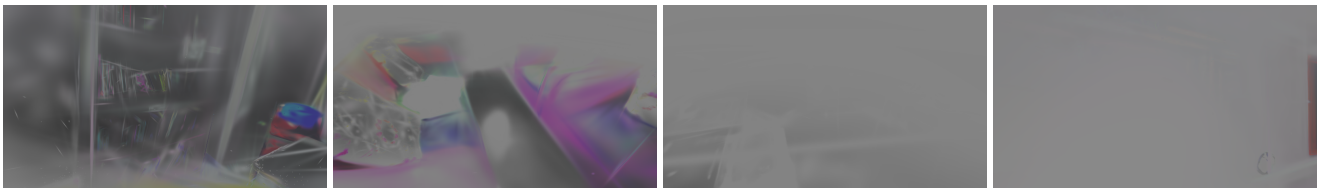

Figure S5. Using only $\mathcal{L}_{\text{photo}}$ (from left to right: *mini-library*, *newspaper-room*, *kitchen*, and *2D artwork*), the method fails to reconstruct all the scenes.

Figure S6. When only $\mathcal{L}_{\text{photo}}$ and $\mathcal{L}_{\text{st}}$ are used, the method fails to reconstruct the *book* scene (first two images rendered from different views). For the *kitchen* scene (last two images rendered from different views), it introduces black color artifacts on the window and wall (highlighted in blue and green).

Figure S7. Without $\mathcal{L}_{\text{st}}$, the method fails across all scenes (from left to right: *mini-library*, *newspaper-room*, *books*, *workdesk*).

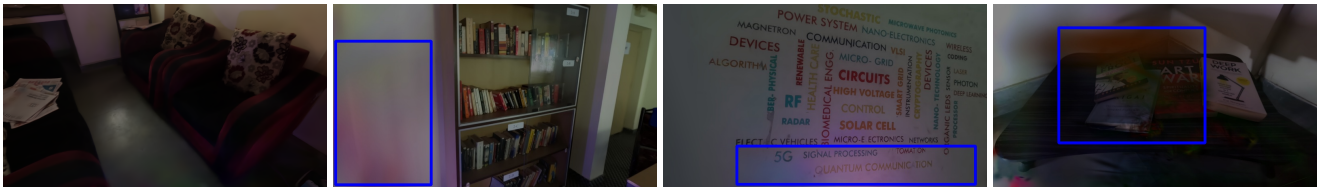

Figure S8. When $\eta = 1$, the method successfully reconstructs the *newspaper-room* scene (first image), but it introduces visible color artifacts in the other scenes, as highlighted by the blue bounding boxes.

Figure S9. Without $\mathcal{L}_{\text{col}}$, the method can reconstruct the *books* scene (first image), but it fails to reconstruct *mini-library* (second image). In addition, it introduces visible color artifacts (pink tint) in *kitchen* and *workdesk*, as highlighted by the blue bounding boxes.

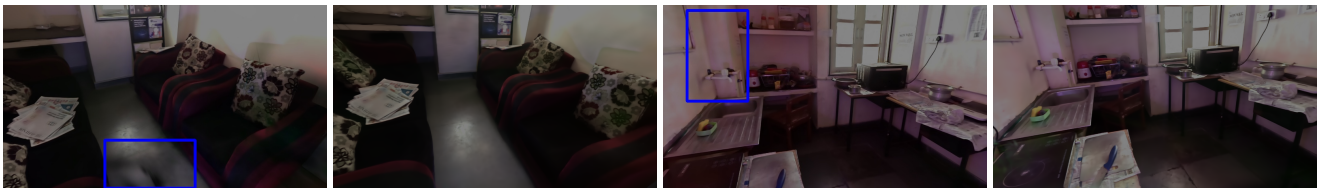

Figure S10. Without $\mathcal{L}_{\text{amb}}$, this configuration leads to inconsistencies in the rendered views. The figure shows the first two images from *newspaper-room* and the last two from *kitchen*. Observe the regions highlighted by the blue bounding boxes, corresponding areas across different views exhibit noticeable photometric variations.

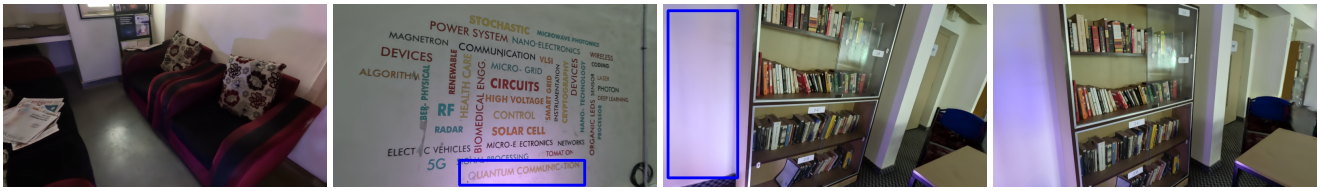

Figure S11. When $I(\underline{v})$ is used as the final output (i.e., removing B), the method is able to reconstruct the *newspaper-room* scene (first image). However, it introduces a slight pink tint in *2D artwork* (second image, highlighted in the blue bounding box) and causes noticeable color shifts in *mini-library* (last two images). The highlighted regions and their corresponding areas across different viewpoints reveal these inconsistencies.

Figure S12. Without $\mathcal{L}_{\text{tv}}$, we observe slight deviations in photometric values for *mini-library* (first two images) and *books* (last two images). The inconsistencies become apparent when comparing the corresponding regions highlighted in the blue bounding boxes across different views.

Figure S13. **Overall Loss $\mathcal{L}$.** Using the full loss $\mathcal{L}$ results in view-consistent renderings. The first two images show renderings of the *mini-library* scene from two different viewpoints, and the last two images show renderings of *books* from two views. Among all settings, $\mathcal{L}$ produces the most stable and consistent results.

References

[1] Jiahui Fan, Fujun Luan, Jian Yang, Milos Hasan, and Beibei Wang. Rng: Relightable neural gaussians. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 26525–26534, 2025. 1

[2] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023. 1

[3] Jonas Kulhanek, Songyou Peng, Zuzana Kukelova, Marc Pollefeys, and Torsten Sattler. Wildgaussians: 3d gaussian splatting in the wild. In *Advances in Neural Information Processing Systems*, pages 21271–21288. Curran Associates, Inc., 2024. 1

[4] Viktor Rudnev, Mohamed Elgharib, William Smith, Lingjie Liu, Vladislav Golyanik, and Christian Theobalt. Nerf for outdoor scene relighting. In *European Conference on Computer Vision (ECCV)*, 2022. 1

[5] Weicai Ye, Shuo Chen, Chong Bao, Hujun Bao, Marc Pollefeys, Zhaopeng Cui, and Guofeng Zhang. Intrinsicnerf: Learning intrinsic neural radiance fields for editable novel view synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 339–351, 2023. 1

[6] Dongbin Zhang, Chuming Wang, Weitao Wang, Peihao Li, Minghan Qin, and Haoqian Wang. Gaussian in the wild: 3d gaussian splatting for unconstrained image collections. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LXXVI*, page 341–359, Berlin, Heidelberg, 2024. Springer-Verlag. 1