Evolving High-Quality Rendering and Reconstruction in a Unified Framework with Contribution-Adaptive Regularization

Supplementary Material

In the supplementary material, implementation details are supplemented in Sec. 1 including the calculation of normals and SDF. In Sec. 2, we report more detailed experimental data to further analyze the CarGS. Sec. 3 shows more visualizations of qualitative results.

1. Implementation Details

1.1. Normal Calculation

Normal from the splat. Following the method described in [1], we align the direction of the minimum scale factor to the Gaussian normal n_i . The final normal map, corresponding to the current viewpoint, is then computed using α-blending:

$$N(p) = \sum_{i \in N} R_c^T n_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j), \tag{1}$$

R_c is defined as the rotation matrix responsible for converting coordinates from the camera frame to the world frame.

Normal from the depth. At a given pixel point p, we sample four neighboring points—up, down, left, and right—and project their corresponding depth values into 3D points, represented as $\{P_j \mid j = 1, \ldots, 4\}$, within the camera coordinate system. The normal of the local plane at p is computed as follows:

$$N_d(p) = \frac{(P_1 - P_0) \times (P_3 - P_2)}{|(P_1 - P_0) \times (P_3 - P_2)|}, \tag{2}$$

1.2. Depth Calculation

The distance between the plane and the camera center can be calculated as:

$$d_i = \left(R_c^T (\mu_i - T_c)\right) R_c^T n_i^T, \tag{3}$$

where T_c is the camera center in the world coordinates, and μ_i represents the center of the Gaussian G_i. Using α-blending, the final distance map corresponding to the current viewpoint is then generated:

$$\mathcal{D} = \sum_{i \in N} d_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j), \tag{4}$$

After obtaining the distance and normal of the plane through rendering, the corresponding depth map can be determined by intersecting rays with the plane:

$$D(p) = \frac{\mathcal{D}}{N(p) K^{-1} \tilde{p}}. \tag{5}$$

	Precision↑	Recall↑	F1↑
Barn	0.58/0.41	0.61/0.58	0.60/0.48
Caterpillar	0.33/0.30	0.61/0.59	0.43/0.40
Courthouse	0.13/0.10	0.26/0.23	0.17/0.14
Ignatius	0.79/0.77	0.77/0.74	0.78/0.75
Meetingroom	0.37/0.23	0.34/0.31	0.35/0.26
Truck	0.62/0.37	0.70/0.54	0.65/0.44
Mean	0.47/0.36	0.55/0.50	0.50/0.41

Table 1. **Comprehensive experiment to demonstrate the effectiveness of Lite-Geo.** We report precision, recall, and F1 scores on the Tanks and Temples dataset, comparing results with (left) and without (right) Lite-Geo.

where $p = [u, v]^T$ refers to the 2D coordinates on the image plane, $\tilde{p}$ represents the homogeneous form of p, and K denotes the intrinsic parameters of the camera.

2. Further Analysis of Lite-Geo

To better understand the influence of Lite-Geo, we conduct detailed experiments using the TNT dataset. As illustrated in Tab.1, the integration of Lite-Geo significantly improves both precision and recall metrics across all reconstructed scenes. Notably, the improvement in precision is more pronounced, which indicates that Lite-Geo effectively reduces errors in structural alignment.

Despite these improvements, certain challenges persist in specific scenarios. For example, in large-scale outdoor scenes such as Courthouse, and complex indoor environments like Meetingroom, both precision and recall remain relatively low. Large-scale outdoor environments often feature high variability in lighting, texture sparsity, and scale changes, which make precise feature matching and alignment more difficult. Similarly, complex indoor scenes with intricate geometries, occlusions, and cluttered details present significant challenges for achieving accurate reconstruction. These limitation serves as a key direction for our future work.

3. Qualitative Results

To further demonstrate the effectiveness of CarGS, we present additional visualizations. As shown in Fig.1, qualitative results on the TNT dataset are provided, showcasing the ground truth, rendering, normal map, and depth map.

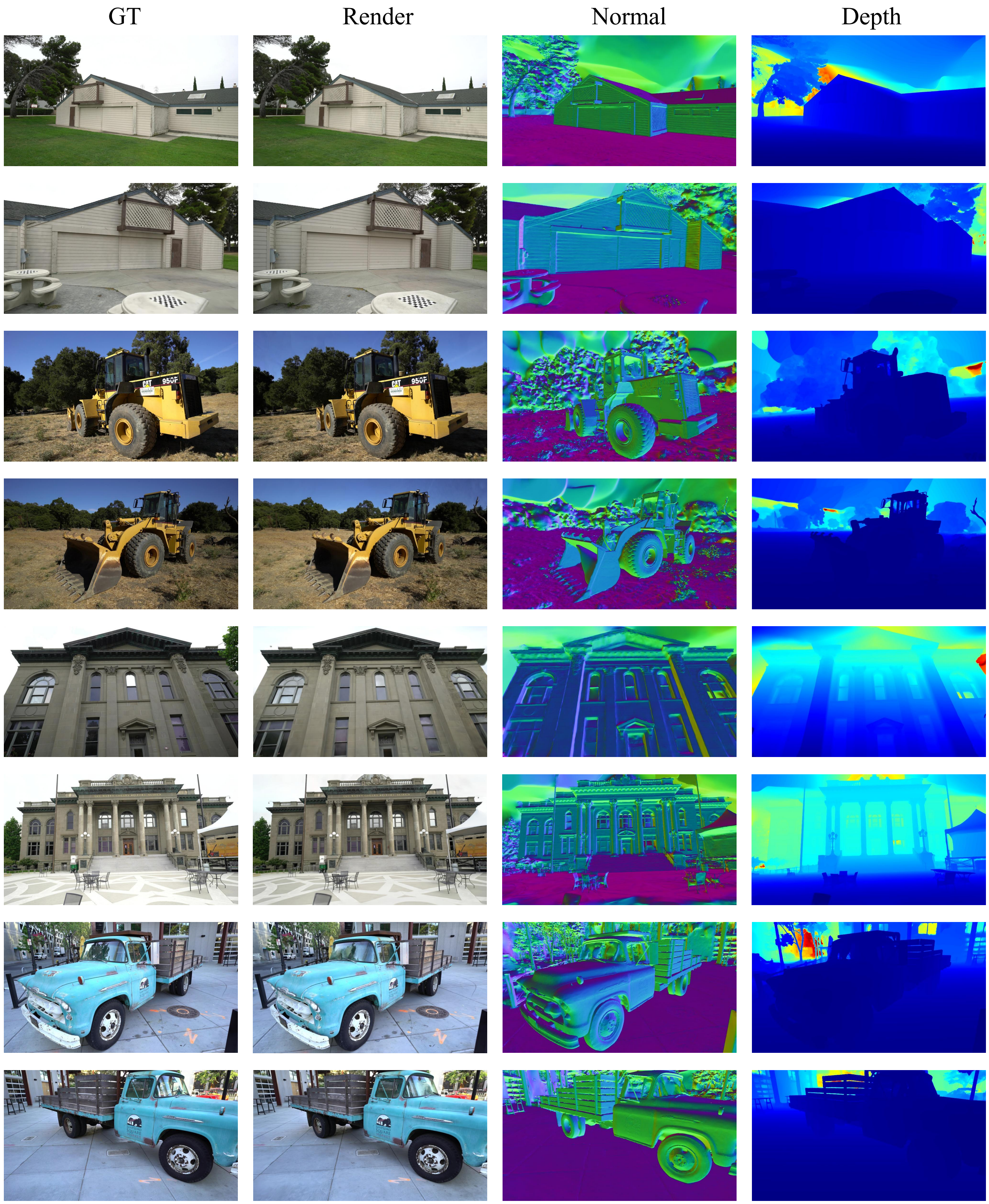

Figure 1. **Qualitative results of the CarGS.** Including the ground truth, rendered outputs, normal maps, and depth maps, offering a detailed visual representation of its performance.

References

[1] Danpeng Chen, Hai Li, Weicai Ye, Yifan Wang, Weijian Xie, Shangjin Zhai, Nan Wang, Haomin Liu, Hujun Bao, and Guofeng Zhang. Pgsr: Planar-based gaussian splatting for efficient and high-fidelity surface reconstruction. *arXiv preprint arXiv:2406.06521*, 2024. 1