Regularizing 3D Gaussian Splatting for Sparse Input

Quanhong Liu, Vanessa Yu
Department of Computer Science, University of Toronto

Motivation

- 3D Gaussian Splatting for Real-Time Radiance Field Rendering [1] is the latest and most powerful technique for novel-view synthesis of scenes. However, it cannot handle sparse inputs. We find out that the geometry regularization introduced by RegNeRF [2] to be a potential solution, but the difference in the model structure and rendering method makes it not easily applicable to 3D Gaussian Splatting, Hence, we want to investigate a way to apply geometry regularization in RegNeRF to 3D Gaussian Splatting to handle spare inputs.

Related Work

- The Neural Radiance Field (NeRF) is an emerging technique that employs a large Multi-Layer Perceptron (MLP) model for synthesizing realistic 3D scenes, optimizing a continuous 5D neural radiance field representation based on a set of input images [3]. Despite its success, NeRF has notable limitations. To address one of these shortcomings, RegNeRF [2] introduced a regularization approach specifically tailored for NeRF, aiming to improve geometry and color consistency with a reduced number of input views.

- 3D Gaussian Splatting [1] tackles challenges related to speed (in both training and rendering), scalability, and handling unbounded scenes. However, it utilizes a rasterizer for Gaussians in rendering which is different from the NeRF models, prompting the need for further exploration into regularization techniques in this context.

References

[1] Kerbl, B., Kopanas, G., Leimkuehler, T., and Drettakis, G. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Trans. Graph. 42, 4, Article 139 (August 2023), 14 pages. https://doi.org/10.1145/3592433, 2023

[2] Niemeyer, M., Barron, J., Mildenhall, B., Sajjadi, M., Geiger, A., and Radwan, N. RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. ArXiv. /abs/2112.00724, 2021

[3] Mildenhall, B., Srinivasan, P., Tancik, M., Barron, J., Ramamoorthi, R., and Ng, R. NeRF: representing scenes as neural radiance fields for view synthesis. Commun. ACM 65, 1 (January 2022), 99–106. https://doi.org/10.1145/3503250, 2021

Method

- We incorporated a geometry regulation term from unobserved viewpoints into 3D Gaussian Splatting to attain sparse input for 3D Gaussian Splatting.

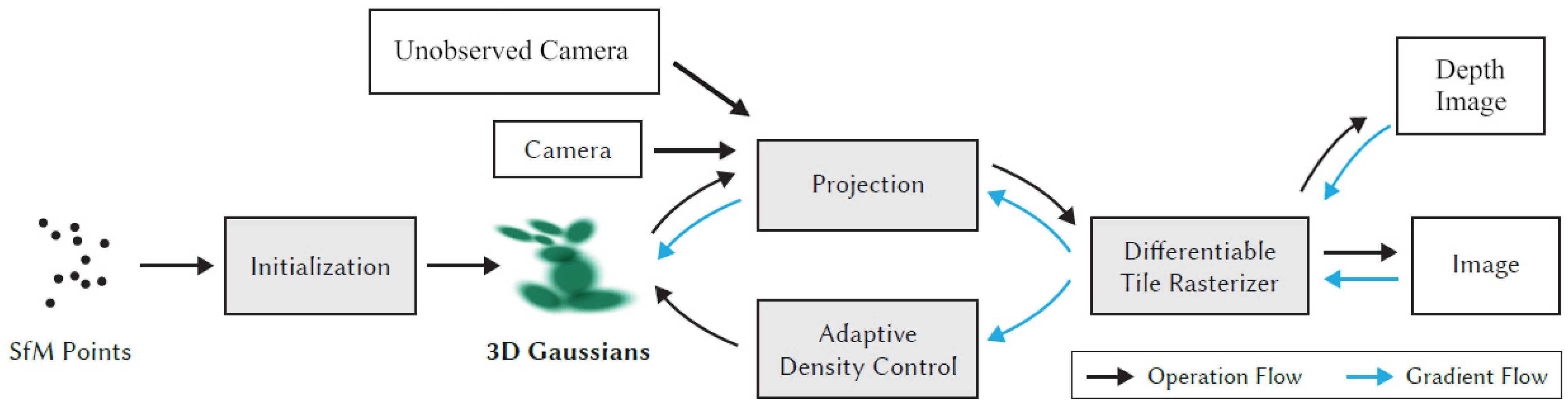


Figure 1: model overview

- For each iteration, we sample an unobserved viewpoint camera, similar to how a pixel's color is rendered, we calculate the expected depth at each pixel as α-blending of each gaussian's depth. Hence, we'll get a rendered depth image. We formulate our depth smoothness loss as:

$$\mathcal{L}_{DS}(\theta, \mathcal{C}_r) = \sum_{\mathbf{r}\in\mathcal{C}_r}\sum_{i,j}(\hat{d}_\theta(\mathbf{r}_{ij}) - \hat{d}_\theta(\mathbf{r}_{i+1j}))^2 + (\hat{d}_\theta(\mathbf{r}_{ij}) - \hat{d}_\theta(\mathbf{r}_{ij+1}))^2$$

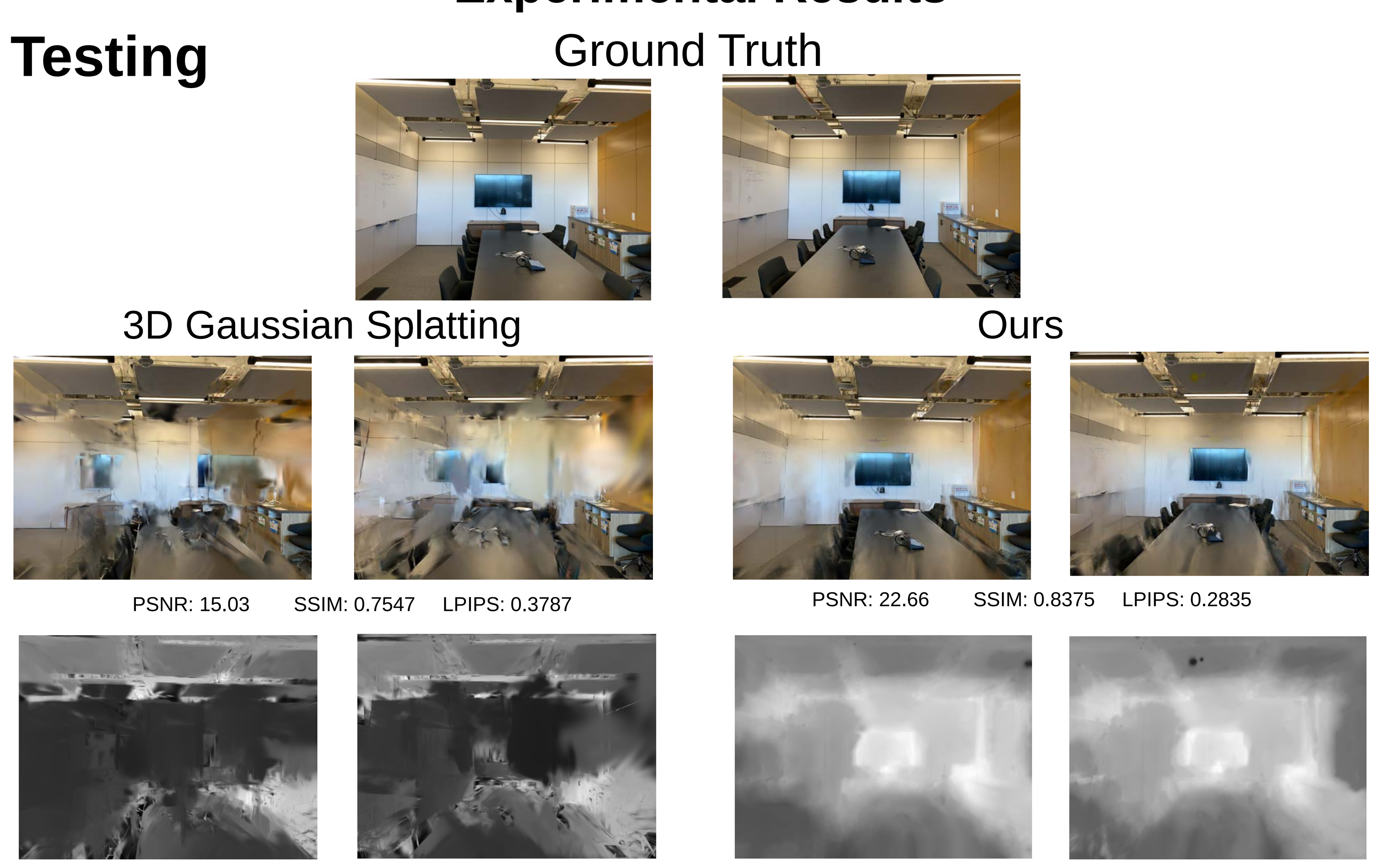
where $\hat{d}_\theta(\mathbf{r}_{i+1j})$ is the expected depth at pixel i, j, θ is the model parameter and $\mathcal{C}_r$ is the randomly sampled unobserved viewpoint camera.

Since the original loss in 3D Gaussian Splatting is $\mathcal{L} = (1-\lambda)\mathcal{L}_1 + \lambda\mathcal{L}_{\text{D-SSIM}}$, we define our total loss to be:

$$\mathcal{L}_{total} = \mathcal{L} + \lambda_D \mathcal{L}_{DS}$$

Experimental Results

Testing



Training