3D Gaussian Splatting for Reconstructing Large Sparse Environments (Student Abstract)

Jonathan Boel Nielsen[1], Xuan Huy Pham[1], Erdal Kayacan[2], Andriy Sarabakha[1]

[1]Department of Electrical and Computer Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark
[2]Department of Electrical Engineering and Information Technology, Paderborn University, Paderborn, Germany
202004778@post.au.dk, huy.pham@ece.au.dk, erdal.kayacan@uni-paderborn.de, andriy@ece.au.dk

Abstract

3D Gaussian splatting (3DGS) has recently demonstrated significant potential in computer vision, enabling high-fidelity 3D scene reconstruction with real-time rendering and fast training times. However, existing methods struggle in large, visually sparse, geometric self-similarity environments due to heavy reliance on image-based feature matching and depth information. In this work, we propose a novel reconstruction pipeline that reduces the dependence on visual features by incorporating IMU and LiDAR data to generate accurate point clouds and robustly localize images within the scene. Global colorization is achieved through 3D-to-2D projections of the localized images, which are then used to supervise 3DGS training. Our results demonstrate that the proposed pipeline significantly enhances the quality of 3D reconstruction for large, sparse scenarios, opening up new opportunities for applications in remote mapping and autonomous inspection.

Introduction

In neural rendering, 3D Gaussian Splatting (3DGS) (Kerbl et al. 2023) has emerged as a state-of-the-art technique, offering significant improvements over previous methods. However, 3DGS pipelines rely on Structure-from-Motion (SfM) to estimate image poses and generate initial sparse point clouds. Thus, SfM-based methods often fail in feature-sparse environments due to the absence of reliable feature correspondences or the presence of large texture-less areas (Pataki et al. 2025), which leads to floating Gaussians. While recent research has proposed enhancements to reduce floating Gaussians and hallucinated geometry (Cheng et al. 2024; Matsuki et al. 2024), the root issue – visual dependency for pose and geometry estimation – remains unresolved.

In this work, we address limitations in image feature dependency by introducing a novel 3DGS pipeline that eliminates the dependency on visual features for both localization and structure estimation. We evaluate our method in simulated and real-world scenarios. Our approach outperforms SfM-based methods in sparse environments by producing more coherent reconstructions with fewer floating Gaussians and less environmental clutter. Although our approach exhibits minor pose estimation errors, the overall visual and structural consistency remains competitive.

Proposed Method

This work proposes a three-step pipeline, illustrated in Figure 1, to achieve SfM-free sparse environment reconstruction. The first step takes RGB images, LiDAR scans and IMU data to generate a point cloud of the scene and estimate camera poses within the scene. Unlike previous work using an image-based SfM approach, we utilize a direct point cloud-based method called FAST-LIO2 (Xu et al. 2022) to estimate poses during run-time. The second step uses 3D-to-2D projection to determine relations between the geometric information of the point cloud and the color information of the images at each estimated pose. We filter, downsample and perform hidden points removal. Then, the colors of the points are registered. The third step globally colors the point cloud using the registered color information. We merge all colors attributed to each point and apply the median color to eliminate outliers. To significantly reduce floating Gaussians in free-space regions, we provide 3DGS with an initial depth estimate for regions with no depth information, such as backgrounds. A spherical region is used to separate objects and background during reconstruction.

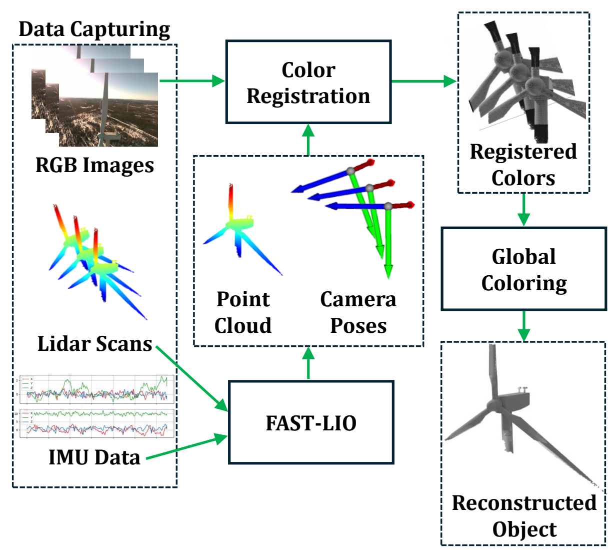

Figure 1: The proposed pipeline.

Method	Dense Environment			Small Sparse Environment			Large Sparse Environment			Real World Environment		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
COLMAP	32.53	0.940	0.163	38.88	0.989	0.022	36.66	0.978	0.065	27.20	0.906	0.152
Ours	28.95	0.906	0.170	30.9	0.938	0.067	35.28	0.983	0.037	21.67	0.677	0.309

Table 1: Evaluation of the reconstruction of different environments using peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and learned perceptual image patch similarity (LPIPS) metrics.

Experimental Results

The proposed method is tested in four different environments: three simulation-based and one real-world. The proposed method is compared with COLMAP's SfM and evaluated to determine under which conditions our method may be most efficient. We evaluate the results using two metrics: quantitative visual fidelity (in Table 1) and qualitative environment coherence (in Figure 2).

Simulation results. In the dense environment, performance is comparable between the two methods. COLMAP achieves slightly sharper images and better metrics; however, our method is still competitive in this environment. In the small sparse environment, COLMAP achieves good metrics but poor coherence. The textures on the object allow good initialization, but the extensive free space surrounding it spawns floating Gaussians that severely degrade performance. In contrast, FAST-LIO struggles with accurate pose estimation due to the object's small size, so our method produces a blurred reconstruction, but with significantly fewer floating Gaussians. In a large sparse environment, COLMAP struggles to create a good initialization with coherent geometry due to sparsity and homogeneity of visual features, and the presence of extensive free space leads to a high number of floating Gaussians. The reconstruction is not coherent, and the geometry of the wind turbine is poor. However, the evaluation metrics are high, as overfitting and floating Gaussians are not penalized. Our method shows better SSIM and LPIPS, and significantly better environment coherence.

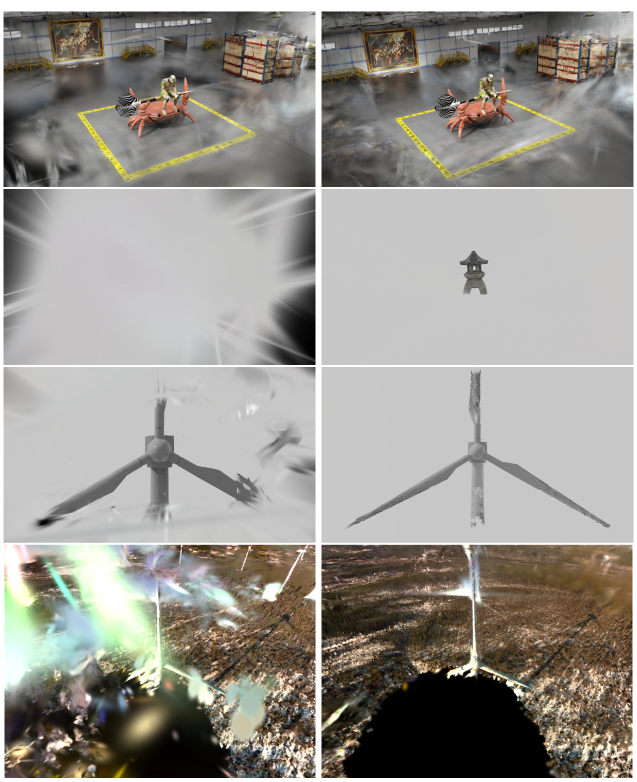

Figure 2: Left: COLMAP-3DGS, right: our method. Environments from top to bottom: dense, small sparse, large sparse, real-world.

Real world results. The real-world data is captured using a drone flying around a wind turbine in a semicircular path facing the wind turbine. The motion is relatively fast, the drone is relatively far from the wind turbine, and the resolution of the camera is relatively low, making the scenario quite challenging. Our method is tested alongside the SfM-based method in this real-world sparse environment. COLMAP is able to provide good initialization and does not have any major pose estimation issues since there are many visual features on the ground. Since COLMAP can reconstruct the ground well, its overall metrics are high. Our method projects the ground and the sky onto a sphere around the environment, which degrades the reconstruction of the ground, but reduces floating Gaussians around the wind turbine. Besides, our method generates a coherent geometry of the wind turbine, whereas COLMAP cannot align the two sides of the top blade, resulting in one side being occluded. This again shows that overfitting can cause issues in poorly initialized SfM-based environments.

Conclusions

By decoupling image localization and scene structure from dependence on visual features, our method unlocks the potential of 3DGS in visually degraded or feature-poor environments. Our experimental results demonstrate that 3DGS can be effectively applied to reconstruct sparse environments where traditional SfM-based methods typically fail. Although the proposed method does not outperform existing approaches in standard reconstruction metrics, primarily due to the absence of a global pose optimization framework, it achieves more coherent reconstructed environments and significantly reduces floating artefacts. In future work, we will incorporate global pose optimization into our pipeline to further enhance performance.

Acknowledgments

This research was partially supported by Aarhus Universitets Forskningsfond (AUFF) and Nationalt Forsvarsteknologisk Center (NFC) under grant 25016.

References

Cheng, K.; Long, X.; Yang, K.; Yao, Y.; Yin, W.; Ma, Y.; Wang, W.; and Chen, X. 2024. GaussianPro: 3D Gaussian splatting with progressive propagation. In *Proceedings of the 41st International Conference on Machine Learning*.

Kerbl, B.; Kopanas, G.; Leimkuehler, T.; and Drettakis, G. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 42(4).

Matsuki, H.; Murai, R.; Kelly, P. H. J.; and Davison, A. J. 2024. Gaussian Splatting SLAM. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 18039–18048.

Pataki, Z.; Sarlin, P.-E.; Schönberger, J. L.; and Pollefeys, M. 2025. MP-SfM: Monocular Surface Priors for Robust Structure-from-Motion. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Xu, W.; Cai, Y.; He, D.; Lin, J.; and Zhang, F. 2022. FAST-LIO2: Fast Direct LiDAR-Inertial Odometry. *IEEE Transactions on Robotics*, 38(4): 2053–2073.