3D gaussian splatting fusion slam

Sun Zewen[1]

[1]**ARC, National University of Singapore**

ABSTRACT

This research presents a real-time SLAM framework that fuses 3D Gaussian Splatting (3DGS) with LiDAR-Inertial-Visual Odometry (LIVO) for photorealistic and robust mapping in challenging environments. Existing 3DGS-based SLAM methods face issues such as high memory usage, poor adaptability to dynamic scenes, and sensitivity to lighting and sparse inputs.
To address these challenges, I propose a hybrid system that integrates tightly coupled multi-sensor fusion via an Iterated Error-State Kalman Filter (IESKF), voxel-wise Gaussian Process Regression (Voxel-GPR) for LiDAR densification, and frequency-aware optimization scheduling for computational efficiency. Additionally, we incorporate mutual-information-based image selection using Shannon Mutual Information (SMI) to discard low-quality frames, enhancing robustness under degraded visual conditions.
Only informative Gaussians are fused into a global map using LIO-estimated poses, ensuring real-time performance with high visual fidelity. The proposed system improves scalability, robustness, and efficiency over existing approaches, making it suitable for deployment on embedded robotic platforms.

Keywords: 3D Gaussian Splatting, SLAM, LIVO

INTRODUCTION AND BACKGROUND

0.1 3D Gaussian Splatting (3DGS) in Robotics

3D Gaussian Splatting (3DGS) is a new method for building 3D scenes. Instead of using neural networks like NeRF, 3DGS uses many small 3D Gaussians to represent scene geometry, color, and transparency. This makes it faster and more efficient for real-time tasks [1]. It also supports differentiable rendering, which is useful for training with gradients.

In robotics, 3DGS helps with accurate and fast scene understanding. Tasks like reconstruction, localization, and navigation benefit from 3DGS because it can create high-quality 3D maps quickly [2]. This makes it a strong candidate for real-time robot perception and mapping.

0.2 3D Gaussian Splatting Fusion SLAM (3DGS + SLAM)

Combining 3DGS with SLAM (Simultaneous Localization and Mapping) allows robots to build dense maps while tracking their position. Compared to NeRF-based SLAM, 3DGS-SLAM works faster and uses explicit 3D structures [3]. Recent systems like Gaussian-SLAM and SplaTAM achieve strong performance in indoor and outdoor tests [4].

Some methods also add semantics, such as object classes, to help robots understand their surroundings better [5]. These advances show 3DGS can support both geometry and high-level perception in robotics.

0.3 Current Drawbacks in 3DGS + SLAM

Despite the advantages, current 3DGS-based SLAM systems have some problems. First, storing many Gaussians uses a lot of memory, which is hard for large-scale scenes [2]. Second, most systems assume static scenes, so they don't work well when objects move. Third, poor lighting or depth input can hurt performance. More research is needed to make 3DGS-SLAM more robust, efficient, and ready for real-world robotics.

RESEARCH OBJECTIVES

The goal of this research is to develop a real-time 3D SLAM framework that integrates 3D Gaussian Splatting (3DGS) with multi-sensor fusion to achieve robust and photorealistic mapping under varying lighting conditions. Specifically, the framework aims to:

- Fuse LiDAR-Inertial-Visual Odometry (LIVO) with 3DGS for accurate and real-time pose estimation [6];
- Incorporate light-robust mechanisms using image selection based on Shannon Mutual Information (SMI), filtering out visually unreliable frames for 3DGS-based visual odometry [7];
- Densify sparse LiDAR point clouds using voxel-wise Gaussian Process Regression (GPR) to generate structured priors for 3D Gaussian initialization [8];
- Introduce frequency-aware resolution scheduling for computationally efficient 3DGS optimization [9].

METHODOLOGY

This research builds upon recent advancements in real-time SLAM using 3D Gaussian Splatting (3DGS), particularly GS-LIVO [6], GS-LIVM [8], DashGaussian [9], and GauSS-MI [7]. The proposed system comprises five integrated components:

1. Real-time Pose Estimation via LIVO

The system employs a tightly coupled LiDAR-Inertial-Visual odometry framework adapted from GS-LIVO [6], which uses an Iterated Error-State Kalman Filter (IESKF) to fuse IMU, LiDAR, and visual inputs. Photometric residuals rendered from the 3D Gaussian map are used to refine pose estimates in a semi-dense visual update stage, thereby ensuring both geometric and photometric consistency.

2. Gaussian Initialization with Voxel-Level GPR

To address LiDAR sparsity and non-uniform point distribution, voxel-wise Gaussian Process Regression (Voxel-GPR) [8] is introduced. For each voxel, principal component analysis (PCA) identifies the dominant axis, along which predictions are generated using a Gaussian process. These predictions provide structured, evenly sampled points along orthogonal axes, which are used to initialize the 3D Gaussian position, scale, and orientation. Covariances estimated during GPR also inform initialization confidence, enabling filtering of unreliable Gaussians.

3. Resolution-Aware Gaussian Optimization

To reduce computational overhead during optimization, the system adopts the DashGaussian strategy [9]. A frequency-based scheduling scheme dynamically adjusts both rendering resolution and Gaussian primitive count across optimization steps. Early stages use low-resolution renderings to fit low-frequency image components; resolution is gradually increased to incorporate finer details. The resolution and primitive count are jointly scheduled to maintain a balance between speed and quality.

4. Mutual-Information-Guided Visual Filtering

Since image quality can vary due to lighting or motion blur, we apply the GauSS-MI framework [7] to assess each frame's informativeness using Shannon Mutual Information (SMI). Frames below a specified MI threshold are excluded from the 3DGS visual update and map integration stages. This filtering mechanism improves robustness and avoids performance degradation caused by low-quality frames.

5. Global Gaussian Map Integration

Selected Gaussians are projected into the world frame using the LIO-estimated transformation:

$$G_{\text{world}} = T_{\text{LIO}} \cdot G_{\text{camera}}$$

Only Gaussians from frames passing the MI threshold are integrated into the global map. This ensures efficient memory usage and maintains photorealistic quality while preserving real-time operation.

EXPECTED CONTRIBUTIONS

1. **A Light-Robust 3DGS-Fusion SLAM System:** A novel SLAM pipeline that conditionally uses vision for Gaussian-based pose refinement.

2. **Voxel-GPR-Based Initialization:** A structured approach to initializing Gaussians using voxel-wise Gaussian process predictions.

3. **Frequency-Aware Optimization Scheduling:** A dynamic scheduler for Gaussian optimization that reduces computation without compromising quality.

4. **Mutual-Information-Guided Image Selection:** A principled visual quality filter based on SMI to discard low-quality frames.

5. **Real-Time SLAM Implementation:** A system deployable on resource-limited platforms (e.g., Jetson Orin NX), validated on diverse datasets.

CONCLUSION

This research proposes a robust and efficient 3D SLAM framework combining the photorealistic rendering of 3D Gaussian Splatting with LiDAR-Inertial-Visual fusion and mutual-information-guided visual selection. It addresses key challenges in outdoor robotics, including LiDAR sparsity, illumination variance, and real-time constraints. The final system aims to be capable of robust deployment in dynamic, unstructured environments, contributing a new direction in real-time photometric SLAM research.

REFERENCES

[1] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3d gaussian splatting for real-time radiance field rendering," *ACM Transactions on Graphics (SIGGRAPH)*, vol. 42, no. 4, pp. 1–14, 2023.

[2] S. Zhu, G. Wang, X. Kong, D. Kong, and H. Wang, "3d gaussian splatting in robotics: A survey," *arXiv preprint arXiv:2410.12262*, 2024.

[3] N. Keetha, J. Karhade, K. Jatavallabhula, G. Yang, S. Scherer, D. Ramanan, and J. Luiten, "Splatam: Splat, track & map 3d gaussians for dense rgb-d slam," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 21 357–21 366.

[4] V. Yugay, Y. Li, T. Gevers, and M. R. Oswald, "Gaussian-slam: Photo-realistic dense slam with gaussian splatting," *arXiv preprint arXiv:2312.10070*, 2023.

[5] S. Zhu, R. Qin, G. Wang, J. Liu, and H. Wang, "Semgauss-slam: Dense semantic gaussian splatting slam," *arXiv preprint arXiv:2403.07494*, 2024.

[6] S. Hong, C. Zheng, Y. Shen *et al.*, "Gs-livo: Real-time lidar, inertial, and visual multi-sensor fused odometry with gaussian mapping," *arXiv preprint arXiv:2501.08672*, 2025.

[7] Y. Xie, Y. Cai, Y. Zhang *et al.*, "Gauss-mi: Gaussian splatting shannon mutual information for active 3d reconstruction," *arXiv preprint arXiv:2504.21067*, 2025.

[8] Y. Xie, Z. Huang, J. Wu, and J. Ma, "Gs-livm: Real-time photo-realistic lidar-inertial-visual mapping with gaussian splatting," *arXiv preprint arXiv:2410.17084*, 2024.

[9] Y. Chen, J. Jiang *et al.*, "Dashgaussian: Optimizing 3d gaussian splatting in 200 seconds," *arXiv preprint arXiv:2503.18402*, 2025.