Online Gaussian Mapping for Missing Video Frames and Novel Views

Background

A moving camera records only a sequence of 2D views, although neighbouring frames observe a shared 3D scene. Simultaneous localisation and mapping (SLAM) can estimate camera motion while building a map. 3D Gaussian splatting provides a map that can also be rendered rapidly from a chosen camera pose [1]. Existing Gaussian SLAM systems build such maps incrementally from video [2, 3].

Problem Specification

When a video frame is dropped, a standard 2D interpolator may struggle with parallax and newly revealed surfaces. A Gaussian map could render the missing view at an estimated camera pose and support modest movement of a virtual camera. Its quality depends on which video frames update the map. Redundant frames waste processing time, while missing viewpoints leave holes or rendering artefacts. This project asks whether selecting frames for view coverage can improve reconstruction of missing frames and nearby novel views under a fixed update budget.

Suggested Proposed Method

- **Start from a working baseline.** Reproduce SplaTAM [2] on a short, mostly static indoor video with colour and depth. Use depth to make the first implementation manageable. Record camera poses and maintain the incremental Gaussian map. Read MonoGS [3] for its monocular and keyframe design.
- **Select and update.** Add a keyframe score that rewards newly visible regions, useful viewpoint change and high render error, while penalising overlap with existing keyframes. Admit selected frames to a small optimisation window; add Gaussians where valid depth reveals unmapped surfaces and locally refine the map. Compare with the baseline keyframe policy and uniform sampling at equal frame and optimisation budgets [4].
- **Render gaps and explore.** Simulate short frame losses by hiding selected colour and depth frames from tracking and map updates. Interpolate their camera poses from adjacent observed poses, then render the map at those poses. Allow small virtual camera offsets. Output a visibility and coverage mask so poorly observed areas can be recognised instead of treated as reliable reconstructions.
- **Evaluate.** Keep target frames out of all map updates. On indoor sequences with colour and depth, such as Replica and the TUM dataset, compare withheld views with recorded images using PSNR, SSIM and LPIPS. Report camera pose error where reference poses exist, as well as update time, rendering time and map size. Inspect failure cases such as large gaps, occlusions and motion.

Expected Outcome and Scope

A reproducible prototype, an ablation of the keyframe score, paired reconstructions of withheld frames, and a short demonstration of virtual camera movement. The core project assumes colour and depth input and largely static scenes. Monocular input or handling moving objects is an extension if time permits. Begin by running [2] on one scene and rendering a withheld frame before changing its selection rule.

Suggested Starting Sources

[1] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," *ACM Transactions on Graphics*, vol. 42, no. 4, 2023. [paper]

[2] N. Keetha et al., "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. 21357–21366. [paper] [code]

[3] H. Matsuki, R. Murai, P. H. J. Kelly, and A. J. Davison, "Gaussian Splatting SLAM," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 18039–18048. [paper] [code]

[4] B. Lee, J. Park, K. T. Giang, and S. Song, "Online 3D Gaussian Splatting Modeling with Novel View Selection," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2025, pp. 1323–1331. [paper]