Gaussian Splatting SLAM

Hidenobu Matsuki[1]* Riku Murai[2]* Paul H. J. Kelly[2] Andrew J. Davison[1]

[1]Dyson Robotics Laboratory, Imperial College London
[2]Software Performance Optimisation Group, Imperial College London

Website: https://rmurai.co.uk/projects/GaussianSplattingSLAM/
Video: https://youtu.be/x604ghp9R_Q/

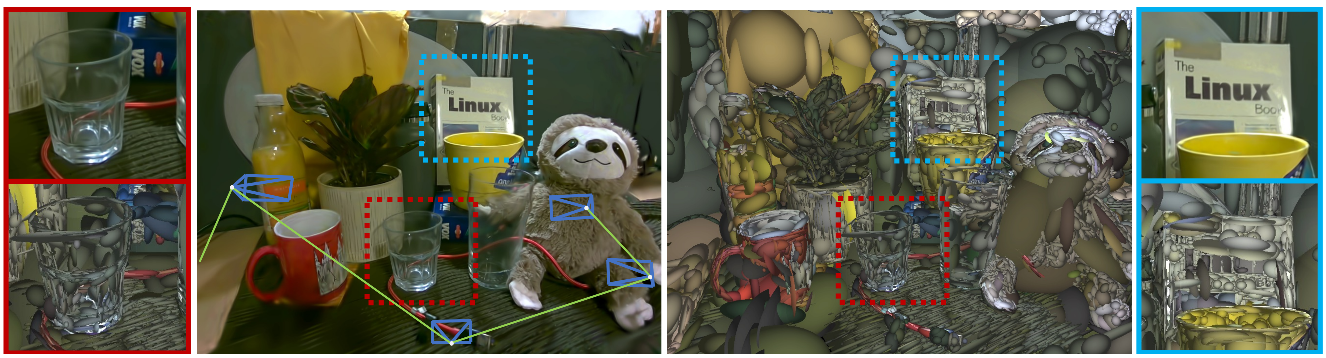

Figure 1. From a single monocular camera, we reconstruct a high fidelity *3D scene* live at 3fps. For every incoming RGB frame, 3D Gaussians are incrementally formed and optimised together with the camera poses. We show both the rasterised Gaussians (left) and Gaussians shaded to highlight the geometry (right). Notice the details and the complex material properties (e.g. transparency) captured. Thin structures such as wires are accurately represented by numerous small, elongated Gaussians, and transparent objects are effectively represented by placing the Gaussians along the rim. Our system significantly advances the fidelity a live monocular SLAM system can capture.

Abstract

We present the first application of 3D Gaussian Splatting to incremental 3D reconstruction using a single moving monocular or RGB-D camera. Our Simultaneous Localisation and Mapping (SLAM) method, which runs live at 3fps, utilises Gaussians as the only 3D representation, unifying the required representation for accurate, efficient tracking, mapping, and high-quality rendering.

Several innovations are required to continuously reconstruct 3D scenes with high fidelity from a live camera. First, to move beyond the original 3DGS algorithm, which requires accurate poses from an offline Structure from Motion (SfM) system, we formulate camera tracking for 3DGS using direct optimisation against the 3D Gaussians, and show that this enables fast and robust tracking with a wide basin of convergence. Second, by utilising the explicit nature of the Gaussians, we introduce geometric verification and regularisation to handle the ambiguities occurring in incremental 3D dense reconstruction. Finally, we introduce a full SLAM system which not only achieves state-of-the-art results in novel view synthesis and trajectory estimation, but also reconstruction of tiny and even transparent objects.

*Authors contributed equally to this work.

1. Introduction

A long-term goal of online reconstruction with a single moving camera is near-photorealistic fidelity, which will surely allow new levels of performance in many areas of Spatial AI and robotics as well as opening up a whole range of new applications. While we increasingly see the benefit of applying powerful pre-trained priors to 3D reconstruction, a key avenue for progress is still the invention and development of core 3D representations with advantageous properties. While many "layered" SLAM systems exist which combine multiple representations, the most in-