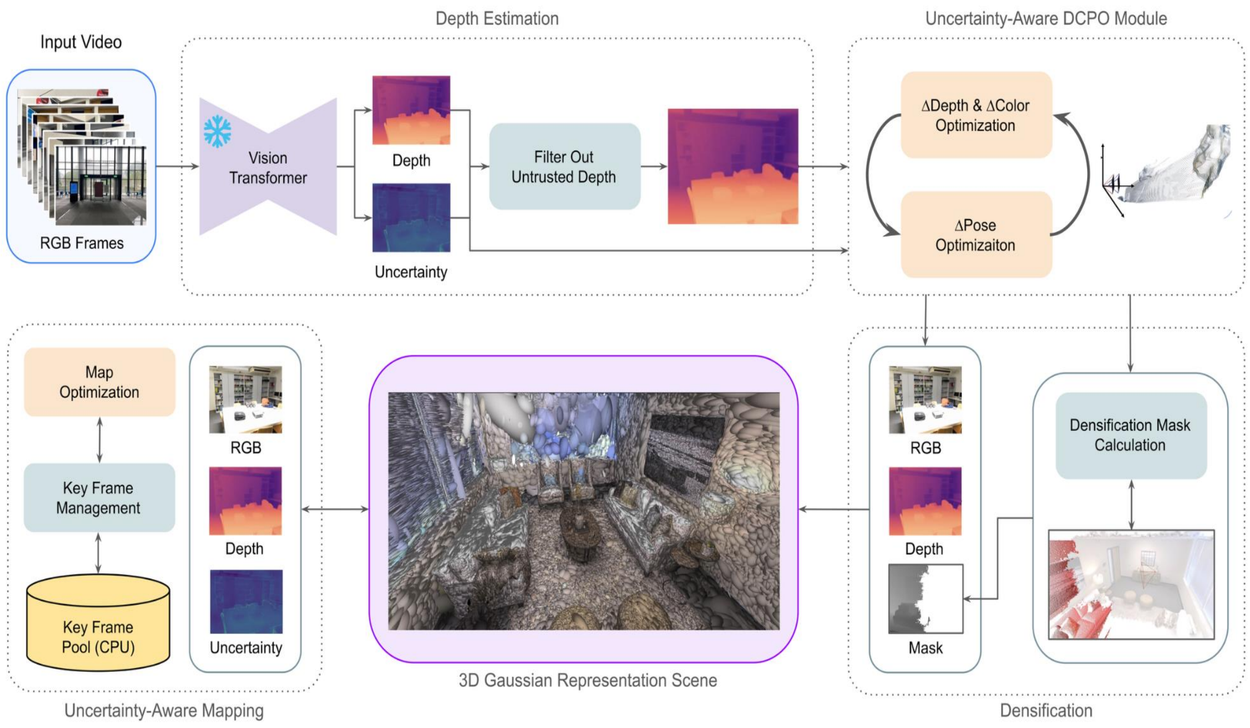
RGB-Only 3D Scene Reconstruction via Uncertainty-Aware 3D Gaussian SLAM

Chih-Yao Chang, Tzu-Chieh Chen, Yan-Ting Li, and Jia-You Xiao

Introduction

- 3D reconstruction is a fundamental task in computer .
- Traditional methods like LiDAR scanning, their prohibitive cost limits accessibility and their effective range is relatively narrow.
- To address these limitations, we propose a novel approach leveraging only RGB video to reconstruct precise 3D scene using 3D Gaussian and SLAM technique.

Method

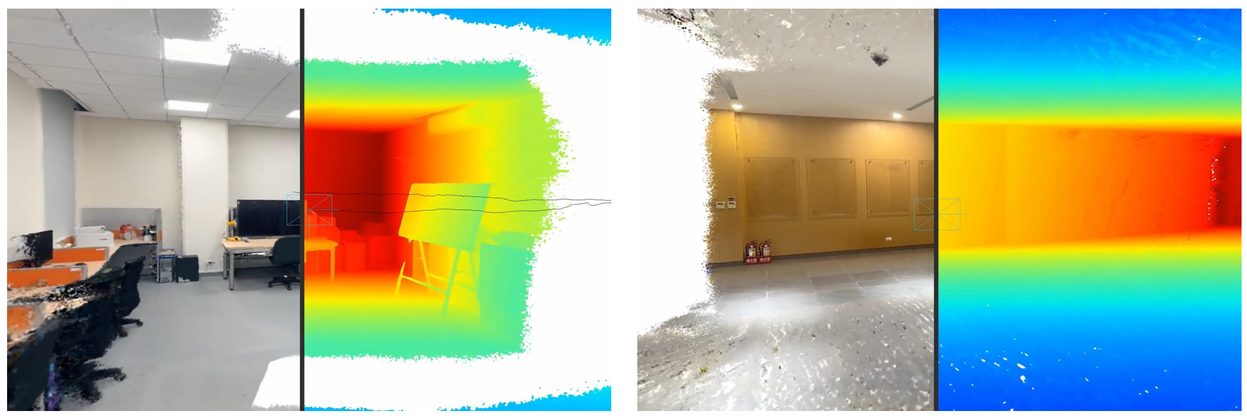


1. **Predict the depth map** and corresponding uncertainty map from the input RGB frames and filter out high-uncertainty pixels to improve the reliability of the estimated depth map.
2. **Iteratively optimize the depth map, color map, and camera pose**, using the uncertainty map to guide the optimization thus reducing the influence of unreliable depth values and improve the overall accuracy.
3. **Inject 3D Gaussians** into the scene at the newly estimated camera pose, using the rendered densification mask along with the optimized depth and RGB maps.
4. Using a set of previously selected keyframes, we **refine the parameters of the 3D Gaussian Map**. Improving overall scene consistency, ensuring the reconstruction remains sharp and coherent over time.

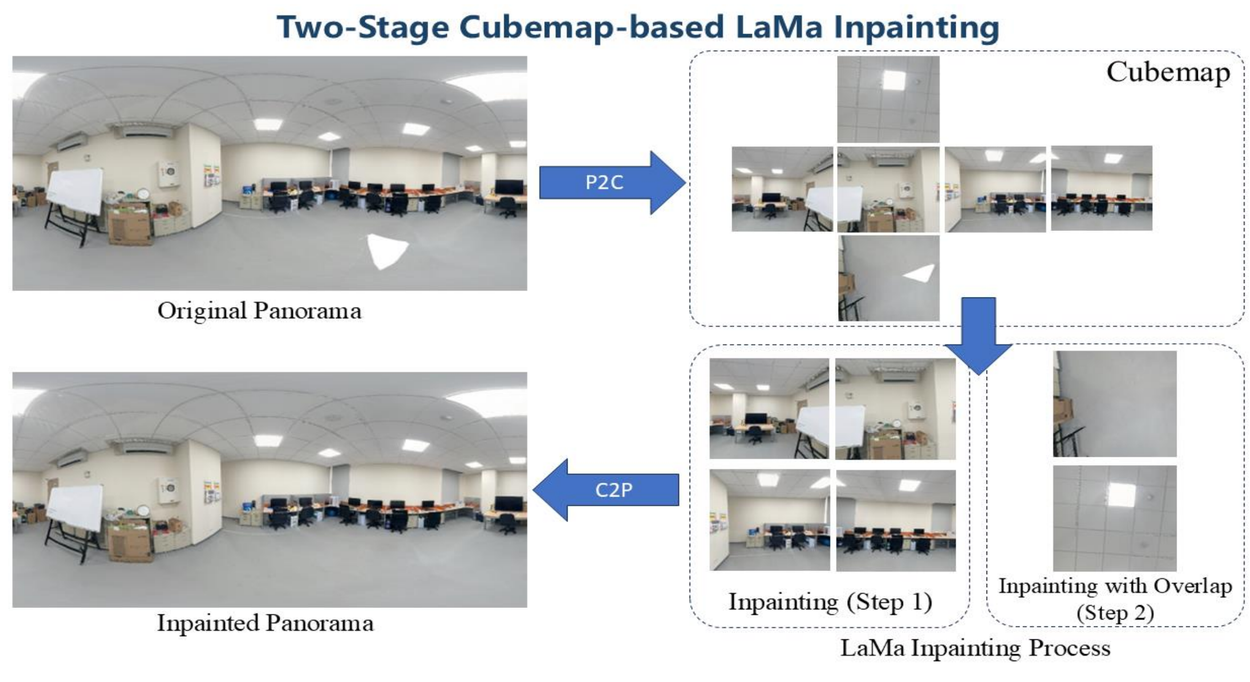
Result

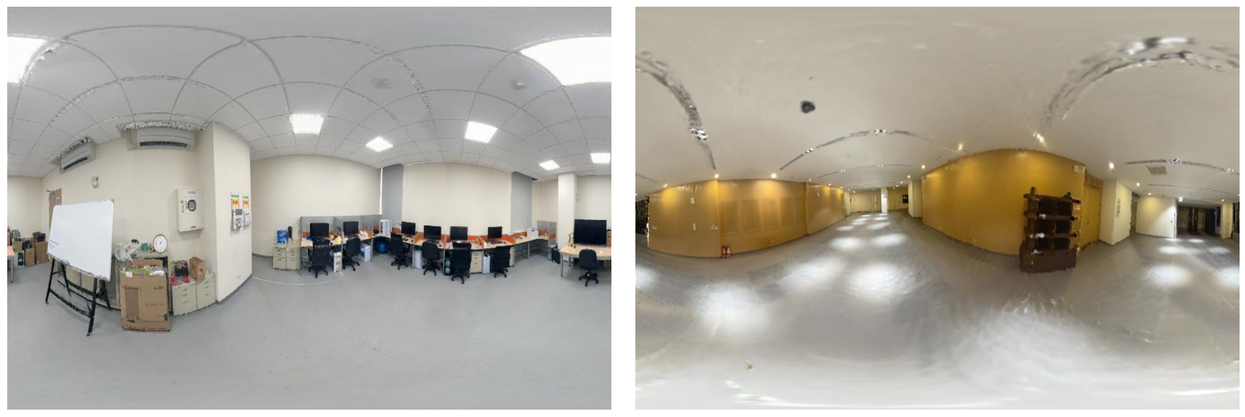
- Reconstructed high-quality, dense 3D scenes from monocular RGB video with minimal visual artifacts.
- Enabled 360° novel view rendering and relocation
- But unseen areas that lack information remain incomplete. To address this, we also propose a novel inpainting method.

Reconstructed Scene

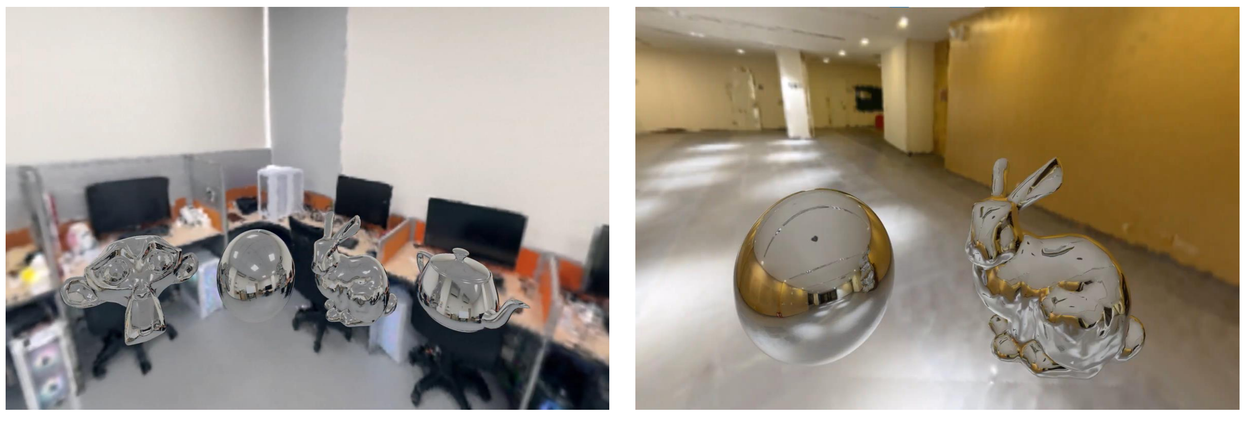


360° Relocation & Inpaint





Environmental Lighting



Conclusion

- Our SLAM system requires only monocular RGB video, eliminating the need for costly LiDAR hardware.
- The reconstructed 3D scene supports full 360° rendering, enabling easier visualization and camera relocation.
- We propose a novel method to inpaint missing regions in 360° images, enhancing completeness and realism.
- The generated 360° images can also be used for scene-aware lighting, improving downstream rendering tasks.