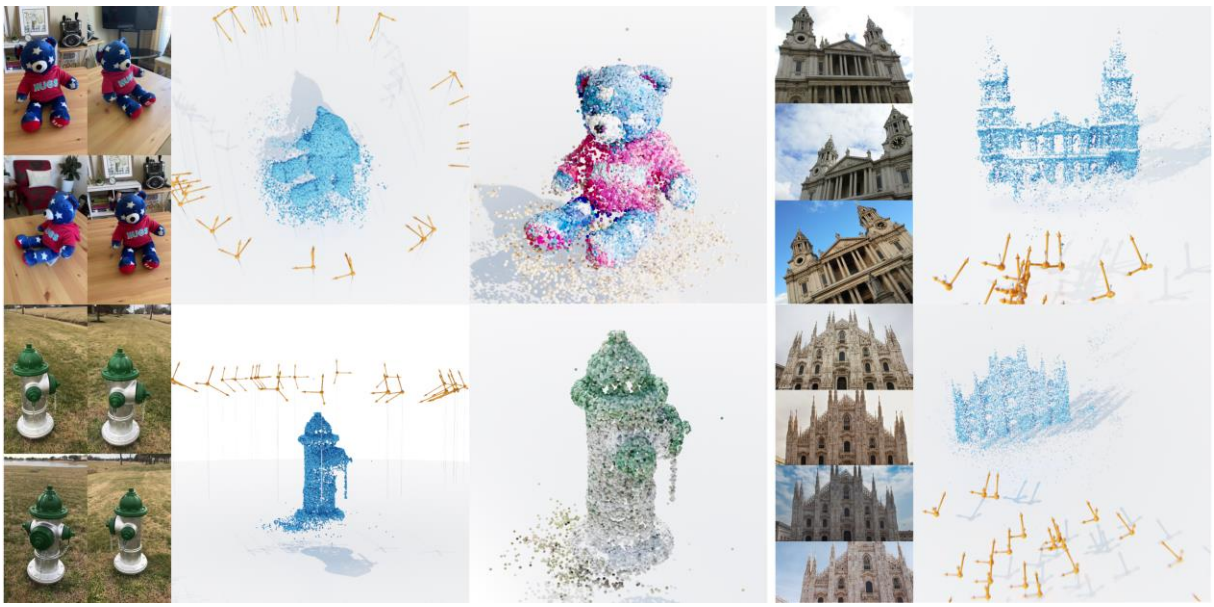
Robust 3D Reconstructions: Learning-Based SfM Initialization for Gaussian Splatting



Background

Structure from Motion (SfM) is a fundamental technique in computer vision, enabling 3D reconstruction by analyzing the motion between images. However, traditional SfM algorithms often require exhaustive feature matching and struggle with texture-less or noisy regions, where identifying consistent feature points becomes challenging. This project aims to address these limitations by exploring alternative point cloud initialization methods that enhance the accuracy of SfM in difficult visual conditions.

Problem Specification

Traditional SfM approaches can fail in areas with limited textures or high noise due to an inability to detect and match feature points across images. This project proposes to evaluate different SfM initialization methods to improve robustness in sparse or low-texture regions. Specifically, it will involve comparing the quality of point clouds generated by traditional SfM techniques with those produced by a learning-based SfM point cloud estimation method. To assess the effectiveness of these methods, trained GSplat models will be used to generate 3D reconstructions, which will then be evaluated for visual quality using PSNR, SSIM, VIF, and other scores against ground truth images.

Suggested Method

The project will implement a pipeline for training and testing GSplat models using point clouds generated by both traditional and learning-based SfM methods. Initially, the SfM point cloud will be

computed using conventional approaches, followed by a learning-based estimation method for comparison. Each point cloud will serve as input to train GSplat models, which will then render 3D reconstructions of the scenes. Performance will be evaluated by comparing PSNR, SSIM, and VIF metrics of the generated images to the ground truth, providing insights into the benefits of each initialization method under different conditions.

Relevant Articles

[1] Wang, J., Karaev, N., Rupprecht, C. and Novotny, D., 2024. VGGSfM: Visual Geometry Grounded Deep Structure From Motion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (GitHub)

[2] Cui, Z. and Tan, P., 2015. Global structure-from-motion by similarity averaging. In Proceedings of the IEEE International Conference on Computer Vision.

[3] Schonberger, J.L. and Frahm, J.M., 2016. Structure-from-motion revisited. In Proceedings of the IEEE conference on computer vision and pattern recognition. (GitHub)

[4] Lindenberger, P., Sarlin, P.E., Larsson, V. and Pollefeys, M., 2021. Pixel-perfect structure-from-motion with featuremetric refinement. In Proceedings of the IEEE/CVF international conference on computer vision. (GitHub)

[5] He, X., Sun, J., Wang, Y., Peng, S., Huang, Q., Bao, H. and Zhou, X., 2024. Detector-free structure from motion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (GitHub)

Project Draft By: Manu Gond (manu.gond@miun.se) | Emin Zerman (emin.zerman@miun.se)