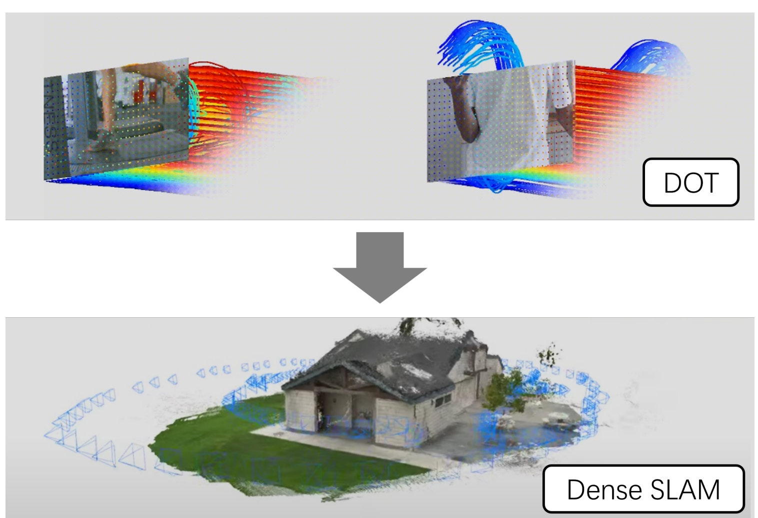
Visual SLAM with Dense Point Tracking

Overview

Visual SLAM aims to recover camera motion and 3D scenes from videos captured with an RGB camera. Existing methods primarily rely on two-view optical flow to infer relative camera pose and 3D geometry, often overlooking the temporal context within the sequence. This oversight can pose challenges in handling occlusions, dynamic scenes, and low-texture regions. However, recent advancements in long-term point tracking have shown potential for significant improvements in this area. By incorporating temporal information, one can more effectively model dynamic motions and occlusions, thereby enhancing Visual SLAM performance [1].

In this project, we will explore a recent advancement in dense point tracking known as DOT [2], a powerful and efficient method for tracking points. Initially, DOT computes rough estimates of a dense flow field and a visibility mask through nearest-neighbor interpolation for given source and target frames. These estimates are then refined using a learnable optical flow estimator designed to explicitly handle occlusions. With the dense point trajectories provided by DOT, our objective is to construct a visual SLAM system that integrates keyframe selection and bundle adjustment optimization [3, 4].

***Preferred background**: Epipolar Geometry, Bundle Adjustment, SLAM



(a) DOT-SLAM

Goals

The goal of this project is to extend DOT into a dense visual SLAM system and evaluate its performance on benchmark datasets and real-world videos.

Tasks

1 Getting familiar with the related literature and implementation details of DOT [1] and bundle adjustment [2]. (~ 50h);

2 Adding a bundle adjustment back-end to DOT to build the dense visual SLAM system (~ 100h);

3 Exploring the integration of additional SLAM features such as keyframe selection and dynamic filtering; (~ 50h);

4 Conducting experiments on existing benchmarks (e.g. KITTI) and real-world data (~ 50h);

5 Writing the project report and preparing presentations (~ 60h).

Contact

Weirong Chen Email: `weirong.chen@tum.de`
Linus Härenstam-Nielsen Email: `linus.nielsen@tum.de`

References

[1] Weirong Chen et al. "LEAP-VO: Long-term Effective Any Point Tracking for Visual Odometry". In: *arXiv preprint arXiv:2401.01887* (2024).

[2] Guillaume Le Moing, Jean Ponce, and Cordelia Schmid. "Dense optical tracking: Connecting the dots". In: *arXiv preprint arXiv:2312.00786* (2023).

[3] Zachary Teed and Jia Deng. "Droid-slam: Deep visual slam for monocular, stereo, and rgb-d cameras". In: *Advances in neural information processing systems* 34 (2021), pp. 16558–16569.

[4] Wang Zhao et al. "Particlesfm: Exploiting dense point trajectories for localizing moving cameras in the wild". In: *European Conference on Computer Vision*. Springer. 2022, pp. 523–542.

[1] https://github.com/16lemoing/dot
[2] https://github.com/princeton-vl/DROID-SLAM