COLAM
AN OFFLINE PYTHON SLAM USING COLMAP

ETH zürich

Computer Vision and Geometry Lab

COLAM - An Offline Python SLAM Using COLMAP

Weirong Chen, Yifei Liu, Kexin Shi, Yidan Gao
Supervised by Sarlin Paul-Edouard

Motivation & Goals

COLMAP[2] is a powerful toolbox for Structure-from-Motion, which estimates camera poses and 3D structures from a set of unordered images. Compared with online SLAM systems:
✔ Robust by learning powerful local features and camera parameters optimization.
✘ Slow on videos since it needs to match images pairs beforehand.
✘ Inefficient due to global bundle adjustment.
Our goal for this project is to:
• Leverage both advantages from COLMAP and ORB-SLAM[1] to build a powerful, accurate, efficient offline SLAM.
• Design simple, extendable and user-friendly python APIs.
• Benchmark on existing SLAM datasets and online videos.

Pipeline

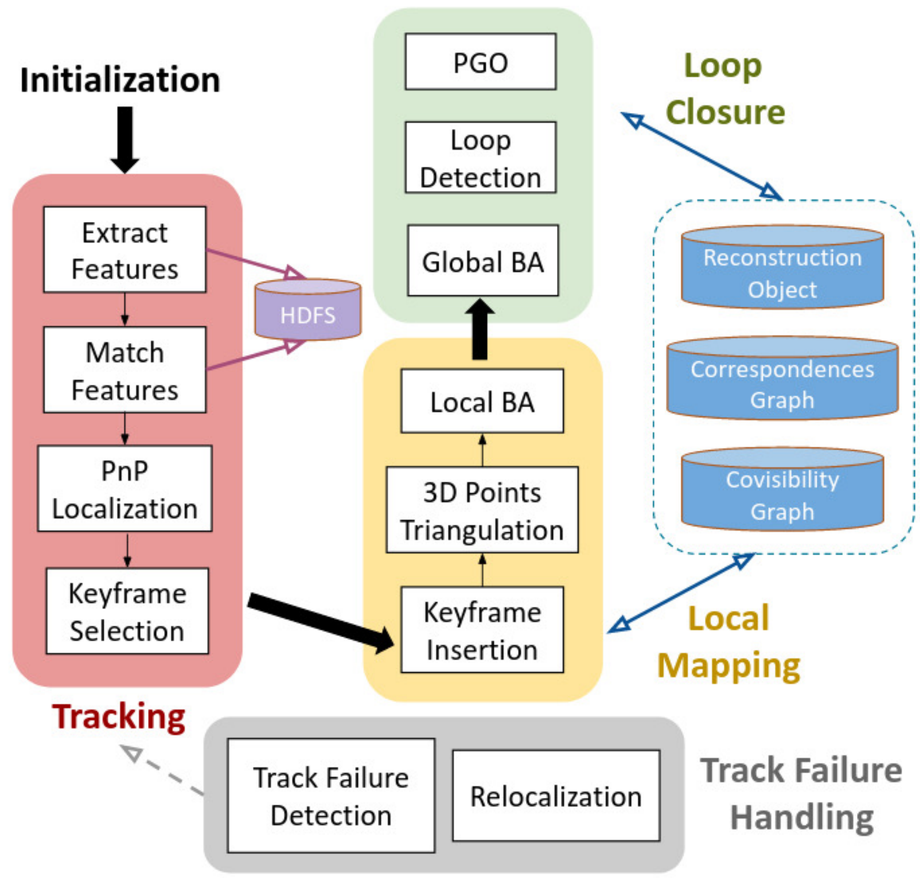

Figure 1:COLAM pipeline.

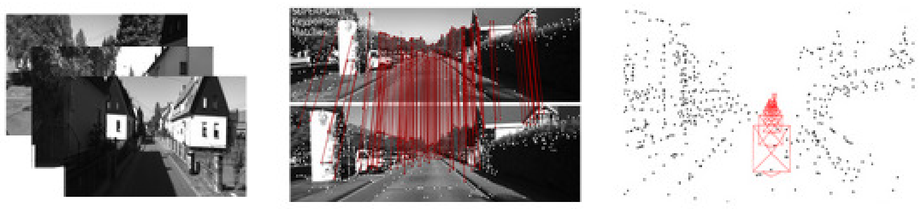

Figure 2:Visualization for tracking.

Key Insights

Major Tools:
• Hloc, PyCOLMAP, PyCeres, Open3D, NetVLAD.
Multiple Criteria for Keyframe Selection:
• Optical flow, ratio and quality of matched features, number of visible 3D points, time duration from the last keyframe.
• Improved ability to track features of variant depths.
Local and Global Bundle Adjustment:
• Refine camera poses and 3D points using the covisibility graph of the current frame.
• More robust against track failure and scale drift.
Loop Closure:
• Use NetVLAD to extract global features and apply Pose Graph Optimization when a loop is detected.
• Better reconstruction quality and trajectory completion.

Quantitative Results

Dataset	COLMAP	ORB-SLAM	Ours
kitti sequence05	> 1 day	5.441	59.130
rgbd_freiburg1_desk	29.367	0.410	7.750
rgbd_freiburg2_desk	> 1 day	4.962	64.304

Table 1:Reconstruction time comparison (min).

Dataset	max	mean	median	min	rmse	sse	std
kitti sequence05	0.858	0.459	0.468	0.175	0.491	191.058	0.174
rgbd_freiburg1_desk	0.549	0.174	0.135	0.029	0.209	3.572	0.116
rgbd_freiburg2_desk	0.194	0.072	0.066	0.008	0.087	1.264	0.047

Table 2:Absolute pose error w.r.t translation part (m)

Qualitative Results

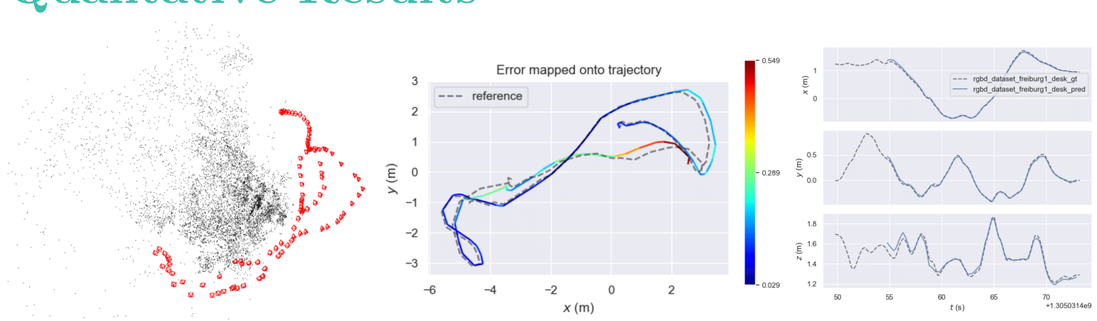

Figure 3:Evaluation on rgbd_freiburg1_desk.

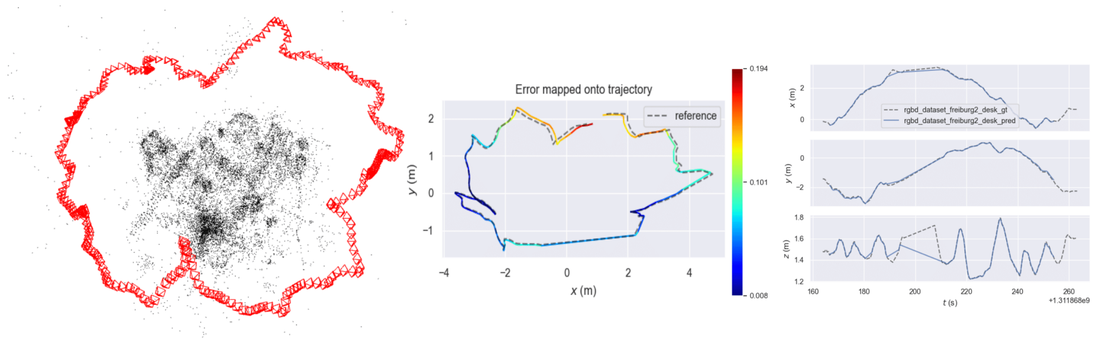

Figure 4:Evaluation on rgbd_freiburg2_desk.

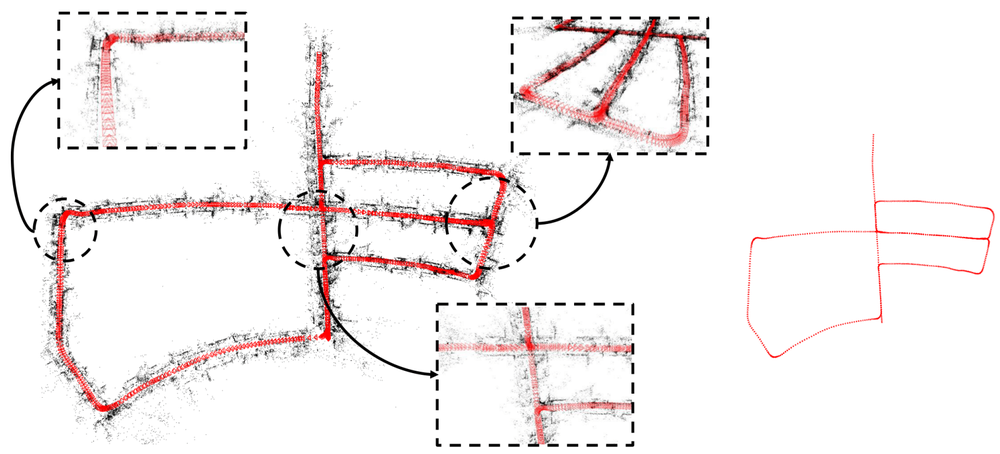

Figure 5:Evaluation on Kitti sequence05.

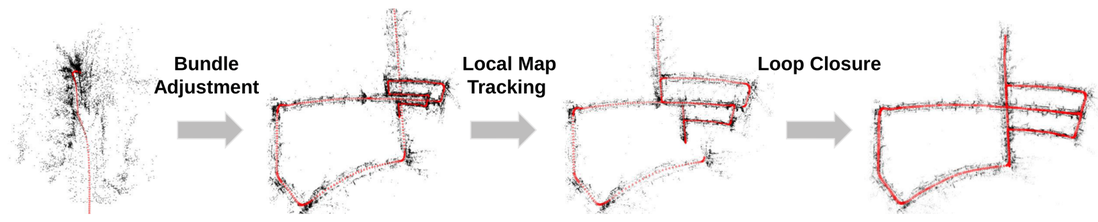

Figure 6:Ablation study on Kitti sequence05.

References

[1] M. J. M. M. Mur-Artal, Raúl and J. D. Tardós. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015. doi: 10.1109/TRO.2015.2463671.

[2] J. L. Schönberger and J.-M. Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.