Paper Reading 1: Short Summaries

Christoper Klammer

March 27, 2023

1 ORB-SLAM

- ORB-SLAM improves upon previous methods using the ORB feature descriptor in order to allow for good illumination and viewpoint invariance of features.

- Map initialization includes model selection and covers fundamental matrix and/or homography depending on the planarity of the two frames, motion is recovered and applied similar to how would be learned in a CV class with BA to correct errors.

- Tracking is for localizing the camera and seeing if a new keyframe should be processed if different enough (or enough time) or local mapping breaks down.

- Once keyframes are created, the covisibility graph adds the keyframe, linking to a node with the most overlap while map points can be removed if incosistent in the frames or keyframes are culled where then pairs of keyframes can be added to insert a map point.

- Loop closure looks for nonadjacent keyframes with high similarity the 3D correspondences for the candidate and if it is similar enough, we optimize this possibility and accept if there are enough inliers.

- From there, fusion takes place where the pose is corrected and propegated to the neighbors, fuse duplicate points, and insert new edges where then the loop can be closed by a pose graph optimization over the essential graph to correct for drift.

2 VLOAM

- VLOAM utilizes both LiDAR and cameras in order to build a map of an unknown environment and assumes a calibrated system between the lidar and the camera.

- The visual odometry has depth information from triangulation and the LiDAR where the motion between frames can be formulated such that features with depth give two equations and without depth gives one that can be stacked and solved by the Levenburg-Marquardt method.

- Features are weighted by their residuals and determined if they are outliers as well, the depth map is stored in a 2D KD-tree based on the angular coordinates from the LiDAR scans, if certain features are tracked then we also can triangulate them.

- The LiDAR odometry refines motion estimates and registers point clouds to the map where a linear motion model is used to correct and correspond edge and planar points between sweeps to estimate the translation T' and then remove the distortion.

- With distortion removed, the two consecutive map clouds are merged to create a new map cloud. From there, the sensor poses are integrated with the ones from visual odometry.

- The experiments show that indoor and outdoor datasets validate the method proposed and that the LiDAR odometry helps dramatically, allowing for a very small drift on the position while also showing that sudden and rapid motion changes are aided by the visual odometry portion proposed, creating more robustness.