[16-833] Paper Summaries 1

Bharath Somayajula

March 27, 2023

Contents

1 ORB-SLAM

1. ORB features are used in all sub-tasks such as tracking, mapping and loop closing due to their relative invariance to rotation and scale

2. The tracked information contains information about map such as 3D coordinate $\mathbf{X}_{w,i}$, viewing direction $\mathbf{n}_i$ and ORB Feature descriptor $\mathbf{D}_i$ and the keyframes such as camera pose $\mathbf{T}_{w,i}$ and camera intrinsics

3. The **Tracking** stage computes incremental pose between frames using either previous frame or relocalization and a decision is made whether to add the frame to list of keyframes

4. The **Local Mapping** stage uses Bundle Adjustment to refine local map estimates and any new keypoints remaining are added to the map

5. The **Loop Closing** stage searches for potential loop closure whenever a new frame is added and if loop closing is detected, the map and camera poses calculated are adjusted to minimize accumulated drift

6. The paper achieves real time performance without using GPU acceleration and is shown to detect loops accurately in NewCollege and KITTI datasets and achieve best keyframe localization on TUM RGB-D dataset

2 VLOAM

1. This work presents a way to combine Image(60Hz) and LIDAR(1Hz) sensor measurements to estimate odometry and map of scene

2. In the visual odometry step, 6DoF pose is estimated by setting up camera motion equations where some feature correspondences between frames have known depth values and some feature correspondences have unknown depth values

3. The depth map used for visual odometry step contains depth and two angles and is constructed using 2D KD-Tree

4. In the LIDAR odometry stage, sweep to sweep refinement stage involves computing modeling drift since different LIDAR point measurements are obtained at different times

5. Once the sweep point cloud is adjusted, it is registered to the point cloud representing the map of environment

6. The method is ranked at the top of KITTI benchmark, generally doesn't require post processing to improve map quality and works with both regular and fish-eye lens cameras.