Reading Assignment: ORB-SLAM/VLOAM

Submitted by: Shahram Najam Syed
Andrew ID: snsyed

ORB-SLAM Overview

In this paper, the authors presented a monocular SLAM approach that offers higher accuracy than previous methods, reporting sub-1cm error in indoor environments and errors of only a few meters in large outdoor videos. This enables the entire system to operate in real-time using only a CPU, without the need for GPU acceleration. ORB-SLAM is a traditional monocular SLAM (Simultaneous Localization and Mapping) system with features that operate at real-time rates, providing users with the capability for mapping and tracking simultaneously. It utilizes ORB (Oriented FAST and Rotated BRIEF) features and implements three parallel threads for tracking, local mapping, and loop closing. It also introduced a unique method for automatic initial relative motion estimation is incorporated within the system. Here, homography is selected for planar motions, while the fundamental matrix is chosen for non-planar motions. Statistically, the system had the capability to detect low-parallax cases and handle planar degeneracies. To reduce mapping complexity, a local covisibility graph is used for tracking and mapping, as well as for localized bundle adjustment. This approach implements a "survival of the fittest" policy regarding keyframes and map points, ensuring system operation as long as necessary. Loop closure is processed using a bag-of-words place recognition method, counting the number of detected features. A new loop closure technique, called the "Essential Graph," achieves similar results at a lower computational cost compared to full bundle adjustment.

V-LOAM Overview

V-LOAM has demonstrated impressive performance on the KITTI dataset, producing a relative position drift of just 0.75%. This robustness is attributed to its efficient use of both visual and lidar (laser) sensors, enabling high performance even under intense motion and challenging environmental conditions. V-LOAM is a real-time, visual-lidar odometry and mapping system that uses a monocular camera and 3D Lidar sensors. It operates through two processes: high-frequency (60Hz) visual odometry, which yields initial motion estimation, and low-frequency (1Hz) lidar odometry, which maintains these estimates over regions. The visual odometry utilizes lidar point cloud depth information to locate and track features, estimating spatial motion for each frame. Meanwhile, the lidar odometry compensates for visual odometry errors by correcting motion integration between lidar sweeps and maintaining registration accuracy. This system leverages geometric features, such as edge and planar points from lidar sensor data, for scan and trajectory extraction. The trajectory extraction process includes validity checks on predictions made by resulting models. Limitations in sweep sections are comprehensively addressed within the system's integrated framework.