Accuracy Evaluation of Appearance-based SLAM with semantic labeled vision features

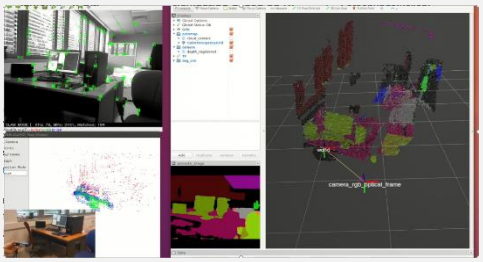
Contact Person: Haoming Zhang

Tel.: +49-241 92 7822 51

Email: haoming.zhang@ifu.rwth-aachen.de

Type of Thesis: Bachelor, Master

Language: German, English



© Real-time voxel based 3D semantic mapping [1]

Scientific Question Statement

The localization of a mobile robot in an unstructured unknown environment is mostly formulated as a smoothing state estimation problem and solved with Simultaneous Localization and Mapping (SLAM). Within the appearance-base SLAM approach, handcrafted visual features like SIFT, FAST, ORB etc. are extracted to enable the bundle adjustment in order to estimate the local camera pose at front-end and detect loop closure at back-end. Nevertheless, since the feature extraction takes all of matched inlier features of the viewed objects into account, which contain only simple geometric information, the state-of-the-art Visual SLAM algorithms might be volatile, such that the robustness of the state estimation cannot be guaranteed in all kinds of environments. One of the ideas to improve the robustness for a long-term SLAM approach is to consider the semantic information of the environment while extracting vision features and smoothing the trajectory. Several approaches [1, 2, 3] have been published, however, the impact of accuracy improvement with the consideration of different object categories for feature extraction and loop closure detection hasn't been researched yet. This should be addressed in this thesis.

Scientific Methodology

In order to integrate semantic information in the SLAM algorithm (Front-End / Back-End), an inference interface of an available semantic segmentation approach will be implemented into the open framework maplab [4], which is used for research in visual-inertial mapping and localization. Subsequently, a semantics-aware feature descriptor will be constructed based on semantic labeled image and available handcrafted visual features for the last step, where the visual features from different object categories will be evaluated while bundle adjusting. Optionally, the semantic information could also be constructed as a factor for graph optimization at the Back-End of the Algorithm.

Goal and expected results

- Interface of semantic segmentation Inference to maplab framework via ROS.
- Concept and implementation of a semantics-aware feature descriptor.
- Validation and Evaluation of new feature extractor in front-end visual odometry.

[1] X. Zhang, D. Filliat, Real-time voxel based 3D semantic mapping, ENSTAParisTech

[2] R. Salas-Moreno, A. Davison, SLAM++: Simultaneous Localization and Mapping at the Level of Objects, CVPR, 2013

[3] L. Nicholson, N. Sünderhauf, QuadricSLAM, IEEE Robotics and automation letters, 2019

[4] T. Schneider, R. Siegwart, maplab: An open framework for research in visul-inertial mapping and localization, CoRR, 2017