Visual Simultaneous Localization And Mapping Evaluation on a mobile robot platform

Data collection and object modeling are key parts of the development and evaluation of any vision-based grasping pipeline. It is very common to first estimate the object pose to infer how to grasp it, which requires an object model, and any vision-based grasping pipeline needs to be evaluated on the specific object of interest.

These tasks are however time-consuming for a human as they require some manual annotation and need to be repeated for every application.

In this work, we want to investigate the feasibility of an automated pipeline for data collection with automated annotation. A core component for such a pipeline is the ability to precisely track the camera pose across time. This general problem is known as simultaneous localization and mapping, or SLAM. Knowing the camera pose is a key component to object modeling, and can greatly reduce the annotation need, as scenes can be annotated once and reprojected in ever frame.

The aim of this work is to compare the accuracy of 2D lidar SLAM [1], various visual SLAM [2], [3] and marker-based SLAM.

Tasks

• Get familiar with ROS
• Select a subset of Visual SLAM methods (Starting with [2], [3], more can be added)
• Capture scenes on the Toyota Human Support Robot
• Use the estimated camera poses to reconstruct the scenes and compare the reconstructed objects models to high-quality object models to evaluate the pose quality.

Workload split

• Research and theory: 20%
• Programming and implementation: 50%
• Writing: 30%

Contact
Jean-Baptiste Weibel, weibel@acin.tuwien.ac.at
Markus Suchi, suchi@acin.tuwien.ac.at

References

