On Sensor Pose Parameterization for Inertial Aided Visual SLAM

Localization of pedestrians is a difficult task when no appropriate infrastructure is available. This is often the case in urban or indoor scenarios, where satellite navigation is hindered due to occlusions or multipath effects. A promising alternative is to combine a small, low-cost IMU with a camera in order to exploit the complementary error characteristics of these devices.

This is often done by estimating the position of landmarks and the pose or the trajectory of the sensor system simultaneously with a stochastic filter. An important aspect of the design of such an estimator is the parameterization of the sensor's pose. Thereby position and orientation are often estimated simultaneously in the filter's state but the relationship between position and orientation is frequently neglected. However, recently it was stated by several authors that it is beneficial to exploit the Lie-algebra structure of rigid body motions for pose parameterization in the context of visual SLAM. This can be done in terms of the twist parameterization, which is the counterpart to the representation of orientation by axis and angle for the group of rigid body motions.

In our work, we show that the twist parameterization of rigid body motions indeed leads to a different error function as compared with the pose parameterization frequently used in INS-GPS integration. We formulate the error state transition and the measurement equations for the twist parameterization as well as for the standard approach. Finally, both approaches are compared on two different simulated trajectories and on an indoor dataset that was obtained with a camera-IMU system that was attached to the torso of a pedestrian. Our results indicate that the twist parameterization does not outperform the standard approach. Although we employ an EKF to estimate the state, the results are of interest for alternative filtering approaches as well.