Semesterthesis – Final Presentation

Welcome everyone to the Final Presentation of my Semsesterthesis where I would like to present an evaluation of Two inhearently different Visual-Inertial Odometry Implementations called Rovio and Okvis.

-

Tasks like windmill inspection and maintenance, will be done by autonomous robots in the future. One key challenge towards this goal is Single Robot Localization. Subsequent tasks like coordinated work, dense reconstruction of a surface or real manipulation rely on an accurate localization – globally, the robot has to fly to the windmill, and locally for the inspection itself.

Visual-Inertial Odometry is one framework talking Localization. By combining the rich structure information of a camera with the short-time accuracy of an Inertial Measurement Unitan accurate Robot Localization can be achieved.

Advantages against other Localization Approaches are the reliance on lightweight and cheap sensors. As passive sensors they have a small power consumption and by relying solely on onboard sensors no external localization system is required.

-

I would like to present 2 main topics today. In the first part I will compare two Visual-Inertial Odometry Implementations by demonstrating their working principles, showing differences and demonstrating their accuracy.

In the second part I will show an analysis regarding the question if Visual-Inertial Odometries can work with non-timesynchronized Hardware.

-

Visual-Inertial Odometry is part of the larger field of Visual Odometry. The Visual-Inertial Odometries can be divided into 2 principally different approaches – Filtering-based and Keyframe-based.

Most Filtering-based Algorithms are based on an extended kalman filter and estimate the robot pose in 2 steps: Firstly they propagate the robot pose between frames based on the IMU measurements. Secondly, as soon as a new image is available, the estimate is updated based on the observation. The camera-observation of a past image is transferred into the current one with the propagation, so via the prior belief of the filter. But a past estimate will never be corrected based on the current observation.

On the other hand the keyframe-based approach estimates the robot pose by performing a nonlinear optimization based on a large number of landmark observations taken from a set of past keyframes and the current frame. Like that poses of the past are corrected based on the current observation as well.

Because of these inhearent differences the Keyframe-based approach often shows better accuracy as it tends to drift less over time.

On the other hand the filtering-based approach is in general less computational expensive while performing, especially locally, surprisingly accurate.

For this work I focused on Rovio and Okvis – two promising Implementations of the two approaches. Both have been developed here at ETH.

-

Rovio stands for Robust Visual-Inertial Odometry and it is using a direct Extended Kalman Filter Approach. What you can see on the slide is a Visualization of Rovio during operation. On the left side we see the current image Rovio is working with. Rovio is using a small number of features – in this moment around 20 – which are reprojected into the image.

Rovio detects new features with a fast corner detector and is working with Multilevel Patch Features. It includes new features into the filter state under certain conditions – they must not lie near an already tracked feature, their Intensity patches has to show reasonable intensity gradients and there needs to be free space in the filter state.

Rovio estimates the Robots Position, Velocity, Attittude, the IMU Biases, the Extrinsics between IMU and Camera and the features themselves. The features are represented in the filter state by their bearing vector wrt the camera frame and their distance to the camera.

On the right side we see a Visualization of Rovios estimates – In the center we see an inertial coordinate frame and the body frame (which is not yet well visible). The white dots represent the filters belief of the Features and the green line denote the 2 sigma bound of the feature uncertainty.

At the initialization all features are initialized with a high uncertainty

PLAY

As we can see, as soon as the camera is moved the uncertainty associated with the features is decreased.

One can observe, that Rovio often is working with edge features, which are not well suited for tracking as they can move along a line.

-

Okvis stands for Optimal Keyframe-based Visual-Inertial SLAM. In the Visualization we can see on top the latest keyframe and on the bottom the current frame.

Okvis detects new features with a multiscale Harris-Corner detector and describes them as Brisk features. In the image, the extracted features are are shown as circles, the green lines highlight matching pairs.

Okvis is estimating the robot pose by performing a nonlinear optimization minimizing an error function containing reprojection error terms and IMU error terms. Like that it fullfills an optimization of all keyframe poses, the newest frame poses, speed and imu-bias terms and all landmark positions.

In order not to grow unbounded, Okvis performs a Marginalization of old keyframes, their landmarks and Speed/Bias Terms.

On the right side of the image you can see a Visualization of the landmarks. All landmarks currently used in the optimization are highlighted in green.

PLAY

One remark regarding Okvis: We have not been able to run it on our machines until today. The OKVIS results I present have been generated some time ago on other laptops than mine.

-