A Keyframe-Based Sliding Window Filter

Jianzhu Huai[1], Yukai Lin[2]

I. System Overview

The system has a feature matching frontend that processes images, a sliding window filter backend that fuses landmark observations and IMU readings. The input to the system are stereo images and IMU readings captured by the snapdragon sensor rig. The system causally processes sensor messages synchronously by reading from rosbags and saves online estimates of poses of the IMU sensor frame relative to a world frame.

II. Frontend Feature Matching

The feature association in the frontend works similarly to the frontend of OKVIS [1]. Back-to-back frame matching, frame-to-keyframe matching, and left-to-right stereo matching are used to associate point landmark observations. All these matching are done by matching BRISK descriptors in brute force.

III. Backend Sliding Window Filter

The backend employs a structureless sliding window filter similar to MSCKF with first estimate Jacobians [2]. The backend is initialized with a few IMU readings. The initial orientation is determined such that the world frame has z-axis along negative gravity. The initial position and velocity are set to zeros. The biases are averaged from the IMU readings assuming the data sessions start from standstill. Camera extrinsic and intrinsic parameters are initialized to calibrated values. To bound computation, redundant navigation states are selected from the sliding window and marginalized depending on whether they are tied to keyframes somewhat similarly to [1].

IV. Timing

A commodity laptop, Dell Insipiron 7590, is used to process all data sessions. It has 16 GB RAM and a Intel Core i7-9750h CPU, with 6 cores operating at 2.60 Ghz. The program runs in Ubuntu 18.04 without GPU acceleration. Because the data are processed synchronously, the whole system works like using one thread. Actually, the feature matching may run in several threads, but the filter update runs in one thread.

The time taken to process each sequence is listed in Table. I.

TABLE I: Processing time for each data session.

Session	Data Duration [s]	Processing Time [s]
indoor_45_3	76.14	21.19
indoor_45_16	46.36	13.06
indoor_forward_11	78.63	31.96
indoor_forward_12	58.50	27.91
outdoor_forward_9	88.16	51.40
outdoor_forward_10	111.07	51.46

V. Parameters

The same set of parameters is used throughout all sequences. The projection and distortion parameters for both snapdragon cameras are fixed in estimation. The rolling shutter effect is ignored as the cameras use global shutters. The extrinsic parameters and temporal delays of the two cameras relative to the IMU are estimated online. The initial standard deviation is 0.02 m for each dimension of the lever arms, 0.01 rad for each dimension of the perturbation of the relative orientations, 0.005 sec for each temporal delay. As for the sliding window queue, at least 5 recent temporally consecutive frames are kept at the tail, and aside from these frames, at most 5 spatially separated keyframes are kept at the head. For the IMU noise driving the system covariance propagation, we use the default parameters which are shown in Table II.

TABLE II: IMU noise parameters.

σ	Gyroscope	Accelerometer
Initial Bias Std Dev	$0.03\mathbf{I}_3\ rad/s$	$0.1\mathbf{I}_3\ m/s^2$
Bias White Noise	$\sigma_{b\omega} = 4 \cdot 10^{-5}$ $rad/s^2/\sqrt{Hz}$	$\sigma_{ba} = 2 \cdot 10^{-3}$ $m/s^3/\sqrt{Hz}$
White Noise	$\sigma_\omega = 0.05$ $rad/s/\sqrt{Hz}$	$\sigma_a = 0.1$ $m/s^2/\sqrt{Hz}$

No loop closure or relocalization is performed in processing the test data sessions.

References

[1] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, “Keyframe-based visualinertial odometry using nonlinear optimization,” *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, Mar. 2015. [Online]. Available: https://doi.org/10.1177/0278364914554813 1

[2] M. Li and A. I. Mourikis, “High-precision, consistent EKF-based visual-inertial odometry,” *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, May 2013. [Online]. Available: https://doi.org/10.1177/0278364913481251 1

[1] J. Huai was with CEGE, The Ohio State University, Columbus, OH

[2] Y. Lin is with ETH Zurich, Zurich, Switzerland