Stereo Visual-Inertial Odometry

Xin Zhang

1. A brief overview of the approach

The algorithm is based on incremental non-linear optimization, with a fixed-length sliding window. It is causal, i.e. does not use information from the future epochs to obtain the pose at current epoch and hence there is no bundle adjustment of the entire history of the poses. No loop closure is used. Key frames are selected to speed up the processing and prepare the algorithm for real-time operations.

2. The data and sensor modalities

Event data is not used.
We use images from the left and right cameras and IMU (Inertial Measurement Unit) onboard the Qualcomm Flight Board, which provides 640x480 grayscale images, hardware-synchronized with the IMU. Although a ROS version was developed for this competition, dataset in .zip files instead of the ROS bags were used to produce the submission. General information about the data is summarized in Table I of [1] and the coordinate frame definition is illustrated in Fig. 4 & 6 of [1].

3. Frontend

OpenCV implementation of FAST (Features from Accelerated Segment Test) [2] detector is used. Other candidates are but are not limited to Shi-Tomasi, Harris, ORB, A-KAZE, BRISK, BRIEF, FREAK, SIFT, and SURF. During tests of the algorithm on other UZH-FPV datasets, FAST stood out as our choice. OpenCV implementation of Lucas-Kanade tracker [3] is used.

4. Backend

The backend is gtsam 4 [4-6]. IMU preintegration factor, an implementation of the 6-way factor family, is utilized to summarize behaviors of consecutive states. Since some of the dataset exhibits a high dynamic motion, the lengths of the sliding window are adapted to reflect this change accordingly.

4.1 Initialization

For each data stream, we only use data after the 50^{th} epoch. Although the first 50 samples are excluded, the remaining samples still allow for a reasonable initialization, with the visual and IMU measurements.

4.2 Total processing time for each sequence

*Table 1 Timing**

	Sequence	**Timing (s)**	Length(s)
1	Indoor_forward_11	**96.06**	85.68
2	Indoor_forward_12	**72.71**	124.07
3	Indoor_45_3	**276.13**	119.82
4	Indoor_45_16	**59.18**	58.72
5	Outdoor_forward_9	**276.46**	314.41
6	Outdoor_forward_10	**383.95**	455.63

* This timing is measured on processing time of the entire data, not from the start time ('Start Time Snapdragon') instructed by the competition portal (https://fpv.ifi.uzh.ch/?sourcenova-comp-post=2019-2020-uzh-fpv-benchmark).

Big timing differences between two similar settings (e.g. indoor_45_3 vs indoor_45_16) can be observed while they have similar length of input data length. This is because in some high dynamic settings, the processing interval should be tuned small to accommodate for the motion changes. This might undermine our efforts to make the algorithm oriented for real-time applications and strategy optimization is anticipated.

5. Exact specifications of the hardware used

The tests were run on a virtual machine (VMware) installed on a Macbook pro 16-inch, allocated with 4-processing (physical) cores (each rated 2.4 GHz) and 16 GB of RAM. The hard disk is a 4TB SSD. The operating system is Ubuntu 16.04.5.

References

[1] J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler and D. Scaramuzza, "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset," *2019 International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, 2019, pp. 6713-6719, doi: 10.1109/ICRA.2019.8793887.

[2] E. Rosten, T. Drummond, "Machine Learning for High-Speed Corner Detection," Leonardis A., Bischof H., Pinz A. (eds) Computer Vision – ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, vol 3951. Springer, Berlin, Heidelberg, 2006., pp. 430-443.

[3] J. Bouguet, "Pyramid Implementation of the Lucas Kanade feature tracker," Intel Corporation, Microcprocessor Research Labs, 2000.

[4] https://github.com/borglab/gtsam

[5] L. Carlone, Z. Kira, C. Beall, V. Indelman, and F. Dellaert, "Eliminating Conditionally Independent Sets in Factor Graphs: A Unifying Perspective Based on Smart Factors," Int. Conf. on Robotics and Automation (ICRA), 2014.

[6] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation," Robotics: Science and Systems (RSS), 2015.