Northeastern University

Final Report

EECE 5554 Robotics Sensing and Navigation

VIO in vSLAM using OpenVINS

Instructor & TA
Dr. Thomas R. Consi
Anish C. Pius
Arunbhaarathi Anbu
Skanda A. Prasanna
Syed M. Asjad
Dhruv D. Shah
Pavan R. Shetty

Group Members
Abdul Al-Badawi
Yuzhi Gu
Haiduo Liu
Abdusattor Sattarov
Luoyan Zhang

December 17, 2022

1 Introduction

This project will demonstrate Visual Inertial Odometry (VIO) techniques to estimate the local position and orientation of an object and velocity of a car relative to its starting position, in Visual SLAM on datasets with mono and stereo camera setups along with an IMU sensor.

The VIO techniques can be used to perform Simultaneous Localization and Mapping (SLAM) to aid robotics navigate previous unseen and unknown environments using visual tracking support from camera inputs and an inertial measurement unit (IMU). There are a large variety of applications for using VIO, such as augmented reality (AR) and indoor navigation that can potentially aid a visually impaired person to navigate indoor and outdoor environments.

2 Background

Compared with other similar Simultaneous Location and Mapping methods, OpenVINS greatly reduces the drift in the monocular case, while it has a smaller impact on the stereo performance. The monocular system in OpenVINS clearly outperforms the current open sourced codebases. [1]

The main algorithm in OpenVINS utilizes Multi-State Constraint Kalman Filter (MSCKF). As a combination of IMU and Vision data under the EKF frame, it can adapt to more intense movements, and it also performs better in tracking.

3 Sensing Principle & Expected Sensor Performance

In order to have OpenVINS work properly, the IMU sensor and Camera must be calibrated. The Kalibr toolbox was recommended to be utilized to calibrate both devices. A calibration board image (Fig. 1) was needed to be printed and utilized to calibrate the camera. Multiple images should have been taken of the board under various angles. IMU sensor was intrinsically calibrated using the same toolbox. A standard Allan Variance plot was needed to be created to calibrate the device. After calibrating each device, the sensors had to be dynamically calibrated. The same calibration board was utilized and its images were taken under various angles, but this time also with a translational motion. However, the data obtained from NUANCE came already with calibration information. This information was used to get OpenVINS running properly with the data.

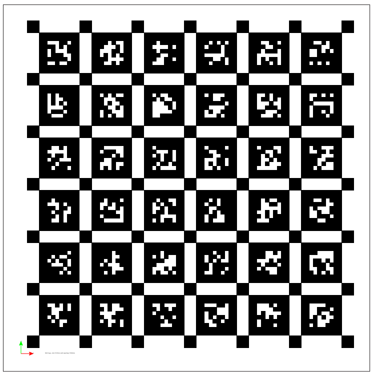

Figure 1: Calibration Board [2]

NUANCE dataset was extremely large and hard to download. To ensure that the program works well with the sensors and data, EuRoC standard dataset was used. This dataset also came with calibration information that was used to run OpenVINS. Nevertheless, when NUANCE data was acquired, it was used as the main dataset.

It was important to calibrate both sensors because even minor errors could result in large biases and disturbances in mapping. The actual IMU sensor was mounted at the roof of the NUANCE car while the camera was set at the front. It needs to be noted that improper mounting could have resulted in poor data or in data with large biases.

4 Data Collection and Running Datasets

The open-source OpenVINS project [1] has the calibration and configuration that supports datasets from 6 different organizations, but it could also run on other datasets like the one from NUANCE as long as we have the proper calibration information and configuration.

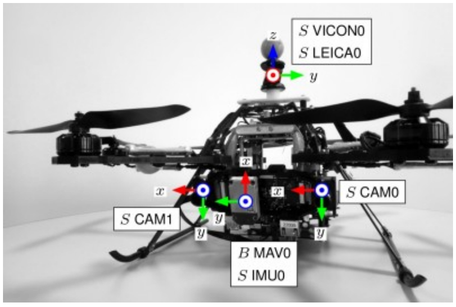

Figure 2: EuRoC MAV Platform [3]

Figure 3: NUANCE Platform [4]

4.1 EuRoC MAV

The ETH ASL EuRoC MAV dataset [3] is one of the most used datasets in the visual-inertial / simultaneous localization and mapping (SLAM) research literature and it is the first dataset our team tried on. The reason for this is the synchronized inertial and camera sensor data and the high quality groundtruth. The dataset contains different sequences of varying difficulty of a Micro Aerial Vehicle (MAV) flying in an indoor room. Monochrome stereo images are collected by two Aptina MT9V034 global shutter cameras at 20 frames per seconds, while an ADIS16448 MEMS inertial unit provides linear accelerations and angular velocities at a rate of 200 samples per second.

4.2 NUANCE

The NUANCE dataset uses a Lincoln Mkz equipped with a sensor suite from NUFR Lab in Northeastern University. It has a linear array consisting of 5 pointgray BlakflyS 5MP cameras, two Velodyne VLP 16 Lidars, a Vectornav VN-100 IMU, and a GPS. [4]

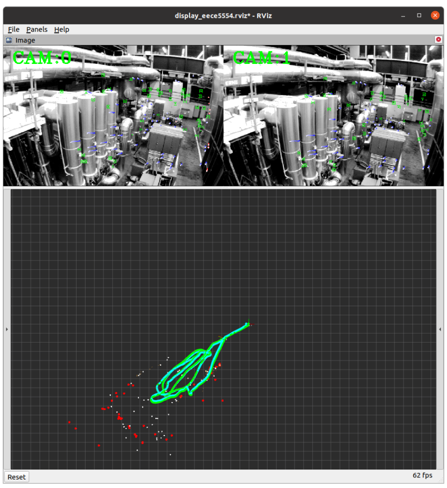

Figure 4: RViz of OpenVINS on EuRoC

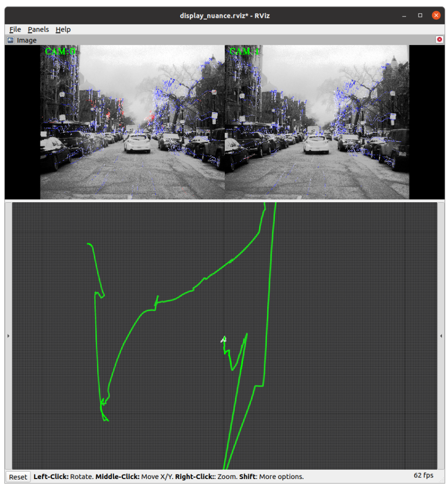

Figure 5: RViz of OpenVINS on NUANCE

5 Algorithm

In the OpenVINS, the main algorithm implanted is Multi-State Constraint Kalman Filter (MSCKF), which is based on Extended Kalman Filter algorithm and designed for real-time vision-aided inertial navigation.

The main structure of the Multi-State Constraint Filter is a loop of 3 steps:

1. **Propagation**: Receive IMU measurement, then propagate the filter state and covariance.

2. **Image Registration**: Every time a new image is recorded,

   a. Augment state and covariance matrix with current camera pose estimate.

   b. Start image processing module.

3. **Update**: Perform EKF update when feature measurements of a given image is available.

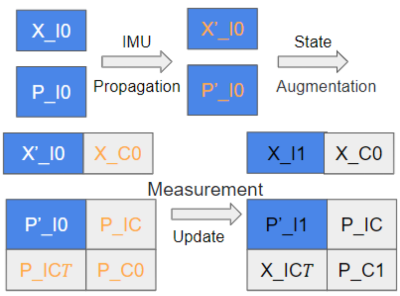

Figure 6: Multi-State Constraint Filter

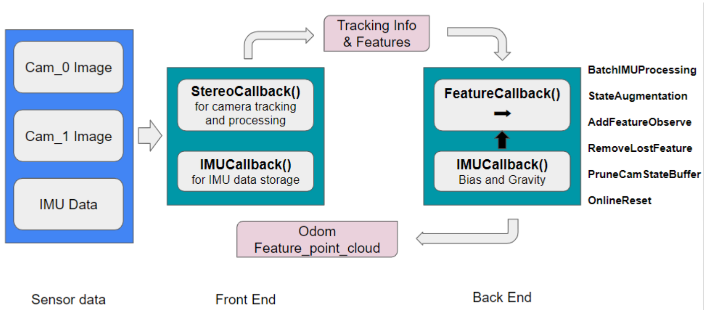

Figure 7: ImageProcessor Front End & MSCFK Back End (with error-state EKF)

OpenVINS provides three different trackers. The feature points extracted from the second image frame will be used to track the feature point from the first image. Then the loss feature points will be deleted and we move on to the next circle. In the next circle the first image frame will be replaced by the second image frame and the second image frame will be replaced by the third one. This is a brief process of tracking.

The main feature of the Multi-State Constraint Kalman Filter [5] is to utilize the localization information provided by multiple measurements of visual features, resulting in a computational complexity linear with the number of features.

As a combination of IMU and Vision data under the EKF frame, it can adapt to more intense movement, while maintaining functioning with texture loss for a certain period of time, therefore being more robust. Compared with other VIO algorithms (VINS, OKVIS), MSCKF adopts a structureless approach where landmark positions are marginalized out of the state vector, therefore, a small number of landmarks are typically tracked to allow real-time operation. As is mentioned in the paper[5], images recorded at 3 Hz can be used for estimation algorithm and generate results at 14 Hz, with a single core of Intel T7200 processor (2GHz clock rate), tracking along 142903 features for EKF updates within a 3km recorded trajectory.

The NUANCE dataset of our project is 50 Gigabytes and is acquired through embedded platforms. In this case, selecting MSCKF is reasonable since it provides faster output and same accuracy as other EKF does, and can be deployed on embedded platforms with limited computing resources.

6 Analysis

Data analysis was performed on EuRoC Dataset in Machine Hall and Vicon Room 1 with medium difficulty. The datasets with a 'medium' difficulty were chosen as the 'difficult' level and the Vicon room 2 'medium' contained an inaccurate estimated trajectory for Machine Hall, there was an error in Vicon room 1 'easy' dataset. The desired quantities to analyze are position and orientation as all ground-truth and trajectory estimation datasets, from all rooms and difficulty levels, containing those quantities in-common in addition to the timestamp (in nanoseconds).

The x and y axes of estimated trajectory and ground-truth were plotted against each other, in the case of the 2D plots, and x, y and z axes are plotted together in the case of the 3D plots. The plots of the estimated trajectory plots provide an indication to how the path that was taken by the MAV compared to the ground truth ground-truth.

To obtain orientation, the attitude in Quaternion format of the estimated trajectory and ground-truth are converted to Euler angles, using the MATLAB built-in function `quat2eul`. From there, the Euler angles from the estimated trajectory and ground-truth were plotted, individually, with their respective timestamps.

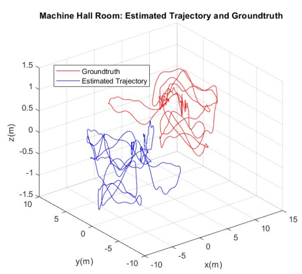

Figure 8: 3D Plot of Traj. in Machine Hall

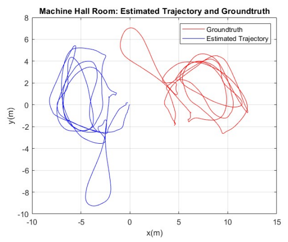

Figure 9: 2D Plot of Traj. in Machine Hall

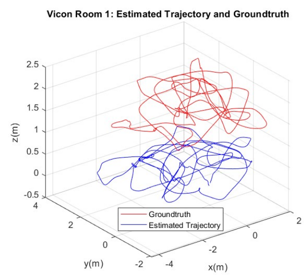

Figure 10: 3D Plot of Traj. in Vicon Hall

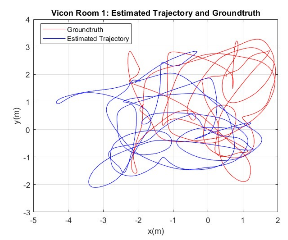

Figure 11: 2D Plot of Traj. in Vicon Hall

7 Conclusion

The output from the Quaternions in the trajectory estimate is relatively different from the quaternions in the groundtruth with respect to each room's datasets. Hence, after conversion from Quaternions to Euler angles, regardless of order of rotation, the Euler angles from the trajectory estimate and groundtruth datasets for each room are also relatively different. Plotting the pose of the MAV, using Euler angles, with respect to each dataset was unintuitive.

The estimated trajectory of the MAV in each room seemed to relatively match that of the groundtruth, however, the orientation of the estimated trajectory MAV seems to be incorrect as it is rotated in a direction that does not match the ground-truth. The reason behind the 'mis-oriented' estimated trajectories in both rooms, is the fact the Euler angles from the estimated trajectory are not the actual representation of how the MAV was oriented whilst navigating the rooms throughout its path. Misalignment errors is a common error in inertial sensors, where an angular difference between the euler angle's axis of rotation and the global frame of the MEMS IMU exist. Lastly, OpenVINS reports that a few of the datasets showcase hugely dynamic motions, which decays measurement accuracy of the laser tracking device. Given the highly accurate nature of visual odometry approaches, the interpretation of the ground-truth may be inaccurate based on stated figures, not mentioned in OpenVINS, provided by the manufacturer.

References

[1] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A research platform for visual-inertial estimation," in *Proc. of the IEEE International Conference on Robotics and Automation*, Paris, France, 2020. [Online]. Available: `https://github.com/rpng/open_vins`.

[2] ethz-asl, *The kalibr visual-inertial calibration toolbox*, 2022. [Online]. Available: `https://github.com/ethz-asl/kalibr`.

[3] M. Burri, J. Nikolic, P. Gohl, *et al.*, "The euroc micro aerial vehicle datasets," *The International Journal of Robotics Research*, 2016. DOI: `10.1177/0278364915620033`. eprint: `http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.full.pdf+html`. [Online]. Available: `http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.abstract`.

[4] F. R. L. in Northeastern University. "Autonomous car | field robotics lab." (), [Online]. Available: `https://fieldroboticslab.ece.northeastern.edu/autonomous-car/`. (accessed: 12.16.2022).

[5] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint kalman filter for vision-aided inertial navigation," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 3565–3572. DOI: `10.1109/ROBOT.2007.364024`.