HILTI CHALLENGE SUBMISSION REPORT

Dr Thien-Minh Nguyen – Nanyang Technological University, Singapore

I. BRIEF OVERVIEW

This report presents the results of VIRAL SLAM with the datasets provided by the HILTI SLAM Challenge [1]. More details on the structure of VIRAL SLAM can be viewed at [2] & [3].

To begin, we shall quickly answer the questions posted on the website:

- A brief overview of the approach:
  - Filter or optimization-based (or else)?

    Answer: The method is optimization-based, specifically optimization of a cost function consisting of lidar, inertia and visual factors on a sliding window, using the ceres solver.

  - Is the method causal? (i.e. does not use information from the future to predict the pose at a given time).

    Answer: Yes. The method does not use any future information. All estimates are derived from the sensor measurements obtained at the present and the past, along with the marginalized information.

  - Is bundle adjustment (BA) used? What type of BA, e.g., full BA or sliding window BA?

    Answer: The *global pose graph optimization* (full BA) is disabled in all experiments. The sliding window optimization (sliding window BA) is carried out at 5Hz or 10Hz rate, and is synchronized with the arrival of messages under `/alphasense/cam0/image_raw` topic. The state estimate in the latest time step on the sliding window is published as the odometry (after each optimization) and is used for evaluating the accuracy.

  - Is loop closing used?

    Answer: No. We find that the sequences mostly capture trajectories within a 100m area, thus no significant drift is observed to require loop closure. In fact, in BASEMENT_2 sequence, DBoW-based loop detection may cause erroneous loop closure between two corridors since visually and geometrically they are very similar.

- Exact sensor modalities used (IMU, stereo or mono, LIDAR data?)

Answer: For all the experiment results submitted, we employ the data from the OS0 Ouster, Livox Avia, the 800Hz ADIS IMU and cam0-cam1 stereo image pair. All sensor data are processed in a tightly coupled framework. The results submitted have been transformed from the 800Hz IMU coordinate frame to the 200Hz IMU coordinate frame.

- Total processing time for each sequence and the used hardware

Answer: Each sequence is processed in real-time, i.e. the process completes when the bag file reaches the end. The experiment is conducted on a consumer laptop with 8-core Intel Core-i7 10875H CPU, 32GB RAM.

- Whether the same set of parameters is used throughout all the sequences

Answer: Yes. We use a single parameter set for all the experiments. The only change is the update rate, as some datasets are recorded in high-texture environments, so the update rate was reduced to keep the real-time performance.

II. RESULT SUMMARY

Table 1 gives a summary of our experiments.

Table 1. ATE and other statistics of the experiments

SEQUENCE	VIRAL ATE (m)	Process Time (s)	Update Rate (Hz)
BASEMENT_1	0.0258	108.386	10
Basement_3	-	325.448	5
Basement_4	0.0694	346.514	5
Campus_1	-	425.517	5
Campus_2	0.0423	371.292	5
Construction_Site_1	-	194.622	10
Construction_Site_2	0.0563	394.925	10
LAB_Survey_2	0.0211	131.6476	5
IC_OFFICE_1	-	195.808	10
OFFICE_MITTE_1	-	259.44	10
PARKING_1	-	578.516	5
uzh_tracking_area_run2	0.1777	85.1641	5

We present the pointcloud maps and the estimated trajectory in the figures bellow. Note that the plane features are marked by cyan, and the edge features are illustrated by orange. For the legend, VIRAL is the normal odometry estimate, VIRAL-KF is the key frames.

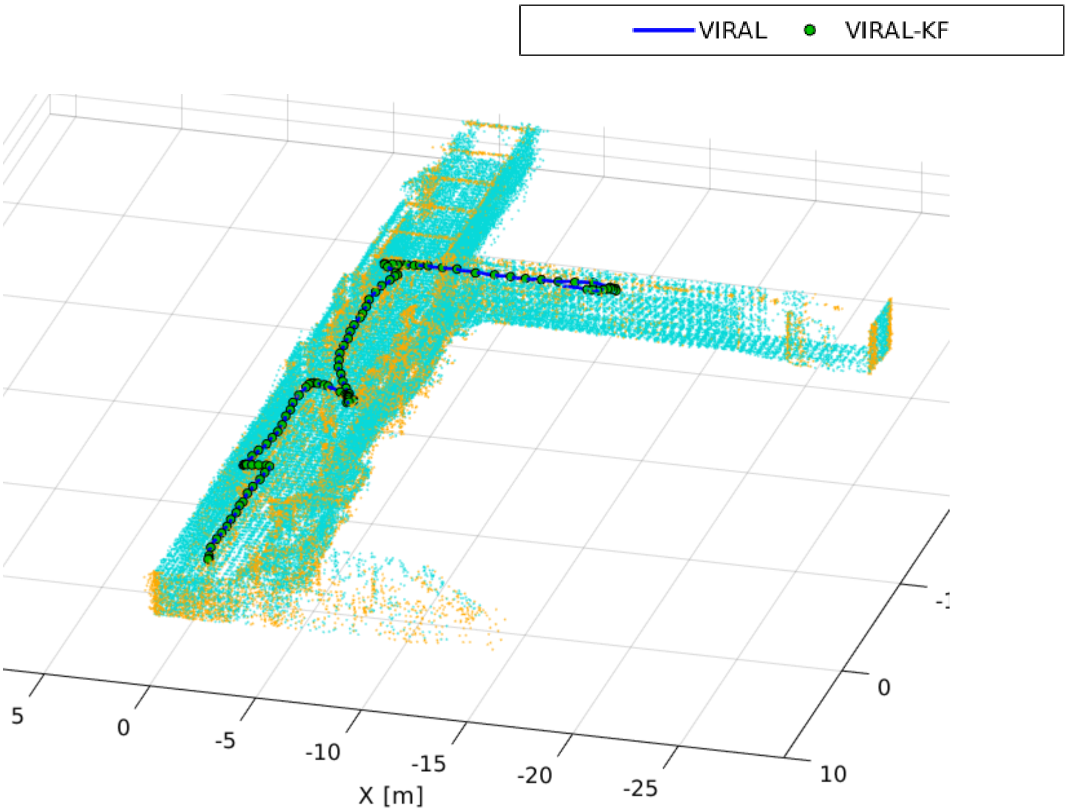

Figure 1. BASEMENT_1

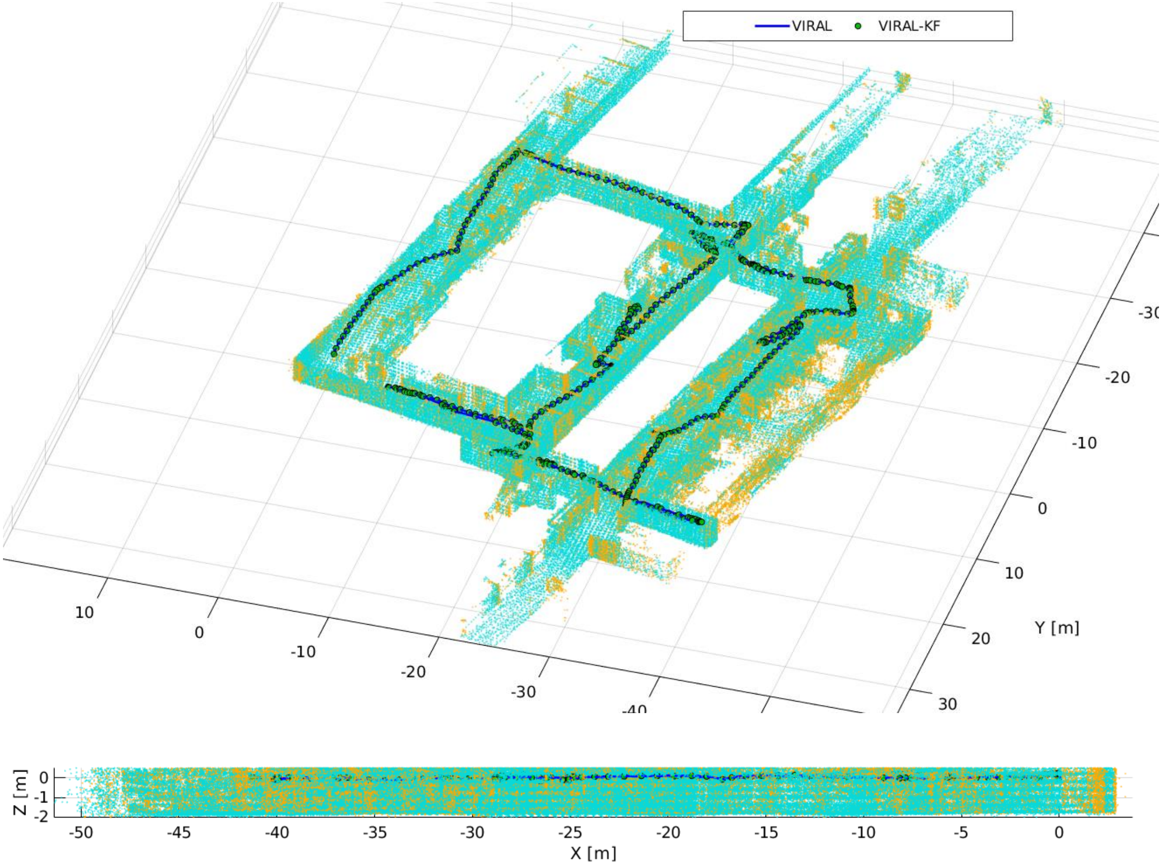

Figure 2. BASEMENT_2. In this sequence we segment the floor using the first few key frames and impose a constraint on the vertical direction based on the floor plane coefficients. The figure at the bottom shows that the vertical drift can be effectively.

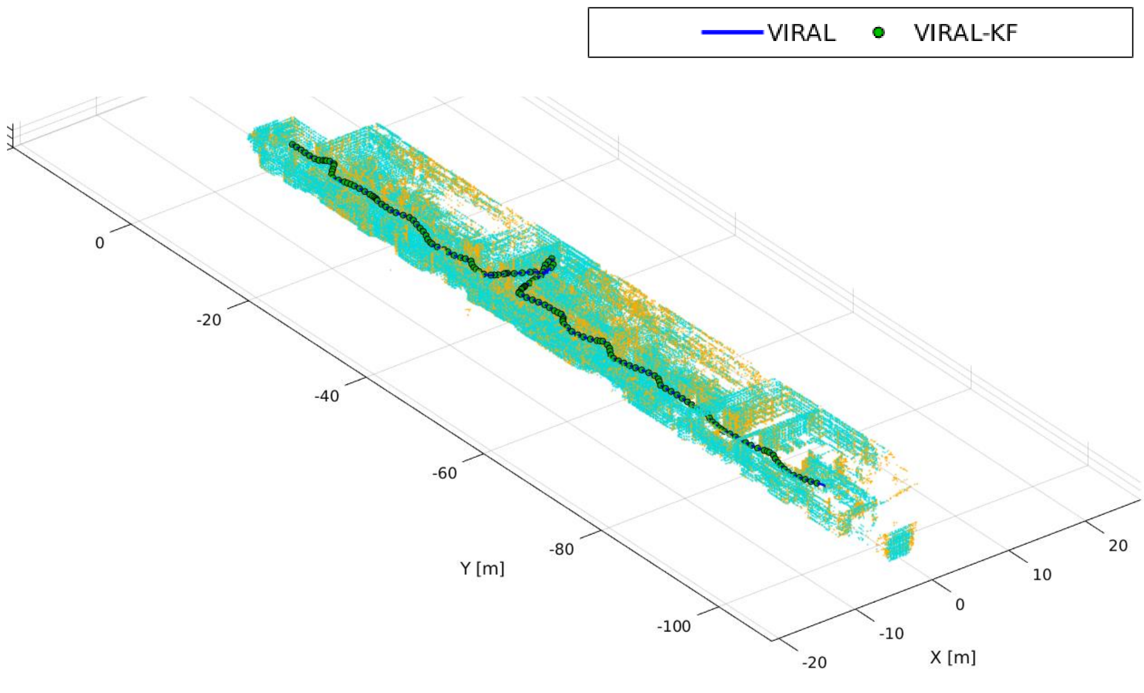

Figure 3. Basement_3

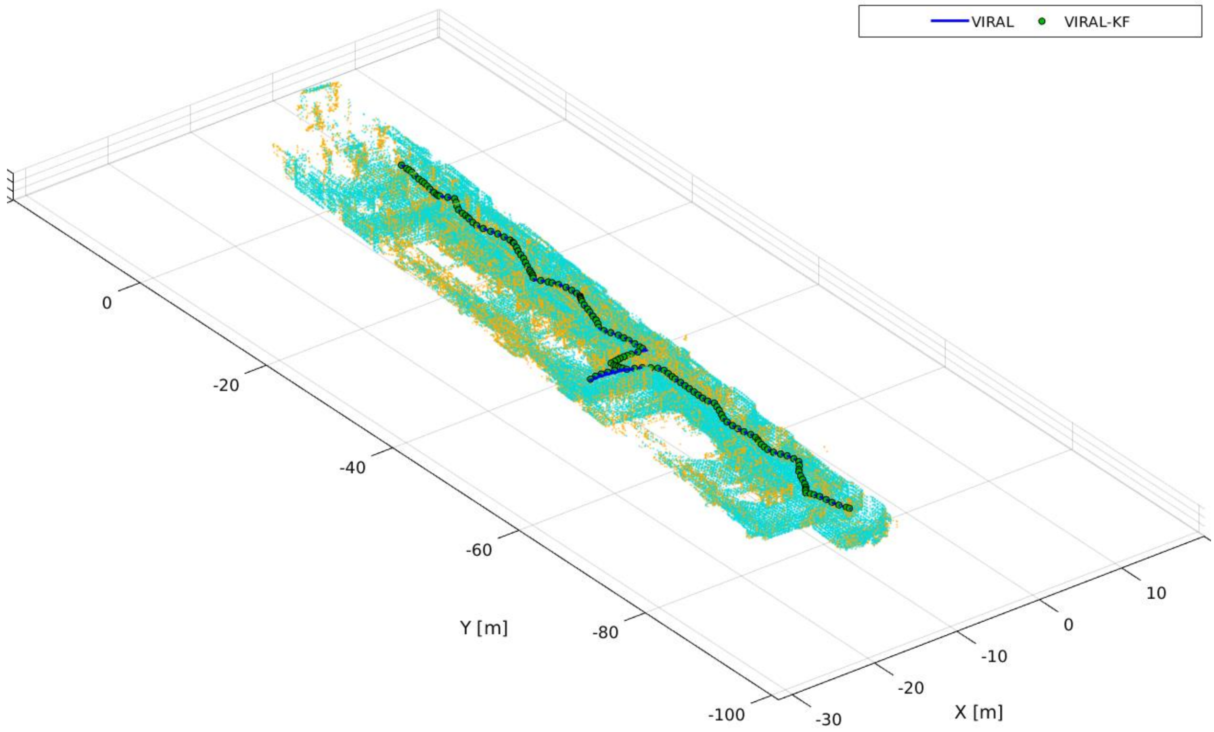

Figure 4. Basement_4

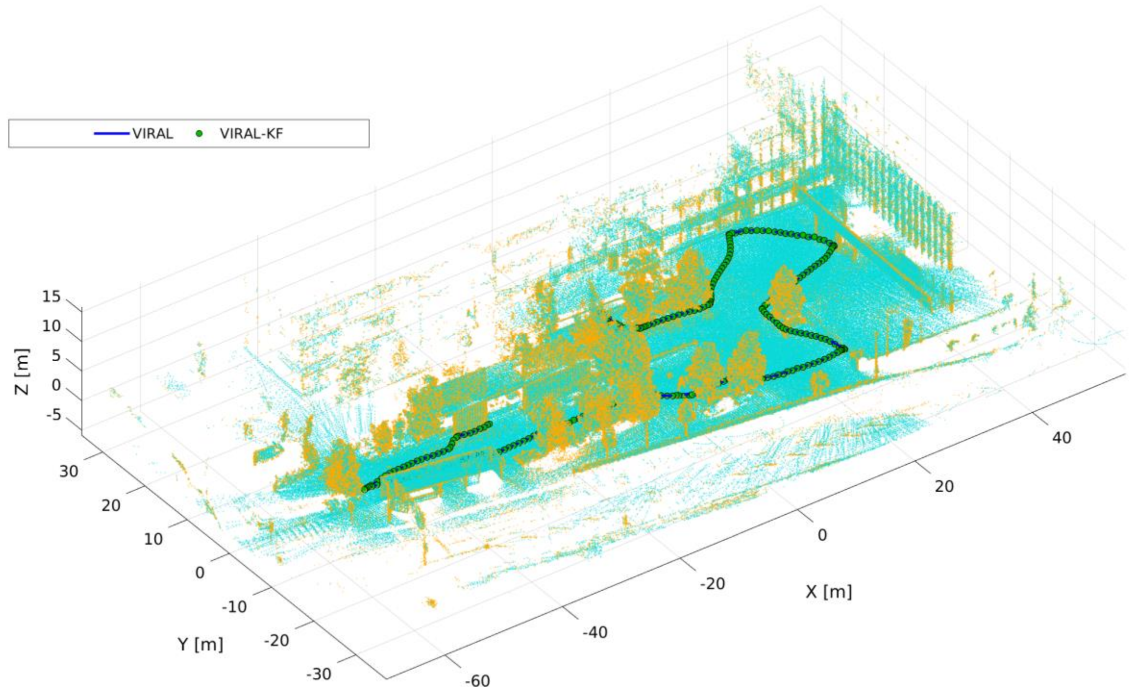

Figure 5. Campus_1

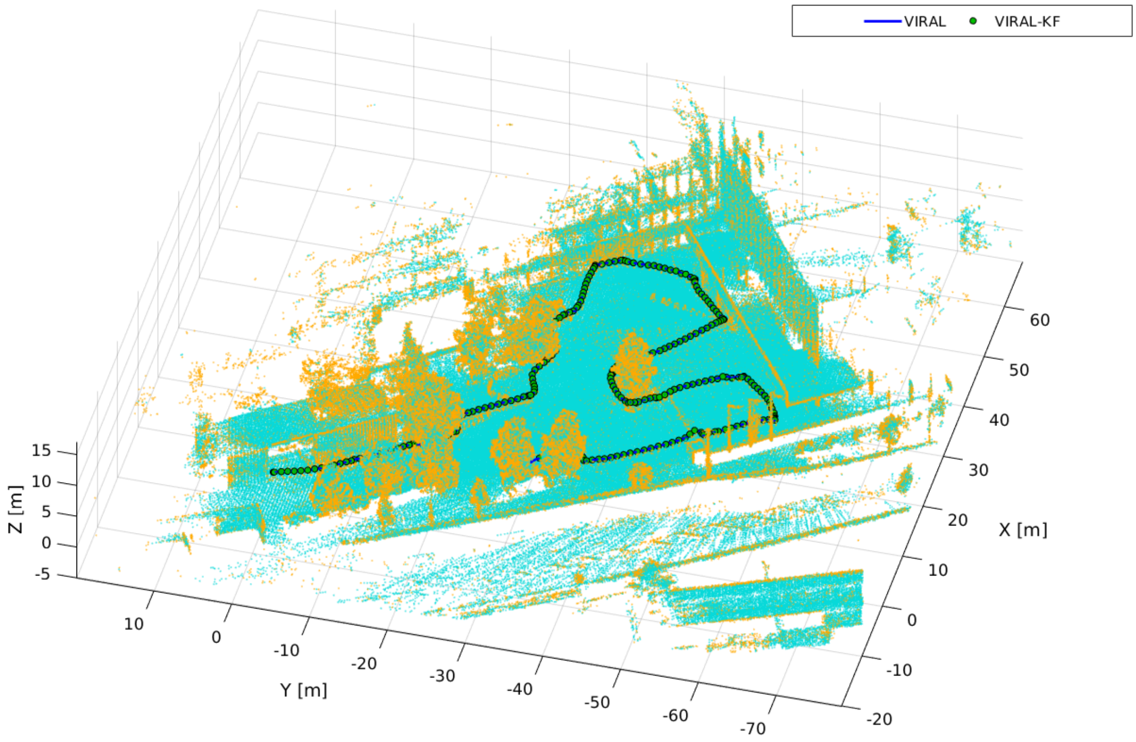

Figure 6. Campus_2

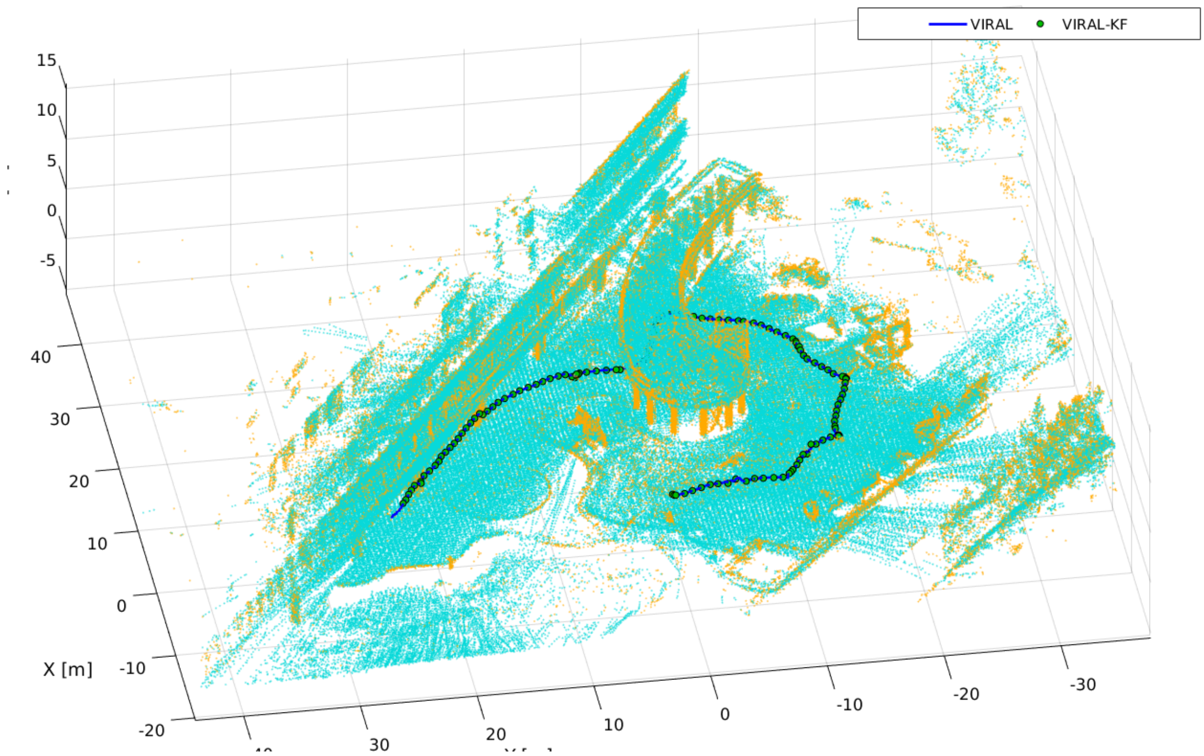

Figure 7. Construction_Site_1

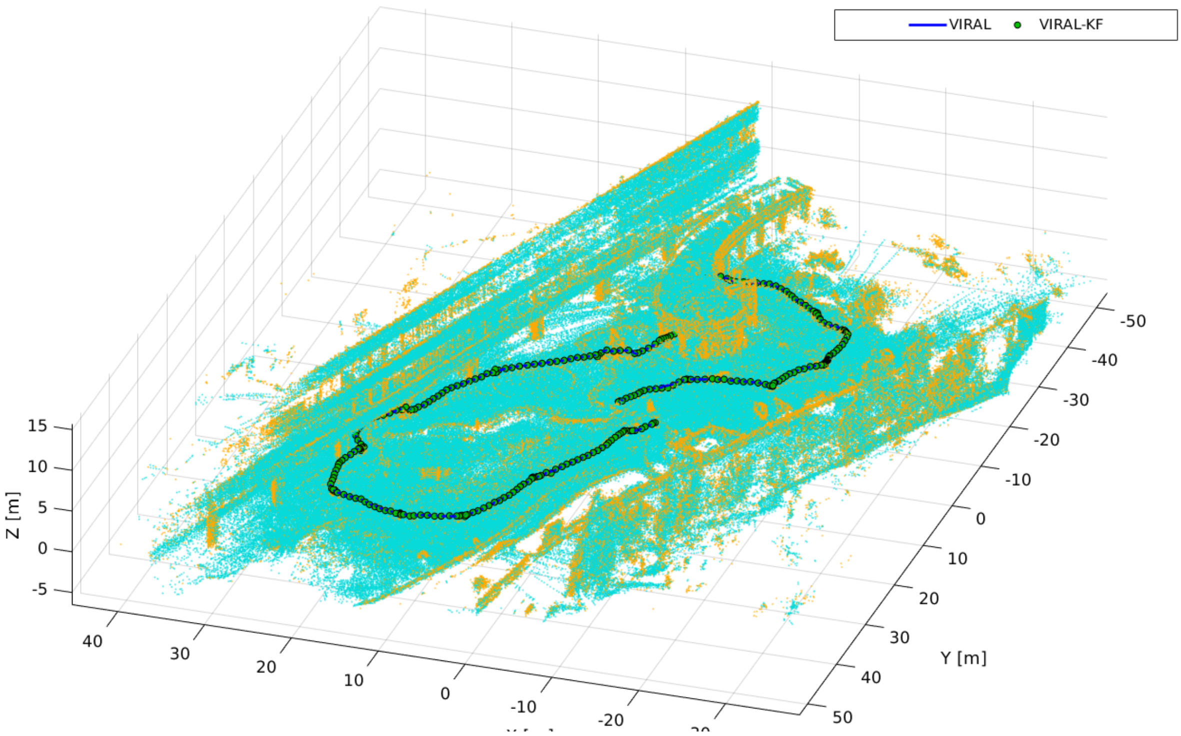

Figure 8. Construction_Site_2

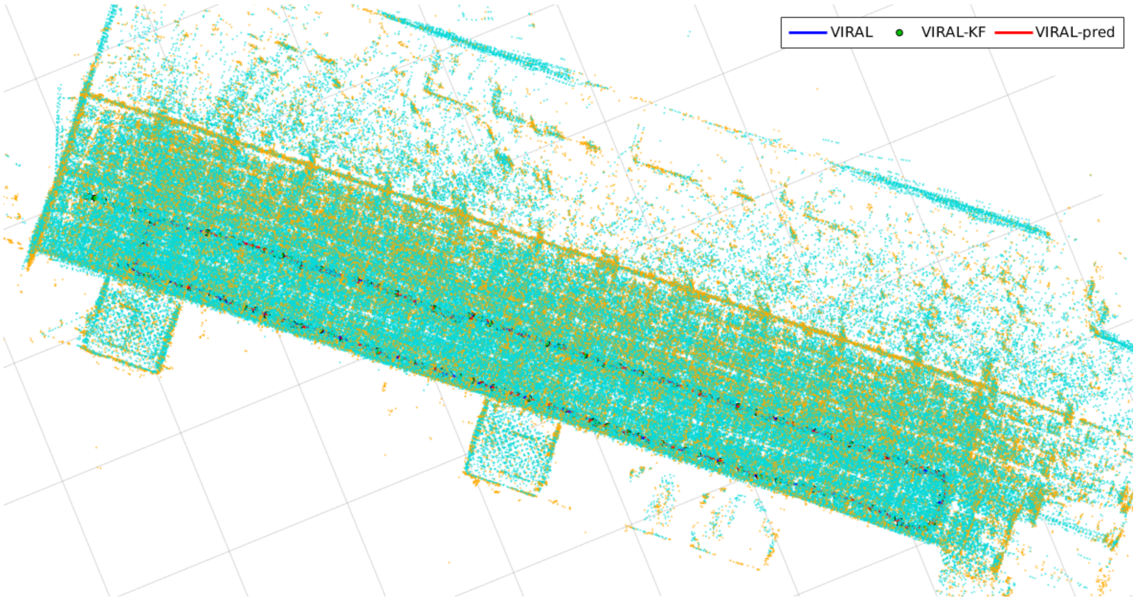

Figure 9. IC_OFFICE_1

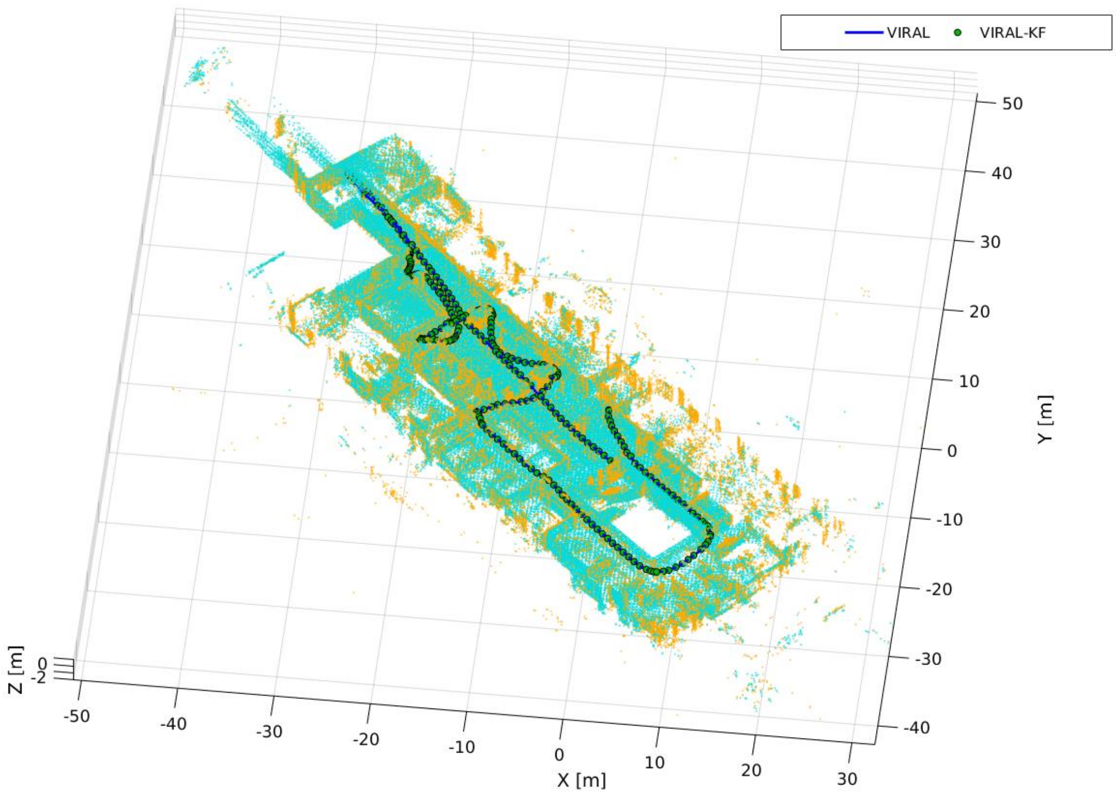

Figure 10. OFFICE_MITTE_1

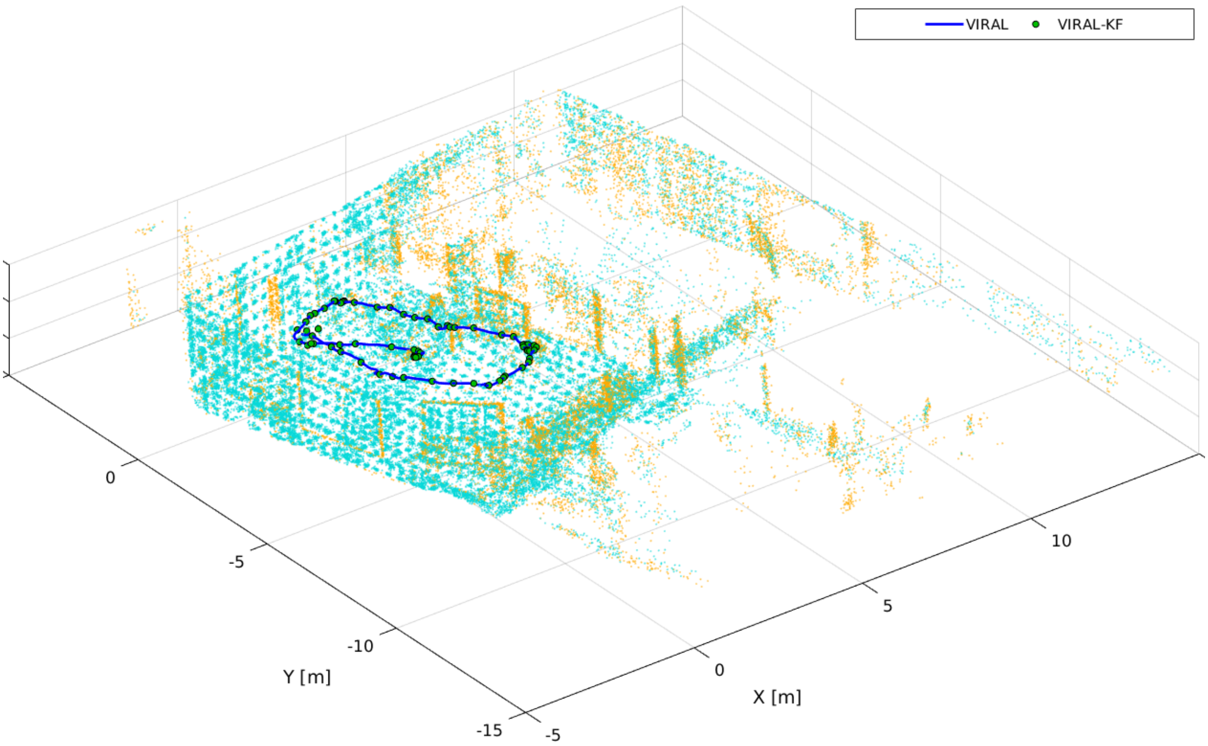

Figure 11. LAB_Survey_1

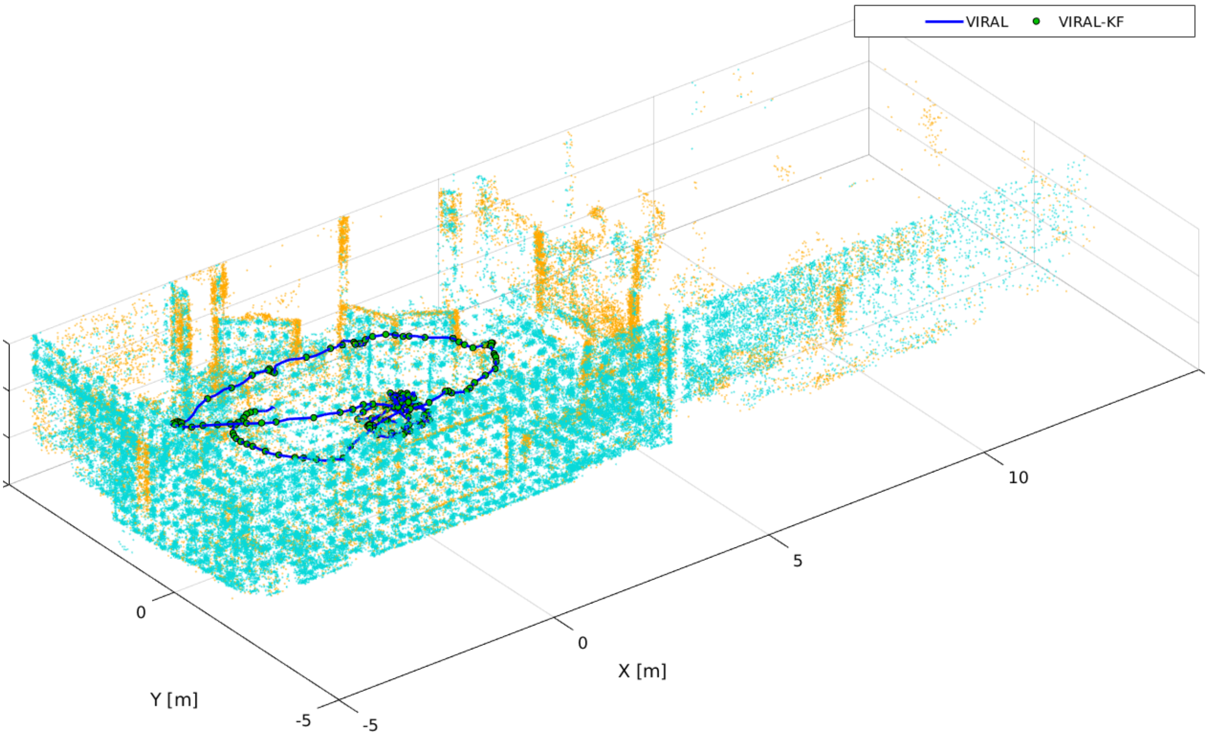

Figure 12. LAB_Survey_2

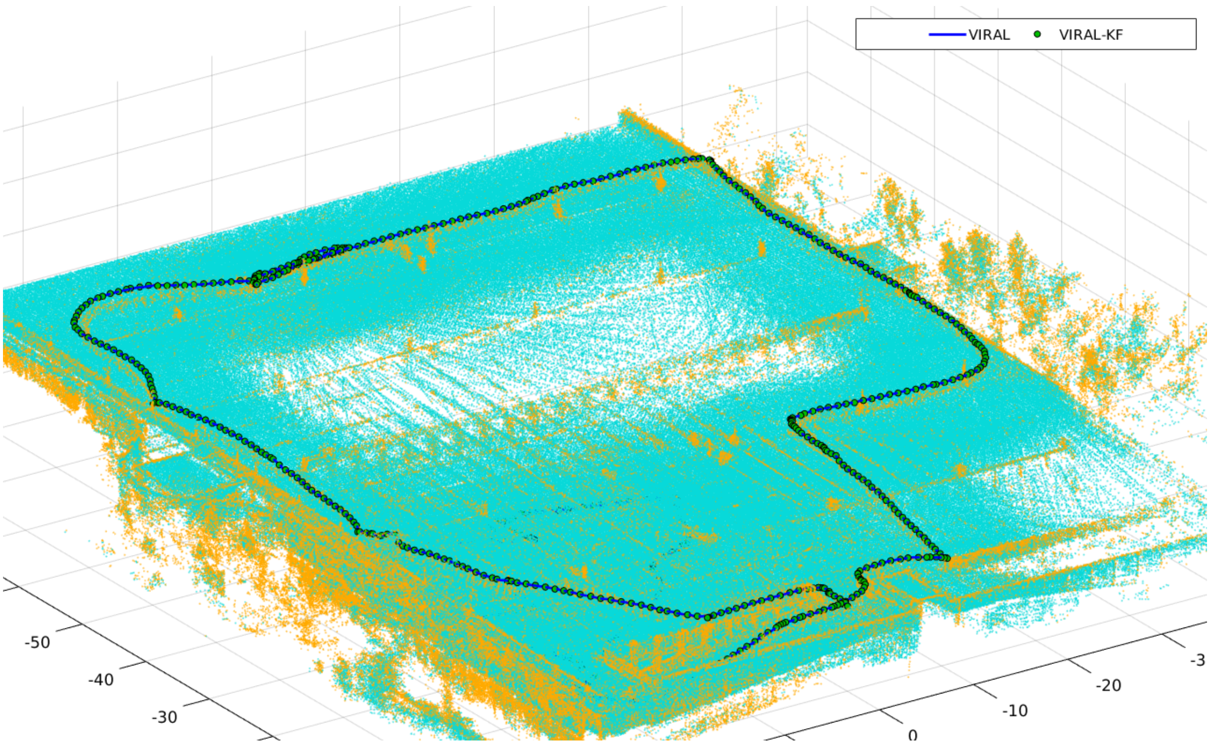

Figure 13. PARKING_1

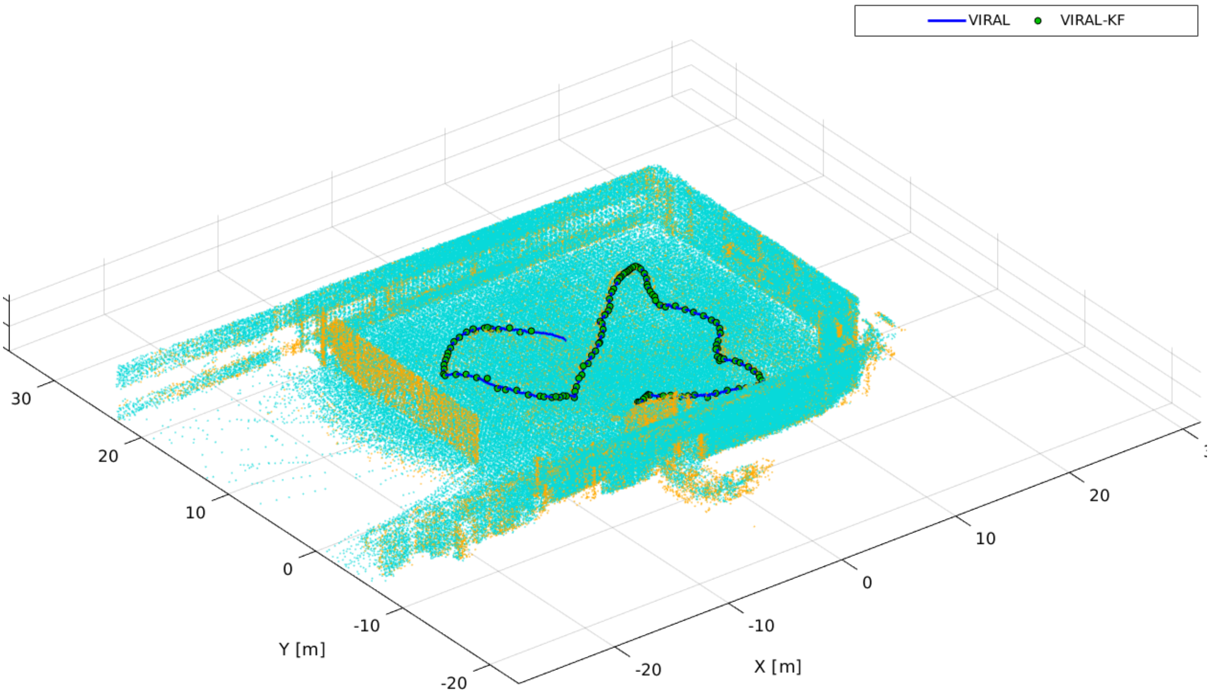

Figure 14. uzh_tracking_area_run2

III. CONCLUSION

From the experiments, we find that VIRAL SLAM can successfully run over all environments and conditions covers in the datasets. The results reinforce the idea that a multi-sensor SLAM system can achieve better accuracy, and most importantly, the capability to work over a variety of greatly different conditions.

We would like to thank the HILTI team for these interesting datasets. The challenge has presented many practical and interesting issues for us to tackle. As we worked on these issues, we have been able to make several improvements and features to our SLAM system.

REFERENCES

[1] Helmberger, Michael, Kristian Morin, Nitish Kumar, Danwei Wang, Yufeng Yue, Giovanni Cioffi, and Davide Scaramuzza. "The Hilti SLAM Challenge Dataset." arXiv preprint arXiv:2109.11316 (2021).

[2] Nguyen, Thien-Minh, Shenghai Yuan, Muqing Cao, Lyu Yang, Thien Hoang Nguyen, and Lihua Xie. "MILIOM: Tightly Coupled Multi-Input Lidar-Inertia Odometry and Mapping." IEEE Robotics and Automation Letters 6, no. 3 (2021): 5573-5580.

[3] Nguyen, Thien-Minh, Shenghai Yuan, Muqing Cao, Thien Hoang Nguyen, and Lihua Xie. "VIRAL SLAM: Tightly Coupled Camera-IMU-UWB-Lidar SLAM." arXiv preprint arXiv:2105.03296 (2021).