Quasi-globally Optimal and Efficient Visual Compass in Urban 3D Structured Environments (Supplementary Material)

Pyojin Kim[1], Haoang Li[2], Kyungdon Joo[3]

***Abstract*— In this supplementary material, we provide additional experimental results on York Urban [1], ICL-NUIM [2], and Tello Urban datasets.**

I. Additional Experimental Results

A. York Urban Dataset

The first row of Fig. 2 shows a representative evaluation result of the proposed method compared to other approaches with the precision, recall, consistency error, and run time on the YUD. While the data sampling-based methods (OLRE and SLRE) show fast computation time and good consistency error, they fail to guarantee precision and recall simultaneously. BnB can obtain all the inlier lines, but it takes more than three seconds per image, which is not suitable for real-time robotics applications. Although QBnB can compute the MF rotations within 0.1 seconds, the existing parameter search-based approaches aim to maximize the number of inliers, not to minimize the consistency (MF rotation) error. The proposed method can obtain the second-lowest consistency error similar to SLRE, and achieve the 100% precision and recall accuracy similar to BnB simultaneously in real-time at 100 Hz.

B. ICL-NUIM Dataset

We measure the mean value of the absolute rotation error (ARE) [4] in degrees, and report the quantitative evaluation results of various methods in Table **??**. We highlight the smallest rotation error for each dataset as bolded. We exclude the parameter search-based methods such as BnB and QBnB because they are not suitable for this kind of continuous 3-DoF rotational motion estimation. We add the result of OPRE [4], the orthogonal plane-based tracking algorithm.

Some of the methods such as OLRE and OPRE depending on multiple lines or planes only sometimes fail to track the MF rotations (marked as $\times$ in Table **??**) due to multiple lines or orthogonal planes not always being visible throughout the entire image sequences. Especially in 'Living Room 0', at one point the camera observes only a single line and plane with very low texture, leading to failure of other approaches. Theoretically, the proposed method only requires at least one plane for vertical dominant direction and one line for horizontal dominant direction to estimate the MF rotations. The proposed method can keep tracking the absolute 3-DoF camera orientation stably and accurately for all image sequences as shown in the first row of Fig. **??**.

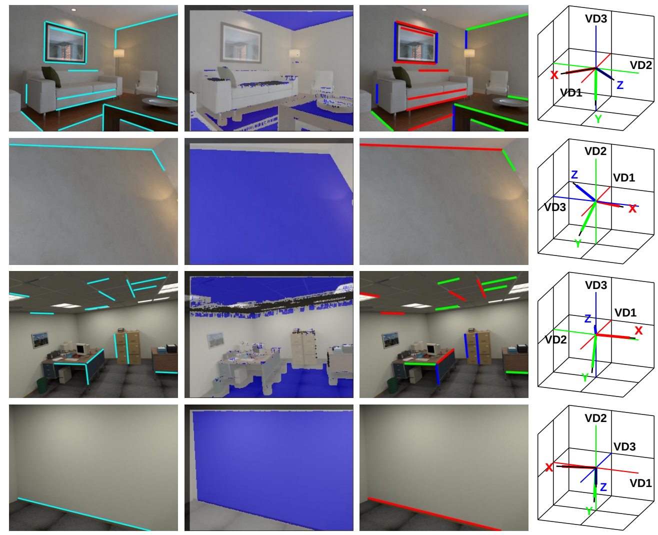


Fig. 1. Representative evaluations on ICL-NUIM [2] dataset. Each row represents a tested frame in living room and office room datasets. Each column denotes extracted image lines by LSD [3], tracked vertical dominant direction (typically ground plane), clustered lines, and the true and estimated 3-DoF camera orientation, respectively.

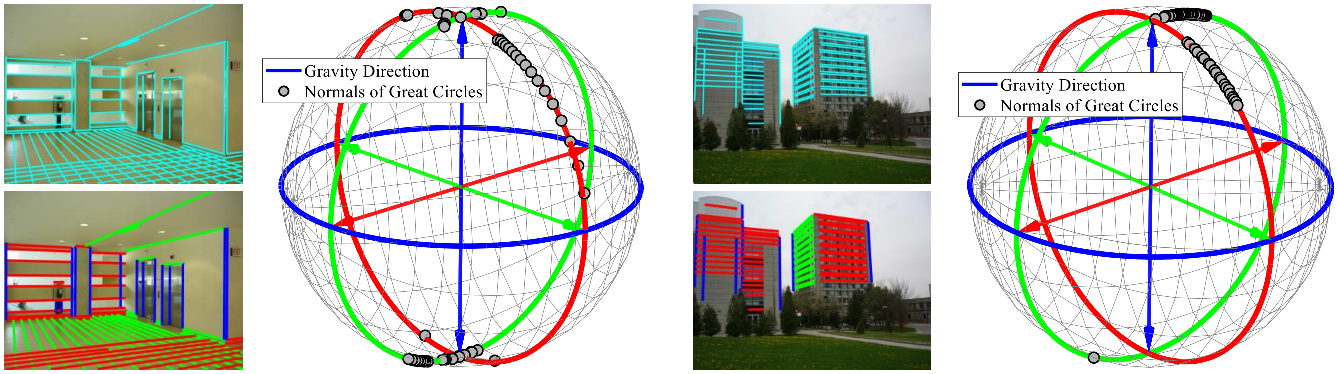
C. Tello Urban Dataset

We evaluate the proposed method on the real-world data from a DJI Tello drone flying in the outdoor urban MW environments as shown in the third row of Fig. **??**. We utilize the gravity direction from the IMU sensor measurements as a vertical dominant direction, and search for the optimal third DoF of horizontal dominant direction with image lines extracted by LSD [3]. Existing approaches based on the depth camera cannot operate in such a drone flight environment. The proposed method can stably track the absolute 3-DoF camera orientations in a challenging outdoor flight environment as shown in the second row of Fig. **??**, showing that it can operate like a drift-free visual-inertial compass for the yaw angle correction.

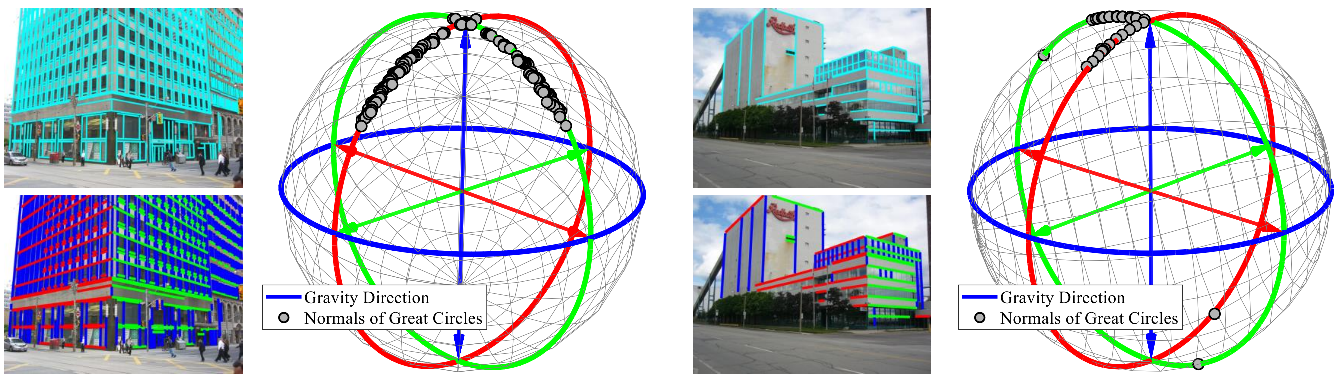
[1]Pyojin Kim is with Department of Mechanical Systems Engineering, Sookmyung Women's University, Seoul, South Korea. {pjinkim}@sookmyung.ac.kr

[2]Haoang Li is with Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, China. {haoang.li.chuk}@gmail.com

[3]Kyungdon Joo is with the Artificial Intelligence Graduate School and the Department of Computer Science, UNIST, Ulsan, South Korea. kdjoo369@gmail.com, kyungdon@unist.ac.kr



(a) Frame 4 (b) Frame 36



(c) Frame 66 (d) Frame 96

Fig. 2. Representative evaluations on York Urban [1] dataset. In each tested frame, we plot the clustered lines and corresponding normals of great circles on the Gaussian sphere. We exploit the mine-and-stab (MnS) to search for the optimal horizontal direction (red and green axis) of the MW rotation, achieving the quasi-global optimality in terms of the number of inlier lines.

References

[1] P. Denis, J. H. Elder, and F. J. Estrada, “Efficient edge-based methods for estimating manhattan frames in urban imagery,” in *ECCV*, 2008.

[2] A. Handa, T. Whelan, J. McDonald, and A. J. Davison, “A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM,” in *IEEE ICRA*, 2014.

[3] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, “Lsd: A fast line segment detector with a false detection control,” *IEEE T-PAMI*, 2008.

[4] P. Kim, B. Coltin, and H. J. Kim, “Visual odometry with drift-free rotation estimation using indoor scene regularities,” in *BMVC*, 2017.