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

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

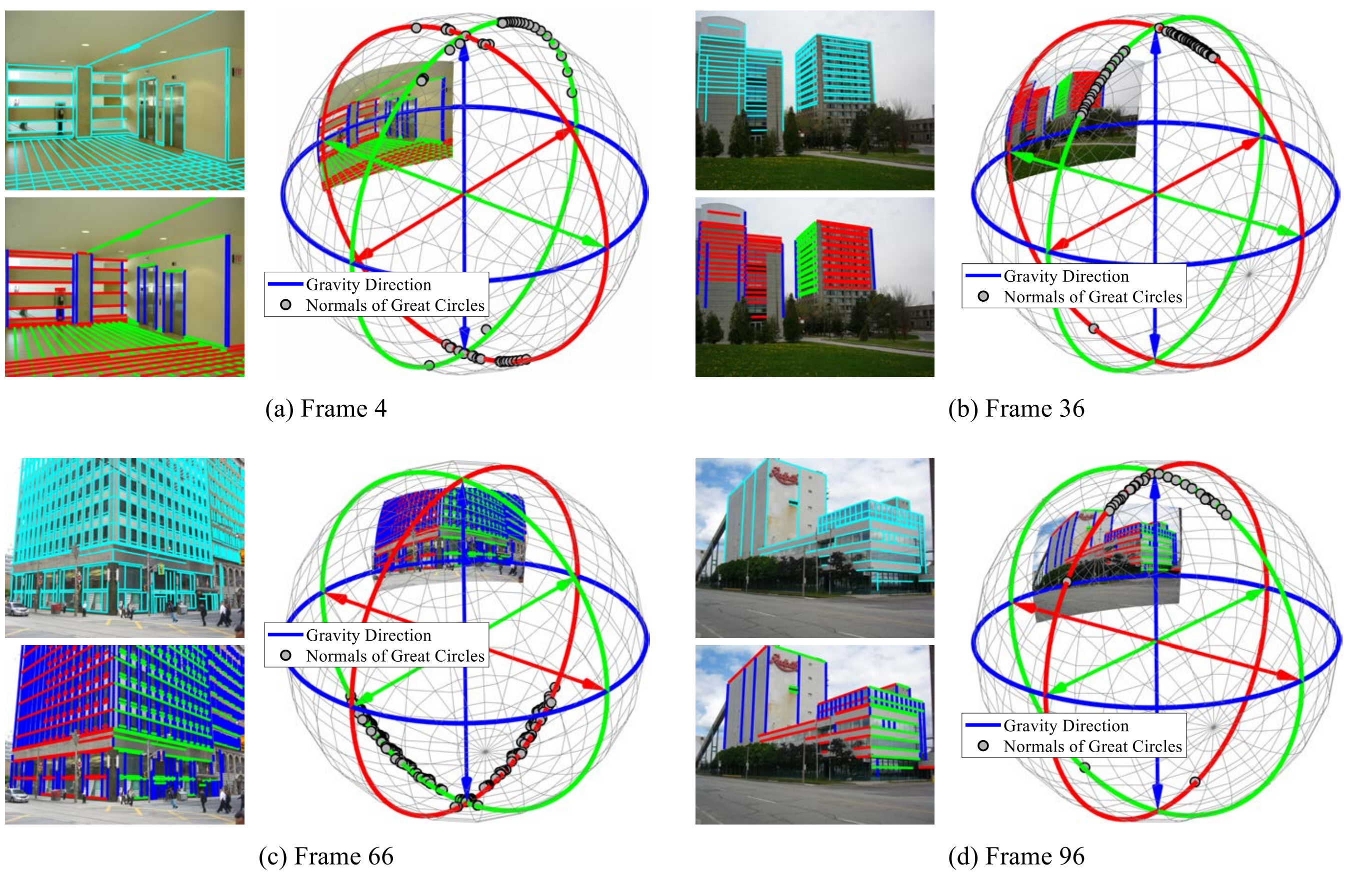


Fig. 1. Representative evaluations on York Urban [1] dataset. In each tested frame, we plot the raw (cyan) and clustered (RGB) image lines with respect to the estimated Manhattan world (MW). We visualize the estimated Manhattan frame (MF) and the corresponding normals of the great circles on the Gaussian 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.

***Abstract*— In this supplementary material, we provide additional experimental results on York Urban [1], ICL-NUIM [2], and our Tello Urban datasets. We present the detailed experimental results for each dataset, and validate the effectiveness of the proposed approach with the geometric analysis on the Gaussian sphere.**

[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

I. Additional Experimental Results

A. York Urban Dataset

The above Fig. 1 shows the drift-free 3-DoF camera orientation estimation results with the proposed mine-and-stab (MnS) approach on the York urban dataset satisfying the Manhattan world (MW). Since the York urban dataset [1] does not provide the gravity direction data (blue in Fig. 1), we obtain the virtual gravity direction vector by utilizing the lines (blue) vertical to the ground plane with the data sampling-based approach [3]. The proposed method can estimate the accurate Manhattan frame (MF) with respect to the camera frame, and it clusters all image lines correctly. Also, as we expected, the normals of the great circles from the image lines (gray dots) are all aligned correctly on the

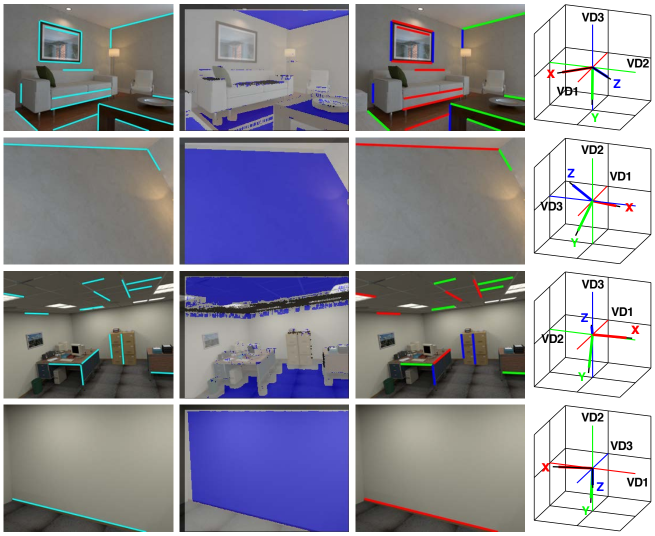


Fig. 2. Representative evaluations on ICL-NUIM [2] dataset. Each row represents a tested frame in the living room and office room datasets. Each column denotes the extracted image lines by LSD [4], tracked vertical dominant direction (typically ground plane), clustered lines in the MW, and the true (black) and estimated (RGB) 3-DoF camera orientation, respectively.

Fig. 3. Representative evaluations on our Tello Urban dataset. We plot the raw image lines (cyan) extracted by LSD [4] and the clustered lines (RGB) for the estimated MF in the image pairs. The top two rows show the image sequences during the indoor flight, and the bottom two rows are the data obtained during the flight in an outdoor urban environment.

horizontal dominant planes. The proposed method shows high efficiency and accuracy by hybridizing the data sampling and parameter search strategies, achieving the quasi-global optimality in terms of maximizing the number of inlier lines.

B. ICL-NUIM Dataset

Fig. 2 shows the detailed experimental results on the ICL-NUIM [2] RGB-D dataset with the proposed approach. The top two rows and bottom two rows in Fig. 2 denote the 3-DoF camera tracking results in the living room and office room, respectively. The first column shows the extracted raw image lines by LSD [4], and the second column visualizes the tracked dominant plane direction (blue) with the surface normals from the depth image sequences. Without loss of generality, the dominant plane does not have to be the ground plane in the indoor RGB-D camera setting. The proposed method continues to track and update the dominant plane that is dominantly observed in the current field of view. The third column is the result of clustering the lines with respect to the estimated Manhattan frame (MF) with the proposed method. In the fourth column, colored thick and thin lines denote the estimated 3-DoF camera frame and the vanishing directions (VDs), and the black lines represent the true pose of the camera. Each of the colored lines and plane in the images corresponds to the VD of the same color.

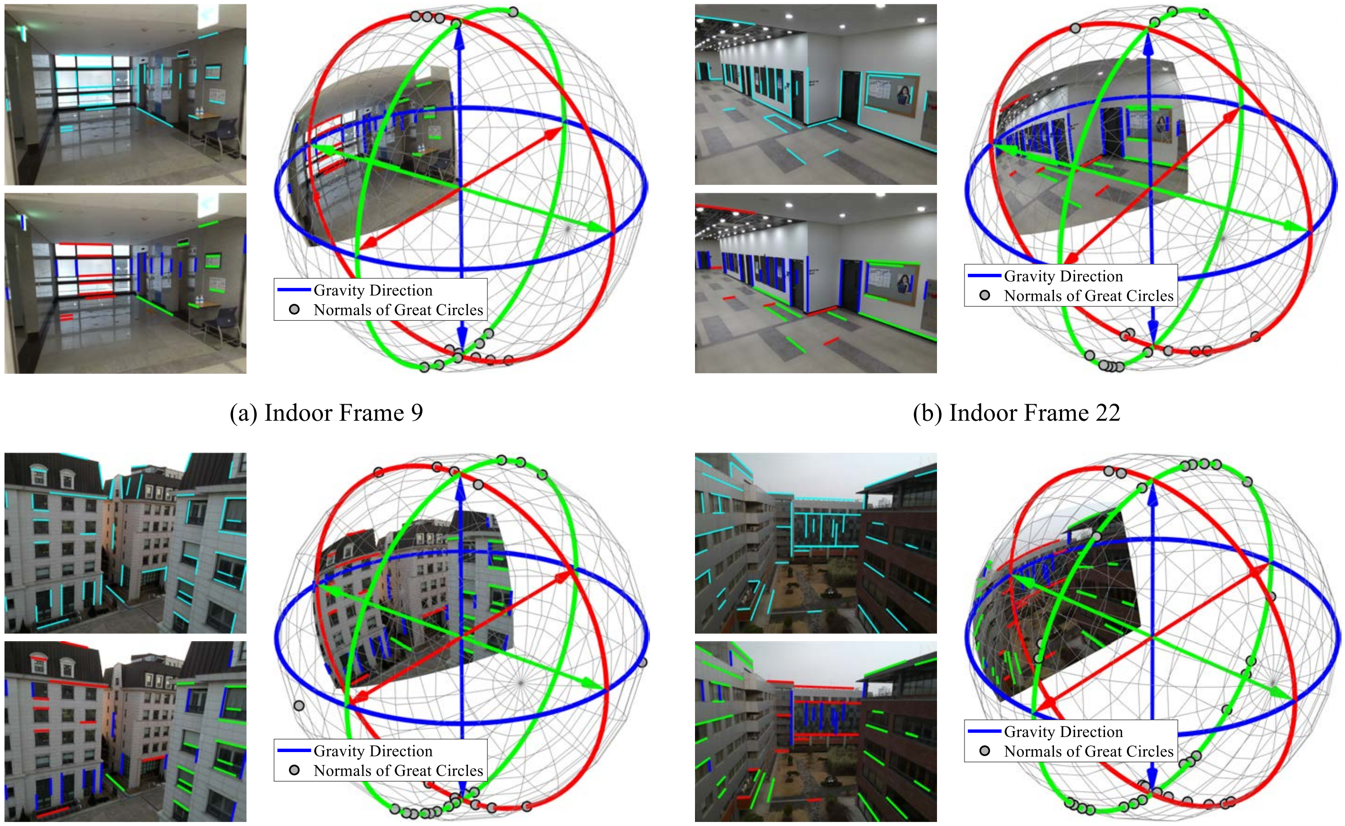
The proposed method can continue tracking the absolute camera orientation stably and accurately as shown in Fig. 2, achieving the average rotation error as 0.26 degrees. In addition, the proposed method can stably track the 3-DoF camera orientation even in a harsh environment with insufficient lines as shown in the second and fourth row of Fig. 2. Theoretically, the proposed method only requires at least one plane for vertical dominant direction and one line for horizontal dominant direction to track the MF rotations. The proposed method can stably track the absolute rotations even when the camera sees only a planar surface with little texture by exploiting the minimal sampling (a single line and plane) to recognize structural regularities.

C. Tello Urban Dataset

Fig. 3 and Fig. 4 show the indoor and outdoor experimental results of applying the proposed approach to the RGB image sequences and gravity direction vectors from an IMU obtained from a DJI Tello drone during the flight. We obtain the time-synchronized RGB image and the gravitational direction vector from an IMU through the ROS DJI Tello driver[1], and additionally refine the gravity direction with the vertical image lines (blue) using the data sampling-based approach [3]. In Fig. 3, each image pair consists of the raw (cyan) and clustered (RGB) image lines, and the top two rows and bottom two rows are the results of 3-DoF camera orientation estimation based on the data obtained during the indoor and outdoor flights, respectively. Existing approaches relying on the depth camera cannot operate in such a drone flight environment due to the effective range of the depth camera and a limited field of view (FoV). The proposed method accurately and precisely clusters all image lines with respect to the estimated Manhattan frame (MF), showing that it can operate like a drift-free visual-inertial compass in a challenging outdoor flight environment.

We additionally visualize the geometric relationships between the image lines and the estimated Manhattan frame (MF) on the Gaussian sphere as shown in Fig. 4. The proposed method clusters the image lines correctly in both indoor and outdoor frames, and the normal vectors of the

[1] http://wiki.ros.org/tello_driver



(a) Indoor Frame 9

(b) Indoor Frame 22

(c) Outdoor Frame 32

(d) Outdoor Frame 26

Fig. 4. Representative evaluations on our Tello Urban dataset with the Gaussian sphere domain. In each tested frame, we plot the raw (cyan) and clustered (RGB) image lines with respect to the estimated Manhattan frame (MF). We visualize the estimated Manhattan frame (MF) and the normal vectors of the great circles from the image lines (gray dots) on the Gaussian sphere together. The normal vectors (gray dots) are aligned well on the estimated Manhattan frame (MF) correctly. 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.

great circles from the image lines (gray dots) are well aligned with respect to the estimated Manhattan frame (MF). We exploit the proposed mine-and-stab (MnS) with Manhattan constraints to search for the optimal third DoF of the camera orientation, achieving the quasi-global optimality in terms of maximizing the number of inlier lines (gray dots).

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] J.-C. Bazin and M. Pollefeys, "3-line RANSAC for orthogonal vanishing point detection," in *IEEE IROS*, 2012.

[4] 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.