Supplementary Material : Robust Large Scale Monocular Visual SLAM

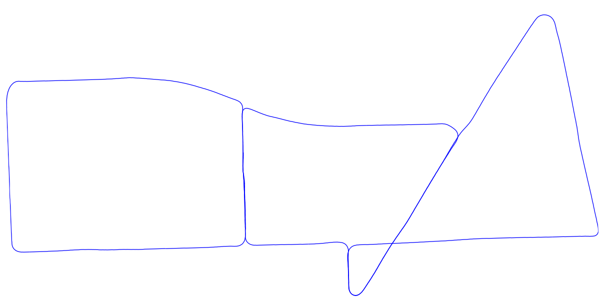
Guillaume Bourmaud Rémi Mégret
Univ. Bordeaux, CNRS, IMS, UMR 5218, F-33400 Talence, France
{guillaume.bourmaud,remi.megret}@ims-bordeaux.fr

Abstract

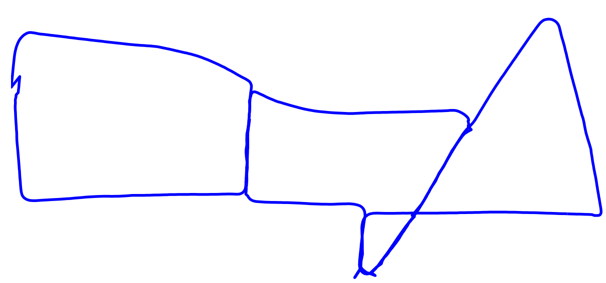
In this supplementary material, we provide additional results on several challenging video sequences.

1. KITTI benchmark : sequences 13 and 15

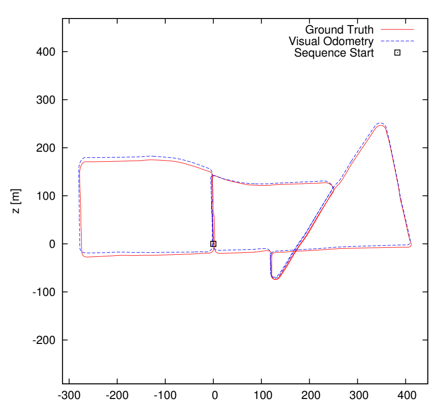
Here we present the result of our approach on two sequences of the KITTI dataset (13 and 15). For those two sequences, the ground truth is not provided. However, a plot of the ground truth is available on the KITTI benchmark website. In Fig.1, our approach outperforms the state of the art monocular VSLAM algorithm [1]. The result of [1] is not available for the sequence 15 (Fig.2) thus we also provide the result of the state of the art algorithm on the KITTI dataset that employs a Velodyne laser scanner.



(a) Our approach

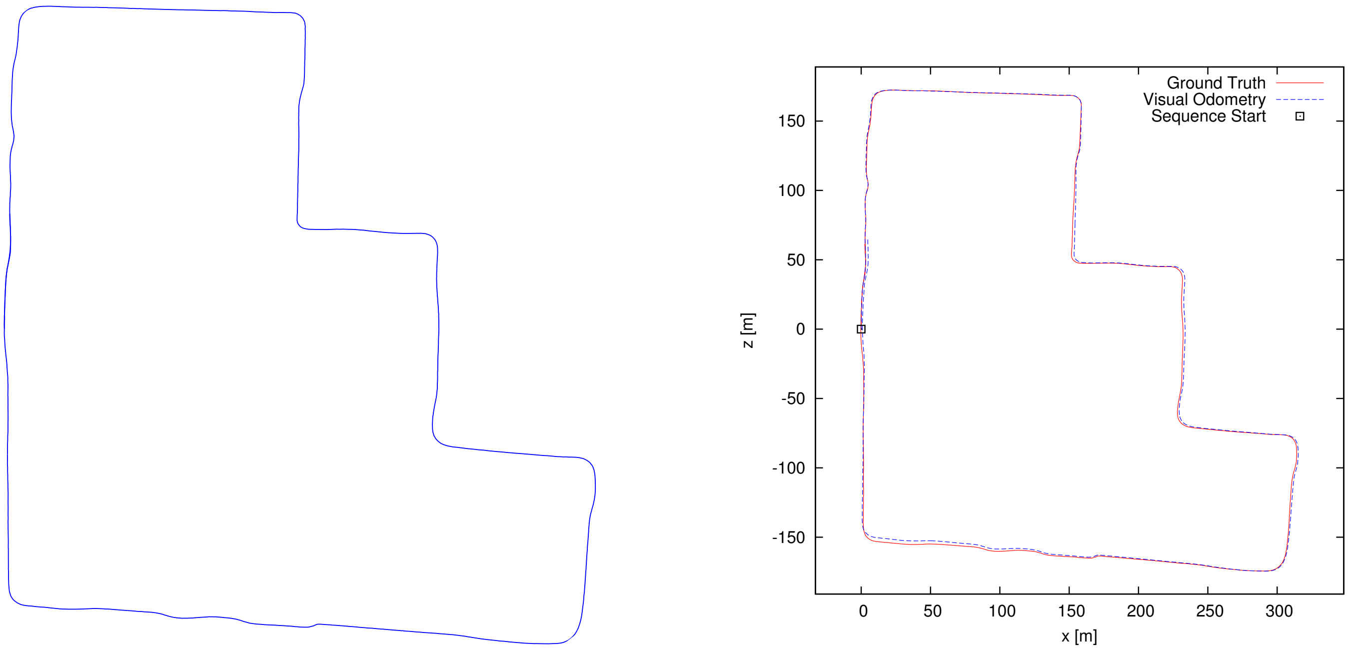


(b) Result of Lim et al. [1]



(c) State of the Art [2] using a Velodyne laser scanner

Figure 1: Sequence 13



(a) Our approach

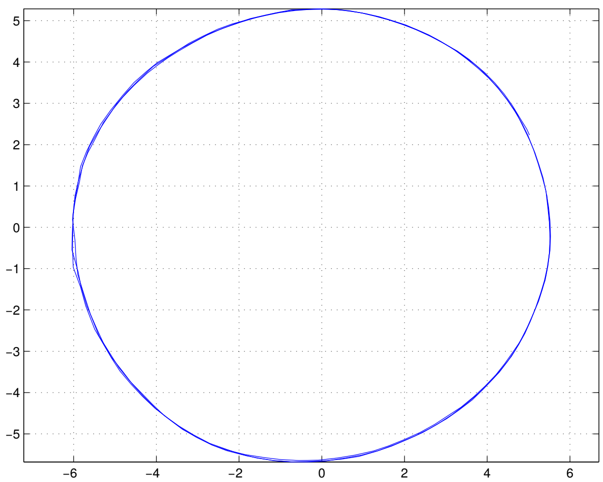
(b) State of the Art [2] using a Velodyne laser scanner

Figure 2: Sequence 15

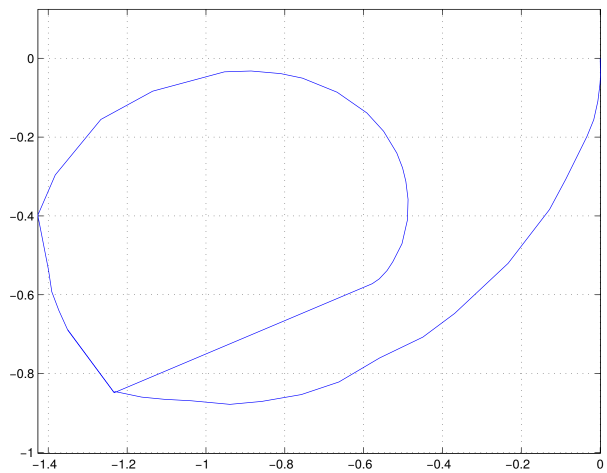
2. A perfect circle

In Fig.3, the estimated trajectory of a camera mounted on a turntable (the camera looks outside) is presented. In this sequence the camera translation is small (the radius circle is 10 cm). The true camera trajectory is a perfect circle. One can see that the camera trajectory estimated with our approach is significantly closer to a perfect circle than the one estimated with [3].

(a) Example of video frames



(b) Our approach



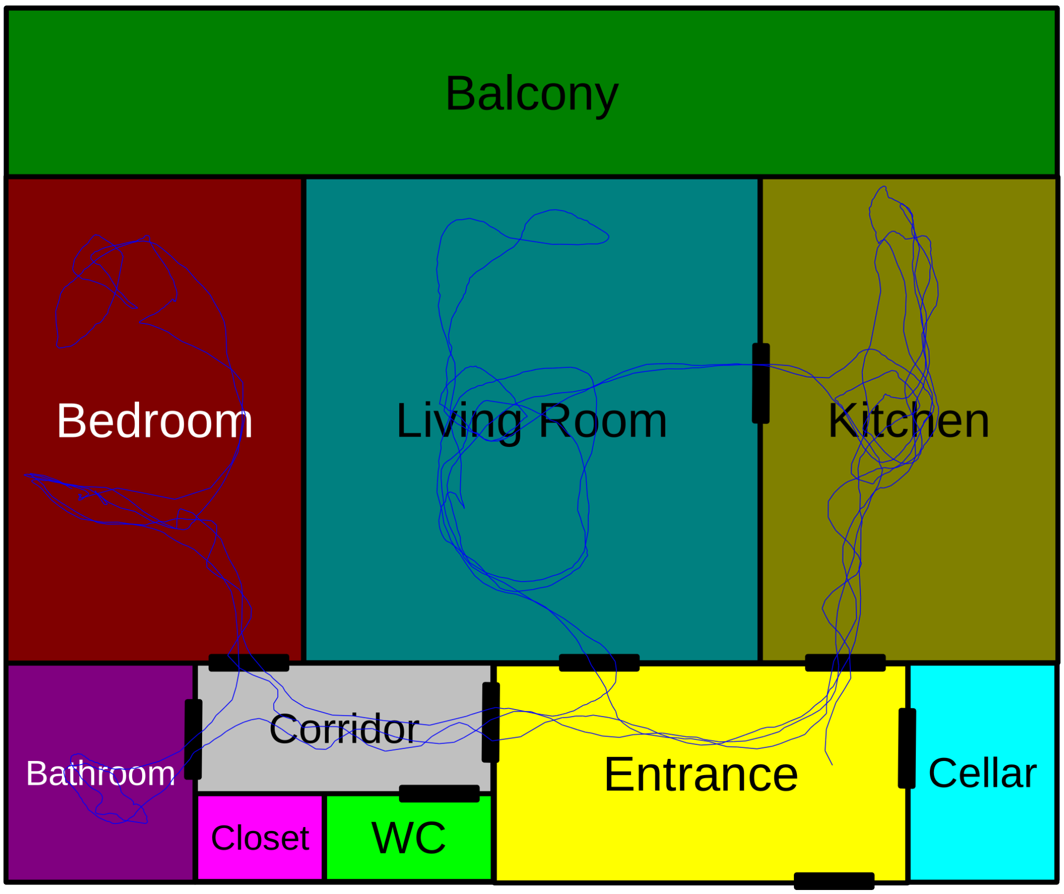
(c) Result of Engel et al. [3]

Figure 3: Result on the “perfect circle” sequence

3. Reconstruction of a whole apartment

Here, we evaluate the performances of our system qualitatively on a video sequence of 10000 frames taken in an apartment. After having applied our automatic framework to the video sequence, we have obtained a set of aligned sub-maps where each sub-map contains a 3D point cloud and a part of the camera trajectory. In order to qualitatively evaluate the result, we manually place on top of the ground plan of the flat the estimated camera trajectory (see Fig. 4). As can be seen, the superimposed camera trajectory is coherent with the ground plan, i.e the camera trajectory goes into (almost) every rooms without crossing the walls and pass through the doors when going from one room to another.

(a) Example of video frames



(b) Superimposed estimated camera trajectory (blue line) with the ground plan of the flat

Figure 4: Reconstruction of a whole apartment

References

[1] H. Lim, J. Lim, and H. J. Kim, "Real-time 6-dof monocular visual SLAM in a large-scale environment," in *ICRA*, 2014. 1, 1b

[2] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: Low-drift, robust, and fast," in *Submitted to IEEE International Conference on Robotics and Automation(ICRA)*, Seattle, WA, May 2015. 1c, 2b

[3] J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large-scale direct monocular SLAM," *ECCV, Lecture Notes in Computer Science*, pp. 834–849, 2014. 2, 3c