Additional Results for Multibody Visual SLAM

Abhijit Kundu, C. V. Jawahar and K Madhava Krishna
abhijit.kundu@gatech.edu, {jawahar, mkrishna}@iiit.ac.in

This document along with the video contains some additional results and better resolution figures, supporting the main paper. Videos of results can be downloaded from the project web-page at http://robotics.iiit.ac.in/projects/mulvslam/. Details of the image sequences used for experiments are listed in Table. 1.

Dataset	Image Resolution	Trajectory Length	Avg. Runtime
Moving Box [4]	320x240	718 images	20Hz
Versailles Rond [2]	760x578	700 images (400m)	7Hz
New College [3]	512x384	1500 images	13Hz
CamVid [1]	480x360 (Resized)	1600 images (0.7km)	11Hz

Table 1. Details of the datasets

Note: The legends (Fig. 1) used here and in the accompanying video are sligtly different from figures shown in the main paper. Motion segmentation results are shown by shade in correponding color of the convex-hull formed of the feature points segmented as independently moving. Reconstructed 3D static world points are colored depending on height from the estimated ground-plane. Trajectories and structure of other moving objects are shown in some color(red/blue). Particles of the BOT filter are shown in green. The images are best viewed on screen.

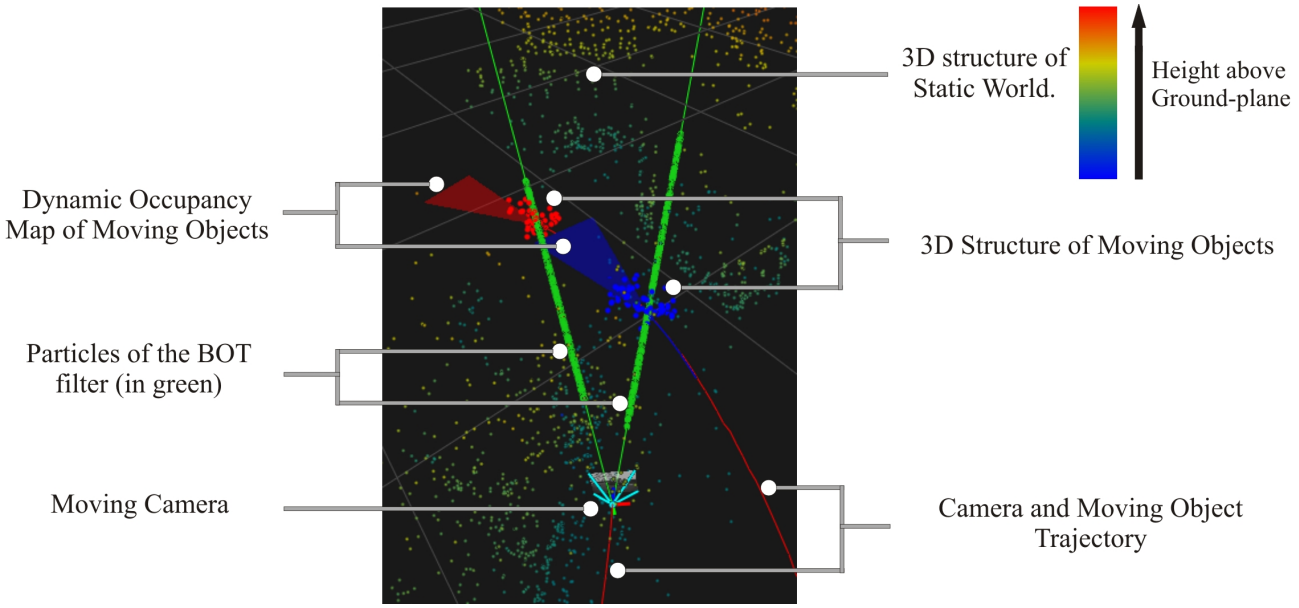

Figure 1. Legends used in the figures

1. New College Sequence

We tested our results on some dynamic parts of the New College dataset [3]. Only left of the stereo image pairs has been used. In this sequence, the camera moves along a roughly circular campus path, and three moving persons passes by the scene. The results on this sequence are highlighted in Fig. 2. It shows the map and camera trajectory with respect to the static world and the final depth estimate from BOT. It is to be noted that, this along with most of the sequences shown in this experiments are generally unobservable. It is only after integration of different cues, we obtain a descent estimate of the moving object location.

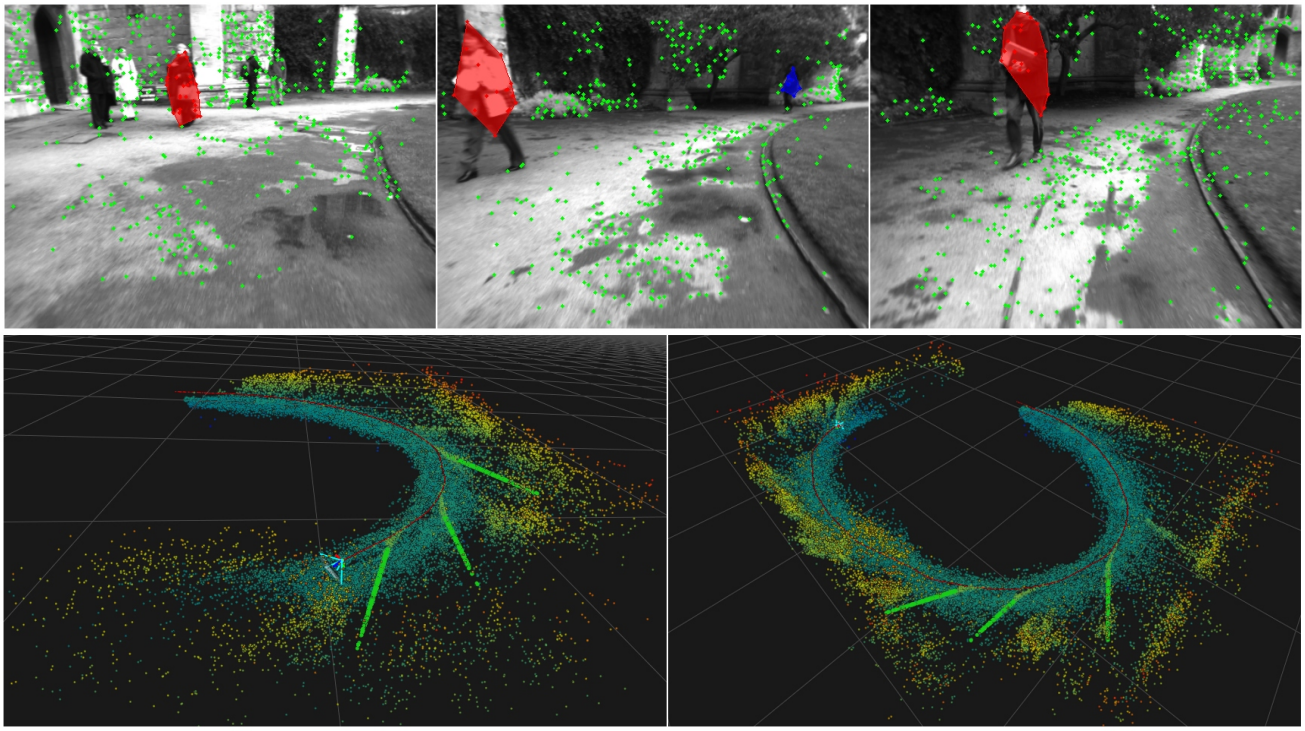

Figure 2. Results for the New College sequence

2. Versailles Rond Sequence

This is an urban outdoor sequence [2] taken from a fast moving car, with multiple numbers of moving objects appearing and leaving the scene. Only left of the stereo image pairs has been used. Fig. 4 shows the results of the integrated map produced by the algorithm. In Fig. 4 and the next figures, online occupancy maps are shown as shaded region, which is estimate the most likely space to be occupied by the moving object in next 16 frames (around 1s). We have shown the estimate for the current best relative scale estimate of the moving object. Considering the uncertainity of the relative scale actually gives a much larger possible occupancy area for the moving object.

Figure 3. Motion Segmentation results on the Versailles Rond Sequence.

3. CamVid Sequence

We have tested our algorithm on subsequence (seq05VD and seq06R0) of the CamVid dataset [1]. Fig. 6 shows the results of our multibody Visual SLAM and the integrated 3D dynamic map it produces. Fig. 5 are the motion segmentation results on the same subsequence. Note high degree of correlation between camera and the car trajectory, which makes it challenging for both motion segmentation and relative scale estimate.

References

[1] G. Brostow, J. Fauqueur, and R. Cipolla. Semantic object classes in video: A high-definition ground truth database. *PRL*, 30(2):88–97, 2009. 1, 2

[2] A. Comport, E. Malis, and P. Rives. Real-time Quadrifocal Visual Odometry. *IJRR*, 29(2-3):245, 2010. 1, 2

[3] M. Smith, I. Baldwin, W. Churchill, R. Paul, and P. Newman. The new college vision and laser data set. *IJRR*, 28(5):595, 2009. 1

[4] S. Wangsirpitak and D. Murray. Avoiding moving outliers in visual slam by tracking moving objects. In *ICRA*, 2009. 1

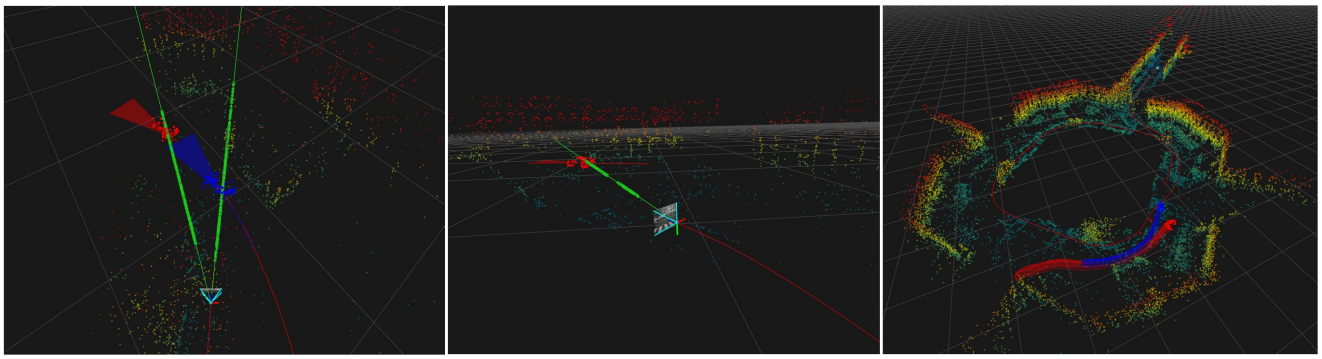

Figure 4. Results on the Versailles Rond Sequence. The left and midlle image shows an instance of the online occupancy map. Shaded region shows the most likely space to be occupied in next 16 frames (around 1s). Right image demonstrates the reconstruction and trajectories of two moving cars.

Figure 5. Motion Segmentation results for the CamVid sequence

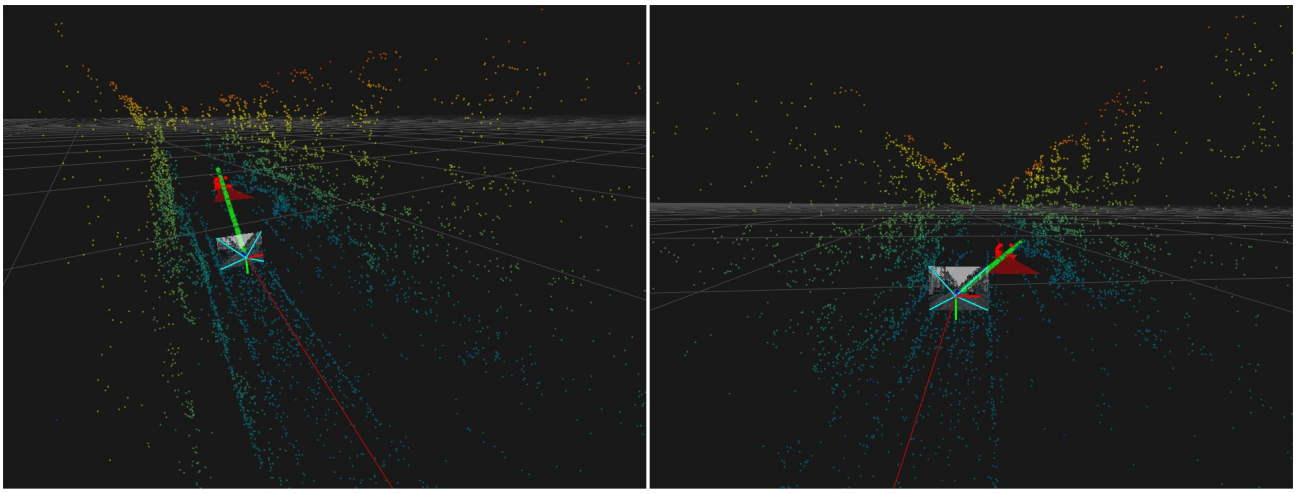

Figure 6. Multibody Visual SLAM results for the CamVid sequence