Trajectory data visualization of sports video based on SLAM

Yifei ZHANG † Hiroshi WATANABE †
† Waseda University

Abstract This paper presents experiments on trajectory data mining from subjective shot sports videos based on SLAM (Simultaneous Localization and Mapping) and wearable camera calibration. The experiment result shows adequate robustness to most common motion clutter in the outdoor long distance sports videos.

1. Introductions

Sports data visualization is of great importance for journalism, professional analysis and visual analytics system. Some sports games, like cricket, tennis, can set a number of high-speed cameras to cover the area of playground, and use the principles of triangulation to calculate the visual images for generating the trajectory of the ball or display a statistically record [1]. However, it is not possible to build similar system for the road-bike or snowboard scenes in which the game path usually is a long distance and single way with complicate circumstances. This paper presents a system based on state-of-the-art ORB-SLAM [2], combined with camera calibration using OpenCV [3] and adjust the ORB (Oriented FAST and Rotated BRIEF) extractor amount, to generate a rough and traceable map using subjective shot videos taken by GoPro camera. Use of the same feature for tracking, mapping, re-localization and loop closing, make the system more efficient even without GPUs. It provide good performance to generate the visual odometry (VO) using some kind of sports video in specifically circumstance. The experiment result shows the possibility to adapt this system to real-time outdoor sports.

2. Materials and Methods

2.1 SLAM background

SLAM is the solution to provide accurate estimates of the moving agent localization and the structure of the surrounding world, gather information by moving exteroceptive senors [4], [5]. The primary stage is based on filter [6], the ground breaking feature-based SLAM is Parallel Tracking and Mapping (PTAM) [7], and ORB-SLAM. Compared with PTAM, ORB-SLAM adapt limited scale operation to avoid unnecessary redundancy, and combine place recognition, scale-aware loop closing [8], and co-visibility information for large scale operation. ORB feature is rotation invariant and resistant to noise, more efficient alternative to SIFT or SURF [9]. The ORB-SLAM system overview can be showed like this:

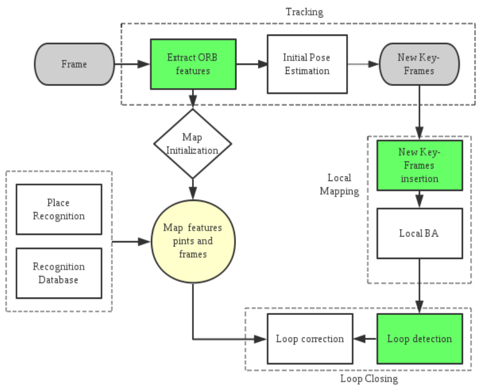

Figure 1. ORB-SLAM system overview

2.2 Camera calibration

The accurate values of the camera intrinsics and radial distortion parameter can be determined by OpenCV pinhole camera model. The camera calibration process determined the exact pixel size, focal length, lens axis offset, and the lens distortion characteristics [3].

2.3 Map Initialization

In the outdoor sports videos with numerous rotation and change of direction, it always turns out fail to extract and match the features between two key-frames. If the tracking is successful for the last frame, a constant velocity motion model will be used to predict the camera pose and search the map point in the last frame; If the match point is not enough, a wider search of the map points around their formal position will be performed, then correct the pose with the newly found point. The generated local map is optimized by the Bundle Adjustment [10], [11] with key-frames.

3. Experimental Apparatus and Setup

3.1 Camera calibration

By using OpenCV camera calibration tools, the camera intrinsic and extrinsic parameters are estimated. Camera.fx, fy (611.18384, 611.06728) are the focal lengths expressed in pixel units; Camera.cx , cy (515.31108,402.07541) are the coordinates of the principal point at the image center; Cmaera.k1, k2, k3 (-0.1054120,-0.1054120,-0.00082381) are radial distortion coeffcients; Camera.p1, p2 (-0.0017121, 0.00142017) are tangential distortion coefficients. With these precise parameters, the images optical distortion are corrected, and the distance of the object from the camera is estimated, the space maps are accurately generated.

3.2 Adjust the ORB extractor amount

During this experiment, the cameraman is acting as the moving agent, and the monocular camera provides frames with scene features. The setting of the ORB Extractor have been updated with a much higher amount to detect the feature of every images, in this experiment, the features was set to 15,000. It results in a more stable tracking with less loss.

4. Results and Discussion

Figure 2 shows the generated map in key-frames, and the video sequence image with feature points. The video of generate 3900 images.

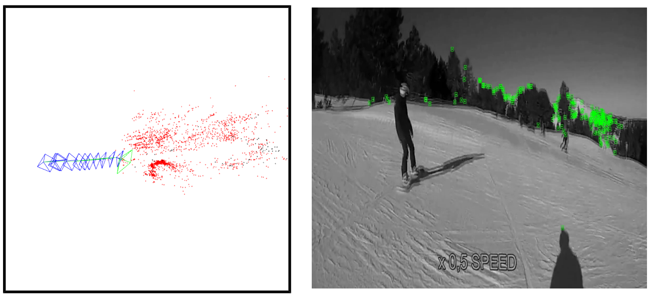

Figure 2. Left:Snowboard video trajectory data with key-frames.
Right: Key-frame feature points in one fames

Two different length sequences are used, referred to different speed of the snowboard speed, both of them worked well and generate data visualization map. Compared with the trajectory generation based on monocular EKF-SLAM using the same video in Figure 2, despite combined with loops detection, it only can show part of trajectory path of the result of ORB-SLAM, and takes more time with more error rate.

Table 3. Result for the experiment of two SLAM system

	Initial Times	Relocalization recall	Lost track images
EKF-SLAM	20s	-	40%
ORB-SLAM	17s	78.4%	20%

5. Conclusion

Trajectory data visualization using ORB-SLAM showed adequate robustness to sports video in specifically circumstance, such as snowboard.

REFERENCES

[1] *http://www.hawkeyeinnovations.co.uk/ Hawkeye*

[2] Raúl Mur-Artal, J. M. M. Montiel and Juan D. Tardós: "ORB-SLAM: A Versatile and Accurate Monocular SLAM System" IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, October 2015

[3] http://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html OpenCV

[4] C. Kerl, J. Sturm and D. Cremers: "Robust odometry estimation for RGB-D cameras", Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA) , 2013

[5] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," IEEE Transations on Pattern Analysis and Machine Intelligence, vol. 29,no. 6, pp. 1052-1067, 2007

[6] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot. Consistency of the ekf-slam algorithm. In Proc. Of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 3562-3568, October 2006

[7] G. Klein and D. Murray: "Parallel tracking and mapping for small AR workspaces", in IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR), pp. 225-234, Nara, Japan, November 2007

[8] Raúl Mur-Artal, J. M. M. Montiel and Juan D. Tardós: "ORB-SLAM: A Versatile and Accurate Monocular SLAM System" IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, October 2015

[9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in IEEE International Conference on Computer Vision (ICCV), pp. 2564-2571, Barcelona, Spain, Novermber 2011

[10] B. Triggs, P. F. McLauchaln, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment a modern synthesis," in vision algorithms: theory and practice, 2000, pp.298-372

[11] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge University Press, 2004

†Graduate School of Fundamental Science and Engineering, Waseda University

〒 169-0072, Shillman Hall 401, 3-14-9 Okubo, Shinjuku, Tokyo

TEL.080-4370-7137 E-mail: zyf_fse@fuji.waseda.jp