2016 International Conference on Artificial Intelligence and Computer Science (AICS 2016)
ISBN: 978-1-60595-411-0

An Improved Monocular ORB-SLAM Method

Xiu-zhen WU, Liu GANG, Wei-si GONG and Yan SHI

Department of Control Engineering, Naval Aeronautical Engineering Institute, Yanta, China.

Keywords: Vision localization, ORB-SLAM, Adaptive threshold, Multiple solution selection.

Abstract. An improved algorithm is proposed for the practical application of monocular vision ORB-SLAM algorithm. When the image FAST feature points are detected, the adaptive gray level threshold is designed to ensure the validity of the feature points detection when the image contrast is decreased due to the illumination changes in the environment, and the robustness of the SLAM algorithm is improved. Initialization of the environment for the plane of the situation, the design of the third frame image verification method to solve the problem of multiple solution affine model. Design hardware platform to validate the proposed method for improving monocular vision ORB-SLAM. The result show that the proposed method is practical and feasible, and it improves the robustness and self-adaptability of the ORB-SLAM algorithm, and has high practical value.

Introduction

With the increase of social demand and the progress of science and technology, more and more mobile robots into everyday life, improve the efficiency of life, save the cost of living, and create more and more social value. In mobile robot technology, the most important problem is the localization of the robot, which is the basis for the robot to complete a specific task.

Because of the ability of Visual Simultaneous Localization and Mapping(V-SLAM) has its own position and attitude estimation of robot in unknown environment and to build a three-dimensional map of the environment, more and more scholars' attention, developing more and more quickly, in the field of robot localization is applied very widely. V-SLAM technology from the initial filtering based approach to gradual development of a key frame based approach, Strasdat [1] have proved that the method based on the key frame is more accurate than the method based on filtering.

Based on the key frame of the V-SLAM process can be divided into three stages: the image pretreatment, the front-end graph construction, the back-end graph optimization, as shown in Fig.1.

(1) The image preprocessing process transforms the dense gray value information of the image to spares feature points, which is convenient for image matching, which provides the basis for the construction of the subsequent front-end graph.

(2) The front-end map construction process transforms the feature points matching information into the position and the attitude of the camera, the map coordinates information, getting the initial solution of the model, the model represents the constraint relations between them.

(3) According to the constraint relation, the optimization process of the back-end is established by the nonlinear least square model. The optimal solution of the model is obtained by the iterative algorithm, and feedback to the front-end map construction process.

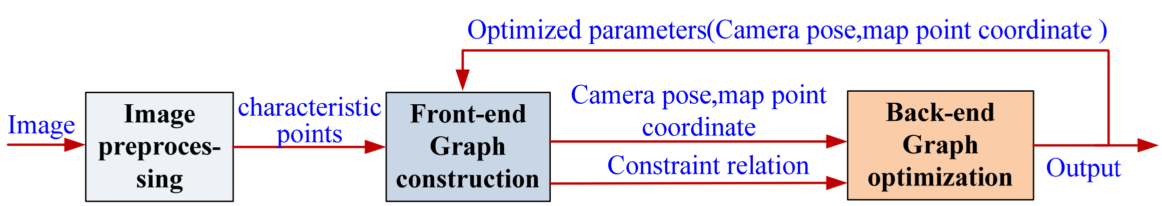


Figure 1. Visual SLAM architecture based on key frame.

The most representative of the method based on key frame is Klein and Murray proposed PTAM algorithm [2]. For the first time, the tracking location and composition are divided into two threads to parallel processing, and the feature points are FAST feature points. Pirker [3] proposed

CD-SLAM algorithm, is a closed loop correction, repositioning and other relatively prefect SLAM system, but did not involve the map initialization method. Song proposed visual odometer algorithm will be used to track the ORB feature points and the optimization of Bundle Adjustment, but there is no global relocation and closed loop correction. Engel [7] proposed LSD-SLAM algorithms to build half density map, relatively sparse feature map has more practical significance, the direct use of pixel density optimization instead of the Bundle Adjustment based on this feature points, but the algorithm still needs the sparse feature points closed loop detection.

R. Mur-Artal [8] and J.D. Tardos [9] proposed a monocular visual ORB-SLAM algorithm, which combines the main idea of the PTAM algorithm and the Strasdat proposed closed loop correction method. Compared with the precious SLAM methods, it has high accuracy and real-time performance, and is the most outstanding monocular visual ORB-SLAM algorithm. But in the practical application, we still have some problems. In order to solve these problems, this paper puts forward the improved scheme, and design an improved monocular visual ORB-SLAM. The organizational structure of this paper is as follows. The first part of the overall introduction of R. Mur-Artal and J.D. Tardos proposed monocular visual ORB-SLAM method. The second part puts forward the problem of monocular visual ORB-SLAM. The third part is the improvement of the design of the problem. The fourth part is the effect of the improved method. The fifth part is the conclusion.

Monocular Visual ORB-SLAM Algorithm

Monocular visual ORB-SLAM process as shown in Fig.2

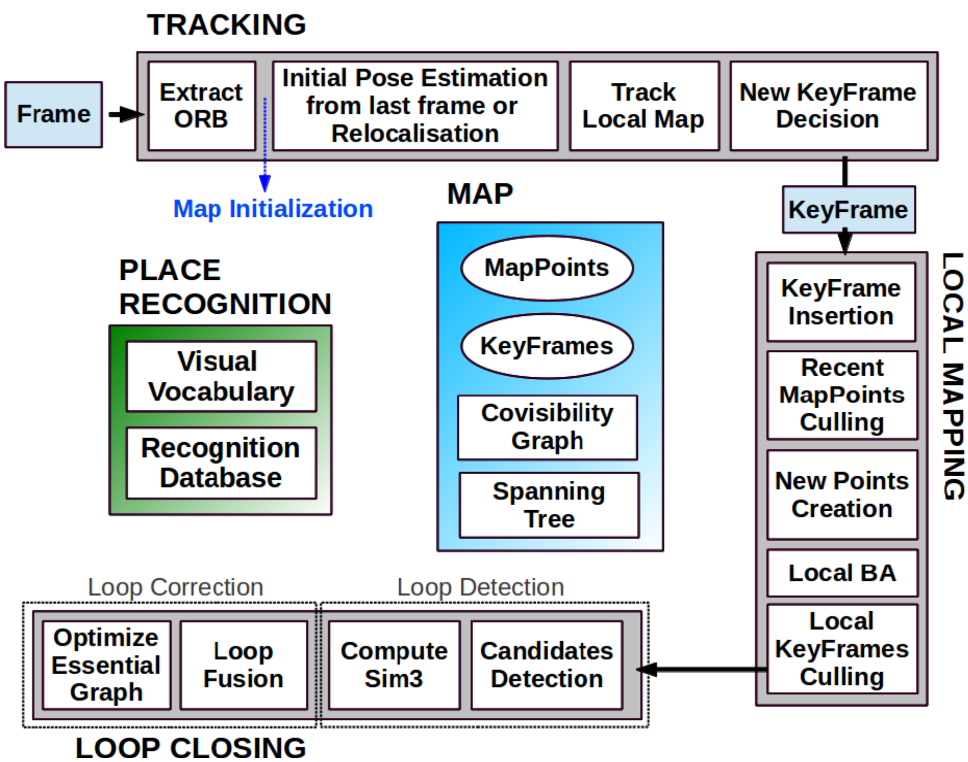


Figure 2. Monocular visual ORB-SLAM overall algorithm process.

As shown in Fig.2. The algorithm is divided into three threads: tracking, patterning and closed loop detection. In the whole process, the ORB-SLAM feature points are used, and the Bundle adjustment (BA) is used to optimize the nonlinear optimization of the three links of the tracking, composition and closed loop detection to obtain accurate position and attitude information of the camera and 3D map coordinates.

Existing Problems

In practical applications, the monocular visual ORB-SLAM algorithm has some problems.

(1) FAST feature points detection using the gray threshold t is fixed, so that the SLAM process due to the light intensity changes lead to changes in contrast, the number of FAST feature points are decreased, the tracking process failed, causing a disruption to the SLAM process.

(2) When the map is initialized, it is the case that the environment is plane, when the corresponding feature points between the two image x and x' are satisfied with the two-dimensional affine transformation:

$$\boldsymbol{x}' = \boldsymbol{H}\boldsymbol{x}, \tag{1}$$

In Equation (1), H represents the affine transformation matrix. Using feature points x and x' to solve the H and then calculate the attitude and position parameters of the camera, there are eight groups of possible solutions, only two of which have physical meaning. In the absence of a priori assumptions, we can not determine which is the optimal solution, only through the disparity size to determine the optimal result is wrong.

Improved Method

To solve the problem of FAST feature points detection, the solution is as follows.

Design self-adaption gray level difference threshold parameter t. Self-adaptation is based on the contrast of the image C:

$$\Delta t = \alpha \cdot C, \tag{2}$$

In Equation (2), α for self-adaption adjustment parameters. The value of α is set according to the test data.

For the initialization problem, the solution is as follows.

Introduction the third image frame, we can carry out the verification in two affine models which have been obtained, which is the optimal solution.

Experimental Verification and Analysis

Experiment Platform Building

Setting up the test platform for sports car experiment improved monocular visual ORB-SLAM algorithm, experimental platform is shown in Fig.3.

Figure 3. Experimental platform.

Main hardware equipment used in the experiment model and parameters as shown below.

(1) Camera

IDS industrial camera:UI-1240LE-C-6. Resolution:1280×1024.

Fixed focus lens: UH0620-10M. Field of view:54.1Hº×41.9Vº.

(2) Differential GPS

Novatel company, FlexPak6, Positioning accuracy: 5cm

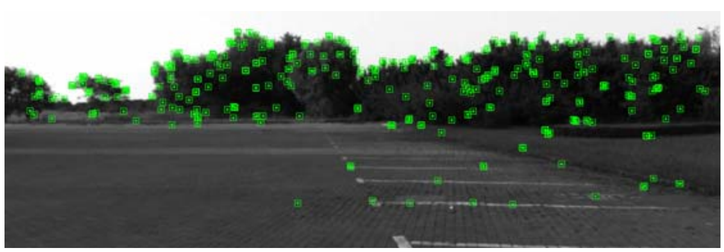
Experiment Procedure

Test site: Yantai Sports Park;
Test time: November 9, 2016
Test procedure: Closed-loop experiment were test within 120m×120m of open space (no large area block, to ensure GPS signal is valid), running a total of 3 laps, the elapsed time is 333 seconds. During the test, the camera and differential GPS data were collected.

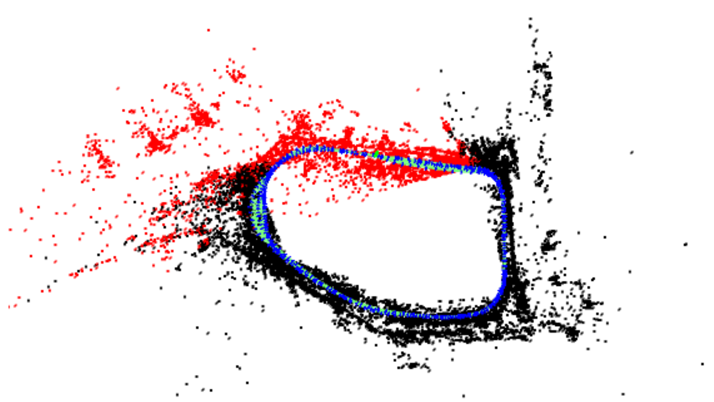
Experiment Results and Analysis

It is assumed that the geographic system is fixed in this range. The geographic coordinate system the beginning of the test is used as the reference coordinate system. The positioning data of the differential GPS is the standard reference value. The latitude and longitude information is converted into the reference coordinate system. In this coordinate system to locate the error analysis and comparison.

The experimental results of the improved monocular visual ORB-SLAM are shown in Fig.4. 4(a) is a feature points detection process, green square represents the FAST feature points on current frame matching. Fig.4(b) represent a single ORB-SLAM map building process and positioning process, black and red dots represent the stored map points, blue dots represent a key frame foe storage, green dots represent the tracks.



(a) ORB-SLAM image feature points.



(b) ORB-SLAM map building and localization process
Figure 4. Experiment result of monocular ORB-SLAM.

The trajectory of the improved monocular visual ORB-SLAM and the reference differential GPS is shown in Fig.5.

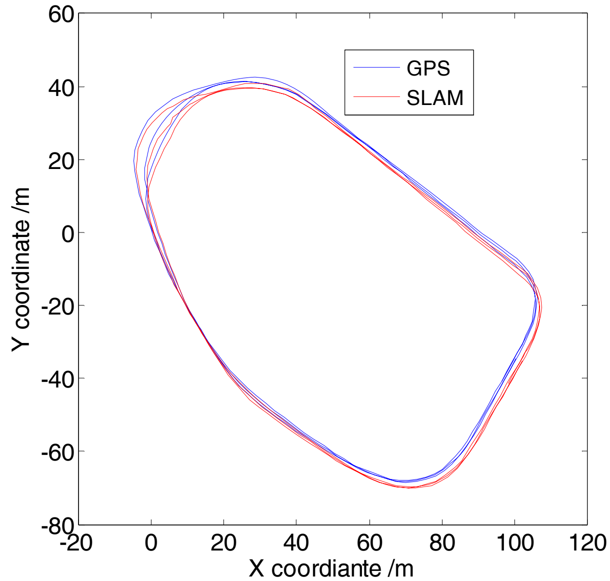


Figure 5. Trajectory comparison of ORB-SLAM and differential GPS.

The navigation error of the improved monocular visual ORB-SLAM is shown in Fig.6.

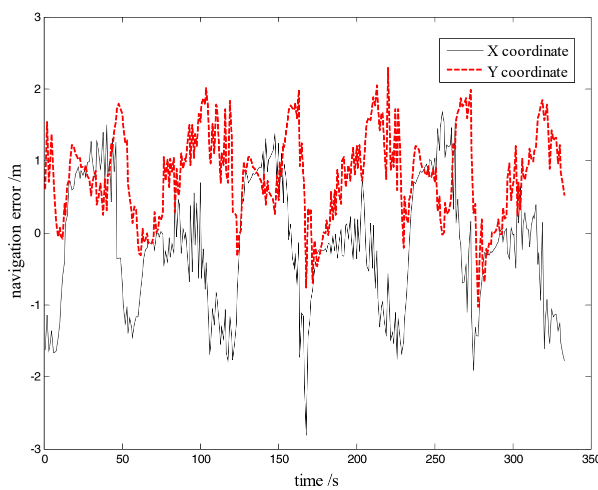


Figure 6. Navigation error of ORB-SLAM.

From the Fig.4 to Fig.6 shows that the improved monocular visual ORB-SLAM method has a higher positioning accuracy. The probability of failure of the monocular visual ORB-SLAM and the improved monocular visual ORB-SLAM method for feature points tracking were calculated by several times for experiments, respectively. Results show that the improved monocular visual ORB-SLAM method has a lower probability of tracking failure by 20 present compared to the previous one. So the improved monocular visual ORB-SLAM method has better environment adaptability.

Conclusion

In this paper, a corresponding solution is proposed to the problem of the monocular visual ORB-SLAM method. The improved monocular visual ORB-SLAM method is proved to have high positioning accuracy and robustness through the car test. Therefore, the improved monocular visual ORB-SLAM method has a high practical value.

References

[1] H. Strasdat, J.M.M. Montiel, and A.J. Davison.Visual Slam: Why filter [J]. Image and Vision Computing, 2012, 30(2): 65–77.

[2] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces [C]//2007 IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR), 2007: 225–234.

[3] K. Pirker, M. Ruther, and H. Bischof. CD SLAM -continuous localization and mapping in a dynamic world [C]//2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2011: 3990–3997.

[4] S. Song, M. Chandraker, and C.C. Guest. Parallel, real-time monocular visual odometry [C]//2013 IEEE International Conference on Robotics and Automation (ICRA), 2013: 4698–4705.

[5] E. Rosten and T. Drummond. Machine learning for high-speed corner detection[C]//2006 European Conference on Computer Vision, 2006:430–443.

[6] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. BRIEF: Binary Robust Independent Elementary Features [C]//2010 European Conference on Computer Vision, 2010: 778–792.

[7] J. Engel, T. Schops, and D. Cremers. LSD-SLAM: Large-scale direct monocular SLAM [C]//2014 European Conference on Computer Vision(ECCV), 2014:834–849.

[8] R. Mur-Artal and J.D. Tardos. Fast relocalisation and loop closing in keyframe-based SLAM [C]//2014 IEEE International Conference on Robotics and Automation (ICRA), 2014: 846–853.

[9] Mur-Artal R., Montiel J.M.M., Tardos J.D. ORB-SLAM: a Versatile and Accurate Monocular SLAM System [J]. IEEE transactions on Robotics, 2015, 1:16.