

# *Research on Improved RGB-D SLAM Algorithm based on ORB Feature*

**Binbin Xu<sup>a</sup>, Pengyuan Liu<sup>b</sup>, and Junning Zhang<sup>c</sup>**

*The Army Engineering University of PLA, Hebei 050003, China.  
<sup>a</sup>xbxmn@163.com, <sup>b</sup>983736124@qq.com, <sup>c</sup>2673500694@qq.com*

**Keywords:** SLAM, Depth camera, Real time, Accuracy.

**Abstract:** The research content of this paper is the contrast experiment of the improved RGB-D SLAM algorithm and the classical RGB-D SLAM algorithm, RGB-D SLAM V2, ORB SLAM algorithm in their real-time feature and accuracy, to verify the effectiveness and superiority of the improved algorithm. Firstly, the improved front-end, loop detection and back-end modules are integrated into the classical RGB-D SLAM framework to form an improved RGB-D SLAM system. Then the TUM data set is simulated, and the evaluation criteria are front-end time, trajectory error, and building visibility. Experimental results on standard datasets show that the improved algorithm can achieve even better results than the existing mature algorithms.

## **1. Introduction**

Image feature information based on monocular or binocular cameras does not contain depth information, so it is necessary to use triangulation method to estimate depth. Depth estimation consumes a lot of computing resources and the error of depth estimation algorithm itself, which affects the real-time and accuracy of visual SLAM system<sup>[1]</sup>. Although laser sensor can obtain depth information, but the price is too high, bringing too high industrial production costs. Depth cameras can solve both problems. The depth information is added compared with the monocular or binocular camera, which saves a lot of computing resources. Compared with the laser sensor price is lower, there is a broad consumer market. However, the depth camera also has its own limitations<sup>[2]</sup>. The measurement distance is short, and the information of the depth obtained contains noise, so the image information cannot be obtained in all directions. Affected by the change of illumination, it is suitable for indoor scenes.

Classic visual SLAM framework adds depth information to form visual RGB-D SLAM<sup>[3]</sup>. In 2010, professor Henry et al. integrated the depth information with SIFT<sup>[4]</sup> feature information and adopted GPU acceleration, which was applied to front-end pose optimization and vision-based loop detection to achieve global consistency and build 3D map. In 2012, two dense 3D maps of large indoor scenes were built using FAST<sup>[5]</sup> key points. Henry's RGB-based environmental map is the first successful attempt to add depth information to traditional visual SLAM

## 2. Improved RGB-D SLAM framework

In this paper, an improved rgb-d SLAM algorithm is proposed. As shown in figure 1. The specific algorithm is as follows: First, the Kinect camera acquires image information and reads and preprocesses it. Second, the image areas are divided into blocks and feature extraction is carried out by using ORB algorithm. The effective feature points are screened according to the response strength of Angle points, and the scale of feature map is formed and maintained. The areas lacking feature points are added once and twice highly efficient feature points from the feature map. The fast approximation nearest neighbor (FLANN) algorithm was used for feature matching and mismatched pairs were eliminated with hamming distance threshold. Thirdly, the motion transformation of key frames and feature map is calculated by RANSAC-ICP<sup>[6]</sup> algorithm to generate the initial pose map based on feature map. Fourth, the improved loop detection module was used to add constraints to the initial pose map (scene segmentation based on the feature extraction, the improved k-mean++ algorithm was used to construct the visual dictionary tree, the similarity was calculated layer by layer from top to bottom, and the error loop was removed from the loop posterior processing). Finally, the modified pose information is obtained by nonlinear optimization of the pose using g2o<sup>[7]</sup>. The moving trajectory of the robot is represented by the camera pose, all camera pose is converted to the world coordinate system and optimized, and the 3D point cloud map with global consistency is obtained.

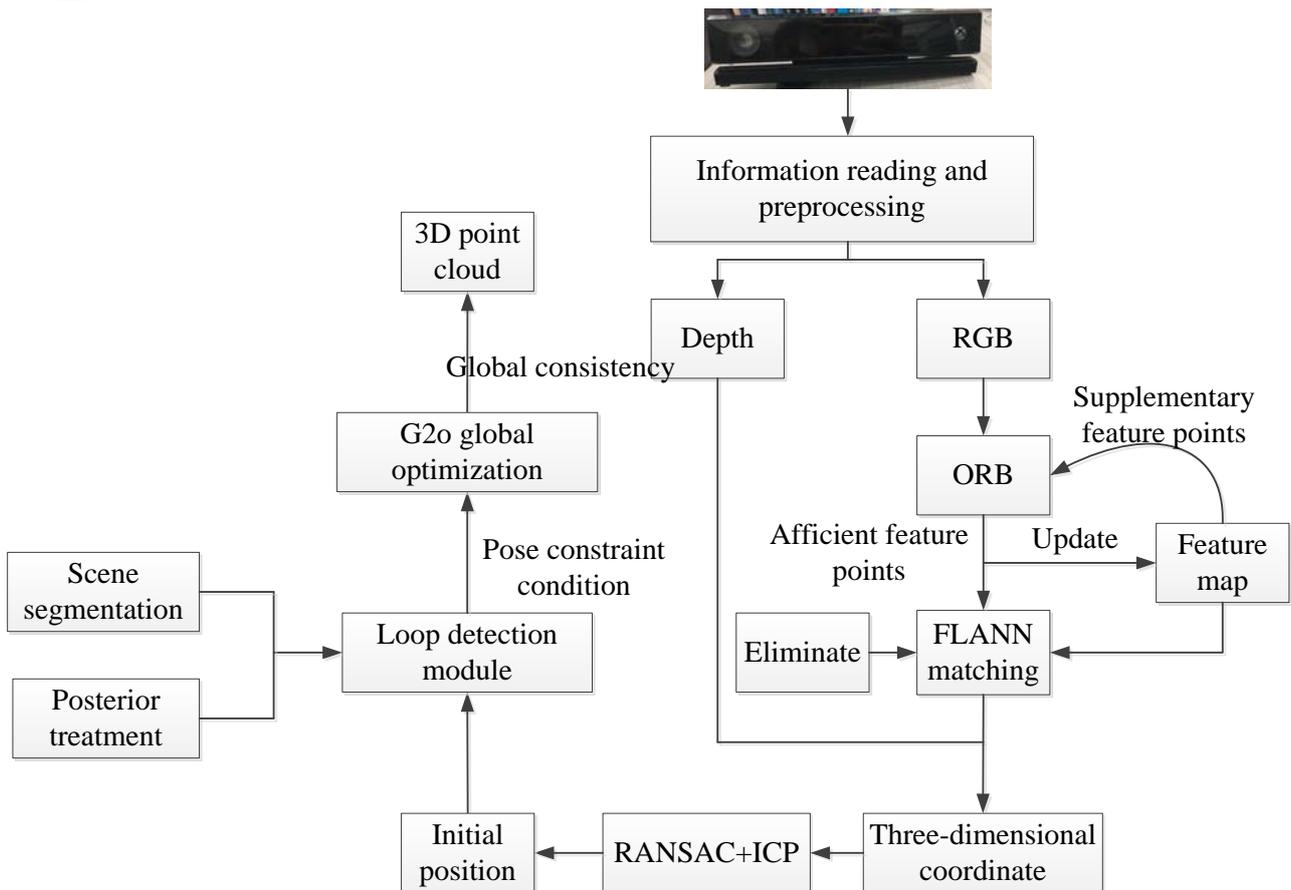


Fig. 1 Algorithm framework

### 3. Contrastive experiments with the TUM dataset

The data set used is the open source standard RGB-D data<sup>[8]</sup> set of the University of Munich, Germany, and fr1/360, fr1/room, fr2/point\_slam are selected as the experimental data sources. The above three data sets are selected from the following aspects: fr1/360 is collecting image information at 360 degrees in a typical office environment. The process of corner is easy to cause the loss of feature points of adjacent frames, which verifies the effectiveness of the feature map to supplement the feature points of the insufficient region. fr2/room starts with the office table as the starting position, and there is a loopback in the scene. Verify the superiority of the improved loopback algorithm. Fr2 /point\_slam is a large data set that mobile robots use Kinect cameras to route through entire factories. Large-scale datasets can better reflect the effectiveness of g2o optimization, and can verify the overall performance of improved RGB-D SLAM.

Table 1 is a comparison of the improved algorithm with the classical RGB-D SLAM<sup>[9]</sup>, RGB-D SLAM V2<sup>[10]</sup> and ORB-SLAM<sup>[11]</sup> based on frame matching in front-end time, standard deviation and mean deviation. From the table, we can see that the improved algorithm is better than the contrast algorithm in front-end time. Compared with RGB-D SLAM and RGB-D SLAMV2, the improved algorithm can achieve the same effect as ORB-SLAM in accuracy. However, the error generally increases in fr3/point\_slam for three reasons: First, the above four algorithms are mainly indoor applications, while point\_slam is large and similar to semi-outdoor environment. Secondly, the camera is mounted on the mobile robot. Through the change of the key frame, it can be seen that the sway of the camera is too big, and the feature information is lost too much. Third, the size of the data set is large, even if the back-end optimization, with the increase of time, the cumulative error is also increasing.

Table 1 Data set experiment

Algorithm name	Data set	Front-end time /ms	Standard deviation /cm	Mean/cm
RGB-D SLAM	fr1/360	92.3	26.2	25.8
	fr1/room	174.9	26.4	24.2
	f2/point_slam2	262.6	45.6	43.9
RGB-D SLAM V2	fr1/360	70.4	17.3	18.4
	fr1/room	155.4	21.3	20.2
	f2/point_slam2	221.3	42.4	35.7
ORB-SLAM	fr1/360	64.7	7.2	8.5
	fr1/room	120.3	4.8	4.2
	f2/point_slam2	201.7	14.6	13.3
Improved algorithm	fr1/360	60.1	6.4	7.8
	fr1/room	105.5	4.6	4.3
	f2/point_slam2	184.3	13.2	10.86

In order to observe the improvement effect of the algorithm more clearly, the trajectories generated by the four algorithms and the 3D point cloud images are carried out, as shown in Fig. 2 and Fig. 3.

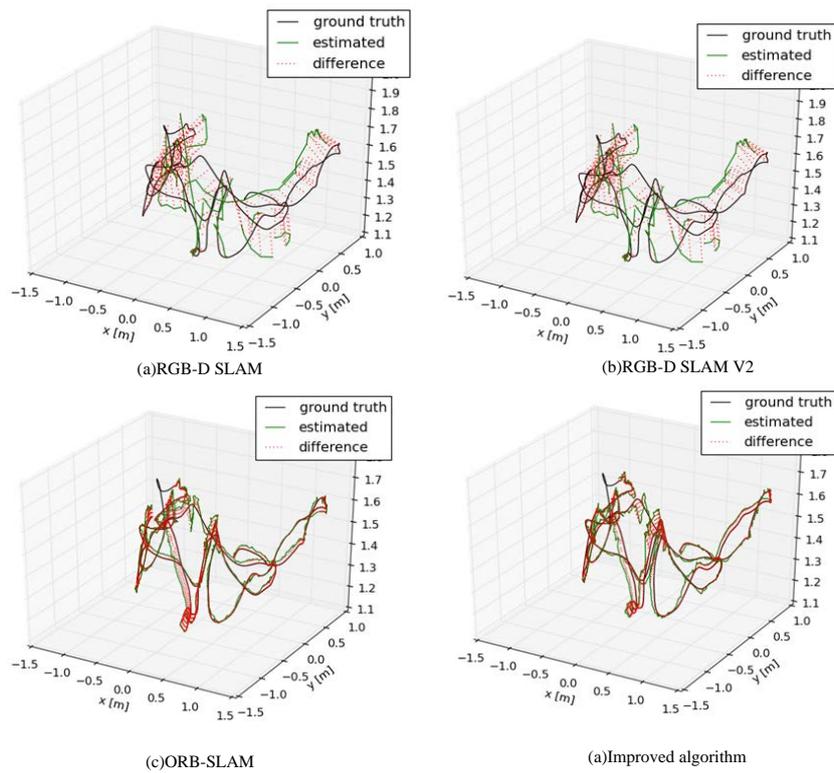


Fig. 2 Trajectory comparison diagram

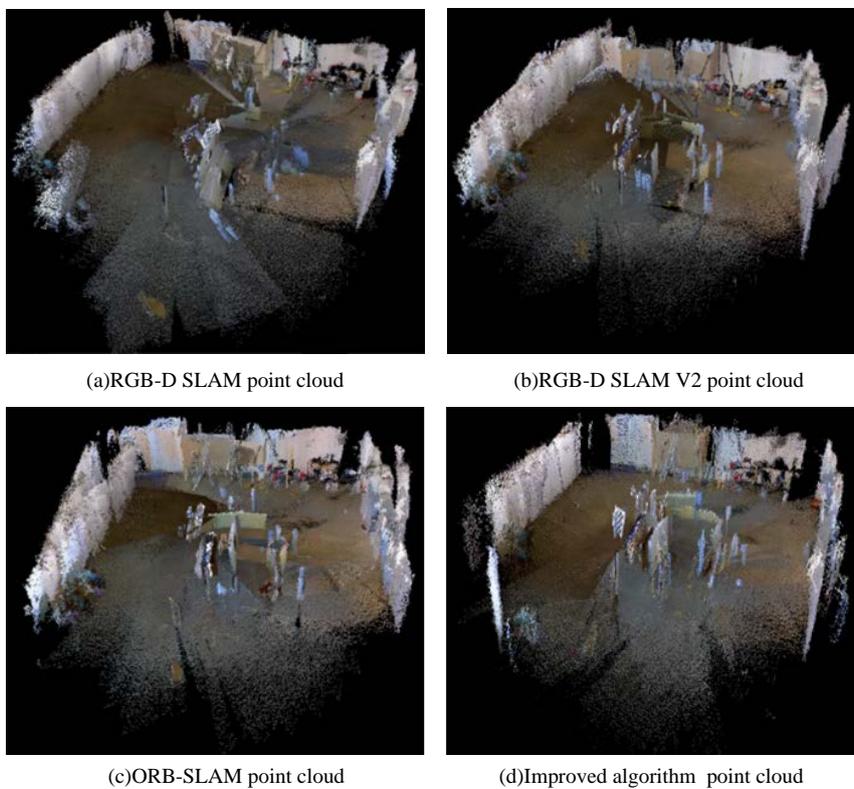


Fig. 3 3D point cloud

Fig. 3 is a 3D point cloud image constructed by the algorithm under the fr2/slam\_point data set. It can be seen that there are a lot of drift points in the graph (a), and overlapping deformation of tables, chairs and calibration boards in the center of the scene. The contours of the objects in the graph (b) and the center of the graph (c) are blurred, and drift points are also found. Compared with the above three point clouds, the drift point of the graph (d) is less, the outline of the object is clear, and the clear calibration board is obviously constructed in the central area, so the improved algorithm has higher precision.

#### 4. Summary

From the experimental results of TUM data sets, we can see that except for the accuracy of large-scale data sets can not be significantly better than ORB-SLAM algorithm, in real-time performance is better than the three comparison algorithms. The experiment verifies the effectiveness of the improved rgb-d SLAM algorithm, estimates the motion posture efficiently in real time and constructs the 3D point cloud map, indicating that the improved algorithm can be applied to the robot.

#### Acknowledgments

Correspondence author: Liu Pengyuan, 983731624@qq.com

#### References

- [1] Jianbin Chen, Jun Li, Yang Xu, et al. A Compact Loop Closure Detection Based on Spatial Partitioning[C].J2017 2nd International Conference on Image, Vision and Computing, IEEE, 2017:71-375.
- [2] Lu Xianwei. Research on SLAM algorithm based on RGB-D data [D].Beijing: Beijing Institute of Technology, 2016.
- [3] WANG Liu-jun CHEN Jia-bin,YU Huan,et al.An Over-view of RGB-D SLAM[J].Navigation Positioning and Timing,2017,4(6):10-18.
- [4]D. G.Lowe.Distinctive image features form scale- nvariant keypoints international Journal of Computer Vi-sion[J].Robotic,2014,4(2): 91 -110.
- [5] C. Forest, M. Pizzoli, and D. Scaramuzza, "Svo : Fast semi - direct monocular visual odometry ," in Robotics and Automation (ICRA), 2014 IEEE International Confer-ence on [C], IEEE, 2014:15-22.
- [6] Eendes F,Hess J, STURM J,et,al.3D mapping with an RGB-D camera[J].IEEE Transactions on Robotics, 2014, 30(1):177-187.
- [7] ZHANG Yi,SHA Jiansong. Visual-SLAM for mobile robot based on graph optimization [J]. CAAI Transactions on Intelligent Systems, 2018, 13(2) :290-295 .
- [8] Silberman N,Fergus R.Indoor scene segmentation using a structured light sensor[C].2011 IEEE International Conference on Computer Vision.2011,601-608.
- [9] G.Klein and D. Murray.Paraller tracking and mapping for small ar workspace [C] .ISMAR, 2007,225-234.
- [10] GAO Xiang. Visual SLAM fourteen: from theory to practice [M]. BeiJing: Publishing House of Electronics Industry, 2016:137-138
- [11] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Trnsactions on Robotics. 2015, 31(5):1147 -1163