2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)
Daejeon, Korea / November 26-29, 2012

GPU-Based Real-Time RGB-D 3D SLAM

Donghwa Lee[1], Hyongjin Kim[1], and Hyun Myung[2]

[1]Department of Civil and Environmental Engineering,
KAIST, Daejeon 305-701, Republic of Korea
(Tel : +82-42-350-5670; E-mail: {leedonghwa,hjkim86}@kaist.ac.kr)
[2]Department of Civil and Environmental Engineering, Robotics Program,
KAIST, Daejeon 305-701, Republic of Korea
(Tel : +82-42-350-3630; E-mail: hmyung@kaist.ac.kr)

Abstract - This paper proposes a GPU (graphics processing unit)-based real-time RGB-D (red-green-blue depth) 3D SLAM (simultaneous localization and mapping) system. RGB-D data contain 2D image and per-pixel depth information. First, 6-DOF (degree-of-freedom) visual odometry is obtained through the 3D-RANSAC (three-dimensional random sample consensus) algorithm with image features. And a projective ICP (iterative closest point) algorithm gives an accurate odometry estimation result with depth information. For speed up extraction of features and ICP computation, GPU-based parallel computation is performed. After detecting loop closure, a graph-based SLAM algorithm optimizes trajectory of the sensor and 3D map.

Keywords - 3D SLAM, RGB-D camera, image features, projective iterative closest point, 3D-RANSAC

1. Introduction

Over the past decade, the SLAM (Simultaneous Localization and Mapping) problem has been studies by many researchers [1–6]. Recently, 3D SLAM techniques using RGB-D (red-green-blue depth) camera have been focused. The RGB-D camera gives color image and per-pixel depth data concurrently. (Fig. 1)

In the 3D SLAM, a variety of image feature extraction algorithms such as SIFT (Scale-Invariant Feature Transform), SURF (Speeded Up Robust Features), and FAST (Features from Accelerated Segment Test) etc. is employed for visual odometry estimation. In addition, ICP (Iterative Closest Point) algorithm is used to improve the accuracy with depth data. And graph-based SLAM core algorithms optimize the full trajectory and 3D model. Loop closure detection also makes use of the features extraction methods. However, the computational complexity of these works is considerable, so that the real-time operation is not easy [6–10].

Microsoft Research presented a KinectFusion algorithm which gives 3D model at 30 Hz using the Kinect sensor and a GPU (Graphics Processing Unit) processor [11]. The GPU processor handles the depth data for estimating 3D movement of the camera and mapping 3D model in real-time. But this work has weakness in drift noise, since they did not have any algorithms for loop closure detection and SLAM optimization.

In this paper, we propose GPU-based RGB-D 3D SLAM system which has real-time processing rate. 3D-RANSAC (RANdom SAmple Consensus) algorithm and the projective ICP algorithm are used for visual odometry estimation with image features and the depth data. A feature manager performs loop closure detection. And then the iSAM (Incremental Smoothing And Mapping) graph-based SLAM algorithm optimizes the full trajectory of the camera and the 3D model. iSAM is a high-speed online SLAM core algorithm based on sparse linear algebra [4].

(a) (b)

Fig. 1 RGB-D sensor data. (a) RGB color image. (b) Per-pixel depth data.

2. Proposed 3D SLAM System

Overview of our 3D SLAM system is illustrated in Fig. 2.

First of all, image features are extracted using SURF algorithm. Recently, the GPU computing has been applied for speed-up feature extraction. In this system, all of the image data from the sensor are processed in real-time. Each 2D feature is located at a point in three-dimensional coordinate space with depth information. 6-DOF visual odometry estimation is performed with feature matching and 3D-RANSAC algorithm.

Second, the projective ICP algorithm estimates 6-DOF odometry accurately with the depth data of the current frame. A point cloud from the depth data finds the nearest point and aligns to previous frame repeatedly. The projective ICP algorithm is also performed on GPU processor.

Third, a feature manager gathers the features and detects loop closure through comparison between the current and the preceding features (Fig. 3). After that, the current frame is matched to the past trajectory of the sensor. 6-DOF constraints to the previous trajectories are also obtained from the feature matching and the 3D-RANSAC algorithm.

Next, the trajectory and constraints of the sensor pose

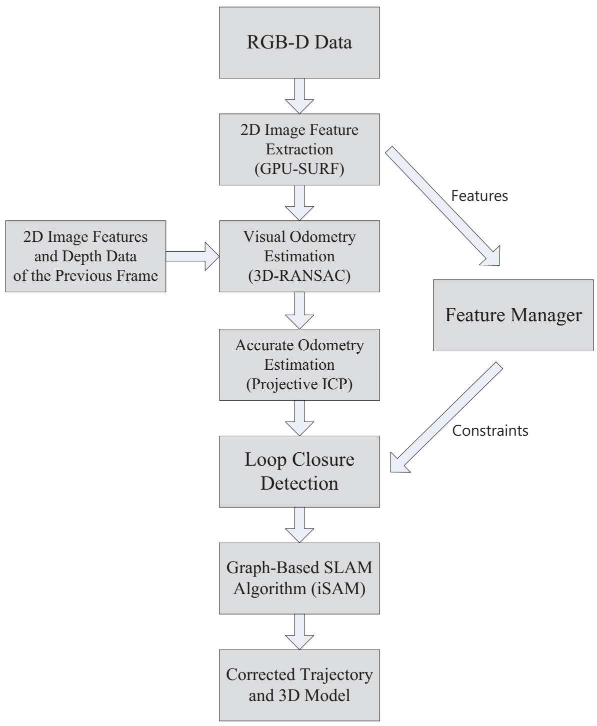


Fig. 2 Overview of the proposed RGB-D 3D SLAM system.

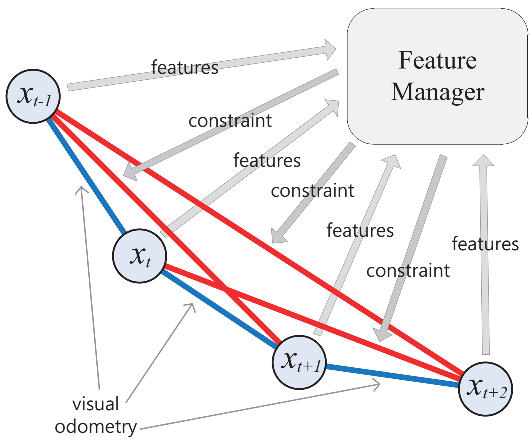
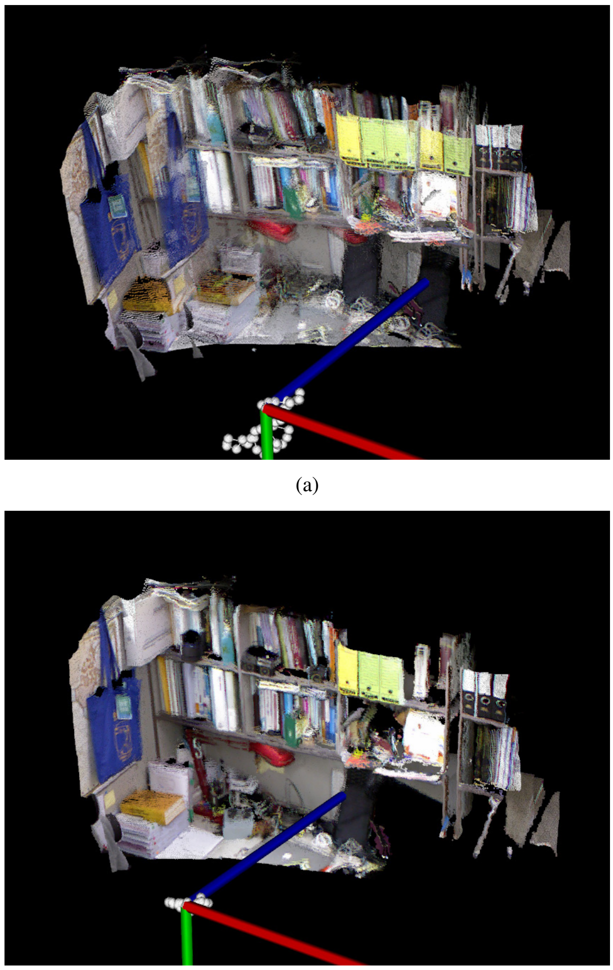


Fig. 3 Feature manager and loop closure detection.

generate graph form. After online Graph-based SLAM algorithm optimizes the constraint graph, the corrected trajectory and the 3D map can be obtained. In this system we use iSAM algorithm which solves graph-based SLAM problem using sparse linear algebra in real-time [4].

3. Experiments

We have performed experiments in a room-size environment. A Mocrosoft Kinect sensor is used for RGB-D data. The sensor has valid depth range from 0.5m to 5m. 2D RGB color image and the depth data are given at 30 frames per second with 640 × 480 resolution. Our 3D



(a)

(b)

Fig. 4 Experimental results. (a) 3D model by only visual odometry. (b) Optimized 3D model by 3D SLAM algorithm.

SLAM system implemented on an Intel Core i7 CPU and an Nvidia GT 560 Ti graphic card supporting GPU programming with CUDA language.

Experimental results are presented in Fig. 4. The whole 3D point cloud data from each node is drawn in 3D space. Fig. 4(a) shows a result with only visual odometry. Odometry estimation noise drifts the sensor trajectory, so that the 3D map is misaligned. In Fig. 4(b), an optimized node graph is shown. The trajectory of the sensor is corrected and the 3D model also aligned.

Average processing times of the proposed system is 45 milliseconds, therefore the rate of our 3D SLAM system is above 20Hz.

4. Conclusion

This paper proposed a GPU-based real-time RGB-D 3D SLAM system using only an RGB-D sensor. The 6-DOF visual odometry is obtained from the 2D image features, the depth data and the 3D-RANSAC algorithm. The projective ICP algorithm improves the accuracy of the odometry estimation. The feature manager handles image features and detects loop closure. Then the graph-based SLAM algorithm optimizes the full trajectory and the 3D model. The GPU computing accelerates operation speed of the 3D SLAM system, and the average processing rate is above 20 Hz. We will evaluate the trajectory and the 3D model with ground truth data in the future.

Acknowledgement

This work was supported partly by the R&D program of the Korea Ministry of Knowledge and Economy (MKE) and the Korea Evaluation Institute of Industrial Technology (KEIT). (The Development of Low-cost Autonomous Navigation Systems for a Robot Vehicle in Urban Environment, 10035354) This work was also financially supported partly by Korea Ministry of Land, Transport and Maritime Affairs (MLTM) as U-City Master and Doctor Course Grant Program.

References

[1] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, 2001.

[2] M. Michael, T. Sebastian, and K. Daphne, "FastSLAM 2.0: An Improved Particle Filtering Algorithm for Simultaneous Localization and Mapping that Provably Converges," *Proc. of The 18th International Joint Conference on Artificial Intelligence*, 2003.

[3] D. Hahnel, W. Burgard, D. Fox, and S. Thrun, "An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," *Proc. of 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 1, pp. 206–211 vol. 1, 2003.

[4] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.

[5] R. Kummerle and G. Grisetti, "g2o: A General Framework for Graph Optimization." in *Proc. of 2011 IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[6] M. Krainin, D. Maturana, and D. Fox, "Visual odometry and mapping for autonomous flight using an rgb-d camera," *Proc. of The 15th International Symposium on Robotics Research (ISRR)*, 2011.

[7] F. Endres, W. Burgard, and D. Cremers, "A Benchmark for the Evaluation of RGB-D SLAM Systems," *Proc. of 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Aug. 2012.

[8] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 647–663, Apr. 2012.

[9] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An evaluation of the RGB-D SLAM system," in *Proc. of 2012 IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 1691–1696.

[10] N. Engelharda, F. Endresa, J. Hessa, J. Sturmb, and W. Burgarda, "Real-time 3D visual SLAM with a hand-held RGB-D camera," *Proc. of the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum, Vasteras, Sweden*, vol. 2011, 2011.

[11] R. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. Davison, P. Kohli, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking," *Proc. of 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 127–136, 2011.