Three-Dimensional Indoor Mobile Mapping With Fusion of Two-Dimensional Laser Scanner and RGB-D Camera Data

Article *in* IEEE Geoscience and Remote Sensing Letters · April 2014

DOI: 10.1109/LGRS.2013.2279872

CITATIONS
11

READS
141

5 authors, including:

Chenglu Wen
Xiamen University
42 PUBLICATIONS **167** CITATIONS
SEE PROFILE

Qingyuan Zhu
Xiamen University
8 PUBLICATIONS **23** CITATIONS
SEE PROFILE

Cheng Wang
Xiamen University
157 PUBLICATIONS **894** CITATIONS
SEE PROFILE

Jonathan Li
University of Waterloo
254 PUBLICATIONS **3,243** CITATIONS
SEE PROFILE

Three-Dimensional Indoor Mobile Mapping With Fusion of Two-Dimensional Laser Scanner and RGB-D Camera Data

Chenglu Wen, Ling Qin, Qingyuan Zhu, Cheng Wang, *Member, IEEE*, and Jonathan Li, *Senior Member, IEEE*

***Abstract*—Three-dimensional mobile mapping in indoor environment, mostly global navigation satellite system-denied space, is to consecutively align the frames to build a global 3-D map of an indoor environment. One of the major difficulties of the current solutions is the failure at the insufficient overlapping between the frames, which is the reality of a lack of correspondences between the frames. To overcome this problem, a 3-D indoor mobile mapping system that integrates a 2-D laser scanner, and an RGB-Depth camera is presented in this letter. In this system, a fusion-iterative closest point (ICP) method, which combines the 2-D mobile platform pose from a Rao–Blackwellized particle filter estimation, an ICP, and a generalized-ICP method, is proposed for the consecutive frame alignment. Fusion-ICP achieves effective frame alignment, particularly in solving the insufficient overlapping frame alignment problem. Comparative experiments were conducted to evaluate the mapping system. The experimental results demonstrate the effectiveness and efficiency of our system for 3-D indoor mobile mapping.**

***Index Terms*—Laser scanner, mobile platform, RGB-D camera, sensor fusion, 3-D indoor mapping.**

I. Introduction

ACCURATE maps of interior spaces created by mobile mapping form the basis for applications such as emergency response, situational awareness, and creation of accurate floor plans. The process of the mobile indoor mapping is nonreliant on global navigation satellite system (GNSS) localization, since these signals cannot penetrate into the interiors of buildings. Building a 3-D map requires the current position and orientation of the mobile platform, while estimating the location of a mobile platform in an environment requires a 3-D map of that environment. This process is referred to as simultaneous localization and mapping (SLAM) [1].

Manuscript received May 14, 2013; revised July 14, 2013 and August 18, 2013; accepted August 24, 2013. Date of publication September 16, 2013; date of current version December 2, 2013. This work was supported in part by the Natural Foundation of Fujian Providence under Project 2011J05159 and in part by the Fundamental Research Funds for Central Universities under Project 2011121045.

C. Wen, L. Qin, and C. Wang are with the School of Information Science and Engineering, Xiamen University, Xiamen 361005, China (e-mail: clwen@xmu.edu.cn; aql.yi5637@163.com; cwang@xmu.edu.cn).

Q. Zhu is with the Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005, China (e-mail: zhuqy@xmu.edu.cn).

J. Li was with the School of Information Science and Engineering, Xiamen University, Xiamen 361005, China. He is now with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2013.2279872

In recent decades, 2-D SLAM solutions based on the 2-D laser scanner have achieved great success by applying the Bayesian probability model. However, such solutions are confined to building 2-D grid maps, whereas the 3-D map brings exponential complexity to the Bayesian probability model. The RGB-D camera, which mainly uses an active vision method or time-of-flight sensing technology to generate the depth image, is a novel sensor that captures an RGB image and a depth image simultaneously. Since there is a corresponding relationship between the pixels in the RGB and depth images, the RGB-D camera can build a dense 3-D map with appearance information. The iterative closest point (ICP) [2] algorithm or its variants are generally used for 3-D frame alignment. However, ICP is a local optimum iterative algorithm that suffers from huge computational consumption that adversely affects the instantaneity of the 3-D mapping. The algorithm also suffers from unstable convergence caused by bad initialization of the rigid transformation that ultimately adversely affects the accuracy of the 3-D map.

This letter presents a 3-D indoor mobile mapping system, with data fused by a 2-D laser scanner and an RGB-D camera. This system can precisely locate the mobile platform and build a 2-D grid and a 3-D dense map. First, a SLAM posterior probability, using a 2-D laser scanner and odometer data, is estimated, and a 2-D grid map and mobile platform pose are achieved. Second, 3-D data frames from the RGB-D camera are consecutively aligned by our proposed fusion-ICP method. Here, the 2-D mobile platform pose from 2-D SLAM and sparse visual features from RGB images are used to accelerate the ICP convergence. Because the pair-wise alignment between frames is not globally consistent, the following are applied to build an accurate global 3-D map: 1) loop closure detection by a particle weight threshold; and 2) pose graph optimization by minimizing nonlinear error function.

II. Related Work

The state-of-the-art 2-D SLAM is generally considered as a posterior probability estimation problem usually solved with a Bayesian probability model. The extended Kalman filter [3] and the particle filter [4] are the two primary solutions for the probability model. The 2-D map represents the environment as the relationship between the mobile platform and the obstacles in a plane, which results in limitations of environmental representation.

There are basically two SLAM solutions based on 3-D information. One is the probability-based model in the 3-D

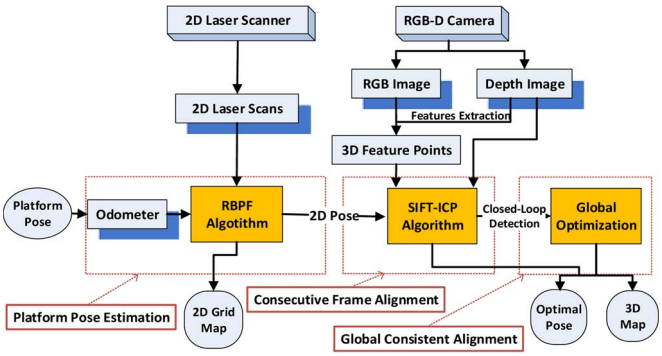


Fig. 1. Framework of the system.

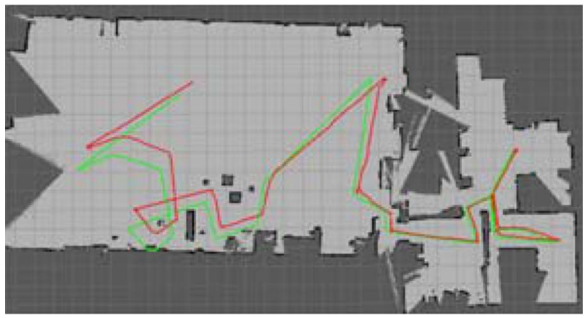


Fig. 2. Simulated grid map built by GMapping. (Green line) Odometer trajectory and (red line) platform trajectory estimated.

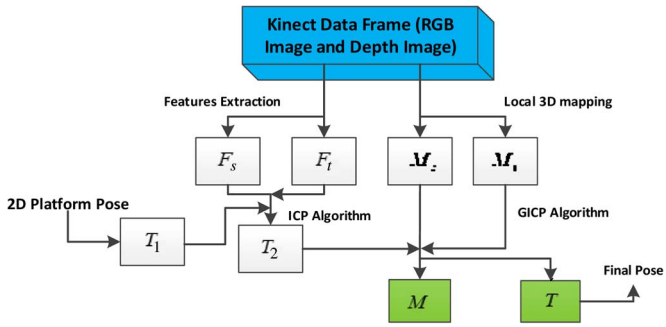


Fig. 3. Fusion-ICP method for frame alignment.

spaces [5], whereas the other is the scan-matching-based algorithm, such as ICP. For example, Bouvrie *et al.* integrated position, inertial measurement unit data, and robot velocity to accelerate the ICP algorithm, thus improving the efficiency of 3-D mapping [6]. Andreasson and Stoyanov used local visual speeded-up robust features and the 3-D-Normal Distributions Transform method for frame alignment [7].

Kinect (an RGD-D camera) brought about a revolutionary breakthrough in 3-D mapping in the fields of indoor mobile mapping and computer vision [8]. Xu *et al.* refined plane parameters of 3-D data based on the segments obtained from the 2-D image of an RGB-D camera, whereas the frame alignment problem was not addressed in the paper [9]. Stückler and Behnke integrated depth and color information for dense multiresolution scene mapping using RGB-D camera, and frames were registered by an improved ICP method [10]. Henry *et al.* built an RGB-D mapping system that applies the scale invariant feature transform (SIFT) feature [11] and ICP algorithm to generate a dense indoor 3-D model [12]. The work that is most similar to the RGB-D mapping system is RGB-D SLAM, which is an open-source system [13]. The sparse visual keypoints are extracted from data frames and used for aligning point clouds by the generalized-ICP (GICP) method [14].

However, most of the above frame aligning methods work only when there is sufficient overlapping between the two frames, in which corresponding sparse visual features between the frames are calculated based on the overlapping part. Moreover, frame alignment performance of the ICP algorithm is dramatically affected by rigid transformation initialization. Du *et al.* estimated the camera pose by a motion capture system through a calibration process, and the camera pose provides initial transformation for ICP-based frame alignment when the movement between the two frames is large [15].

III. Proposed Method

Our proposed 3-D indoor mobile mapping system (as shown in Fig. 1) is divided into three modules as follows: mobile platform pose estimation, consecutive frame alignment, and globally consistent alignment.

A. Two-Dimensional Laser Scanner-Based Platform Pose Estimation

The particle filter is a serialized Monte Carlo filtering method that uses particle sets to approximate the posterior probability of the SLAM problem. Given the mobile platform observations $Z_{0:t}$, the mobile platform odometry measurement $U_{0:t}$, and the initial mobile platform location x_0, the posterior about potential mobile platform trajectories $x_{0:t}$ and maps m is represented as

$$p(X_{0:t}, m|Z_{0:t}, U_{0:t}, x_0) = p(m|X_{0:t}, Z_{0:t}) \times p(X_{0:t}|Z_{0:t}, U_{0:t}, x_0). \quad (1)$$

To estimate the pose of the mobile platform, we apply the GMapping solution, an open-source approach that adaptively reduces the particle number of the Rao–Blackwellized particle filter (RBPF) method to solve the SLAM problem in the robot operating system [16]. By decoupling the state space of the Markov chain, RBPF simplifies the steps required to estimate the probability and greatly reduces the number of particles. GMapping greatly reduces the number of samples of the particles by approximating the sampling weights using a Gaussian distribution. Subsequent to implementing 2-D SLAM with GMapping, a 2-D grid map for indoor environment is built, and the mobile platform trajectory is estimated. An example of a 2-D grid map with a mobile platform trajectory built by GMapping using simulated data is shown in Fig. 2. Because the frame alignment is processed in 3-D space, the 2-D mobile platform pose must be converted into a 4×4 matrix, T1, as the initial estimation of the frame alignment.

B. Consecutive Alignment With Fusion-ICP

The proposed fusion-ICP method includes two separate iterations, namely, an ICP-based iteration and a GICP-based iteration (see Fig. 3). The detailed method is executed in the following steps.

1) Step 1: The RGB images I_s and I_t and two local 3-D point cloud maps M_s, and M_t are obtained from two adjacent data frames. SIFT descriptors are separately extracted from the RGD images I_s and I_t, and the

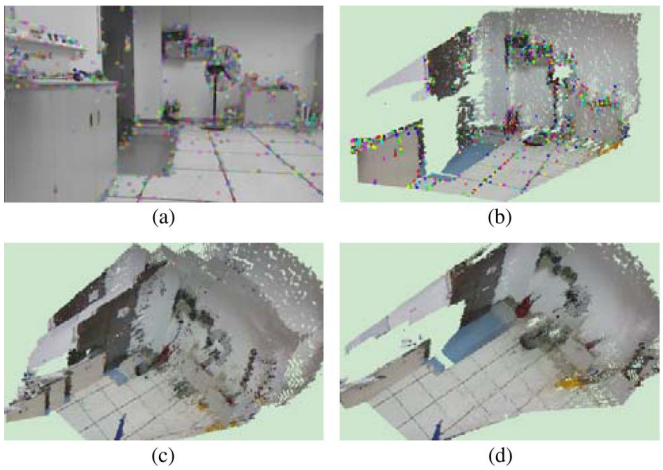


Fig. 4. Consecutive alignment. (a) 2-D SIFT features. (b) Corresponding 3-D SIFT points. (c) Before alignment. (d) Frame aligned by fusion-ICP method.

matched SIFT pairs in the two images are obtained by a fast nearest neighbor searching method [11]. Then, the corresponding SIFT 3-D point sets, i.e., F_s and F_t, are obtained. Finally, using the original ICP algorithm with point sets F_s and F_t as the input and mobile platform pose T_1 as the initial rigid transformation, the initial rigid transformation for adjacent frames, i.e., T_2, is computed.

2) Step 2: The final rigid transformation, i.e., T, between maps M_s and M_t is estimated using the GICP algorithm with rigid transformation T_2 for the initial iterative computation. The optimal rigid transform T, consisting of rotation and translation component (R and t), is used for registering the point cloud frames into a global 3-D coordinate system.

To deal with frame alignment with insufficient correspondence or nonoverlapping frames, we compare the number of the 3-D corresponding feature points to a specified threshold right after step (1). If sufficient 3-D correspondences exist, the iterative convergence of T_2 by ICP is accelerated by an initialization transformation T_1 from a platform pose estimation. Then, an optimal rigid transformation T is estimated by GICP. If insufficient correspondence exists, T_1 will directly replace T, since the platform and the RGB-D camera are considered as rigidly assembled and with the same pose change.

Fig. 4 shows an example of a consecutive alignment using the fusion-ICP method. We extract SIFT features [see Fig. 4(a)] from the RGB image and the corresponding 3-D feature points [see Fig. 4(b)] from the point cloud. Then, we align two frames [see Fig. 4(c)] to create a common coordinate system [see Fig. 4(d)].

C. Globally Consistent Alignment

By aligning frames consecutively, it is expected that a global 3-D map should finally be built. However, misalignment among consecutive frames, or noise in the depth data, results in the mobile platform pose estimation deviating over time. In particular, when exploring a large-scale environment, the cumulative error in frame alignment causes a series of unpredictable errors, which is known as the closed-loop problem.

Since the 2-D grid map and the 3-D point cloud map geometrically represent the real world, a 3-D map is consistent with a 2-D map. Therefore, the loop closure detection method in RBPF can be applied to recognize previously visited locations in a 3-D map. Along with exploration of the environment, the particle weight of RBPF changes over time. Once the variance of the particle weight values falls within a given threshold, it is considered to meet a closed loop. At this time, RBPF resamples its particles, resets all particles' weights, and corrects the pose and 2-D map. At this point, the loop closure information is shared with the 3-D map building process.

Creation of a globally consistent trajectory and an accurate 3-D map requires globally consistent alignment, which can be treated as a nonlinear least-square problem. One solution is to represent the 3-D map as a graph structure, in which each node of the graph is a mobile platform pose-related state variable to be optimized, and each edge between the two nodes represents a geometric constraint. Then, the global consistent alignment problem is converted to a pose graph optimization process. We employ the g^2o framework to optimize the pose graph [17]. The g^2o framework efficiently reduces the errors in the pose graph, where each node in the pose graph is parameterized by the rigid transformation, and each edge between the two nodes depicts constraints calculated based on associated covariance matrices. When returning to a known area in a 3-D map, accumulative errors are minimized and global optimal poses are generated, and a global accurate map is achieved.

Pose graph optimization can be solved by minimizing the error function, i.e.,

$$F(x) = \sum_{\langle i,j\rangle \in C} e(x_i, x_j, z_{ij})^T \Omega_{ij} e(x_i, x_j, z_{ij}) \tag{2}$$

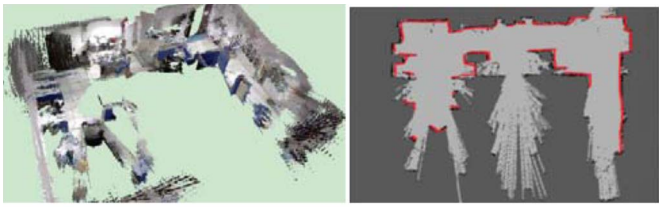
$$x^* = \arg\min_x F(x) \tag{3}$$

where $x = (x_1^T, \ldots, x_n^T)^T$ is a vector of state variables, where each state variable x_i represents a pose or rigid transformation, Ω_{ij} and z_{ij} depict the information matrix and the mean of a constraint of the state variables x_i and x_j, respectively. $e(x_i, x_j, z_{ij})$ is a vector error function that measures how much the state variables x_i and x_j match the constraint z_{ij}. When x_i and x_j perfectly satisfy z_{ij}, the estimated trajectory exactly matches the real path has visited.

D. Map Modeling

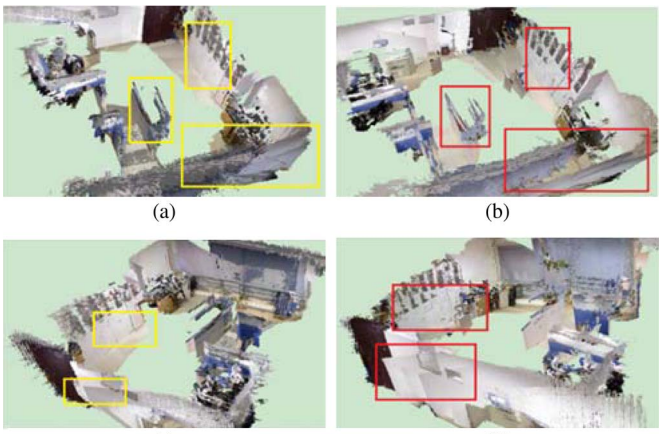
The global 3-D map achieved after pose graph optimization contains a large amount of data, which is redundant and requires computation. Surface reconstruction of a 3-D map is usually the optimal choice. However, mobile mapping raises some extra requirements, including robustness to noises, nearly real-time speed, and incremental surface mapping.

To accelerate modeling process, a fast greedy triangulation surface reconstruction method [18], which is a greedy algorithm-based method, is used for building a map model where the triangular edges are directly written and never deleted. The points are consecutively connected to form triangles by pruning points according to the following criteria: visibility, maximum angle, and minimum angle.



(a) (b)

Fig. 5. 2-D and 3-D maps of Robot Lab 1 in CSD building after pose graph optimization. (a) 2-D map built by 2-D laser scanner. (b) Corresponding 3-D map.



(a) (b) (c) (d)

Fig. 6. Example of a built 3-D map. (a) 3-D map in one perspective by our system. (b) 3-D map in corresponding perspective to (a) by RGB-D SLAM system. (c) 3-D map in other perspective by our system. (d) 3-D map in corresponding perspective to (c) by RGB-D SLAM system.

IV. Experiments and Results

The hardware of our 3-D mapping system consists of the following: a mobile robot Pioneer 3-AT, a 2-D laser scanner SICK LMS-100, and a Kinect camera. Robot pose, 2-D grid map, and 3-D global map are obtained while the robot is exploring the environment. Several 3-D mapping experiments under an indoor environment were conducted to evaluate system performance. We also compared our system with RGB-D SLAM system.

A. Globally Consistent Alignment

Two data sequences were used to evaluate global optimization in our mapping system. The two sequences were obtained from Robot Lab 1 and a corridor in the Cognitive Science Department (CSD) building of Xiamen University. Fig. 5 shows a more complicated 3-D map of Robot Lab 1. The 2-D grid map of Robot Lab 1 is built by the 2-D laser scanner [see Fig. 5(b)]; the Robot Lab path is approximately 13 m long. Fig. 5 demonstrates that 2-D maps are highly consistent with 3-D maps in our 3-D mapping system.

B. Comparison With RGB-D SLAM

We compared our 3-D mapping system with the RGB-D SLAM system by building 3-D maps of a closed loop approximately 16 m long within the CSD building of Xiamen University (Robot Lab 2).

Fig. 6 shows some detailed mapping results between the two systems. Compared with the 2-D gird map, the two systems both produce 3-D dense maps with few misalignments. Some details are marked by blocks in the 3-D maps built by our system and the RGB-D SLAM system. The detailed comparisons show our system achieves a more accurate and robust mapping performance in reducing alignment errors. We used 2-D robot pose to estimate the initial rigid transformation, whereas RGB-D SLAM used sparse visual feature matching.

TABLE I
Euclidean Norms for Consecutive Frames Alignment Using Different ICP Variations in Meters

	Robot Lab1-1	Robot Lab1-2	Robot Lab2-1	Robot Lab2-2	Avera-ge
ICP(Avg.)	**0.067**	**0.127**	**0.049**	**0.045**	**0.072**
ICP(Max.)	0.104	0.238	0.077	0.079	0.125
GICP(Avg.)	**0.032**	**0.099**	**0.044**	**0.035**	**0.053**
GICP(Max.)	0.082	0.172	0.062	0.049	0.091
SIFT+ICP(Avg.)	**0.021**	**0.039**	**0.051**	**0.015**	**0.032**
SIFT+ICP(Max.)	0.049	0.047	0.075	0.019	0.048
Fusion-ICP(Avg.)	**0.027**	**0.039**	**0.071**	**0.015**	**0.038**
Fusion-ICP(Max.)	0.043	0.045	0.073	0.017	0.045

TABLE II
Timing Results per Frame for Consecutive Frames Alignment Using Different ICP Variations in Seconds

	Robot Lab1-1	Robot Lab1-2	Robot Lab2-1	Robot Lab2-2	Average
ICP	0.975	1.116	1.059	0.697	0.962
GICP	4.472	3.975	4.288	3.654	4.097
SIFT+ICP	1.497	1.764	1.904	1.054	1.555
Fusion-ICP	1.048	1.626	1.940	0.980	1.399

C. Frame Alignment Comparison With ICP Variations

We compared the proposed fusion-ICP method with ICP and two other ICP variations, namely, GICP and SIFT + ICP. We evaluated them by aligning consecutive frames for four different indoor scenes and analyzed the error and timing results. We collected the following four data sets: a) Robot Lab 1-1 (approximately 20 m long) within Robot Lab 1; b) Robot Lab 1-2 (approximately 30 m long) within Robot Lab 1; c) Robot Lab 2-1 within Robot Lab 2, achieved by rotating the robot approximately 100°; and d) Robot Lab 2-2 (approximately 15 m long) within Robot Lab 2.

Some of the above four data sets have a few objects in their environment, whereas some data sets have many objects. Consequently, the environment differs from data set to data set. For the four methods, Euclidean norms of consecutive frame alignment are compared. In most cases, our fusion-ICP method performed well compared with the other three methods (see Table I). In particular, our method outperformed SIFT + ICP method regarding the maximum alignment error, and this usually happens when the mobile platform is making turns at the corner or there is lack of features between the frames.

We also evaluated timing results according to the above data sequences. As shown in Table II, for four data sequences, the fusion-ICP method achieved acceptable timing results per frame. In particular, the average running time of the fusion-ICP method is 1.399 s/frame, which is faster than the GICP method but slower than the ICP method, because the fusion-ICP method includes platform pose estimation, ICP, and GICP steps. In fusion-ICP, the GICP step is used to align the frames after a good estimation of initial rigid transformation is achieved in the ICP step, thus accelerating the aligning process.

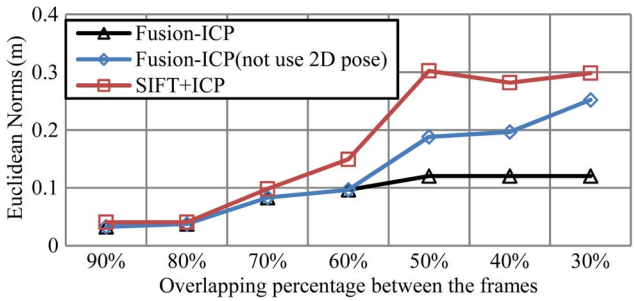


Fig. 7. Alignment accuracy comparison with different overlapping percentages.

(a) (b)

Fig. 8. Nonoverlapping frame alignment. (a) Two local data frames required. (b) Aligning nonoverlapping frames by robot pose.

D. Nonoverlapping Frame Alignment

A visual sparse feature-based frame alignment method is valid only when sufficient correspondence between frames exists; the method cannot work when only little or no correspondence exists because of the insufficient corresponding features for ICP. Our 3-D mapping system is designed to deal with frame alignment with little correspondence, even nonoverlapping frame alignment. The robot pose obtained by 2-D SLAM initializes the rigid transformation between the consecutive frames when sufficient feature correspondence exists; it aligns the frames directly when insufficient correspondence exists or nonoverlapping frames are aligned.

We acquired the laser scanner and RGB-D data frame at one position and then rotated the robot through a certain angle to acquire another data frame to achieve frames with different overlapping percentages. The Euclidean norms of aligning frames are summarized in Fig. 7 with different overlapping percentages for fusion-ICP method and SIFT + ICP method.

We observed that the alignment accuracy of both methods decreased when frame overlapping percentage decreased, and the fusion-ICP method (triangle icons in Fig. 7) outperformed the SIFT + ICP method (square icons in Fig. 7) overall. The fusion-ICP method triggers the process of aligning the frames by the current estimated 2-D pose of the platform when the Euclidean norms are larger than a threshold of 0.1 (diamond icons in Fig. 7). Under this situation, the SIFT + ICP fails to align the two frames, according to the large Euclidean norms. For example, when the overlapping percentage between the frames is 50%, the Euclidean norms of fusion-ICP is 0.120 m, and it performs much better than the SIFT + ICP of 0.302 m, where the latter is considered to fail in aligning frames.

Data frame 1 (dark purple point cloud, marked as 1) and data frame 2 (red point cloud, marked as 2) in Fig. 8(a), respectively, are obtained at two positions and the two local data frames share no correspondence. Thus, 3-D mapping approaches using the data association algorithm based on sparse visual features, for example, the SIFT + ICP method, are invalid for aligning the frames. With fusion-ICP, frames were aligned using no data association but a 2-D platform pose [as shown in Fig. 8(b)].

V. Conclusion

A 3-D indoor mobile mapping system with fusion data from a 2-D laser scanner and an RGB-D camera has been introduced for mapping of GNSS-denied spaces. Accurate mobile platform positioning is achieved by laser scanner-based 2-D SLAM, and no inertial measurement unit is required. The proposed fusion-ICP method accelerates the alignment process and improves aligning accuracy, and it particularly deals with frame alignment when insufficient correspondence exists or nonoverlapping frames alignment. Experimental results showed the effectiveness of our system to build a 3-D global and accurate indoor map.

Acknowledgment

The authors would like to thank the associate editor and anonymous reviewers for their valuable comments to improve this letter.

References

[1] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Proc. 2nd Annu. Conf. Uncertainty Artif. Intell.*, 1986, pp. 435–461.
[2] P. Besl and N. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
[3] M. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. Robot. Autom.*, vol. 17, no. 3, pp. 229–241, Jun. 2001.
[4] S. Thrün, D. Fox, and W. Burgard, "A real-time algorithm for mobile robot mapping with application to multi robot and 3-D mapping," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2000, pp. 321–328.
[5] K. Pathak, N. Vaskevicius, J. Poppinga, M. Pfingsthorn, S. Schwertfeger, and A. Birk, "Fast 3D mapping by matching planes extracted from range sensor point-clouds," in *Proc. IEEE/RSJ Int. Conf. Intell. Rob. Syst.*, Oct. 2009, pp. 1150–1155.
[6] B. D. Bouvrie, "Improving RGBD Indoor Mapping with IMU data," M.S. thesis, Dept. Softw. Technol., Delft Univ. Technol., Delt, The Netherlands, 2011.
[7] H. Andreasson and T. Stoyanov, "Real time registration of RGB-D data using local visual features and 3D-NDT registration," presented at the Semantic Perception, Mapping and Exploration Workshop of IEEE Int. Conf. Robot. Autom., May 2012.
[8] [Online]. Available: http://www.xbox.com/en-US/kinect and http://www.primesense.com/
[9] K. Xu, L. Qin, and L. Yang, "RGB-D fusion toward accurate 3D mapping," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2011, pp. 2618–2622.
[10] J. Stuckler and S. Behnke, "Integrating depth and color cues for dense multi-resolution scene mapping using RGB-D cameras," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, Sep. 2012, pp. 162–167.
[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Computer Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[12] 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," *Int. J. Robot. Res.*, vol. 31, no. 5, pp. 647–663, Apr. 2012.
[13] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An evaluation of the rgb-d slam system," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 1691–1696.
[14] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robot.: Sci. Syst.*, Jun. 2009, vol. 25, pp. 26–27.
[15] J. Du, Y. Ou, and W. Sheng, "Improving 3D indoor mapping with motion data," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2012, pp. 489–494.
[16] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Trans. Robot. Autom.*, vol. 23, no. 1, pp. 34–46, Feb. 2007.
[17] R. Kuemmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A General framework for graph optimization," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3607–3613.
[18] Z. C. Marton, R. B. Rusu, and M. Beetz, "On fast surface reconstruction methods for large and noisy datasets," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 3218–3223.