Feature-based RGB-D SLAM with Dense Terrain Mapping for a Walking Robot

Dominik Belter, Piotr Skrzypczyński
Institute of Control and Information Engineering, Poznań University of Technology
ul. Piotrowo 3A PL-60-965 Poznań, Poland
Email: {name.surname}@put.poznan.pl

I. Introduction

Although advanced motion planning algorithms allow legged robots to traverse autonomously challenging terrains [5], in order to achieve full autonomy these robots need to have both environment mapping and self-localization implemented on-board. There are two issues that make the Simultaneous Localization and Mapping (SLAM) problem particularly hard if a legged robotic platform in unstructured environment is considered: (i) the environment maps employed by typical SLAM algorithms rarely are suitable for supporting motion planning functions (foothold selection, leg trajectories planning, etc.), (ii) legged robots that plan their motion require very accurate pose estimates. Our earlier experiments demonstrated [2] that it is necessary to achieve localization accuracy of about the diameter of our robot's foot to register the range data into a local elevation map that enables reliable foothold planning. Moreover, an autonomous walking robot requires a dense representation of the terrain for planning. While some state-of-the-art 3-D SLAM systems build voxel-based maps, they usually rely on massive parallel processing to handle the huge amount of range data [13]. Such approach is not suitable for most legged robots, due to the power consumption of a high-end GPGPU. Therefore, an autonomous legged robot needs a SLAM algorithm that is very accurate, runs in real-time without hardware acceleration and is coupled with a dense mapping method that can reflect the uncertainty of both the range measurements and the robot's pose estimate.

In the workshop presentation we demonstrate the algorithms and techniques we used to create a system that meets the aforementioned requirements and is implemented on our autonomous walking robot Messor II. Moreover, we discuss the methods used to evaluate the system, in both the qualitative and quantitative aspects. We present results of experiments in controlled and natural environments and point out the need to create benchmark data sets appropriate for evaluation of SLAM/mapping in challenging environments.

II. System Architecture

We introduce the PUT SLAM system, which is a SLAM solution based on the factor graph optimization approach and a persistent, scalable map of 3-D point features. This system follows the non-linear optimization approach to state estimation, which is currently considered to be the state-of-the-art method in SLAM. PUT SLAM implements a number of novel ideas, such like the fast visual odometry subsystem [3] that predicts the sensor (robot) pose with respect to the large map of features, thus enabling more robust feature matching. The SLAM system allows the RGB-D data processing front-end and the optimization back-end to work fully in parallel, thus achieving real-time performance on a multi-thread CPU with no GPGPU. The back-end is based on the g^2o general graph optimization library [9].

This research is financed by the National Science Centre under decision DEC-2013/09/B/ST7/01583

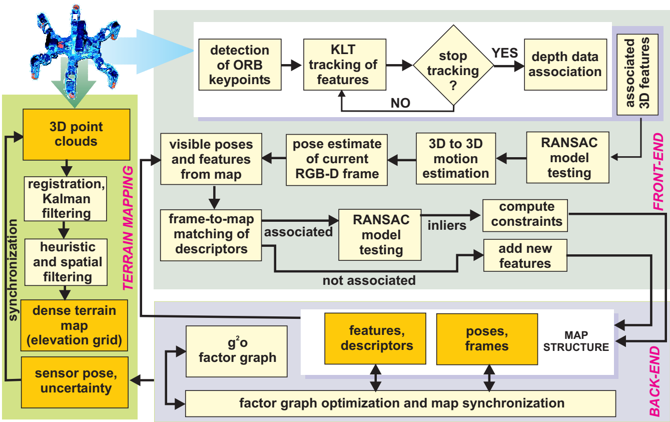

Fig. 1. Block scheme of the PUT SLAM with dense mapping system

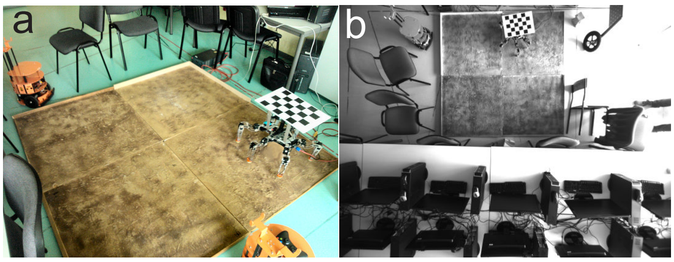

Fig. 2. Messor II robot with Asus Xtion and the chessboard marker (a), and a view of the experimental setup from an overhead camera

The SLAM system is coupled with a dense terrain mapping component. This component is based upon our earlier elevation mapping algorithm [1] that models the grid map as a Markov Random Field of order 0, where the state of each individual cell can be estimated as an independent random variable. Therefore, we use a scalar Kalman filter at each cell to efficiently handle the sparse and uncertain range data when

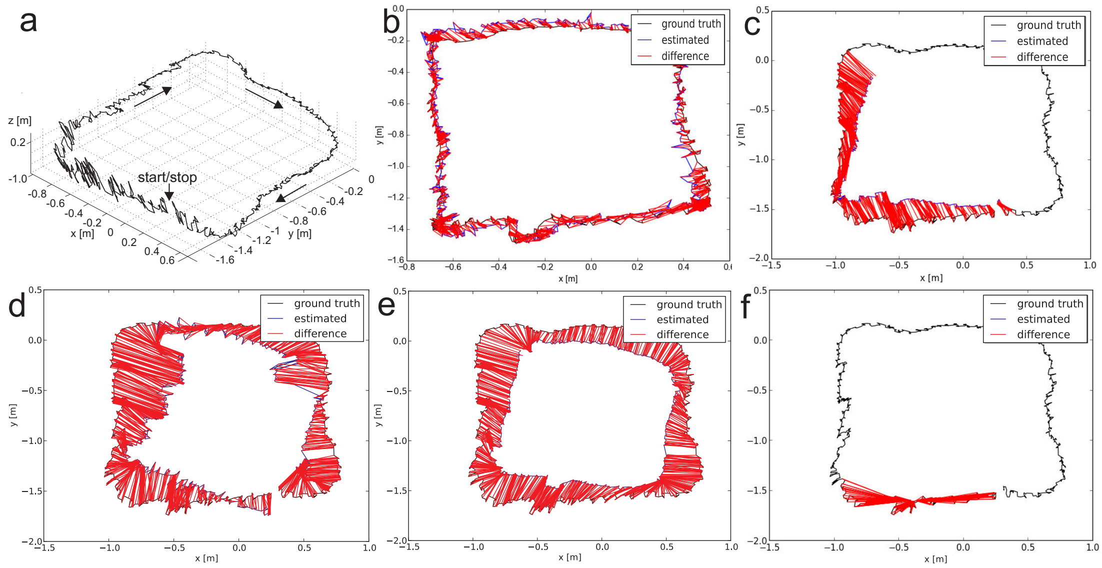

Fig. 3. Trajectories for an exemplary Messor II sequence: ground truth (a), PUT SLAM (b), ORB-SLAM2 (c), CCNY_RGBD (d), RGB-D SLAM v2 (e), KinFu LS (f)

we update the elevation values. Moreover, we implement a number of heuristics (e.g. line-of-sight visibility checking) to cope with spurious measurements. When we have a robot pose estimate from the PUT SLAM system we can also take into account the sensor location uncertainty, in a way similar to the one proposed in [8]. This allows us to reduce the drift of the local maps with respect to the fixed global coordinate system.

III. Experimental Evaluation

The proposed SLAM and terrain mapping approach is currently investigated and evaluated using various data sets and robotic platforms. The PUT SLAM was tested on publicly available data sets in order to ensure that our results are verifiable. On the TUM RGB-D benchmark [12] we achieved the sensor trajectory recovery accuracy that is equal or even better than the best results published so far [6]. However, the available benchmark data sets do not allow us to test the system under conditions that are specific to walking robots, such like limited field of view of the sensor and vibrations of the platform. Therefore, we tested various versions of our system on the real legged machine Messor II with an Asus Xtion PRO Live RGB-D sensor. To obtain quantitative results demonstrating the accuracy of PUT SLAM we have performed a series of experiments on a small terrain mockup (Fig. 2a). The ground truth trajectories were obtained using the PUT Ground Truth system, which uses ceiling-mounted, high-resolution passive cameras (Fig. 2b). The Xtion RGB-D frames were registered synchronously with the overhead camera images at the frequency of 15 Hz, due to frame rate limit of the cameras [11]. Evaluation of the SLAM systems is based on the Absolute Trajectory Error (ATE) and Relative Pose Error (RPE) metrics [12].

The ATE compares the distance between the estimated and ground truth trajectories, and is computed from the Root Mean Square Error (RMSE) for all nodes of the ground truth

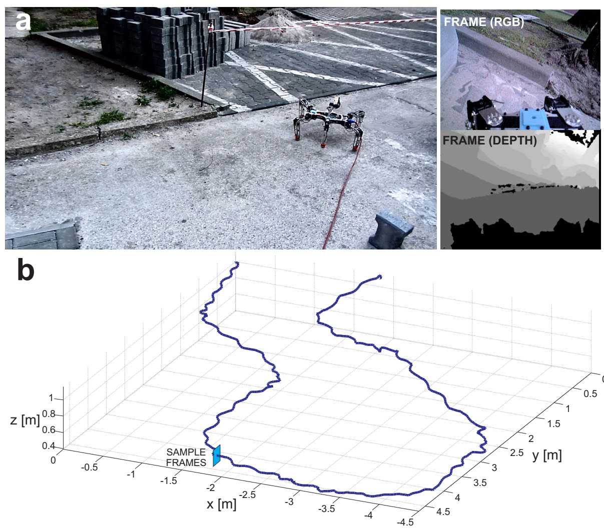

Fig. 4. Outdoor experiment on rough terrain – semi-loop trajectory: example view of the site with the Messor II robot and sample RGB-D frames (a), trajectory recovered by the PUT SLAM system (b)

and the estimated trajectory, whereas RPE corresponds to the local drift of the trajectory. Figure 3 demonstrates exemplary results we obtained in these experiments. The Messor II used its default tripod gait at 45% of the maximum speed. The trajectory resembled a rectangle of the lap size equal to 2 m (Fig. 3a). Using PUT SLAM we got a reasonably precise trajectory estimate (Fig. 3b), whereas other open-source SLAM systems we have tested on the data set obtained with the Messor II robot produced much worse results, like the Kalman-filter-based CCNY_RGBD and the pose-graph-based RGB-D SLAM v2 [7]. The monocular ORB-SLAM2 [10], which yields very precise pose estimates on typical data sets was unable to complete the experiment loosing tracking on blurred RGB images. Also the KinFu Large Scale (LS) software, an open-source re-implementation of the Kintinuous algorithm [13] failed on all the trajectories registered from the walking robot, apparently due to problems caused by fast rotations and too low frame rate of the depth images.

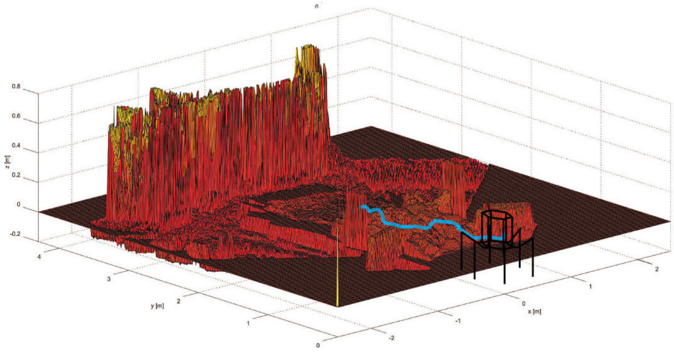

Fig. 5. Example of an elevation map built on-line from Asus Xtion data and used for motion planning

We have also tested the qualitative behavior of the PUT SLAM system in conditions that resembled a real construction site (under proper lighting conditions), as shown in Fig. 4 [4]. The mapping component was tested on the real robot as well, indoors (Fig. 5) and outdoors [5].

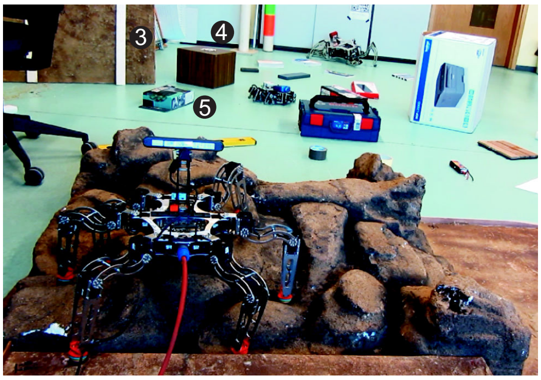

Fig. 6. Terrain mapping experiment on a rocky terrain mockup

However, in this case the problem is the lack of publicly available benchmarks that provide both trajectory and dense environment map ground truth. We consider development of such a benchmark data set as necessary to foster research on fully autonomous ground robots. Without the map and trajectory ground truth we can test the integration of the terrain mapping subsystem in PUT SLAM only quantitatively, as demonstrated in Fig. 6 and Fig. 7. The precise pose estimate obtained from the feature-based RGB-D SLAM allows the elevation map to correctly represent objects, as seen in Fig. 7b, whereas the map built using only the pose estimate from proprioceptive sensing (Fig. 7a) is blurred.

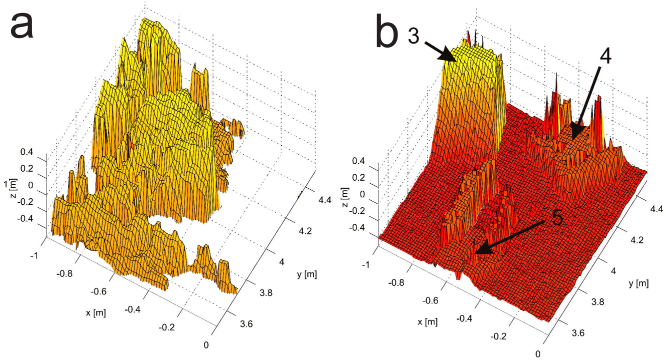

Fig. 7. Elevation maps obtained using proprioceptive sensing (a), and pose estimates obtained from the RGB-D SLAM (b)
. Numbers indicate objects seen in Fig. 6

References

[1] D. Belter, P. Łabęcki, P. Skrzypczyński, "Estimating terrain elevation maps from sparse and uncertain multi-sensor data", Proc. IEEE Int. Conf. on Robotics and Biomimetics, Guangzhou, 2012, 715-722.

[2] D. Belter, P. Skrzypczyński, "Precise self-localization of a walking robot on rough terrain using parallel tracking and mapping", Industrial Robot: An International Journal, 40(3), 2013, 229–237.

[3] D. Belter, M. Nowicki, P. Skrzypczyński, "On the performance of pose-based RGB-D visual navigation systems", Computer Vision – ACCV 2014 (D. Cremers et al., eds.), LNCS Vol. 9004, Springer, 2015, 407–423.

[4] D. Belter, M. Nowicki, P. Skrzypczyński, K. Walas, J. Wietrzykowski, "Lightweight RGB-D SLAM system for search and rescue robots", Progress in Automation, Robotics and Measuring Techniques, (R. Szewczyk, et al., eds.), AISC 351, Springer, 2015, 11–21.

[5] D. Belter, P. Łabęcki, P. Skrzypczyński, "Adaptive motion planning for autonomous rough terrain traversal with a walking robot", Journal of Field Robotics, 33(3), 2016, 337–370.

[6] D. Belter, M. Nowicki, P. Skrzypczyński, "Improving accuracy of feature-based RGB-D SLAM by modeling spatial uncertainty of point features", Proc. IEEE Int. Conf. on Robotics & Automation, Stockholm, 2016, 1279–1284.

[7] D. Belter, A, Kostusiak, M. Nowicki, P. Skrzypczyński, "Real-time SLAM from RGB-D data on a legged robot: An experimental study", Advances in Cooperative Robotics (M. O. Tokhi, G. S. Virk, eds.), Singapore, World-Scientific, 2016, 320–328.

[8] P. Fankhauser, M. Bloesch, C. Gehring, M. Hutter, R. Siegwart, "Robot-centric elevation mapping with uncertainty estimates", Mobile Service Robotics (K. Kozlowski et al., eds.), World-Scientific, 2014.

[9] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, W. Burgard, "g2o: A general framework for graph optimization", Proc. IEEE Int. Conf. on Robotics & Automation, Shanghai, 2011, 3607–3613.

[10] R. Mur-Artal, J. M. M. Montiel, J. D. Tardós, "ORB-SLAM: A versatile and accurate monocular SLAM system", IEEE Trans. on Robotics, 31(5), 2015, 1147–1163.

[11] A. Schmidt, A. Kasinski, M. Kraft, M. Fularz, Z. Domagala, "Calibration of the multi-camera registration system for visual navigation benchmarking", Int. Journal Advanced Robotic Systems, 11(83), 2014.

[12] J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems", Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems,Vilamoura, 2012, 573–580.

[13] T. Whelan, H. Johannsson, M. Kaess, J. J. Leonard, J. B. McDonald, "Robust real-time visual odometry for dense RGB-D mapping", Proc. IEEE Int. Conf. on Robot. and Automat., Karlsruhe, 2013, 5704–5711.