Robotics and Autonomous Systems 84 (2016) 88–96

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Improving multimodal data fusion for mobile robots by trajectory smoothing

Vladimír Kubelka [a,b,*], Michal Reinstein [a], Tomáš Svoboda [a,b]

[a] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic
[b] Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Czech Republic

HIGHLIGHTS

- The problem of fusing sensor modalities with significantly different sampling rates in a mobile robot localization system is addressed.
- A heuristic approach to include a low-rate position increment modality is proposed.
- The proposed approach is grounded with respect to a standard Rauch–Tung–Striebel smoother for the Kalman filter.
- Performance of the proposed approach is experimentally evaluated and selected fail-cases are discussed.

ARTICLE INFO

Article history:
Received 13 July 2015
Received in revised form
19 May 2016
Accepted 25 July 2016
Available online 2 August 2016

Keywords:
Field robots
Sensor fusion
Search and rescue robots

ABSTRACT

Localization of mobile robots is still an important topic, especially in case of dynamically changing, complex environments such as in Urban Search & Rescue (USAR). In this paper we aim for improving the reliability and precision of localization of our multimodal data fusion algorithm. Multimodal data fusion requires resolving several issues such as significantly different sampling frequencies of the individual modalities. We compare our proposed solution with the well-proven and popular Rauch–Tung–Striebel smoother for the Extended Kalman filter. Furthermore, we improve the precision of our data fusion by incorporating scale estimation for the visual modality.

1. Introduction

For successful deployment of mobile robots to complex dynamically changing environments, such as those typical for Urban Search & Rescue (USAR), reliable localization is crucial. In modern mobile robots, a popular solution lies in the combination of proprioceptive sensors, usually in form of an integrated Inertial Navigation System (INS), that captures the body dynamics at high rate, and an external source of aiding, using either vision [1] or range measurements [2]. Since most of the solutions are based on the well-proven Extended Kalman filter (EKF) [1,2], the state estimation architecture we designed for our platform (see Fig. 1) is based on the error state EKF framework as well.[1]

As we have shown in [3] (results of this work are summarized in Sections 3.1 and 3.2), performing data fusion of various modalities – such as in our case the inertial data, track odometry, visual odometry, and laser-based mapping – provides satisfactory results even when exposed to harsh environmental conditions, which can cause some of the modalities to fail. There is a number of well known problems connected with each named modality. First, the track odometry is strongly susceptible to high slippage, especially in skid-steer robots such as ours [4]. Second, it is the drift of the inertial sensors caused primarily by integrating the sensor noise, misalignment and instrumental errors. Third, the sensitivity to illumination and lack of scene texture influence the visual odometry performance [5]. And fourth, the laser-based mapping is sensitive to dynamic changes and to the overall geometric structure of the environment [6,7]. We addressed all these issues in [3] and introduced a failure-case methodology for evaluation of our multimodal data fusion. In this methodology we invoke challenging conditions that cause different modalities to fail on purpose and hence allow us to properly evaluate the robustness of localization.

However, our currently published results [3] raised a question that motivated us into a more in-depth research of the critical

* Corresponding author at: Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Czech Republic.
E-mail addresses: kubelvla@fel.cvut.cz, kubelvla@fel.cvut.cz (V. Kubelka), reinstein.michal@fel.cvut.cz (M. Reinstein), svobodat@fel.cvut.cz (T. Svoboda).

[1] TRADR: Long-Term Human–Robot Teaming for Robot-Assisted Disaster Response www.tradr-project.eu.

http://dx.doi.org/10.1016/j.robot.2016.07.006
0921-8890/

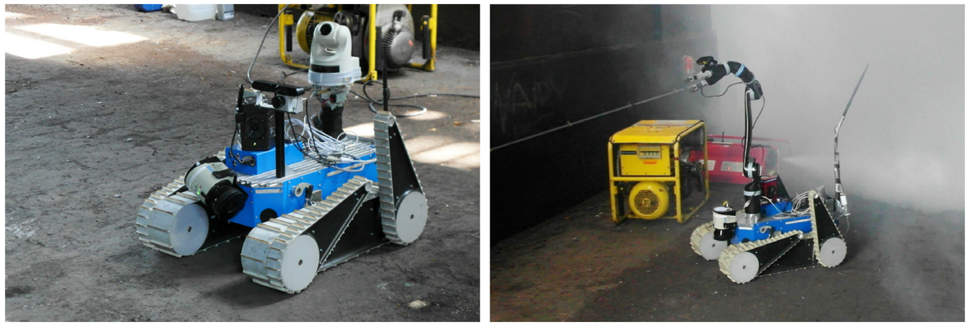

Fig. 1. TRADR[1] robotic platform at the USAR training site of the fire brigade of Dortmund, Germany. Robot sensor suite is highly configurable, includes Point Grey Ladybug3 omni-directional camera and SICK LMS-151 laser range-finder, can be extended further by various cameras (left) or a robotic arm (right).

issue of significantly different sampling frequencies of individual modalities. This is in general relevant to all multimodal data fusion algorithms, regardless a platform type. In this paper we therefore present our most recent results and compare them to the popular and commonly used standard smoother for Kalman filters. We have chosen the Rauch–Tung–Striebel smoother (RTS) [8] for the EKF as a representative of the Kalman smoothers family. By its definition, RTS best fits our data fusion scenario since it allows to recompute past position estimates based on information introduced by low-rate position increment measurements. We therefore exploited the RTS as benchmark for our multimodal data fusion architecture. For this purpose, we exploit our multimodal dataset[2] [3], which includes precise ground truth for both position and orientation obtained using a Vicon tracking system.

Our contribution is twofold. Due to experimental comparison and analysis, we were able to ground our novel approach to fusing multiple modalities at significantly different sampling rates with respect to the RTS smoother for EKF (described in Section 3.4). We hence offer our solution as an alternative to this popular RTS smoother, whether intended for robotics application or multimodal data fusion in general. Secondarily, with respect to our previous results, we improved the multimodal data fusion by incorporating velocity information from the visual odometry and resolved the scale problem for processing panoramic images (Section 3.3).

The paper is structured as follows: Section 2 introduces the related work, Section 3 sums up our previous work, describes our new proposed solutions and presents them in the context of smoothers for the EKF. Section 4 summarizes the experimental evaluation and Section 5 concludes the implications of our work.

2. Related work

Regarding the multi-modal data fusion, we built on our previous results described in [3], especially the design of the EKF error models [9–11]—even though the later work concerned a legged robot.

If long-term reliability and good accuracy are required, dead-reckoning solutions – such as those based on IMU and odometry – need other exteroceptive aiding modalities. In [12] it is shown that an IMU based dead reckoning system can be realized and successfully combined with the visual odometry to produce a reliable navigation system. We include visual odometry measurements into the EKF fusion scheme as well, yet directly in a form of angular and translational velocities computed by a more general implementation of visual odometry [5] designed for an omni-directional camera (note that in [12], the problem of tracking visual features is simplified by using a marker for planar homography).

Besides the visual odometry, another typical sensor for aiding is the laser range-finder. The laser range-finders are usually used for estimating vehicle motion by matching consecutive laser scans and thus creating a metric map of the environment [6,7]. Examples of successful deployment can be found for indoor – without IMU but combined with vision [13] – as well as for outdoor—relying only on the IMU [2]. The most popular approach of scan matching is based on the Iterative Closest Point (ICP) algorithm, which was first proposed by [14,15]. Later, [16] proposed a 6D Simultaneous Localization and Mapping (SLAM) system relying primarily on the ICP. Work of [17] proposed a localization system combining a 2D laser SLAM with a 3D IMU/odometry-based navigation subsystem. Contrary to the later publications realized in the context of SLAM, we only consider the output of the ICP algorithm[3] as a local pose measurements—similarly as with the visual odometry, we treat the laser localization module as a *velocity sensor*.

Solutions exploiting the EKF for fusing the dead reckoning with exteroceptive sensors are very popular [1,2,19–21]; our fusion scheme is based on the EKF as well. Still, a number of problems arise in multimodal data fusion. The problem of utilizing several sensors for localization, which may provide contradictory measurements, is discussed in [22]. The authors use Bayes filters to estimate sensor measurement uncertainty and hence evaluate the sensor validity. We separately addressed this problem in [10], where we utilized machine-learning techniques to detect anomalous measurements.

Since we aim for grounding our approach with respect to the smoothers for Kalman filters (in order to smooth the trajectory estimates), we have chosen the well established RTS [8] for the EKF as our base reference for benchmarking. Smoothers like RTS are well proven in the context of localization. In [23] a network of time-of-flight *Cricket* sensors provide measurements with a slight delay; the authors utilize an interacting multiple model fixed lag smoother to incorporate these delayed measurements. In [24] indoor localization problem is used to demonstrate properties of a smoother for the Unscented Kalman Filter. And finally in [25], the RTS smoother is actually utilized for the SLAM problem. Smoothing in Kalman filtering can be applied to wide range of problems, e.g. work of [26] applies the RTS to improve state estimation of a dynamic power system.

Several modifications of the RTS smoother have been proposed [27–29] since its original publication [8]. They mainly aim on better numerical stability and performance of the filter when deployed on computers with limited precision of number representation. We compare our algorithm to the RTS smoother using

[2] This dataset has been already released to the robotics community at https://sites.google.com/site/kubelvla/public-datasets/nifti-zurich-2013.

[3] We use the *libpointmatcher* implementation of the ICP algorithm [18].