Conference Paper

Collaborative Navigation for Flying and Walking Robots

Author(s):
Fankhauser, Péter; Blöesch, Michael; Krüsi, Philipp; Diethelm, Remo; Wermelinger, Martin; Schneider, Thomas; Dymczyk, Marcin; Hutter, Marco; Siegwart, Roland

Publication Date:
2016

Permanent Link:
https://doi.org/10.3929/ethz-a-010687710

Originally published in:
http://doi.org/10.1109/IROS.2016.7759443

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Collaborative Navigation for Flying and Walking Robots

Péter Fankhauser\textsuperscript{1,2}, Michael Bloesch\textsuperscript{1}, Philipp Krüssi\textsuperscript{1}, Remo Diethelm\textsuperscript{1}, Martin Wermelinger\textsuperscript{2}, Thomas Schneider\textsuperscript{1}, Marcin Dymczyk\textsuperscript{1}, Marco Hutter\textsuperscript{2}, Roland Siegwart\textsuperscript{1}

Abstract—Flying and walking robots can use their complementary features in terms of viewpoint and payload capability to the best in a heterogeneous team. To this end, we present our online collaborative navigation framework for unknown and challenging terrain. The method leverages the flying robot’s onboard monocular camera to create both a map of visual features for simultaneous localization and mapping and a dense representation of the environment as an elevation map. This shared knowledge from the flying platform enables the walking robot to localize itself against the global map, and plan a global path to the goal by interpreting the elevation map in terms of traversability. While following the planned path, the absolute pose corrections are fused with the legged state estimation and the elevation map is continuously updated with distance measurements from an onboard laser range sensor. This allows the legged robot to safely navigate towards the goal while taking into account any changes in the environment. In this setup, our approach is independent of external localization, relative observations between the robots, and does not require an initial guess about the pose of the robots. The presented methods are fully integrated and we demonstrate their capabilities in an experiment with a hexacopter and a quadrupedal robot.

I. INTRODUCTION

Flying and ground vehicles, such as walking robots, possess complementary beneficial properties for autonomous navigation tasks. Flying robots can quickly cover large areas and have a favorable viewpoint for situational assessment. On the other hand, a ground vehicle has longer autonomy, can carry substantial payload, and can actively interact with the environment. Working with a heterogeneous team of flying and ground vehicles opens new possibilities to combine these unique features. In this paper, we present our work on a fully integrated approach on collaborative navigation for flying and walking robots in unknown and challenging terrain. We are interested in particular in legged robots, because of their superior capability to traverse challenging terrain which can be found for example in search and rescue scenarios. Our approach is to have the (faster) flying robot explore and map the environment (see Fig. 1), while the (slower) walking robot utilizes the shared data to localize itself and plan a global path to the goal position.

In our work, we introduce a range of methods which interact as follows: A flying robot (e.g. a hexacopter or similar) is guided to fly over an unknown area of interest in order to collect visual data with a monocular camera. The visual data is used to create a global map of visual features (landmarks) and simultaneously localize itself against it (SLAM). Furthermore, the collected monocular image stream is used to create depth images with help of a semi-dense reconstruction pipeline. These depth-images are used to generate a probabilistic elevation map representation of the terrain. The information about the terrain profile is leveraged by the walking robot to plan a global path to a desired goal position. This is achieved by interpreting the elevation map as a traversability map, judging the easiness/safety to walk over a certain region. While following the planned path with a walking gait, the pose of the robot relative to the global landmark map is tracked with an onboard camera. The absolute pose references and the high-rate data from joint states and Inertial Measurement Unit (IMU) are fused in the legged state estimator providing accurate, fast, and drift-free pose estimations for the locomotion controller. As the legged robot moves through the environment, the original elevation and traversability maps are continuously updated with help of a laser range sensor. The walking robot performs continuous replanning of the path to account for the new data acquired in the map. This allows the walking robot to navigate through a previously unknown environment while preventing collision with unforeseen obstacles.

To the best of our knowledge, the only work that focuses on a collaboration of a legged and a flying robot for navigation was discussed in [1]. However, it was limited to the application of receiving an alternative camera viewpoint by visually servoing a quadrotor with attached markers.

Prior work in collaboration between aerial and ground robots often addressed the navigation problem in large-scale outdoor environments [2, 3, 4, 5, 6]. These approaches rely on the (at least partial) availability of satellite-based localization, which precludes their application in indoor environments and GPS-denied outdoor areas such as dense

\textsuperscript{1} Autonomous Systems Lab (ASL), ETH Zurich (pfankhauser@ethz.ch)
\textsuperscript{2} Robotic Systems Lab (RSL), ETH Zurich

The research leading to these results has received funding from Google’s project Tango. This work was supported in part by the Swiss National Science Foundation (SNSF) through project 200021_149427 / 1, the National Centre of Competence in Research Robotics, and the armasuisse S+T UGV research program.
forests or narrow urban canyons. Alternatively, several approaches have been proposed for GPS-independent relative localization of aerial and ground robots. A flying robot is tracked using a sensor mounted on a ground robot \cite{7}, or vice versa \cite{8, 9}, while mapping is accomplished by only one of the agents. The two robots are required to stay close to each other at all times, since the tracked agent can only be localized as long as it is in the field-of-view of the other robot. In a navigation scenario, this prevents the aerial robot from exploring areas further away from the ground robot, which is essential for efficient global guidance. Moreover, tracking-based systems may not easily be scalable to groups of more than two robots.

Our localization and mapping system is independent of both external localization (such as GPS) and relative observations between the robots. In contrast, the robots build and localize within a single map of visual features using on-board cameras. Hence, the system is readily extendable to groups of more than two robots. The approaches of Michael et al. \cite{10} and Forster et al. \cite{11} are conceptually similar to ours, yet based on geometric registration, which is known to be sensitive to initial guesses. We experimentally demonstrate that our feature-based visual localization and mapping framework enables reliable operation of both aerial and ground robots using a single map — despite the strongly varying viewpoints of the different agents, and without requiring prior information about the relative poses of the robots.

Moreover, unlike \cite{10} and \cite{11}, we integrate localization and mapping with a complete system for autonomous navigation of the ground robot, comprising collaborative terrain mapping using both cameras and laser scanners (multi-modal mapping), as well as traversability estimation, path planning, and motion control. To this end, we also integrate the high-level localization and mapping framework into our high-bandwidth but drift-affected Extended Kalman Filter-based (EKF) state estimation fusing kinematics and inertial measurements. Together, this provides the control and planning framework with an accurate, drift-free, global state estimation, with sufficient bandwidth such that the legged robot can stabilize itself and precisely navigate through the environment.

In this work, we contribute with the design and implementation of a navigation framework for ground robots in rough terrain that is enhanced by the help of flying robots. The framework is designed for a collaborative navigation in the sense that data is actively shared by the robots in order to increase the planning breadth and thus enhance the capability to navigate in previously unknown terrain. The proposed methods comprise all aspects of navigation such as localization, mapping, map interpretation, path planning, state estimation, and control in a tightly coupled manner. To this end, we present an experimental evaluation showcasing the advantages of our approach.

The remainder of the paper is structured as follows. Section II introduces the different methods applied in the navigation framework and describes their interrelation. Section III presents the experimental setup and evaluation of the proposed framework. In Section IV, we extend the sequential deployment of the flying and walking robot with a concurrent operation of the platforms. Section V concludes our work.
frame of reference for local SLAM landmarks. Only a fixed-number of keyframes is kept within the optimization window to keep the computational complexity tractable. Local SLAM landmarks get marginalized once the corresponding keyframe leaves the temporal window. This approach enables accurate visual-inertial localization with respect to a global reference frame. At each localization step, the SLAM landmarks are queried in the database of the global map landmarks to establish the feature matches. This means that no initial guess about the pose of the walking robot is required.

B. Legged State Estimation

Locomotion with force-controlled legged robots requires permanent motion stabilization using an on-board feedback controller which computes corrective joint torques based on the current estimation of the robot’s pose and velocity (e.g. [12]). The discussed localization and mapping framework provides an accurate pose estimation (see [14] for an evaluation), but the obtained bandwidth (5 Hz) is insufficient for the feedback controller for stabilizing the system.

In order to overcome the above shortcoming, an EKF-based state estimator is implemented on the legged robot as presented in [17]. This filter fuses the measurements from the dedicated on-board IMU with data it receives from the joint encoders at a frequency of 400 Hz in order to guarantee a high bandwidth feedback. While the local egomotion can be estimated accurately, this filter is prone to drift in the position and the yaw angle (due to the IMU roll and pitch are fully observable) [17].

To overcome this limitation, we extend the existing on-board legged state estimation by enabling the incorporation of external 6 Degree of Freedom (DoF) pose measurements between an arbitrary inertial coordinate frame (not necessarily gravity aligned) and a body-fixed coordinate frame. The motivation is twofold. First, the additional pose measurements improve the quality of the legged state estimator by limiting the drift of the unobservable states and by facilitating the estimation of online calibration parameters like IMU biases. Second, it provides the controller and planner processes with a unified, high accuracy, high bandwidth, gravity aligned estimate of the robot pose and its derivatives.

Directly fusing the external 6 DoF pose measurements with the internally estimated pose can lead to inconsistencies, especially if the frame of reference of the external pose is not properly aligned with gravity. Thus, we decouple the system by introducing two inertial frames, one for the ‘internal’ fusion of kinematics and IMU which is always aligned with gravity, and one for the reference for the ‘external’ pose measurements. The filter state is then augmented with the transformation between both inertial frames, which is modeled as a random walk process with small covariance parameter. This so-called loose coupling of both frameworks yields a state estimation which exhibits a high accuracy and drift-free localization with respect to the global map while ensuring a sufficiently high bandwidth. All derivatives are always represented w.r.t. the body frame (robo-centric formulation) and taken w.r.t. the ‘internal’ odometry frame in order to avoid discontinuities.

C. Elevation Mapping

We choose an elevation map representation to approximate the geometry of the terrain. In this model, a two-dimensional regular grid covers the mission area and stores for each cell the height of the terrain at that position. Although different terrain representations exist, elevation maps have the advantage that they can accurately capture the terrain profile while allowing for simple and efficient data handling.

The elevation map is generated by populating the grid cells with range measurements and fusing new measurements
D. Traversability Estimation

Finding a safe path for the walking robot through the environment requires to judge the suitability of the terrain for safe locomotion. To this end, we interpret the generated elevation map in terms of traversability, a local measure expressing the safety to pass the terrain at different locations. The range measurements originate either from estimating the image depth from the camera motion of the flying robot, or directly from the distance measurements from the rotating laser range sensor on the walking robot. In both cases, we are interested in the terrain profile below the robot and remove all distance measurements which are above the robot’s current base.

For the flying robot, we employed a semi-dense reconstruction pipeline. Assuming an accurate estimation of the camera trajectory, the pipeline performs a patch warping and alignment between selected camera frames and thereby estimates the corresponding depth. Given a frame of interest, a matching frame is heuristically selected in order to allow a good estimation of the disparity. The baseline of the selected camera frames is adapted to the mean scene distance which improves depth estimation for varying flying altitudes of the robot.

For all data generated by the flying and the walking robot, we use the elevation mapping method presented in [18] to probabilistically fuse multiple measurements that fall into the same cell. To this end, we extend the elevation map representation with a height variance information for each cell. The noise of the sensor measurements is approximated with the measurement errors models from [19] for the laser range sensor and an error model based on [20] for the depth estimation from the camera. With these models, the sensor noise in direction of the measurement and lateral direction is expressed as a quadratic function of the measurement distance. With the knowledge of the current pose of the robot, the sensor noise is propagated to the corresponding elevation measurement noise. A recursive filter fuses the received height measurement to estimate the height and variance for each cell. We use the Mahalanobis distance to reject measurements that fall below the highest elevation measurement. This is important to correctly reflect the height of objects such as boxes and tables, where multiple heights fall into the same cell. Note that because in this work the robots are able to localize themselves accurately and drift-free, we do not have to propagate the robot pose uncertainty into the elevation map in contrast to [18].

E. Navigation Planning

We employ a sampling-based RRT* planner to find the shortest valid path from the walking robot’s start pose to the goal as presented in [22]. The planner searches for a trajectory of the x- and y-position of the robot’s footprint. Since a legged robot works omni-directionally, we ignore the yaw-orientation of the robot in the planning process and adapt the yaw-angle such that the robot walks forward towards the next node on the path. To guarantee a rotation-independent traversability of the terrain, we approximate the footprint conservatively with a circumscribed circle. When checking the connectivity between two nodes in the planning process, we check if all cells within the convex hull of the two circles of the footprint are traversable. We define the path cost between two nodes as the weighted sum of the length cost and the traversability cost. The cost length is computed as the Euclidean distance between the two nodes. The traversability cost is given by the inverse of the sum of the traversability values in the convex hull of the two circular footprints normalized with the path length. The weighting parameters in the summation of the total costs determine how much short paths over safe paths are preferred or vice versa.

Since the map is permanently updated, the planner is automatically restarted on a regular basis to update the plan from the current position. In the recurring planning steps, new information in the updated map is taken into account which makes sure that previously unmapped obstacles are avoided and potential shortcuts are taken.

F. Path Execution

The robot receives walking commands by iterating through the nodes of the planned path as the robots walks. A PID controller determines the desired velocity from the error between the current and the desired pose. The desired velocity contains the translational velocities in x- and y-direction and the yaw rotational velocity. For long distances, we prefer to walk forward and orient the robot towards the goal pose while for short distances, the controller is also allowed to step sideways. The commands are executed by the walking robot by placing its feet in a static walking gait corresponding to the desired velocity [23].

As the navigation planning updates the path on a regular basis, the planning time might cause the new path to start behind the current robot location. To tackle this issue, a simple algorithm iteratively removes the first pose on the path as long as the angle between its connections to the second pose and current robot pose is smaller than 90°.
III. EXPERIMENTAL EVALUATION

We have successfully verified and analyzed our proposed approach with multiple experiments involving the entire algorithmic pipeline. In the following, we describe our results for the mission shown in Fig. 3. The walking robot is deployed in a start area from which the direct line of sight to the desired goal position is blocked by a big obstacle. The only feasible path to the goal consists of a scaffolding with narrow passages (90 cm width) and a ramp which leads down to the floor. During the execution of the mission with the walking robot, a previously unmapped obstacle is put on the pathway and forces the walking robot to maneuver around it.

A. Localization and Mapping

For mapping the environment with the flying robot, we let the UAV fly multiple rounds over our experimental area to detect and track the visual landmarks (see Fig. 4a). The captured data are transmitted to the walking robot. With its higher computational capacity (Intel Core i7 4500U dual-core 1.8 GHz and 8 Gb RAM), the generation of the global localization map takes several seconds. After the exploration with the flying robot from the air, the walking robot is able to localize itself within the global map. We have tested different start poses of the legged robot without an initial guess for the localization. The proposed method was able to reliably localize even though the landmarks are observed from a different view point (see Fig. 4b) than originally recorded.

Once the robot starts walking, the localization against the global map continuously tracks the pose of the robot (localization runs in real-time with ~20 ms processing time per query). Figure 5 shows the successfully matched features of the walking robot as it moves through the global map. The pose correction updates to the legged state estimator helped to safely track the path on the scaffolding without causing any instabilities of the walking controller.

B. Elevation Mapping

Similar to the localization map, the data from the UAV are used to generate dense vision depth estimates which are fused in an elevation map on the computer of the walking robot (processing in approx. real-time). Figure 6a shows a snapshot of the elevation map creation during the exploration phase of the flying robot. The dense reconstruction from the camera of the flying robot delivers accurate depth measurements at a frame rate of 5 Hz. This enables to generate an elevation map with a grid size of 4 cm with only few holes. After the exploration flight, the entire area is well covered as elevation map (see Fig. 6b).

As the legged robot moves through the environment, the elevation map is continuously updated from the rotating laser range sensor as shown in Fig. 6c. The moved obstacle is successfully captured by the map updates (A in Fig. 6). Errors in the pose estimation of the walking robot can cause an offset between the previous map from the flying robot and new data from the walking robot. While this has a bigger effect on map regions further away from the walking robot (such as B in Fig. 6), these offsets diminish for close regions and the important area in front of the robot is represented with more detail in comparison to the original map (C in Fig. 6).

C. Traversability Estimation and Navigation Planning

Figure 7 shows the interpreted elevation map (after the full mission) as traversability map. We can observe that all safe areas including the ramp are correctly marked as traversable while the borders of the scaffolding are correctly interpreted as dangerous. Thanks to the high resolution of the map and the locally bounded interpretation of the traversability, the
navigation planner can successfully find a feasible path even in narrow regions.

The first plan based on the map from the exploration of the flying robot (no updates from the walking robot yet) is shown in Fig. 8a. The planner is able to find a short path while maintaining a minimal distance to untraversable regions. As the robot starts tracking the planned path, the planner continuously replans the path to take into account the refined map and a changing environment. Figure 8b shows how the updated plan successfully guides the robot around the previously unmapped obstacle. Furthermore, we can observe that our navigation controller approach smoothly guides the robot along the path without causing irregularities because of plan updates.

IV. CONCURRENT OPERATION OF WALKING AND FLYING ROBOTS

In the experimental evaluation of Section III, we have let the flying robot first fully explore the environment before sharing the data with the walking robot (sequential operation). In the following, we describe the main differences that have to be considered for an operation of the platforms at the same time (concurrent operation).

We enable a simultaneous exploration and navigation of the robots by updating the shared map data between the robots in a batch-based approach. After the flying robot has transferred the first batch of the (incomplete) map, the walking robot starts to localize against the global localization map and uses the provided terrain data to plan its path and start execution. As the operation continues, new map
data is provided by the flying robot and is merged with the existing data to extend the range in which the legged robot can operate within the global frame.

A. Localization and Mapping

As the flying robot operates, the environment is explored on an increasing scale. To maintain global consistency, we apply bundle-adjustment (see Section II-A) on the trajectory which was previously included in the last batch and the newly acquired trajectory data. We maintain a computationally scalable map by pruning features with the least visual information.

As the global map grows with each new batch, the legged robot is able to localize in more areas. Figure 9 shows an example of how the walking robot is able to localize with increased certainty as the map coverage and quality increases with each batch update A–D. We quantify the global localization quality as the number of feature correspondences between the global map and the features from the current view of the walking robot (vertical axis of Fig. 9). The lower limit for the number of features for a (long-term) reliable localization is 4, and for performance reasons we threshold the max. number at 25. Although some viewpoints of the walking robot contain large areas with very little visual features (e.g. at time 80 s and 155 s), the localization can still be performed reliably.

B. Legged State Estimation

With each update of the global map from the flying robot, the localization of the walking robot can vary. Figure 10 shows a top-down view on the trajectories for the localization with the different map batches A–D. For each update A–D, the resulting trajectory for the walking robot transits to the latest (and more accurate) localization. This means that at each update A–D the perceived robot’s position can ‘jump’ instantaneously. We decouple these discontinuities from the control framework with the loosely coupled approach described in Section II-B.

C. Mapping and Navigation Planning

The walking robot can utilize the map updates from the flying robot to increase the range and precision of the path planning. At each map update, empty cells in the global terrain map are filled in where new data is available, while existing cells are not changed. Where no data is yet provided from the flying robot and is not yet visible by the onboard sensors, our planner assumes a traversability with value 0.1. Figure 11 shows snapshots of the concurrent mapping and navigation planning process of the two platforms. The terrain and the walking robot’s position changes slightly with each localization and terrain map update from the flying robot. These changes are taken into account in our navigation planning method (Section II-E) which adapts the path locally to changes in the environment and replans the global path on a regular basis.

V. Conclusion

We have presented the design and methods for a collaborative localization and mapping framework for a heterogeneous team of flying and ground robots. The methods applied to the flying and ground robot tightly integrate with each other and allow the ground robot to navigate with significantly increased performance than without the help of the flying robot. We have experimentally validated our approach with a full integration with a hexacopter and a quadrupedal robot,
in which the hexacopter enables the quadrupedal robot to navigate in previously unknown, rough terrain in an environment of $\sim 9 \times 5 \text{m}$. Our implementation demonstrates a successful functional interaction of many complex navigation tasks, which is rarely demonstrated in other work. We have additionally shown that our framework naturally extends to concurrent operation of the robots.

In current work, the goal is to extend the presented localization method to a larger team of robots (e.g. [24]), but reliable wireless communication is still a major issue. In the future, we will evaluate our approach in more extended and complex environments and adapt our strategy where necessary. Furthermore, we are working on extending the rough terrain capabilities of our legged robot by taking foothold selection and whole body climbing maneuvers into account. Eventually, the legged robot could also be equipped with a landing platform, where the flying robot could rest and charge until its action is required.

Fig. 11: The terrain map and associated motion plans for different execution stages (A→D). From left to right the coverage of the semi-dense visual map is increasing (grey) as the hexacopter flies over the area (smooth black line). Simultaneously, the legged robot adds its own perception data (laser, blue) to the elevation map. The estimated traversability (brightness of the terrain) as well as the planned (colored coordinate frames) and executed path (zigzag black line) of the legged robot are depicted.


REFERENCES


