

A Fully-Integrated Sensing and Control System for High-Accuracy Mobile Robotic Building Construction

Abel Gawel*, Hermann Blum*, Johannes Pankert†, Koen Krämer†, Luca Bartolomei‡, Selen Ercan§
Farbod Farshidian†, Margarita Chli‡, Fabio Gramazio§, Roland Siegwart*, Marco Hutter† and Timothy Sandy†

Abstract— We present a fully-integrated sensing and control system which enables mobile manipulator robots to execute building tasks with millimeter-scale accuracy on building construction sites. The approach leverages multi-modal sensing capabilities for state estimation, tight integration with digital building models, and integrated trajectory planning and whole-body motion control. A novel method for high-accuracy localization updates relative to the known building structure is proposed. The approach is implemented on a real platform and tested under realistic construction conditions. We show that the system can achieve sub-cm end-effector positioning accuracy during fully autonomous operation using solely on-board sensing.

I. INTRODUCTION

The available workforce in the construction industry is decreasing, and safety regulations are rising, causing a strong demand for innovative solutions on construction sites, e.g., digitization of the construction process and utilization of robots. However, robotic fabrication is traditionally associated with automated assembly lines, where fixed positioning and constant conditions determine the role that the robot may undertake in the fabrication process. Unlike such stationary robotic facilities, construction sites are spatially complex and cluttered environments. A construction robot must localize in its environment reliably, while being able to work with conditions deviating from as-planned building information. While humans use established tools to achieve required mm accuracy in construction tasks, mobile robots usually achieve only cm relative accuracy using on-board sensing [1]. This is not sufficient for many construction tasks such as drilling of holes, and installation of fixtures. Moreover, as every building project is unique, a key challenge is to find strategies

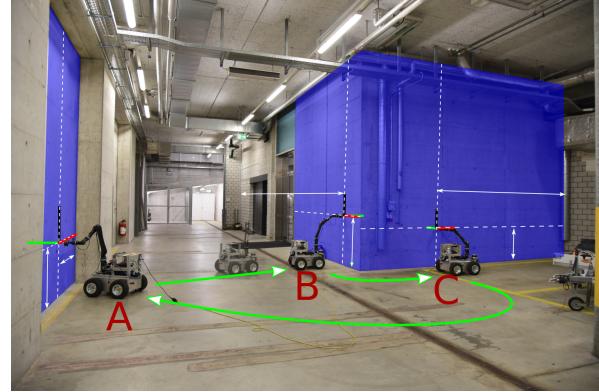


Fig. 1: The proposed system enables high-accuracy interaction of a mobile robot with the environment using only on-board sensing. Here, three task locations are highlighted between which a robot traversed in our experiments yielding sub-cm errors.

to cope with the inaccuracies of the building materials and components [1]. Robot world representations are usually referenced to one globally consistent map, or a number of sub-maps. Yet, the execution of building tasks primarily depends on the as-built status of existing installations, i.e., a set of local references and dependencies. Finally, motion planning and execution on construction sites is also a challenging problem. The typical kinematic redundancy of a mobile manipulator needs to be leveraged appropriately in order to coordinate locomotion and manipulation, while taking into account the physical constraints of the environment and uncertainties in the sensing [2].

While existing solutions rely on external sensing systems, or augmentation of the environment with specific markers to achieve the accuracy required in construction tasks, this work presents an autonomous mobile robotic solution for highly accurate building task execution, solely using on-board sensing. The system is based on the comprehensive integration of software solutions on a mobile platform to perform the workflow from extracting tasks from building plans to task execution. It automatically derives construction tasks from a 3D building model, enables the robot to localize within the model using LiDAR sensors, uses local references for highly-accurate localization, plans trajectories, dynamically avoids obstacles, and implements a whole-body motion planning regime distributing the movement between the mobile base and robotic manipulator. We evaluate the robot's performance towards discrete building tasks, e.g. drilling, under realistic conditions. This work presents the following contributions:

- A novel optimization-based method for high-accuracy

* Autonomous Systems Lab, ETH Zurich

† Robotic Systems Lab, ETH Zurich

‡ Vision for Robotics Lab, ETH Zurich

§ Gramazio-Kohler Research, ETH Zurich

Luca Bartolomei, Margarita Chli contributed to the motion planning {lbartolomei, mchli}@ethz.ch.

Hermann Blum, Abel Gawel, Roland Siegwart contributed to the state estimation and high accuracy localization {blumh, gawela, rsiegwart}@ethz.ch.

Timothy Sandy contributed to the state estimator, and motion control tsandy@ethz.ch.

Koen Krämer, Johannes Pankert, Farbod Farshidian, Marco Hutter contributed to the Motion planning and control {kokraeme, pankertj, farshidian, mahutter}@ethz.ch.

Selen Ercan, Fabio Gramazio contributed to the building task interface {ercan, gramazio}@arch.ethz.ch.

This work was partially supported by the Swiss National Science Foundation (SNF), within the National Centre of Competence in Research on Digital Fabrication and by the HILTI group.

robot localization using a set of plane intersections between laser distance sensor measurements and a 3D building model.

- LiDAR-based localization within a 3D building model using 3D scan matching.
- Whole-body model-predictive planning and control for end-effector pose tracking of a wheeled mobile manipulator.
- Experimental validation of the proposed system on a prototype robot in a realistic construction-like environment.

The remainder of this paper is organized as follows. In Section II, we review the related work on construction robotics, and related sensing solutions. Section III introduces the sensing and control system developed within this paper followed by an experimental evaluation of the system performance in Section IV, and concluding remarks in Section V.

II. RELATED WORK

Construction robotics and digital fabrication are emerging topics in the robotics research community [1]. Existing solutions towards construction automation range from fixed automation solutions [3] over innovative systems, e.g. using flying robots [4], to multi-purpose mobile ground robotic systems [4], [5], [6]. The proposed system is within the last category. Multi-purpose ground-robotic solutions are converging on typical design aspects, consisting of robotic mobile bases, manipulators, task-specific end-effectors, and sensing suite capabilities [5], [7], [8], [9], [6], [10]. Localizing such robots within the construction environment can be a difficult task, as perceptual conditions are challenging for cameras, and environments cannot easily be augmented with external localization markers or beacons, as used in [8], [7]. To enable automatic task generation from digital building models and enable robots to autonomously travel between them, it is further required to not only identify the working location, but also localize within the building model of the environment. [11] match corner-, and door plate-detection with cameras against 2D maps extracted from 2D floor plans. [12], [13] localize within floor plans abstracted to 2D grid maps using G-ICP. They further robustify their localization routine using a pose-graph SLAM system that enables the robot to account for largely cluttered environments as compared to the clean floor plan map. A solution using an initial reference scan of the environment using 3D LiDAR scans is presented in [10]. Similarly, [14] and [15] use scan matching and MCL in 2D using 2D LiDAR sensors. However, no domain shift in localization between environment model and LiDAR scans is performed, as the work location is inserted into the environment represented by the reference scan. Hence, the accuracy is limited by the reference scanner characteristics. This paper, however, presents a solution facilitating 3D building models for localization of 3D LiDAR scans. The routine performs ICP between LiDAR scans and the point-cloud sampled from a 3D CAD triangular mesh model.

Another key challenge for construction robots is high-accuracy end-effector positioning [1]. This requires highly

accurate knowledge about the end-effector position, e.g., via on-board sensing [8], [7], [6], [10] or external positioning systems [4], [9]. Additionally, the positioning of the end-effector and task execution require either stiff structure [6] or high control frequency [9]. Using 2D LiDAR sensors, it is possible to perform both the localization in the previously recorded model and high-accuracy localization using the same scan-matching routine [10], [14], [15]. Yet, high-accuracy on-board localization can also be realized with local visual markers and their detection using cameras [8], [7], reaching accuracies well below 1 cm. [14] and [15] report localization accuracy using scan matching and MCL below 1 cm in 2D using 2D LiDAR sensors in static environments, with an accuracy only mildly suffering under changed scenarios. [6] use 3D LiDAR scans to precisely localize in 3D against the known model of a work-piece (in contrast to localizing within the environment), given a localization prior, achieving an average accuracy of $\pm 5\text{mm}$.

In this work, we extend this concept facilitating the prior from 3D LiDAR localization, and inexpensive laser distance measurement sensors to perform local localization at mm scale via measurement intersections with the known building model.

Many of the previous works have considered executing construction tasks with mobile manipulators. Such robots must locate tools mounted on the manipulator's end-effector with high accuracy. In order to do so, robot motions are typically decomposed into two sequential steps of first repositioning the robot's base, then moving the arm to bring the tool to the desired location [10], [6], [9], [7]. To the best of our knowledge, this is the first work that adaptively plans and tracks whole-body motions for the high-accuracy building task execution. In doing so, our system can smoothly position its tool from optimal body postures.

III. SENSING AND CONTROL SYSTEM

In this section, we present the sensing and control system for high-accuracy mobile manipulation. The system covers the full range of functionalities to enable a mobile robot to receive tasks from a building model, autonomously localize within it, reach indicated task locations, and position an end-effector tool with mm-accuracy. The system solely relies on on-board sensing and requires no human intervention, however allows human intervention if necessary. Key software components are illustrated in Fig. 2.

A central role is given to the high-level manager that allocates tasks received from a building task interface to various software components. We consider building tasks which require the high-accuracy positioning of an end-effector-mounted tool at discrete locations in place, such as drilling, anchoring, and pick-and-place tasks. Expanding the system to support tasks which require either control of the tool's interaction force with the environment (e.g., chiseling) or the high-accuracy tracking of the tool along a continuous trajectory (e.g., plastering) is planned for future work. The mode of the system is switched depending on the distance from the current to the desired tool pose. The state

is estimated via a moving horizon estimator (MHE), fusing LiDAR localization updates, IMU measurements, and wheel odometry [16]. The estimated state is used in feedback by a whole-body motion controller which tracks base and end-effector reference trajectories. These references are generated by a Model Predictive Controller (MPC) strategy running at ~ 100 Hz. When the tool is far from the target pose, a stochastic navigation planner is used to find a collision-free trajectory to bring the robot near the target location. This base trajectory is used to seed the MPC. Once sufficiently close, the MPC iteratively plans and tracks a whole-body trajectory to complete the task. Finally, to reach highest accuracy, an additional localization routine is performed near task locations using the localization of the state estimator as an initial guess, refined with an independent sensor set.

A. State Estimation

State estimation of the robot's base-pose \mathbf{x}_{t_n} at time t_n is performed using the *ConFusion* [16] moving horizon estimator, minimizing an error over the various inputs in the least-square sense. Individual errors e_j are calculated using measurements \mathbf{z}_j connecting to one or more states $\mathbf{x}_{t_{i-j}:t_i}$, and static parameters \mathbf{s} , such as calibrations within the function \mathbf{h} , i.e.,

$$e_j = \mathbf{h}(\mathbf{x}_{t_{i-j}:t_i}, \mathbf{s}, \mathbf{z}_j) \quad (1)$$

We thus optimize for the states $\mathbf{x}_{t_o:t_n}$, and static parameters \mathbf{s} , forming an optimization problem over n state instances, i.e.,

$$\{\mathbf{x}_{t_o:t_n}, \mathbf{s}\} = \underset{\mathbf{x}_{t_o:t_n}, \mathbf{s}}{\operatorname{argmin}} \sum_{i=0}^n \sum_{\mathbf{z}_j \in \mathcal{Z}_{t_{i-j}:t_i}} \|e_j\|^2 \quad (2)$$

Different sensor modalities contribute error terms to the optimizer. Our system implements two types of state constraints, i.e., pose constraints ($j = 0$), and relative pose constraints ($j = 1$).

1) *3D LiDAR Localization*: We facilitate 3D LiDAR to localize the robot within the building model of the environment. The map \mathcal{M} is represented by a sampled point-cloud from the 3D CAD triangle mesh model M . Motion compensated 3D LiDAR scans are continuously matched against the model using ICP. ICP performs a separate minimization, of the error between new LiDAR points $\mathbf{p}_k \in \mathcal{L}_{t_n}$, and matched map points $\mathbf{q}_k \in \mathcal{M}$ with associated surface normals \mathbf{n}_k , facilitating a point-to-plane error metric, i.e.,

$$\hat{\mathbf{x}}_{t_n}^{ICP} = \underset{\hat{\mathbf{y}}_{t_n}}{\operatorname{argmin}} \sum_{k=0}^K \|(\mathbf{p}_k - \mathbf{q}_k)\mathbf{n}_k\| \quad (3)$$

The resulting robot base pose updates $\hat{\mathbf{x}}_{t_n}^{ICP} \subseteq \mathbf{x}_{t_n}$, are then sent to the state estimator, assuming that the pose updates are corrupted by a constant normally-distributed measurement noise. While we focus on localization within the building model in this work, the localization module is also able to build a map by incrementally fusing consecutively matched scans, and initialized at the starting location of the robot. An approximate initial guess within the environment is given by

the known starting location. A global localization scheme is outside the scope of this paper.

2) *Wheel Odometry*: Rotary encoders mounted on a robot's wheels are used to measure the angular velocity of the wheels. Considering a simple differential drive kinematic model, these measurements provide a process constraint on the evolution of the estimated base pose. The modeled confidence in this constraint is much lower when the wheels are turning than when they are stationary, since the simple kinematic model does not capture wheel slippage and non-flat ground profiles. This constraint therefore mostly helps enforce that the base remains stationary when the wheels are not turning.

3) *IMU*: Angular velocity and linear acceleration measurements from an inertial measurement unit (IMU) provide an additional process constraint on the evolution of the base pose. The linear velocity, gyroscope bias, and accelerometer biases are also estimated online as is typically done for inertial sensor fusion. The full model and conventions used are described in a previous work [17]. Because of their high rate and low latency, the IMU measurements are additionally used to compensate for the latency of the LiDAR pose updates. When estimates are received by the robot controller, they are forward propagated through the more recent IMU measurements up to the current controller time. Since the underlying MHE problem is solved using non-linear optimization, as opposed to more commonly used filter-based methods, fusion of the multiple process measurements does not require augmenting the state with additional terms, as in [18].

B. High Accuracy Localization (HAL)

Both, the measuring accuracy of the 3D LiDAR and discrepancies between the building model and the as-built conditions of the environment, limit the accuracy of the state estimate. For building task execution, the system therefore implements a HAL strategy against the task-relevant references, i.e., walls around the task space. Initialized with the state estimator's current pose, we localize the static robot base directly in the building model, represented as a 3D triangular mesh, using high accuracy laser distance sensors mounted at the end-effector.

This is done by optimizing the error between expected and measured distances of the laser sensors for several end-effector poses $\mathbf{y}_{t_n}^{(i)}$ (6 in our experiment, for ease of notation, the index (i) will be omitted in the following). At each iteration, we find the planes against which the laser distance sensors are measuring using ray-tracing in the building model from the current estimate of the sensor poses at the end-effector. Given the plane support \mathbf{t}_p and the plane normal \mathbf{r}_p of the intersected triangle in the mesh, we perform an optimization over the locally differentiable error function $e_{t_n}^{HAL}(\mathbf{y}_{t_n})$:

$$e_{t_n}^{HAL}(\mathbf{y}_{t_n}) = \sum_i c \left(\left\| z_i - \frac{(\mathbf{t}_p(\mathbf{y}_{t_n}) - \mathbf{y}_{t_n}^t) \mathbf{r}_p(\mathbf{y}_{t_n})}{\mathbf{y}_{t_n}^R \mathbf{r}_p(\mathbf{y}_{t_n})} \right\| \right)$$

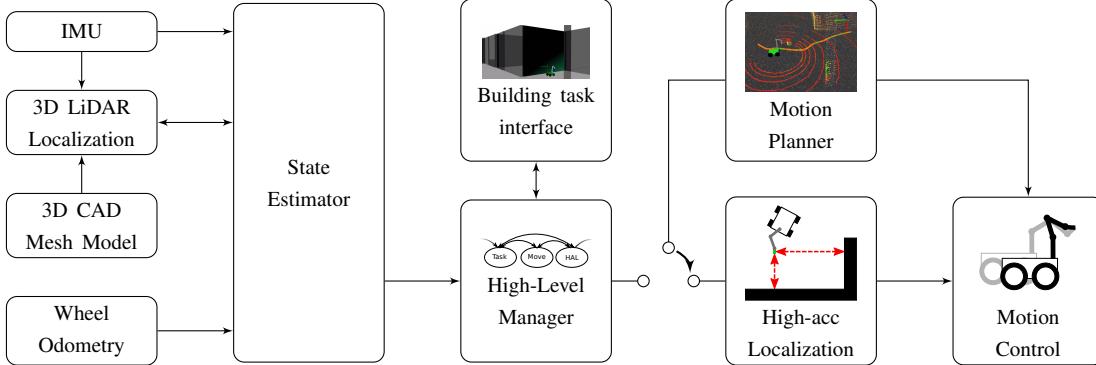


Fig. 2: Software system overview of the mobile construction robot: The building task interface provides task locations within the building model of the environment to the high-level robot manager. The high-level manager requests the current robot state fused from ICP-based localization within the building model, IMU measurements, and wheel odometry. Depending on the robot state, the manager switches between planning robot base trajectories to the task location, and a HAL procedure depending on the distance to the target. Once the robot is close to the target, a HAL procedure using laser distance sensors is triggered. The commands are executed using a whole body motion controller.

$$\mathbf{y}_{t_n}^{HAL} = \underset{\mathbf{y}_{t_n}}{\operatorname{argmin}} e_{t_n}^{HAL}(\mathbf{y}_{t_n})$$

with $\mathbf{y}_{t_n}^t$, and $\mathbf{y}_{t_n}^R$ the translational, and rotational components of \mathbf{y}_{t_n} respectively. To account for problematic measurements against clutter or windows and errors in plane retrieval due to initial pose uncertainty, we use a Cauchy robust cost function c . The relative transformations between the end-effector poses are assumed to be known by the manipulator's joint encoder readings. The procedure is formalized in Algorithm 1.

During the development of the system, we found that the estimate of the orientation with this routine was not as reliable and accurate as the initial estimate from aligning the much richer LiDAR scan with the building model. We therefore keep the orientation component fixed to the initial estimate and only optimize over the position. Doing so further increases the precision of the corrected position.

These local pose corrections $\mathbf{y}_{t_n}^{HAL}$ are only used for the local task execution, and not part of the general state estimation framework.

C. Building Task Interface

The building task interface bridges the gap between the design and planning environment containing the complex building information, and the robotic system. On one side, this serves as a suitable interface for users of the robotic system facilitating a building process that allows simple task-level commanding. On the other side, it seamlessly integrates the workflow between building construction and the robotic system for executing building tasks.

A 3D building model is built within the open-source, Python-based computational framework COMPAS¹ and implemented in Rhino Grasshopper to simulate robot state and task status in relation to the building model. The kinematic model of the robot is visualized using the `compas_fab`

Algorithm 1: Proposed algorithm for HAL

```

1 function localize ( $\mathbf{y}_{t_n}, \mathcal{S}, m, M$ );
Input : Initial guess for pose of end-effector  $\mathbf{y}_{t_n}$ ,
        extrinsic calibration of sensors  $\mathcal{S}$ , iterations
         $m$ , and 3D triangle mesh model  $M$ 
Output: corrected pose of end-effector  $\mathbf{y}_{t_n}^{HAL}$ 
2  $e_{t_n}^{HAL}(\mathbf{y}_{t_n}) = 0$ 
3  $\mathbf{y}_{t_n}^{(i)} \leftarrow \mathbf{y}_{t_n}$ 
4 for  $i = 1 : n$  do
5    $\mathbf{y}_{t_n}^{(i)} \leftarrow \text{moveSensorHead}(\mathbf{y}_{t_n}^{(i)})$ ;
6   for  $j = 1 : \text{num\_sensors}$  do
7      $\mathbf{y}_{t_n}^{(j)} \leftarrow \text{retrievePose}(\mathbf{y}_{t_n}^{(i)}, \mathcal{S})$ 
8      $z_j \leftarrow \text{measureDistance}()$ 
9      $\mathbf{t}_p, \mathbf{r}_p \leftarrow \text{raytrace}(\mathbf{y}_{t_n}^{(j)}, M)$ 
10     $e_j \leftarrow c \left\| z_j - \frac{(\mathbf{t}_p(\mathbf{y}_{t_n}^{(j)}) - \mathbf{y}_{t_n}^{(j),t}) \mathbf{r}_p(\mathbf{y}_{t_n}^{(j)})}{\mathbf{y}_{t_n}^{(j),R} \mathbf{r}_p(\mathbf{p}_j)} \right\|$ 
11     $e_{t_n}^{HAL}(\mathbf{y}_{t_n}) += e_j$ 
12 end
13 end
14  $\mathbf{y}_{t_n}^{HAL} \leftarrow \underset{\mathbf{y}_{t_n}}{\operatorname{argmin}} e_{t_n}^{HAL}(\mathbf{y}_{t_n})$ 
15 return  $\mathbf{y}_{t_n}^{HAL}$ ;

```

package of the COMPAS framework. Communication between high-level manager and building environment is established using robotic middle-ware `roslibpy`². This interface between building task and high-level manager enables automatic task generation, monitoring, and user intervention. Building tasks defined within the building model are thus converted to end-effector poses represented in the reference frame of the robot and sent to the high-level manager.

While the current system uses feed-forward task generation from the building model, the system is designed to receive updates from the high level-manager for the as-built status of the construction environment to update the building model and robot tasks. Furthermore, planned for future work, it will enable leveraging task-specific references,

¹ <https://compas-dev.github.io/>

² <https://roslibpy.readthedocs.io>

such as relative distances to selected building components, and relative tolerances between consecutive building tasks.

D. Motion Planning

The motion planning module generates robot state reference trajectories to bring the end-effector to the target pose commanded by the building task interface. A typical task consists of: globally navigating the wheeled base through the workspace to reach the task execution area, jointly controlling base and arm motion to get the end-effector near the commanded pose, executing the HAL routine based on the laser distance sensors on the end-effector, then finally fine adjusting the end-effector to the commanded pose.

The planning module is made up of three components in order to address these subtasks. When the robot is far away from its target, a stochastic path planner computes planar base trajectories for navigation through the workspace towards the target. The planner takes into account both prior-knowledge of the building structure as well as LiDAR data, which enables re-planning for dynamic obstacle avoidance. An MPC module is responsible for generating whole-body motion references. As an input it takes either the base motion plan from the navigation planner, or, when the robot is close enough to its target, a desired end-effector pose. In both cases a continuous optimal whole-body motion plan, comprised of both base and arm motion, is computed to track the desired references. In order to accurately bring the end-effector to the desired poses required for the HAL routine and execution of the building task, the base is assumed to be stationary and a pose interpolation approach is used, combined with an iterative inverse kinematics algorithm, in order to compute desired joint positions. These three components of the planning module are described in more detail below.

1) *Base Navigation Planning:* The objective of the base navigation planner is to drive the robot close to the area where the tasks need to be executed. Collision-free trajectories are computed in the workspace by running the RRT* algorithm [19], while checking for collisions by means of Octomap [20]. The planner adopts an *optimistic* behavior by assuming the unknown space to be free. The initial occupancy map of the environment is generated using the point-cloud sampled from the 3D CAD triangle mesh model of the robot's environment and directly used for planning. The initial map is then updated incrementally using the LiDAR scans acquired during navigation, in order to identify obstacles which are not captured by the building model or to clear free areas which were previously considered as occupied. In case the building model is not accessible, the occupancy map is built incrementally from scratch during the navigation. Collision checks are performed along the planned trajectory every time the occupancy map is updated considering a rectangular bounding box around the robot. This allows for fast reactions to obstacles that appear and may block the path during navigation. Every time a potential collision is identified along the current planned path, the

planner is triggered and a new trajectory that considers the updated map of the environment is computed.

2) *Whole-Body MPC:* The MPC module, implemented with the OCS2 framework [21], uses a kinematic plant model of the robot. The robot state consists of the base position and orientation in the horizontal plane and the arm joint angles $\mathbf{x} = (x_b, y_b, \theta_b, q_0, \dots, q_5)^T$. The module's outputs are the base linear velocity, the steering rate in the horizontal plane and the arm joint angular velocities $\mathbf{u} = (v_b, \dot{\theta}_b, \dot{q}_0, \dots, \dot{q}_5)^T$. The output is forwarded to the inverse dynamics tracking controller.

Two different types of references are used. To track the base motion plan, the references $\tilde{\mathbf{x}}_j$ are the base position and orientation and default joint angles for the arm. For generating whole body motion trajectories the desired end-effector pose \mathbf{x}_{ee} is used as a reference. The controller tracks both references with different cost functions. A unified cost function is set up as the sum over both individual costs to track both references with a single MPC strategy:

$$J(\mathbf{x}, \tilde{\mathbf{x}}, \mathbf{u}) = \int_{t_0}^{t_0+T} (1 - \alpha) C_j + \alpha C_{ee} + \mathbf{u}^T \mathbf{R} \mathbf{u} d\tau + \alpha \Phi_{ee} \quad (4)$$

$$C_j = (\mathbf{x} - \tilde{\mathbf{x}}_j)^T \mathbf{Q}_j (\mathbf{x} - \tilde{\mathbf{x}}_j) \quad (5)$$

$$C_{ee} = (\xi(\mathbf{x}) \boxminus \tilde{\mathbf{x}}_{ee})^T \mathbf{Q}_{ee} (\xi(\mathbf{x}) \boxminus \tilde{\mathbf{x}}_{ee}) + (\mathbf{x} - \tilde{\mathbf{x}}_d)^T \mathbf{Q}_d (\mathbf{x} - \tilde{\mathbf{x}}_d) \quad (6)$$

$$\Phi_{ee} = (\xi(\mathbf{x}(T)) \boxminus \tilde{\mathbf{x}}_{ee}(T))^T \mathbf{Q}_{ee,\Phi} (\xi(\mathbf{x}(T)) \boxminus \tilde{\mathbf{x}}_{ee}(T)) \quad (7)$$

The parameter $\alpha \in \{0, 1\}$ determines which cost is active and negotiates a smooth transition between the two operating modes. The cost C_j penalizes deviations of the current robot state $\mathbf{x}(t)$ to the base motion plan $\tilde{\mathbf{x}}_j$ with $\mathbf{Q}_j = \mathbb{W}$. C_{ee} penalizes end-effector tracking errors with $\mathbf{Q}_{ee} = \mathbb{W}$ and an additional weaker cost encourages maintaining a default arm configuration with $\mathbf{Q}_d = 0.01 \text{ diag}(0, 0, 0, 3, 10, 10, 0, 5, 0)$. The small weight on the default configuration helps to avoid self-collisions and collisions of the base with the wall. $\mathbf{R} = \mathbb{W}$ puts costs on the control inputs \mathbf{u} . An additional terminal cost Φ_{ee} with $\mathbf{Q}_{ee,T} = 10 \mathbb{W}$ penalizes deviations from the desired end-effector pose at the end of the MPC horizon. The transformation function ξ from the robot state $\mathbf{x}(t)$ to the end-effector pose is autogenerated by the RobCoGen library [22]. The code is templated over the scalar type such that the automatic differentiation toolbox CppAd [23], used in OCS2, can generate derivatives of the cost function to solve the MPC problem [24]. With the distance measure \boxminus , we subtract the positions of the two poses and use a quaternion distance measure for the orientations [25]. The MPC is solved over a time horizon of $T = 1$ s. If the desired target cannot be reached within the current time horizon, the references are interpolated and for each roll-out a reference is chosen such that it can be reached in the given time, assuming some nominal speed.

3) End-effector Pose Interpolation and Inverse Kinematics: For the situations in which the base is stationary, smooth motion reference paths between desired end-effector poses are obtained using pose interpolation. This is done by defining a smooth interpolation profile with zero velocity and acceleration at the start and end, and applying that to linearly interpolate between position vectors and using spherical linear interpolation for the orientation quaternions. To compute the arm joint positions for the desired end-effector poses, we use an iterative inverse kinematics scheme as described in [26].

E. Whole-Body Motion Control

The motion references generated by the components of the planning module (described in Sec. III-D) are tracked by a motion controller. Two separate motion controllers are used for tracking the whole-body MPC motion references and for precisely tracking end-effector pose references when the base is stationary.

During navigation and whole-body end-effector motion tracking, the motion references come from the MPC in the form of velocities for the mobile base in the horizontal plane and the six arm joints. The velocity references for the wheeled base are translated into feed-forward wheel velocity references v_{wheel}^{ff} using a two-wheeled differential drive model of the robot. Because the actual base has four wheels and depends on a skid-steering mechanism which introduces a significant amount of wheel slip, an additional term v_{wheel}^i is added based on the integrated error between the desired and measured linear and angular base velocities. The final wheel velocity references are computed as $v_{wheel}^{ref} = v_{wheel}^{ff} + v_{wheel}^i$ and are tracked by the base motor controller. For the arm joints, the reference velocities are numerically integrated to obtain position references, and numerically differentiated and low-pass filtered to obtain acceleration references. These three levels of motion references are then tracked using torque-control. This allows for compliant and safe interaction when the arm comes in contact with humans or its environment. The required joint torques are computed using a numerical inverse dynamics controller as proposed in [27]. Computation of the joint-torque references takes into account the system dynamics, desired joint motions, and the actuator torque limits. To compensate for unmodeled dynamics, such as joint friction, we update these torque references with additional desired torques based directly on the joint position and velocity errors. These joint torque references are then tracked by the arm's internal controller.

When executing the HAL scanning routine and moving to the final task pose the arm actuators are switched to position-control mode. In this situation, the joint position references from the inverse kinematics solver (Sec. III-D.3) are commanded directly.

IV. EXPERIMENTS

We evaluate the system on a prototype robotic platform within a realistic construction environment.

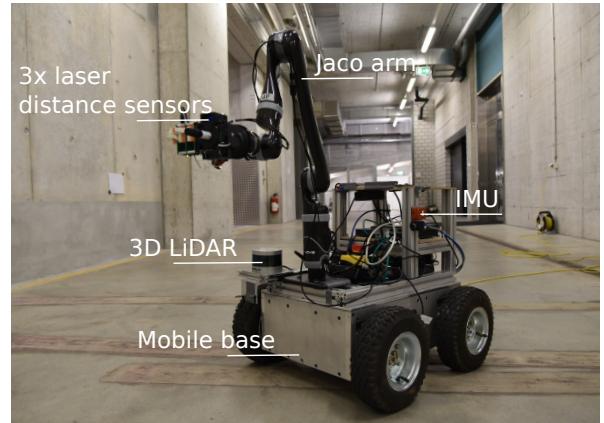


Fig. 3: The mobile construction robot research platform *Waco* used in this work, equipped with 6 DoF arm, and sensors.

A. Hardware Setup

The robot used for the experiments is depicted in Fig. 3. The hardware configuration consists of a variety of stock components, and custom built installations, thus forming a research platform suitable for the testing of software solutions for highly accurate mobile manipulation. The robotic platform comprises a mobile base (*Inspector Bots Super Mega Bot*³) with mounted Kinova Jaco robot arm⁴, a custom 3D-printed end-effector, sensor suite for localization, and computing hardware. The robot is named *Waco*, derived from "Wheeled Jaco". As a tool, the end-effector is equipped with a small spring in series with a marker, in order to perform non-destructive discrete tool positioning tasks. For on-board localization, the robotic platform is equipped with a Velodyne VLP-16 sensor, an Xsens MTi-100 IMU, wheel encoders, and a sensor-head consisting of three orthogonal laser distance measurement sensors at the end-effector. The on-board computing hardware is based on an Intel Core i7-6700 CPU @ 3.4 GHz, and 16 GB RAM. A second off-board computer, communicating over WLAN, runs the building task interface for the operator.

B. Experimental setup

The system is tested in a realistic construction site environment with unfinished concrete walls, and some clutter, see Fig. 1, and Fig. 4. Using the building task interface, we specify interaction tasks on three different walls, which are the locations to be marked by the custom-made tool. The system accuracy is first tested by repeatedly commanding different targets in a 3×3 -dot pattern on each wall, as illustrated in Fig. 7. The robot then autonomously approaches the task locations from random locations several meters away, and places a single dot. This procedure is thereafter repeated. For later ground-truth evaluation, the commanded task locations and the final placements of the end-effector are measured with a Leica Nova TM50 Total Station. Furthermore, we demonstrate the entire system performing fully autonomous loops between three task locations (one target

³www.inspectorbots.com

⁴www.kinovarobotics.com

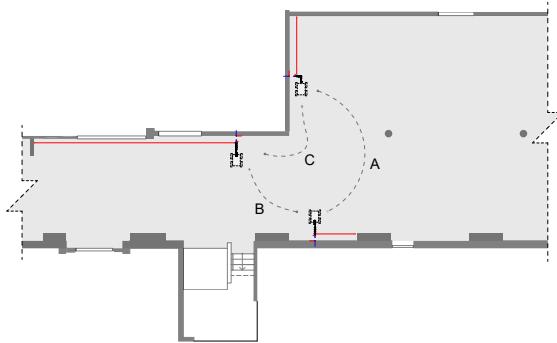


Fig. 4: Plan view of the experimental setup: The robot is illustrated in the three task locations A, B, and C. The red lines indicate the lateral HAL references, the vertical reference is measured against the floor.

on each wall), as depicted in Fig. 1 and Fig. 4. Moreover, we vary the experimental setup by dynamically adding obstacles in the planned path of the robot, demonstrating adaptive re-planning of the platform. In the experiments, we use a building model of the environment with minor deviations between the as-planned and as-built status. However, notable deviations between the reality and the building plan persist, such as smooth surfaces instead of brick walls, unmodeled HVAC components such as pipes, and small deviations in walls.

C. Results

The integrated system experiment shows reliable performance throughout 30 individual experiments. The robot concludes multiple cycles between the three task locations without human intervention, also adaptively avoiding dynamic obstacles⁵.

1) Positioning evaluation: The robot performance in the interaction experiment at the three task locations is reported in Fig. 5 and Table I. Here, we report the absolute errors in global coordinates, with respect to the building model origin used to derive the robot tasks. Furthermore, the relative errors within the drawing pattern of the robot are depicted. For this, all pair-wise combinations are evaluated. The plots show the distribution and total spread of all measurements. Overall, we observe mean relative errors between 3.3 mm and 5.9 mm. The absolute errors for the three task locations are below 1 cm for location A, and between 22 mm and 37 mm for locations B and C respectively. The lateral (x) and vertical components (z) of the errors are split in Fig. 6. Notable observations are a high lateral offset for locations B and C, as well as a high vertical offset at task location C, while both error components are low at location A. A sample 3×3 dot pattern created by the robot at location B is depicted in Fig. 7.

2) Discussion: This section offers a discussion of the experimental results. Firstly, it is noteworthy that the system achieves generally low relative errors facilitating locally

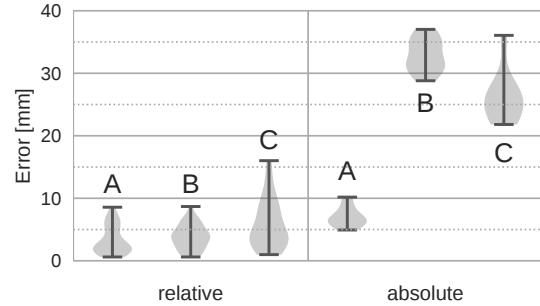


Fig. 5: Absolute and relative errors in point position for a nine dot pattern at the three different task locations.

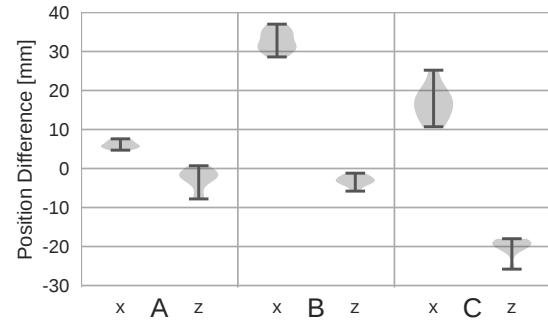


Fig. 6: Difference between commanded and executed point position over the different axes in the nine dot pattern.

repeatable references. This is especially important for locally accurate construction tasks, e.g., installation of several screws in a specific pattern to mount a lamp or rail. However, the absolute error is presently highly dependent on the task location. While the principal solution of the HAL routine is theoretically sound, we identified sources of error leading to increased absolute errors particularly for locations B and C.

In both locations, we note a large lateral absolute error. This can be best attributed to our current experimental set-up of taking reference walls with large distance in lateral direction as localization reference for the HAL routine. Here, the distances between task locations, and reference walls are 1.7 m, 11.5 m, and 4 m for locations A-C respectively. Slight rotational errors in the external calibration of the laser distance sensors are magnified by these distances. Furthermore, the floor at location C has an inclination that is not accurately reflected in the building model, adding

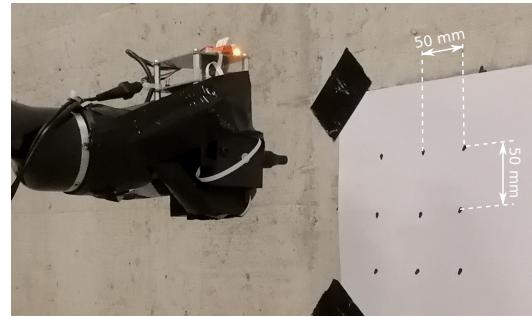


Fig. 7: Nine dot task pattern at location B: The robot is commanded to perform a 50 mm-spaced 3×3 -grid of interactions in three locations.

⁵Please refer to the accompanying video for a full demonstration: https://youtu.be/O182Gh_1T9w.

TABLE I: Mean positioning errors.

	A	B	C
Mean absolute error [mm]	7.0	32.9	26.1
Mean relative error [mm]	3.3	3.8	5.9

error components to the absolute error both laterally, and vertically. Also, the laterally referenced walls for B and C are rough brick walls, instead of the load-bearing concrete walls in the building plan.

Lastly, we also tested the system with a simplified building model of the environment having several significant deviations from the as-built conditions, notably in the inclination of the floor, and the distances between the walls. Testing in this scenario led to a large increase in errors especially in relative errors for location C by factors of 2 – 3, as most measurements were rejected as outliers given the discrepancy from the as-planned building model. While the system performs well in locations of the building model that reflect the real conditions, a promising avenue of future research is handling of discrepancies between as-planned and as-built status of construction environments as well as handling of highly cluttered environments containing a large amount of unmodelled objects.

V. CONCLUSIONS

In this work, we presented a system for high-accuracy interactions of a mobile robot for construction sites. The system performs all state estimation using on-board sensing only. Notable features are a tight integration with the digital building model for task allocation, and robot localization. It is implemented on a robotic platform, and is able to reliably reach high-accuracy interactions under realistic conditions. Presently, the HAL routine is hand-designed. In future work it would be interesting to investigate task-adaptive strategies to yield the most reliable measurements. Furthermore, we wish to close the feedback loop between the perception and building task interface, automatically updating the building model perceived from the as-built conditions.

REFERENCES

- [1] H. Ardiny, S. Witwicki, and F. Mondada, “Are autonomous mobile robots able to take over construction? a review,” *International Journal of Robotics, Theory and Applications*, vol. 4, no. 3, pp. 10–21, 2015.
- [2] M. Schwartz and J. Park, “Robotic control strategies applicable to architectural design,” *Technologyl Architecture+ Design*, vol. 1, no. 1, pp. 73–82, 2017.
- [3] P. Bosscher, R. L. Williams II, L. S. Bryson, and D. Castro-Lacouture, “Cable-suspended robotic contour crafting system,” *Automation in Construction*, vol. 17, no. 1, pp. 45–55, 2007.
- [4] J. Willmann, F. Augugliaro, T. Cadalbert, R. D’Andrea, F. Gramazio, and M. Kohler, “Aerial robotic construction towards a new field of architectural research,” *International journal of architectural computing*, vol. 10, no. 3, pp. 439–459, 2012.
- [5] V. Helm, S. Ercan, F. Gramazio, and M. Kohler, “Mobile robotic fabrication on construction sites: Dimrob,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2012, pp. 4335–4341.
- [6] T. Sandy, M. Gifthaler, K. Dörfler, M. Kohler, and J. Buchli, “Autonomous repositioning and localization of an in situ fabricator,” in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 2852–2858.
- [7] X. Zhang, M. Li, J. H. Lim, Y. W. D. Tay, H. Pham, and Q.-C. Pham, “Large-scale 3d printing by a team of mobile robots,” *Automation in Construction*, vol. 95, pp. 98–106, 2018.
- [8] C. Feng, Y. Xiao, A. Willette, W. McGee, and V. R. Kamat, “Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites,” *Automation in Construction*, vol. 59, pp. 128–138, 2015.
- [9] S. J. Keating, J. C. Leland, L. Cai, and N. Oxman, “Toward site-specific and self-sufficient robotic fabrication on architectural scales,” *Science Robotics*, vol. 2, no. 5, 2017.
- [10] K. Dörfler, T. Sandy, M. Gifthaler, F. Gramazio, M. Kohler, and J. Buchli, “Mobile robotic brickwork,” in *Robotic Fabrication in Architecture, Art and Design*. Springer, 2016, pp. 204–217.
- [11] R. C. Luo, Y.-C. Lin, and C.-C. Kao, “Autonomous mobile robot navigation and localization based on floor plan map information and sensory fusion approach,” in *IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2010, pp. 121–126.
- [12] F. Boniardi, T. Caselitz, R. Kümmel, and W. Burgard, “Robust lidar-based localization in architectural floor plans,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3318–3324.
- [13] ———, “A pose graph-based localization system for long-term navigation in cad floor plans,” *Robotics and Autonomous Systems*, vol. 112, pp. 84–97, 2019.
- [14] J. Röwekämper, C. Sprunk, G. D. Tipaldi, C. Stachniss, P. Pfaff, and W. Burgard, “On the position accuracy of mobile robot localization based on particle filters combined with scan matching,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 3158–3164.
- [15] G. Vasiljević, D. Miklić, I. Draganjac, Z. Kovačić, and P. Lista, “High-accuracy vehicle localization for autonomous warehousing,” *Robotics and Computer-Integrated Manufacturing*, vol. 42, pp. 1–16, 2016.
- [16] T. Sandy, L. Stadelmann, S. Kerscher, and J. Buchli, “ConFusion: Sensor fusion for complex robotic systems using nonlinear optimization,” *IEEE Robotics and Automation Letters*, 2019.
- [17] T. Sandy and J. Buchli, “Dynamically decoupling base and end-effector motion for mobile manipulation using visual-inertial sensing,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017, pp. 6299–6306.
- [18] V. Kubelka, L. Oswald, F. Pomerleau, F. Colas, T. Svoboda, and M. Reinstein, “Robust data fusion of multimodal sensory information for mobile robots,” *Journal of Field Robotics*, vol. 32, no. 4, pp. 447–473, 2015.
- [19] S. Karaman and E. Frazzoli, “Sampling-based algorithms for optimal motion planning,” *International Journal of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011.
- [20] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, “OctoMap: An efficient probabilistic 3D mapping framework based on octrees,” *Autonomous Robots*, 2013.
- [21] F. Farshidian, E. Jelavic, A. Satapathy, M. Gifthaler, and J. Buchli, “Real-time motion planning of legged robots: A model predictive control approach,” in *IEEE-RAS International Conference on Humanoid Robotics (Humanoids)*, 2017, pp. 577–584.
- [22] M. Frigerio, J. Buchli, D. G. Caldwell, and C. Semini, “RobCoGen: a code generator for efficient kinematics and dynamics of articulated robots, based on Domain Specific Languages,” vol. 7, no. 1, pp. 36–54, 2016.
- [23] B. M. Bell, “Cppad: a package for c++ algorithmic differentiation,” *Computational Infrastructure for Operations Research*, vol. 57, p. 10, 2012.
- [24] M. Gifthaler, M. Neunert, M. Stäuble, M. Frigerio, C. Semini, and J. Buchli, “Automatic differentiation of rigid body dynamics for optimal control and estimation,” *Advanced Robotics*, vol. 31, no. 22, pp. 1225–1237, 2017.
- [25] B. Siciliano, Ed., *Robotics: modelling, planning and control*, ser. Advanced textbooks in control and signal processing. London: Springer, 2009.
- [26] A. Goldenberg, B. Benhabib, and R. Fenton, “A complete generalized solution to the inverse kinematics of robots,” *IEEE Journal on Robotics and Automation*, vol. 1, no. 1, pp. 14–20, 1985.
- [27] C. D. Bellicoso, F. Jenelten, P. Fankhauser, C. Gehring, J. Hwangbo, and M. Hutter, “Dynamic Locomotion and Whole-Body Control for Quadrupedal Robots,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.