

Received April 13, 2022, accepted April 23, 2022, date of publication April 25, 2022, date of current version May 17, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3170408

Vision-Based UAV-UGV Collaboration for Autonomous Construction Site Preparation

OREN ELMAKIS¹, TOM SHAKED¹, AND AMIR DEGANI¹, (Member, IEEE)

Technion Autonomous Systems Program, Technion - Israel Institute of Technology, Haifa 3200003, Israel

Corresponding author: Amir Degani (adegani@technion.ac.il)

This work was supported in part by the Technion Autonomous Systems Program, and in part by the Ministry of Science & Technology, Israel.

ABSTRACT Construction site preparation tasks rely on experienced operators and heavy machinery for clearing debris, earthmoving, leveling, and soil stabilization. These actions require complex collaboration between human teams to survey the site, estimate the material condition, and guide the operators accordingly. In recent years there has been a critical labor shortage due to increasing demands in construction. Integrating autonomous systems can mitigate this gap by replacing traditional methods with robotic solutions. However, while ideal conditions for automatic systems are static and highly controlled, construction sites are dynamic and unstructured environments. The ability of autonomous systems to overcome these conditions during outdoor construction site preparation tasks relies on their capacity to map the material on-site and continuously perform localization. This study suggests a solution to these problems by collaborating between an Unmanned Aerial Vehicle (UAV) and an Unmanned Ground Vehicle (UGV). In this method, the UAV produces a material map and monitors the UGV's location relative to known static landmarks. These measurements are then sent to the ground vehicle and are added to the onboard sensors using the Extended Kalman Filter (EKF) approach. Thus, the UAV enhances the operation of the UGV by providing an accurate localization and mapping from the air and allowing it to perform a site-preparation task beyond mere sensing. This approach is examined with simulation and validated by outdoor experiments. Additionally, this method is integrated within Shepherd, a custom-developed plugin for computer-aided design applications.

INDEX TERMS Autonomous systems, multi-robot systems, pose estimation, robotics and automation, system implementation.

I. INTRODUCTION

The labor shortage is one of the biggest hurdles in the construction industry today [1], with more than eighty-one percent of construction firms recently reporting difficulty filling different positions [2]. The global COVID-19 pandemic exacerbates this problem by reducing access to labor. This situation emphasizes the need for automated solutions to construction-related tasks that currently rely on manual labor [3]. Recent advancements in robotics provide a possible solution to this problem by exploring multi-agent collaboration in the construction site. Multi-agent collaboration allows performing complex construction tasks such as assemblies of complex architectural buildings using precise localization [4] or creating free-form embankments

with agile control [5]. Incorporating robotic platforms in construction can not only compensate for the shortage of workers but also aid in performing complex tasks that require high skill, such as preparing wire mesh reinforcements for concrete or carving stone [6], [7].

Nevertheless, as construction usually begins with preliminary site preparation activities, such as clearing debris, earthmoving, leveling, and soil stabilization [8] need to be met. These preparation tasks are traditionally performed by human operators driving bulldozers, excavators, loaders, and dump trucks. As these tasks are complex and laborious, they rely on experienced human operators for their successful accomplishment. The limited access to skilled workers, recently worsened by the global pandemic, requires developing alternative approaches to site preparation that employ remote control and autonomy using Unmanned Ground Vehicles (UGV) (Fig. 1).

The associate editor coordinating the review of this manuscript and approving it for publication was Abderrahmane Lakas¹.

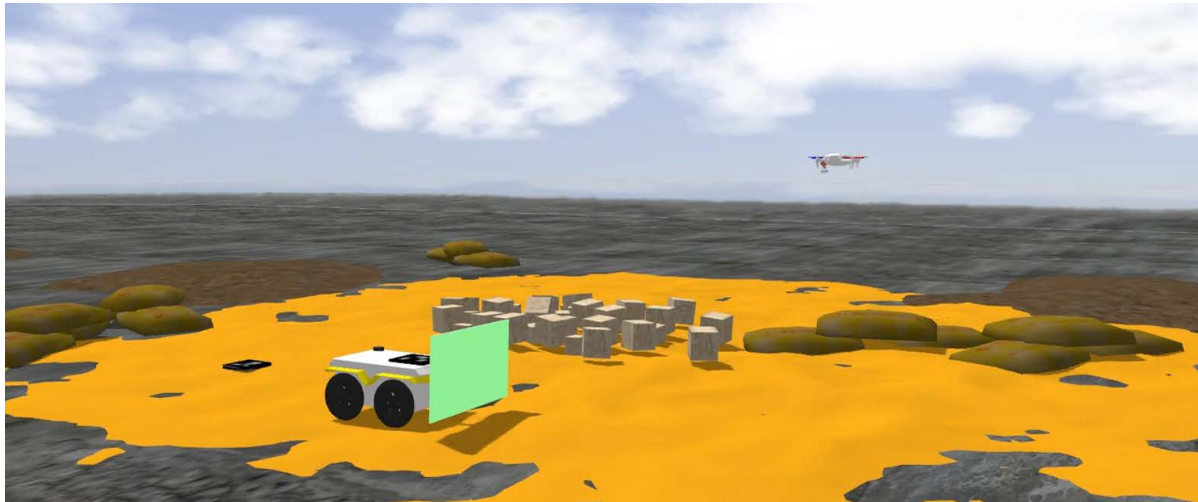


FIGURE 1. A team of a UGV and a UAV performing construction site preparation simulated in Gazebo.

In the recent decade, there has been a growing interest in research of Unmanned Aerial Vehicles (UAVs) as assisting platforms for various kinds of ground vehicles. This paradigm can potentially improve the utility of ground

vehicles by providing them with an additional pair of “eyes in the sky” [9], [10]. In this context, the motivation for this research is to enable human-guided autonomous construction site preparation. Achieving this goal requires mapping the material on-site and performing accurate localization throughout the process.

The novelty of the presented research is in developing a computer vision method for adaptive material mapping and a collaborative Extended Kalman Filter (EKF) for localizing the robotic platform on the ground using a UAV. These capacities are crucial since interactions with the material lead to changes in the environment and errors in the localization of the UGV. By overcoming these challenges, the paper contributes to UAV-UGV collaborative schemes with specific applications to autonomous construction site preparation.

The outcome of this research provides a method for eliminating navigation errors caused by sharp maneuvers and material pushing [11], as well as generating an online grid map of the material on-site. This method is introduced as a new capacity inside Shepherd, a custom-developed plugin integrated with Rhinoceros 3D computer-aided design application (CAD) and Grasshopper visual programming user interface (UI). Shepherd provides an interface for human-robot collaboration for site preparation with only simple guidance from a human operator [12].

II. RELATED WORK

This section provides the state-of-the-art in the field of autonomous robotics, focusing on localization and state estimation as well as the gaps that hinder autonomy in robotic applications for construction sites.

A. AUTONOMOUS ROBOTICS IN CONSTRUCTION

In recent years, accelerated technological advances have allowed robots to operate with greater speed and precision in indoor environments, while significant challenges regarding outdoor environments remain unsolved. In this context, recent studies have focused on achieving autonomy in construction in on-site conditions [13]. These studies span from designing construction application algorithms for building motion-supporting structures such as ramps [14] to the assembly of dry stone walls [15]. Others present algorithms for autonomous safety implementation using LiDAR measurements [16] or robust planning and control approach for excavation [17]. While existing research focuses on enabling robotic applications in highly controlled static environments, achieving autonomy on-site requires overcoming dynamic and unstructured conditions. A fundamental capacity for achieving on-site autonomy relies on localization and material mapping - the ability to determine the location of the robot with respect to its environment during the earth moving task while mapping the material in the work environment in real-time.

B. UGV-UAV COLLABORATION

The growing interest in UAV-UGV collaboration is demonstrated by research on collaborative path planning, localization, obstacle avoidance, and mapping [9], [10]. Previous work on the topic provides a control scheme to coordinate ground and aerial vehicles for locating moving targets in a given area [18]. As this research focuses on distributed control and obstacle avoidance, it does not employ UAVs to assist the UGVs. Research on decentralized aerial and ground cooperation schemes uses a vision-based tracking controller for object transportation in unsafe industrial areas [19]. Here, visual data collected by the UAV is used for assisting the UGVs, but it is limited to obstacle avoidance.

Recent research expands collaboration schemes by providing a framework for vision-based cooperative path-planning using an optimal A* algorithm [20]. However, for localization, the study relies on an ultra-wideband indoor positioning system that requires deploying multiple sensors on-site. An alternative approach employs a UGV-UAV team for collaborative Simultaneous Localization and Mapping (SLAM) [21]. However, this approach assumes a static environment where all robots perform the same mapping task. In contrast, autonomous construction site preparation requires specialized robots to perform specific tasks that change the environment in which they work. Hence, the presented research requires methods and tools that enable localization and material mapping in dynamic environments.

C. MATERIAL MAPPING

Material mapping is defined here as a process allowing for tracking the location of the gravel on-site. This process is essential for autonomous site-preparation tasks as it enables planning the earthmoving action according to the current location and dispersion of the material. This process requires distinguishing the material from its surroundings – a capacity relying on computer vision. Existing computer vision methods perform such tasks using semantic segmentation based on advanced deep learning that relies on pre-existing data sets [22]. However, due to large variability in the used material (gravel), this method requires building a comprehensive data set and performing exhaustive model training.

In contrast, human operators can easily recognize objects on-site without prior knowledge, while robots can accurately map the material during the task [23], [24]. Additionally, integrating material mapping and path planning with human operation increases the system's capabilities, making it more flexible and robust to various site preparation scenarios [25]. Therefore, in the presented research, a human performs initial material recognition while the robotic system is used for online material mapping throughout the task.

D. ROBOT LOCALIZATION

Localization during outdoor construction tasks is challenging due to unstructured environments, moving objects, and dynamic efforts that are highly complicated to model. For instance, as wheel odometry estimates the robot's motion according to the rotary encoder's measurements and the wheel shape [26], its output suffers from accumulated errors when performed in rough terrain. Over time, these errors accumulate and lead to severe drifting of the localization. In contrast, skidding does not affect visual odometry as it relies on tracking the geometrical features in a frame to assess the camera motion [27]–[29]. Nevertheless, visual odometry requires a static environment and cannot recognize similar features while performing sharp maneuvers.

Currently, a common solution for achieving accurate outdoor localization relies on real-time kinematic positioning (RTK) to correct for common errors in satellite navigation

systems (GNSS). The RTK-GNSS measurements are satisfactory in many autonomous tasks in which the environment is static. However, these measurements can be distorted by weather conditions or disruptions in communication [30]. In addition, construction tasks require reliable and constant localization for accurate implementation. Furthermore, the robot needs a relative localization that matches the construction plans, requiring a challenging and expensive global mapping process to coordinate the axis systems. The presented research uses static landmarks in the construction zone to enable localization relative to other objects by performing on-site state estimation.

E. STATE ESTIMATION METHODS

The state estimation problem is typically solved using methods that integrate data from a variety of sensors to minimize localization error [31] - for example, the non-parametric approach for state estimation such as Adaptive Monte Carlo Localization (AMCL). The AMCL algorithm is a filter initialized by placing scattered particles on the map and examining their weight by comparing the expected and actual measures. Implementing the AMCL requires prior knowledge of the environment, such as a static map. While a static map can be generated from georeferenced aerial images [32]–[34], using an offline map is not possible for construction tasks in which robotic platforms actively change their environment. This phenomenon can be seen by moving bricks around during robotic wall building [35] or by altering the entire topography during autonomous robotic excavation [5].

Another common method for state estimation is the two-stage Kalman filter approach, using the UGV model to predict the motion and then correct the state by utilizing onboard sensors data [36]. However, the Kalman filter is adequate only for linear systems, which are only a few in relative to the non-linear systems. In contrast, the Extended Kalman Filter (EKF) uses a linearization technique to handle non-linear systems. This allows the EKF to be deployed in multiple systems, providing the state estimation problem fast and simple solution [37]. Implementing onboard sensor data is limited and may reduce the efficiency of the system and result in unreliable localization. For example, using an onboard laser range finder and matching the data requires slow base motion, thus increasing the time to finish a task [35]. In other cases, such as using symmetric maps without relative measurements, the localization may not converge at all [38]. Therefore, research can benefit from the use of an additional agent providing accurate online localization by monitoring the UGV.

As the study of robotic collaboration is not entirely covered yet, different agents can improve the performance of the task and provide complementary capabilities. The studies are focused on developing algorithms for the autonomous collaborative process [39], [40], assuming perfect localization or available RTK-GNSS measurements [41]. Other studies consider the localization problem but either use

similar agents [42] or enable work in static environments [21]. The presented research employs collaboration to address these gaps using a quadrotor UAV to perform Site Localization – estimating the location of a UGV on-site in real-time during construction tasks. The research suggests using the Extended Kalman Filter (EKF) on the UGV while integrating accurate data of online localization from the UAV. The localization acquired by the EKF is local and relative to static objects inside the construction site.

III. FRAMEWORK AND METHODS

The following section provides the research framework and methods that enable vision-based UAV-UGV collaboration for localization and material mapping, supporting site preparation tasks that involve granular and soil interactions. These capacities allow a UAV to accurately use prior knowledge on the location of static landmarks to locate a UGV on-site. In this process, Shepherd is used for designing and implementing the site preparation action (Fig. 2). The framework is presented through four main topics: (1) system

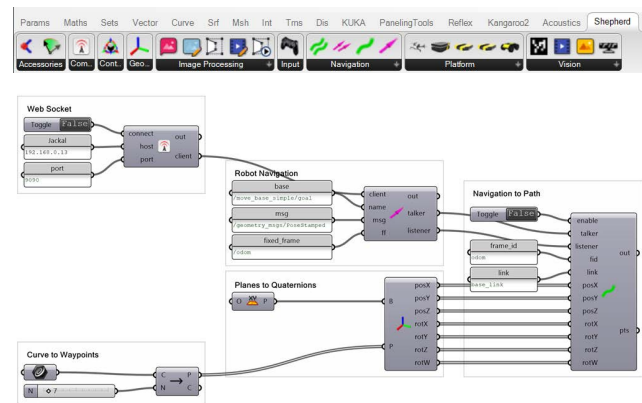


FIGURE 2. The Shepherd user interface in Rhino 3D Grasshopper plugin and a sample visual code of converting a Rhino 3D curve to a robot path.

overview, (2) site visualization, (3) material mapping, and (4) site localization.

A. SYSTEM OVERVIEW

The proposed system is assembled from three components: Shepherd a human-robot interface, a UGV, and a UAV (Fig. 3). The collaboration between the components allows to perform the task and enhances the overall capabilities of the system. The robot team used in this research consists of a quadrotor UAV and a four-wheeled UGV. The UGV is equipped with a custom front shovel covering the entire front of the platform for material pushing tasks (Fig.4). The UGV’s field of view is limited, and its maneuverability depends on the site conditions. Therefore, a quadrotor equipped with a monocular downward-facing camera is used to provide an online top-view of the construction site. The top-view images allow closing the loop of the autonomous task by visually monitoring the UGV as well as additional objects in the construction zone (Fig. 5).

The monitoring loop allows real-time visual analysis by providing information relevant to the task. Specific information crucial to site preparation is the localization of the UGV. The view provided by the quadrotor allows to reliably measure the position of the UGV in relation to static objects on-site. For this purpose, the system includes landmarks inside the construction zone. The landmarks used are ArUco markers [43], which are placed in known locations in the map, and on top of the UGV.

Controlling the UGV is done using Shepherd, a custom tool developed for simulating and controlling mobile robotic platforms using the Rhinoceros 3D Grasshopper UI [12]. The tool allows non-expert operators to (1) rapidly explore alternative paths by employing parametric motion planning, (2) simulate them within the Rhinoceros 3D modeling environment, and (3) execute them using Grasshopper UI to control a Robotic Operation System (ROS) running robotic

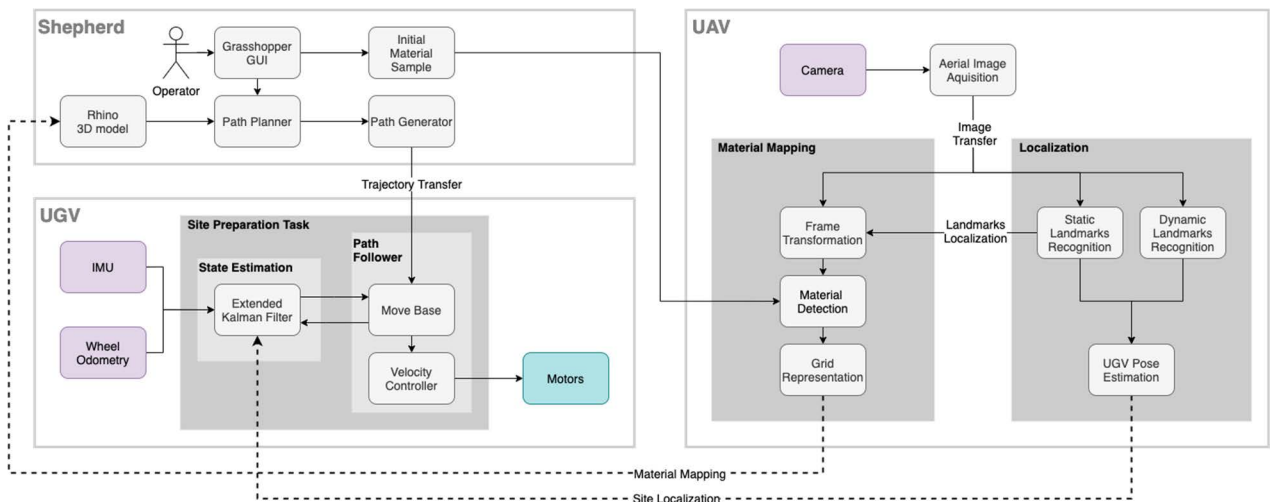


FIGURE 3. High-level model of the system describing the function of the three main components: Shepherd, the UGV and the UAV.



FIGURE 4. The UGV agent jackal of clearpath with a custom blade.

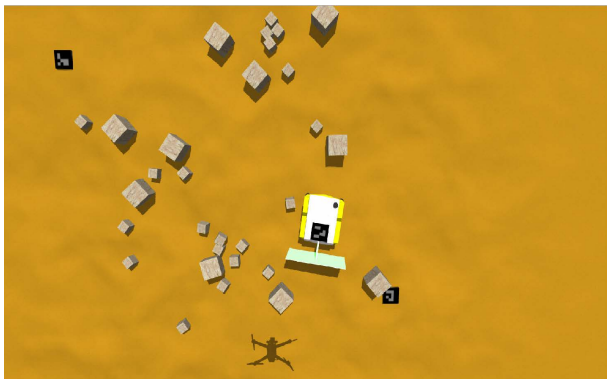


FIGURE 5. Illustration of the top-view image of the construction zone from the quadrotor, in black the ArUco tags as a static landmarks.

platforms in real-time. Its capacities include (1) generating paths using primitive geometries, (2) simulating the behavior of robotic platforms in the Gazebo open-source robot simulator, and (3) controlling robotic platforms running ROS. Shepherd enables publishing and subscribing to ROS topics, navigating between waypoints, and performing image processing for peripherals such as RGB and depth cameras positioned on the robot or external platforms such as UAVs.

The communication between Shepherd and the robot is performed using a server-client software model, which enables controlling multiple robotic platforms in parallel (Fig. 3). This communication is achieved by implementing the Rosbridge library, which provides an API to ROS functionality for non-ROS programs [44], and Roslibpy – a library enabling robot control using Python and IronPython without running a ROS interface on the server [45].

Robot control using Shepherd is performed in one of three main approaches: (1) sending a direct speed and direction command (employing a velocity topic), (2) navigating to a specific point in space (using a goal topic), or (3) navigating along a path. This research implements the third approach, in which a curve in Rhinoceros 3D represents a path in space. The path is divided into a set of points that are translated to a list of goals iteratively sent to the robot. The robot arrival at a goal is determined in relation to a predefined distance, measured as a radius from the point on the curve. Following each successful arrival of the robot to the goal, it continues moving to the following position in the list.

During the study we focus on three types of trajectories: Slalom, Spiral, and Fork. We produced these trajectories using Shepherd, demonstrating the simplicity of using Shepherd to generate trajectories for robotic systems (Fig. 6).

B. SITE VISUALIZATION

We propose using top-view online images taken by a UAV to localize the UGV position on-site. The site localization is presented as a relative location of the UGV within the site. Typically, conservative localization is initially begun by defining the origin of the local coordinate system on the robot’s position with onboard sensors. However, using only wheel odometry with IMU leads to positioning errors which rapidly increase over time due to cumulative errors of the sensors. Therefore, we provide an approach to estimate the position of the UGV relatively to local static landmarks while avoiding drifting errors.

The UGV’s position measurement is based on prior knowledge about static landmarks’ location on the site. The origin is set in a location that allows continuous relative measurement of the UGV’s position throughout the operation. This measurement does not depend on the camera’s movement and allows defining measurements relative to the static landmarks. This means the process is invariant to the motion of the UAV, thus improving the localization robustness. Using the position of the landmark provides stable boundaries and features in the frame and

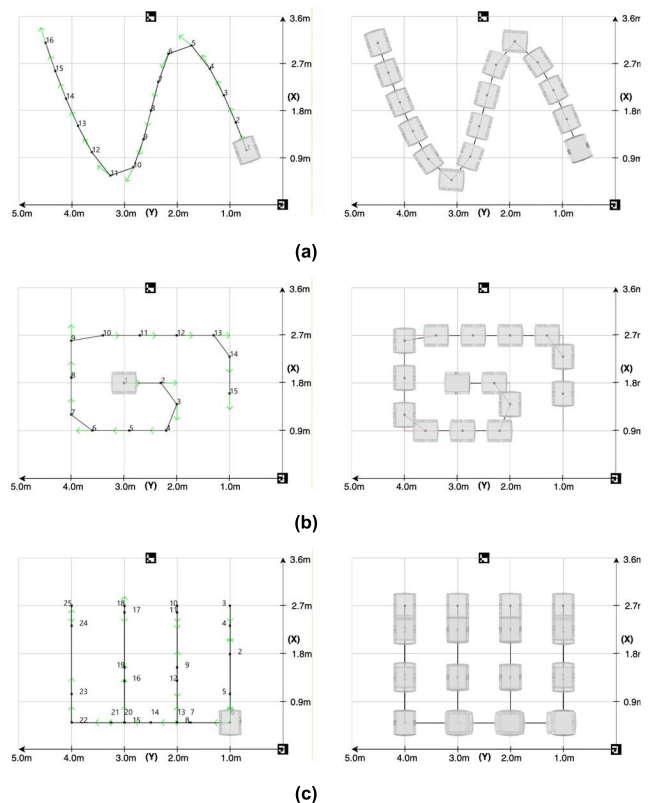


FIGURE 6. The Slalom, Spiral, and Fork trajectories generated using Shepherd in Grasshopper (left) and simulated in Rhinoceros 3D (right).

enables interpretation from pixel coordinates to the actual relative location of the UGV.

In the first step, the UAV needs to be placed where the top-view image contains the UGV and the landmarks. Then, the pixel coordinate location of the landmarks is transformed into a metric representation based on the knowledge of the actual location of the landmarks. The transformation between image coordinates to local frame coordinates is performed by using the scale transformation:

$$f_X = \frac{X_{L,l_i} - X_{L,l_j}}{x_{L,l_i} - x_{L,l_j}} \quad (1)$$

$$f_Y = \frac{Y_{L,l_i} - Y_{L,l_j}}{y_{L,l_i} - y_{L,l_j}} \quad (2)$$

where f defines the relative scale, X , YX , Y are the local plane landmarks coordinates, x , y are the image plane location, and l_i , l_j are two arbitrary static landmarks in the camera frame.

We then use the scale transformation to calculate the relative location of the UGV as follows:

$$X_L = (X_{UGV} - X_{L,l_i})f_X \quad (3)$$

$$Y_L = (Y_{UGV} - Y_{L,l_i})f_Y \quad (4)$$

where X_L , Y_L and X_L , Y_L are the site's local coordinates of the UGV.

C. MATERIAL MAPPING

Specific focus is given to developing a capacity for the material mapping process. Mapping the material enables tracking the gravel location during the site preparation task and thus improves the planning of the autonomous action. This method is implemented using data derived from the UAV site visualization by applying computer vision techniques that enable segmenting different materials in the construction site, as demonstrated in Fig. 8. This functionality allows monitoring the process advancement while considering the site S as an area that includes all the scattered material piles $S \in \{O_i | i=1, \dots, n\}$. The material pile is defined as a minimal bounding box according to $O_i = ((x_{min}, y_{min}), (x_{max}, y_{max}))$, where (x_{min}, y_{min}) is the top left corner of the pile and (x_{max}, y_{max}) is the bottom right corner. The initial mapping process operation involves a human operator to draw a bounding box O_{User} that contains the material. The bounding box functions as prior visual knowledge for the segmentation process and is defined as $O_{User} \in \{(R_i, G_i, B_i) | i \in 1, \dots, n\}$ where (R_i, G_i, B_i) are the red, green, blue channels respectively. Then we use O_{User} as a representative sample for recognizing the material.

Assuming a Gaussian distribution of the image color as $RGB \sim (\mu, \sigma^2)$, we calculate the mean color vector as follows:

$$\vec{\mu} = \frac{\sum RGB_i}{n} \quad (5)$$

where RGB_i is the value vector of colors in cell i , and $\vec{\mu}$ is the estimated mean color vector.

Then we calculate the color Standard Deviation (SD):

$$\vec{\sigma} = \sqrt{\frac{\sum (RGB_i - \vec{\mu})^2}{n}} \quad (6)$$

where $\vec{\sigma}$ is the vector of the color SD, which we use as an estimated color value for finding the material on-site. Then, manually chosen by the operator k_{low} , and k_{high} determines the segmented area according to $[\vec{\mu} - k_{low} \vec{\sigma}, \vec{\mu} + k_{high} \vec{\sigma}]$.

D. VISION-BASED SITE LOCALIZATION

In order to leverage the site localization for UGV state estimation, we integrate an EKF filter with the site localization data. The EKF was developed for a generic UGV. The dynamic model of the robot is presented as a set of non-linear equations:

$$x_t = f(u_t, x_{t-1}) + \omega_t \quad (7)$$

where state at time t is defined as x_t , the action command u_t , and the process noise is represented by ω_t . In the same manner, we describe the sensors' measurement as follows:

$$z_t = h(x_t) + e_t \quad (8)$$

where z_t define the measurement, $h(x_t)$ defines the measurement process, and e_t the measurement noise.

Both the process and the measurement noise are assumed as zero-mean multivariate Gaussian noises with the covariances Q_t and R_t respectively, written as:

$$\omega_t \sim \mathcal{N}(0, Q_t) \quad (9)$$

$$e_t \sim \mathcal{N}(0, R_T) \quad (10)$$

The EKF algorithm is a double-stage Gaussian filter approach that relies on the dynamic model and sensors measurements to estimate the robot state. The first stage of the EKF is called State Prediction. This stage calculates the predicted state x_t and estimates the corresponding process covariance according to the following equations:

$$x_{t|t-1}^p = f(x_{t-1|t-1}^c, u_t) \quad (11)$$

$$\Sigma_t^p = G_{t-1} \Sigma_t^c G_{t-1}^T + R_t \quad (12)$$

The Jacobian matrix G_{t-1} is set as $G_{t-1} = \frac{\partial f}{\partial x_{t-1|t-1}^c}$, and can be calculated using the first order Taylor expansion.

In the second stage, called State Correction, the EKF uses the sensors' data to correct the state prediction, which can be written as:

$$K_t = \Sigma_t^p H_t^T (H_t \Sigma_t^p H_t^T + Q_t)^{-1} \quad (13)$$

$$x_{t|t}^c = x_{t|t-1}^p + K_t (z_t - h(x_{t|t-1}^c)) \quad (14)$$

$$\Sigma_t^c = (I - K_t H_t) \Sigma_t^p \quad (15)$$

The observation matrix H_t is the partial derivatives of the measurement model with respect to the state variables, defined as $H_t = \frac{\partial h}{\partial x_{t|t-1}^c}$, the Kalman gain is K_t , and $x_{t|t}^c$, Σ_t^c are the corrected state and the covariance.

IV. AUTONOMOUS SITE PREPARATION

The following section describes the simulations and the experiments in detail. In both cases, the system is similar: a Clearpath Jackal UGV equipped with a custom 40 by 50 cm front shovel tool emulating a dozer blade and a Parrot Bebop 2 UAV equipped with a 1080p camera using a 180-degree wide-angle lens. The size of the worksite is 3.6 by 5 meters, as shown in Fig. 7.

A. SIMULATION SETUP

The software system communicates using ROS and is divided into simulative and experimental sub-systems. A realistic simulation was established based on the dynamic Gazebo engine, which compared the drone top-view localization performance with the actual experiments. The simulation consists of rough terrain produced using the Blender 3D modeling software alongside multiple aggregates in different

weights and sizes produced in Gazebo. These are used to illustrate the localization performance in a realistic scenario of autonomous site preparation with UGV. The baseline used in the setup is a simple IMU with wheel odometry for local EKF state estimation. This is then compared with the collaborative visual-based site-localization EKF. Both were implemented synchronously throughout the simulations and compared with the ground truth.

Using Shepherd, we designed three trajectories: Slalom, Spiral, and Fork. The Slalom demonstrates simple motion over the site requiring moderate orientation maneuvers. The Spiral trajectory demonstrates changing radius maneuvers over the entire site. Lastly, the Fork is the most challenging trajectory, as it requires performing sharp rotation maneuvers in place. These trajectories are shown in simulations and experiments in Fig. 7 as top view images of the UGV taken by the UAV.



FIGURE 7. The UGV performing the (a) Slalom, (b) Spiral, and (c) Fork trajectories simulated in Gazebo (left), and the experiments (right). Both are respectively shown as top view image sequences of the simulated and actual UAV.

B. EXPERIMENTAL SETUP

The experimental setup addresses the performance contribution of online communication with the drone. To isolate the additional factors of the estimation reliability, we implemented two similar EKF's, working synchronously, while only one of them is listening to the drone localization. The experiment was conducted on a high grip surface, causing skid steering vehicles to produce unreliable orientation estimations. The localization was examined on the same trajectories of the simulation (shown in Fig. 7).

The communication between the Jackal UGV, Shepherd, and the Bebop UAV is performed using the bebop_autonomy package [46], which provides an online wireless communication framework. The Jackal motion control is achieved by implementing the ROS navigation stack [47] MoveBase package, which provides the motion control for the Jackal. Fig. 7 shows the UGV performing the trajectories designed in Shepherd. Setting up the material mapping is divided into 3 stages: (1) sampling the material color for tracking, (2) defining the color segmentation bounds, and (3) selecting the region of interest (ROI) for mapping on the aerial image (Fig. 8). Following this, the UAV autonomously maps the material in the ROI and continuously updates the map which is sent to Shepherd. The material mapping was examined on the Fork trajectory since it is a strategy commonly used for earthmoving [48].

V. RESULTS

The evaluation of the localization has been performed with different initial conditions. A hundred simulations tested each trajectory with various aggregate configurations to illustrate the moving construction site materials. The results of the simulations were summarized in Fig. 9 representing the mean of 100 runs $\bar{x} = \frac{\sum x_i}{100}$. During the runs, the error of each measurement was calculated as $\bar{e} = \bar{X} - \bar{x}$ while \bar{X} is the ground truth. The results for the trajectories experiment presented in Fig. 10 show a significant difference between the EKF estimator using the drone measurements and the baseline EKF. The poor results of the baseline

EKF in comparison to the site-localization EKF are caused by several reasons: (1) The slip of the wheels causing idle wheel rotation, (2) the estimated wheel radius is not accurate, (3) the terrain is not known and assumed as a 2D plate, and (4) the interaction with the aggregates - causing a change in the dynamics of the UGV motion. The estimated trajectory of the EKF's is represented in Fig. 9 (d1-d3).

The results of the baseline EKF show a shorter path than the ground truth, inaccurate orientation estimation, and accumulation of errors. The shorter path may be caused by an effective wheel radius assumption in the estimated model, which results in an unreliable motion model. The orientation error is a result of wheel slip and rough terrain, which causes an idle rotation of the wheels. These errors lead to error accumulation in the baseline EKF during the operation. In contrast, the site-localization EKF is overcomes these errors and performs accurate localization during operation with an absolute error of less than 0.2m. This is implemented using the site visualization from the UAV, creating the site-localization EKF.

In contrast to the simulation, the UGV performed the experiment only once for each trajectory, in which we continuously repeat the trajectory multiple times. The experiments presented in Fig. 10 validate the results obtained in the simulation. We ran the UGV along the trajectory multiple times during the experiment while comparing the estimators.

The site-localization estimator shows accurate results similar to the simulation, while the baseline EKF diverges. In contrast to the simulation, the experiment was implemented on a high grip plate terrain. The results present an increased error in the orientation, as expected from the skid steering platform, which depends on slipping for rotation maneuvers. But unlike the simulations, the estimated path of the baseline EKF is longer, which may cause by an error in wheel radius, wheel slipping, or inaccurate velocity estimation. The site-localization EKF presents stable and accurate UGV localization estimation during

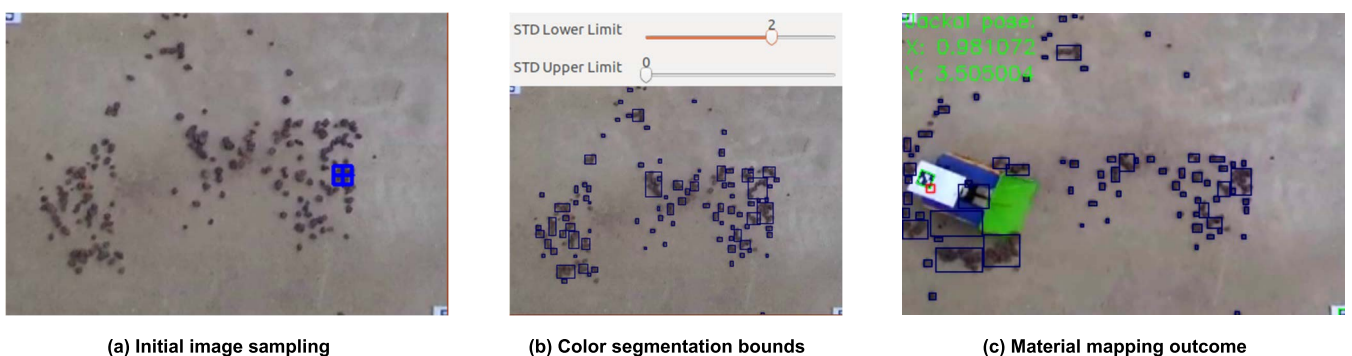


FIGURE 8. The setup stages of the material mapping are displayed in a preliminary experiment: (a) an initial sample of the material color for tracking, (b) defining the color segmentation bounds, and (c) autonomous mapping by the UAV within the region of interest for mapping.

the runs in the experiments (see supplementary video for reference).

Incorporating the material mapping capacity in the experimental stage enabled real-time tracking of the specific

location and relative dispersion of the aggregates on-site (Fig. II). Therefore, it provides an account of the progression of the task and can potentially indicate whether it is successful or not according to predefined measures.

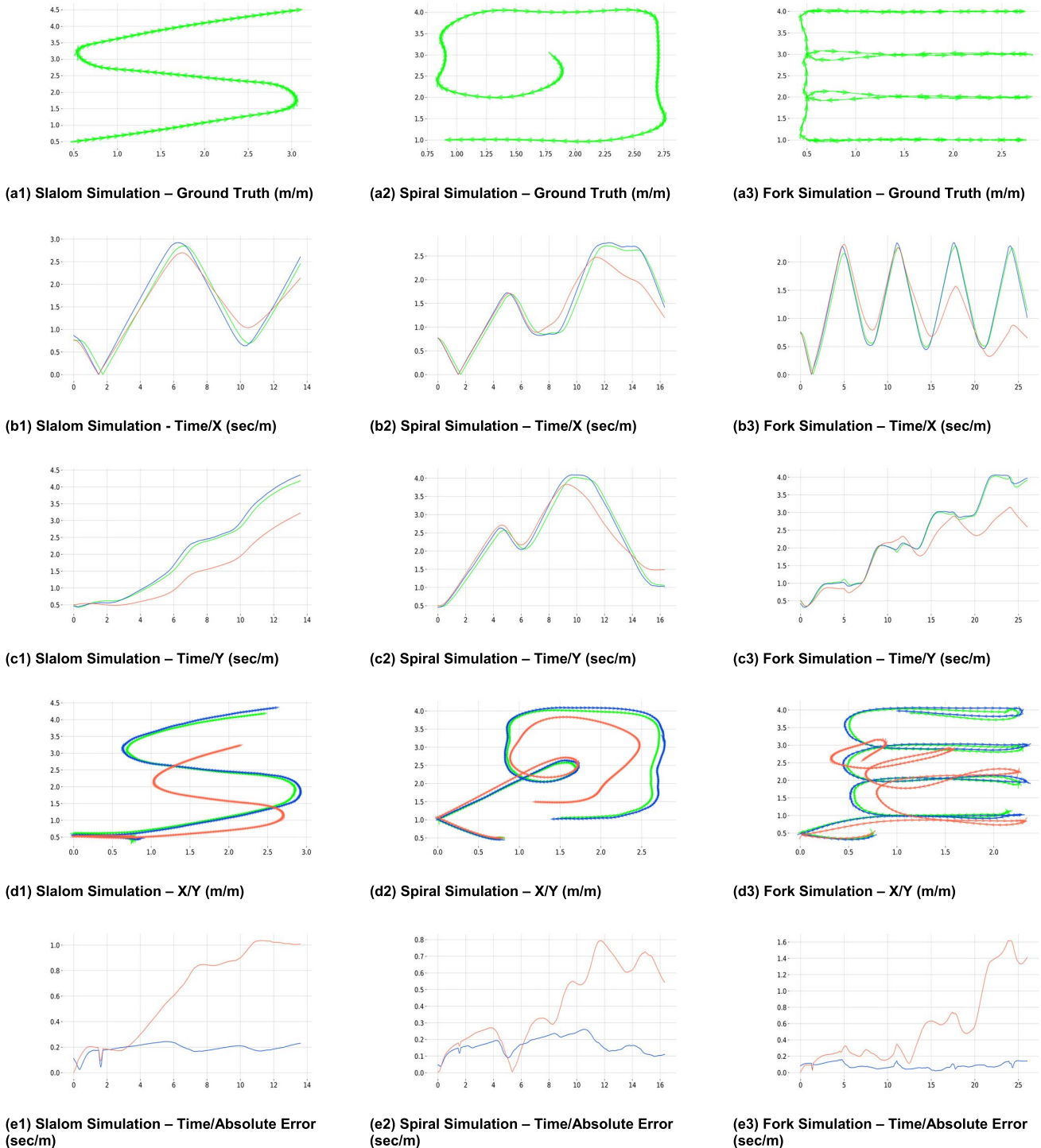


FIGURE 9. The results of the simulations, for Slalom (a1-e1), Spiral (a2-e2), and Fork trajectories (a3-e3). Presenting the results of the baseline EKF in red color, Site-localization EKF in blue color, and the ground-truth in green color.

VI. LIMITATION AND CONCLUSIONS

As shown in Section V, the collaborative site visualization significantly improves the EKF estimation. As presented in the results, the maximal error of the site-localization EKF is 0.2m. Additionally, the suggested method allows for overcoming fault wheel radius, wheel slipping, and inaccurate velocity estimation of the UGV.

Nevertheless, additional factors should also be considered, specifically regarding the UAV’s flight. These factors include drifting in the location of the UAV due to external forces and limited flight time. In this context, as long as the UGV and the landmarks are kept inside the UAV’s camera frame, the site visualization is invariant to shifts in the UAV’s location. This is demonstrated in the experiments by the successful localization of the UGV despite wind disturbances. Lastly, as the task duration is limited by the UAV’s flight time, its use is preferable to a static camera since it can cover larger areas while locating multiple UGVs across the site.

VII. FUTURE WORK

As this is ongoing research, the methods and tools are continuously developed. Future work will therefore focus on addressing current limitations, improving the feedback in the system, and expanding its collaborative capabilities.

As construction site preparation tasks can benefit from increasing the accuracy of the outcome with respect to the desired location and shape of the material, a future iteration of this research will include adaptive material shaping. In this process, the UGV is given a placement task - arranging material in specific shapes. This goal requires understanding the material location after each pushing sequence and adapting the path accordingly. The material location changes throughout the process, and thus the material mapping method updates the map for the path planner before each pushing sequence.

While currently the research assumes that the site is free from static obstacles, future work will focus on safe path planning in environments that contain static and dynamic obstacles such as existing infrastructure, human collaborators, or other vehicles. Additionally, while the UAV is stationary in the presented experiments, there is no technical limitation regarding the suggested algorithm for moving the UAV during the experiments. Therefore, future work will incorporate a “smart camera” method in which an optimal path for the UAV will be explored. Furthermore, future research will expand the collaborative capabilities of the system by exploring the use of multiple UGV’s to perform a site preparation task. This will require developing strategies for distributing the task between the platforms and is expected

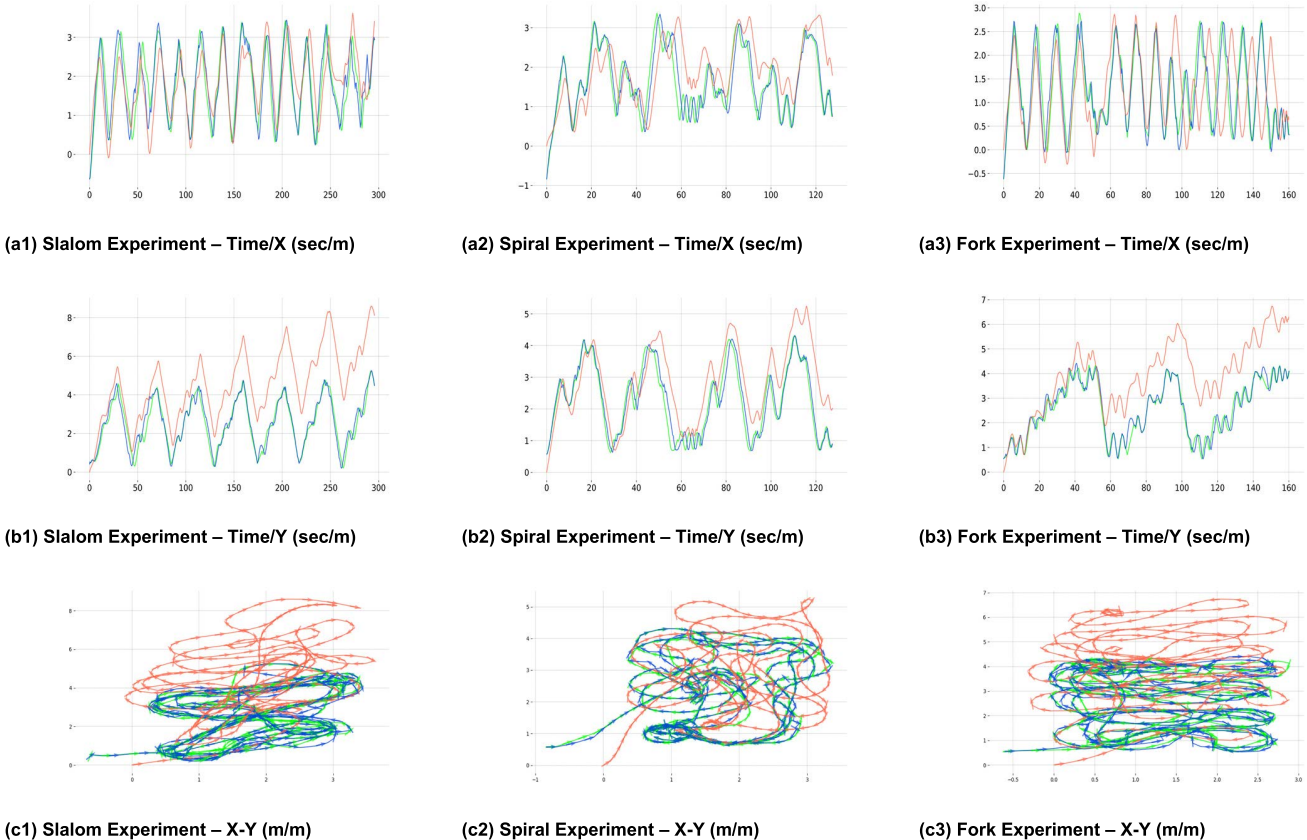


FIGURE 10. The results of the experiments, for Slalom (a1-c1), Fork (a2-c2), and Fork trajectories (a3-c3). Presenting the results of the baseline EKF in red color, Site-localization EKF in blue color, and the ground-truth in green color.

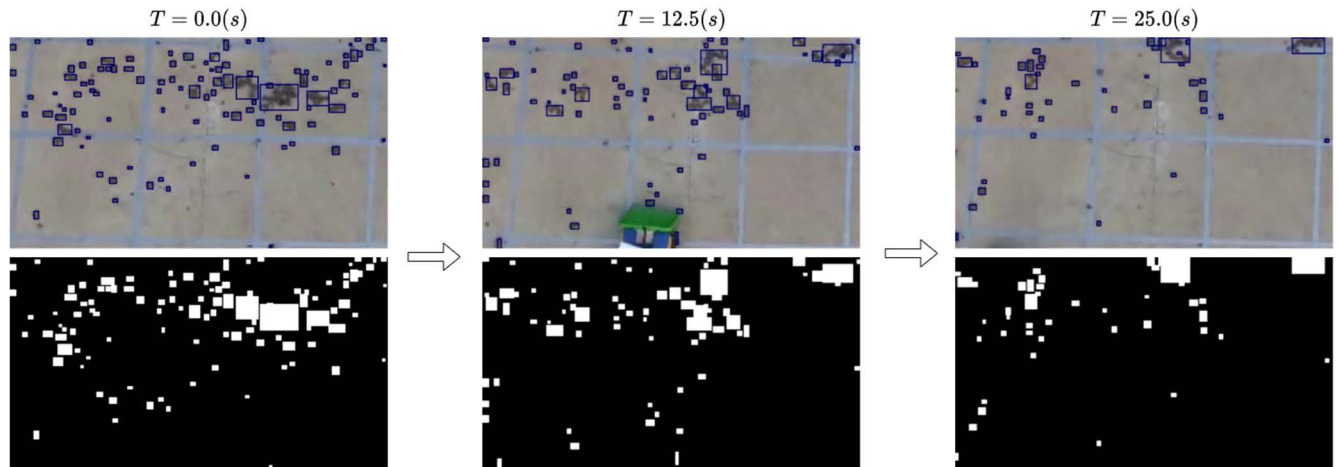


FIGURE 11. The results of the material mapping process during the experiment: on the top is the site with the generated bounding boxes of the aggregate material, and on the bottom is the material binary grid.

to increase the resilience of the entire system by providing redundancy and reducing the time it takes to perform a similar task using only a single UGV.

REFERENCES

- [1] H. Karimi, T. R. B. Taylor, G. B. Dadi, P. M. Goodrum, and C. Srinivasan, "Impact of skilled labor availability on construction project cost performance," *J. Construct. Eng. Manage.*, vol. 144, no. 7, 2018, Art. no. 04018057, doi: [10.1061/\(asce\)co.1943-7862.0001512](https://doi.org/10.1061/(asce)co.1943-7862.0001512).
- [2] 2020 Sage Construction Hiring and Business Outlook Survey | Associated General Contractors of America. Accessed: Jul. 31, 2021. [Online]. Available: <https://www.agc.org/news/2019/12/18/2020-sage-construction-hiring-and-business-outlook-survey>
- [3] A. Alsharif, S. Banerjee, J. Uddin, A. Albert, and E. Jaselskis, "Pandemic on the United States construction industry," *Int. J. Environ. Res. Public Health*, vol. 18, no. 4, p. 1559, Feb. 2021, doi: [10.3390/ijerph18041559](https://doi.org/10.3390/ijerph18041559).
- [4] F. Augugliaro, S. Lupashin, M. Hamer, C. Male, M. Hehn, M. W. Mueller, J. S. Willmann, F. Gramazio, M. Kohler, and R. D'Andrea, "The flight assembled architecture installation: Cooperative construction with flying machines," *IEEE Control Syst.*, vol. 34, no. 4, pp. 46–64, Aug. 2014, doi: [10.1109/MCS.2014.2320359](https://doi.org/10.1109/MCS.2014.2320359).
- [5] D. Jud, I. Hurkxkens, C. Giroto, and M. Hutter, "Robotic embankment," *Construct. Robot.*, vol. 1, pp. 1–13, Jun. 2021, doi: [10.1007/s41693-021-00061-0](https://doi.org/10.1007/s41693-021-00061-0).
- [6] M. Lussi, T. Sandy, K. Dorfler, N. Hack, F. Gramazio, M. Kohler, and J. Buchli, "Accurate and adaptive *in situ* fabrication of an undulated wall using an on-board visual sensing system," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 3532–3539, doi: [10.1109/ICRA.2018.8460480](https://doi.org/10.1109/ICRA.2018.8460480).
- [7] T. Shaked, K. L. Bar-Sinai, and A. Sprecher, "Adaptive robotic stone carving: Method, tools, and experiments," *Autom. Construct.*, vol. 129, Sep. 2021, Art. no. 103809, doi: [10.1016/j.autcon.2021.103809](https://doi.org/10.1016/j.autcon.2021.103809).
- [8] R. L. Peurifoy, W. B. Ledbetter, C. J. Schexnayder, R. L. Schmitt, and A. Shapira, *Construction Planning, Equipment, and Methods*. New York, NY, USA: McGraw-Hill, 2018.
- [9] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges," *Comput. Commun.*, vol. 149, pp. 270–299, Jan. 2020, doi: [10.1016/j.comcom.2019.10.014](https://doi.org/10.1016/j.comcom.2019.10.014).
- [10] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: State-of-the-art in applications and challenges," *J. Netw. Comput. Appl.*, vol. 166, Sep. 2020, Art. no. 102706, doi: [10.1016/j.jnca.2020.102706](https://doi.org/10.1016/j.jnca.2020.102706).
- [11] J. Stüber, C. Zito, and R. Stolkin, "Let's push things forward: A survey on robot pushing," *Frontiers Robot. AI*, vol. 7, pp. 1–18, Feb. 2020, doi: [10.3389/frobt.2020.00008](https://doi.org/10.3389/frobt.2020.00008).
- [12] T. Shaked and A. Degani, "Shepherd: A fabrication-oriented tool for simulation and control of mobile robotic platforms for collaborative earthworks," in *Proc. ACADIA Realignments, Toward Crit. Comput.*, 2021.
- [13] K. H. Petersen, N. Napp, R. Stuart-Smith, D. Rus, and M. Kovac, "A review of collective robotic construction," *Sci. Robot.*, vol. 4, no. 28, Mar. 2019, Art. no. eaau8479, doi: [10.1126/SCIROBOTICS.AAU8479](https://doi.org/10.1126/SCIROBOTICS.AAU8479).
- [14] M. Saboia, V. Thangavelu, and N. Napp, "Autonomous multi-material construction with a heterogeneous robot team," *Robot. Auton. Syst.*, vol. 121, Nov. 2019, Art. no. 103239, doi: [10.1016/J.ROBOT.2019.07.009](https://doi.org/10.1016/J.ROBOT.2019.07.009).
- [15] R. L. Johns, M. Wermelinger, R. Mascaro, D. Jud, F. Gramazio, M. Kohler, M. Chli, and M. Hutter, "Autonomous dry stone: On-site planning and assembly of stone walls with a robotic excavator," *Construct. Robot.*, vol. 4, nos. 3–4, pp. 127–140, Dec. 2020, doi: [10.1007/s41693-020-00037-6](https://doi.org/10.1007/s41693-020-00037-6).
- [16] A. Rasul, J. Seo, K. Oh, A. Khajepour, and N. Reginald, "Predicted safety algorithms for autonomous excavators using a 3D LiDAR sensor," in *Proc. IEEE Int. Syst. Conf. (SysCon)*, Aug. 2020, pp. 1–7, doi: [10.1109/SYSCON47679.2020.9275854](https://doi.org/10.1109/SYSCON47679.2020.9275854).
- [17] D. Jud, G. Hottiger, P. Leemann, and M. Hutter, "Planning and control for autonomous excavation," *IEEE Robot. Autom. Lett.*, vol. 2, no. 4, pp. 2151–2158, Oct. 2017, doi: [10.1109/LRA.2017.2721551](https://doi.org/10.1109/LRA.2017.2721551).
- [18] H. G. Tanner, "Switched UAV-UGV cooperation scheme for target detection," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 3457–3462, doi: [10.1109/ROBOT.2007.364007](https://doi.org/10.1109/ROBOT.2007.364007).
- [19] E. H. C. Harik, F. Guerin, F. Guinand, J.-F. Brethe, and H. Pelvillain, "UAV-UGV cooperation for objects transportation in an industrial area," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Mar. 2015, pp. 547–552, doi: [10.1109/ICIT.2015.7125156](https://doi.org/10.1109/ICIT.2015.7125156).
- [20] A. Lakas, B. Belkhouche, O. Benkraouda, A. Shuaib, and H. J. Alasmawi, "A framework for a cooperative UAV-UGV system for path discovery and planning," in *Proc. Int. Conf. Innov. Inf. Technol. (IIT)*, Nov. 2018, pp. 42–46, doi: [10.1109/INNOVATIONS.2018.8606028](https://doi.org/10.1109/INNOVATIONS.2018.8606028).
- [21] S. Hood, K. Benson, P. Hamod, D. Madison, J. M. O'Kane, and I. Kleititis, "Bird's eye view: Cooperative exploration by UGV and UAV," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2017, pp. 247–255, doi: [10.1109/ICUAS.2017.7991513](https://doi.org/10.1109/ICUAS.2017.7991513).
- [22] M. H. Mughal, M. J. Khokhar, and M. Shahzad, "Assisting UAV localization via deep contextual image matching," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2445–2457, 2021, doi: [10.1109/JSTARS.2021.3054832](https://doi.org/10.1109/JSTARS.2021.3054832).
- [23] Y.-J. Zheng, Y.-C. Du, H.-F. Ling, W.-G. Sheng, and S.-Y. Chen, "Evolutionary collaborative human-UAV search for escaped criminals," *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 217–231, Apr. 2020, doi: [10.1109/TEVC.2019.2925175](https://doi.org/10.1109/TEVC.2019.2925175).
- [24] Y.-J. Zheng, Y.-C. Du, Z.-L. Su, H.-F. Ling, M.-X. Zhang, and S.-Y. Chen, "Evolutionary human-UAV cooperation for transmission network restoration," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1648–1657, Mar. 2021, doi: [10.1109/TII.2020.3003903](https://doi.org/10.1109/TII.2020.3003903).
- [25] L. F. Bertuccelli and M. L. Cummings, "Scenario-based robust scheduling for collaborative human-UAV visual search tasks," in *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, Dec. 2011, pp. 5702–5707, doi: [10.1109/CDC.2011.6160994](https://doi.org/10.1109/CDC.2011.6160994).

- [26] F. Liu, X. Li, S. Yuan, and W. Lan, "Slip-aware motion estimation for off-road mobile robots via multi-innovation unscented Kalman filter," *IEEE Access*, vol. 8, pp. 43482–43496, 2020, doi: [10.1109/ACCESS.2020.2977889](https://doi.org/10.1109/ACCESS.2020.2977889).
- [27] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).
- [28] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571, doi: [10.1109/ICCV.2011.6126544](https://doi.org/10.1109/ICCV.2011.6126544).
- [29] H. Zhou, H. Fan, K. Peng, W. Fan, D. Zhou, and Y. Liu, "Monocular visual odometry initialization with points and line segments," *IEEE Access*, vol. 7, pp. 73120–73130, 2019, doi: [10.1109/ACCESS.2019.2920453](https://doi.org/10.1109/ACCESS.2019.2920453).
- [30] S. Misra, B. Wang, K. Sundar, R. Sharma, and S. Rathinam, "Single vehicle localization and routing in GPS-denied environments using range-only measurements," *IEEE Access*, vol. 8, pp. 31004–31017, 2020, doi: [10.1109/ACCESS.2019.2963286](https://doi.org/10.1109/ACCESS.2019.2963286).
- [31] M. B. Alatise and G. P. Hancke, "A review on challenges of autonomous mobile robot and sensor fusion methods," *IEEE Access*, vol. 8, pp. 39830–39846, 2020, doi: [10.1109/ACCESS.2020.2975643](https://doi.org/10.1109/ACCESS.2020.2975643).
- [32] O. Shalev and A. Degani, "Canopy-based Monte Carlo localization in orchards using top-view imagery," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2403–2410, Apr. 2020, doi: [10.1109/LRA.2020.2970975](https://doi.org/10.1109/LRA.2020.2970975).
- [33] A. Yol, B. Delabarre, A. Dame, J.-E. Dartois, and E. Marchand, "Vision-based absolute localization for unmanned aerial vehicles," in *Proc. IEEE/RSS Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 3429–3434, doi: [10.1109/IROS.2014.6943040](https://doi.org/10.1109/IROS.2014.6943040).
- [34] Y. Xu, L. Pan, C. Du, J. Li, N. Jing, and J. Wu, "Vision-based UAVs aerial image localization: A survey," in *Proc. 2nd ACM SIGSPATIAL Int. Workshop AI Geograph. Knowl. Discovery*, Nov. 2018, pp. 54–63, doi: [10.1145/3281548.3281556](https://doi.org/10.1145/3281548.3281556).
- [35] T. Sandy, M. Gifthalder, K. Dörfler, M. Kohler, and J. Buchli, "Autonomous repositioning and localization of an *in situ* fabricator," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2016, pp. 2852–2858.
- [36] R. Diversi, R. Guidorzi, and U. Soverini, "Kalman filtering in extended noise environments," *IEEE Trans. Autom. Control*, vol. 50, no. 9, pp. 1396–1402, Sep. 2005, doi: [10.1109/TAC.2005.854627](https://doi.org/10.1109/TAC.2005.854627).
- [37] S. Yang and M. Baum, "Extended Kalman filter for extended object tracking," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 4386–4390, doi: [10.1109/ICASSP.2017.7952985](https://doi.org/10.1109/ICASSP.2017.7952985).
- [38] H. Jo and E. Kim, "New Monte Carlo localization using deep initialization: A three-dimensional LiDAR and a camera fusion approach," *IEEE Access*, vol. 8, pp. 74485–74496, 2020, doi: [10.1109/ACCESS.2020.2988464](https://doi.org/10.1109/ACCESS.2020.2988464).
- [39] M. Krizmancic, B. Arbanas, T. Petrovic, F. Petric, and S. Bogdan, "Cooperative aerial-ground multi-robot system for automated construction tasks," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 798–805, Apr. 2020, doi: [10.1109/LRA.2020.2965855](https://doi.org/10.1109/LRA.2020.2965855).
- [40] W. Gao, J. Luo, W. Zhang, W. Yuan, and Z. Liao, "Commanding cooperative UGV-UAV with nested vehicle routing for emergency resource delivery," *IEEE Access*, vol. 8, pp. 215691–215704, 2020, doi: [10.1109/ACCESS.2020.3040790](https://doi.org/10.1109/ACCESS.2020.3040790).
- [41] H. Kandath, T. Bera, B. Bardhan, and S. Sundaram, "Autonomous navigation and sensorless obstacle avoidance for UGV with environment information from UAV," in *Proc. 2nd IEEE Int. Conf. Robotic Comput. (IRC)*, Jan. 2018, pp. 266–269, doi: [10.1109/IRC.2018.00056](https://doi.org/10.1109/IRC.2018.00056).
- [42] D. Erez, S. Arzeti, and D. Zarruck, "A novel simple two-robot precise self-localization method," *IEEE Access*, vol. 7, pp. 154044–154055, 2019, doi: [10.1109/ACCESS.2019.2944222](https://doi.org/10.1109/ACCESS.2019.2944222).
- [43] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.*, vol. 47, no. 6, pp. 2280–2292, Jun. 2014, doi: [10.1016/j.patcog.2014.01.005](https://doi.org/10.1016/j.patcog.2014.01.005).
- [44] C. Crick, G. Jay, S. Osentoski, B. Pitzer, and O. C. Jenkins, "Rosbridge: ROS for non-ROS users," in *Proc. 15th Int. Symp. Robot. Res. (ISRR)*, H. I. Christensen and O. Khatib, Eds. Cham, Switzerland: Springer, 2017, pp. 493–504, doi: [10.1007/978-3-319-29363-9_28](https://doi.org/10.1007/978-3-319-29363-9_28).
- [45] G. Casas, M. Lütke, B. Lytle, A. Jeandeau, and H. Obinata. (2019). *Roslibpy: ROS Bridge Library*. Accessed: Jul. 1, 2021. [Online]. Available: <https://roslibpy.readthedocs.io/en/latest/>
- [46] *Bebop_Autonomy—ROS Driver for Parrot Bebop Drone (Quadrocopter) 1.0 & 2.0—Bebop_Autonomy Indigo-Devel Documentation*. Accessed: Nov. 12, 2021. [Online]. Available: <https://bebop-autonomy.readthedocs.io/en/latest/>
- [47] E. Marder-Eppstein, E. Berger, T. Foote, B. Gerkey, and K. Konolige, "The office marathon: Robust navigation in an indoor office environment," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2010, pp. 300–307, doi: [10.1109/ROBOT.2010.5509725](https://doi.org/10.1109/ROBOT.2010.5509725).
- [48] M. Hirayama, J. Guivant, J. Katupitiya, and M. Whitty, "Path planning for autonomous bulldozers," *Mechatronics*, vol. 58, pp. 20–38, Apr. 2019, doi: [10.1016/j.mechatronics.2019.01.001](https://doi.org/10.1016/j.mechatronics.2019.01.001).



OREN ELMAKIS received the B.Sc. degree in mechanical engineering from the Technion – Israel Institute of Technology, Haifa, Israel, in 2019, where he is currently pursuing the Ph.D. degree with the Civil, Environmental, and Agricultural Robotics Laboratory (CEAR).

His research interests include robotic multi-agents decision making, motion control, and collaborative state estimation in civil and environmental applications.



TOM SHAKED received the B.Arch. and M.A. degrees from Tel-Aviv University, in 2012 and 2014, respectively, and the Ph.D. degree from the Technion – Israel Institute of Technology, in 2021, focusing on developing robotic tools for adaptive fabrication in construction.

He was a Lecturer and a Digital Fabrication Laboratory Manager with Tel-Aviv University. He co-founded CTwiz real estate analytics start-up and was the Head of Education at Infolio, an Apple

Mobility Partner Company (2014–2016). He was a Lead Researcher at The David & Barbara Blumenthal Israel Center for Innovation and Research in Textiles (CIRTex), Shenkar College of Engineering (2016–2017). He is currently a Visiting Lecturer at the School of Engineering and Design, Technical University of Munich (TUM) and a Postdoctoral Researcher at the Civil, Environmental, and Agricultural Robotics Laboratory (CEAR), Faculty of Civil and Environmental Engineering, Technion IIT. His research interests include collective robotic construction and autonomous systems in unstructured environments. He was a two-time recipient of the Azrieli Research Fellowship in 2012 and 2018. He is a Registered Architect.



AMIR DEGANI (Member, IEEE) received the B.Sc. degree (*summa cum laude*) in mechanical engineering from the Technion – Israel Institute of Technology, Haifa, Israel, in 2002, and the M.S. and Ph.D. degrees in robotics from Carnegie Mellon University, Pittsburgh, PA, USA, in 2006 and 2010, respectively. From 2011 to 2019, he was an Assistant Professor of civil and environmental engineering, Technion, and the Director of the Civil, Environmental and Agricultural

Robotics (CEAR) Laboratory researching robotic legged locomotion and autonomous systems in civil and agriculture applications. In 2019, he was promoted to a Tenured Associate Professor. He is currently an Associate Professor at the Technion – Israel Institute of Technology. His research interests include mechanism analysis, synthesis, control and motion planning and design with emphasis on minimalistic concepts and the study of nonlinear dynamic hybrid systems. He has six patents in the robotics field and has received the Best Paper Award at the IEEE BioRob Conference in 2006, the Best Video Award at the IEEE ICRA Conference in 2010, and the JTCF Novel Technology Paper Award in IEEE IROS 2015. He is also an Associate Editor of the IEEE ROBOTICS AND AUTOMATION LETTERS. He has been an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS and of the IEEE ICRA and IROS conferences.

...