

A Shared-Autonomy Construction Robotic System for Overhead Works

David Minkwan Kim^{*1}, K. M. Brian Lee^{*2}, Yong Hyeok Seo^{*1}, Nikola Raicevic^{*2}, Runfa Blark Li^{*2}, Kehan Long², Chan Seon Yoon³, Dong Min Kang³, Byeong Jo Lim³, Young Pyoung Kim³, Nikolay Atanasov², Truong Nguyen², Se Woong Jun¹, Young Wook Kim¹

Abstract—We present the ongoing development of a robotic system for overhead work such as ceiling drilling. The hardware platform comprises a mobile base with a two-stage lift, on which a bimanual torso is mounted with a custom-designed drilling end effector and RGB-D cameras. To support teleoperation in dynamic environments with limited visibility, we use Gaussian splatting for online 3D reconstruction and introduce motion parameters to model moving objects. For safe operation around dynamic obstacles, we developed a neural configuration-space barrier approach for planning and control. Initial feasibility studies demonstrate the capability of the hardware in drilling, bolting, and anchoring, and the software in safe teleoperation in a dynamic environment.

I. INTRODUCTION

Construction faces worsening labor shortages due to the tough and risky nature of the job [1], often referred to as “dirty, dull, and dangerous (DDD)” tasks. This perception contributes to an aging workforce and, in turn, a higher risk of injuries [2]. A growing number of commercial systems, such as the Hilti Jaibot [3] and CSC Robo Drillcorprio [4], are being developed to assist construction work. We aim to advance this effort by deploying state-of-the-art techniques from mobile and field robotics to practical systems that improve job safety and comfort.

This paper introduces the ongoing development of a robotic system designed for overhead ceiling work, one of the most physically straining jobs in construction. We developed a mobile lift system (Fig. 1) equipped with a torso that carries up to two manipulators, each with a drilling end effector and sensing payloads to perform overhead tasks such as drilling on ceilings or attaching light fixtures. A teleoperation system is designed for intuitive operation, to best utilize the expertise of experienced on-site workers in robot operation.

A key challenge is ensuring safety and situational awareness when teleoperating in real construction sites with many moving objects and occlusions. For situational awareness, we use our prior work on DynaGSLAM [5] to build a 3D Gaussian splat (3DGS) reconstruction of dynamic environments

This work was supported by the Ministry of Trade, Industry and Energy (MOTIE), Korea, under the Strategic Technology Development Program, supervised by the Korea Institute for Advancement of Technology (KIAT) [Grant No. P0026052]. *These authors contributed equally to this work.

¹Intelligent Robotics Research Center, Korea Electronics Technology Institute (KETI), Seongnam, Korea. {kmk9846, syh4661, daniel.kimyo}@keti.re.kr

²Dept. of Electrical and Computer Engineering, University of California, San Diego (UCSD), La Jolla, CA, USA. {kmblee, nmarinoraitsevs, runfa, k3long, natanasov, tqn001}@ucsd.edu

³ITONE Inc., Seoul, Korea. {parcae, kangdm, limbj, mrkims}@it-1.kr

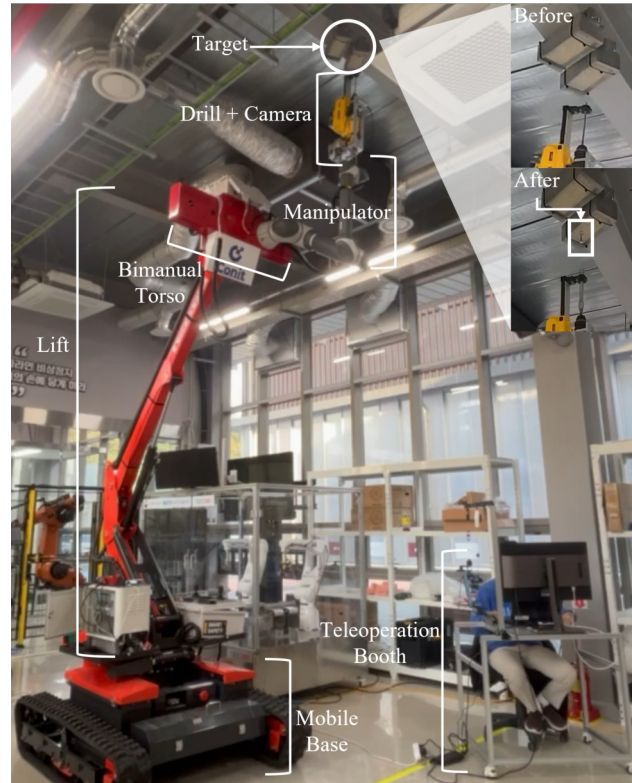


Fig. 1: Our system performing drilling, bolting, and anchoring tasks in a laboratory setting. Video available¹.

to allow novel-view synthesis for beyond-visual-line-of-sight (BVLOS) teleoperation as is common in aerial robots [6]. For safe operation in dynamic environments, our prior work in neural configuration-space barriers (NCSB) [7] is used to plan joint trajectories to desired end-effector poses, and to ‘filter’ both teleoperation and autonomy commands for collision avoidance against dynamic obstacles.

We present an initial feasibility study with a single-arm prototype that shows the capability of the hardware platform to drill, anchor and bolt on ceilings. Further experimental results demonstrate safe planning and control in dynamic environments for collision avoidance using NCSB, and 3D reconstruction using DynaGSLAM. We view the main contribution of this paper as the initial demonstration of state-of-the-art robotics techniques applied in a practical construction robotic system.

¹Video: www.kmblee.dev/videos/construction_workshop

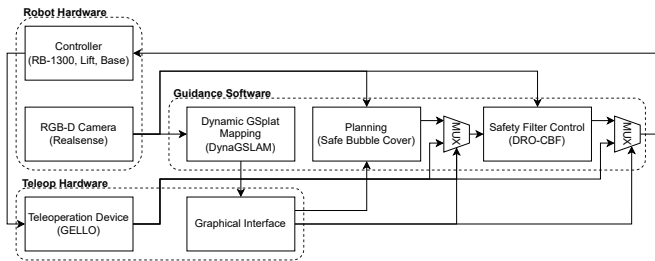


Fig. 2: System overview

II. HARDWARE PLATFORM

We designed and built a robotic system and a teleoperation booth as shown in Fig. 1 to enable BVLOS teleoperation for construction tasks. We briefly describe the system.

A. Robotic System

The robot hardware consists of a mobile base, a two-stage lift, and a manipulator. The two-stage lift enables vertical motion of a torso from 437 mm to 3,680 mm. The torso is designed to accept up to two manipulators. Currently, a Rainbow Robotics RB10-1300 manipulator is horizontally mounted on the torso and supports up to 10 kg of payload. A custom end-effector (Fig. 3) is attached for drilling, bolting, and anchoring, with a damping module placed between the drill and manipulator to reduce load impact. To support teleoperation, a pan-tilt VR camera on top of the torso provides real-time 3D visualization, while a RealSense D435 camera above the drill offers local visual feedback during tasks. The mobile base measures 2,660 mm \times 1,350 mm \times 1,710 mm and reaches speeds up to 2.4 km/h including the lift and manipulator. Both the base and lift can be remotely controlled via UART communication. The system is optimized for ultra-low-latency control.

B. Teleoperation Booth

The teleoperation booth is equipped with an operator PC, a monitor for visual inspection of the site, as well as a haptic device for control. The control commands are transmitted from the operator PC to the robot PC via WebRTC [8] over a 5G network, enabling low-latency operation even in BVLOS settings. For intuitive manipulation, the monitor offers photorealistic visual feedback of the site using DynaGSLAM [5] (Sec. III-A), as well as physical feedback and commands through a haptic device based on GELLO [9] (Sec. III-C).

III. SOFTWARE

An overview of the system is shown in Fig. 2. Given RGB-D images from the Realsense D435 camera, we construct a 3D Gaussian splat for visual teleoperation using DynaGSLAM [5], which is used for visual feedback in operator graphical interface. An NCSB [7] is constructed from the RGB-D images for safe planning and control.

A. Dynamic Gaussian Splat Mapping

We use DynaGSLAM [5] from our prior work to provide photorealistic visual feedback of dynamic, cluttered construction sites. Given RGB-D images of a dynamic scene

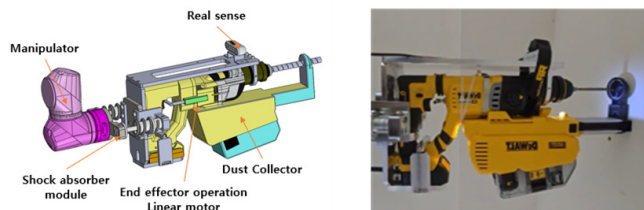


Fig. 3: A custom drilling end-effector assembly.

from the Realsense D-435 camera, DynaGSLAM reconstructs the scene as a set of Gaussian blobs defined as $\mathcal{G} = \{(\mathbf{m}_i(t), \Sigma_i, \alpha_i, \mathbf{sh}_i)\}$, with means $\mathbf{m}_i(t)$, covariances Σ_i , transparency α_i , and spherical harmonics coefficients \mathbf{sh}_i . Importantly, DynaGSLAM produces photorealistic rendering of moving objects by formulating the means $\mathbf{m}_i(t)$ as time-varying and modeling their temporal evolution using cubic Hermite splines. For computational efficiency, the interpolation parameters of $\mathbf{m}_i(t)$ are directly initialized using a novel management strategy [5, Sec. 5.2] that fuses previous interpolation parameters with the current optical flow and RGB-D pointcloud. The Gaussians are rendered and optimized against the RGB-D images using alpha blending:

$$I(\mathbf{v}, t) = \sum_{g_i \in \mathcal{G}(t)} c_i f(\mathbf{v}, g_i) \prod_i (1 - f(\mathbf{v}, g_i)), \quad (1)$$

where c_i is the color of Gaussian g_i given spherical harmonics \mathbf{sh}_i , and $f(\mathbf{v}, g_i) = \alpha_i N_{2D}(\mathbf{v}; P\mathbf{m}_i(t), P\Sigma_i P^T)$ is the contribution of Gaussian g_i at pixel \mathbf{v} and time t . P is the affine projection transform to the image plane, and the Gaussians are ordered relative to the query viewpoint.

B. Neural Configuration-Space Barrier

Given an RGB-D pointcloud \mathcal{P} , we consider the problem of planning and executing a manipulator trajectory that achieves a desired end-effector pose while avoiding dynamic obstacles. This safety requirement is encoded as an NCSB, which is defined using the configuration-space distance field (CDF) [10]. Let $\mathbf{q} \in \mathbb{R}^N$ be the manipulator's configuration, and $\mathcal{B}(\mathbf{q}) \subset \mathbb{R}^3$ the manipulator's geometry in the workspace. The CDF $d : \mathbb{R}^N \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is the closest distance in the configuration space (i.e. joint angles) to collision with a workspace point [10]:

$$d(\mathbf{q}, \mathbf{p}) = \min_{\mathbf{q}^*} \|\mathbf{q} - \mathbf{q}^*\|, \text{ s.t. } \mathbf{p} \in \partial\mathcal{B}(\mathbf{q}^*). \quad (2)$$

The CDF is approximated with a neural model $\hat{d}_\theta \approx d$ with weights θ . Given the pointcloud \mathcal{P} , the NCSB is $h_\theta(\mathbf{q}, \mathcal{P}) = \min_{\mathbf{p} \in \mathcal{P}} \hat{d}_\theta(\mathbf{q}, \mathbf{p})$, so that $h_\theta(\mathbf{q}, \mathcal{P}) \geq 0$ represents safety.

To plan a trajectory that satisfies the NCSB constraint, we use the safe bubble cover [11] algorithm, which exploits the Lipschitz property of the NCSB h_θ to construct spherical safe regions around sampled configurations. Doing so obviates collision checking within the safe regions, dramatically reducing computation time for faster re-planning.

Meanwhile, the plan does not consider uncertainties in the pointcloud \mathcal{P} and the CDF parameters θ . To execute actions that are robustly safe against these uncertainties, we

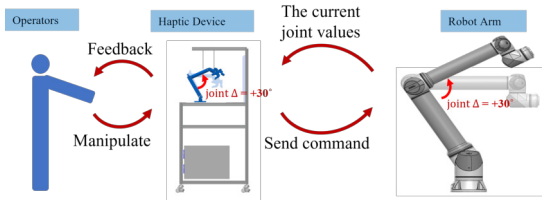


Fig. 4: Flow diagram of teleoperation using the haptic device.

use the distributionally robust control barrier function (DR-CBF) formulation from our prior work [12]:

$$\begin{aligned} \mathbf{u}^* &= \min_{\mathbf{u}} \|\mathbf{u} - \mathbf{u}_{nom}\|, \\ \text{s.t. } \inf_{\mathbb{P} \in \mathcal{M}(\mathcal{P}, \theta)} \mathbb{P} \left(\frac{\partial h_{\theta}}{\partial \mathbf{q}} \mathbf{u} + \frac{\partial h_{\theta}}{\partial t} + \alpha h_{\theta} \geq 0 \right) &\geq 1 - \epsilon, \end{aligned} \quad (3)$$

where \mathbf{u}_{nom} is a nominal control action from either the teleoperation device or the plan, and $\alpha \geq 0$ is a parameter. The probabilistic constraint ensures safety $h_{\theta}(\mathbf{q}, \mathcal{P}) \geq 0$ with probability $\geq 1 - \epsilon$ over the manipulator’s trajectory, with respect to all distributions $\mathcal{M}(\mathcal{P}, \theta)$ over the pointcloud \mathcal{P} and CDF parameters θ that are within a given Wasserstein distance away from the observations. The constraint can be relaxed to a linear inequality over samples of \mathcal{P} and θ [12].

C. Teleoperation

Teleoperating a manipulator can be counterintuitive due to the kinematic differences between humans and robots. BVLOS teleoperation is even more challenging due to a lack of tactile feedback. Thus, we build on GELLO [9] to develop a haptic device that mirrors the kinematic structure of the manipulator, and provides proprioceptive feedback. We customized the GELLO hardware [9] for the RB10-1300 manipulator, and developed a custom software illustrated in Fig. 4. When the operator manipulates the haptic device, the software generates relative joint position commands for the manipulator by comparing the changes in the haptic device’s joint angles to the current joint angles of the manipulator. This allows the operator to fully control all six joints of the manipulator. For proprioceptive feedback, actuator currents are applied to the haptic device that are proportional to those measured in the manipulator, replicating the corresponding torques. Thus, the device provides proprioceptive feedback that emulates the actual physical interaction, enhancing realism and intuition for effective operation.

IV. FEASIBILITY STUDY

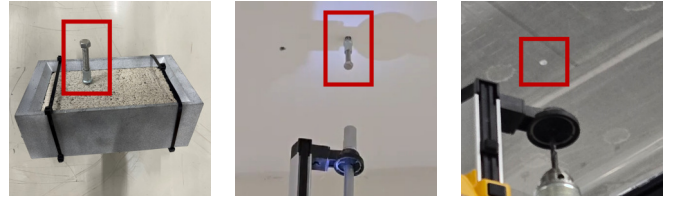
A. Robot Drilling

We evaluate the feasibility of the hardware platform for BVLOS teleoperation by performing drilling, anchoring, and bolting on ceilings, which are representative of overhead works. We considered ceilings made of concrete, gypsum, and acrylic materials in three different sites, as shown in Fig. 5(a), for a total of 9 experiments.

As shown in Fig. 5(b), all tasks were successfully performed on concrete and gypsum materials in all three sites. For acrylic, only drilling succeeded, whereas anchoring and



(a) Three different experimental sites.



(b) Results from each site

Fig. 5: Experimental sites and results.

TABLE I: Result of the experiments conducted with three different material types in all sites in Fig. 5(a)

Method	Experimental Materials		
	Concrete	Gypsum	Acrylic
Drilling	✓	✓	✓
Anchoring	✓	✓	X
Bolting	✓	✓	X

bolting failed in all sites. The failure of the two operations is because they rely on frictional forces, which acrylic does not provide due to its smoothness. In contrast, the rougher surfaces of concrete and gypsum offered enough friction for stable anchoring and bolting. We are exploring bimanual operation to address this limitation.

B. Planning and Control

We present an initial evaluation of the suitability of the planning and control modules in dynamic environments using a 6-DoF xArm6 [13] in a laboratory environment. The manipulator was assigned a goal end-effector position to navigate towards in a cluttered environment with both static and dynamic obstacles as depicted in Fig. 6. A single Realsense D-435 camera was used to provide observations of the obstacles from a fixed location. We used an Aruco marker to track the velocity of dynamic obstacles, which may be replaced with DynaGSLAM.

The results in Fig. 6 show that the manipulator follows a nominal trajectory to reach the desired goal configuration, while deviating from the nominal trajectory to avoid incoming obstacles as necessary. We observed that the safe bubble cover [11] significantly reduces the number of collision checks, by up to tenfold compared to other baselines, while yielding comparable path lengths.

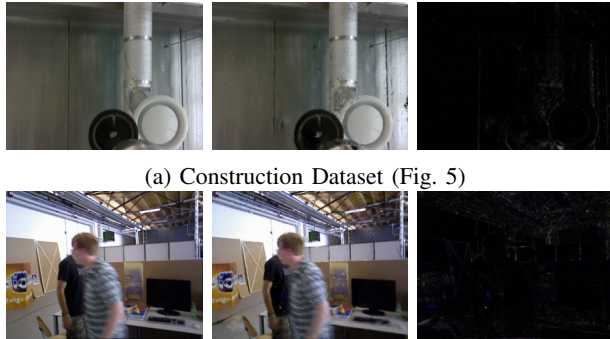
TABLE II: Image Reconstruction Accuracy of DynaGSLAM

Dataset	PSNR (dB)↑	LPIPS (%)↓	SSIM (%)↑
TUM [14]	27.2	20.0	94.5
Construction (Fig. 5)	26.8	44.7	84.4



(a) Planned configurations (b) Response to dynamic obstacle (c) Response to dynamic obstacle (d) Response to dynamic obstacle

Fig. 6: Bubble-CDF planner and DR-CBF control in a laboratory experiment with a 6-DoF xArm robot.



(a) Construction Dataset (Fig. 5)

(b) TUM Dataset [14]

Fig. 7: Qualitative results of DynaGSLAM. Left: ground truth. Center: rendered. Right: error.

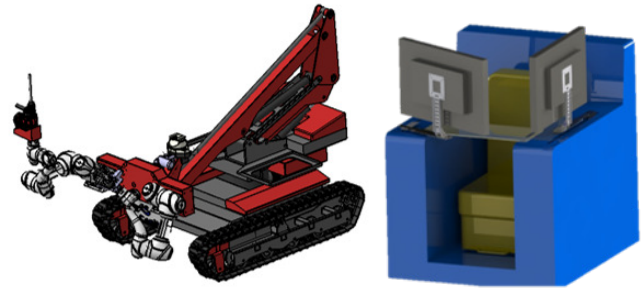
C. Dynamic Mapping

We evaluate dynamic scene reconstruction with DynaGSLAM in two datasets, TUM [14] and one collected during ceiling drilling experiments (Fig. 5). The qualitative results in Fig. 7 and the quantitative results in Table II show that DynaGSLAM renders RGB images accurately. This is owing to accurate reconstruction of dynamic environments despite moving objects and featureless ceilings. Such ability of handling moving objects allows photorealistic visual feedback even as the manipulator or work pieces are moving.

V. ONGOING AND FUTURE WORK

We presented an initial hardware prototype and the experimental results of individual software modules in isolation in laboratory settings. We are currently upgrading the hardware platform based on initial findings, as envisioned in Fig. 8, and are integrating the software modules for a full demonstration in a real construction site. Specifically, we will mount an additional manipulator to the torso, which will enable simultaneous handling of fasteners and fixtures used in ceiling installations. The planning and control module will be accordingly extended to jointly consider bimanual operations and the mobile base. We are also integrating higher-payload manipulators to support more demanding tasks, and a more permanent booth will be built for greater ergonomic comfort.

We are also making further improvements to the guidance software for greater autonomy. In particular, the teleoperation interface will allow the operator to specify a goal end-effector pose for planning. To support planning, we are working on constructing the NCSB from Gaussian splat maps.



(a) Dual-arm construction robot

(b) Teleoperation booth

Fig. 8: Envisioned hardware improvements

REFERENCES

- [1] K. S. Anderson, “Survey: 65% of construction workers are fatigued on the job,” *Construction Dive*, 2018.
- [2] N. V. Schwatka, L. M. Butler, and J. R. Rosecrance, “An aging workforce and injury in the construction industry,” *Epidemiologic Reviews*, vol. 34, pp. 156–167, 2012.
- [3] Hilti Group, “Hilti jaibot – semi-autonomous drilling robot for ceiling work,” 2025, accessed: 2025-04-17. [Online]. Available: <https://www.hilti.com/content/hilti/W1/US/en/business/business/trends/jaibot.html>
- [4] CSC Robotic Engineering, “Drillcorpio – ceiling construction robot,” 2025, accessed: 2025-04-17. [Online]. Available: <https://cscrobotic.com/en/our-products/>
- [5] R. B. Li, M. Shaghghi, K. Suzuki, X. Liu, V. Moparthi, B. Du, W. Curtis, M. Renschler, K. M. B. Lee, N. Atanasov, and T. Nguyen, “DynaGSLAM: Real-time Gaussian-splatting SLAM for online rendering, tracking, motion predictions of moving objects in dynamic scenes,” *arXiv preprint arXiv:2503.11979*, 2025.
- [6] S. X. Fang, S. O’Young, and L. Rolland, “Development of small UAS beyond-visual-line-of-sight flight operations: System requirements and procedures,” *Drones*, vol. 2, no. 2, 2018.
- [7] K. Long, K. M. B. Lee, N. Raicevic, N. Attasseri, M. Leok, and N. Atanasov, “Neural configuration-space barriers for manipulation planning and control,” *arXiv preprint arXiv:2503.04929*, 2025.
- [8] A. Tiberkak, A. Hentout, and A. Belkhir, “WebRTC-based MOSR remote control of mobile manipulators,” *International journal of intelligent robotics and applications*, vol. 7, no. 2, p. 304–320, 2023.
- [9] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel, “Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators,” 2023, *arXiv preprint arXiv:2307.01234*.
- [10] Y. Li, X. Chi, A. Razmjoo, and S. Calinon, “Configuration space distance fields for manipulation planning,” in *Robotics: Science and Systems (RSS)*, 2024.
- [11] K. M. B. Lee, Z. Dai, C. L. Gentil, L. Wu, N. Atanasov, and T. Vidal-Calleja, “Safe bubble cover for motion planning on distance fields,” *arXiv preprint arXiv:2408.13377*, 2024.
- [12] K. Long, Y. Yi, Z. Dai, S. Herbert, J. Cortés, and N. Atanasov, “Sensor-based distributionally robust control for safe robot navigation in dynamic environments,” *arXiv preprint arXiv:2405.18251*, 2024.
- [13] UFactory, “xArm,” 2025, accessed: 2025-04-17. [Online]. Available: <https://www.ufactory.us/xarm>
- [14] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, “A benchmark for the evaluation of RGB-D SLAM systems,” in *Proc. of the International Conference on Intelligent Robot Systems (IROS)*, Oct. 2012.