Leveraging the USTC FLICAR Dataset to Prepare Robots for Heavy-Duty Aerial Work Tasks in Construction Environments

Ziming Wang, Yujiang Liu, Yifan Duan, Xingchen Li, Xinran Zhang, Jianmin Ji, Erbao Dong and Yanyong Zhang, *Fellow IEEE*

***Abstract*— The author's previous journal article [1] introduced the USTC FLICAR dataset, which aims to enhance the autonomy of aerial robots in construction environments through advanced sensor fusion. We discuss the integration of LiDAR, camera, and inertial measurement units on an autonomous aerial platform—specifically a bucket truck retrofitted with sophisticated robotics—to address safety and efficiency in aerial work. This integration addresses key challenges in construction robotics such as navigating complex, dynamic environments and ensuring the safety of human workers. Our approach demonstrates the potential to significantly reduce risks and improve operational accuracy in high-risk environments. All resources of the dataset is available for download at: `https://ustc-flicar.github.io/`.**

Fig. 1. Typical construction aerial work scenes using bucket trucks

I. INTRODUCTION

In the construction industry, where the risk of accidents and the demand for precision are high, traditional manual methods are increasingly being supplemented by autonomous systems. Aerial work, involving tasks such as inspection, maintenance, and repair in hard-to-reach areas, is particularly perilous and highlights the need for robust autonomous solutions. Figure 1 shows some typical aerial work scenes in the construction industry. The integration of robotics into these tasks can not only enhance safety by reducing human exposure to dangerous conditions but also increase the efficiency and accuracy of the operations conducted.

Despite their potential, the deployment of autonomous robots in such environments faces substantial hurdles, including the need for precise navigation and robust perception in highly unstructured, cluttered settings. In recent years, numerous public datasets have played significant roles in the advancement of autonomous cars and UAVs. However, these two platforms differ from aerial work robots: UAVs are limited in their payload capacity, while cars are restricted to two-dimensional movements. The ground datasets represented by KITTI [2] features a wealth of sensors and centimeter-level accuracy ground truth obtained through RTK-GPS or LiDAR SLAM. While the aerial datasets represented by EuRoC [3] typically have a limited types and number of sensors due to payload capacity limitations. And millimeter-level accuracy ground truth is generated using a laser tracker. What about combine the strengths of EuRoC and KITTI to create a new aerial dataset which has both millimeter-level accuracy ground truth and rich sensors? Here comes USTC FLICAR. We create the "Giraffe" mapping robot based on a new heavy-duty aerial platform bucket truck, which is equipped with a variety of well-calibrated and time synchronized sensors: four 3D LiDARs, two stereo cameras, two monocular cameras, Inertial Measurement Units (IMUs), and a GNSS/INS system. A laser tracker is used to record the millimeter-level ground truth positions. We also make its ground twin, the "Okapi" mapping robot, to gather data for comparison. This dataset facilitates the development and testing of simultaneous localization and mapping (SLAM), crucial for effective autonomous navigation and task execution in dynamic construction sites.

Through this paper, we address the challenge of integrating and calibrating a comprehensive array of sensors on an autonomous aerial work platform. Our work focuses on the practical application of these technologies in a construction context, demonstrating how advanced sensor fusion enhances the robot's ability to perceive its environment and execute tasks with high precision.

This introduction sets the stage for discussing the methodologies involved in sensor integration, the challenges overcome in robust perception, and the significant benefits these autonomous systems bring to construction environments. The subsequent sections will delve into the specific technologies and data analysis methods used, showcasing the real-world applicability and effectiveness of our approach.

*This work was not supported by National Key R&D Program of China (No. 2018YFB1307403)

Ziming Wang is with School of Computer Science and Technology, University of Science and Technology of China, Hefei, China `zimingwang@mail.ustc.edu.cn`

Ziming Wang is also with the Department of Mechanical Engineering, The University of Hong Kong, Pokfulam Hong Kong SAR, China `zmw2001@hku.hk`

TABLE I

THE SENSOR MODEL SPECIFICATIONS AND DATA INFORMATION IN THIS DATASET.

No	Sensor	Model	ROS Topic	Message type	Rate
1	IMU/INS	Xsens MTi-G-710	/imu/data	sensor_msgs/Imu	400Hz
2	Horizontal LiDAR 1	Velodyne HDL-32E	/velodyne_points_HDL32	sensor_msgs/PointCloud2	5/10Hz (rotate at 10Hz)
3	Horizontal LiDAR 2	Ouster OS0-128	/os_cloud_node/imu /os_cloud_node/points /img_node/reflect_image /img_node/signal_image	sensor_msgs/Imu sensor_msgs/PointCloud2 sensor_msgs/Image sensor_msgs/Image	100Hz 10Hz 10Hz 10Hz
4	Horizontal LiDAR 3	LiVOX Avia	/livox/lidar /livox/imu	livox_ros_driver/CustomMsg sensor_msgs/Imu	10Hz 200Hz
5	Vertical LiDAR 1	Velodyne VLP-32C	/velodyne_points_VLP32	sensor_msgs/PointCloud2	10Hz
6	Stereo Camera front	PointGrey Bumblebee xb3	/camera/left/image_raw /camera/center/image_raw /camera/right/image_raw	sensor_msgs/Image	10-16Hz 10-16Hz 10-16Hz
7	Stereo Camera back	PointGrey Bumblebee xb2	/cam_xb2/left/image_raw /cam_xb2/right/image_raw	sensor_msgs/Image	10-20Hz 10-20Hz
8	Mono Camera 1	Hikvision MV-CB016-10GC-C	/hik_camera/image_raw	sensor_msgs/Image	20Hz
9	Mono Camera 2	Hikvision MV-CE060-10UC	/right_camera/image	sensor_msgs/Image	20Hz

Fig. 2. "Giraffe" and "Okapi" acquisition systems:
"Giraffe" aerial system: (a), (b) and (c).
"Okapi" ground system: (a), (b) and (d)
(a) multi-sensor data collection platform (Fig. 3), (b) laser tracker ground truth system, (c) bucket truck, (d) ground robot.

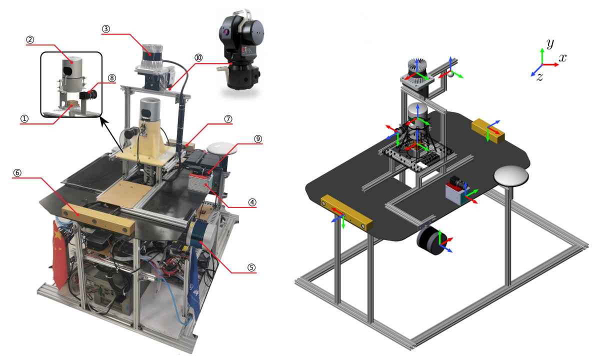

Fig. 3. Sensor setup on a multi-sensor platform.

A. Data Acquisition

II. METHODOLOGY

The primary focus of our methodology involves the integration and calibration of a diverse array of sensors on an autonomous aerial platform to support robust navigation and perception in construction environments. This section outlines the sensor setup, data acquisition strategies, and the calibration techniques employed, which are critical for ensuring the precision and reliability of the data collected. The data was acquired using two different systems, the "Giraffe" and "Okapi" systems, as depicted in Figure 2. The "Giraffe" system is an aerial platform consisting of a multisensor data collection platform (a), a laser tracker ground truth system (b), and a bucket truck (c). On the other hand, the "Okapi" system is a ground-based system similar to an autonomous vehicle, equipped with the same sensors (a) and a ground truth recording system (b), and mounted on a ground robot (d) for the acquisition of ground-level data for comparison with the data collected by the aerial system.

A. Sensors Setup

Our aerial platform, referred to as the "Giraffe," is equipped with an array of sensors on a bucket truck that includes four 3D LiDARs, two stereo cameras, two monocular cameras, inertial measurement units (IMUs), and a GNSS/INS system . This configuration is designed to capture a comprehensive set of data that supports robust perception and navigation capabilities. The LiDAR sensors provide

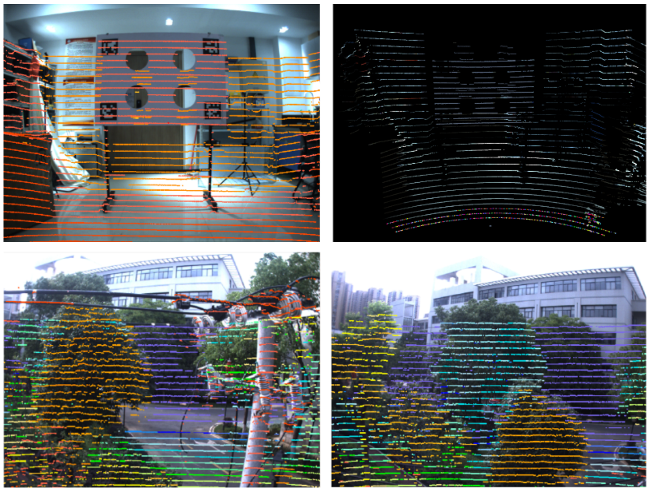

Fig. 4. Velo2Cam [5] camera-LiDAR calibration. *Up-Left*: special calibration board, project LiDAR point cloud to the image. *Up-Right*: colorize LiDAR point cloud with image.*Down*: Aerial scenes LiDAR points fusion with images.

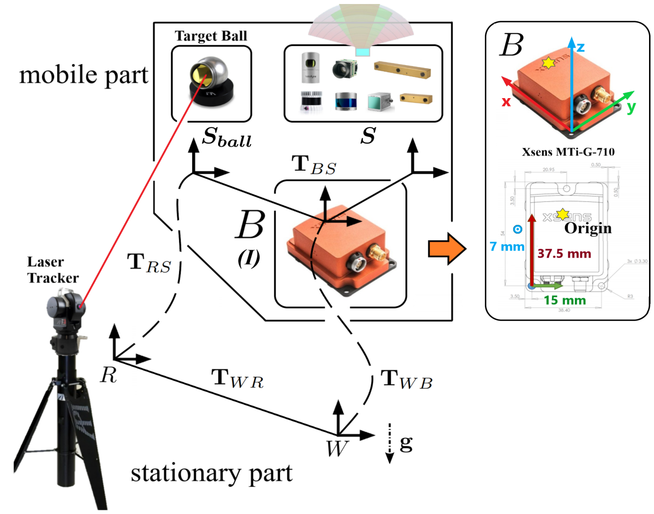

Fig. 5. The sensor system used to capture the datasets consists of multiple sensors, each reporting measurements in its own reference frame $\boldsymbol{S}$. The datasets also include raw data from ground truth instruments, reported in the target ball frame $\boldsymbol{S_{ball}}$ and Laser tracker frame $\boldsymbol{R}$. The body frame $\boldsymbol{B}$ is aligned with the IMU sensor frame $\boldsymbol{I}$. Calibration information for all extrinsic parameters linking the sensors to the body frame $\boldsymbol{B}$ and intrinsic parameters is included in the dataset.

spatial mapping and object detection capabilities, crucial for navigating complex construction sites, while the camera systems capture detailed visual and textural information, facilitating object recognition and scene interpretation.

Data acquisition is conducted using both the aerial platform and its ground counterpart, the "Okapi," to ensure a diverse environmental dataset. The aerial platform performs tasks in a variety of settings including urban construction sites and industrial environments, simulating real-world operational conditions. The datasets include different times of day and weather conditions to test the system's robustness under variable lighting and climatic influences.

B. Calibration Techniques

Proper calibration of sensors is imperative to achieve accurate data fusion and reliable results. We employ several calibration methods tailored to each sensor type:

- LiDAR Calibration: We use a combination of manual alignment techniques followed by automated point cloud registration to ensure the LiDARs are precisely calibrated. This process is vital for accurate 3D reconstruction of the environment.
- Camera Calibration [4]: The stereo and monocular cameras are calibrated using a standard checkerboard calibration method to determine intrinsic parameters and to rectify lens distortions. This step is crucial for accurate depth estimation and object detection.
- Sensor Fusion Calibration: A key aspect of our methodology is the fusion of LiDAR and camera data, which is achieved through a sophisticated calibration process that aligns the data streams in time and space. This ensures that the visual and spatial data complement each other, enhancing the system's ability to navigate and interpret complex environments. As shown in Figure 4.

C. Ground Truth Validation

To validate the accuracy of our autonomous localization and mapping, we use laser trackers and other ground truth systems to provide a baseline against which our robotic system's performance can be measured. This validation is critical for assessing the precision and reliability of the autonomous system under real-world conditions.

III. Results

We evaluated state-of-the-art visual and LiDAR SLAM algorithms using different sensor suites and data sequences to analyze our dataset's characteristics and challenges. The absolute trajectory error (ATE [6]) was used as a performance metric. Toolbox evo [7] is used. We ensured parameter consistency across algorithms and sensors to achieve optimal results for each.

For visual SLAM, we tested algorithms like ORB-SLAM3 [8] and VINS-Mono [9] on various configurations, including monocular, monocular-inertial, and stereo-inertial systems. These evaluations highlighted differences in performance based on camera characteristics and lighting conditions. For instance, the Bumblebee-XB3 camera performed better in low light but struggled with glare, affecting the accuracy of ORB-SLAM3. Conversely, VINS-based methods showed greater robustness in extreme lighting conditions, maintaining accuracy by leveraging environmental data from the Bumblebee-XB3 camera.

In LiDAR SLAM testing, we used A-LOAM [10] and LeGO-LOAM [11] on Velodyne LiDARs and tested LIO-SAM [12] and FAST-LIO [13] on LiDAR-Inertial systems. LiDAR SLAM generally provided reliable results, with an ATE around 0.1 m under typical conditions. However, challenges arose with vertical LiDAR due to its limited field

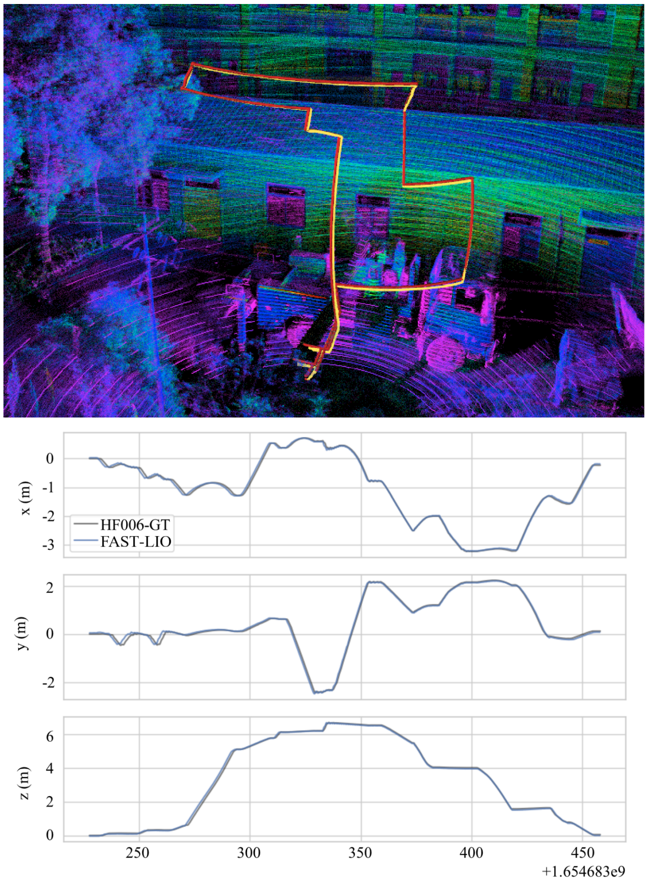

Fig. 6. The ground truth is marked in red and the SLAM trajectory is marked in yellow. The figure shows the running results of Fast-LIO [13] on the hf006 sequence.

of view and issues in texture-poor environments, leading to increased ATE in scenarios with large rotations or low feature overlap.

Overall, the study confirmed the efficacy of integrating well-calibrated IMUs to enhance SLAM accuracy and robustness, particularly under challenging conditions. The results varied by sensor type and environmental factors, with ORB-SLAM3 excelling in well-lit conditions and VINS-based methods performing consistently across varied lighting scenarios.

IV. Conclusion and Future Work

A. Conclusion

The research presented in this extended abstract demonstrates the effectiveness of integrating advanced sensor technologies in autonomous aerial work robots for construction environments. Through rigorous evaluation of various SLAM algorithms using the USTC FLICAR Dataset, we have substantiated the significant advantages of multimodal sensor fusion over single-sensor systems. Our findings confirm that the hybrid approach not only enhances the accuracy and robustness of autonomous navigation and mapping but also substantially improves the system's operational efficiency and safety in complex, dynamic construction sites.

The successful deployment of these technologies in real-world settings illustrates their potential to revolutionize construction operations by reducing risks and increasing efficiency. The robustness of the integrated SLAM systems in diverse environmental conditions underscores the practical viability of this technology for widespread industrial applications. For more detailed information, we invite interested readers to consult our full journal paper [1] and visit our project website.

B. Fulture Work

Moving forward, our research will focus on refining data fusion and SLAM algorithms. A key aspect of our future work will involve developing more robust and practical algorithms by integrating the Fast-LIO [13] framework, the USTC FLICAR dataset [1], and the inherent features and constraints of robotic arms. This approach is aimed at advancing construction robotics and making meaningful contributions to the community. Collaborative efforts with industry and academic partners will continue to be crucial in optimizing these technologies and broadening their impact across various high-risk industries.

References

[1] Wang Z, Liu Y, Duan Y, et al. USTC FLICAR: A sensors fusion dataset of LiDAR-inertial-camera for heavy-duty autonomous aerial work robots. The International Journal of Robotics Research. 2023;42(11):1015-1047. doi:10.1177/02783649231195650

[2] Geiger A, Lenz P, Stiller C, et al. (2013) Vision meets robotics: the KITTI dataset. The International Journal of Robotics Research 32(11): 1231–1237.

[3] Burri M, Nikolic J, Gohl P, et al. (2016) The EuRoC micro aerial vehicle datasets. The International Journal of Robotics Research 35(10): 1157–1163.

[4] Zhang Z (2000) A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11): 1330–1334.

[5] Beltrán J, Guindel C, de la Escalera A, et al. (2022) Automatic extrinsic calibration method for LiDAR and camera sensor setups. In: IEEE Transactions on Intelligent Transportation Systems. New York, USA: IEEE.

[6] Sturm J, Engelhard N, Endres F, et al. (2012) A Benchmark for the Evaluation of RGB-D SLAM Systems. In: 2012 IEEE/RSJ international conference on intelligent robots and systems, Vilamoura-Algarve, Portugal, 7–12 Oct. 2012, pp. 573–580. IEEE.

[7] Grupp M (2017) evo: Python package for the evaluation of odometry and SLAM. https://github.com/MichaelGrupp/evo

[8] Campos C, Elvira R, Rodríguez JJG, et al. (2021) ORB-SLAM3:an accurate open-source library for visual, visual–inertial, and multimap SLAM. IEEE Transactions on Robotics 37(6): 1874–1890.

[9] Qin T, Li P and Shen S (2018) VINS-mono: a robust and versatile monocular visual-inertial state estimator. IEEE Transactions on Robotics 34(4): 1004–1020.

[10] Zhang J and Singh S (2014) LOAM: lidar odometry and mapping in real-time. In: Robotics: Science and Systems. Berkeley, CA, USA: University of California, Vol. 2, pp. 1–9.

[11] Shan T and Englot B (2018) LeGO-LOAM: lightweight and ground-optimized lidar odometry and mapping on variable terrain. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS), Madrid, Spain, 1–5 Oct. 2018, pp. 4758–4765. IEEE.

[12] Shan T, Englot B, Meyers D, et al. (2020) LIO-SAM: tightly coupled lidar inertial odometry via smoothing and mapping. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), Las Vegas, NV, USA, Oct 2020, pp. 255135–295142. IEEE

[13] Xu W and Zhang F (2021) FAST-LIO: a fast, robust LiDAR inertial odometry package by tightly-coupled iterated kalman filter. IEEE Robotics and Automation Letters 6(2): 3317–3324.