Development of a Simulated UAV Platform for Sensor-Fusion-Based SLAM

Muhammad Fairuz Mummtaz[*], Jaejun Yoo[†], Ida Bagus Krishna Yoga Utama[*], Irzal Zaini[*], and Yeong Min Jang[*]

[*]*Department of Electronics Engineering*
Kookmin University
Seoul, South Korea
muhammad.fairuz91@ui.ac.id, ibkyutama@ieee.org, irzal.zaini00@gmail.com, yjang@kookmin.ac.kr

[†]*DNA+ Drone Platform Research Center*
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, South Korea
jjryu@etri.re.kr

***Abstract*—Accurate positioning is crucial for autonomous navigation of unmanned aerial vehicles (UAVs), as it provides reliable pose information to the perception and planning-control modules. However, when global navigation satellite system (GNSS) signals are unavailable or degraded, the accumulated drift of dead-reckoning and inertial estimates significantly reduces localization accuracy, especially in complex 3D environments. To address this issue, this paper proposes a simulation-based sensor fusion based simultaneous localization and mapping (SLAM) framework for robust UAV positioning without relying on GNSS. The framework directly utilizes sequential 3D LiDAR scans to estimate the UAV pose and incrementally build a dense map of the environment, while high-resolution camera streams provide contextual information for the segmentation task. In addition, a modular simulation pipeline is constructed to support configurable sensor models, flight trajectories, and scene geometries, enabling systematic evaluation under GNSS-denied scenarios. The proposed approach is validated using a simulation experiments on a complex environment, and the results demonstrate that the 3D LiDAR SLAM system achieves accurate and stable localization in the simulation environment.**

***Index Terms*—UAV, Simultaneous Localization and Mapping, Segmentation, Sensor Fusion, Deep Learning**

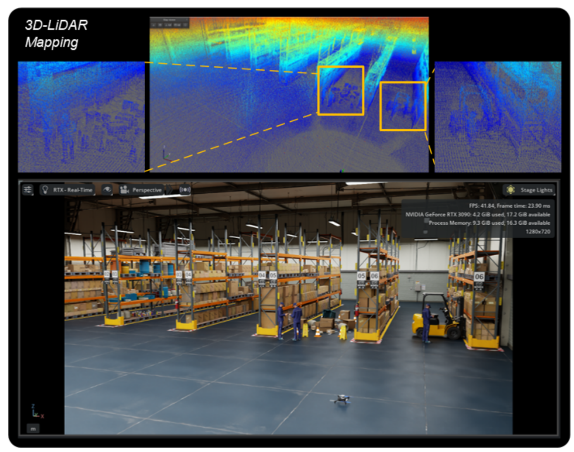

Fig. 1. The 3D map illustrates the environment reconstructed using the SLAM algorithm, where LiDAR data was fused with estimated poses to generate an accurate spatial representation. The 3D map results demonstrates how the algorithm effectively captures structural details and spatial consistency as the UAV navigates through the scene.

I. Introduction

Recent research on UAV SLAM and navigation has gradually shifted to reliably operating in GPS-denied or degraded environments [1]. This shift reflects the growing demand for autonomous aerial systems capable of maintaining robust situational awareness and precise localization in complex scenarios such as urban areas, indoor spaces, or disaster zones where satellite signals are unavailable. Advances in vision-based perception, multi-sensor fusion, and deep learning have further accelerated this trend, enabling UAVs to navigate with increasing autonomy and resilience [2].

Recent surveys on UAV-based SLAM illustrate that many of the sensing modalities, such as cameras, LiDAR, radar, and IMUs, are now being combined with the aim of solving key challenges like aggressive motion, large-scale 3D environments, and dynamic scenes [3],[4],[5]. However, their performances still depend heavily on the structure of the environment, the quality of sensor calibration, and careful algorithm design. Indeed, recent works have illustrated that the combination of complementary modalities allows for enhanced localization accuracy and robustness in cluttered indoor, urban, and agricultural scenarios [6],[7],[8].

Furthermore, the works related to UAV navigation in confined indoor spaces prove that the integration of LiDAR, cameras, and IMUs, possibly enhanced by deep learning, allows for more reliable mapping and obstacle avoidance in GPS-denied conditions [9]. However, challenges persist in achieving real-time performance, maintaining long-term consistency, and handling dynamic elements across diverse environmental conditions [10],[11].

As one of the key challenge in implementing effective multi-sensor fusion lies in achieving precise extrinsic calibration between sensors, which allows their data streams to mutually complement each other for reliable task performance [12]. This can be accomplished either by developing a controlled environment for the calibration process or by employing targetless extrinsic calibration methods in standard environments [13], [14]. Both approaches successfully yield the necessary extrinsic parameters for the sensors involved.

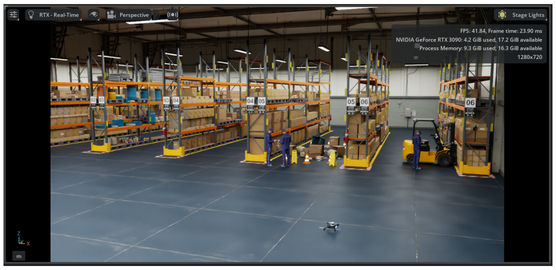

Fig. 2. Warehouse scene environment.

To support rapid development and evaluation of such sensor-fusion SLAM systems, there is an increasing interest in high-fidelity simulation platforms that offer realistic multi-modal sensor data and complex 3D environments [15]. On some recent platforms, such as end-to-end UAV SLAM simulators and telepresence-based UAV simulation frameworks, have made synthetic datasets, configurable sensor suites, and ground-truth trajectories available, thus enabling systematic benchmarking of SLAM pipelines before their deployment on real hardware [16],[15],[17]. However, most of the existing tools tend to target specific sensing configurations or application scenarios; therefore, there is still a need for more versatile simulated UAV platforms that are specifically targeted at integrated sensor-fusion-based SLAM research.

Fig. 3. Sensor fusion integrated drone 3D design.

The simulation provides a practical means to achieve real-time performance, agile maneuvering, high-resolution sensing, and reduced development cost. There are several available simulator environment and platform such as ROS-Gazebo, Airsim-W, XTDrone, and SmrtSwarm, but these tools exhibit limitations in faithfully replicating complex real-world conditions. Developed by NVIDIA, the Isaacsim platform is introduced. In particular, it leverages PhysX, Integrated ML support, and ROS2 support to create a real-time, interactive framework that offers a dynamic and visually realistic virtual environment [18]. By using this framework, we developed a end-to-end LiDAR-SLAM simulation using drone and virtual environment inside of the Isaacsim software itself.

II. Methodology

For our implementation, we used NVIDIA Isaacsim software as the main framework to develop the UAV-based LiDAR-SLAM algorithm. There are several other software that support our simulation, such as PegasusSimulator, PX4-Autopilot, ROS2, etc. Furthermore, for the SLAM algorithm, we take reference from MOLA as one of the robust and modular SLAM algorithm [19].

A. Simulation Software

The main software used for developing the simulation is called Isaac Sim which is NVIDIA's robotics simulation platform built on Omniverse, used to design, test, and train AI-powered robots in physically realistic 3D environments with photorealistic rendering and accurate physics. It supports importing robots and environments from standard formats such as URDF, MJCF, and CAD, and provides built-in tools for sensor simulation (cameras, LiDAR, IMU), control, motion planning, and synthetic data generation for perception models.

B. Simulation Environment

Several key features of the Isaac Sim platform were utilized in this implementation, including the scene environment, robotic models, various sensors, and PhysX-based physics simulation. These components collectively enable the simulation to closely replicate real-world scenarios, ensuring realistic interaction dynamics and sensor responses.

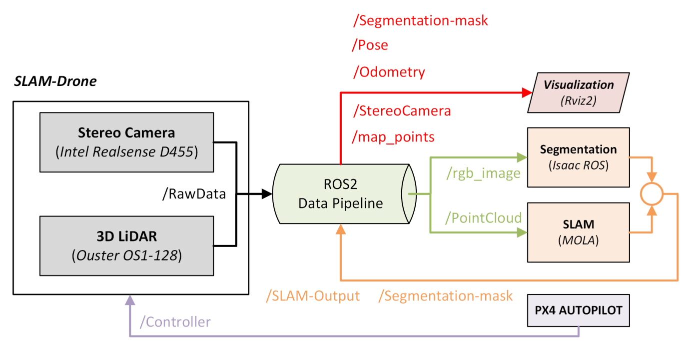

Fig. 4. Overall SLAM system architecture.

The warehouse scene template in NVIDIA Isaac Sim provides a realistic industrial environment for UAV simulation and SLAM testing, featuring modular assets like warehouse buildings, shelving units, racks with stacked boxes, human-form workers and structural columns as shown in Figure 2. This template also supports robotics research by offering customizeable layouts through extensions like warehouse creator for procedural wall/column placement, making it ideal for validating multi-sensor fusion algorithms in GPS-denied warehouse scenarios typical of autonomous drone navigation.

The 3D-designed drone implementation utilizes Universal Scene Description (USD) assets in Isaac Sim, featuring a modular quadrotor frame with articulated propellers called Iris drone from PegasusSimulator [20], RealSense D455 camera,

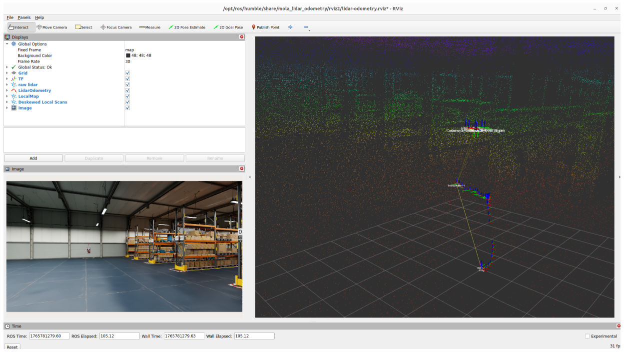

Fig. 5. On-simulation rviz2 visualization.

and Velodyne LiDAR sensors positioned for optimal sensor fusion coverage, all scaled to real-world metrics with PhysX-enabled rigid bodies and joint drives for realistic flight dynamics as shown in Figure 3. The accompanying ROS2 data pipeline, integrates several sensors data streams through (LiDAR and Camera), fuses them via tf2 transforms into a unified /odom frame, feeds into LiDAR SLAM nodes (MOLA), and publishes /map and /pose estimates for downstream navigation. This end-to-end pipeline enables rapid prototyping of multi-sensor SLAM for UAVs in warehouse environments, bridging simulation with real-time ROS2 middleware for seamless transition to hardware deployment.

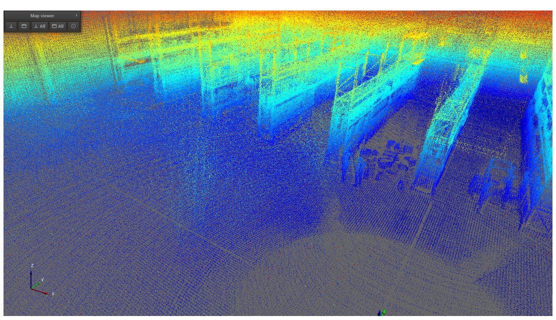

Fig. 6. 3D map results.

C. Overall Algorithm Architecture

The overall architecture of the proposed algorithm consists of several components, as illustrated in Figure 4. The system begins with UAVs equipped with two sensors which are a stereo camera and a 3D LiDAR that continuously publish their data to the ROS2 data pipeline. The implemented SLAM algorithm, MOLA (Modular System for Localization and Mapping), processes the LiDAR point cloud data to perform simultaneous localization and mapping. The output from the MOLA SLAM module is then published back to the ROS2 pipeline for visualization using the RViz2 plugin. Additionally, based on Isaac-ROS library, using the image data from D455 Intel Realsense camera, we also implemented segmentation algorithm to accurately classify the human-form object inside of simulation. Lastly, the UAV's motion is controlled through the PX4-Autopilot library, which interfaces with an external joystick controller.

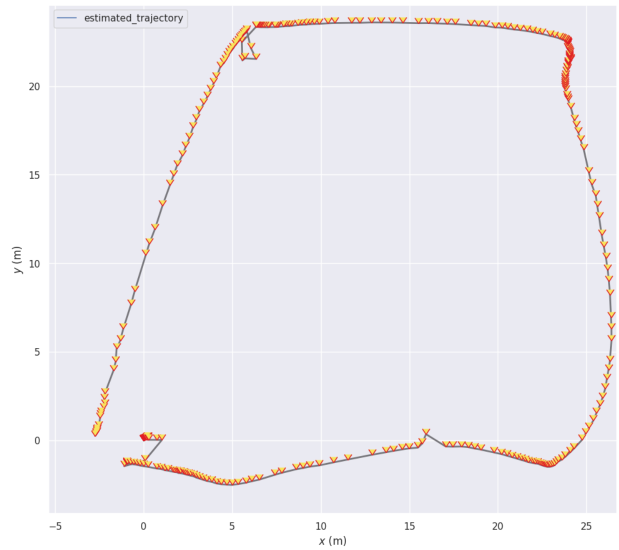

Fig. 7. 2D representation of the UAV's simulated flight path.

Once the simulation begins, the UAV can be manually controlled using an external joystick to scan the entire environment. As the UAV navigates through the scene, the LiDAR sensor continuously collects point cloud data from different viewpoints, which can be visualized in real time using RViz2, as shown in Figure 5. All collected data are stored and integrated within the MOLA SLAM framework. Upon completing the scanning process, the MOLA algorithm generates a comprehensive 3D map of the simulated environment.

III. Results and Discussion

After executing the simulation and deploying the SLAM algorithm within the developed simulation environment, we evaluated its performance based on the quality of the generated 3D map, the reconstructed UAV trajectory and segmentation mask. The quality of the 3D map served as a measure of the algorithm's mapping accuracy and spatial consistency, while the trajectory analysis allowed us to assess localization stability and drift behavior as the UAV scanned the entire simulated environment.

A. Simulation Results

Based on the 3D map results shown in Figure 6, the SLAM model demonstrates sufficient capability in reconstructing the overall environment within the simulation. The generated map clearly captures key structural elements, including the factory shelves, multiple boxes, and the surrounding walls that define the boundaries of the scene. The spatial environment consistency across the scene indicates that the algorithm is able to maintain reliable pose estimation throughout the scanning process. Minor dissimilarity appear at the edges of some objects such as the boxes in the middle of the factory, which

Fig. 8. D455 intel realsense camera image stream.

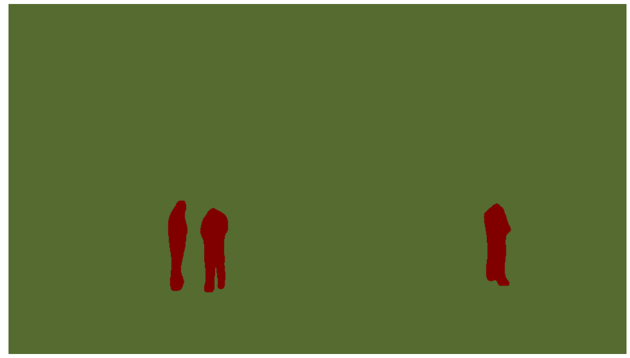

Fig. 9. Segmentation result.

may be attributed to sensor noise or limitations in depth perception during fast UAV motion.

During the overall scene scanning process in the simulation, the SLAM algorithm simultaneously recorded the UAV's trajectory based on the odometry data it generated. This trajectory data, shown in Figure 7, confirms that the UAV successfully covered the entire factory environment during the mapping task. The trajectory visualization reveals that the flight path maintained consistent coverage across different regions with minimal deviation from the planned route. This indicates that the SLAM algorithm provided stable pose estimation and effective spatial awareness throughout the scanning process.

As for the texture information, we implemented the segmentation based on the image data from D455 Intel Realsense camera attached together with LiDAR sensor on the top of the UAVs. We feed those images stream as shown in Figure 8 directly to Isaac ROS based segmentation model. The segmentation model then will resulting a new image which contain the segmentation mask of the human-form object inside of the simulation scene as shown in Figure 9.

For texture information, we implemented semantic segmentation using image data from the Intel RealSense D455 camera, which is attached together with the LiDAR sensor on the top of the UAV. These image streams, as shown in Figure 8, are directly fed into an Isaac ROS-based segmentation model, which generates a segmentation mask for human-form objects within the simulation scene, as illustrated in Figure 9. The generated segmentation mask precisely localizes human-form objects within the simulation environment.

IV. Conclusion and Future Work

In this study, we implemented a SLAM algorithm for UAV applications within a controlled simulation environment using Isaac Sim. The proposed system successfully generated an accurate map of the entire warehouse scene and effectively documented each UAV's translational and rotational movements throughout the operation via several graphs explained before.

In contrast, implementing SLAM using only LiDAR point clouds presents limitations in recognizing objects within the scene. Although LiDAR data provides high geometric accuracy, it lacks textural and semantic information necessary for reliable object identification. Therefore, integrating LiDAR and camera data enables a more comprehensive understanding of the environment. This cooperative approach leverages the precise geometric mapping of LiDAR and the rich visual details from the camera. The point cloud captures the structural layout, while the camera data, processed through segmentation algorithms, distinguishes individual objects, which is human on this simulation. Both algorithms have been tested and validated within the developed simulation environment.

In future work, we plan to extend the current simulation-based implementation to real-world environments to evaluate the algorithm's performance under practical conditions. Additionally, we aim to integrate LiDAR and camera data into a unified, feature-rich map that combines accurate geometric information with detailed visual features for enhanced environmental understanding. Furthermore, we intend to incorporate segmented dynamic objects within the scene such as UAVs, ground robots, and other moving entities to enable more comprehensive perception analysis.

V. Acknowledgement

This work is supported by the Korea Agency for Infrastructure Technology Advancement grant funded by the Ministry of Land, Infrastructure and Transport (RS-2022-00141819, Development of Advanced Technology for Absolute, Relative, and Continuous Complex Positioning to Acquire Ultra-precise Digital Land Information)

References

[1] F. Yao, C. Lan, L. Wang, and et al., "Gnss-denied geolocalization of uavs using terrain-weighted constraint optimization," *International Journal of Applied Earth Observation and Geoinformation*, vol. 135, p. 104277, 2024.

[2] J. Wang, Z. Yu, and et al., "Vision-based deep reinforcement learning of unmanned aerial vehicle (uav) autonomous navigation using privileged information," *Drones*, 2024.

[3] Y. Ren, Y. Cai, H. Li, , and et al., "A survey on lidar-based autonomous aerial vehicles," *IEEE/ASME Transactions on Mechatronics*, pp. 1–17, 2025.

[4] G.-J. Park, J.-H. Noh, and S.-J. Oh, "Low-rank approximation for underwater drone localization," *ICT Express,*, vol. 10, pp. 22–27, 2024.

[5] K. Wang, L. Kooistra, , and et al., "Uav-based simultaneous localization and mapping in outdoor environments: A systematic scoping review," *Journal of Field Robotics*, vol. 41, no. 8, pp. 1949–1975, 2024.

[6] A. James, A. Seth, and et al., "Sensor fusion for autonomous indoor uav navigation in confined spaces," *International Conference on Sensing Technology (ICST)*, 2023.

[7] Z. Zhan, Y. Ming, and et al., "Agrilira4d: A multi-sensor uav dataset for robust slam in challenging agricultural fields," *arXiv*, 2025.

[8] J. Jiang, T. Zhang, and et al., "Lidar-based 3d slam for autonomous navigation in stacked cage farming houses: An evaluation," *Computers and Electronics in Agriculture*, 2025.

[9] A. Seel, F. Kreutzjans, and et al., "Deep reinforcement learning based uav for indoor navigation and exploration in unknown environments," *International Conference on Control, Automation and Robotics*, 2022.

[10] T. Deng, Y. Zhou, and et al., "Multi-modal uav detection, classification and tracking algorithm—technical report for cvpr 2024 ug2 challenge," *arXiv*, 2024.

[11] T. He, J. Hou, and et al., "Multimodal uav target detection method based on acousto-optical hybridization," *Drones*, 2025.

[12] M. F. Mummtaz, J. Yoo, and et al., "Snow-calib: Deep learning-based camera-lidar extrinsic calibration under snowy weather conditions," *IEEE Access*, vol. 13, pp. 182 751–182 762, 2025.

[13] S. Jeong, S. Kim, and et al., "O3 lidar–camera calibration: One-shot, one-target and overcoming lidar limitations," *IEEE Sensors Journal*, vol. 24, pp. 18 659–18 671, 2024.

[14] N. C. Ou, Hanyu, and et al., "Targetless lidar-camera calibration via cross-modality structure consistency," *IEEE Transactions on Intelligent Vehicles*, vol. 9, pp. 2636–2648, 2023.

[15] Q. Li, C. Wu, and et al., "Mssp : A versatile multi-scenario adaptable intelligent robot simulation platform based on lidar-inertial fusion," *arXiv*, 2024.

[16] I. Hroob, R. Polvara, and et al., "Benchmark of visual and 3d lidar slam systems in simulation environment for vineyards," *Towards Autonomous Robotic Systems*, pp. 168–177, 2021.

[17] B. Udugama, "Mini bot 3d: A ros based gazebo simulation," *arXiv*, 2023.

[18] F. P. Audonnet, A. Hamilton, and et al., "A systematic comparison of simulation software for robotic arm manipulation using ros2," *International Conference on Control, Automation and Systems (ICCAS)*, 2022.

[19] J. L. Blanco-Claraco, "A flexible framework for accurate lidar odometry, map manipulation, and localization," *The International Journal of Robotics Research*, 2022.

[20] M. Jacinto, J. Pinto, and et al., "Pegasus simulator: An isaac sim framework for multiple aerial vehicles simulation," *International Conference on Unmanned Aircraft Systems (ICUAS)*, 2024.