Multi-Drone SLAM for Infrastructural Inspection*

*Note: Sub-titles are not captured in Xplore and should not be used

Khushal Hemant Sharma
SCAI
Arizona State University
Tempe, USA
Email: kshar114@asu.edu

Sakshi Sanjay Khade
SCAI
Arizona State University
Tempe, USA
Email: skhade5@asu.edu

Mainak Malay Saha
SCAI
Arizona State University
Tempe, USA
Email: msaha4@asu.edu

Mrunmayee Nitin Valunj
SEMTE
Arizona State University
Tempe, USA
Email: mvalunj@asu.edu

***Abstract*—Efficient and accurate inspection of large infrastructures is vital in domains such as civil engineering, energy distribution, and industrial maintenance. Traditional manual inspections are labor-intensive, pose safety risks, and often result in incomplete or inconsistent data. This paper proposes a multi-drone simultaneous localization and mapping (SLAM) framework for efficient and scalable infrastructure inspection. By deploying multiple drones equipped with RGB-D sensors, integrated with RTAB-Map-based SLAM and MAVROS-enabled flight control, the system reduces mission duration and improves overall mapping accuracy. Each drone autonomously navigates using predefined trajectories, collects spatial and visual data, and generates local maps. These individual maps are later fused offline using RTAB-Map's multi-session capabilities to produce a unified, metrically accurate 3D reconstruction. Field experiments and simulated trials demonstrate the effectiveness of this approach, highlighting its scalability, robustness, and potential for widespread adoption in automated inspection tasks.**

I. Introduction

Infrastructure inspection is a critical process that ensures the continued safety, functionality, and longevity of civil structures such as bridges, buildings, pipelines, tunnels, and power stations. These inspections are vital not only for detecting structural anomalies, degradation, or defects but also for adhering to safety regulations and performing preventive maintenance. However, traditional inspection techniques often involve extensive scaffolding, cranes, and manual labor, which are not only costly and labor-intensive but also expose human workers to potentially hazardous environments.

The advent of unmanned aerial vehicles (UAVs), commonly known as drones, has significantly transformed the landscape of infrastructure inspection. UAVs offer a safe, fast, and cost-effective alternative by enabling high-resolution data collection from difficult-to-access or dangerous locations without requiring direct human involvement. With advancements in onboard computing and sensor technologies, drones can now carry RGB-D cameras, LiDAR systems, and other specialized sensors, thereby making them powerful platforms for autonomous 3D mapping.

Despite their advantages, single-drone inspection systems face notable challenges. These include limited battery life, which restricts mission duration and coverage area; increased cumulative drift in SLAM systems over time and distance; and the inefficiency of sequentially covering large infrastructures. These limitations often result in longer mission times, gaps in coverage, and reduced spatial resolution.

To address these issues, this paper proposes a collaborative and distributed approach using multiple UAVs equipped with RGB-D sensors and onboard SLAM capabilities. The system is designed around a divide-and-conquer strategy, where the inspection area is partitioned into subregions and assigned to different drones operating in parallel. Each UAV performs autonomous flight and real-time SLAM using the RTAB-Map framework. RTAB-Map is chosen for its ability to support multi-session SLAM, loop closure detection, and graph optimization.

For autonomous control and navigation, the MAVROS package is utilized to interface with the PX4 autopilot system. MAVROS allows the drones to follow predefined waypoint trajectories with precision while maintaining stable flight dynamics. During the mission, each drone collects visual and depth data, processes local maps in real time, and stores them for post-mission integration.

Following data collection, all local maps are merged offline using RTAB-Map's multi-session processing capabilities. This fusion process detects inter-session loop closures, minimizes accumulated drift, and produces a unified, metrically accurate 3D map of the inspected structure. The offline nature of the fusion also helps overcome real-time constraints such as limited bandwidth, unsynchronized clocks, and CPU/GPU load on embedded platforms.

In summary, the proposed system enables scalable, efficient, and accurate 3D infrastructure inspection by combining the strengths of RTAB-Map for SLAM, MAVROS for UAV control, and collaborative data fusion. This approach not only reduces overall inspection time but also enhances the reliability and detail of the final map, paving the way for broader adoption of autonomous drone-based inspection in industrial and civil engineering applications.

II. Related Work

Numerous methodologies have been explored for enabling autonomous SLAM (Simultaneous Localization and Mapping)

This work was carried out as part of a course project. No external funding was involved.

in aerial robotics. Among the most widely recognized systems are ORB-SLAM and Cartographer. ORB-SLAM, in particular, is known for its feature-based tracking and real-time performance using monocular, stereo, or RGB-D cameras. It employs keyframe-based visual SLAM and provides accurate localization with loop closure capabilities. Cartographer, developed by Google, extends SLAM capabilities to 2D and 3D LiDAR data and is particularly suited for indoor mobile robotics. However, both systems are designed for single-agent SLAM and require significant modification to support multiple UAVs operating concurrently.

Efforts to scale SLAM to multiple agents have led to the development of multi-robot SLAM systems. These systems aim to improve mapping efficiency and robustness by sharing sensory data among agents. Approaches such as centralized pose graph fusion, inter-agent loop closures, and collaborative front-end data sharing have been proposed. However, such systems face challenges related to communication bandwidth, time synchronization, data association complexity, and computational load balancing. Most multi-agent SLAM frameworks depend on real-time data exchange and tight coupling, which can be impractical in outdoor or bandwidth-constrained environments.

RTAB-Map (Real-Time Appearance-Based Mapping) has emerged as a versatile SLAM framework that supports multi-session operation, a key advantage for multi-drone systems. RTAB-Map is compatible with visual, RGB-D, and LiDAR sensors, and incorporates robust loop closure detection, graph optimization, and 3D map generation. Its multi-session capability enables the post-processing fusion of independently generated maps, thereby eliminating the need for real-time inter-agent communication. This makes it especially well-suited for offline integration of data from multiple UAVs.

Complementing SLAM capabilities, MAVROS serves as the de facto middleware bridge between the Robot Operating System (ROS2) and MAVLink-based flight controllers such as PX4. MAVROS facilitates UAV control through services and topics related to position tracking, waypoint navigation, flight mode switching, and sensor feedback. While MAVROS has seen widespread adoption in single-UAV applications, its utility in coordinating multiple autonomous drones for collaborative SLAM remains relatively underexplored in both academic research and practical deployment.

The integration of RTAB-Map and MAVROS in a multi-UAV inspection framework offers a promising solution to the challenges of scalability, accuracy, and operational autonomy. However, to date, the combined use of these technologies for large-scale infrastructure inspection—especially in a decentralized, post-fusion context—has received limited attention in existing literature. This study aims to bridge that gap by demonstrating a practical and effective implementation of collaborative SLAM for infrastructural mapping using commercially available tools and platforms.

III. System Overview

The proposed system architecture is built to support scalable, collaborative SLAM using multiple drones, each outfitted with a set of hardware and software components designed for autonomous navigation and real-time map generation.

Each UAV is equipped with:

- An RGB-D camera for capturing aligned depth and color imagery essential for accurate 3D SLAM.
- A PX4-compatible flight controller to handle low-level flight dynamics and execute waypoint-based missions.

The simulation-based setup used for validation was based on the Assignment 3 - exttx500_depth_mono drone configuration.

The UAVs are orchestrated using ROS2, which facilitates inter-process communication across sensor, control, and SLAM nodes. ROS2 provides a reliable and modular framework that supports real-time data streaming, service calls, and message passing between various system components.

Key software elements in the architecture include:

- **RTAB-Map (Real-Time Appearance-Based Mapping):** Used for performing real-time visual SLAM, generating local maps, detecting loop closures, and supporting multi-session map fusion.
- **MAVROS:** Serves as the middleware bridge between ROS2 and PX4 flight controllers, providing services for waypoint navigation, state estimation, telemetry monitoring, and mission control.
- **RTAB-Map Database (.db):** After each mapping session, RTABMap writes a SQLite.db file containing all map and sensor data. This database can be used to: Export point clouds to .ply, Visualize the map directly (e.g., via RTAB-Map's GUI or RViz) and Query, merge, or post-process sessions (loop closures, map fusion, etc.).

Each UAV is assigned a specific subregion of the target infrastructure for inspection. Waypoints are generated prior to the mission based on the structural layout, and uploaded to the UAVs via MAVROS. During flight, the UAVs autonomously navigate their respective sectors, capturing synchronized RGB-D frames and building local maps in real time using RTAB-Map. These maps are stored as RTAB-Map databases and bag files.

Upon mission completion, all local SLAM outputs are transferred to a centralized workstation for offline processing. This central system, equipped with an Intel Core i5 8th generation processor and an NVIDIA GTX 1050 Ti GPU, performs multi-session fusion using RTAB-Map's post-processing tools. The fusion process detects inter-session loop closures, aligns overlapping regions, and produces an optimized global 3D map. This unified point cloud can be exported in standard formats such as PCD or PLY for further analysis or visualization.

This modular and distributed architecture not only reduces mission time by parallelizing coverage but also maintains high accuracy through map overlap and offline optimization. It offers flexibility, allowing the system to scale from two to several drones depending on the inspection scope and available hardware.

IV. Methodology

The multi-drone SLAM system follows a structured pipeline consisting of five key stages: pre-mission trajectory planning, SLAM initialization, autonomous data collection, offline map fusion, and point cloud export. Each stage is designed to optimize efficiency, enhance SLAM accuracy, and ensure seamless data integration across multiple UAV platforms.

Before adopting RTAB-Map as the primary SLAM solution, we initially explored the use of ORB-SLAM3 due to its advanced capabilities and support for visual-inertial SLAM. While we were able to build the necessary packages and dependencies, we encountered persistent issues when attempting to simulate the pipeline. Despite extensive debugging efforts and experimentation with ROS2 integration, we were ultimately unable to execute a working simulation of ORB-SLAM3 in our multi-agent context. A significant portion of project time was spent attempting to resolve these issues. Given the time constraints and limited documentation available for ROS2 compatibility, we made the decision to pivot to RTAB-Map. This framework provided robust out-of-the-box ROS2 support, active community maintenance, and ready-to-use multi-session SLAM tools—making it a more practical and reliable choice for our objectives.

A. Trajectory Planning

The planning phase involves generating mission-specific flight paths for each drone based on a pre-analyzed model or schematic of the infrastructure. The inspection site is subdivided into overlapping subregions using spatial segmentation algorithms or manual zoning in a 3D environment. Each subregion is assigned to a specific UAV, ensuring minimal inter-drone interference while maintaining sufficient overlap for inter-session loop closures.

The trajectory that is generated follows a boustrophedon or lawn-mower pattern that has been rotated vertically and curved to effectively wrap around the structure of interest. This pattern ensures consistent and dense coverage while avoiding blind spots. A representative visualization of this 3D zigzag path is included in the results section.

Waypoint trajectories are generated using tools such as QGroundControl or ROS-based planners, which support exporting waypoint files compatible with PX4 flight controllers. These waypoints account for obstacles, structural geometry, line-of-sight constraints, and sensor field of view. The planned flight paths are tested in simulation using Gazebo or PX4 SITL to validate flight dynamics, mission duration, and expected SLAM coverage.

B. SLAM Initialization

RTAB-Map is configured and launched on each UAV at mission start. Critical initialization parameters include:

- Working memory and time thresholds for map updates
- Minimum number of features for keyframe insertion

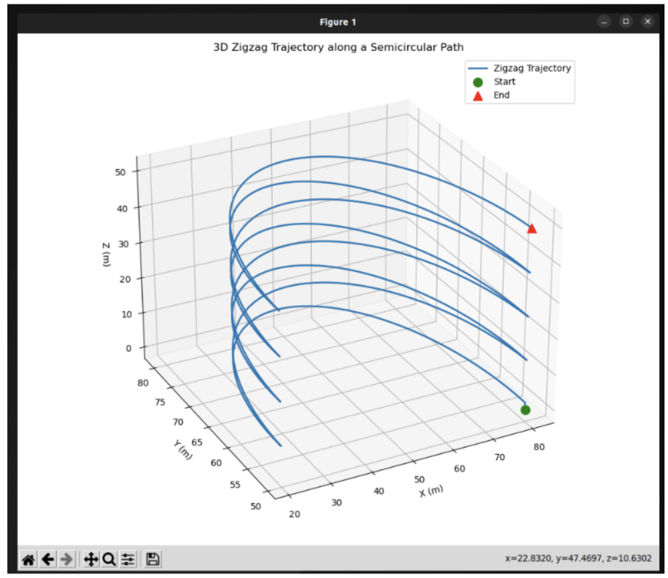

Fig. 1. 3D Zigzag Trajectory along a Semicircular Path. This vertically rotated lawn-mower pattern ensures dense and consistent coverage for infrastructure inspection.

- Loop closure detection thresholds
- Maximum depth range for RGB-D filtering

Initial sensor calibration and IMU alignment are performed to ensure accurate data capture. The system begins collecting synchronized RGB and depth data as the UAV takes off. Local SLAM maps are constructed in real time through visual odometry, feature-based graph expansion, and local loop closures. RTAB-Map's database is continuously updated with new nodes (keyframes) and edges (transforms) in the pose graph.

C. Data Collection

Each UAV autonomously follows its assigned trajectory using MAVROS in Offboard mode. MAVROS publishes trajectory setpoints and reads sensor feedback, ensuring stable flight. RGB-D data is collected at approximately 15–30 FPS depending on the sensor and computational capability. Each UAV continuously logs:

- RGB and depth image streams
- Inertial measurements
- Pose estimates from SLAM
- Environmental transforms (TF data)
- RTAB-Map's internal graph structure

All data is stored in ROS2 bag files, ensuring reliable offline access for fusion. SLAM status and UAV telemetry are monitored in real-time using Rviz2 and companion ground stations. Safety mechanisms such as geofencing and failsafe landings are integrated into the control logic.

D. Offline Fusion

After all missions are completed, the recorded data is offloaded from each UAV's storage and aggregated on a central processing workstation. Each UAV's RTAB-Map database is individually loaded, and RTAB-Map's multi-session import functionality is employed to fuse the maps.

The fusion process consists of the following steps:

- Identification of overlapping keyframes between sessions
- Detection of inter-session loop closures using visual similarity metrics (e.g., bag-of-words)
- Global graph optimization using pose graph solvers like TORO or GTSAM
- Consistency checks for spatial alignment and error minimization

This fusion step is crucial for producing a metrically consistent map, correcting accumulated drift, and integrating partial reconstructions into a coherent global model.

E. Point Cloud Export

Following graph optimization, the complete 3D map is exported in point cloud format (e.g., .ply, .pcd). RTAB-Map provides built-in tools to export dense or filtered maps with adjustable resolution, voxel size, and depth cutoffs.

The exported map undergoes additional processing using software such as MeshLab or CloudCompare to:

- Apply voxel grid filtering for downsampling
- Remove outliers and sensor noise
- Perform surface reconstruction for mesh generation
- Visualize spatial features for structural assessment

These outputs can be used to generate inspection reports, calculate volumetric or structural metrics, and enable augmented reality overlays for engineers in the field.

This comprehensive methodology ensures a robust and scalable workflow for infrastructure inspection, leveraging modular UAV deployments, robust SLAM algorithms, and efficient data fusion techniques. The approach is adaptable to different inspection scenarios, sensor payloads, and mission complexities, supporting real-world deployment in varied environmental conditions.

Author Contributions

Khushal Hemant Sharma
Khushal focused on designing and implementing the autonomous flight trajectory strategy. He researched various coverage methods and created a vertically rotated boustrophedon (lawn-mower) path tailored for infrastructure inspection. He developed visualization scripts to verify path accuracy and ensured smooth integration into ROS2-compatible flight planners. Alongside Mainak, he configured the RTAB-Map SLAM stack for distributed mapping, tuning key parameters for memory efficiency, loop closure performance, and graph quality. He was also involved in testing mapping accuracy and generating trajectory and pose graph visualizations.

Sakshi Sanjay Khade
Sakshi led the development of the multi-UAV spawning system in Gazebo Harmonic. Working closely with Mrunmayee, she ensured each drone was initialized with clean namespace isolation and synchronized SLAM nodes. She developed the launch infrastructure needed for simultaneous multi-drone deployment and monitored ROS2 bag recordings for data integrity. Her work ensured consistent RGB-D streaming across all agents, and she played a key role in troubleshooting SLAM initialization issues to ensure smooth and parallel operation.

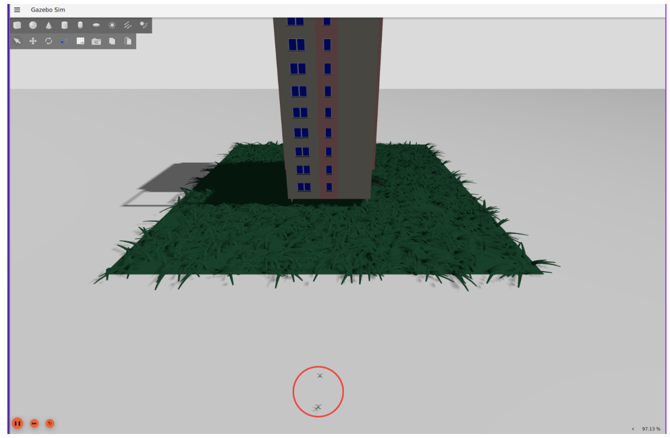

Fig. 2. Gazebo Harmonic simulation environment featuring a high-rise infrastructure mock-up. Two drones are initialized in front of the structure for inspection mission simulation.

Mainak Malay Saha
Mainak led the map merging and post-processing pipeline. He was responsible for offline fusion of independently generated maps using RTAB-Map's multi-session features. He handled cross-session loop closure detection, global graph optimization, and the resolution of spatial misalignments in the fused output. Alongside Khushal, he tuned SLAM parameters for accuracy and performance. Mainak also processed the final point cloud using tools like MeshLab and CloudCompare, and took the lead on writing the report, organizing documentation, and managing project deliverables.

Mrunmayee Nitin Valunj
Mrunmayee worked on orchestrating multiple UAVs within the simulation environment and ensuring the system's stability. In collaboration with Sakshi, she developed and tested the drone spawning system in Gazebo Harmonic. She configured MAVROS for each agent, verified offboard control activation, and managed trajectory execution across multiple drones. Her responsibilities included managing URDF/SDF configuration files, parameter loading, and resolving mission execution issues. She also helped maintain consistent SLAM operation under simultaneous multi-agent load.

V. Results and Discussion

The proposed multi-drone SLAM system was validated entirely through simulation, using the Gazebo environment to create controlled testing scenarios. This simulation platform enabled detailed experimentation, including mission rehearsal, evaluation of spatial mapping performance, and testing of the overall pipeline in structured and semi-structured virtual environments.

In a two-drone simulated setup, the mission duration was effectively halved compared to a single-drone execution due to parallel inspection of subregions. By assigning discrete yet overlapping zones to each drone, the system ensured sufficient data redundancy to support reliable inter-session loop closures.

Map fusion was conducted offline using RTAB-Map's multi-session capabilities. The merged global map showed coherent alignment between independently collected maps, confirming the robustness of the SLAM system under simulation. Overlapping regions played a key role in drift correction during global graph optimization.

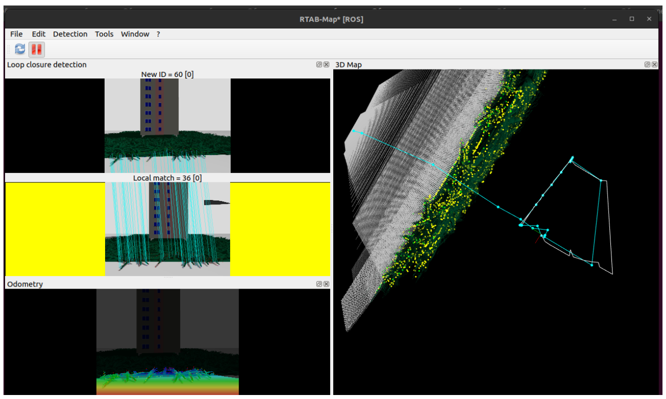

Fig. 3. Merged Point Cloud and Pose Graph Visualization in RTAB-Map. This figure illustrates loop closure detection and post-fusion alignment achieved through multi-session SLAM.

Key Performance Indicators (KPIs)

- **Mapping Accuracy:** Visual inspection of the merged point cloud indicated minimal misalignment.
- **SLAM Consistency:** Offline multi-session merging using RTAB-Map was successful and required minimal manual intervention, highlighting the reliability of automatic loop closure detection and graph optimization mechanisms.

Challenges Encountered

- **Time Synchronization:** ROS2 nodes running independently in simulation maintained separate time clocks. This occasionally resulted in timestamp mismatches during data logging, which required alignment corrections during post-processing.
- **Resource Limitations:** Although not tested on physical hardware, it is anticipated that real-time performance may be affected by the compute limitations of embedded platforms in future implementations.
- **Bandwidth Constraints:** While simulation did not replicate Wi-Fi bandwidth limitations, real-world implementation would need to consider communication latency and throughput, especially for online map merging.

Despite these constraints, offline map fusion was successfully demonstrated within the simulation pipeline. Our goal was to implement online map fusion—allowing simultaneous merging of SLAM data during mission execution—but this was not achieved due to challenges associated with inter-agent synchronization and communication latency. However, RTAB-Map's existing support for multi-session fusion presents a promising foundation for future real-time or semi-real-time integration efforts.

In conclusion, the simulation results validate the technical feasibility and scalability of the proposed approach. The system demonstrates strong potential for application in real-world scenarios once adapted to physical UAV platforms, with further development required to address online fusion and system-wide real-time constraints.