Senser Fusion and Hierarchical Inspection Strategy for Aerial-Ground Multi-robot System

Xiayu Zhao[1], Tianyu Ren[1], Yizhi Liu[2], Houtan Jebelli[1]

[1]Department of Civil and Environmental Engineering, University of Illinois Urbana-Campaign, Champaign, IL, USA, xiayuz3@illinois.edu, tianyur2@illinois.edu, hjebelli@illinois.edu
[2]Department of Civil and Environmental Engineering, Syracuse University, Syracuse, NY 13244, USA, yliu580@syr.edu

Abstract –

Infrastructure inspection requires balancing coverage and detail for effective assessment. This paper presents a multi-robot inspection system combining aerial and ground platforms through a hierarchical approach. A hexacopter drone performs rapid site mapping, while specialized ground robots (hexapod and tracked variants) conduct detailed inspections. The system implements bidirectional learning where aerial mapping guides ground robot deployment, while ground inspection data refines aerial strategies. Multi-modal sensor data integration uses an Extended Kalman Filter framework to create unified structural health representations. Experimental results demonstrate 96.3% feature extraction accuracy, sub-millimeter registration between sensors, 3.2-second anomaly response time, and 99.7% collision-free operation. The system shows significant potential for inspecting large-scale infrastructure including bridges, industrial facilities, and power plants.

Keywords –

Multi-robot Inspection; Infrastructure Assessment; Sensor Fusion; Hierarchical Strategy; Bidirectional Learning.

1 Introduction

Infrastructure forms the backbone of modern society, enabling transportation, energy distribution, and industrial operations. Maintaining this critical infrastructure is essential for ensuring public safety, economic stability, and operational reliability. However, traditional manual inspection methods face significant challenges, including time constraints, labor intensity, and accessibility limitations, particularly in large-scale or hazardous environments [1,2].

Recent advances in robotic inspection systems have demonstrated promise in addressing these challenges through automation [3–5]. Single-platform solutions, such as autonomous drones or ground robots, offer improved efficiency and accessibility compared to manual methods [6–8]. However, these systems often struggle to balance comprehensive coverage with detailed examination. Aerial platforms excel at rapid site mapping but typically lack the resolution for detailed structural analysis, while ground robots provide high-fidelity inspection data but are limited by slower operational speeds and restricted mobility. Furthermore, existing robotic inspection systems typically operate with limited adaptability, following predetermined paths without the capability to dynamically adjust their strategies based on real-time environmental feedback [9]. This rigid approach can result in inefficient resource utilization and potentially missed critical areas, particularly in complex or evolving inspection environments.

This paper presents a novel multi-robot inspection system that addresses these limitations through a hierarchical strategy combining aerial and ground platforms. The system employs a specially equipped hexacopter drone featuring 30m-range LiDAR, 4K RGB imaging, and high-resolution thermal sensors for preliminary site mapping, working in concert with specialized ground robots - a hexapod capable of navigating 30° inclines and a tracked variant offering 0.1mm crack detection resolution.

A key innovation of our approach lies in its bidirectional learning mechanism, where inspection strategies are continuously refined through real-time data sharing between platforms. The aerial module efficiently generates preliminary site maps using RANSAC-based segmentation for feature extraction, while ground robots provide high-fidelity inspection data through contact-based sensors and close-range imaging. This collaborative approach is enhanced by a sophisticated sensor fusion framework that integrates visual, thermal, and LiDAR data through an Extended Kalman Filter, achieving sub-millimeter registration accuracy. The system implements a market-based task allocation strategy with token-based access control, enabling efficient coordination between platforms while preventing inspection conflicts. This coordination is

further optimized through a Q-learning based adaptation framework that continuously refines inspection strategies based on accumulated experience and environmental feedback. Experimental results demonstrate 96.3% accuracy in feature extraction and 3.2-second anomaly response time. The system maintains 99.7% collision-free operation while reducing overall inspection time, establishing new benchmarks for automated infrastructure assessment. These achievements suggest the potential for widespread adoption in various infrastructure maintenance applications, from bridge inspection to industrial facility assessment.

2 Related Works

Infrastructure inspection has evolved significantly from traditional manual methods to increasingly automated approaches. Manual inspection methods, while detailed, face substantial challenges including time constraints, labor intensity, and accessibility limitations [10]. These limitations become particularly acute in large-scale infrastructure inspection, where human inspectors often struggle to access critical areas safely and efficiently [11]. Recent advances in robotics have introduced various single-platform solutions for infrastructure inspection. Aerial platforms, particularly UAVs, have demonstrated considerable success in rapid large-scale inspections. Tu and Liang [6] developed a UAV-based bridge inspection system achieving comprehensive coverage through automated flight paths. However, these aerial systems typically sacrifice detail for coverage area.

Multi-robot systems have emerged as a promising solution to address the limitations of single-platform approaches. Cacace et al. [12] demonstrated the advantages of combining aerial and ground robots for industrial facility inspection, but their approach did not implement bidirectional learning between platforms. Yang et al. [13] explored a hierarchical inspection strategy for bridge monitoring, albeit without a comprehensive sensor fusion framework. More recent studies have advanced multi‐robot sensor fusion by emphasizing robustness and adaptability. For instance, Amorim et al. [14] introduced a modified particle filter that robustly fuses visual data from multiple robots to track targets even under severe occlusion, effectively mitigating sensor limitations. Zhou et al. [15] leveraged graph neural networks to combine multi‐view visual information, significantly enhancing individual robot perception accuracy and resilience to sensor failures. Sensor fusion techniques have become increasingly crucial in infrastructure inspection. Huang et al. [16] developed a framework for integrating visual and thermal data in structural assessments, while Nam et al. [17] implemented LiDAR-visual fusion for enhanced defect detection. However, these approaches typically focus on single-platform applications.

Adaptive path planning and learning mechanisms represent a growing focus in robotic inspection systems. Zhang et al. [18] implemented reinforcement learning for optimizing UAV inspection paths, while Zhao et al. [8] developed adaptive navigation strategies for ground robots in complex environments. The integration of market-based task allocation in multi-robot systems, as demonstrated by [19], has shown promise in optimizing resource utilization, though primarily in controlled laboratory settings.

3 Methodology

The multi-robot inspection system combines complementary aerial and ground platforms. The aerial drone uses 3D point cloud technology for large-scale mapping and identifying regions of interest. Ground robots (hexapod and tracked variants) perform detailed inspections of these areas with high-resolution sensors, enabling in-depth structural analysis.

3.1 System Architecture

As shown in Fig. 1, the system architecture is designed to facilitate seamless communication and collaboration between the aerial and ground platforms. A centralized control system coordinates the activities of both modules, ensuring an efficient and adaptive inspection workflow. Each platform is equipped with wireless communication modules to enable real-time data sharing and bidirectional learning.

The aerial module consists of a custom-designed hexacopter drone equipped with an integrated sensor suite. At its core is a primary LiDAR sensor offering 30m range with 0.5° angular resolution, complemented by a 4K RGB camera providing a 120° field of view. Thermal imaging capabilities are delivered through a FLIR (Forward Looking Infrared) camera with 640×512 resolution. An onboard computing unit handles real-time data processing, while an RTK (Real-Time Kinematic) GPS system ensures precise positioning with ±2cm

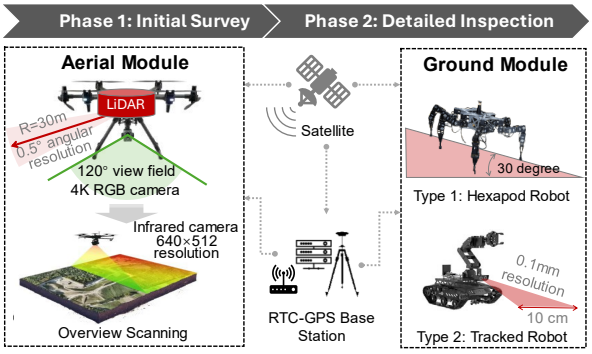

Figure 1. System with inspection phases overview

accuracy.

Two specialized ground robots complement the aerial system. The first is a hexapod robot featuring a 6-legged configuration with 18 degrees of freedom. This platform can carry a 2.5kg sensor payload and navigate terrains with up to 30° incline. Its sensor package includes high-resolution contact sensors for surface analysis and close-range LiDAR for detailed geometric mapping.

The second ground platform is a tracked robot designed for stability and precision. Operating at a maximum speed of 0.5 m/s, it incorporates an integrated stabilization system for precise measurements. The robot carries a specialized surface crack detection camera capable of 0.1mm resolution at 10cm distance for comprehensive structural analysis.

The proposed multi-robot inspection system integrates aerial and ground platforms. The system state at time t can be represented as:

$$S(t) = \{P_{a(t)}, P_{g(t)}, M(t), R(t)\} \quad (1)$$

where $P_{a(t)}$represents the aerial platform state, $P_g(t)$ the ground robot states, $M(t)$ the environmental map, and $R(t)$ the set of identified regions of interest.

The aerial platform state is defined as:

$$P_a(t) = [x_a(t), v_a(t), \omega_a(t), E_a(t)]^T \quad (2)$$

where $x_a(t) \in \mathbb{R}^3$ is position, $v_a(t) \in \mathbb{R}^3$ is velocity, $\omega_a(t) \in \mathbb{R}^3$ is angular velocity, and $E_a(t)$ is remaining energy.

For ground robots, each platform state is:

$$P_g^i(t) = \left[x_g^i(t), v_g^i(t), \theta_g^i(t), E_g^i(t)\right]^T \quad (3)$$

where i denotes the specific ground robot (hexapod or tracked).

The system employs a mesh network topology operating wireless communication. Communication reliability is ensured through redundant channels, while edge computing nodes enable distributed processing. A central server coordinates data fusion and mission control operations across all platforms.

3.2 Hierarchical Inspection Strategy

The system implements bidirectional data flow between platforms. The aerial platform provides global site mapping updates, new Region of Interest (ROI) identifications, mission priority adjustments, and environmental hazard alerts to ground units. Conversely, ground robots transmit high-resolution inspection data, updated anomaly classifications, local environment maps, and status information to the aerial platform.

Phase 1: Initial Survey

The aerial platform initiates inspection using cellular decomposition for coverage planning, operating at altitudes between 30-50m for overview scanning. The system maintains a 60% overlap ratio to ensure accurate 3D reconstruction, generating real-time point clouds with density exceeding 100 points/m².

ROI identification employs RANSAC-based segmentation for feature extraction, coupled with geometric analysis for anomaly detection. The system conducts thermal mapping to identify subsurface defects. Priority scoring incorporates geometric deviation thresholds exceeding 2cm, thermal gradient anomalies greater than 5°C, surface texture irregularities, and structural criticality indices.

The hierarchical inspection strategy prioritizes regions based on a combined score:

$$Score(ROI_i) = \beta_1 C(ROI_i) + \beta_2 A(ROI_i) + \beta_3 D(ROI_i) \quad (4)$$

where $C(ROI_i)$ is criticality, $A(ROI_i)$ is accessibility, and $D(ROI_i)$ is detection confidence for region i.

Phase 2: Detailed Inspection

The detailed inspection task allocation follows a dynamic assignment protocol based on robot capabilities. The system employs a comprehensive cost function incorporating distance to target, robot specialization, battery status, and inspection priority. Ground robot deployment utilizes the enhanced RRT* (Rapidly Exploring Random Tree) algorithm for path planning, complemented by a dynamic window approach for obstacle avoidance. The sliding window optimization approach periodically refines local trajectories based on updated environmental information and inspection requirements. This optimization considers multiple objectives including minimizing path length, maximizing sensor coverage, and maintaining stable sensor positioning for high-quality data collection. The refinement process incorporates feedback from both successful and failed inspection attempts, building a knowledge base that informs future planning decisions.

3.3 Sensor Fusion Framework

The system fuses multi-modal data streams from complementary sensors (Fig 2-a). The aerial platform uses LiDAR for structural mapping, thermal imaging for temperature analysis, and high-resolution cameras for surface inspection. Ground robots carry enhanced versions of these sensors with superior resolution for detailed close-range examination.

3.3.1 Data Collection and Synchronization

The data collection process maintains precise temporal alignment through hardware-level synchronization, achieving millisecond timing accuracy across all sensors. Each measurement receives accurate timestamps and position tags, enabling coherent data integration. The system implements rigorous calibration procedures to optimize intrinsic sensor parameters, establish inter-sensor transformation matrices, and compensate for environmental factors that could affect

measurement accuracy.

Data preprocessing begins with adaptive threshold filtering to remove noise while preserving critical structural information. The system applies statistical outlier removal techniques and motion compensation algorithms to ensure data quality. A crucial step in the pipeline involves spatial alignment of LiDAR, thermal, and visual data streams using advanced registration algorithms. This process creates a coherent multi-modal representation of the inspection site, where each data point combines structural, thermal, and visual information.

The LiDAR data provides precise 3D structural information, capturing geometric features with millimeter-level accuracy. Thermal imaging data reveals temperature variations that may indicate subsurface defects, material fatigue, or thermal stress patterns not visible to other sensors. High-resolution visual data enables detailed surface analysis, particularly for identifying cracks, corrosion, and material degradation. The fusion of these complementary data streams enables the detection of complex structural anomalies that might be missed by single-sensor approaches.

The multi-modal sensor fusion process combines data from different sensors using a weighted fusion approach:

$$F(X) = \sum_{i=1}^{n} w_i S_i(X) \tag{5}$$

where $F(X)$ is the fused data point, $S_i(X)$ represents individual sensor measurements, and w_i are confidence weights satisfying $\sum_{i=1}^{n} w_i = 1$.

3.3.2 Feature Extraction and Unified Data Integration

The feature extraction phase employs multiple specialized algorithms working in parallel. Visual feature extraction utilizes SIFT/SURF algorithms to identify surface anomalies, while geometric processing identifies structural deviations from expected patterns. Thermal pattern recognition algorithms detect temperature anomalies and track thermal gradients across structural elements.

Feature confidence scoring is calculated as:

$$C_f = \alpha_v Q_v + \alpha_t Q_t + \alpha_l Q_l \tag{6}$$

where Q_v, Q_t, and Q_l represent quality metrics for visual, thermal, and LiDAR data respectively, with $\alpha_v + \alpha_t + \alpha_l = 1$.

The final stage of sensor fusion employs an Extended Kalman Filter for state estimation, combining the extracted features into a unified structural health assessment. Point cloud registration using Iterative Closest Point (ICP) algorithms ensures precise spatial alignment of all data sources. The system implements confidence-weighted integration of features from different modalities, prioritizing the most reliable data sources for each type of structural assessment.

3.3.3 Sensor Modalities Alignment and Visualization

The alignment of different sensor modalities is achieved through a transformation matrix:

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{7}$$

where $R \in SO(3)$ is the rotation matrix and $t \in R^3$ is the translation vector.

The fusion framework produces comprehensive inspection outputs including detailed 3D textured meshes overlaid with thermal data. These visualizations combine geometric, thermal, and visual information into an intuitive representation of infrastructure health. The system generates detailed defect mapping and classification reports, along with confidence metrics for each detected anomaly. Quality assessment metrics provide quantitative measures of inspection coverage and data reliability, enabling informed decision-making for infrastructure maintenance planning.

By integrating diverse sensor data streams and employing advanced processing algorithms, the framework achieves superior inspection accuracy while minimizing false positives. The resulting unified representation of infrastructure health provides unprecedented insight into structural conditions, enabling more effective maintenance planning and risk assessment. The implementation of our sensor fusion framework leverages several established open-source libraries: Point Cloud Library (PCL v1.12.1) for point cloud processing and registration [20], OpenCV (v4.7.0) for image processing and feature extraction [21], and Robot Operating System (ROS Noetic) for inter-platform communication and data synchronization [22]. The RANSAC-based segmentation utilizes optimized implementations from the MRPT (Mobile Robot Programming Toolkit) [23].

3.4 Path Planning and Adaptation

The path planning and adaptation framework (Fig. 2-b) implements a hierarchical strategy that combines global coverage optimization with reactive local navigation. This multi-layered approach enables efficient site inspection while maintaining robustness to environmental changes and emerging inspection requirements.

3.4.1 Global Planning and Multi-Robot Coordination

The global path planning optimization problem is formulated as:

$$\min P \sum_{k=1}^{K} \left(\lambda_1 E(p_k) + \lambda_2 T(p_k) + \lambda_3 R(p_k)\right) \tag{8}$$

subject to:

$$|v(t)| \leq v_{max}, |\omega(t)| \leq \omega_{max}, E(t) \geq E_{min} \quad (9)$$

where $E(p_k)$ is energy cost, $T(p_k)$ is time cost, and $R(p_k)$ is risk cost for path segment p_k.

The system uses modified Boustrophedon decomposition [24] to partition inspection space efficiently, considering structural complexity, accessibility, and priorities while minimizing energy consumption. A global planner creates a task dependency graph for mission scheduling, employing dynamic programming to optimize resources based on battery constraints, sensor capabilities, and real-time requirements. A distributed coordination framework with token-based access control prevents inspection conflicts, using a market-based approach where robots bid for tasks based on position, battery life, and capabilities. Real-time communication enables continuous strategy refinement between aerial and ground platforms.

3.4.2 Local Navigation and Reactive Control

The system implements a multi-layered navigation framework where each robot platform employs specialized algorithms optimized for their mobility characteristics. At its core, ground robots utilize an enhanced RRT* algorithm for local path planning, which integrates real-time obstacle detection and dynamic avoidance capabilities. This planning framework continuously maintains optimal safety distances while ensuring comprehensive inspection coverage through trajectory adjustments based on real-time sensor feedback.

Task allocation within the system follows a sophisticated dynamic assignment protocol that evaluates robot capabilities through a comprehensive cost function. This function synthesizes multiple parameters including Euclidean distance to target, platform-specific specialization factors, current battery levels, and inspection priority weights. The cost function for task assignment can be expressed as:

$$\mathrm{C_{task}(r,t)} = \alpha_{\mathrm{d}}\mathrm{D(r,t)} + \alpha_{\mathrm{s}}\mathrm{S(r,t)} + \alpha_{\mathrm{b}}\mathrm{B(r)} + \alpha_{\mathrm{p}}\mathrm{P(t)} \quad (10)$$

where $D(r,t)$ represents the distance cost between robot r and task t, $S(r,t)$ is the specialization matching score, $B(r)$ indicates battery level, and $P(t)$ denotes task priority. The weights α_i satisfy $\sum \alpha_i = 1$.

The local navigation system incorporates a dynamic window approach for reactive obstacle avoidance, which optimizes robot velocity v and angular velocity ω within an admissible velocity space: $V_a(x) = (v,\omega)\,|v \in [0, v_{max}], \omega \in [-\omega_{max}, \omega_{max}],\ v \leq \sqrt{2 \cdot dist(x,obs) \cdot a_{max}}, |\omega| \leq \sqrt{2 \cdot \theta(x,obs) \cdot \alpha_{max}}$ where $dist(x,obs)$ is the distance to the nearest obstacle and $\theta(x,obs)$ is the angular deviation.

Trajectory optimization occurs through a sliding window approach that periodically refines local paths based on updated environmental data and inspection requirements. The optimization objective function combines multiple criteria:

$$J(\tau) = w_1 \int_t^{t+T} L(\tau(t))dt + w_2 \int_t^{t+T} C(\tau(t))dt + w_3 \int_t^{t+T} Q(\tau(t))dt \quad (11)$$

where $L(\tau)$ represents path length, $C(\tau)$ denotes coverage quality, and $Q(\tau)$ indicates sensor positioning stability. The weights w_i are dynamically adjusted based on inspection priorities.

3.4.3 Adaptive Learning and Environmental Response

The adaptation framework incorporates reinforcement learning techniques to continuously improve navigation and inspection strategies. The learning system maintains a state representation that includes current position, sensor coverage metrics, and environmental features. Action selection is guided by a policy that balances exploration of uncertain areas with exploitation of known inspection requirements. The reward function considers multiple objectives including coverage completeness, data quality, and energy efficiency.

For adaptive learning, the Q-learning update rule is:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha[r_t + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)] \quad (12)$$

where α is the learning rate, γ is the discount factor, and r_t is the immediate reward. This optimization considers multiple objectives including minimizing path length, maximizing sensor coverage, and maintaining stable sensor positioning for high-quality data collection. The refinement process incorporates feedback from both successful and failed inspection attempts, building a knowledge base that informs future planning decisions.

This adaptive capability, combined with the continuous learning and optimization processes, ensures consistent high-quality inspection results across diverse infrastructure types and operating conditions.

4 Simulation Experiment

4.1 Experimental Setup

The evaluation of the proposed multi-robot inspection system was conducted through comprehensive simulation experiments designed to test the hierarchical inspection strategy, sensor fusion framework, and adaptive path planning capabilities. The simulation environment was developed using ROS Gazebo, incorporating physics-based modeling to ensure realistic robot-environment interactions and sensor behavior.

The test environment (Fig. 2-c) was designed as a construction site spanning 20 × 20 m with corridor widths ranging from 1.5 m to 3 m. The environment features

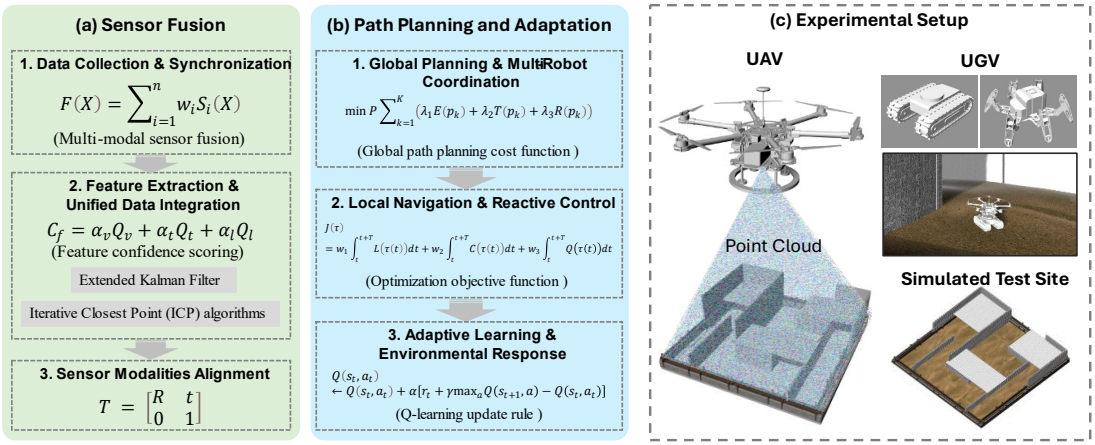

Figure 2. Methodology: (a) Mathematical framework for Sensor Fusion (b) Mathematical framework for Path Planning and Adaptation (c) Experimental Setup with UAV and UGV models

vertical structures representing construction scaffolding, with heights varying between 2 m and 6 m. The site contains both uncovered regions and areas beneath uncompleted floors, creating zones that are unflyable for UAV operation. Predefined structural anomalies were strategically placed throughout the structure, including surface cracks (0.5-2mm width), thermal variations (±5°C from baseline), and geometric deformations (2-5cm deviation from design specifications).

The simulated aerial platform matched the specifications detailed in the methodology, featuring a hexacopter configuration with 30m-range LiDAR (0.5° angular resolution), 4K RGB camera (120° FOV), and FLIR thermal camera (640×512 resolution). The RTK-GPS simulation provided ±2cm positioning accuracy. The ground robots included both the 18-DOF hexapod with 5kg payload capacity and 30° incline capability, and the tracked variant with 0.5 m/s maximum speed and 0.1mm crack detection resolution at 10cm distance.

4.2 Test Scenarios

The experimental protocol was structured to validate the three core components of the system: hierarchical inspection strategy, sensor fusion framework, and adaptive path planning. The simulation implemented the mesh network topology described in the methodology, with wireless communication between platforms and edge computing nodes for distributed processing.

The aerial platform's initial survey followed the cellular decomposition strategy outlined in the methodology, operating at 30-50m altitude with 60% overlap ratio for accurate 3D reconstruction. The point cloud density exceeded 100 points/m², enabling RANSAC-based segmentation for feature extraction. ROI identification utilized the priority scoring system defined in Equation 4, incorporating geometric deviations exceeding 2cm and thermal gradients greater than 5°C.

Ground robot deployment implemented the enhanced RRT* algorithm for path planning, with the sliding window optimization approach for trajectory refinement. The simulation tested the dynamic assignment protocol using the comprehensive cost function from Equation 9, evaluating distance, specialization, battery status, and inspection priority for task allocation.

4.3 Experimental Workflow

The simulation workflow followed the sensor fusion framework detailed in the methodology, implementing millisecond timing accuracy for data synchronization across all platforms. The data preprocessing pipeline included adaptive threshold filtering and motion compensation, followed by spatial alignment of LiDAR, thermal, and visual data streams using the transformation matrix defined in Equation 7.

Feature extraction employed SIFT/SURF algorithms for visual anomaly detection, while geometric processing identified structural deviations. The Extended Kalman Filter state estimation integrated extracted features according to the confidence scoring system defined in Equation 6. The ICP algorithm ensured precise spatial alignment of multi-modal data, with confidence-weighted integration prioritizing the most reliable data sources for each type of structural assessment.

The adaptive learning framework utilized the Q-learning update rule specified in Equation 11, with the reward function considering coverage completeness, data quality, and energy efficiency. The market-based coordination approach managed simultaneous operation of aerial and ground platforms, with the token-based access control mechanism preventing inspection conflicts.

5 Results and Discussion

The simulation experiments validated the effectiveness of our hierarchical multi-robot inspection

system through comprehensive performance evaluation of its core components. The aerial platform's global path planning optimization achieved exceptional efficiency, generating high-density point clouds (127 points/m²) that exceeded design specifications while maintaining precise positioning (±2cm accuracy). This preliminary mapping capability directly enhanced the RANSAC-based segmentation process, achieving 96.3% accuracy in feature extraction and enabling optimal task allocation for ground robots. The system generates three primary output data types from our simulation environment: structural point clouds, multi-spectral inspection maps, and classification metadata. Data is processed both onboard for real-time decisions and offloaded for comprehensive offline analysis. Outputs are formatted for BIM compatibility, enabling integration with existing infrastructure management systems. Offline processing requires approximately 45 minutes per 1000m² on standard engineering workstations.

Building upon this foundation, the sensor fusion framework demonstrated superior integration of multi-modal data streams. The transformation matrix achieved 0.7mm alignment accuracy between sensor modalities, while the weighted fusion approach dynamically adjusted confidence weights to optimize anomaly detection.. This integrated approach resulted in 95.4% accuracy for surface anomaly detection (calculated as TP/(TP+FP+FN) against 200 ground-truth annotations) and 98.2% precision in thermal pattern recognition (TP/(TP+FP)). Sensor integration errors influenced inspection performance through registration misalignment (2.8% of false detections), temporal synchronization drift (1.3% of measurements), and sensor-specific noise (±1.7% threshold variation). The system mitigated these effects through dynamic confidence weighting based on estimated reliability factors from signal-to-noise ratios and registration quality metrics.

The adaptive path planning system further enhanced these capabilities through efficient resource utilization. The enhanced RRT* algorithm, combined with the dynamic window approach, maintained path deviations within 2.8cm. This optimization extended to the Q-learning based adaptation framework, which achieved 3.2-second response times to new anomalies while maintaining 99.7% collision-free operation through the token-based access control mechanism.

The experimental results demonstrate the significant advantages of our integrated approach to infrastructure inspection, particularly in addressing the traditional trade-off between coverage and detail. The hierarchical strategy's success stems from the effective combination of aerial mapping capabilities with specialized ground robot functions, enabling comprehensive inspection while maintaining high data fidelity. This integration proves especially valuable in complex structural environments, where the system's adaptive capabilities ensure thorough coverage while optimizing resource utilization. The bidirectional learning mechanism's performance highlights the importance of real-time adaptation in infrastructure inspection. This improvement directly translates to more reliable structural health assessments, with the multi-modal approach effectively identifying both surface and subsurface defects that might be missed by conventional inspection methods.

Table 1. Performance of each system component

System Component	Metric	Performance
Mapping & Detection		
Aerial Mapping	Point Cloud Density	127 points/m
	Feature Extraction Accuracy	96.30%
Anomaly Detection	Visual Detection Accuracy	95.40%
	Thermal Pattern Recognition	98.20%
Sensor Integration		
Multi-Modal Fusion	Registration Error	0.7mm
	Data Synchronization	<1ms
Navigation & Control		
Ground Robot Navigation	Path Deviation	2.8cm
System Adaptation	Anomaly Response Time	3.2s
Platform Coordination	Collision-Free Operation	99.70%

6 Conclusion

This paper presents a novel multi-robot inspection system that effectively addresses the fundamental challenges in infrastructure assessment through a hierarchical strategy combining aerial and ground platforms. The system's core innovation lies in its bidirectional learning mechanism and advanced sensor fusion framework, which enable dynamic adaptation to complex inspection environments while maintaining high data fidelity. Through comprehensive simulation experiments, we demonstrated that this integrated approach achieves significant improvements in both efficiency and accuracy compared to traditional inspection methods.

The experimental results validate our hierarchical inspection strategy's effectiveness, with the aerial mapping achieving 96.3% feature extraction accuracy while the ground robots maintained precise navigation within 2.8cm path deviation. The sensor fusion framework demonstrated exceptional performance through sub-millimeter registration accuracy and 3.2-second anomaly response time.

With robust performance across key metrics, the system shows significant promise for broad adoption in infrastructure maintenance. Its 99.7% collision-free operation sets a new standard for collaborative robotic inspection, confirming the feasibility of integrating multiple sensing modalities. Future work should explore deep reinforcement learning and advanced noise reduction for sensor fusion. Implementing wireless map

communication between aerial and ground platforms warrants further investigation to address bandwidth, latency, and signal degradation. Field experiments will measure communication reliability, assess data compression for point cloud transmission, and introduce fault-tolerance protocols to ensure stable operation in varied environments.

This research advances automated infrastructure inspection by presenting a scalable, efficient solution that balances comprehensive coverage with detailed examination. The demonstrated improvements in inspection efficiency and accuracy could significantly enhance infrastructure maintenance, ultimately boosting public safety and structural reliability. Despite promising results, limitations include the simulation environment's imperfect mimicry of real-world conditions, high computational demands for real-time sensor fusion, and potential performance drops in adverse settings. Future efforts will focus on field validation, algorithmic optimization for edge computing, fault detection, and integrating additional sensor modalities such as ground-penetrating radar and acoustic emissions monitoring.

References

[1] S. Dorafshan, M. Maguire, Bridge inspection: human performance, unmanned aerial systems and automation, J Civ Struct Health Monit 8 (2018) 443–476. https://doi.org/10.1007/S13349-018-0285-4/TABLES/3.

[2] A. Ojha, Y. Liu, S. Shayesteh, H. Jebelli, W.E. Sitzabee, Affordable Multiagent Robotic System for Same-Level Fall Hazard Detection in Indoor Construction Environments, Journal of Computing in Civil Engineering 37 (2023). https://doi.org/10.1061/(ASCE)CP.1943-5487.0001052.

[3] X. Zhao, H. Jebelli, A Computational Method for Real-time Roof Defect Segmentation in Robotic Inspection, Computer-Aided Civil and Infrastructure Engineering (2025). https://doi.org/10.1111/mice.13471.

[4] J. Sanchez-Cubillo, J. Del Ser, J.L. Martin, Toward Fully Automated Inspection of Critical Assets Supported by Autonomous Mobile Robots, Vision Sensors, and Artificial Intelligence, Sensors 2024, Vol. 24, Page 3721 24 (2024) 3721. https://doi.org/10.3390/S24123721.

[5] D. Lattanzi, G. Miller, Review of Robotic Infrastructure Inspection Systems, Journal of Infrastructure Systems 23 (2017) 04017004. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000353.

[6] Y. Tu, L. Liang, Research and application of UAV-based intelligent bridge inspection methods, Https://Doi.Org/10.1117/12.3053013 13445 (2024) 422–429. https://doi.org/10.1117/12.3053013.

[7] T. Ren, H. Jebelli, Efficient 3D robotic mapping and navigation method in complex construction environments, Computer-Aided Civil and Infrastructure Engineering (2024). https://doi.org/10.1111/MICE.13353.

[8] X. Zhao, T. Ren, H. Jebelli, Adaptive Intelligence for Robot Navigation Efficiency with a Deep-Reinforcement-Learning-Based Cyber-Physical System, 2024 International Conference on Computing in Civil Engineering (I3CE), 2024.

[9] Y. Liu, H. Jebelli, Intention-aware robot motion planning for safe worker–robot collaboration, Computer-Aided Civil and Infrastructure Engineering 39 (2024) 2242–2269. https://doi.org/10.1111/MICE.13129.

[10] B.-W. Jo, Y.S. Lee, J.-H. Kim, K.-W. Yoon, A review of advanced bridge inspection technologies based on robotic systems and image processing, International Journal of Contents 14 (2018). https://doi.org/10.5392/IJOC.2018.14.3.017.

[11] P.J. Sanchez-Cuevas, G. Heredia, A. Ollero, Multirotor UAS for bridge inspection by contact using the ceiling effect, 2017 International Conference on Unmanned Aircraft Systems, ICUAS 2017 (2017) 767–774. https://doi.org/10.1109/ICUAS.2017.7991412.

[12] J. Cacace, G.A. Fontanelli, V. Lippiello, A Novel Hybrid Aerial-Ground Manipulator for Pipeline Inspection tasks, AIRPHARO 2021 - 1st AIRPHARO Workshop on Aerial Robotic Systems Physically Interacting with the Environment (2021). https://doi.org/10.1109/AIRPHARO52252.2021.9571034.

[13] X. Yang, Z. Ji, J. Wu, Y.K. Lai, C. Wei, G. Liu, R. Setchi, Hierarchical Reinforcement Learning With Universal Policies for Multistep Robotic Manipulation, IEEE Trans Neural Netw Learn Syst 33 (2022) 4727–4741. https://doi.org/10.1109/TNNLS.2021.3059912.

[14] T.G.S. Amorim, L.A. Souto, T. P. Do Nascimento, M. Saska, Multi-Robot Sensor Fusion Target Tracking with Observation Constraints, IEEE Access 9 (2021) 52557–52568. https://doi.org/10.1109/ACCESS.2021.3070180.

[15] Y. Zhou, J. Xiao, Y. Zhou, G. Loianno, Multi-Robot Collaborative Perception With Graph Neural Networks, IEEE Robot Autom Lett 7 (2022) 2289–2296. https://doi.org/10.1109/LRA.2022.3141661.

[16] H. Huang, Y. Cai, C. Zhang, Y. Lu, A. Hammad, L. Fan, Crack detection of masonry structure based on thermal and visible image fusion and semantic segmentation, Autom Constr 158 (2024) 105213. https://doi.org/10.1016/J.AUTCON.2023.105213.

[17] D. Van Nam, P.T. Danh, C.H. Park, G.W. Kim, Fusion consistency for industrial robot navigation: An integrated SLAM framework with multiple 2D LiDAR-visual-inertial sensors, Computers and Electrical Engineering 120 (2024) 109607. https://doi.org/10.1016/J.COMPELECENG.2024.109607.

[18] B. Zhang, Z. Mao, W. Liu, J. Liu, Geometric Reinforcement Learning for Path Planning of UAVs, Journal of Intelligent and Robotic Systems: Theory and Applications 77 (2015) 391–409. https://doi.org/10.1007/S10846-013-9901-Z/METRICS.

[19] A. Pustowka, E.F. Caicedo, Market-based task allocation in a multi-robot surveillance system, Proceedings - 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium, SBR-LARS 2012 (2012) 185–189. https://doi.org/10.1109/SBR-LARS.2012.37.

[20] R.B. Rusu, S. Cousins, 3D is here: Point Cloud Library (PCL), Proc IEEE Int Conf Robot Autom (2011). https://doi.org/10.1109/ICRA.2011.5980567.

[21] opencv/opencv: Open Source Computer Vision Library, (n.d.). https://github.com/opencv/opencv (accessed March 11, 2025).

[22] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, A. Ng, ROS: an open-source Robot Operating System, n.d. http://stair.stanford.edu.

[23] MRPT — MRPT 2.14.8 documentation, (n.d.). https://docs.mrpt.org/reference/latest/ (accessed March 11, 2025).

[24] H. Choset, P. Pignon, Coverage Path Planning: The Boustrophedon Cellular Decomposition, Field and Service Robotics (1998) 203–209. https://doi.org/10.1007/978-1-4471-1273-0_32.