Research Statement

Jianhao Jiao (ucacjji@ucl.ac.uk), Department of Computer Science, University College London.

I. Vision: Lifelong SLAM and Embodied Intelligence for Field Robots

A. Background

Robots are leaving labs for unstructured environments where they must perceive, reason, and act under uncertainty. Resilient autonomy underpins disaster response, mining, infrastructure inspection, agriculture, ports, and home services, improving productivity and reducing risk. This transition exposes core gaps in robustness, adaptation, and safety. Environmental complexity grows as robots leave structured spaces for GPS-denied, dynamic, and harsh terrains with smoke, rain, or poor lighting. Task diversity expands from executing single skills to multi-objective missions that combine locomotion, manipulation, search, and resource optimization. System sustainability becomes essential, demanding safe, long-term operation with resilience to weather, degradation, and shifting mission requirements. The result is a gap in online adaptation, multi-objective planning, physics-consistent control, and distributed coordination.

Despite progress, current "foundation" navigation and manipulation policies remain difficult to field at scale. They are compute- and energy-intensive for mobile robots, heavily reliant on supervised datasets instead of self-supervised multi-modal learning, brittle under distribution shift, and loosely coupled to physics- and safety-aware control. As a result, systems often exhibit slow execution, unstable motion, and low success rates in the open world. Closing the lab-to-field gap requires lifelong, resource-aware adaptation and robust decision making under uncertainty.

B. Five-Year Objectives and Scope

My five-year goal is to develop **lifelong, human-level navigation** by designing self-adaptive and resilient algorithms that maintain robust performance across diverse, changing environments. The systems will be implemented efficiently and validated rigorously on real platforms (*e.g.*, unmanned ground vehicles and legged robots). Central is lifelong learning: robots should adapt to new scenes and tasks with few-shot data while preserving safety and reliability. Concretely, "human-level" means executing everyday missions with the reliability of a routine 15-minute commute, regardless of weather, lighting, or crowding. As illustrated in Fig. 1, this vision translates into five objectives:

- *Large-scale localization and mapping.* Maintain accurate maps over large areas under compute, memory, and bandwidth limits; remain stable as environments change.
- *Appearance change.* Remain reliable under short-term dynamics and long-term variations by extracting persistent features.
- *Mobility.* Achieve stable traversal on varied terrain using few-shot traversability estimation and adaptive planning.
- *Scene understanding.* Interpret objects, human intent, and social norms to enable safe, purposeful interaction.
- *Robustness.* Remain resilient to shocks, interference, sensor degradation, and software faults through fault-tolerant interfaces and adaptive algorithms.

II. Past Research: Robust SLAM in Large-Scale and Diverse Environments

A. Problem Statement

Simultaneous Localization and Mapping (SLAM) estimates robot states while building a 3D map from sensor data. A typical pipeline has a *front end* (feature extraction and data association) and a *back end* (optimization over a geometric model). The estimation problem can be written as

$$\hat{\mathcal{X}} = \arg\min_{\mathcal{X}} f(\mathcal{X}; \mathcal{Z}), \tag{1}$$

where $f(\cdot)$ is the objective and $\mathcal{Z}$ is the set of measurements. With multiple sensors, SLAM couples with extrinsic calibration:

$$\hat{\mathcal{X}},\, \hat{\mathcal{Y}} = \arg\min_{\mathcal{X},\mathcal{Y}} f(\mathcal{X}, \mathcal{Y}; \mathcal{Z}), \tag{2}$$

where $\mathcal{Y} = \{\mathbf{y}_i\}_{i=1}^{N}$ are sensor extrinsics. This yields a classic "chicken-and-egg" coupling: accurate fusion needs good calibration, while optimal calibration requires reliable SLAM. The challenge is central to modern autonomous systems—mobile robots, mapping rigs, and AR/VR devices, whose robustness hinges on multi-sensor fusion.

Fig. 1. Real-world challenges in navigation: large-scale localization and mapping, appearance change, mobility, scene understanding, and robustness.

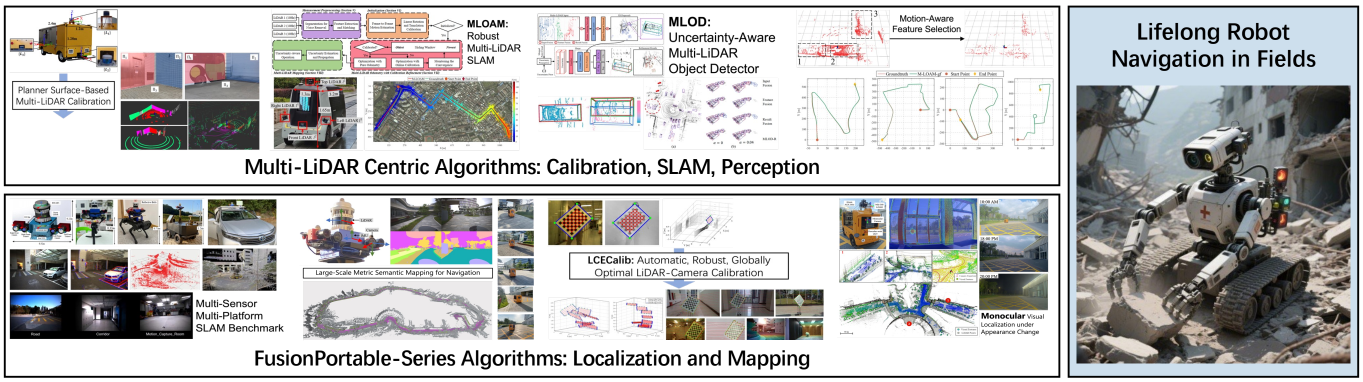

Fig. 2. Overview of past research, including multi-LiDAR centric algorithms and FusionPortable-series algorithms.

B. Contributions to Multi-LiDAR SLAM and Perception

I led the pioneering development of Simultaneous Localization and Mapping (SLAM) and online calibration algorithms for multi-LiDAR rigs, which are critical for increasing sensing range and density in fields like autonomous driving. As shown in Fig. 2 (top), key contributions include:

1) M-LOAM (Multi-LiDAR Odometry and Mapping): This extensible framework solves the joint calibration and SLAM problem, featuring online self-calibration, resource-aware feature selection (*e.g.,* greedy-based feature selection), and a two-stage mapping approach that rigorously models sensor noise and calibration error. M-LOAM achieved state-of-the-art (SoTA) performance, including localization drift under $30cm$ over $700m$ indoors, and less than 2.5% drift over $3.4km$ in urban settings.
2) Technology Transfer and Community Impact: Our work on motion-based calibration was incorporated into MATLAB's LiDAR Toolbox. M-LOAM remains widely adopted with significant open-source traction.
3) Uncertainty Integration (MLOD): I integrated the rigorous uncertainty estimation from M-LOAM into deep learning, proposing MLOD (Multi-LiDAR 3D Object Detection). This work demonstrated the capacity to enhance neural network robustness by using quantifiable geometric uncertainty to handle unreliable points, boosting average detection accuracy significantly.

C. Foundations for Field Robotics and Large-Scale Validation

To bridge the gap between lab algorithms and real-world deployment, I designed and executed systems for flexible, scalable operation in unstructured environments:

1) FusionPortable Series: I introduced the FusionPortable and FusionPortableV2 datasets, a comprehensive multi-sensor SLAM benchmark spanning 40+ sequences, $50km$, and diverse platforms (wheeled, legged, handheld). This work provides critical standardized resources for developing generalized navigation algorithms.
2) Globally Optimal Calibration (LCE-Calib): This system provided precise extrinsic calibration between LiDARs, frame cameras, and event cameras, utilizing an uncertainty-aware data association scheme and extending the QPEP solver for globally optimal accuracy, laying the technical groundwork for complex multi-modal fusion.
3) Semantic Mapping: I developed a real-time metric-semantic mapping system that generates a global mesh map for large-scale outdoor environments, enabling terrain assessment and autonomous navigation by providing crucial long-term consistency and human-readable context beyond traditional metric-only maps.

These works result in 1 IJRR, 3 TRO, 4 T-Mech, and 7 ICRA/IROS publications, demonstrating expertise in multi-modal fusion, geometric estimation, and rigorous large-scale validation across diverse robotic platforms.

TABLE I
LIST OF REPRESENTATIVE PUBLICATIONS.

Venue	Count
The International Journal of Robotics Research (IJRR) [1]	1
IEEE Transactions on Robotics (TRO) [2]–[4]	3
IEEE/ASME Transactions on Mechatronics (TMech) [5]–[8]	4
IEEE Transactions on Automation Science and Engineering (TASE) [9]	1
IEEE Robotics and Automation Magazine (RAM)	1
IEEE International Conference on Robotics and Automation (ICRA) [10]–[13]	7
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) [14]–[16]	3

III. FUTURE RESEARCH AGENDA (FIVE-YEAR PLAN)

A. Motivation

Rapid advances in AI and the emergence of capable humanoids and quadrupeds broaden the scope of real deployments in mining, disaster response, and public infrastructure. Meeting these opportunities requires robots that navigate and interact with the world with human-level reliability. My five-year agenda builds on geometric SLAM and learning-based perception to create systems that adapt over time, are safe under uncertainty, and learn efficiently from multi-modal experience. The aim is to move beyond short trials toward persistent operation that combines navigation, manipulation, tactile perception, and memory.

B. Proposed Works

1) ***Direction 1: Navigation and SLAM for Dynamic Real-World Environments****:* I will develop a next generation navigation stack that maintains accuracy and safety across large, changing environments and long durations on humanoids and quadrupeds.

- *Lifelong localization and mapping.* Design spatial-temporal map representations and uncertainty-aware estimators that update online, preserve past knowledge, and enable zero- or few-shot adaptation as lighting, structure, and routes evolve.
- *Multi-modal self-supervised fusion.* Learn to combine vision, LiDAR, IMU, and other signals with test-time adaptation and resource-aware computation so performance degrades gracefully when sensors fail or conditions worsen.
- *Safety-integrated navigation.* Couple SLAM with risk-aware planning and control, including runtime monitors and certification tools that respect physical constraints and enable operation near people and hazards.
- *Field-scale validation.* Establish long-horizon trials in mines, campuses, and mock disaster sites, including multi-robot coordination for search and inspection, with metrics on reliability, recovery time, and energy per task.

2) ***Direction 2: Embodied Manipulation, Tactile Perception, and Long-term Memory Management****:* I will endow robots with interactive skills that complement navigation so they can act purposefully in complex settings.

- *Visuo-tactile representations.* Develop foundation-model style pretraining for touch and vision to infer contact geometry and material properties, enabling generalizable, contact-rich manipulation on irregular objects and terrain.
- *Skill libraries and closed-loop control.* Build reusable manipulation and locomotion skills that compose from language or goal specifications, and use high-rate tactile feedback for force regulation, slip prevention, and recovery from failure.
- *Long-term memory for action.* Create episodic and semantic memory that links navigation and manipulation, supports planning over extended tasks and horizons, and accumulates experience to improve efficiency and safety over time.
- *Human-aware deployment.* Integrate compliant control and transparent decision making for safe collaboration, with demonstrations in rubble clearing, autonomous material handling, and assistive service tasks.

C. Potential Funding Opportunities and Collaborations

1) Funding: My prior research has been supported from several sources, including the UKRI Future Leaders Fellowship, Hong Kong Innovation and Technology Fund (ITF), the Research Grants Council's General Research Fund (GRF), the National Natural Science Foundation of China (NSFC) Young Scientists Fund (PI), and Meta Project Aria project. Aligned with my research agenda, I am actively pursuing new support industry partners such as Tencent, Huawei, and Meta. My proposals span foundational topics in robotics, embodied AI, and computer vision, together with application-driven projects in autonomous driving, home-service robotics, factory inspection, disaster response, and surveying.

2) Collaborations: I plan to collaborate with roboticists, computer-vision and machine-learning researchers. Selected collaborators from top-tier institutions include:

- Professor Dimitrios Kanoulas, University College London, UK. Topic: legged robot learning.
- Professor Simon Julier, University College London, UK. Topic: robot vision and navigation.
- Professor Martin Magnusson, Örebro University, Sweden. Topic: SLAM.
- Professor Jiwen Lu, Tsinghua University, China. Topic: embodied AI.
- Professor Rui Fan, Tongji University, China. Topic: visual navigation.

REFERENCES

[1] H. Wei, **Jiao, Jianhao**, X. Hu, J. Yu, X. Xie, J. Wu, Y. Zhu, Y. Liu, L. Wang, and M. Liu, "Fusionportablev2: A unified multi-sensor dataset for generalized slam across diverse platforms and scalable environments," *The International Journal of Robotics Research*, vol. 44, no. 7, pp. 1093–1116, 2025.

[2] **Jiao, Jianhao**, H. Ye, Y. Zhu, and M. Liu, "Robust odometry and mapping for multi-lidar systems with online extrinsic calibration," *IEEE Transactions on Robotics*, vol. 38, no. 1, pp. 351–371, 2021.

[3] P. Yin, J. Jiao, S. Zhao, L. Xu, G. Huang, H. Choset, S. Scherer, and J. Han, "General place recognition survey: Towards real-world autonomy," *IEEE Transactions on Robotics*, 2025.

[4] J. Wu, Y. Zheng, Z. Gao, Y. Jiang, X. Hu, Y. Zhu, J. Jiao, and M. Liu, "Quadratic pose estimation problems: Globally optimal solutions, solvability/observability analysis, and uncertainty description," *IEEE Transactions on Robotics*, vol. 38, no. 5, pp. 3314–3335, 2022.

[5] S. Zhang, J. He, B. Yang, Y. Zhu, J. Wu, **Jiao, Jianhao**, and J. Yuan, "Vircap: Virtual camera exposure control based on image photometric synthesis for visual slam application," *IEEE/ASME Transactions on Mechatronics*, 2024.

[6] X. Hu, L. Zheng, J. Wu, R. Geng, Y. Yu, H. Wei, X. Tang, L. Wang, **Jiao, Jianhao**, and M. Liu, "Paloc: Advancing slam benchmarking with prior-assisted 6-dof trajectory generation and uncertainty estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 29, no. 6, pp. 4297–4308, 2024.

[7] B. Yang, J. Cheng, B. Xue, **Jiao, Jianhao**, and M. Liu, "Efficient global navigational planning in 3-d structures based on point cloud tomography," *IEEE/ASME Transactions on Mechatronics*, vol. 30, no. 1, pp. 321–332, 2024.

[8] **Jiao, Jianhao**, F. Chen, H. Wei, J. Wu, and M. Liu, "Lce-calib: automatic lidar-frame/event camera extrinsic calibration with a globally optimal solution," *IEEE/ASME Transactions on Mechatronics*, 2023.

[9] **Jiao, Jianhao**, R. Geng, Y. Li, R. Xin, B. Yang, J. Wu, L. Wang, M. Liu, R. Fan, and D. Kanoulas, "Real-time metric-semantic mapping for autonomous navigation in outdoor environments," *IEEE Transactions on Automation Science and Engineering*, 2024.

[10] **Jiao, Jianhao**, R. Fan, H. Ma, and M. Liu, "Using dp towards a shortest path problem-related application," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8669–8675.

[11] **Jiao, Jianhao**, Y. Zhu, H. Ye, H. Huang, P. Yun, L. Jiang, L. Wang, and M. Liu, "Greedy-based feature selection for efficient lidar slam," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 5222–5228.

[12] J. He, H. Huang, S. Zhang, **Jiao, Jianhao**, and M. Liu, "Apmp: Accurate prior-centric monocular positioning with offline lidar fusion," *2024 International Conference on Robotics and Automation (ICRA)*, 2024.

[13] S. Zhang, J. He, B. Xue, J. Wu, P. Yin, **Jiao, Jianhao**, and M. Liu, "An image acquisition scheme for visual odometry based on image bracketing and online attribute control," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 381–387.

[14] **Jiao, Jianhao**, Y. Yu, Q. Liao, H. Ye, R. Fan, and M. Liu, "Automatic calibration of multiple 3d lidars in urban environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 15–20.

[15] **Jiao, Jianhao**, P. Yun, L. Tai, and M. Liu, "Mlod: Awareness of extrinsic perturbation in multi-lidar 3d object detection for autonomous driving," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 556–10 563.

[16] **Jiao, Jianhao**, H. Wei, T. Hu, X. Hu, Y. Zhu, Z. He, J. Wu, J. Yu, X. Xie, H. Huang *et al.*, "Fusionportable: A multi-sensor campus-scene dataset for evaluation of localization and mapping accuracy on diverse platforms," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 3851–3856.