SLAM for Robotics: Towards Lifelong Spatial Memory and Navigation in Unstructured Worlds

Jianhao Jiao

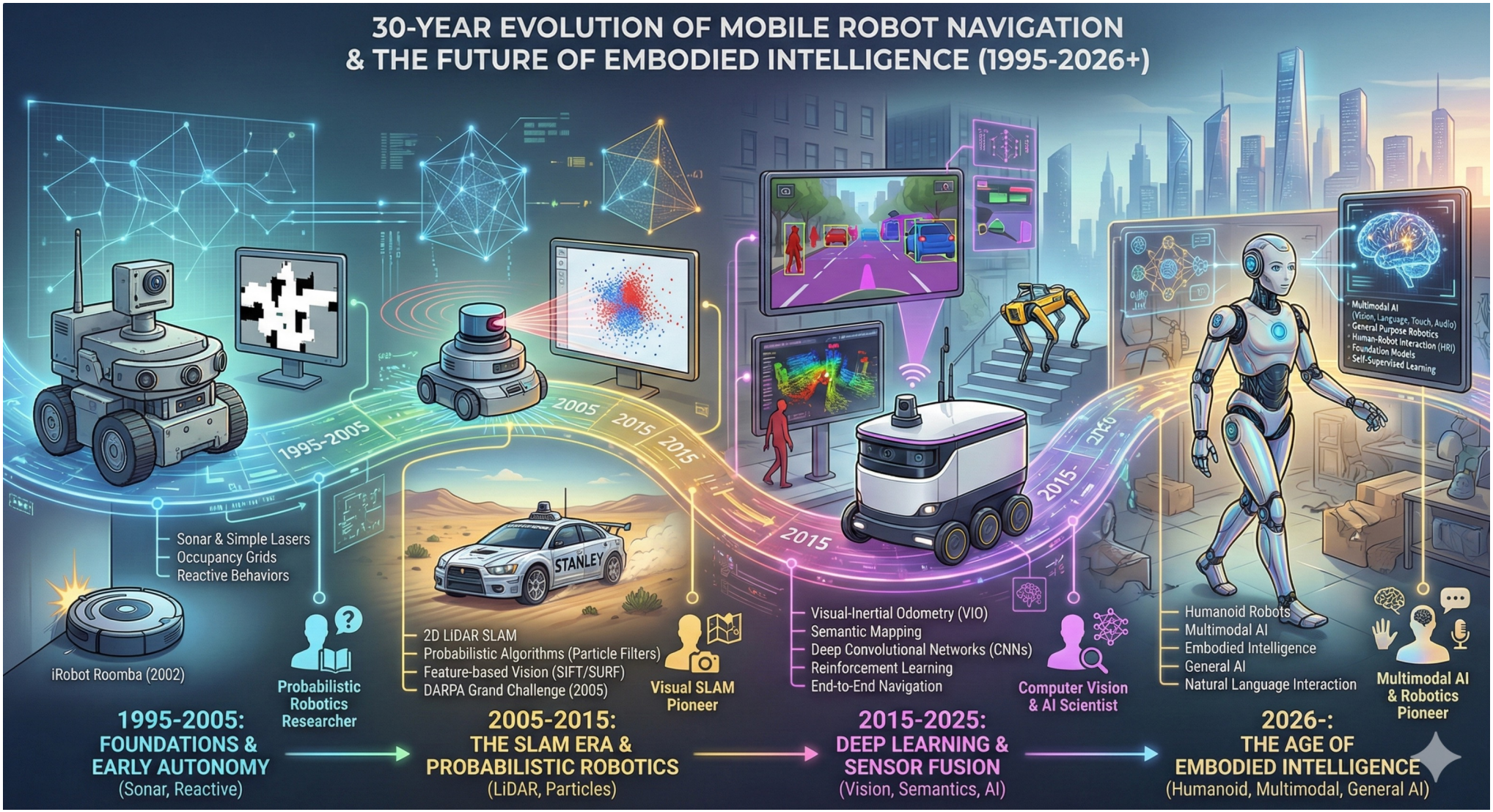
Senior Research Fellow

University College London-HK PolyU

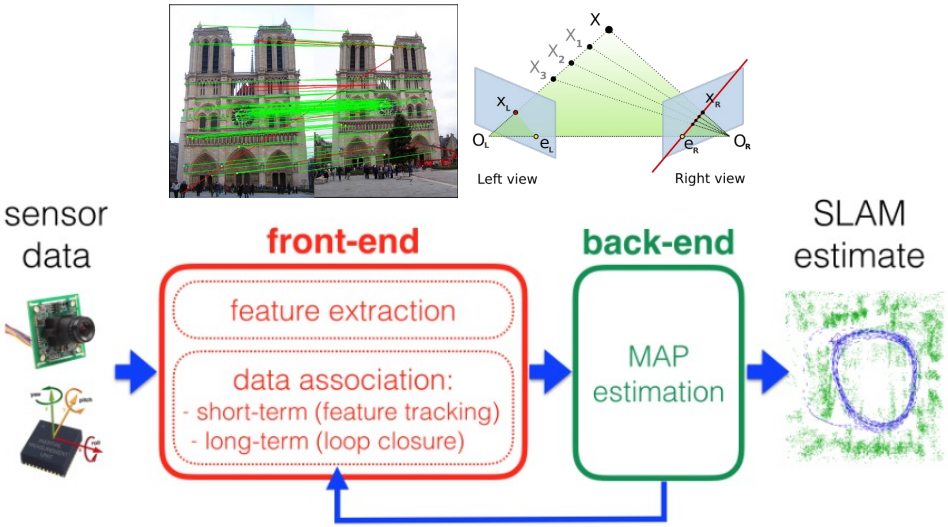
Email: jiaojh1994@gmail.com | Website: https://gogojjh.github.io

Invited by Prof.Junfeng Wu, CUHK-SZ

2026-01-30


30-YEAR EVOLUTION OF MOBILE ROBOT NAVIGATION
& THE FUTURE OF EMBODIED INTELLIGENCE (1995-2026+)
1995-2005
2005
2015
2015
2015
2015-
2026
Sonar & Simple Lasers
Occupancy Grids
Reactive Behaviors
iRobot Roomba (2002)
Probabilistic Robotics Researcher
STANLEY
2D LiDAR SLAM
Probabilistic Algorithms (Particle Filters)
Feature-based Vision (SIFT/SURF)
DARPA Grand Challenge (2005)
Visual SLAM Pioneer
Visual-Inertial Odometry (VIO)
Semantic Mapping
Deep Convolutional Networks (CNNs)
Reinforcement Learning
End-to-End Navigation
Computer Vision & AI Scientist
Multimodal AI
(Vision, Language, Touch, Audio)
General Purpose Robotics
Human-Robot Interaction (HRI)
Foundation Models
Self-Supervised Learning
Humanoid Robots
Multimodal AI
Embodied Intelligence
General AI
Natural Language Interaction
Multimodal AI & Robotics Pioneer
1995-2005:
FOUNDATIONS & EARLY AUTONOMY
(Sonar, Reactive)
2005-2015:
THE SLAM ERA & PROBABILISTIC ROBOTICS
(LiDAR, Particles)
2015-2025:
DEEP LEARNING & SENSOR FUSION
(Vision, Semantics, AI)
2026-:
THE AGE OF EMBODIED INTELLIGENCE
(Humanoid, Multimodal, General AI)

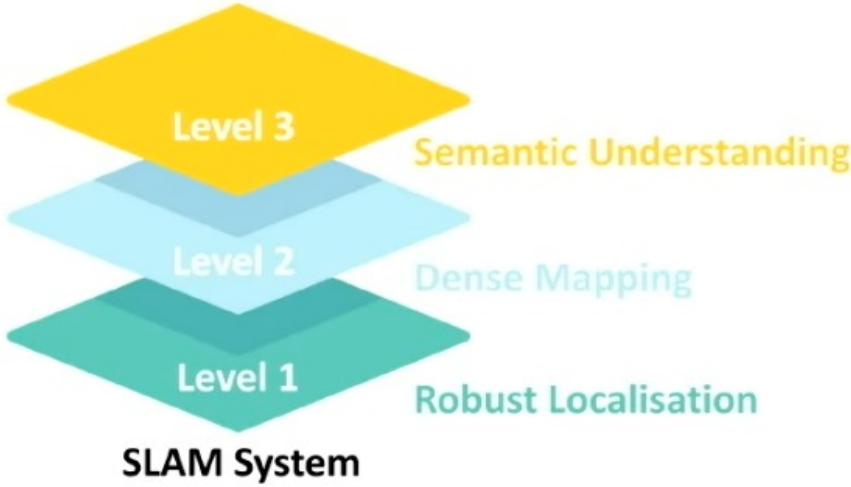
What is SLAM?



➢ SLAM comprises the simultaneous estimation of the **state** of a robot equipped with on-board sensors, and the construction of a model (the **map**) of the environment that these sensors are perceiving.

--- Past, present, and future of SLAM: Toward the robust-perception age, T-RO2017

What is SLAM?



➢ **Spatial AI** is the online problem where vision is to be used, usually alongside other sensors, as part of the AI which permits an embodied device to interact usefully with its environment.

➢ **SLAM**, a continuously evolving and broadening field with progress marked by real-time system-building, demos and open source, is the basis for working towards Spatial AI.

--- Prof.Andrew Davison, talk at MIT 2025

SLAM Examples

Gaussian Splatting SLAM [Matsuki2024CVPR]

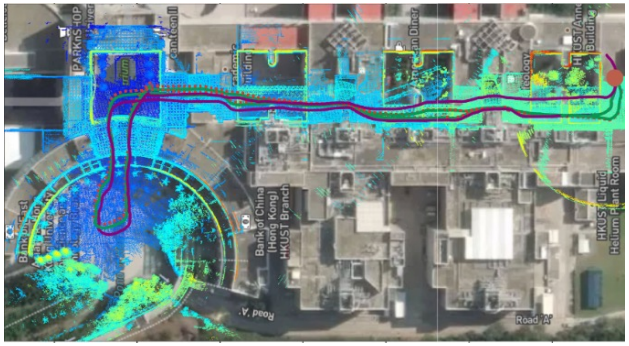
Scene Graph Reconstruction [Hughes2022RSS]

SLAM for Robotics: Mobile Autonomy Outside the Lab

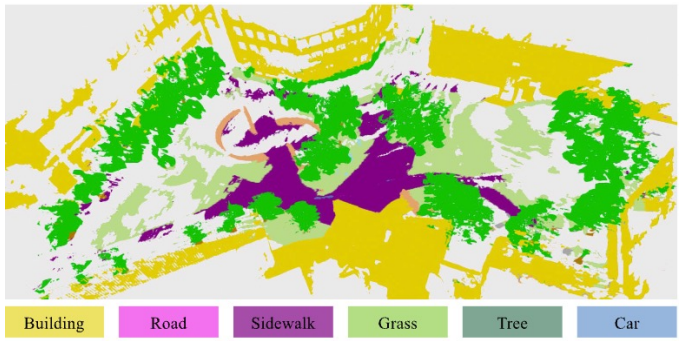
Vision-Language Navigation [Zantout2025IROS]

High-Speed Navigation [Ren2025SR]

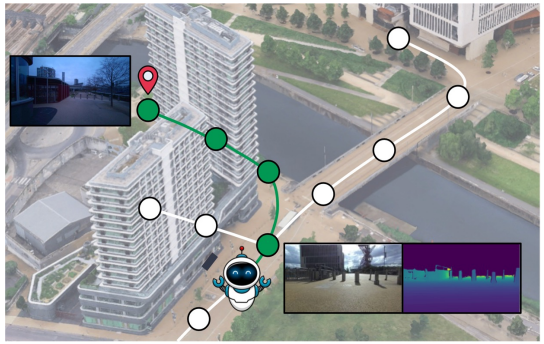
From Classic SLAM to Lifelong Navigation



LiDAR SLAM [Jiao2021TRO]

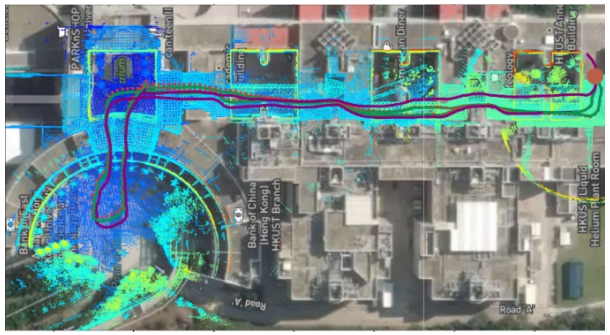


Metric-Semantic Mapping [Jiao2024TASE]

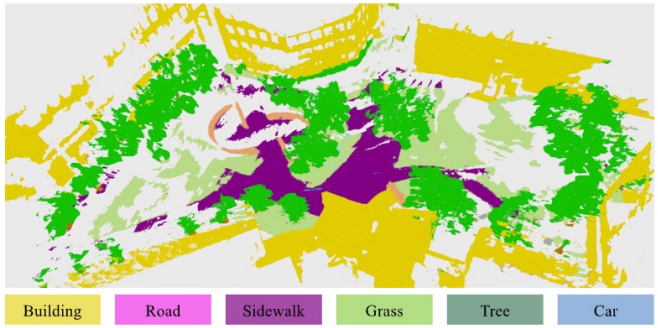


Lifelong Navigation [Under Review]

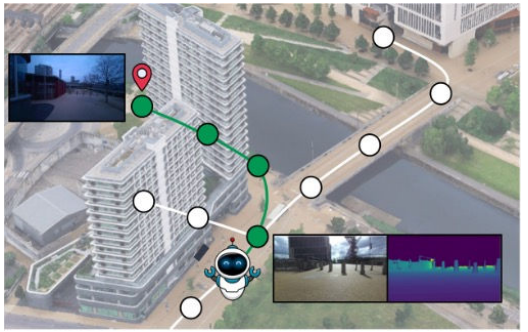
From Classic SLAM to Lifelong Navigation



LiDAR SLAM [Jiao2021TRO]



Metric-Semantic Mapping [Jiao2024TASE]

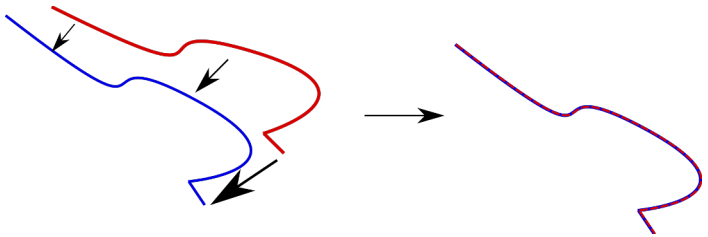


Lifelong Navigation [Under Review]

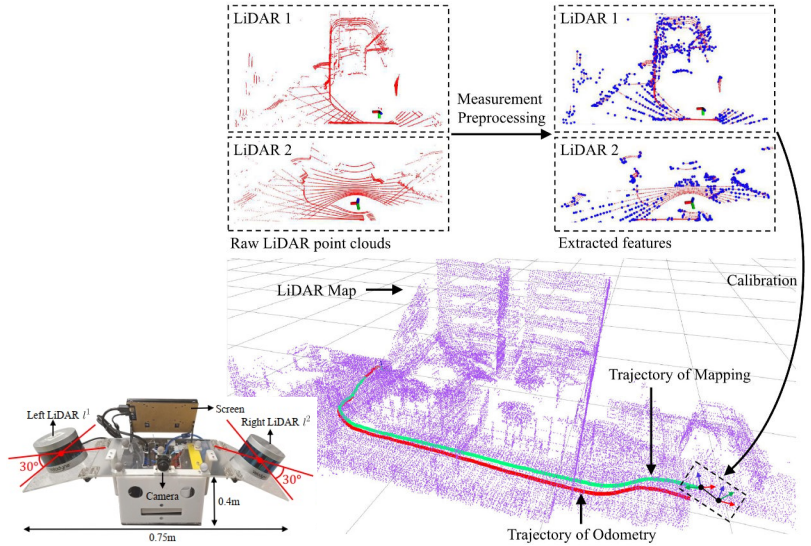
Some questions:

- Why does this transition happen?
- What are the new focus in lifelong navigation?
- What are the best candidates for scene representations?

Classic LiDAR SLAM



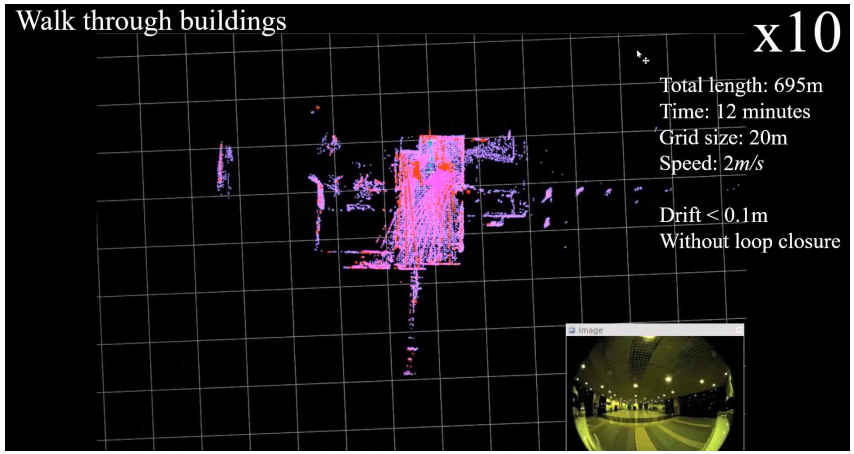
Iterative Closest Point [Arun1987PAMI]

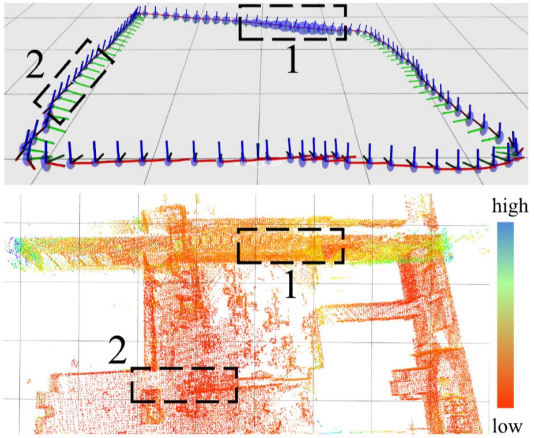


M-LOAM: Multi-LiDAR SLAM with uncertainty-ware mapping [Jiao2021TRO]

- Extrinsics as one of estiamted states
- Propagate errors (noise, extrinsic error, degeneracy) into map

Classic LiDAR SLAM (2018-2021)

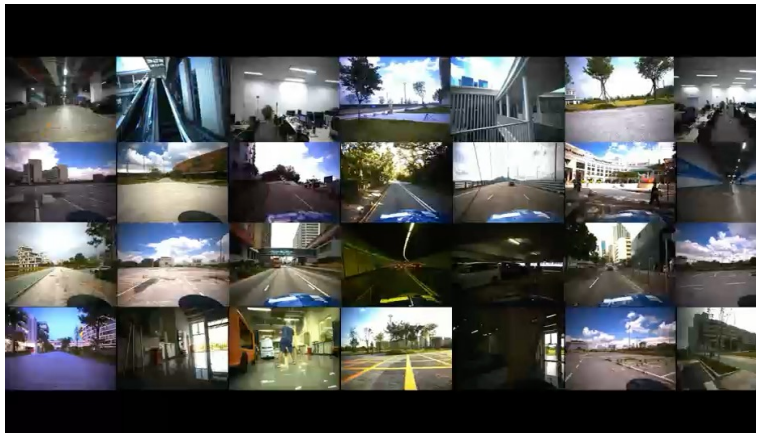




M-LOAM: Multi-LiDAR SLAM with uncertainty-ware mapping [Jiao2021TRO]

- Extrinsics as one of estiamted states
- Propagate errors (noise, extrinsic error, degeneracy) into map

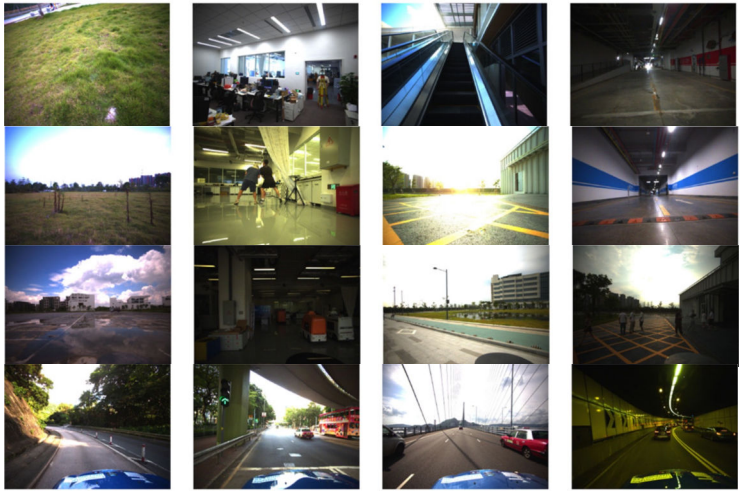
Challenging SLAM Datasets (2021-2023)

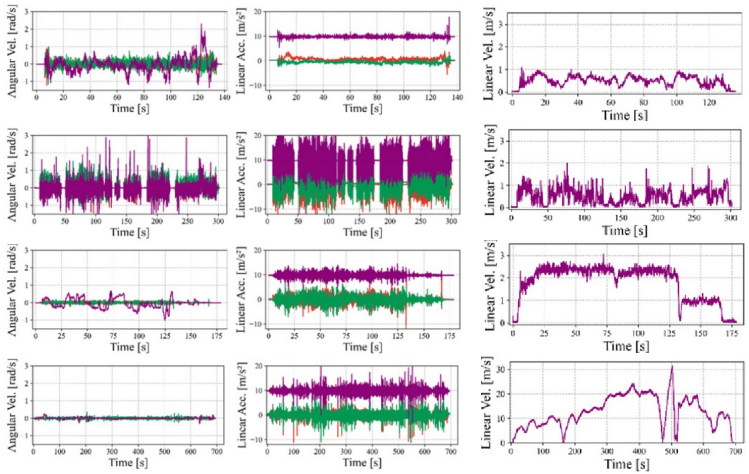


FusionPortable Dataset for SLAM Evaluation [Jiao2022IROS] [Wei2025IJRR]

- **Open questiones:** is it possible to use one SLAM for all platforms and all enviroments?
- Data: **42km trajectories** on **4 platforms**, covering **10 challenging scenarios**
- Motion patterns for different types of robots vary significantly

Challenging SLAM Datasets (2021-2023)





fusionportable.github.io/dataset/fusionportable_v2/

FusionPortable Research Dashboard | Home | Calibration | Localization & Mapping | Perception | Datasets | Search | Previous | Next

- News
- Overview
- Usage Steps
- Sensors
  - Definitions of Coordinate Frame
  - Explanation of ROS Topic and Message
- Various Platforms and Scenarios
- Ground-Truth Devices
- Third-View of Data Collection
- Dataset Details and Download
- Sensor Calibration Tutorial
- Sensor Calibration Files
- Experiments
- Tools
- Known Issues
- Related Works
- Publications
- Contact

FusionPortable V2

From Campus to Highway: A Unified Multi-Sensor Dataset for Generalized SLAM Across Diverse Platforms and Scalable Environments

News

- (2025-09-28) The list of Related Works that have utilized the FusionPortable dataset has been updated.
- (20250410) Some rosbags are extracted as individual files and converted into the KITTI format. Click here to try.
- (20240629) The tutoial of senosr calibration (intrinsics and extrinsics) is provided. Click here to try.
- (20240508) Groundtruth poses of all vehicle-related sequences are postprocessed: eliminate poses characterized by high uncertainty.
- (20240422) Data can be downloaded from Baidu Wang Pan with the code **byj8**.
- (20240414) All sequences, ground-truth trajectories, and ground-truth maps have been publicly released. If you find issues of GT trajectories and maps, please contact us or report here.
- (20240413) A small simulated navigation environment is provied.
- (20240408) The development tool has been initially released.
- (20240407) Data can be downloaded from Google Drive.

Overview

Usage Steps

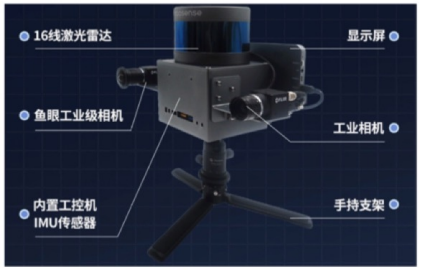
1. Read through the overview of the FusionPortableV2 dataset: *sensors, coordinate frames, and definitions of ROS topics and message*.
2. Download data from this link.
3. Check examples of using the dataset from this link.

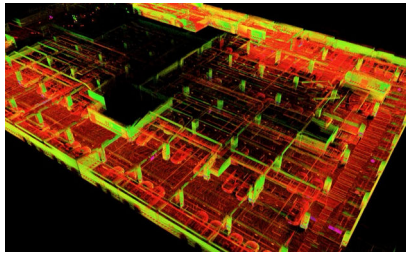
Sensors

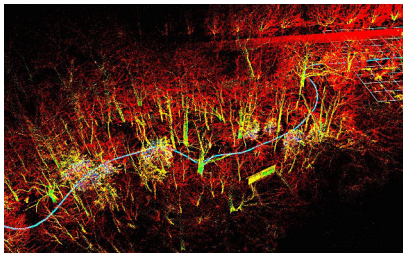
- **Handheld Sensor**:
  - 128-beam Ouster LiDAR (OS1, 120m range);
  - Stereo FLIR BFS-U3-31S4C cameras;
  - Stereo DAVIS346 cameras;
  - STIM300 IMU;
  - 3DM-GQ7-GNSS/INS
- **UGV Sensor**: Omron E6B2-CWZ6C wheel encoder
- **Legged Robot Sensor**: Built-in joint encoders, contact sensors, and IMU of the Unitree A1

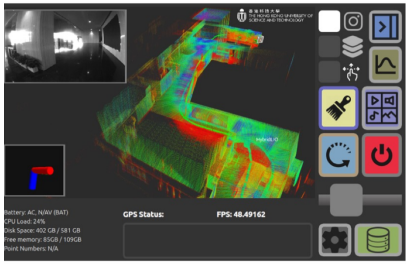
Sensor	Characteristics	ROS Topic	ROS Message Type	Rate (Hz)
3D LiDAR	Ouster OS1-128, 45° vert. × 360° horiz. FOV	/os_cloud_node/points	sensor_msgs/PointCloud2	10
	IMU: ICM20948, 9-axis MEMS	/os_cloud_node/imu	sensor_msgs/Imu	100
	Range, near-ir, reflectivity, signal images	/os_image_node/{range,nearir,...}_image	sensor_msgs/Image	10
Frame Camera	Stereo FILR BFS-U3-31S4C, global shutter			
	66.5° vert. × 82.9° horiz. FOV	/stereo/frame_{left,right}/image_raw	sensor_msgs/CompressedImage	20
	1024 × 768 resolution			
Event Camera	Stereo DAVIS346, 67° vert. × 83° horiz. FOV	/stereo/davis_{left,right}/events	dvs_msgs/EventArray	30
	346 × 240 resolution			
	Images that capture color data	/stereo/davis_{left,right}/image_raw	sensor_msgs/CompressedImage	20
	IMU: MPU6150, 6-axis MEMS	/stereo/davis_{left,right}/imu	sensor_msgs/Imu	1000
IMU	STIM300, 6-axis MEMS	/stim300/imu	sensor_msgs/Imu	200
INS	3DM-GQ7-GNSS/INS	/3dm_ins/nav/odom	nav_msgs/Odometry	10
	Dual-antenna, RTK-enabled INS	/3dm_ins/gnss_{left,right}/fix	sensor_msgs/NavSatFix	10
		/3dm_ins/imu	sensor_msgs/Imu	200
Wheel Encoder	Omron E6B2-CWZ6C, 1000P/R	/mini_hercules/encoder	sensor_msgs/Jointstate	100
Legged Sensor	Built-in joint encoders and contact sensors	/unitree/joint_state	sensor_msgs/JointState	50
	Built-in IMU	/unitree/imu	sensor_msgs/Imu	50
	Out-of-the-box kinematic-ineratial odometry	/unitree/body_odom	nav_msgs/Odometry	50

Impact of the FusionPortable Dataset

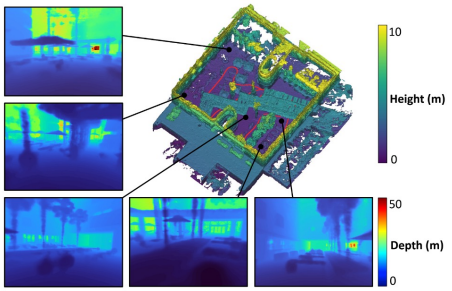




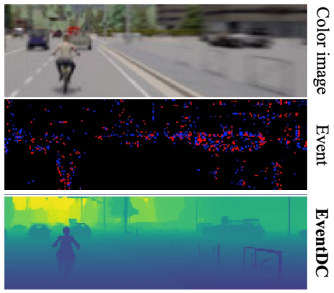




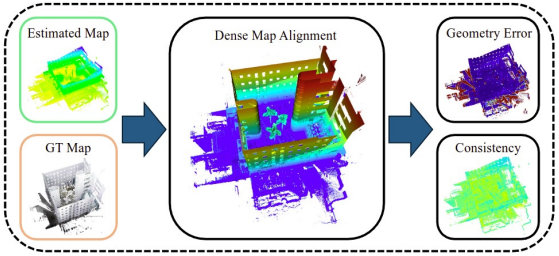
UCL 香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY 香港科技大学(广州) THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU) KAISA 佳兆业



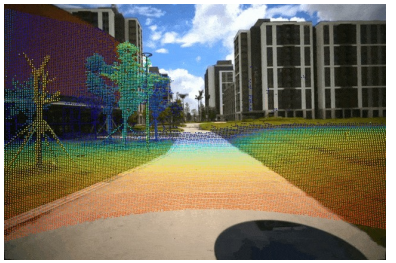
LONER [RAL2023]



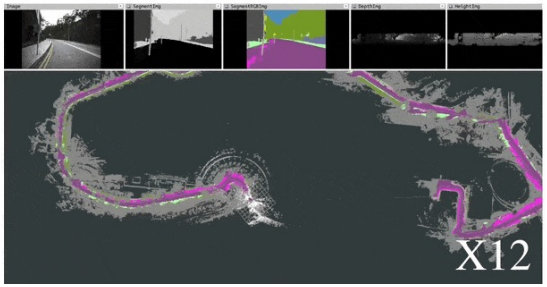
Event-based Depth Completion [NeurIPS2025]



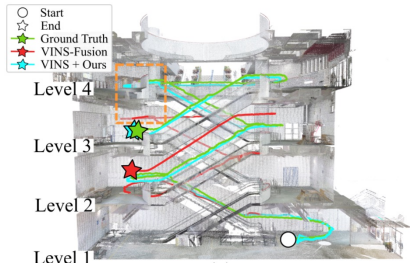
Mapping Evaluation [RAL2025]



LCE-Calib [T-MECH2022]



Metric-Semantic Mapping [TASE2024]



Verification of Loop Closure [Under Review]

SLAM-Enabled Products and Systems

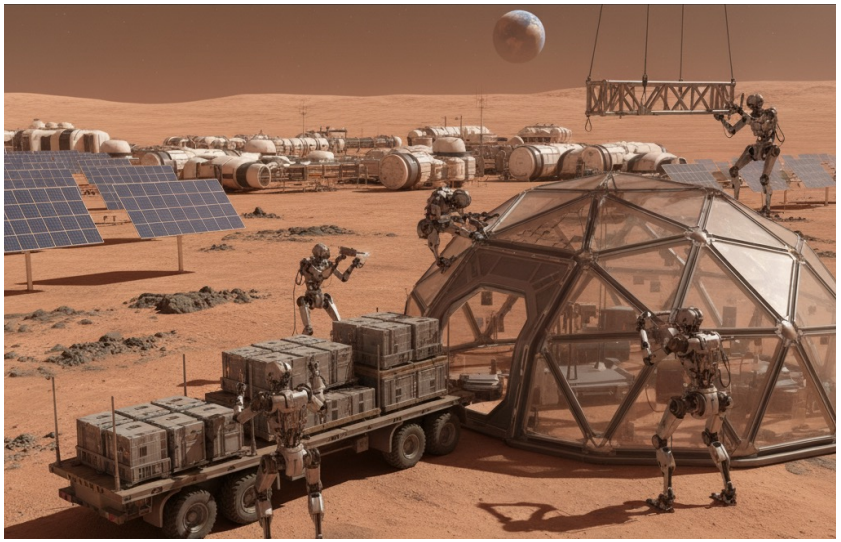
Drones

AR Glasses

Sweeping Robots

- **Mature Geometric Perception:** Localization and sparse/semi-dense serve as foundational technologies for many real-world products.
- **But Autonomy Gap:** systems remain fragile and lack context as well as interaction, requiring human intervention for long-term operation or recovery.

Next-Gen SLAM: Spatial Memory for Embodied Navigation

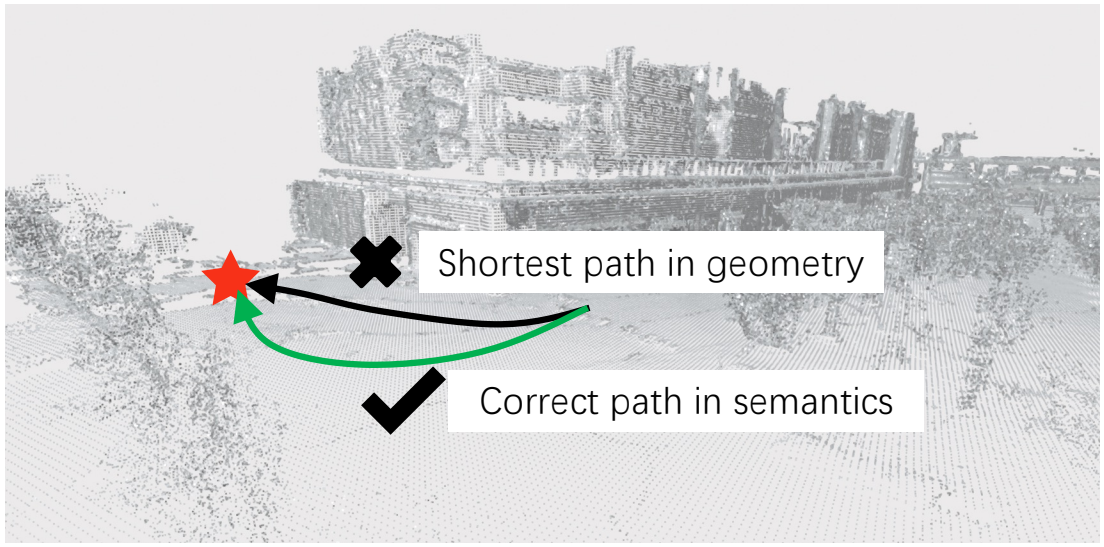


Semantic Understanding

Task Versatility and Learning

Lifelong Autonomy

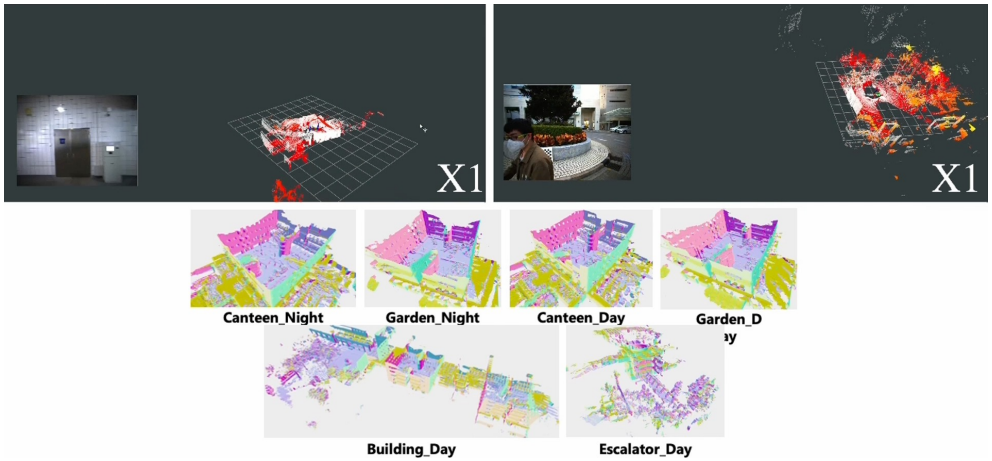
Metric-Semantic Mapping



Real-time metric-semantic mapping [Jiao2024TASE]

- GPU-accelerated volumetric mapping
- Bayesian update for noisy semantic labels
- Enhance the safety of robot navigation in unstructured environments

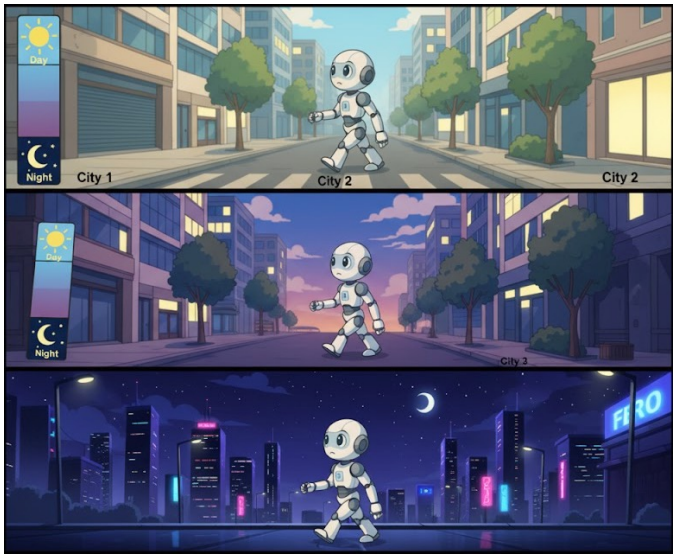
Metric-Semantic Mapping

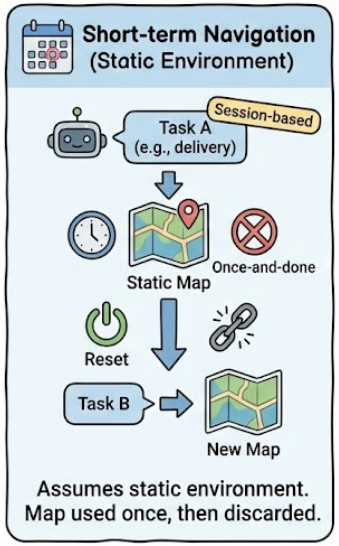


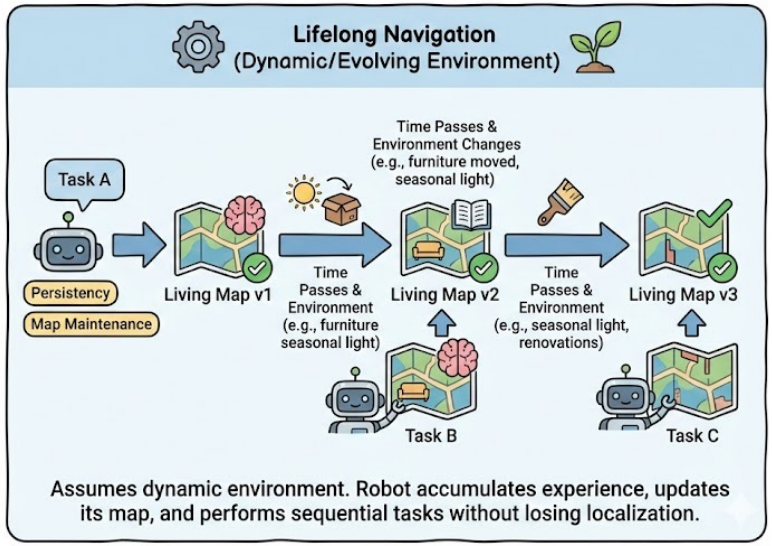
Real-time metric-semantic mapping [Jiao2024TASE]

- GPU-accelerated volumetric mapping
- Bayesian update for noisy semantic labels
- Enhance the safety of robot navigation in unstructured environments

New Focus: Lifelong Navigation







What lifelong navigation focuses:

- **Working** in highly dynamic real-world scenarios
- **Problems**: long-term localization, semantic understanding, dynamic mapping, etc.
- **Beyond accuracy evaluation**: success rate, memory growth rate, etc.

Coincidence vs. Inevitability: Map as a type of Memory

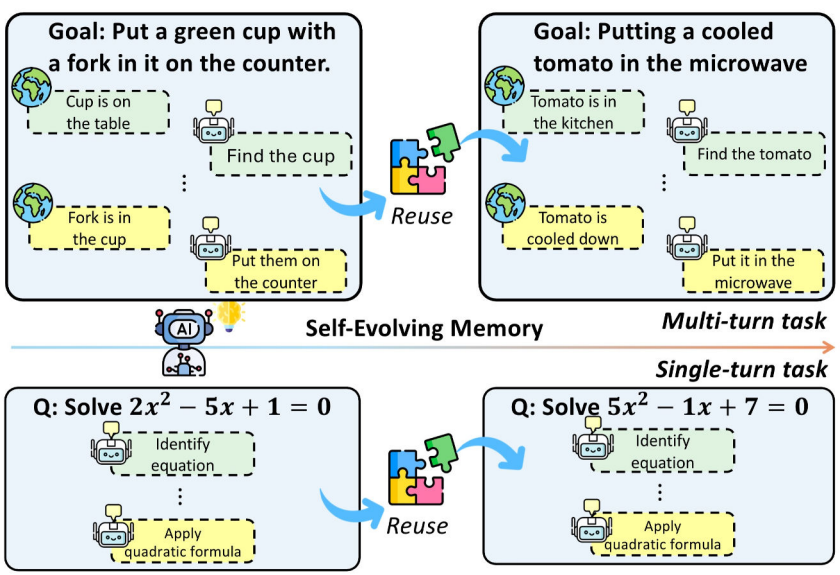
Google DeepMind

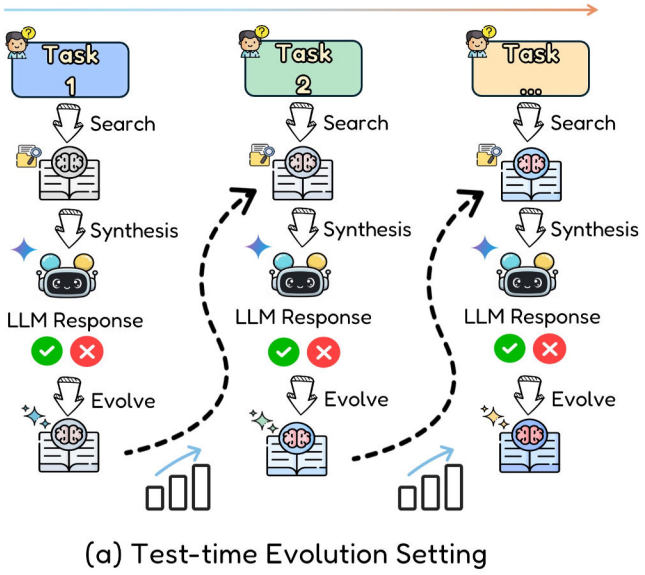
2025-11-27

Evo-Memory: Benchmarking LLM Agent Test-time Learning with Self-Evolving Memory

Tianxin Wei[†,1], **Noveen Sachdeva**[2], **Benjamin Coleman**[2], **Zhankui He**[2], **Yuanchen Bei**[1], **Xuying Ning**[1], **Mengting Ai**[1], **Yunzhe Li**[†,1], **Jingrui He**[1], **Ed H. Chi**[2], **Chi Wang**[2], **Shuo Chen**[2], **Fernando Pereira**[2], **Wang-Cheng Kang**[2] and **Derek Zhiyuan Cheng**[2]

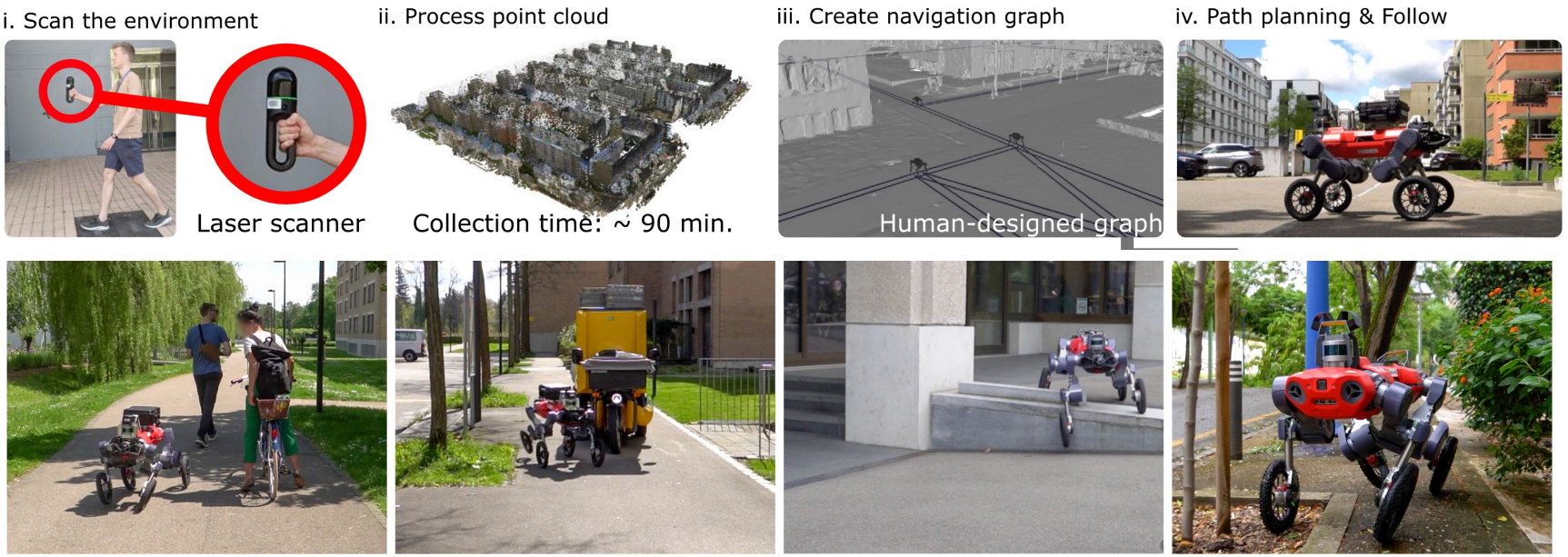
[†]Work done while at Google DeepMind, [1]University of Illinois Urbana-Champaign, [2]Google DeepMind





(a) Test-time Evolution Setting

Existing Solutions: Dense Mapping and Graph Extraction



[Lee2024Science Robotics]

- ✓ Accurate results
- ✓ Well-defined pipeline
- ✗ Carefully data capture
- ✗ High storage overload
- ✗ Time-consuming reconstruction

Existing Solutions: Teach and Repeat

[Qiao2025ICRA]

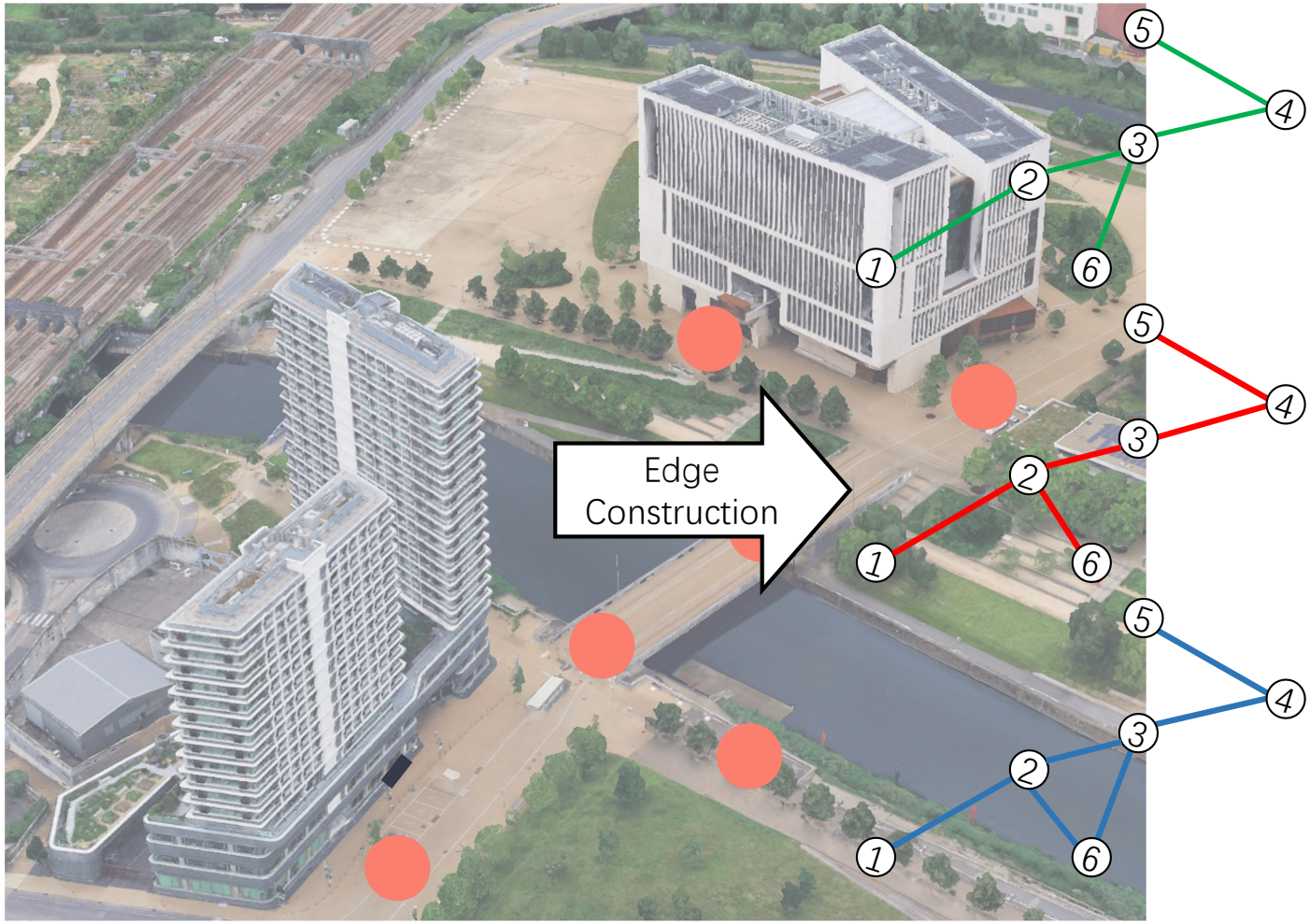
- **Teach:** user drives a robot which stores sparse visual information in a relative pose graph
- **Repeat:** robot localizes by matching live visual data to map and steers to stay on path

✓ Well-defined pipeline
✓ Relaxing global consistency

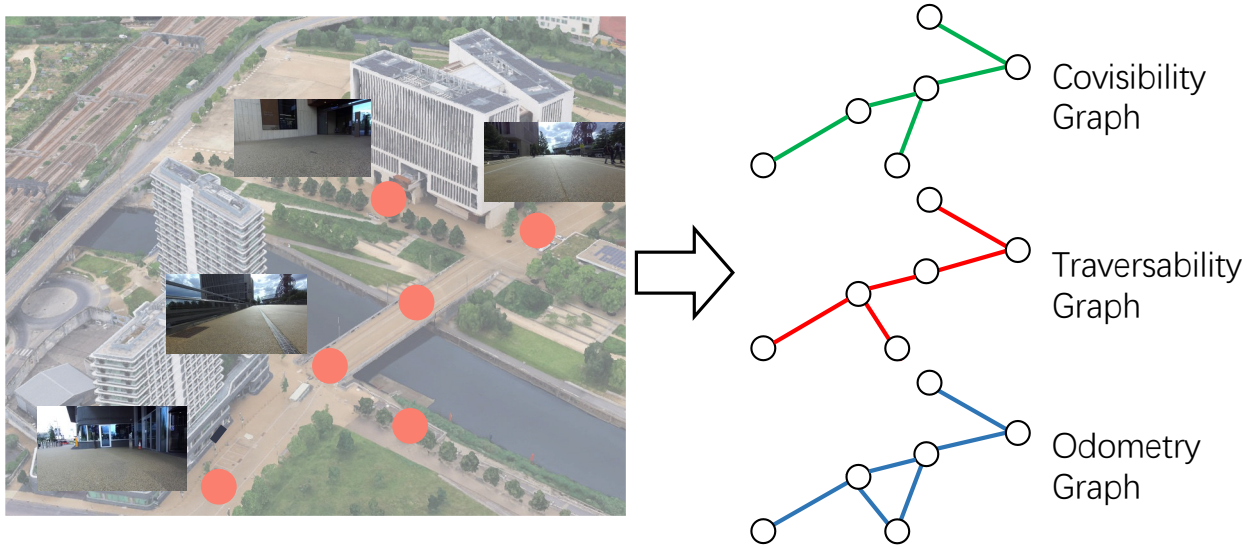
✗ Limited to the teach path
✗ Time-consuming data collection

Our Solution: Sparse Map Representation

Our Solution: Sparse Map Representation

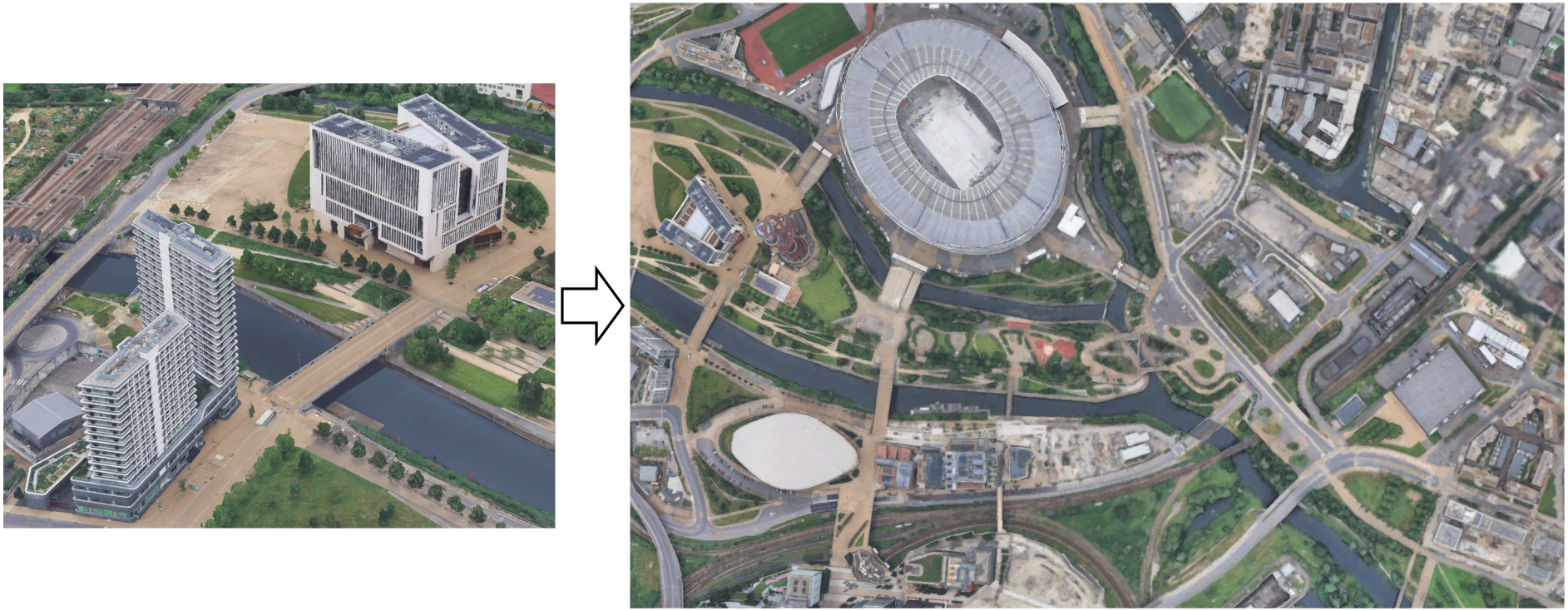


Our Solution: Sparse Map Representation



- **Map** is represented as a three-layer topometric map.
- **Submap** is constructed by individual mobile devices integrated with odometry.
- **Collaborative localization** with 3D geometric foundation model

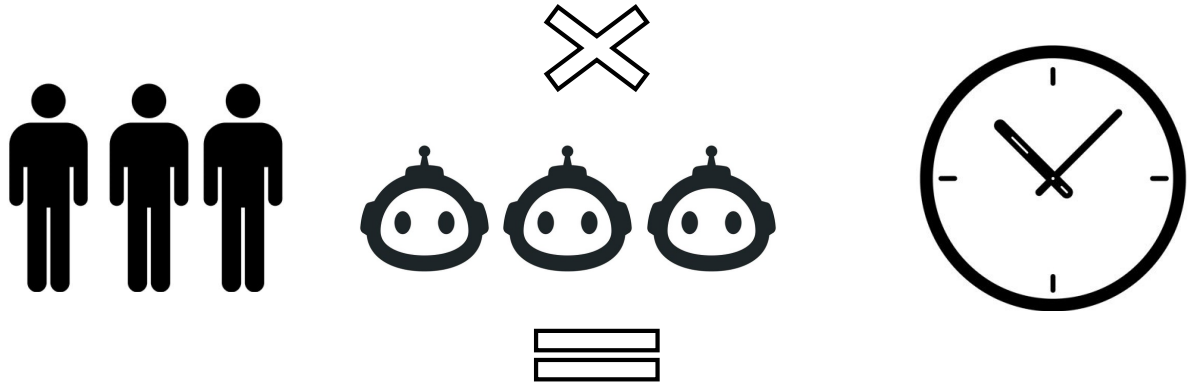
How to Extend the Mapping Scale?



Our Solution: Crowdsourcing Mapping

Relax

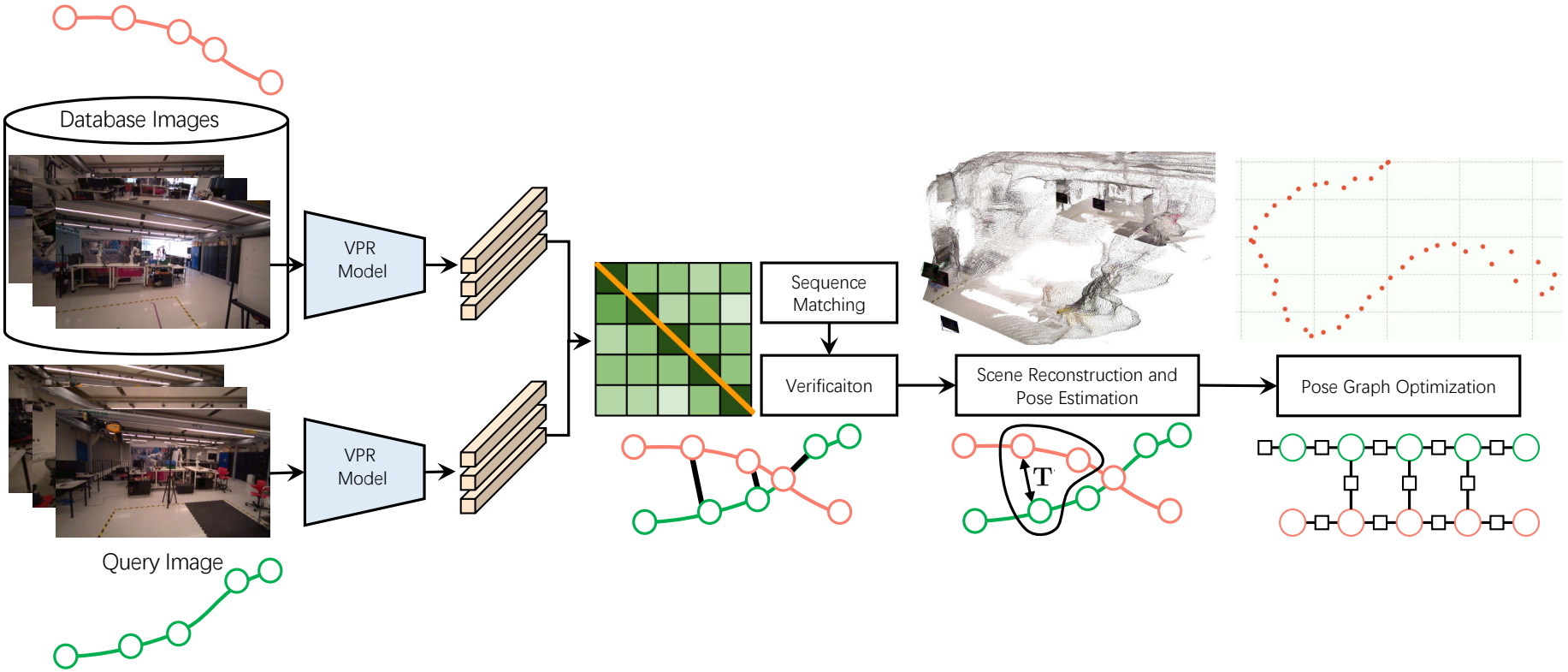
Our Solution: Crowdsourcing Mapping



Large amount of **Fresh** mapping data from **Public Users**

Our Solution: Crowdsourcing Mapping

OpenNavMap System: Open Navigation Map



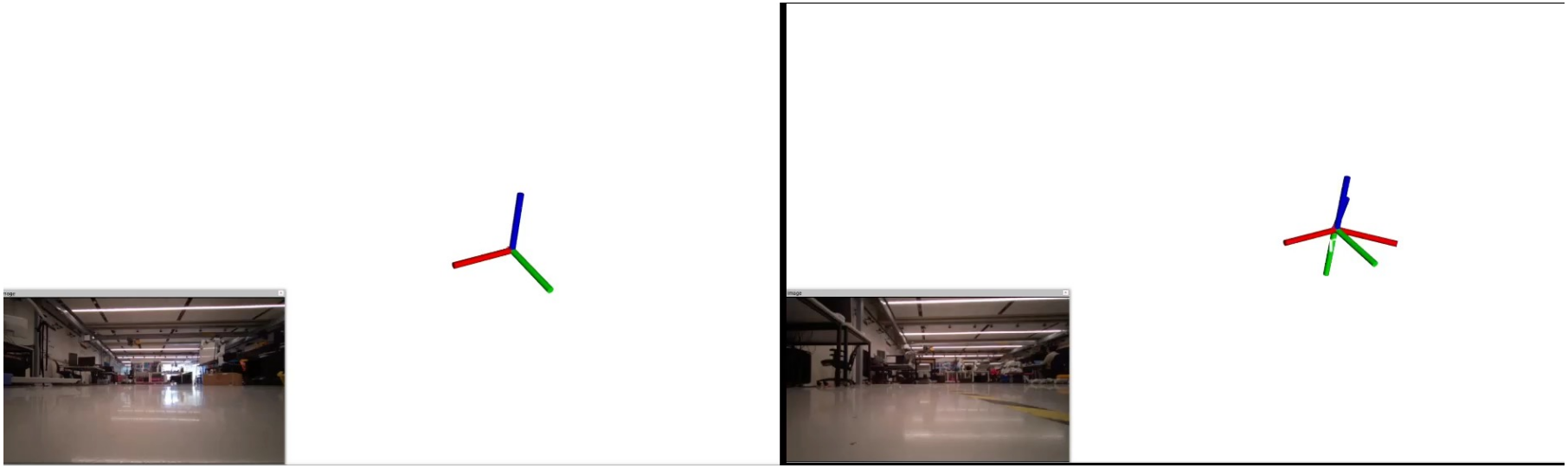
Collaborative localization Pipeline by Gathering Multi-Source Submaps

IROS2025 OWN Workshop Best Paper

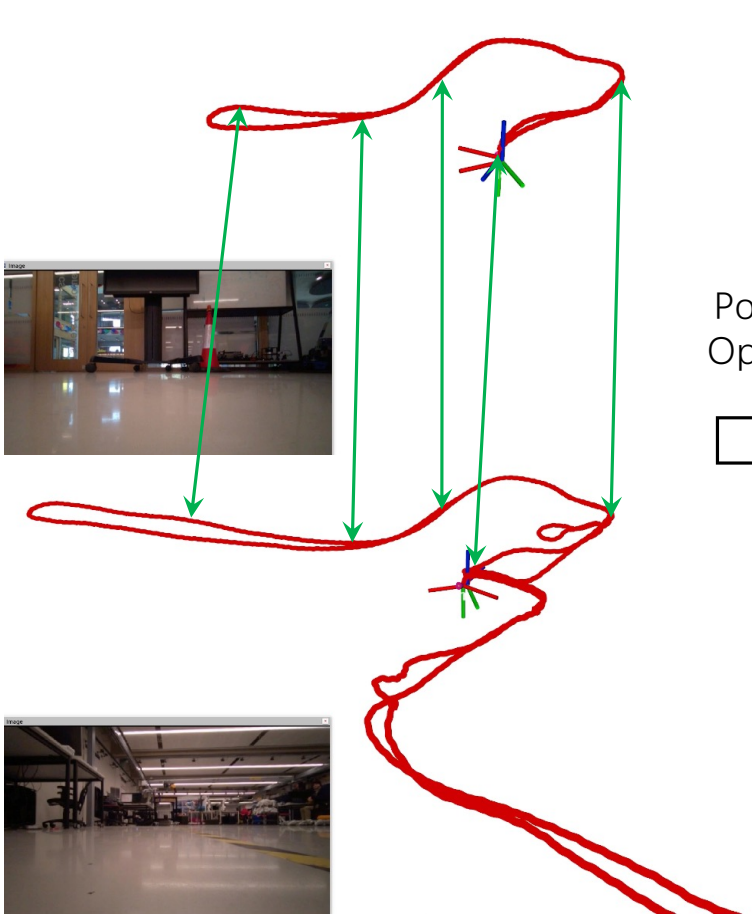
Experiments

Dataset evaluation: 9 months, 37 sequences, 19km trajectories

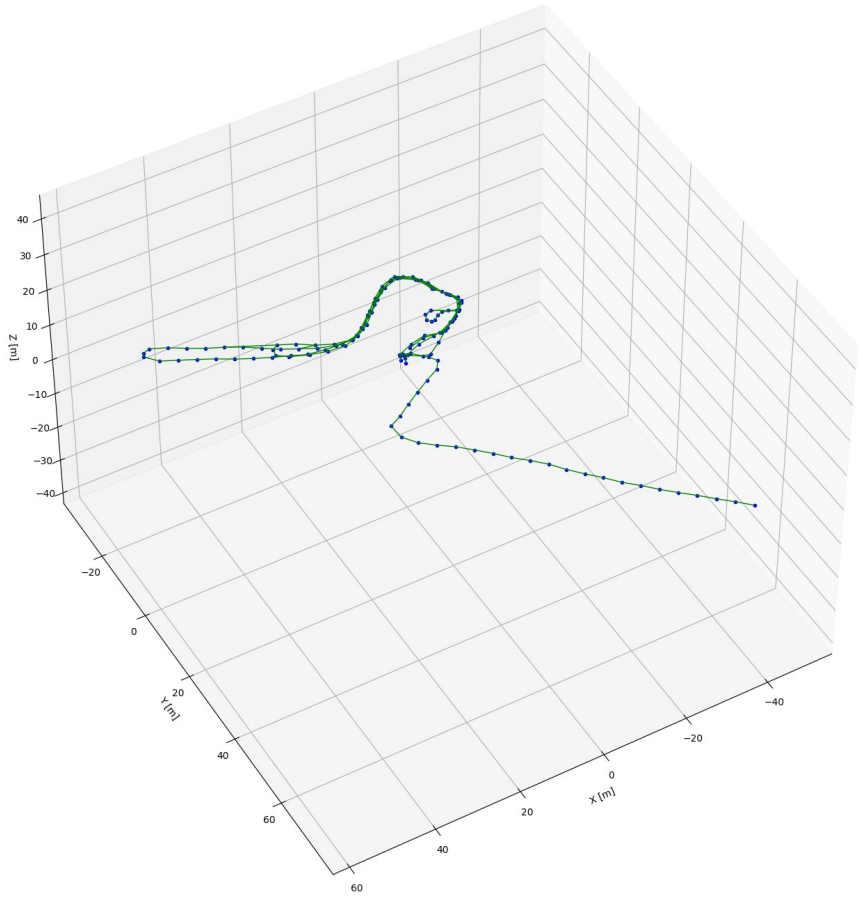
Submap Construction



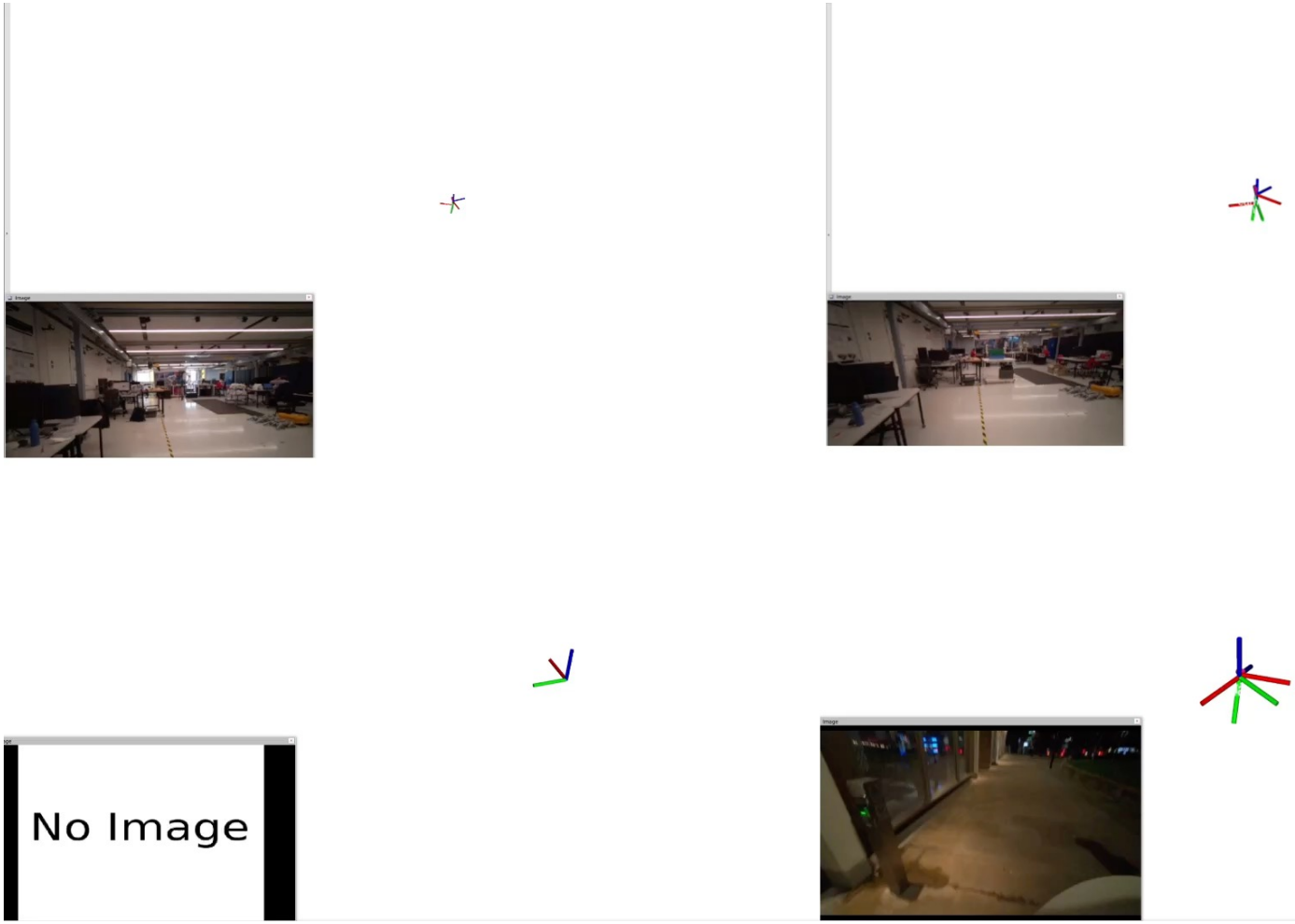
Collaborative Localization and Map Merging



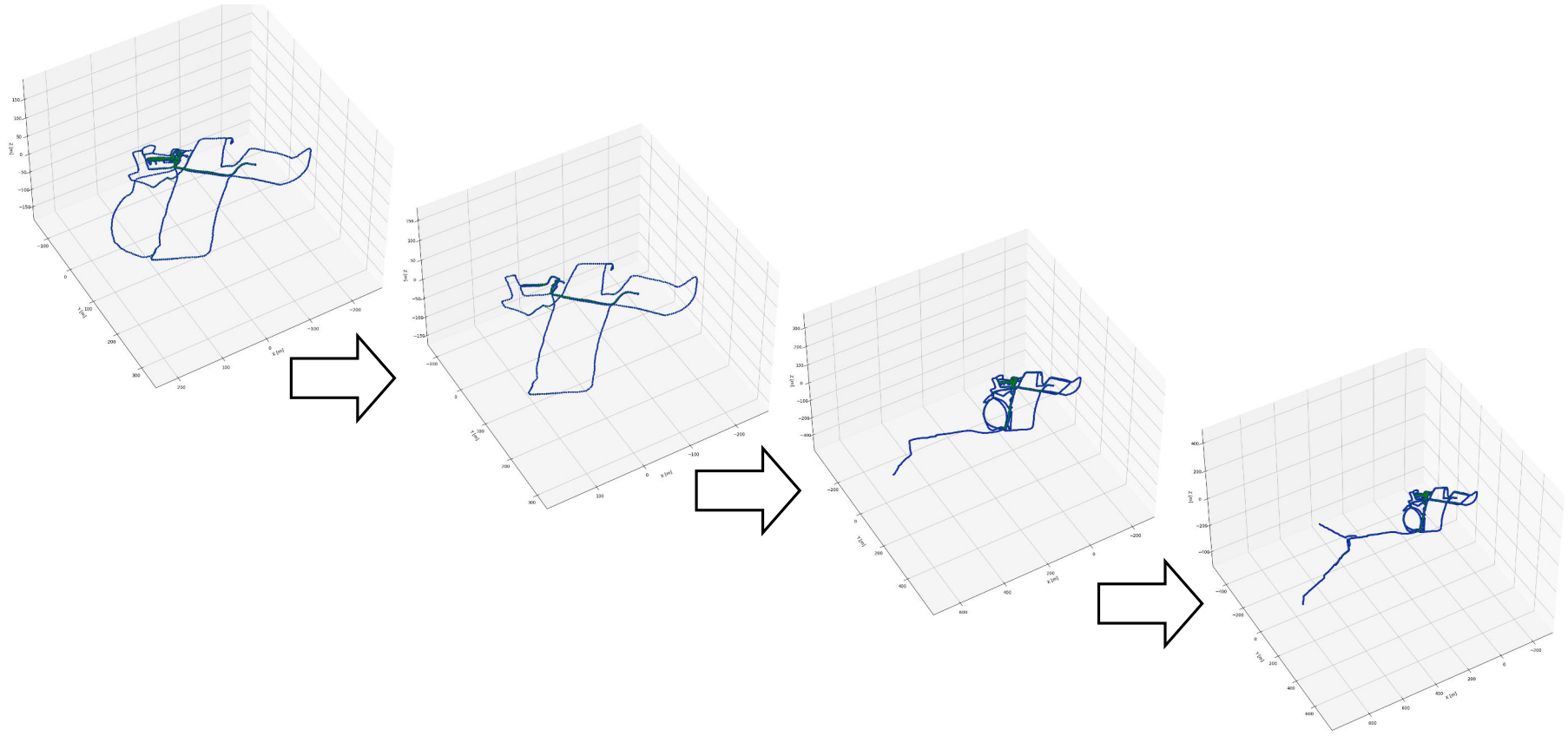
Pose Graph
Optimization



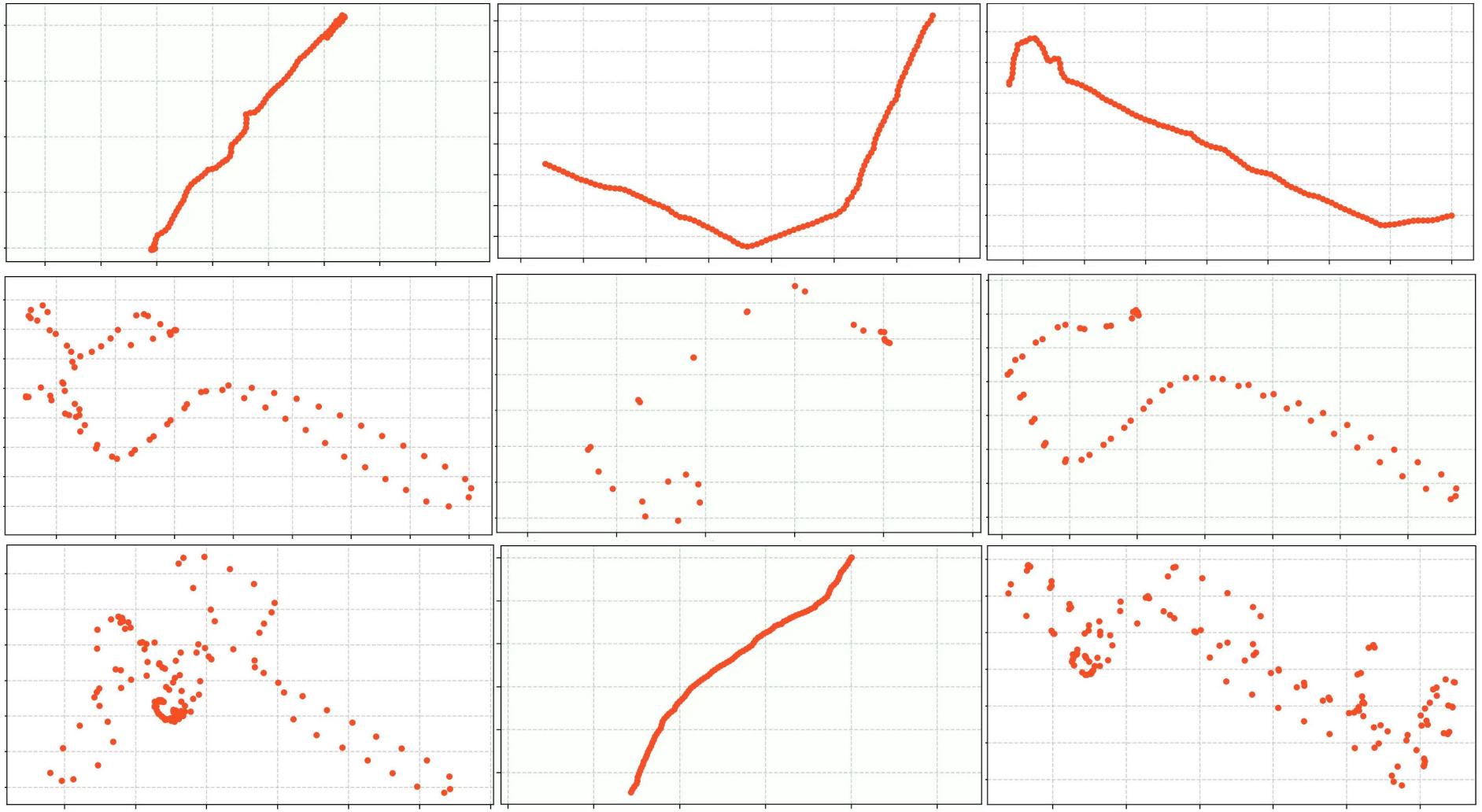
Map Merging with More Submaps



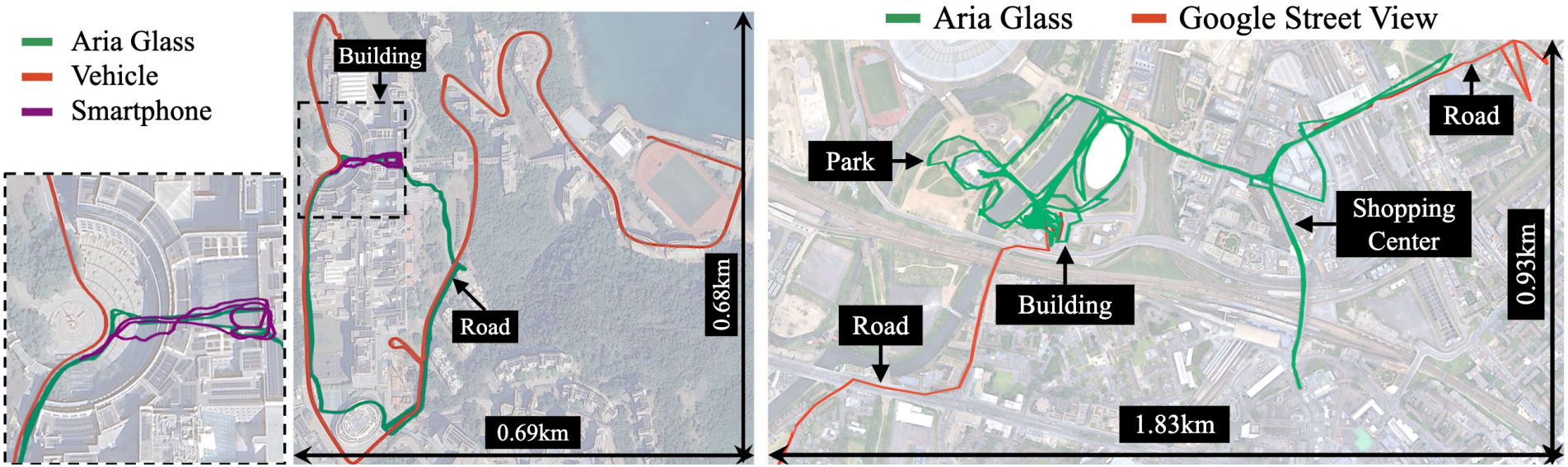
Map Merging with More Submaps



Map Merging with Shuffle Submap Orders



Final Map Construction



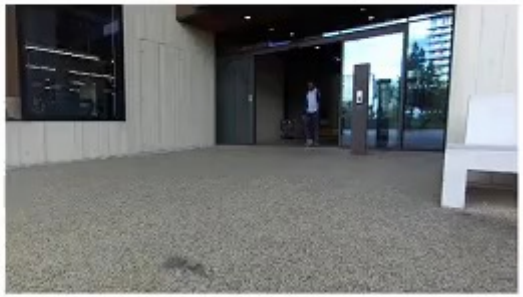
LiteVloc: Enable Visual Navigation with the OpenNavMap

Map-Lite Visual Localization for Image Goal Navigation [Jiao2025ICRA]

- **Map**: sparse and discrete map eliminats the dense, metrically-precise map construction.
- **Hierarchical Vloc:** first retrieve place, then estimate the relative pose.

LiteVloc: Enable Visual Navigation with the OpenNavMap

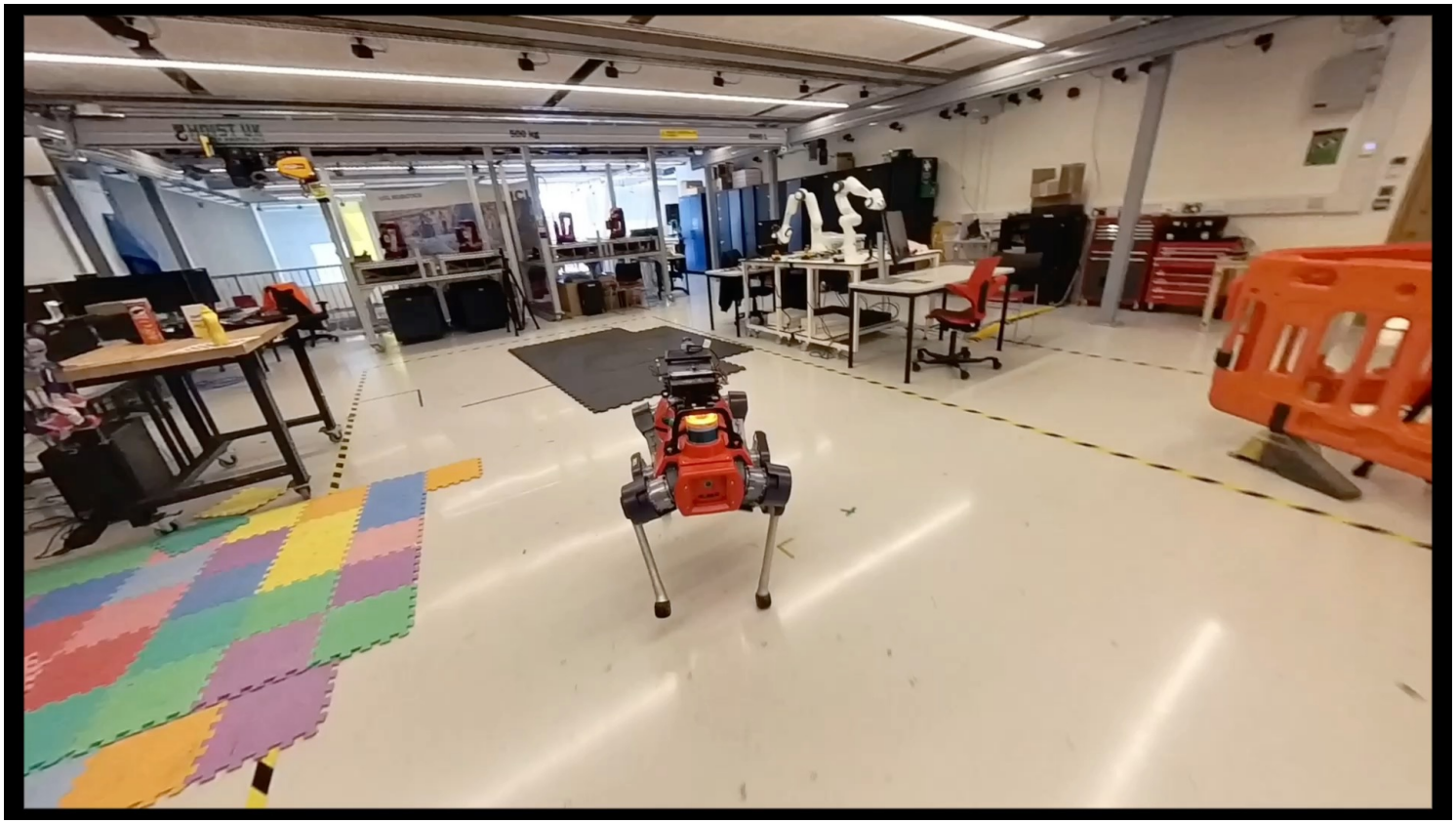
Introduction



Goal image

Imagine the robot is navigating its usual route

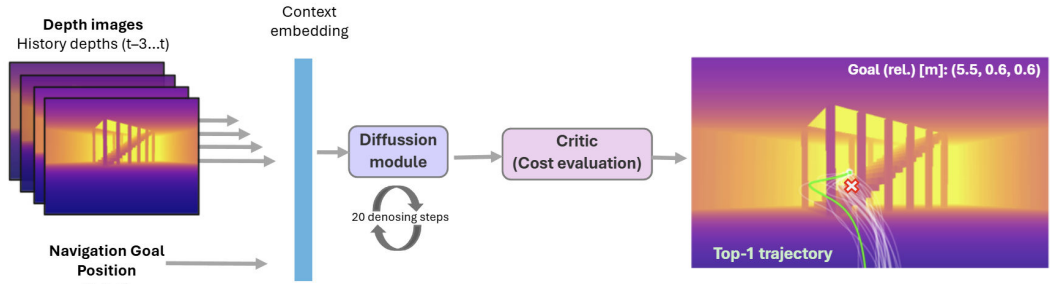
Demonstration of Image-Goal Navigation

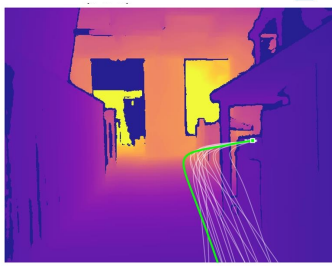


Zero-shot deployment

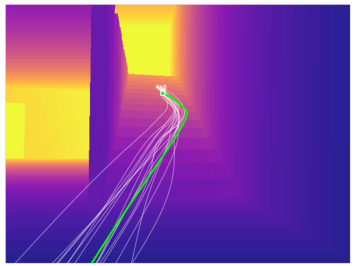
SanD-Planner: Visual Path Planning

- **Map-less** and **diffusion-based** local path planning for multi-layer scenarios
- The structural inductive bias provided by **B-splines** drastically improves sample efficiency.

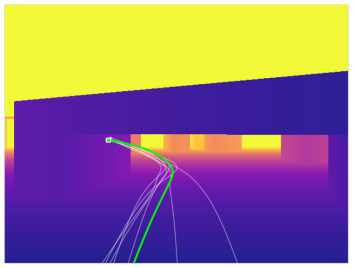




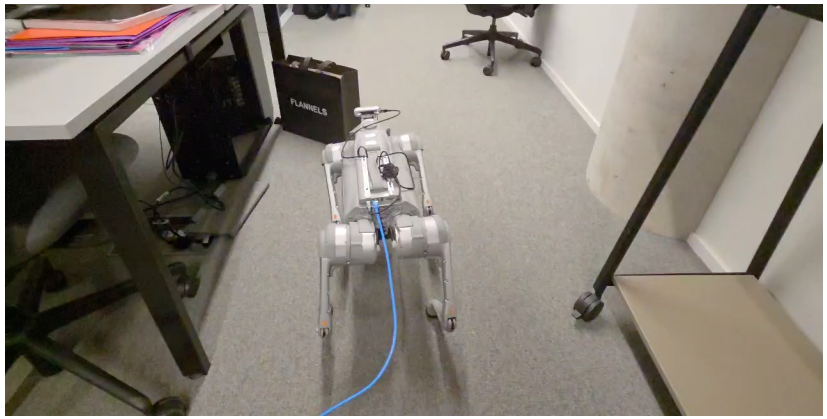
Collision avoidance

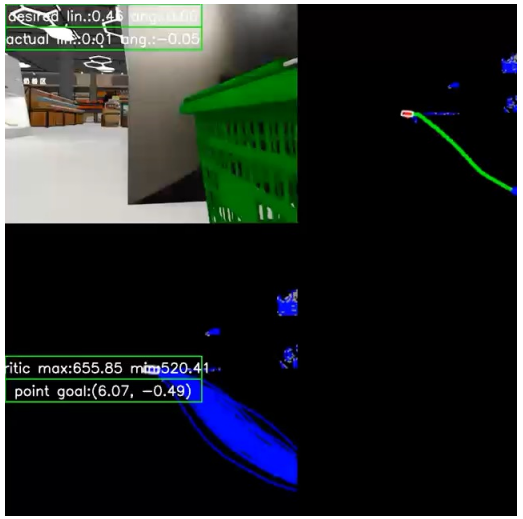


Climbing stairs



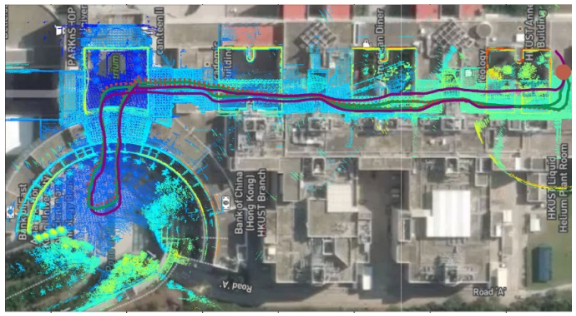
Passing obstacles



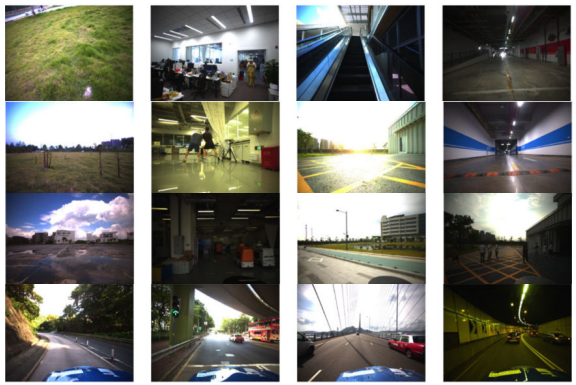


Next-Gen SLAM: Spatial Memory for Embodied AI

Geometric Representation

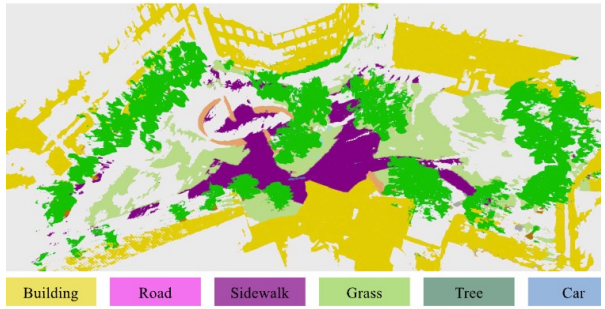


M-LOAM: multi-LiDAR SLAM



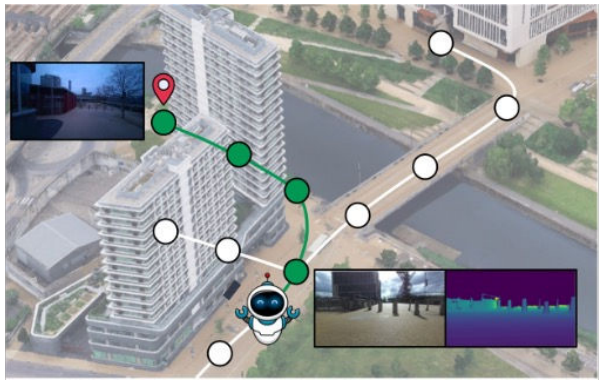
FusionPortable Dataset

Semantic



Metric-Semantic Mapping

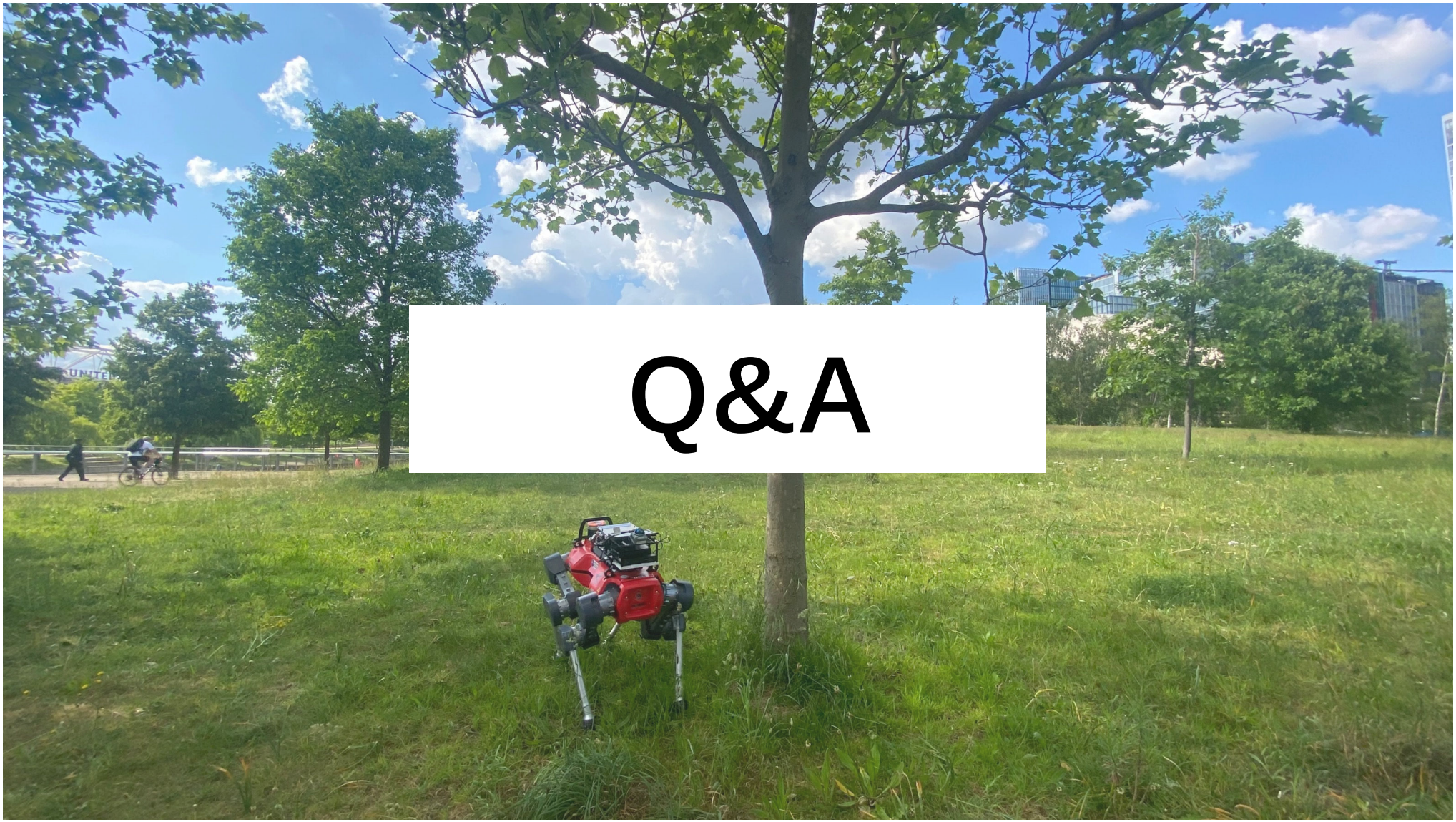
Lifelong



OpenNavMap and LiteVLoc

Lifelong Field Autonomy

Goal


Q&A