



From Autonomous Exploration to Embodied Q&A

Grzegorz Malczyk, Kostas Alexis
Autonomous Robots Lab, NTNU

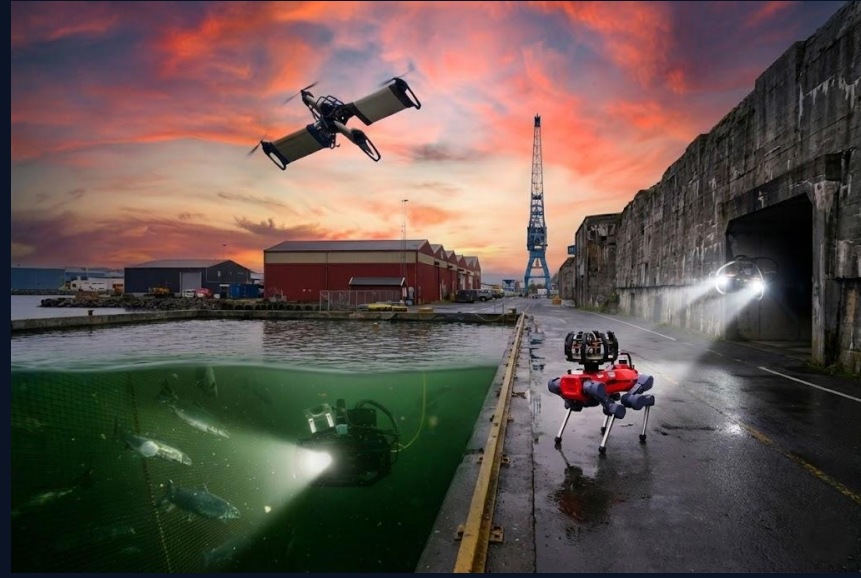
Talk Themes

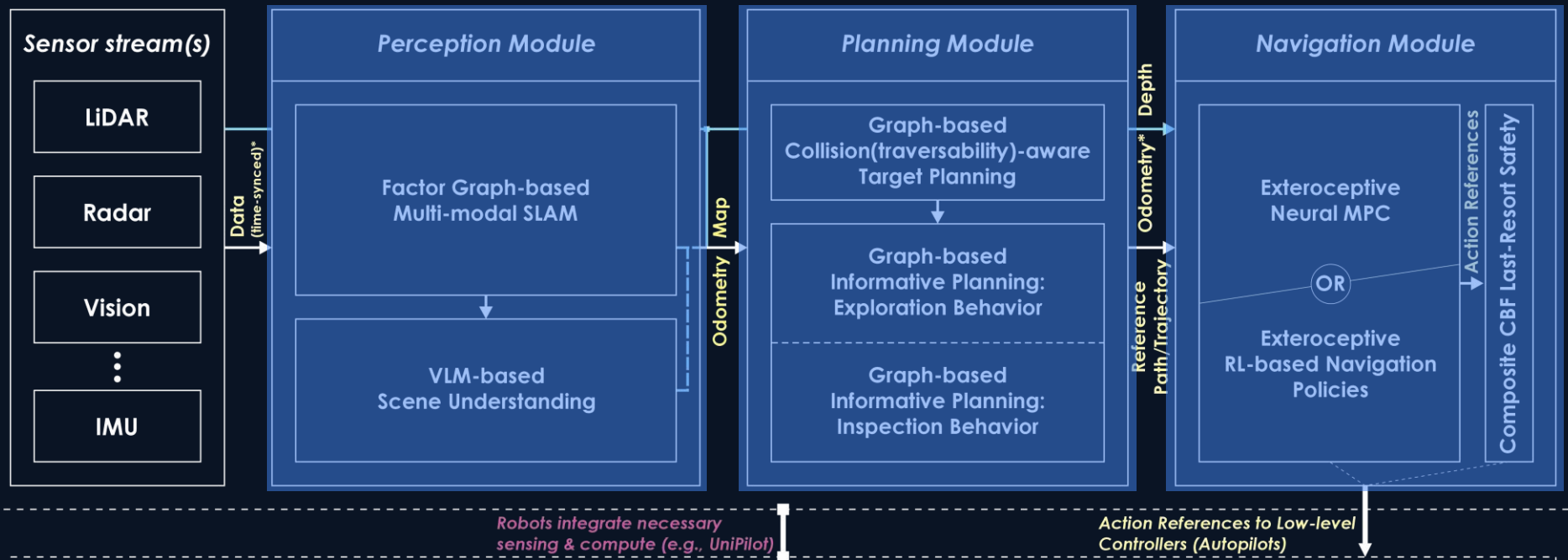
1. Core Autonomy: The Unified Autonomy Stack
2. Autonomous Exploration across Morphologies: OmniPlanner
3. Scene Reasoning and Embodied Q&A

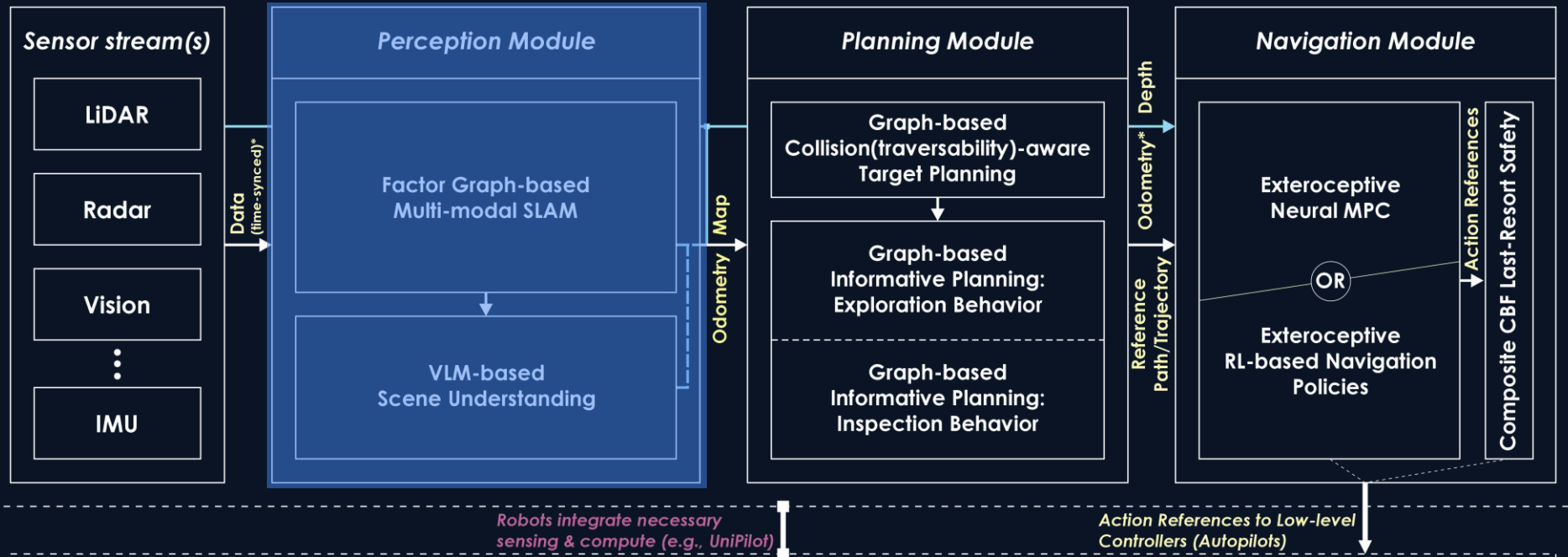
Unified Autonomy - Vision

Autonomous all-domain operations
scaling robotic autonomy in denied,
contested, degraded, and
unconventional environments.

Unified autonomy for Physical AI
across air, land, and sea.

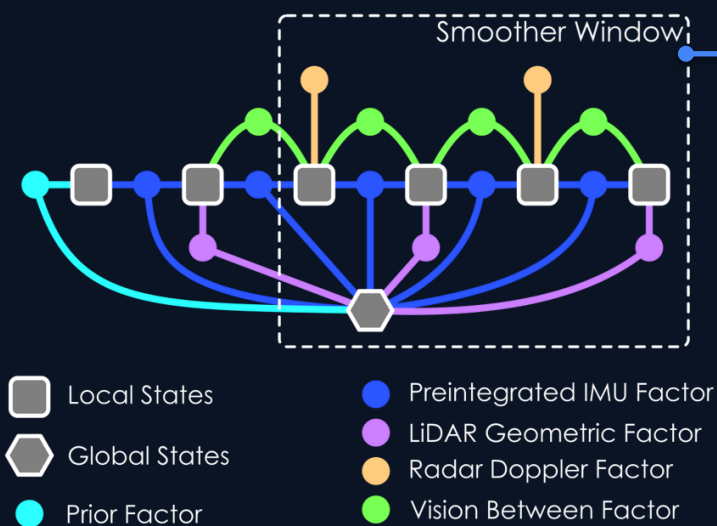




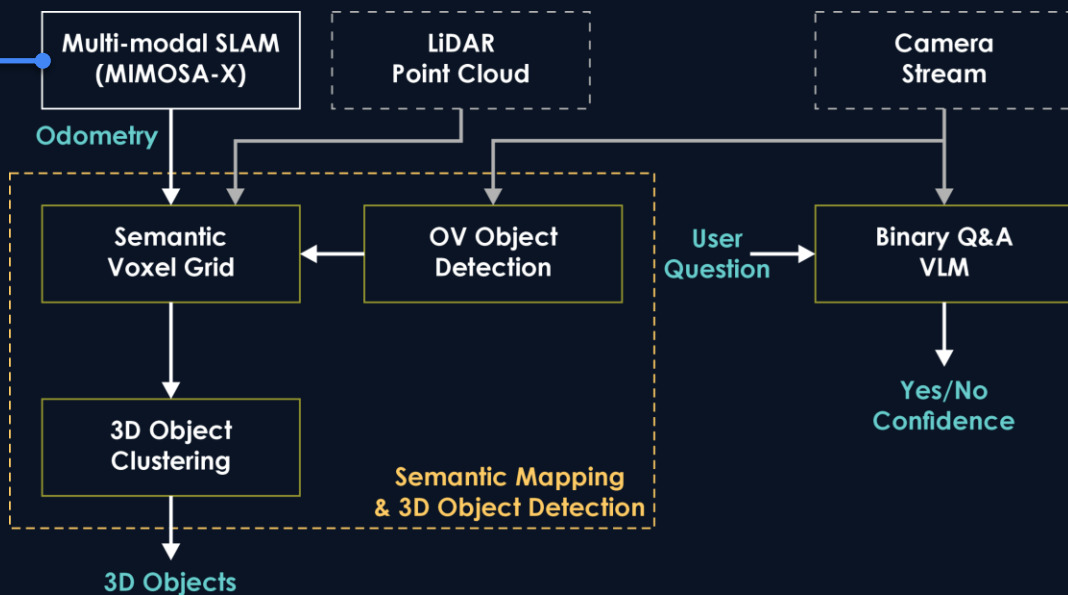


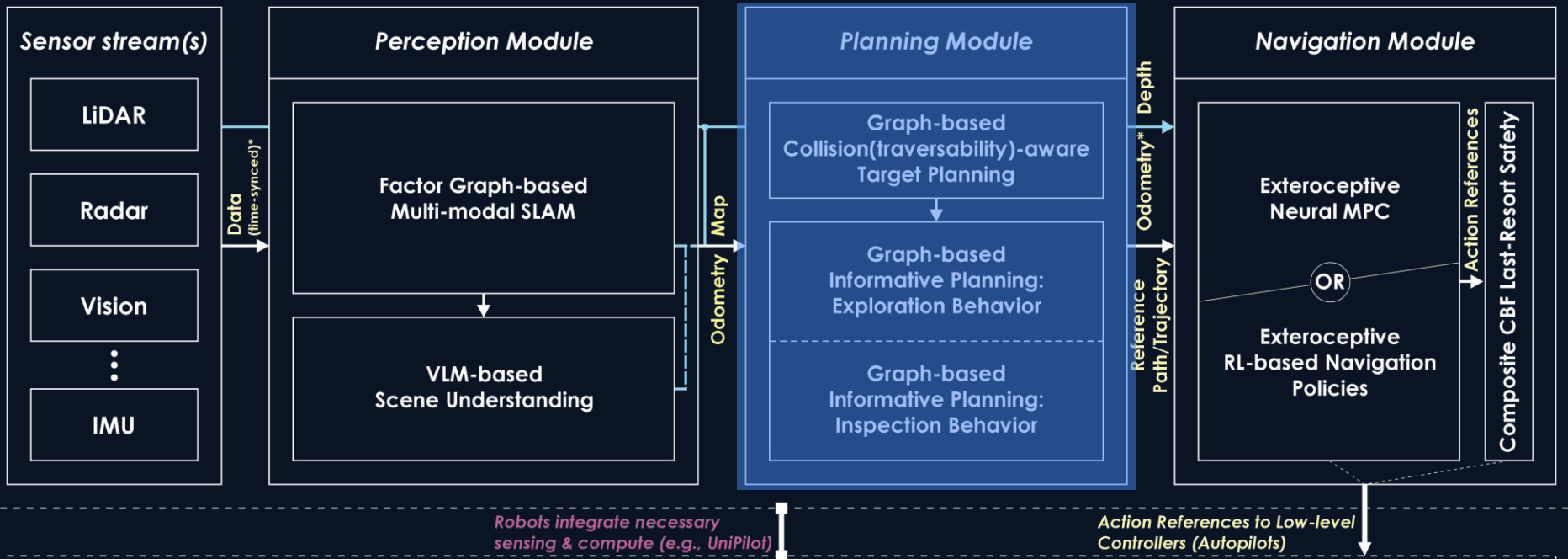
Multi-modal Domain-configured Perception

▪ Geometric Scene Reasoning



▪ Object-centric Scene Reasoning



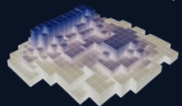


Omni-morphology Multi-behavior Planning

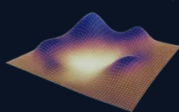
Embodiments



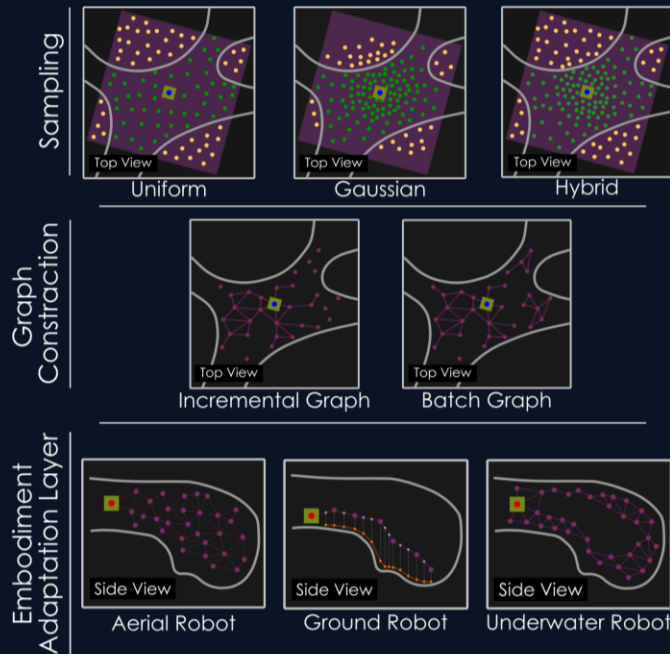
Volumetric Map



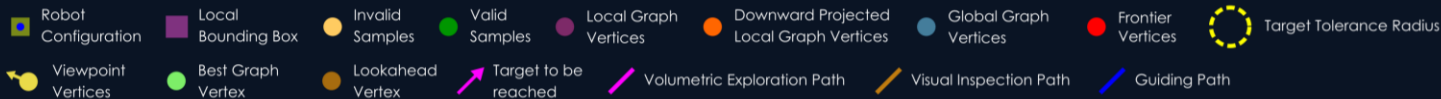
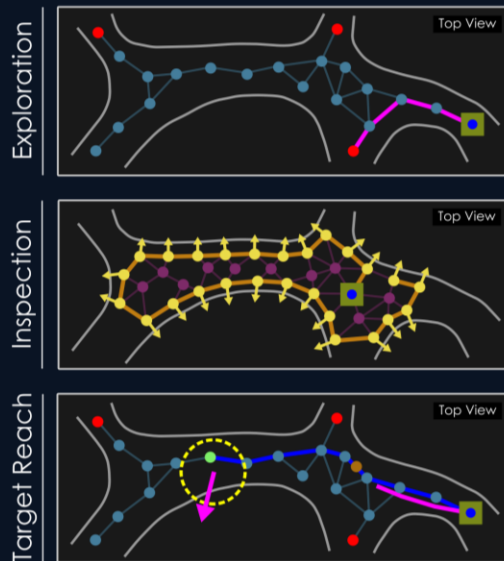
Elevation Map

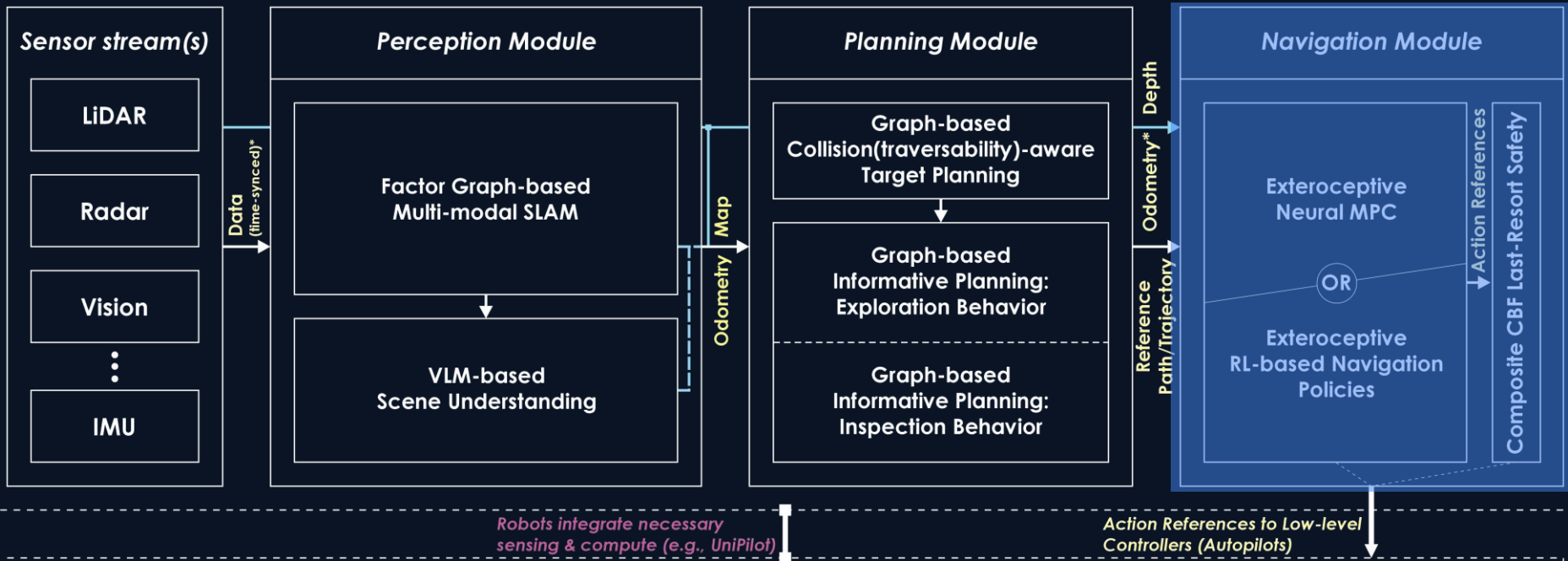


Planning Kernel

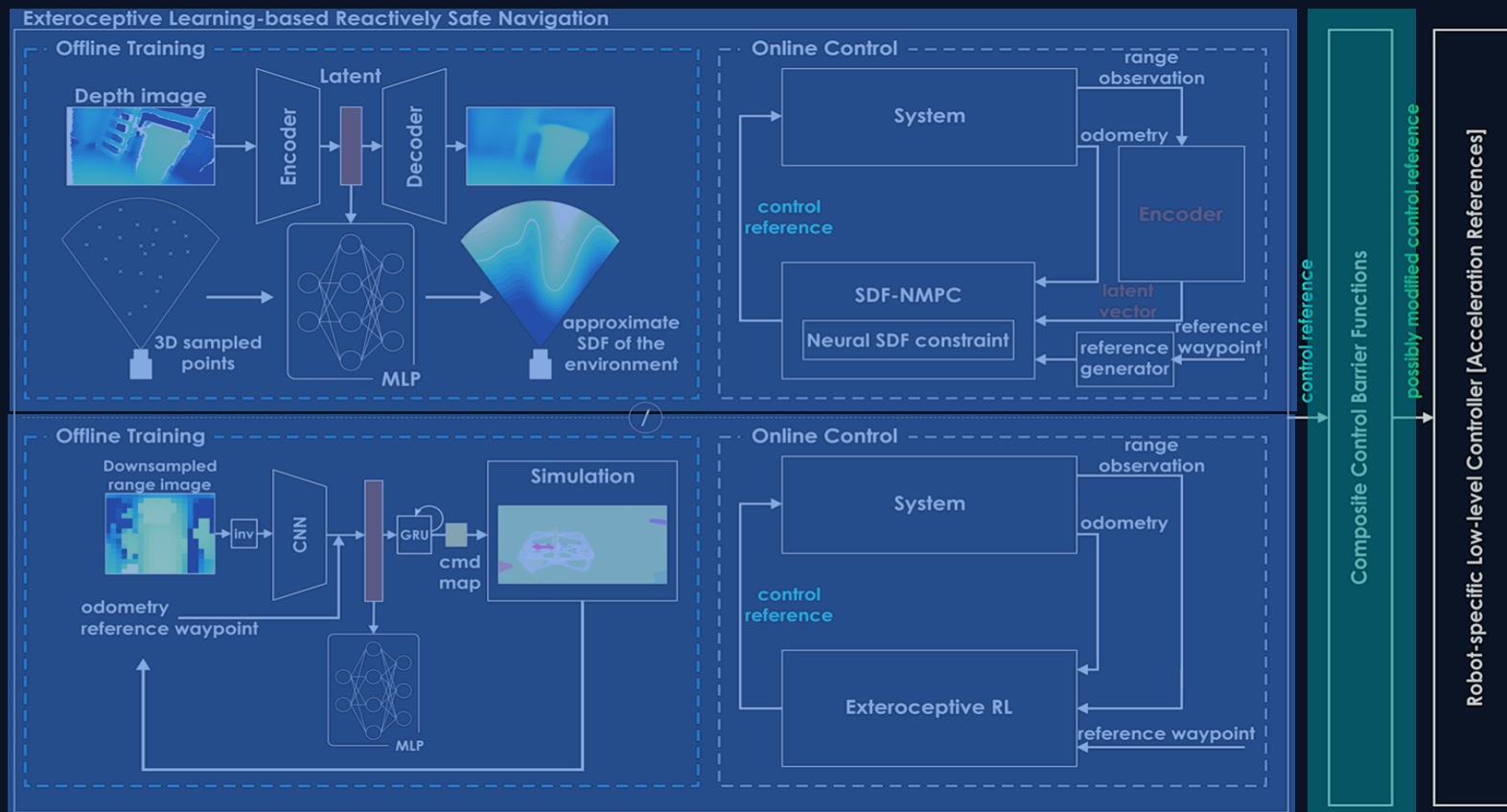


Behaviors





Multi-layered Navigation



Utilized Robot Morphologies



AR-1



AR-2



GR-1

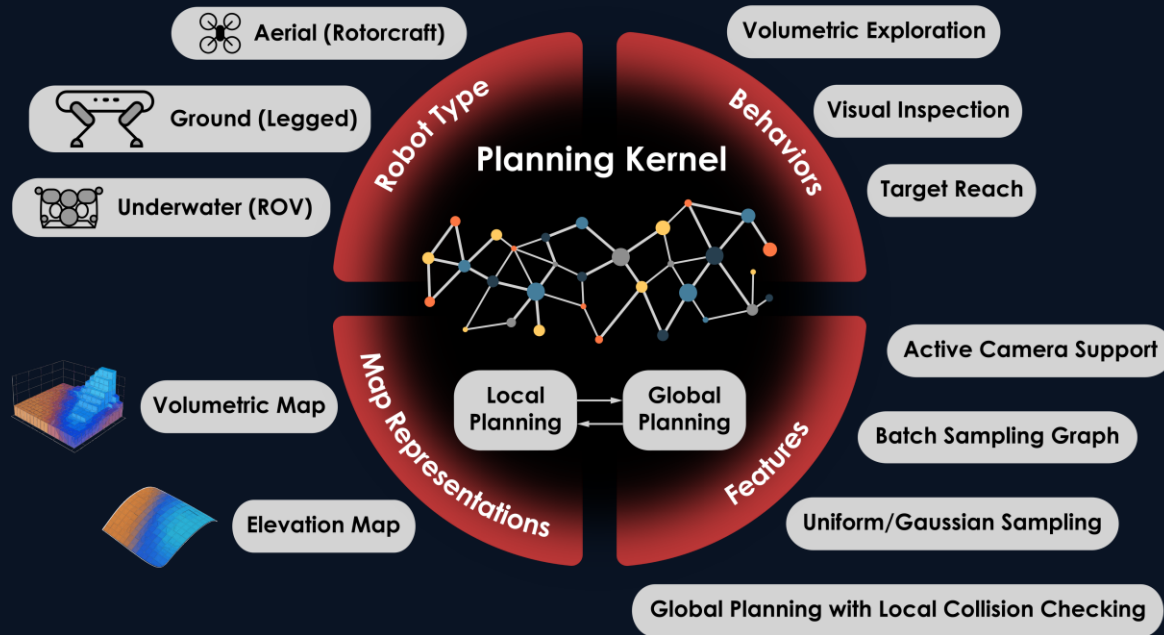


UniPilot

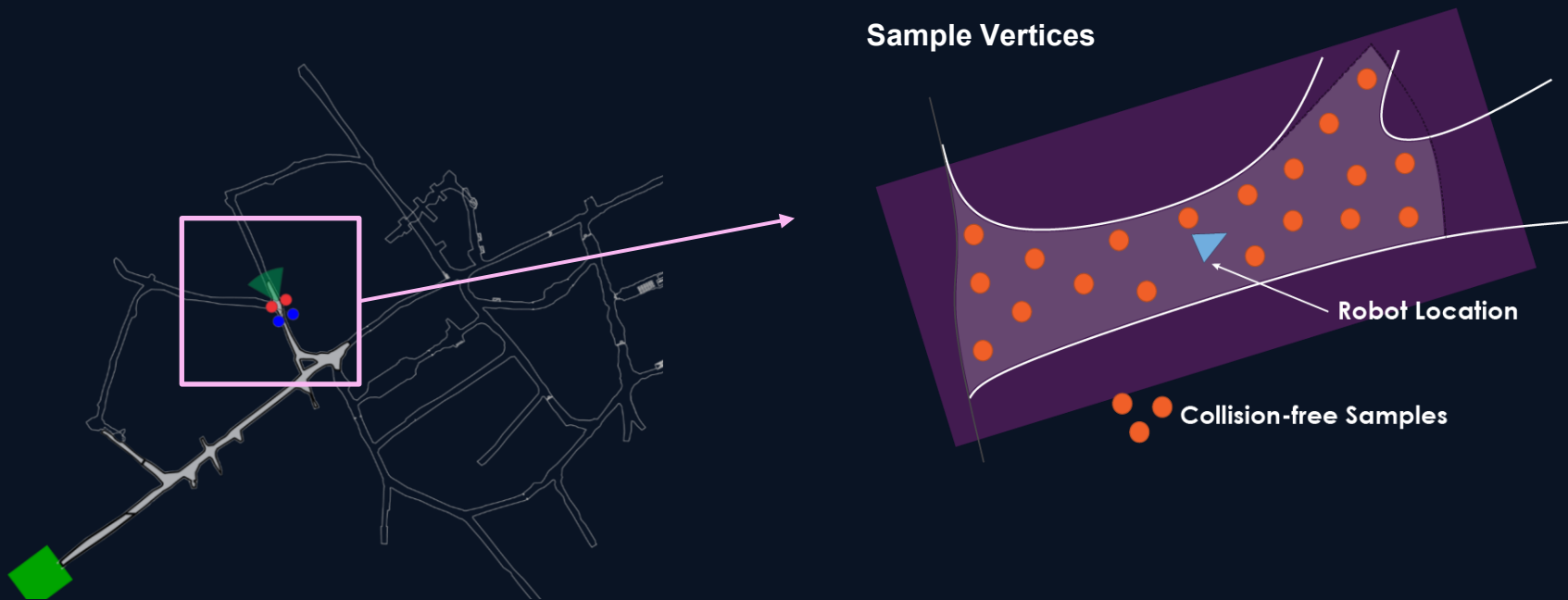
LiDAR	Ouster OS0-128
Radar	TI IWR6843AOP
Vision	FLIR Blackfly S 0.4MP
IMU	VectorNav VN-100
Compute	Khadas VIM4

LiDAR	RoboSense Airy or Hesai JT-128
Radar	TI IWR6843AOP-based unit
Vision	3 x MIPI Vision Components IMX296
IMU	VectorNav VN-100
Compute	NVIDIA OrinNX

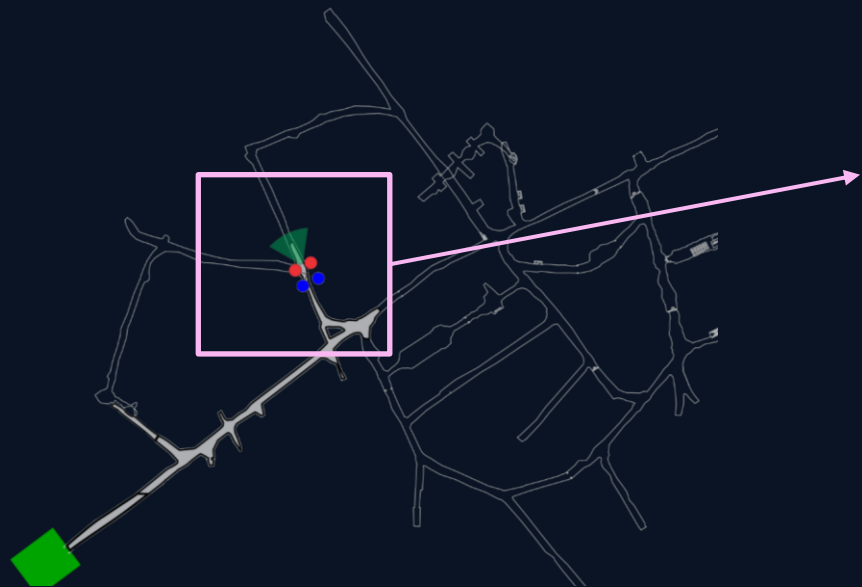
Universal Exploration and Inspection Path Planning across Robot Configurations



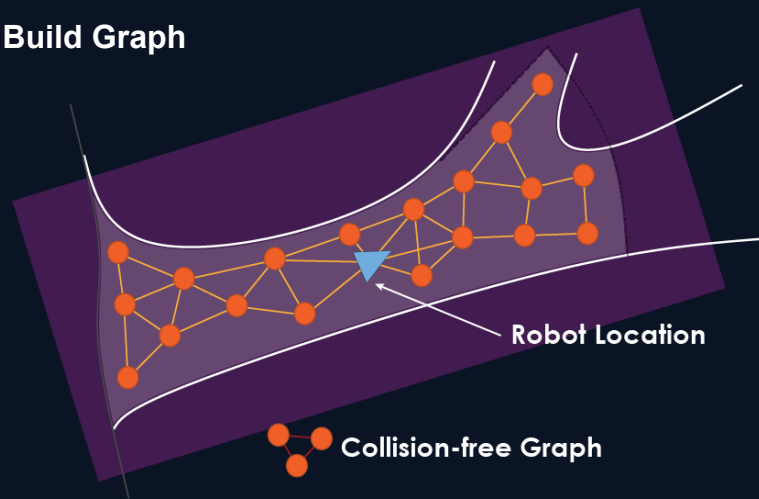
Planning Kernel – Local Planning Module



Planning Kernel – Local Planning Module



Build Graph

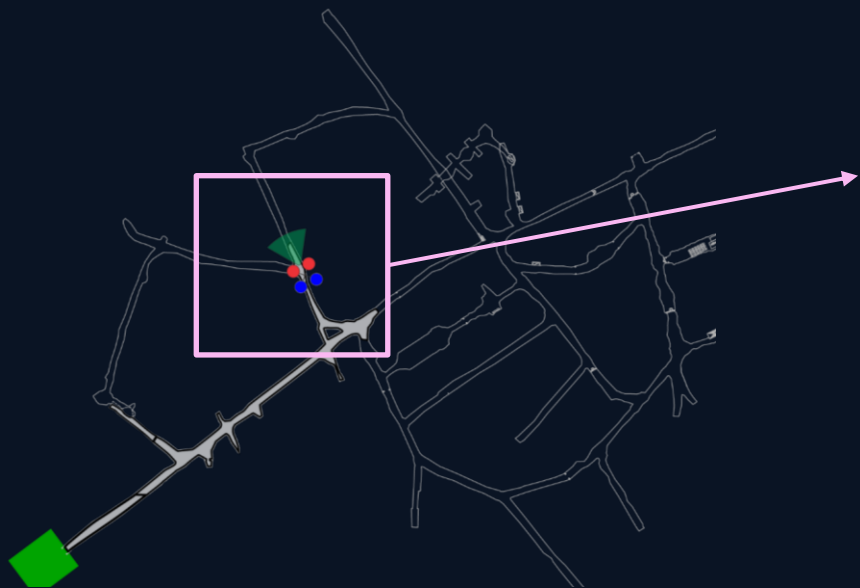


Planning Kernel – Global Planning Module

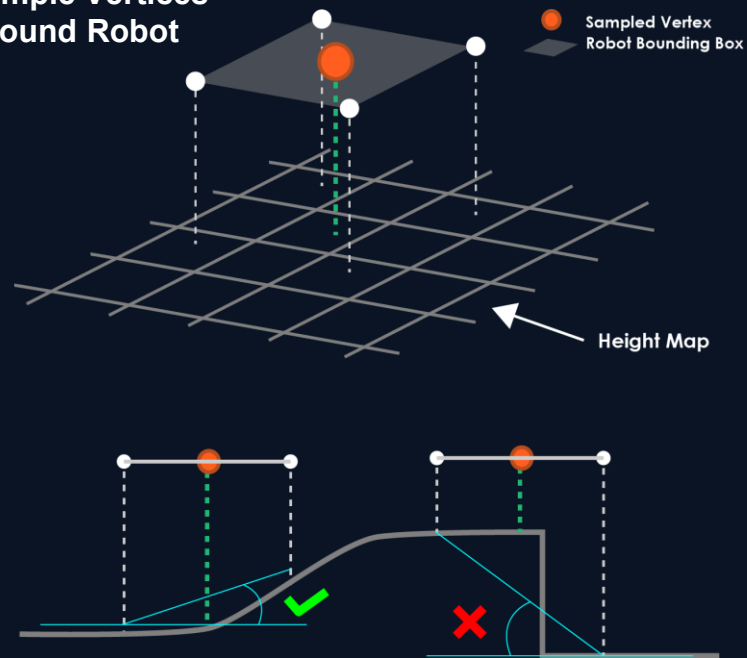


- Maintains sparse **Global Graph** in explored space
- Tracks exploration frontiers:
Vertices with high Exploration Gain
$$\Gamma_E^L \geq \Gamma_{E,thr}^L$$
- Tracks time remaining to return to home

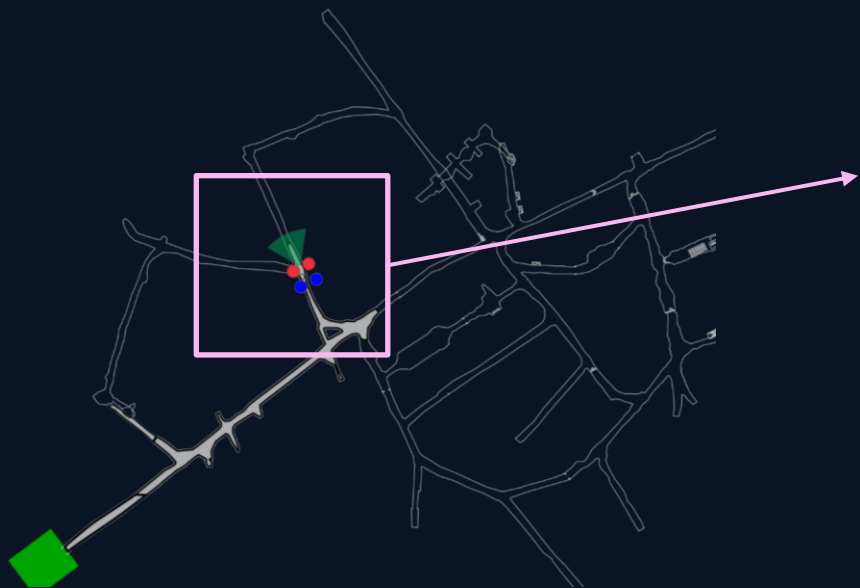
Ground Robot Adaptation



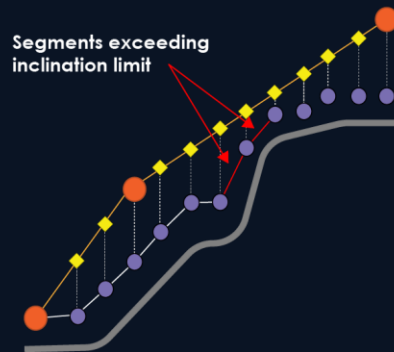
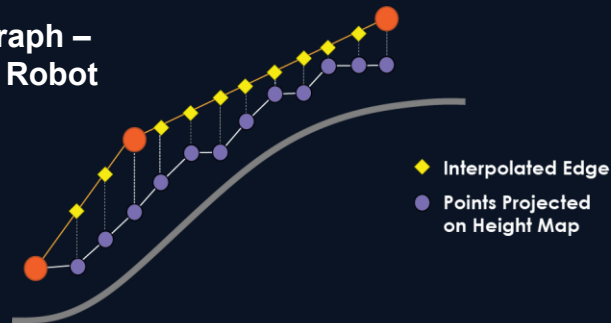
Sample Vertices – Ground Robot



Ground Robot Adaptation



Build Graph – Ground Robot



Behavior 1 – Volumetric Exploration

- Exploration of unknown volume using depth sensor
- Map representation: Volumetric map + Elevation map (ground robots)
- Bifurcated Local-Global planning architecture
- **Local Planner:** Exploration within a local space around the robot
- **Global Planner:** Repositioning to exploration frontiers. Safe return to home

Volumetric Exploration

Local Exploration Planner

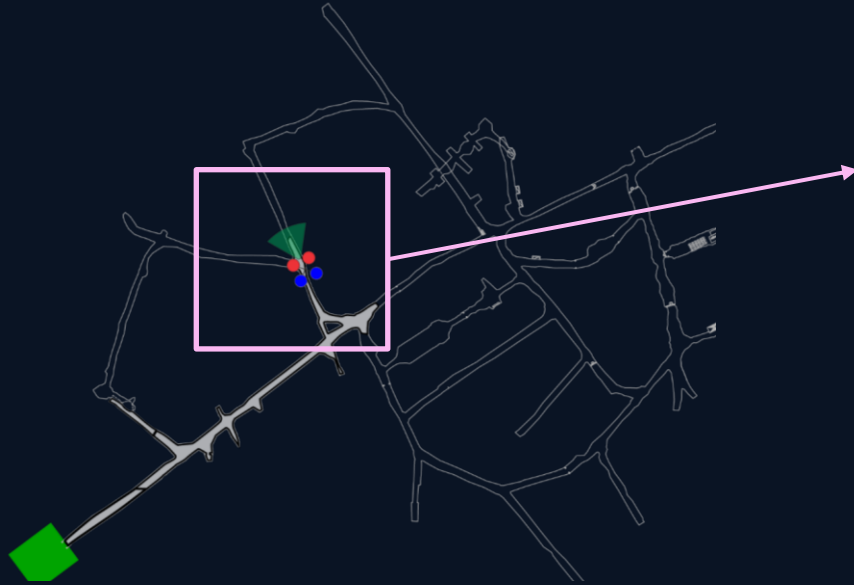
Rapid, sensor and motion
constraint aware exploration

Global Exploration Planner

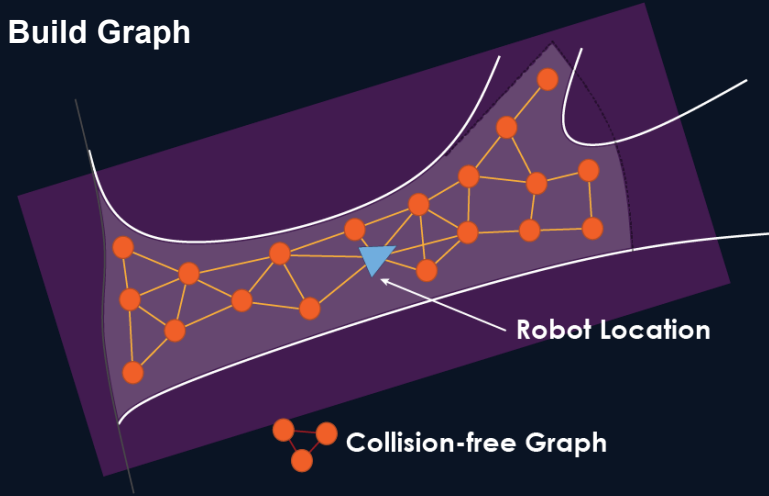
Global Repositioning

Autonomous Homing

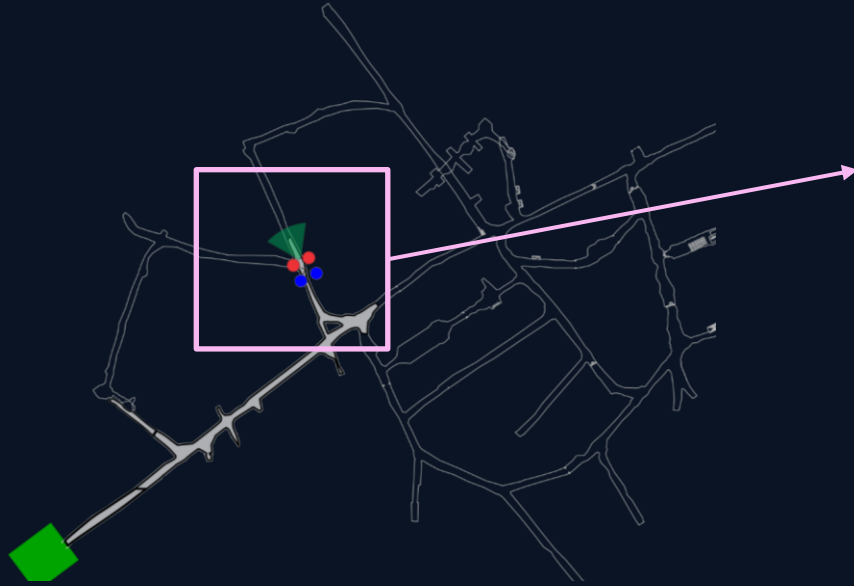
Behavior 1 – Volumetric Exploration – Local Planner



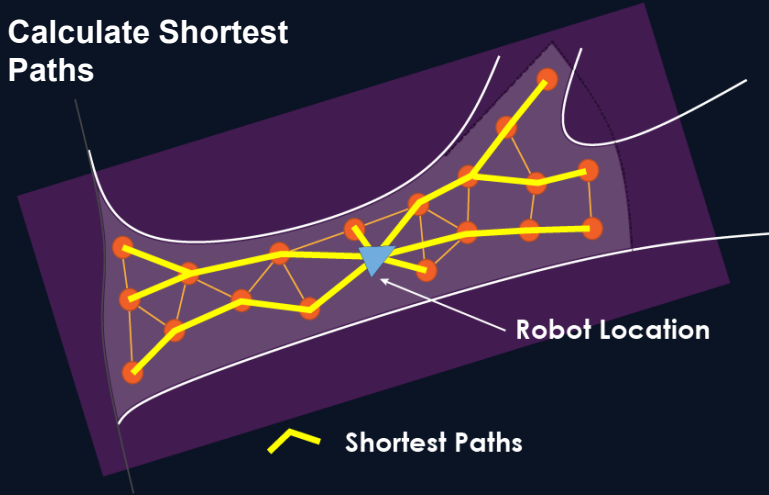
Build Graph



Behavior 1 – Volumetric Exploration – Local Planner

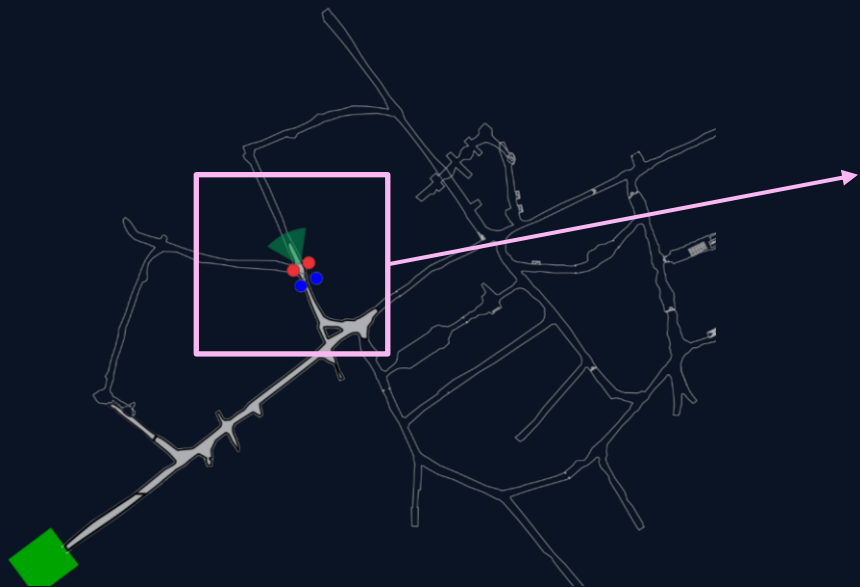


Calculate Shortest Paths



Calculate the set Σ_L of shortest paths from current robot location to each vertex

Behavior 1 – Volumetric Exploration – Local Planner



Calculate Exploration Gain

For path σ_i :

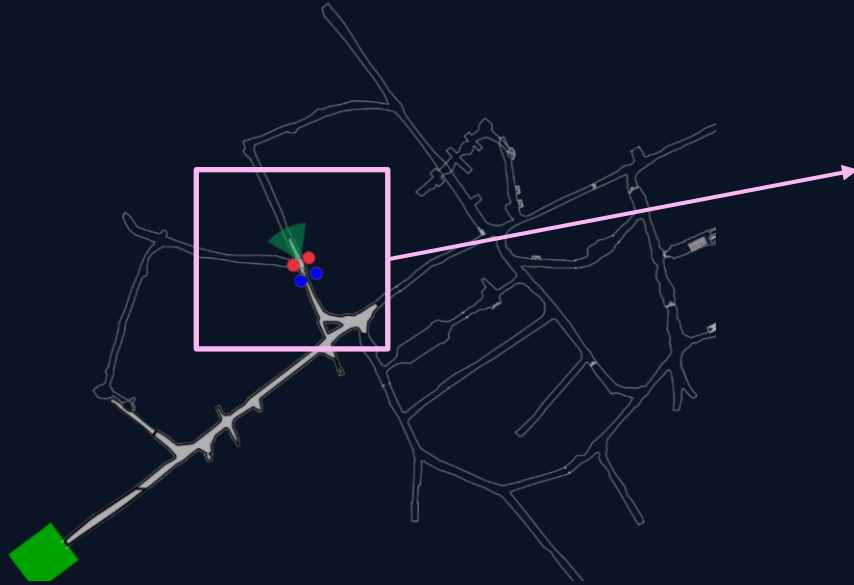
$$\Gamma_E^L(\sigma_i) = e^{-\zeta \mathcal{Z}(\sigma_i, \sigma_e)} \sum_{j=1}^{m_i} \Lambda(\nu_j^i) e^{-\lambda \mathcal{D}(\nu_1, \nu_j^i)}$$

Path direction penalty

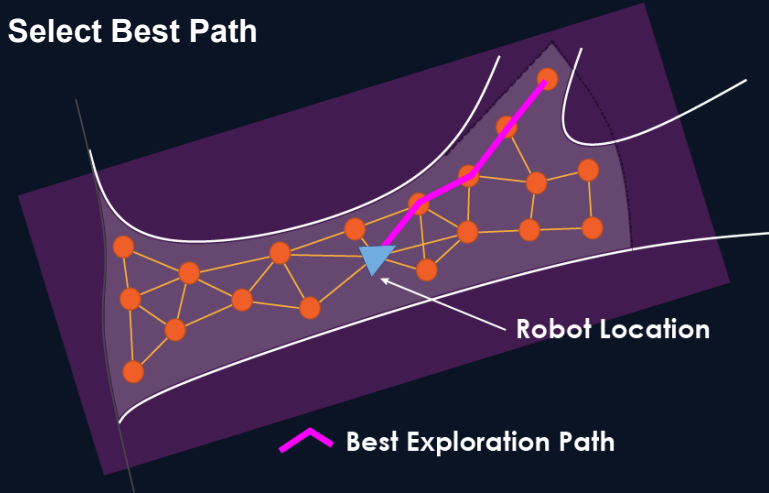
Volumetric Gain of vertex

Distance to vertex

Behavior 1 – Volumetric Exploration – Local Planner



Select Best Path



Behavior 1 – Volumetric Exploration – Global Planner

Global Repositioning



- Calculate time to home. Return to home if remaining time $\leq$ time to home
- Select frontier vertex $\nu_{G,i}^{\mathcal{F}}$ with highest gain

$$\Gamma_E^G(\nu_{G,i}^{\mathcal{F}}) = \underbrace{\mathcal{T}(\nu_{G,cur}, \nu_{G,i}^{\mathcal{F}})}_{\text{Exploring time remaining after reaching the frontier}} \underbrace{\Lambda(\nu_{G,i}^{\mathcal{F}})}_{\text{Volumetric Gain of the frontier}} \underbrace{e^{-\epsilon_D \mathcal{D}(\nu_{G,cur}, \nu_{G,i}^{\mathcal{F}})}}_{\text{Distance to the frontier (along the graph)}}$$

Exploring time remaining
after reaching the frontier

Volumetric Gain of
the frontier

Distance to the frontier
(along the graph)

Behavior 1 – Volumetric Exploration

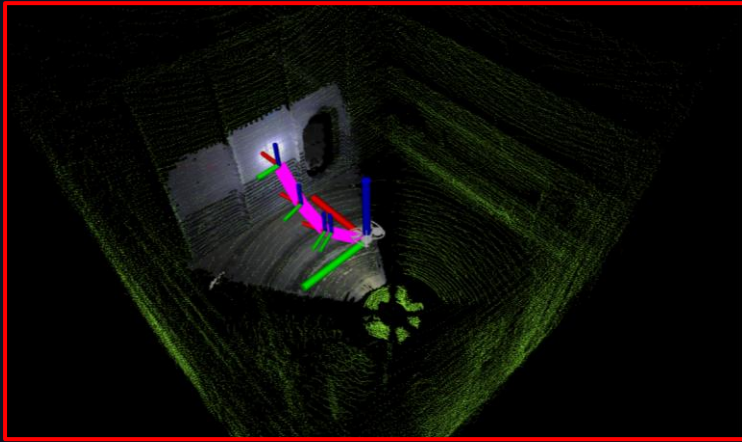


Aerial Robot

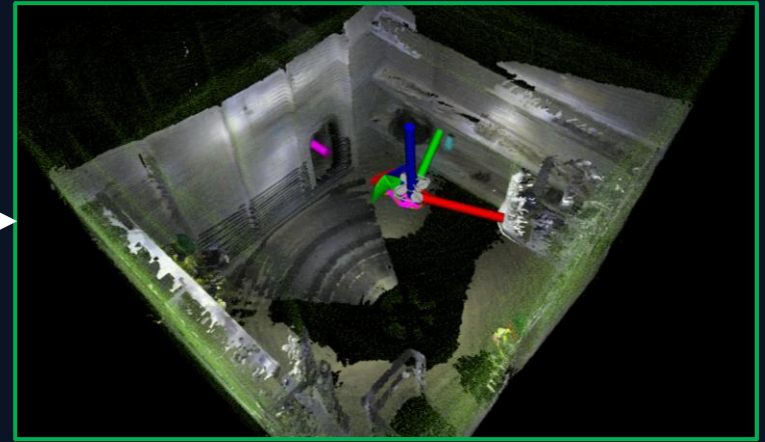
Legged Robot

Behavior 2 – Visual Inspection

Visual inspection of occupied surfaces of a volumetric map



Mapped Volume

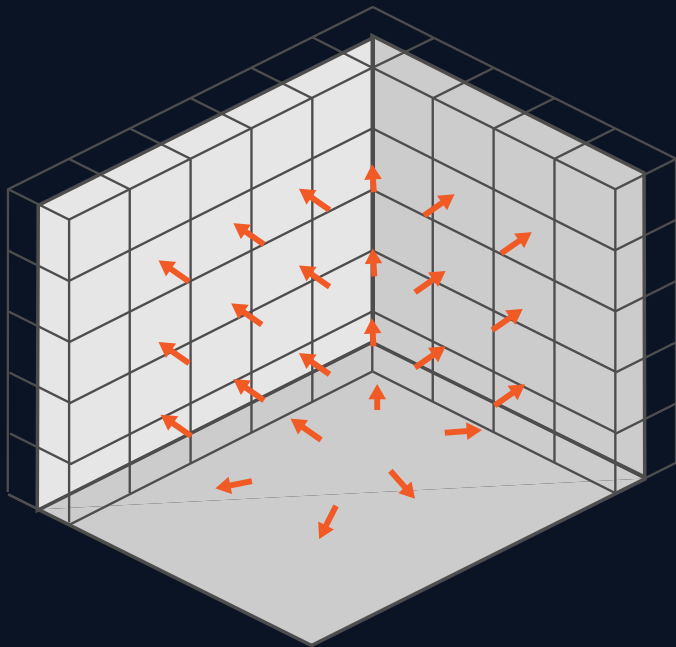


After inspection

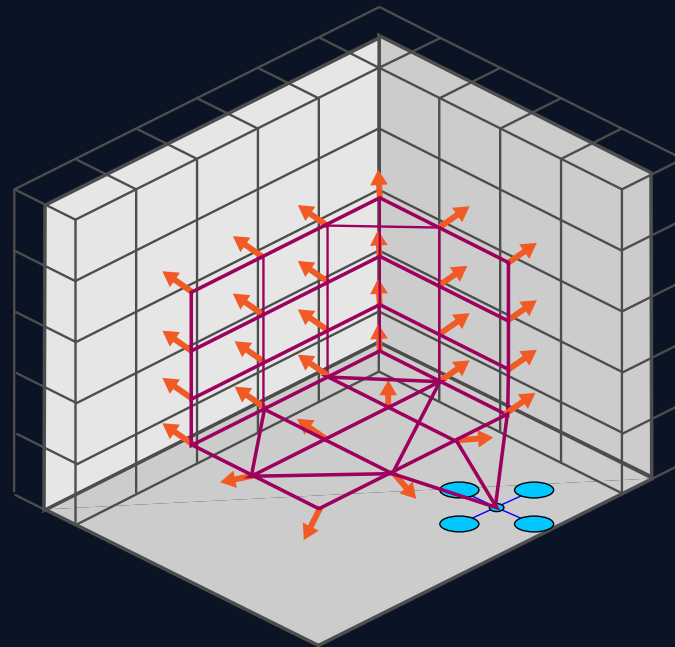
Green: Mapped but not inspected
Color: Mapped and inspected

Behavior 2 – Visual Inspection

Generate Viewpoints



Build Graph



Viewpoint Graph

Selected Viewpoints

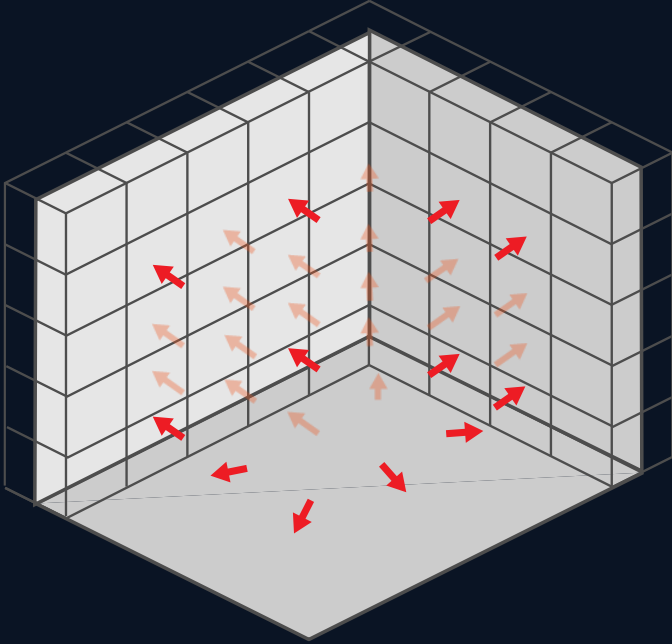
GVI Path



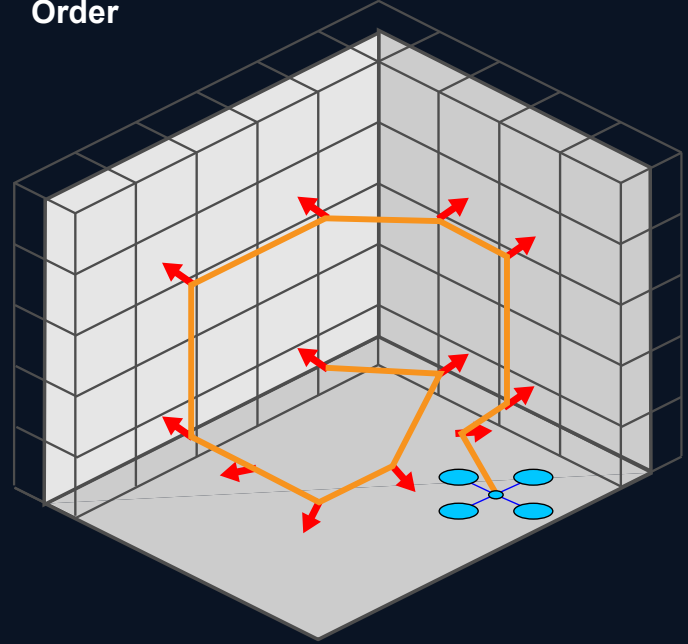
Current Robot Location

Behavior 2 – Visual Inspection

Select Viewpoint Set



Find Connecting Order



Viewpoint Graph

Selected Viewpoints

GVI Path



Current Robot Location

Behavior 2 – Visual Inspection



Aerial Robot

Autonomous Underwater Exploration & Inspection In Visually-degraded Environments

Mohit Singh, Mihir Dharmadhikari, and Kostas Alexis
Norwegian University of Science and Technology

The Research
Council of Norway

This work was supported by Research Council of Norway under projects
RESIFARM(NO-327292) and SENTIENT(NO-321235)



Underwater Robot



Behavior 3 – Target Reach

Goal: Find collision-free path to target location in unknown space

Approach:

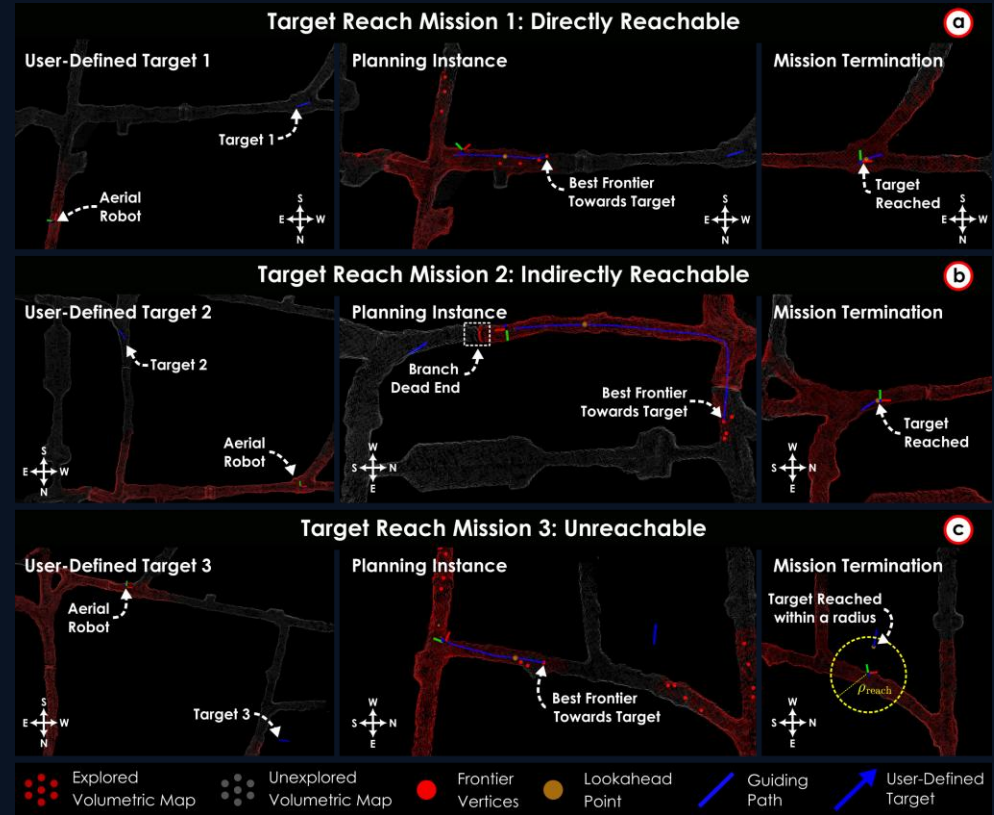
- Find best frontier to visit on the global graph:

$$\nu_{\text{best}} = \arg \min_{\nu_f \in \mathcal{F}} d_f^l + \lambda_{\text{bal}} d_f^u \quad \lambda_{\text{bal}} = 1 + \frac{d_f^u}{d_f^u + d_f^l}$$

Path length to frontier

Distance from frontier to target

- Get shortest path towards ν_{best} along the global graph $\rightarrow$ guiding path
- Iteratively, use Local Planning Module to advance along the guiding path



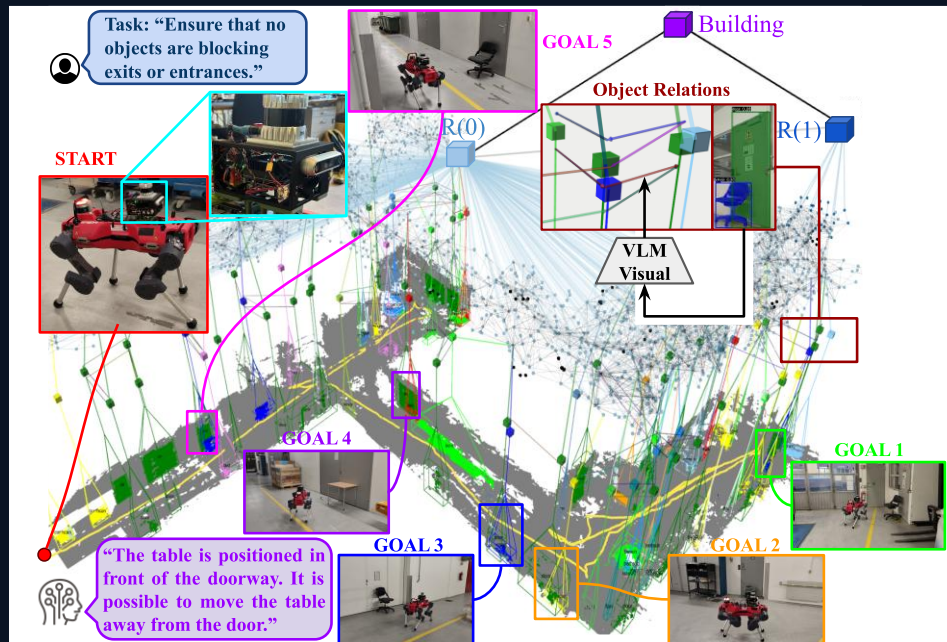
Behavior 3 – Target Reach



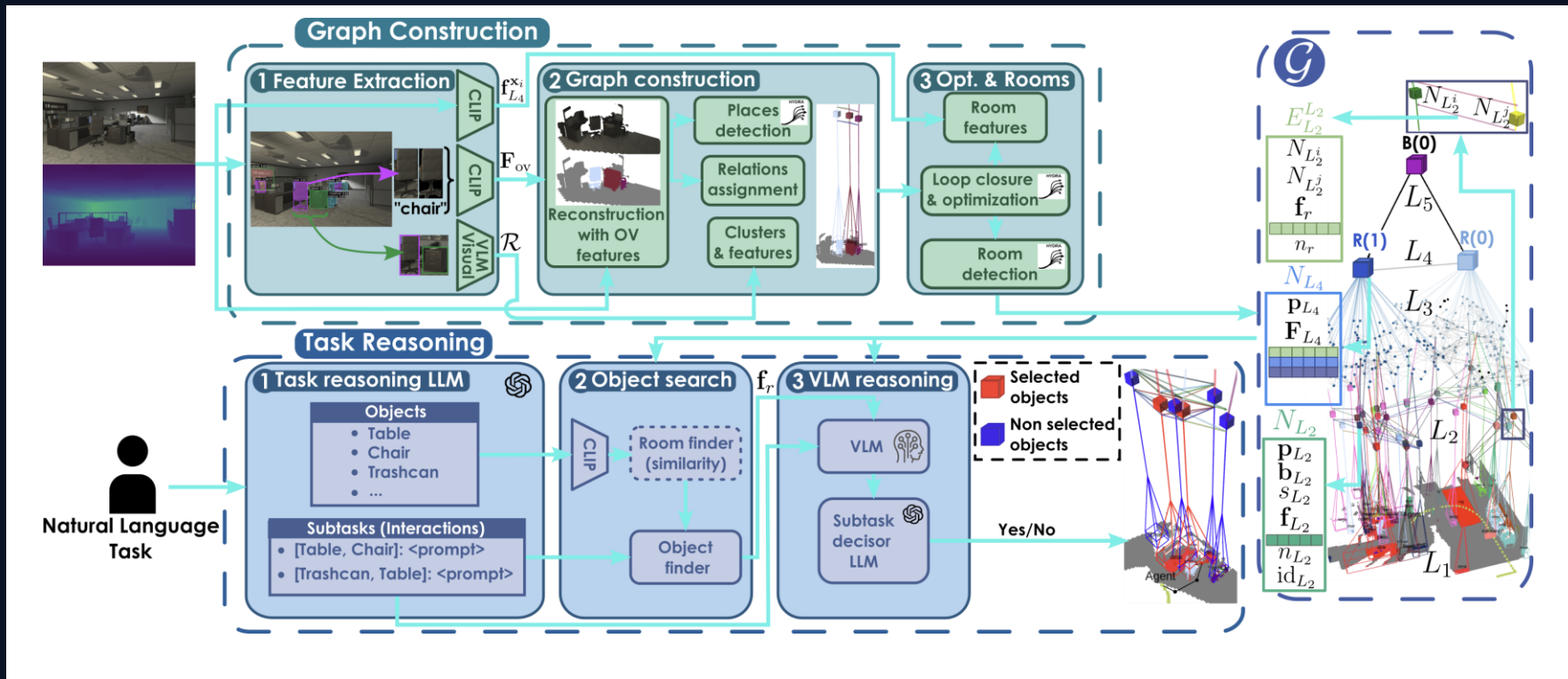
Relationship-Aware Hierarchical 3D Scene Graph

ReasoningGraph:

- Incremental hierarchical 3D scenegraph with open-vocabulary features.
- Object-relational features extracted with Vision Language Models.
- Reasoning about tasks:
- Interpret natural-language tasks.
- Identify relevant objects.
- Evaluate interaction feasibility through a VLM.



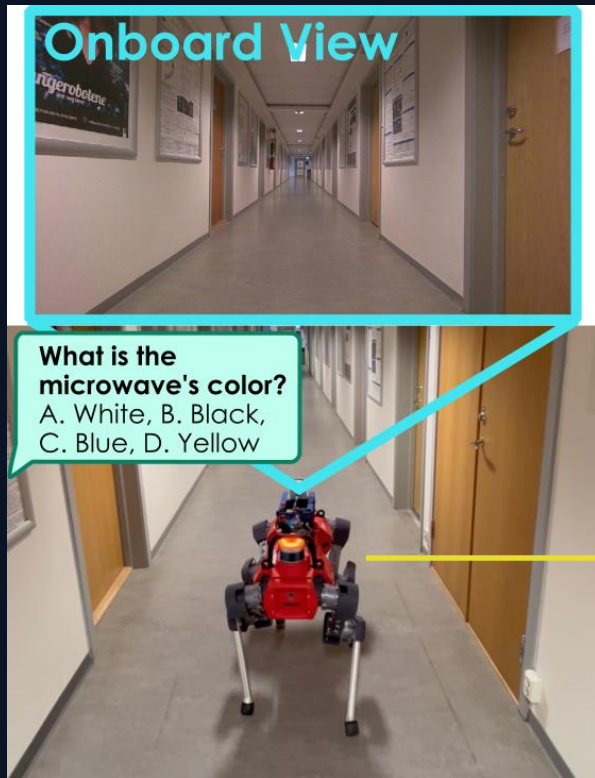
ReasoningGraph: Method Overview



Experimental Results

Task 2: Ensure Exits And Entrances
Are Not Blocked

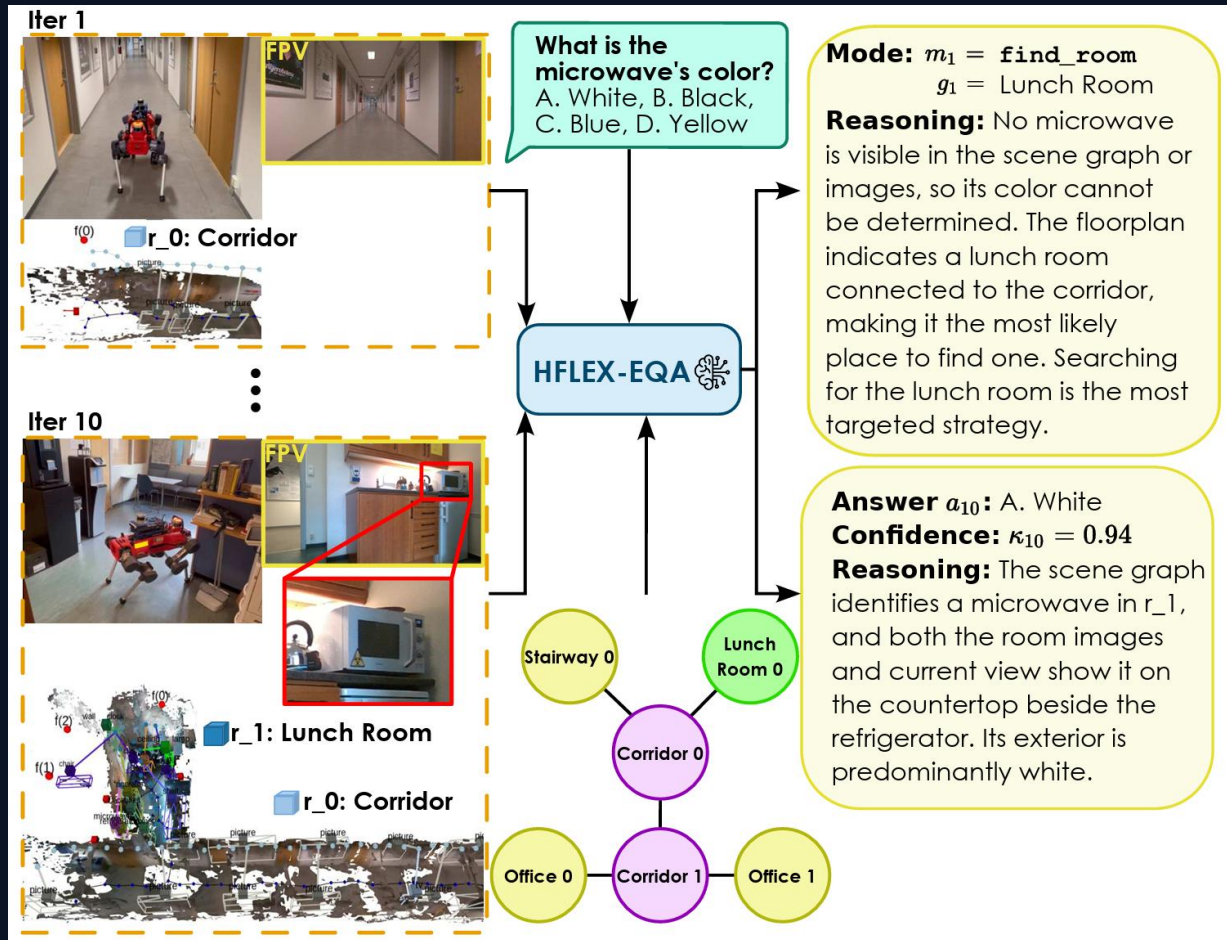
Vision-Language Exploration for Embodied Q&A



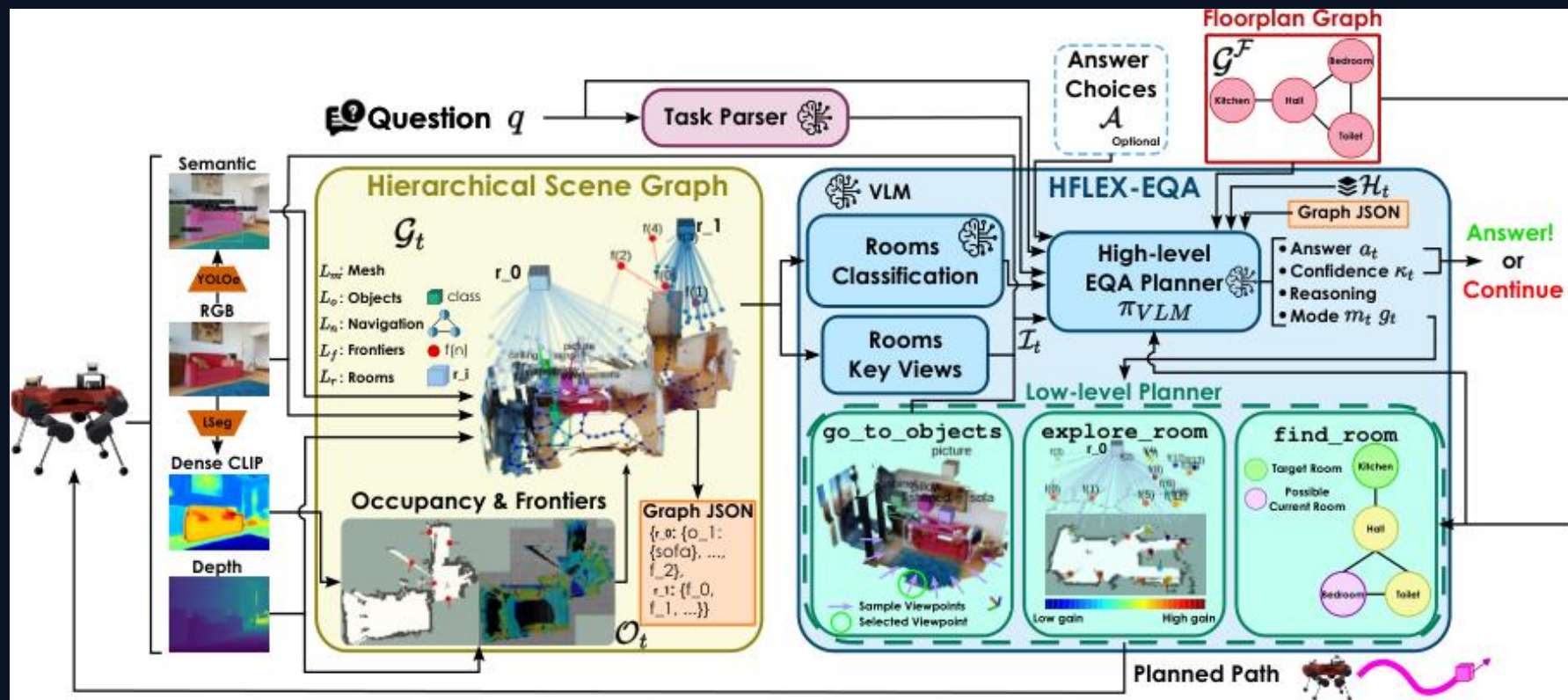
No microwave is visible in the scene graph or images, so its color cannot be determined. The floorplan indicates a lunch room connected to the corridor, making it the most likely place to find one. **Searching for the lunch room is the most targeted strategy.**

HFLEX-EQA*

Hierarchical Floorplan-Guided Vision-Language Exploration for Embodied Question Answering



HFLEX-EQA



HFLEX-EQA: Experiments

Are the fridge and chair the same color?

A. Yes, B. No



Where can put my jacket?

A. On the hanger in the office, B. In the cabinet, C. On the chair



Is there any visual support in the meeting room?

A. No, B. Yes



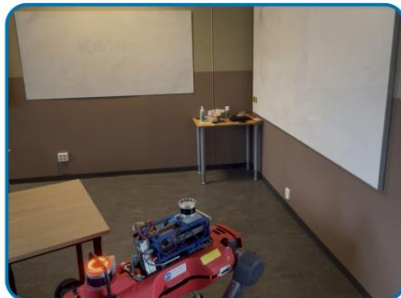
How many chairs are there to have lunch?

A. Three, B. Five, C. Seven, D. Two



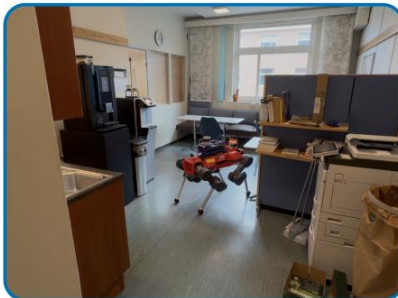
What is written on the whiteboard?

A. Robotics, B. Deep learning,
C. Hello world, D. University



What is the microwave's color?

A. White, B. Black,
C. Blue, D. Yellow



Where can I find my laptop?

A. In the cabinet,
B. On the chair,
C. On the desk



Real-world Experiments

Are the fridge and chair the same color?

A. Yes, B. No



Where can put my jacket?

A. On the hanger in the office, B. In the cabinet, C. On the chair



Is there any visual support in the meeting room?

A. No, B. Yes



How many chairs are there to have lunch?

A. Three, B. Five, C. Seven, D. Two



What is written on the whiteboard?

A. Robotics, B. Deep learning,
C. Hello world, D. University



What is the microwave's color?

A. White, B. Black,
C. Blue, D. Yellow



Where can I find my laptop?

A. In the cabinet,
B. On the chair,
C. On the desk

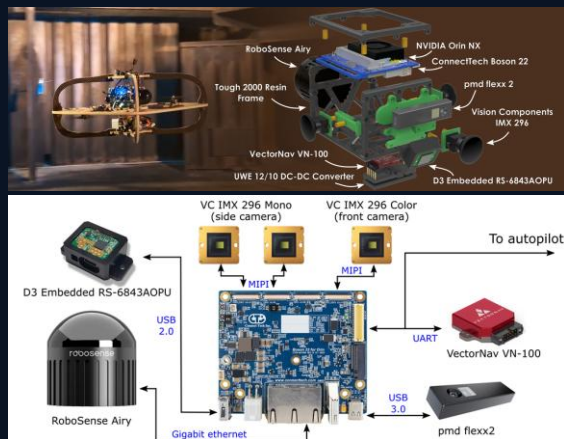


Open-Source & Community-ready

Open-source Code

- https://github.com/ntnu-arl/unified_autonomy_stack
- https://github.com/unr-arl/gbplanner_ros
- https://github.com/ntnu-arl/reasoning_hydra
- https://github.com/ntnu-arl/hflex_eqa
- <https://github.com/ntnu-arl/sdf-nmpc>
- https://github.com/ntnu-arl/aerial_gym_simulator
- <https://github.com/ntnu-arl/mimosa>
- <https://github.com/ntnu-arl/reaqrovio>
- <https://github.com/ntnu-arl/DeepVL>
- https://github.com/unr-arl/mbplanner_ros
- https://github.com/ethz-asl/mav_control_rw
- <https://github.com/ethz-asl/nbvplanner>

Reference Hardware



Sim Infrastructure

- Aerial Gym Simulator
- Port into official NVIDIA Isaac Lab
- Complete simulation + navigation policy training pipeline

