Task-Driven Navigation: Leveraging Experience using Deep Learning

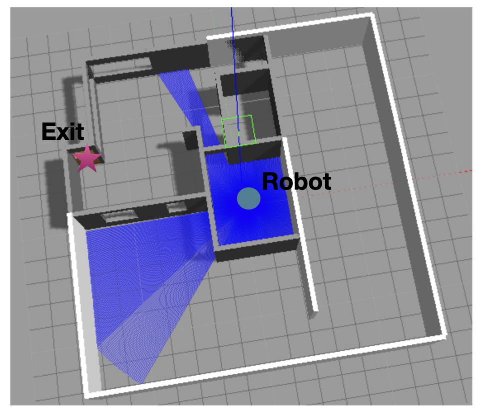
Zili Wang[1], Sean B. Andersson[1,2], Roberto Tron[1,2]

[1] Division of Systems Engineering, [2] Department of Mechanical Engineering
Boston University, MA 02155

Introduction

Problem Formulation

Given a mobile robot equipped with a laser range scanner and a limited-footprint camera in an unknown bounded two-dimensional environment, create an algorithm that ideally minimizes the total path traveled and the total number of camera measurements to steer the robot from a random initial position to the exit of a building.



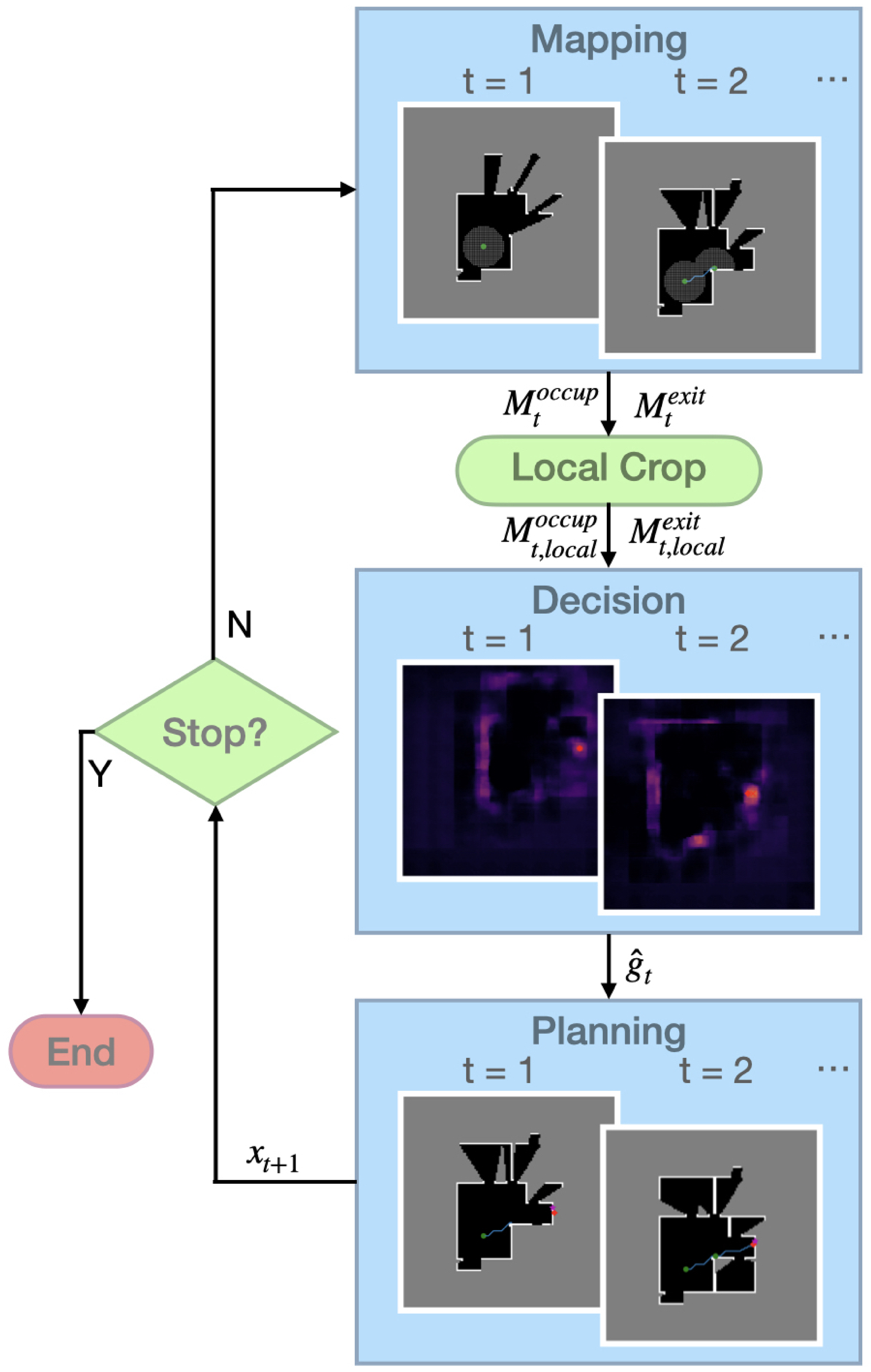
Other Work

- **Task-based robot navigation**
  Expensive computation, heavy sensors, or requires the complete floorplan
- **Frontier-based exploration**
  Navigation only aims to complete the map

Our Contribution

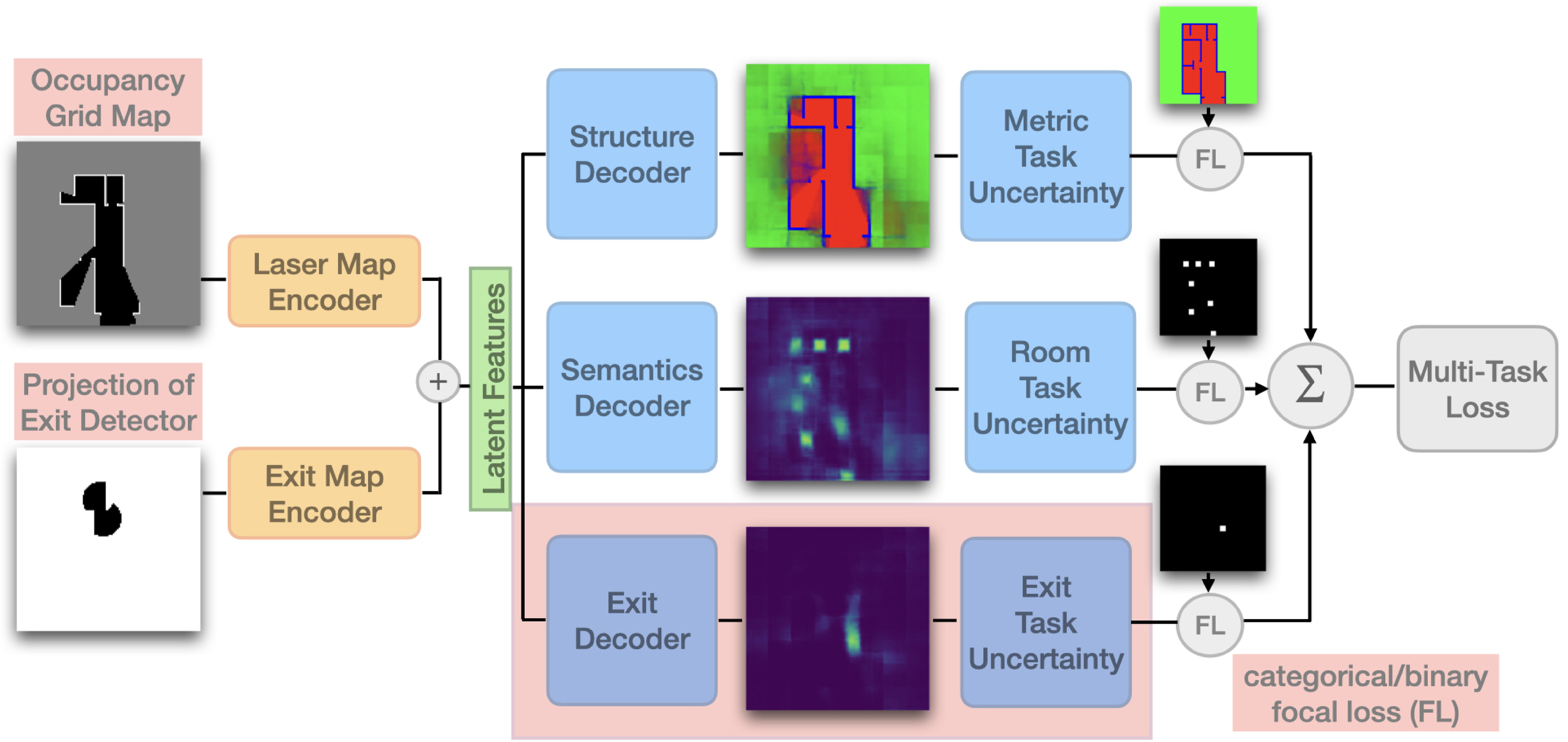
- We trained a multi-task auto-encoder deep neural network to predict the exit of the unknown building from a partial local map.
- We provided a more efficient search strategy that combines data-driven estimation and traditional motion planning to drive the robot to the target.

System Overview



- **Mapping**
  Acquire laser and optional camera measurements, and accumulate them into two occupancy maps:
  1) laser map - raycast the laser measurement
  2) exit map - project the exit detector result from the camera measurement
- **Decision**
  Pass a local extraction of the maps to an auto-encoder network that predicts a probability heatmap of the the exit, then estimate and select the interim goal accordingly.
- **Planning**
  Use the A^* algorithm to find a feasible trajectory from the current robot position to the interim goal.

Exit Prediction Network



To predict the exit location, we use a multi-task auto-encoder architecture with skip connections, local laser map and exit map as the inputs, and focal loss.

- **Deep Neural Network (DNN) auto-encoder**: leverage prior experience with similar environments to predict likely exit locations
- **Multi-task auto-encoder**: a shared latent space representation that includes both metric and semantic information of the current local environment
- **Dataset**: exit navigation with known exit locations in more than 380 residential floorplans.

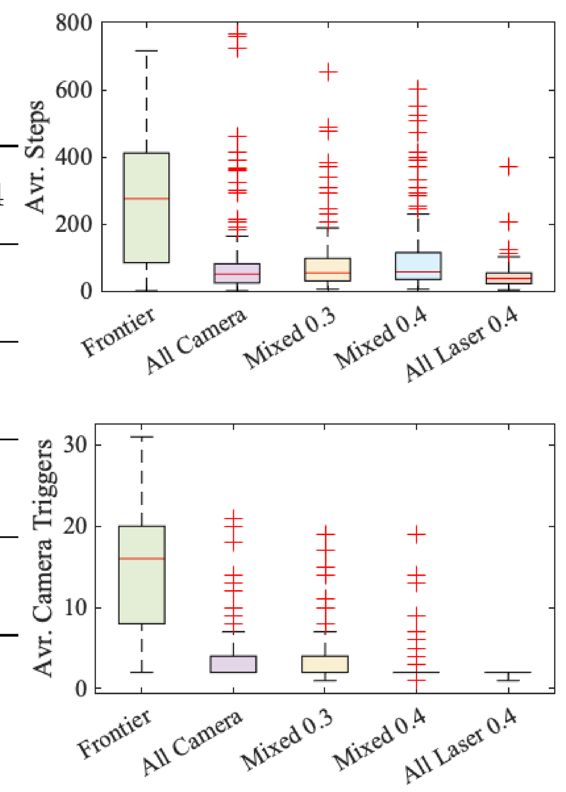
Camera Measurements and Termination

	Frontier	All Laser +Cameras	Mixed Laser +Camera 0.3	Mixed Laser +Camera 0.4	All Laser 0.4
Thresh. Camera Trigger			✓	✓	
Camera Termination	✓	✓	✓	✓	
Threshold Termination					✓

- **Smoothed exit density**: convolution of the exit heatmap with a 7 × 7 box kernel
- **Threshold**: a value decided upon the smoothed exit density, to trigger the camera or terminate

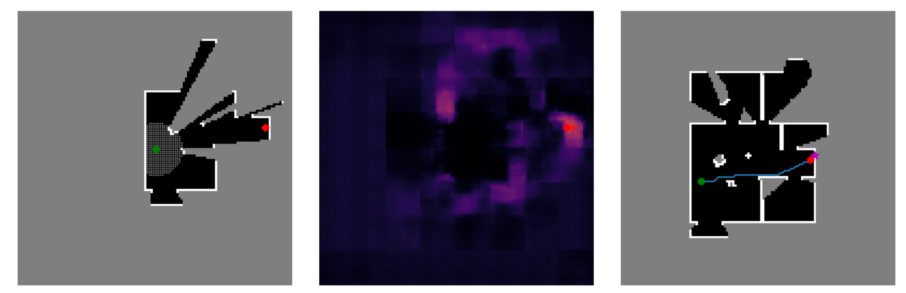
Results

	Frontier	All Laser +Camera	Mixed Laser +Camera 0.3	Mixed Laser +Camera 0.4	All Laser 0.4
Success Rate	100	100	100	100	57.8
Avr. Steps	265.8 (274, 190.4)	**86.6 (50, 132.1)**	123.3 (53.5, 238.5)	142.7 (56.5, 295.8)	41.5 (36.5, 35.7)
Avr. Success in Steps	265.8 (274, 190.4)	**86.6 (50, 132.1)**	123.3 (53.5, 238.5)	142.7 (56.5, 295.8)	43.9 (38, 38.6)
Avr. Camera Triggers	13.9 (15, 7.4)	4 (3, 3.2)	2.5 (2, 2.8)	**2.2 (1, 2.4)**	**0 (0, 0)**



- **Simulation**: we ran the exploration algorithms on 36 new environments, with the robot starting from the same initial random positions for each method.
- **Completeness** : the termination criteria of all the cases guarantee the exit will eventually be detected by the camera, except All Laser 0.4.
- Higher threshold leads to greater average steps and less camera triggers.
- Our methods are always superior to the baseline frontier-based method.

Robustness of Network



- **Simulation**: we tested on the environments containing random obstacles that were not in the training.
- Result shows slightly worse metrics.

Future Work

- Introduce lightweight models to improve efficiency.
- Apply our method in high level hierarchical path planning.
- Conduct ablation study to show the necessity of multi-task.