Training Deep Probabilistic Models for Semantic Mapping with Mobile Robots

Kousuke Ariga* Kaiyu Zheng*

Advisor: Dr. Andrzej Pronobis

Introduction

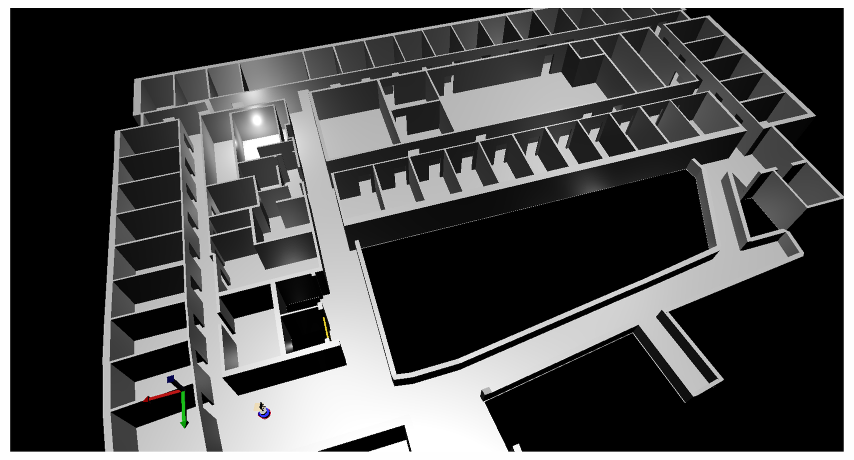
Background An autonomous agent needs to understand the semantics of an environment in order to perform complex actions. The goal of our project is to enable mobile robots to learn a unified deep genertive model that spans from low-level sensory input to high-level human semantics of the environment. Training such model requires a large dataset.

Contributions We created a large dataset of fully-annotated localization, sensory information and *virtual scans*, a low-level geometry representation of robot-centric environment. Along the way, we also made the following contributions:

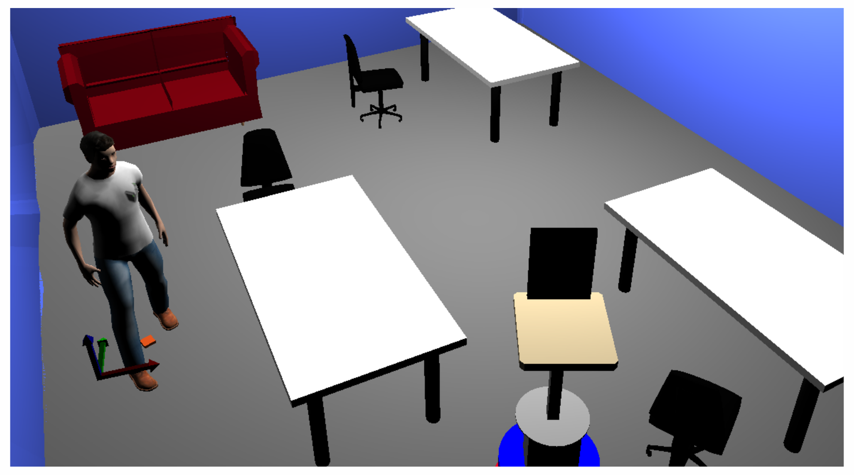
1. Developed an algorithm to generate randomly furnished simulated world.
2. Improved the robustness of the navigation system to allow automatic data collection.
3. Implemented an algorithm to generate virtual scans.
4. Developed a pipeline to generate the dataset from laser data and odometry.

We show how this dataset can be used for learning deep models.

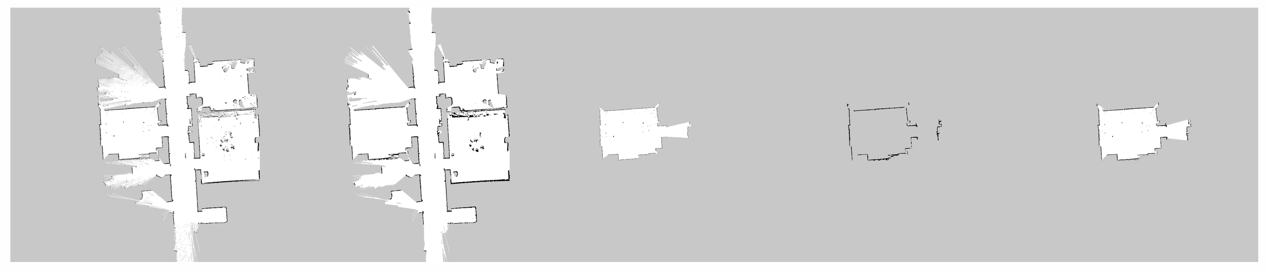
Generating Simulated Worlds



What: A system for sampling objects from a pre-defined probability distribution of each room class to furnish the simulated world.
Why: We can collect data efficiently as simulation can be run in parallel and without human supervision.

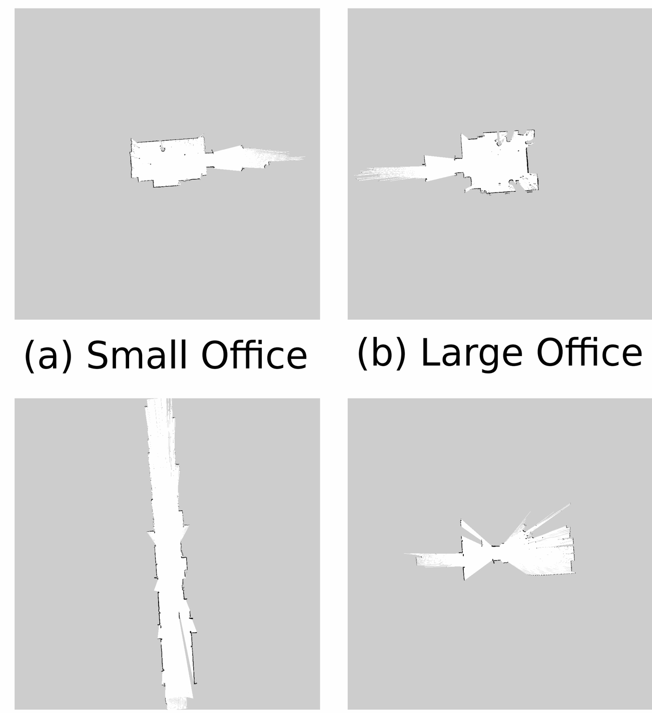


Representing Local Environments



What: An algorithm that extracts a local, robot-centric sensory representation (i.e. virtual scan) of the immediate robot environment.
Why: We want to model the geometry and semantics of a local environment only.

Our robot DUB-E and our computing platform gerlach.

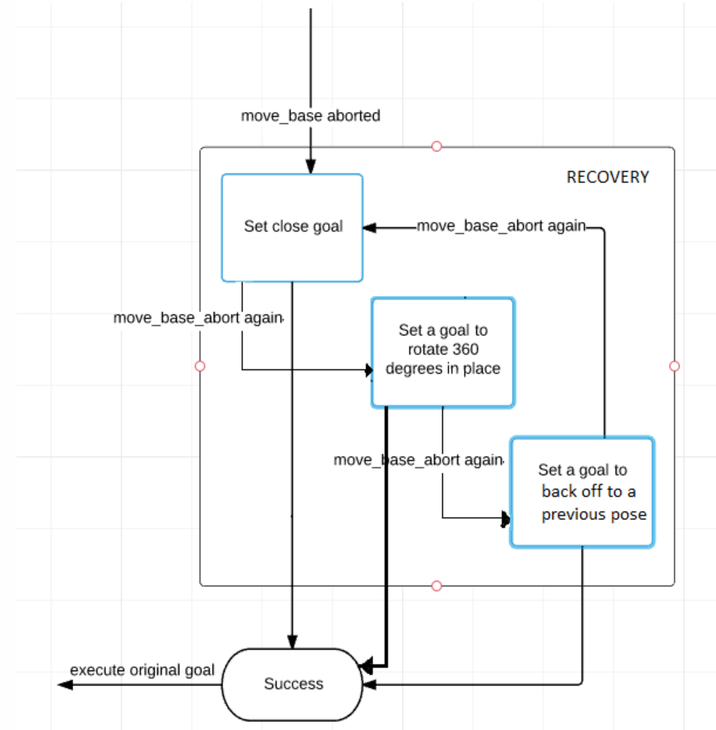


(a) Small Office (b) Large Office

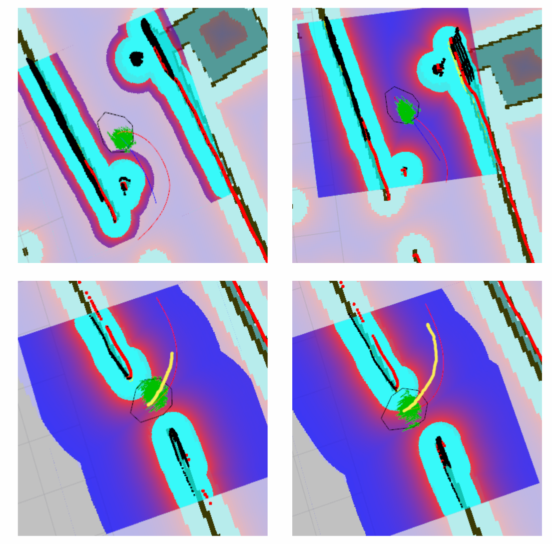
(c) Corridor (d) Doorway

Navigation

- We use the Robotic Operating System (ROS) navigation stack.
- We fine-tuned the parameters for global planner, local planner, and the adaptive Monte-Carlo Localization algorithm (AMCL).
- We designed an FSM to monitor navigation behavior and handle recovery.
- We contributed an extensive guide to navigation algorithm tuning for the ROS community.



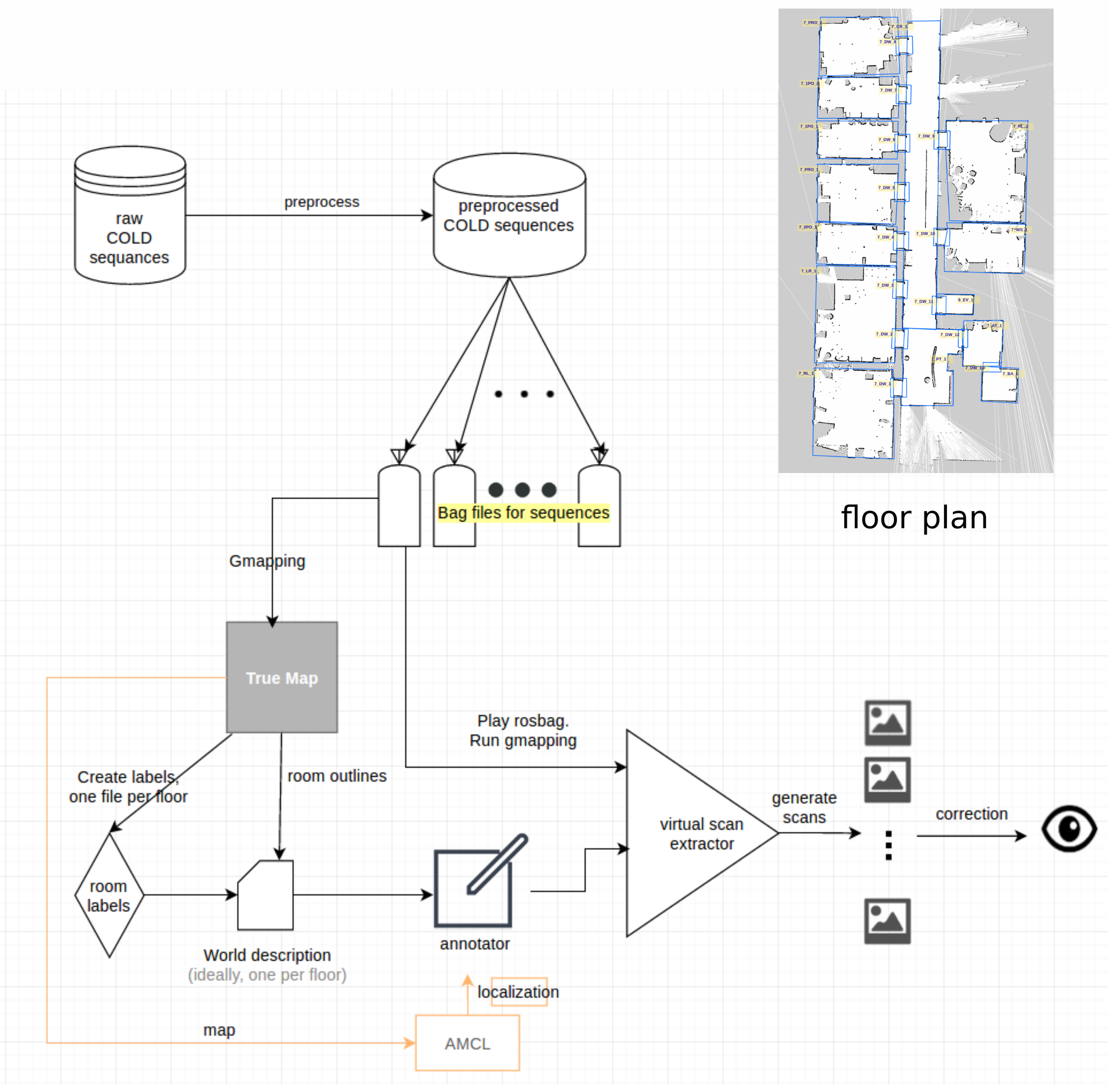
Finite State Machine for Recovery Behaviors



Examples of Parameter-tuning for DWA local planner
Top: costmap inflation; Mid: sim_time

ROS

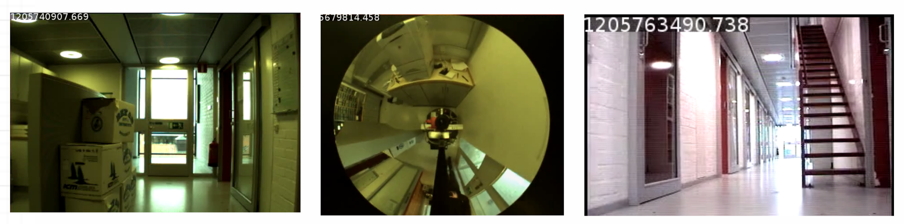
Data Collection Pipeline

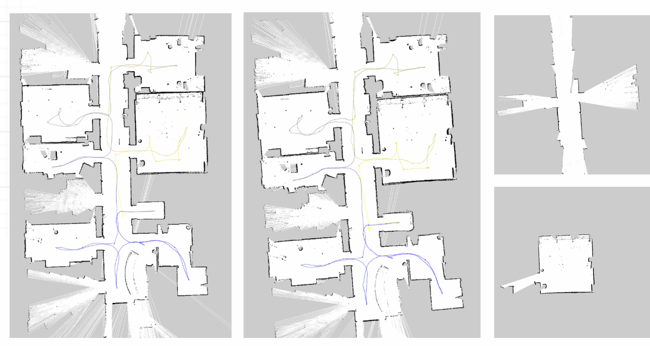


Data-collection pipeline

floor plan

- We designed a data-collection pipeline for collecting annotated localization, SLAM,[1] virtual scan, and RGB data.
- It is capable of collecting data in real time, or from datasets of raw laser scanner and odometry readings.
- We used it to process the COLD database, a large-scale database consisting of raw laser range and visual data. The resulting dataset can be used as a benchmark and for training networks that learn the environment at different scales.



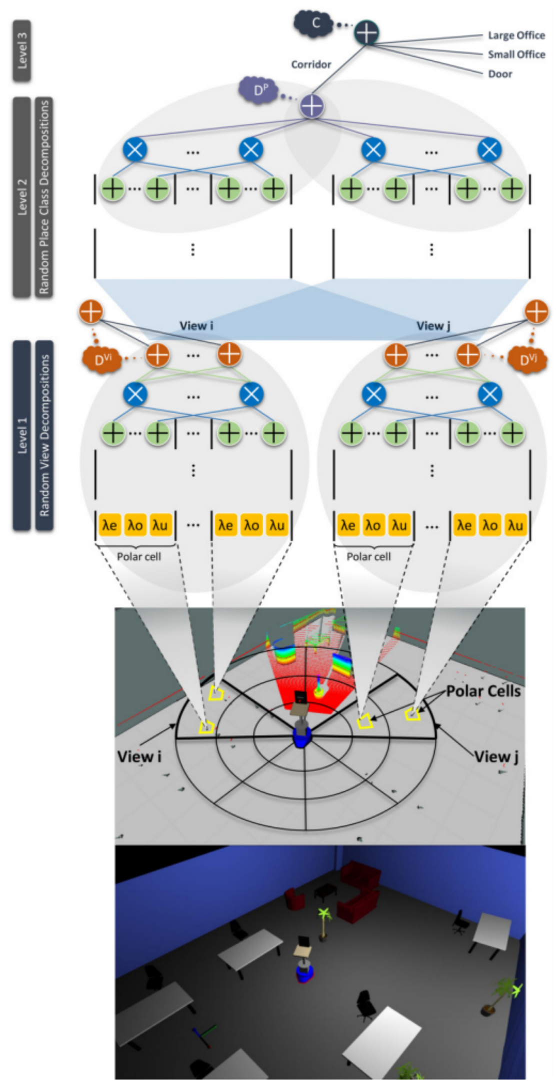


Visualization of data in a sample processed sequence.

[1] A. Pronobis and B. Caputo. "COLD: COsy Localization Database", IJRR, 28(5), May, 2009.

1. SLAM stands for simultaneous localization and mapping.

Applications of the Dataset



An example deep learning model that learns DASH: the Deep Generative Spatial Model [3].

The fully-annotated large dataset supports the learning of various kinds of data-driven models.

For example, the Deep Affordance Spatial Hierarchy (DASH) [2] is a hierarchical spatial representation that was learned using this dataset:

- Laser readings, odometry, and RGB images are input to the perceptual layer.
- The virtual scans are converted to polar occupancy grids used in the peripheral layer.
- The full floor maps and robot poses provide the information for generating topological maps.
- The annotations provide groundtruth human semantics.

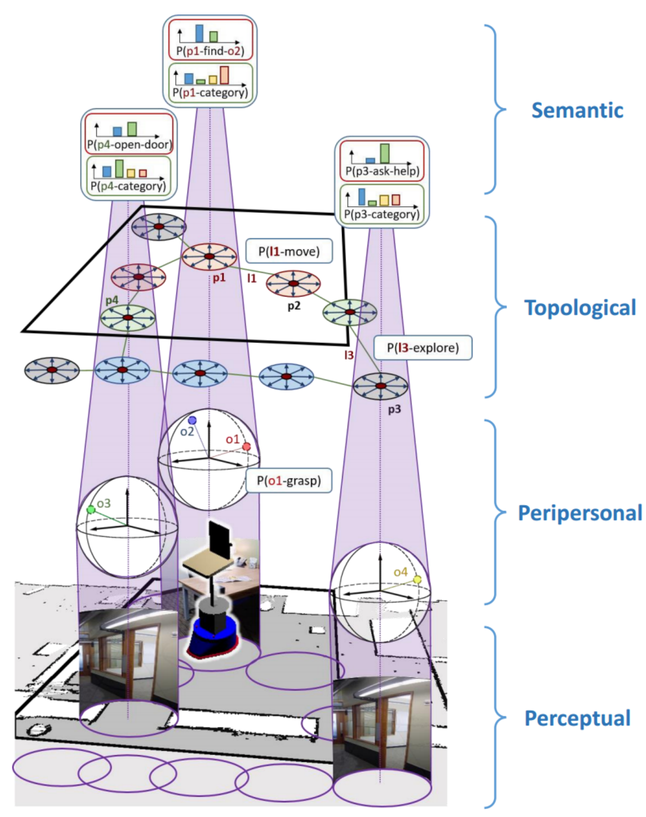


Illustration of DASH

[2] A. Pronobis, F. Riccio, R. P. N. Rao. "Deep Spatial Affordance Hierarchy: Spatial Knowledge Representation for Planning in Large-scale Environments". In ICAPS'17 Workshop on Planning and Robotics.

[3] A. Pronobis, R. P. N. Rao. "Learning Deep Generative Spatial Models for Mobile Robots". arXiv:1610.02627

Conclusion

- We created a large dataset of fully-annotated localization, sensory information, and virtual scans.
- The dataset is expected to be used for a wide range of mobile robotic research.

Future Work

- We plan to use this dataset to train deep learning networks which are designed to model the environment at global scale (e.g. an entire floor).
- We also plan to use the data-collection pipeline to perform large-scale experiments on the deep generative spatial model. Specifically, we investigate the benefits of semi-supervised learning and simulation-based initialization of the model.

Acknowledgements

We sincerely appreciate the Robotics State-Estimation Lab to provide the space and equipments for this research project.

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

*: Kousuke Ariga and Kaiyu Zheng contributed equally to this project.