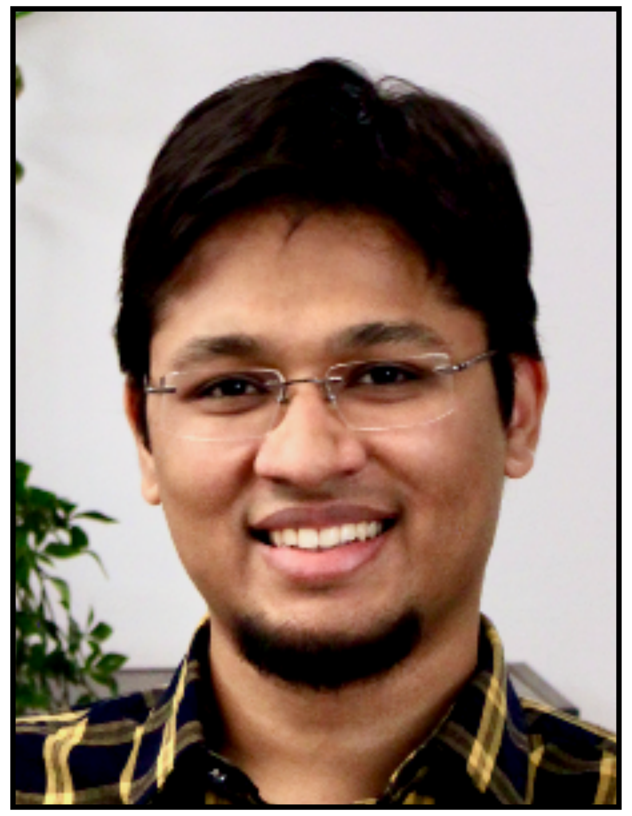
Neural Topological SLAM for Visual Navigation

CVPR-2020

Webpage: https://devendrachaplot.github.io/projects/Neural-Topological-SLAM




Devendra Singh Chaplot

Ruslan Salakhutdinov

Abhinav Gupta

Saurabh Gupta

Carnegie Mellon University

facebook Artificial Intelligence Research

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN

Semantic Priors and Common-Sense

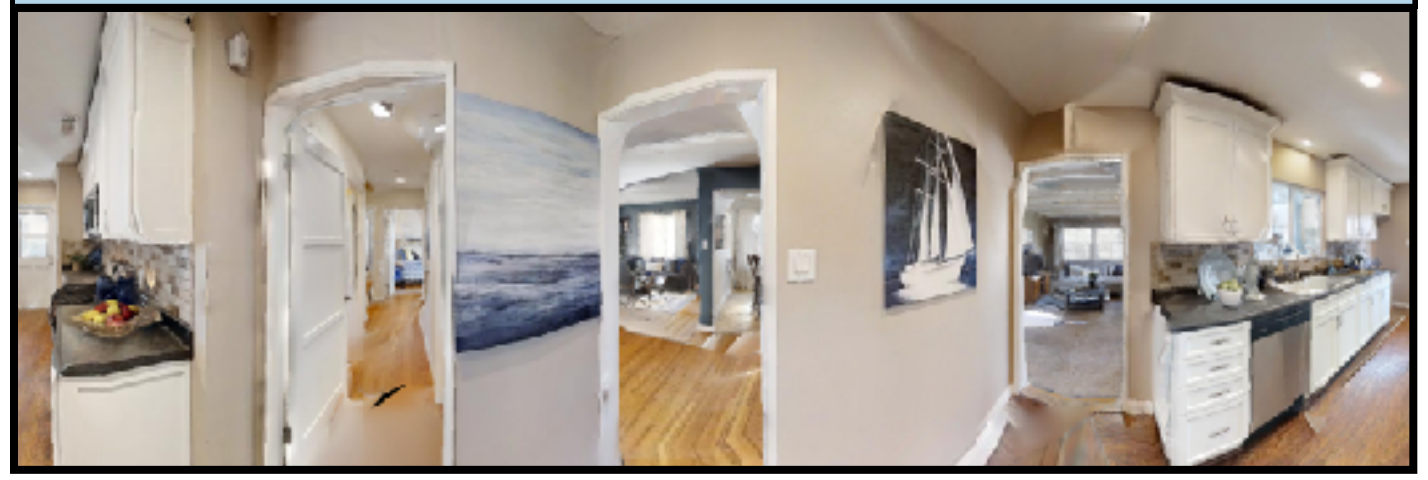
- Humans use semantic priors and common-sense to explore and navigate everyday
- Most navigation algorithms struggle to do so

Semantic Priors and Common-Sense

- Humans use semantic priors and common-sense to explore and navigate everyday
- Most navigation algorithms struggle to do so

Image Goal Task

Source Image (I_S)

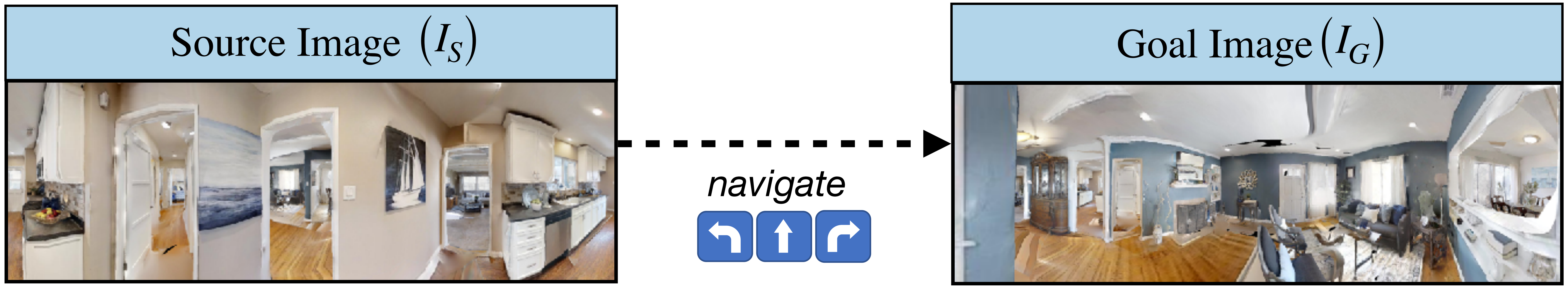


Goal Image (I_G)

Image Goal Task

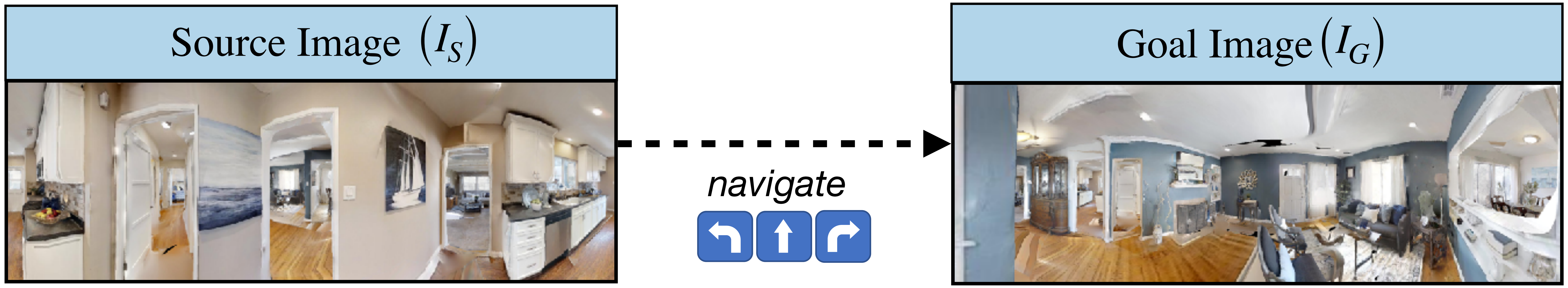
- Agent observations are panoramic images

Image Goal Task



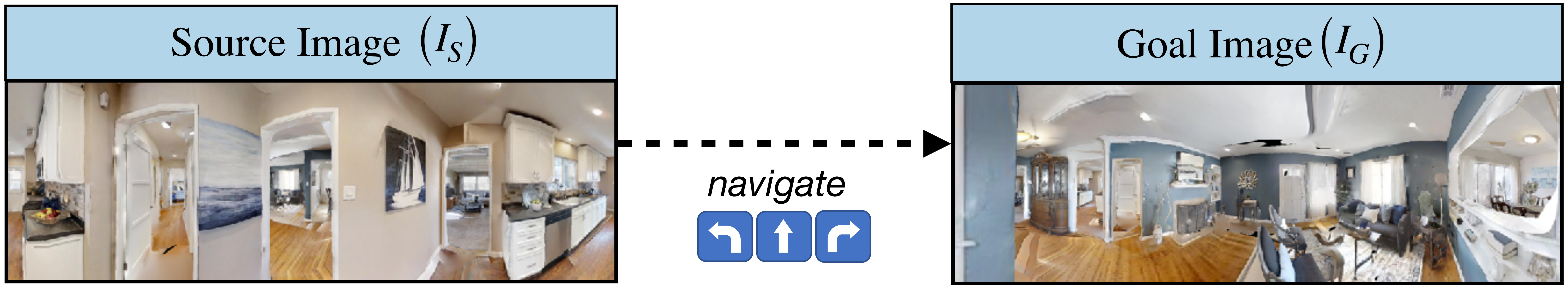
- Agent observations are panoramic images
- Take actions to navigate to the goal location

Image Goal Task



- Agent observations are panoramic images
- Take actions to navigate to the goal location
- Take the `stop' action at the goal location

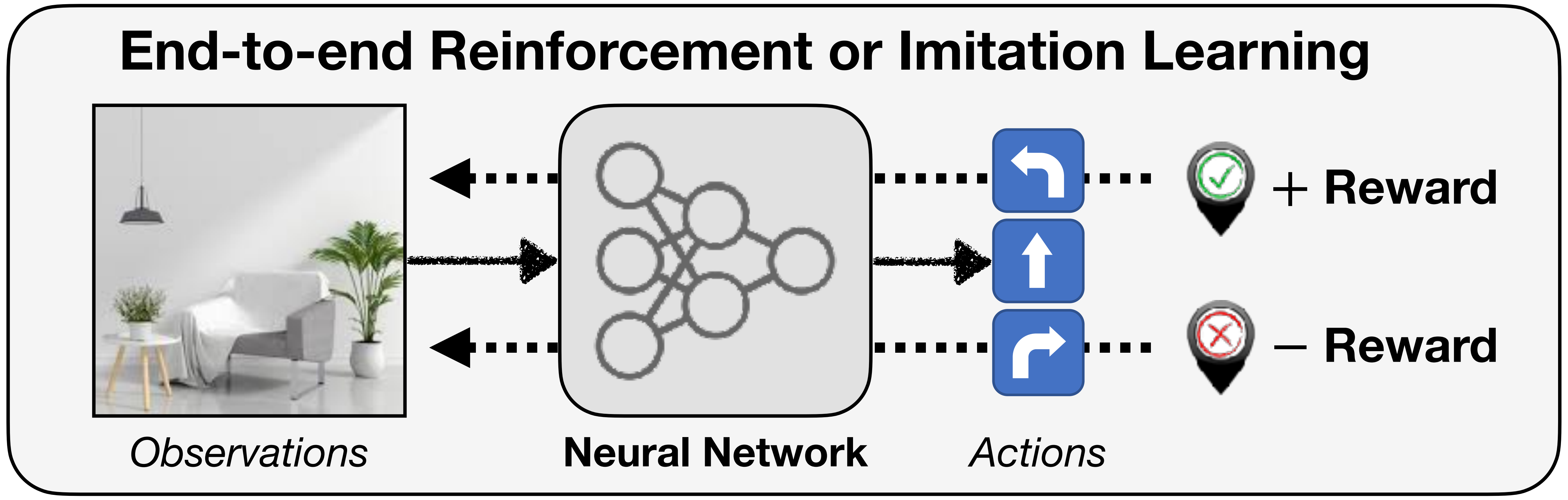
Image Goal Task



- Agent observations are panoramic images
- Take actions to navigate to the goal location
- Take the `stop' action at the goal location
- Sequential goals

Prior work

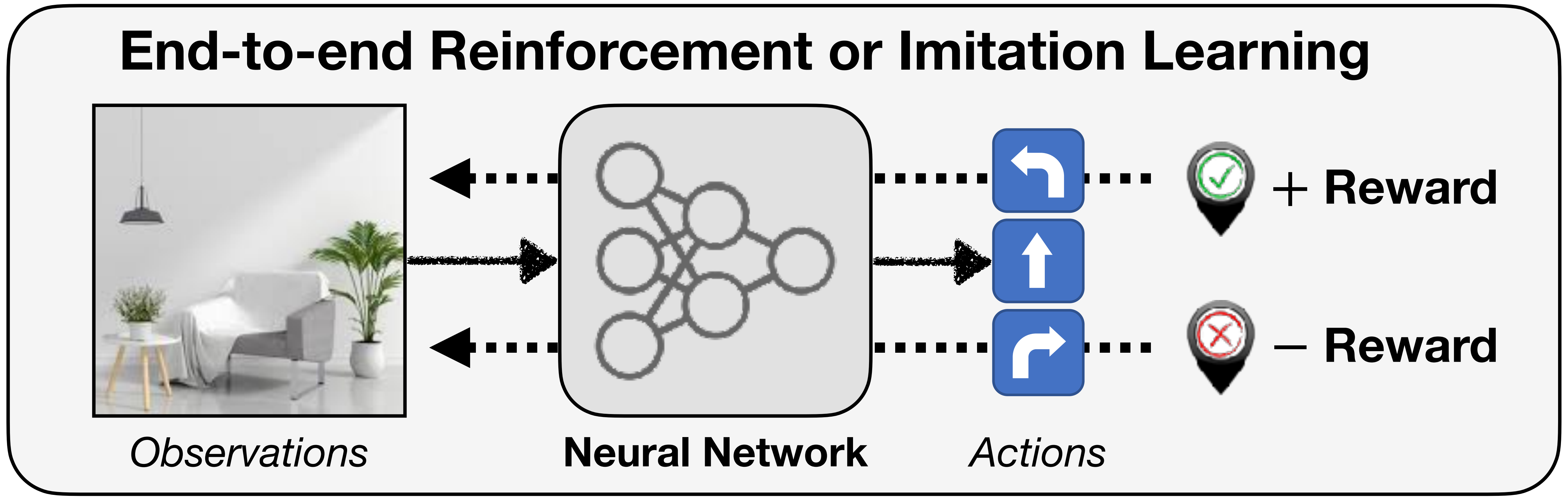
Prior work

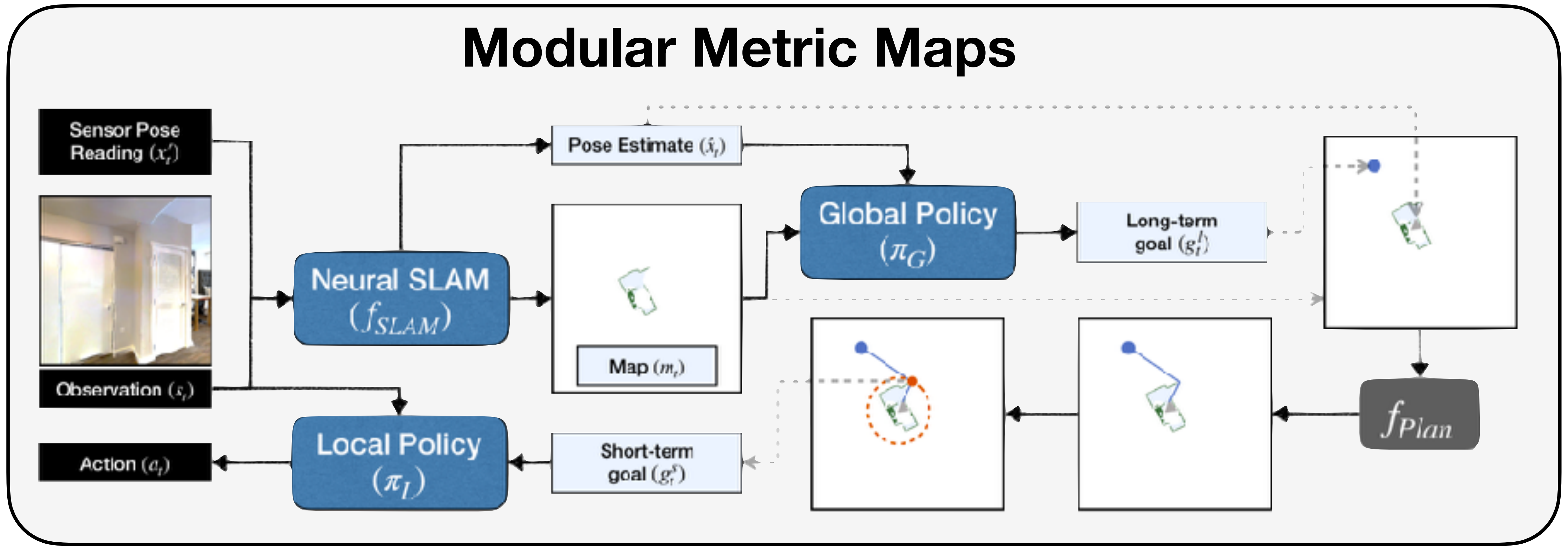


End-to-end Learning

- High sample complexity
- Ineffective in large environments

Prior work





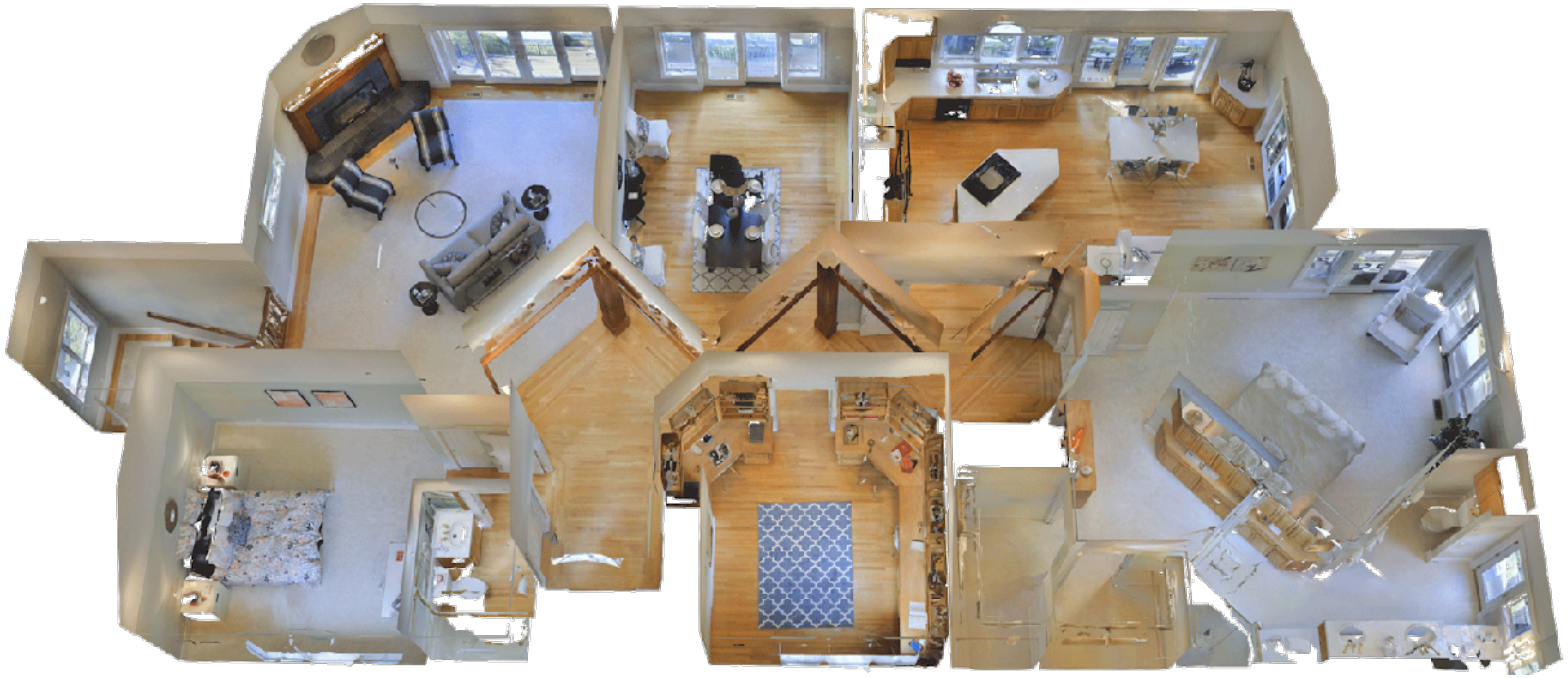
End-to-end Learning

- High sample complexity
- Ineffective in large environments

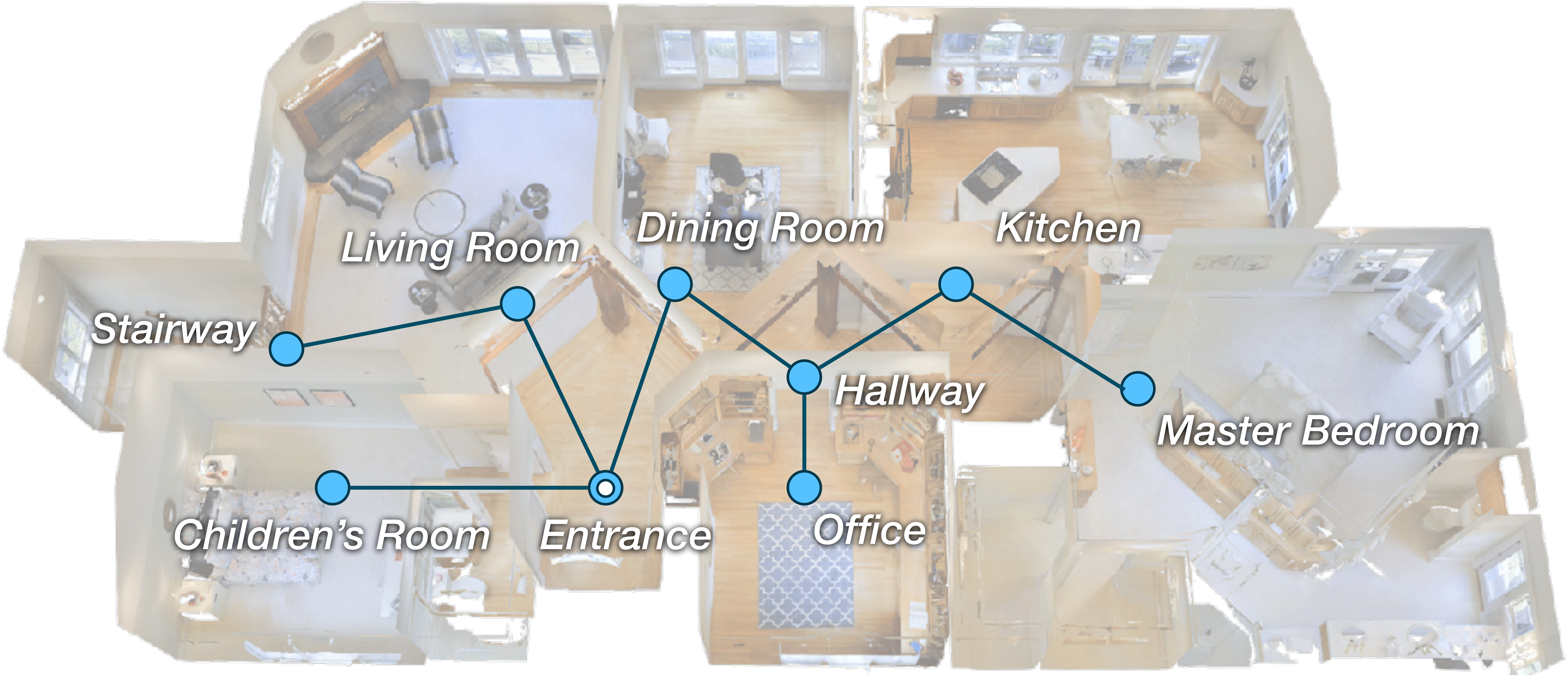
Modular Metric Maps

- Can not learn semantic priors
- Pose error accumulation

Topological Maps

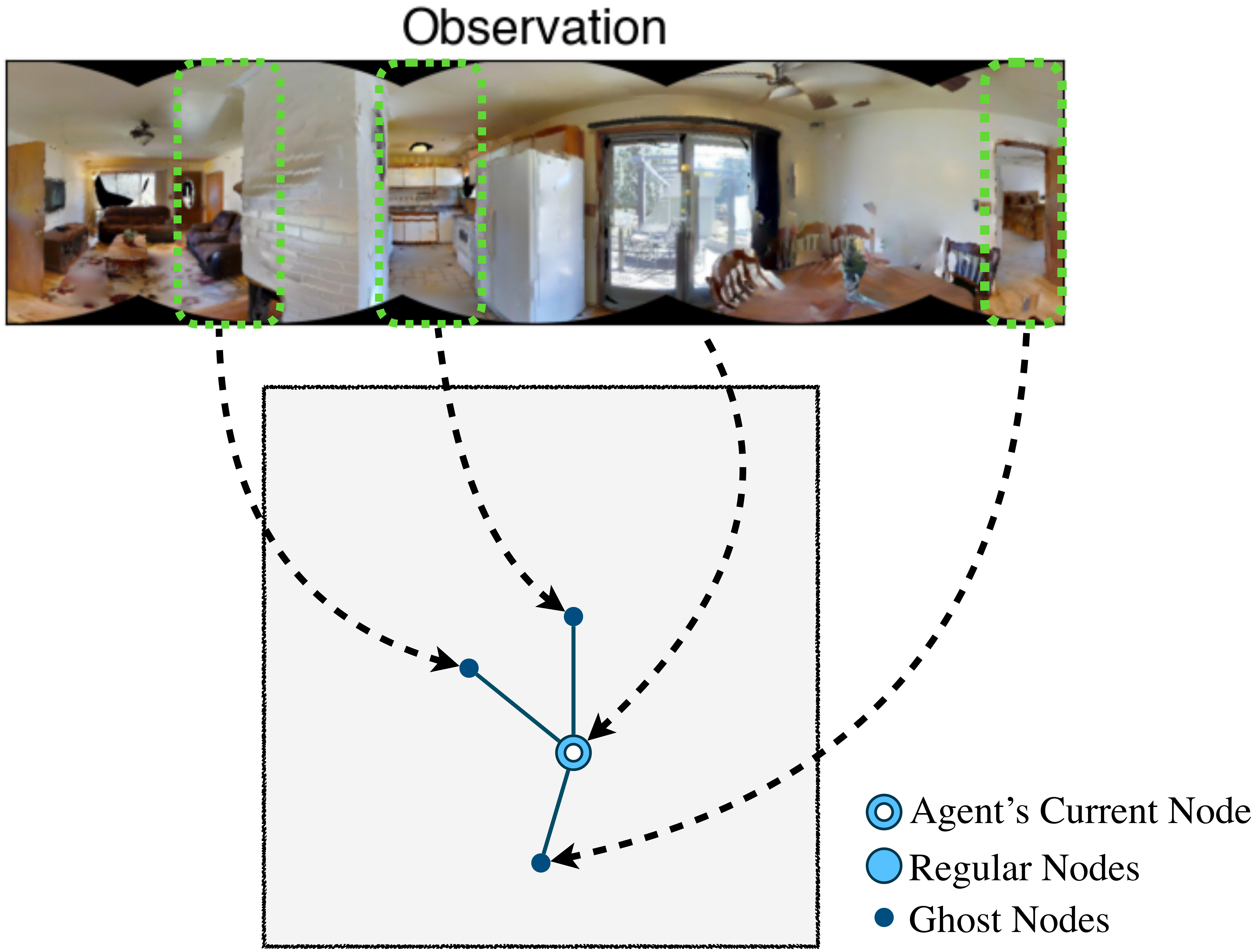


Topological Maps



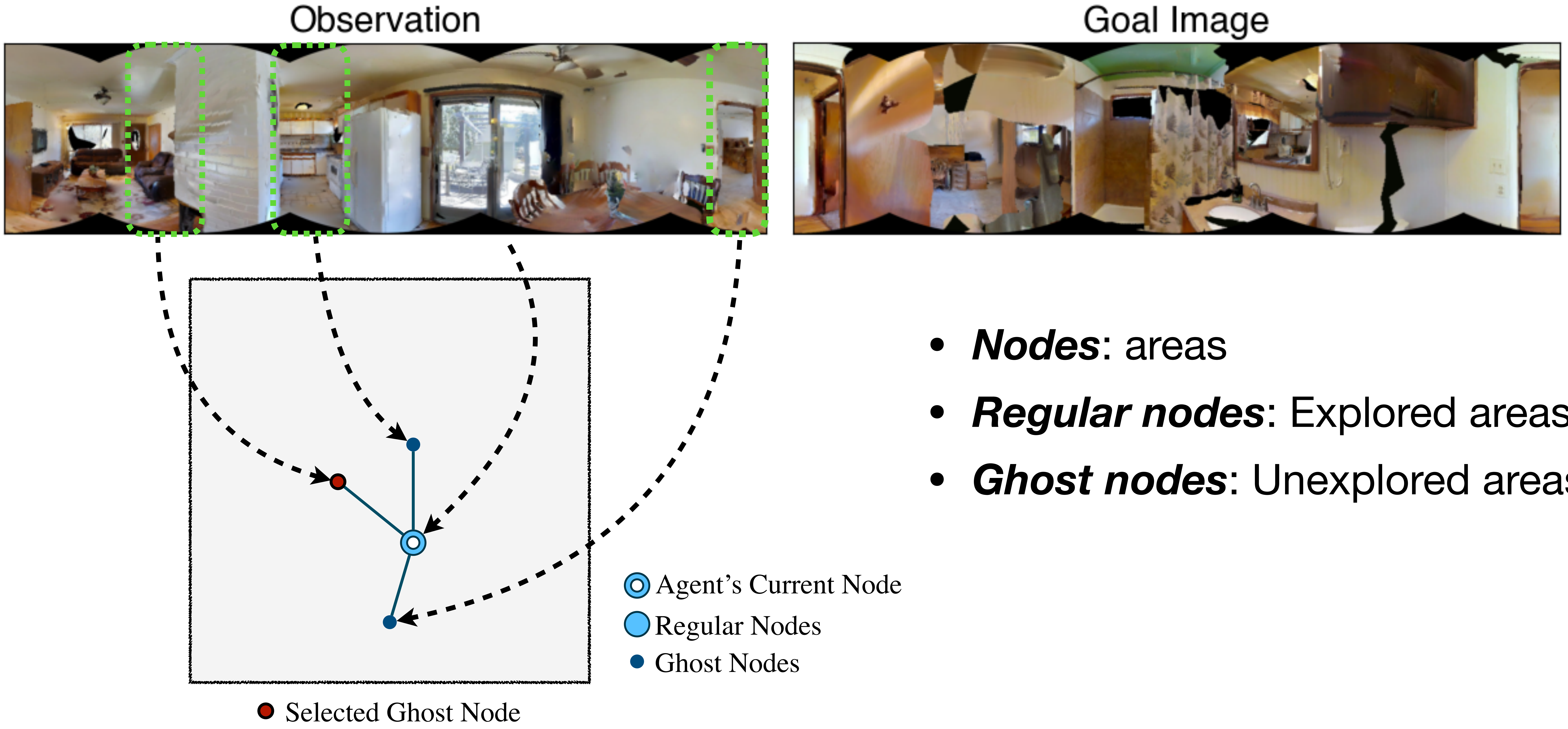
Topological Graph Representation

Topological Graph Representation



- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas

Topological Graph Representation



- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas

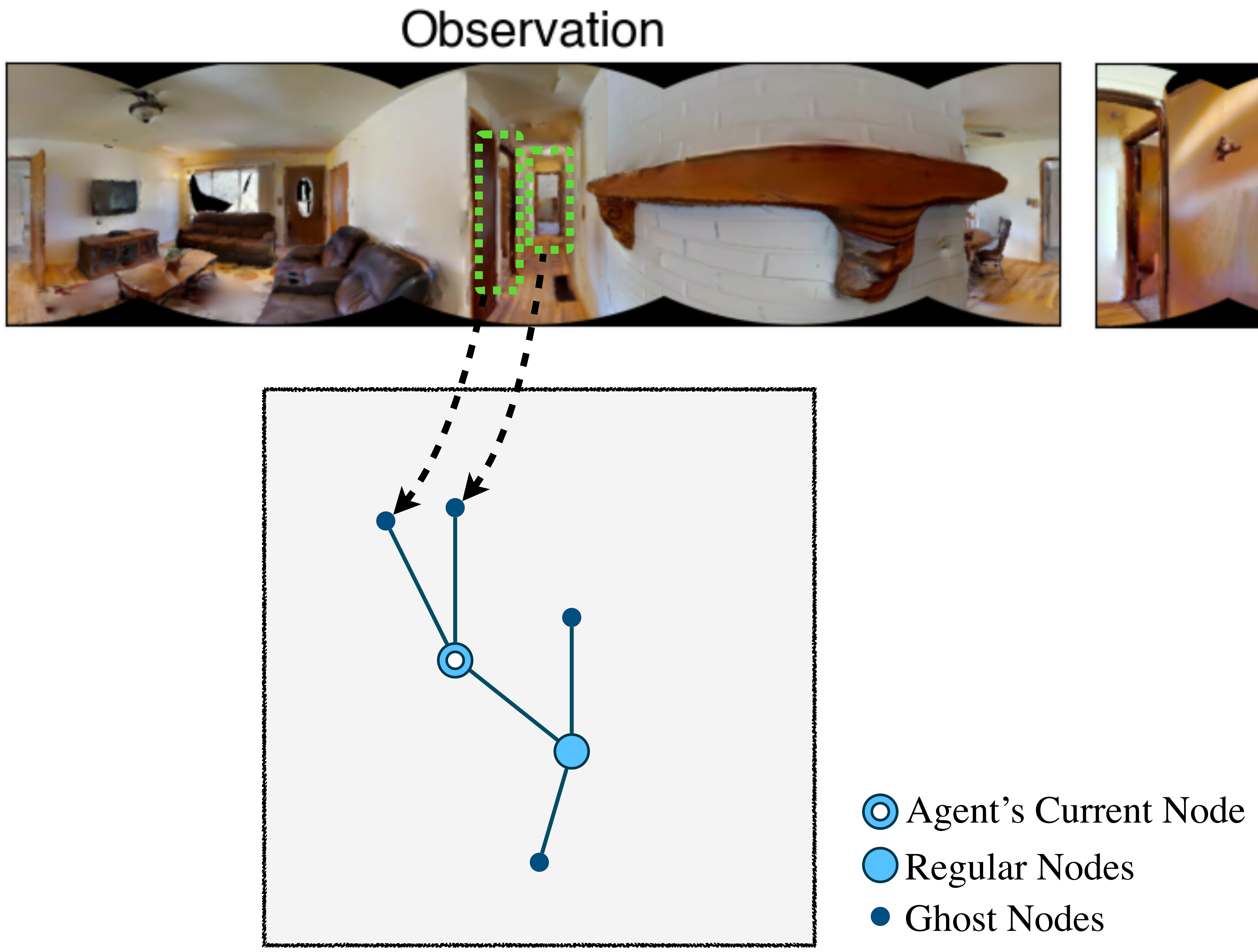
Topological Graph Representation

Observation

Goal Image

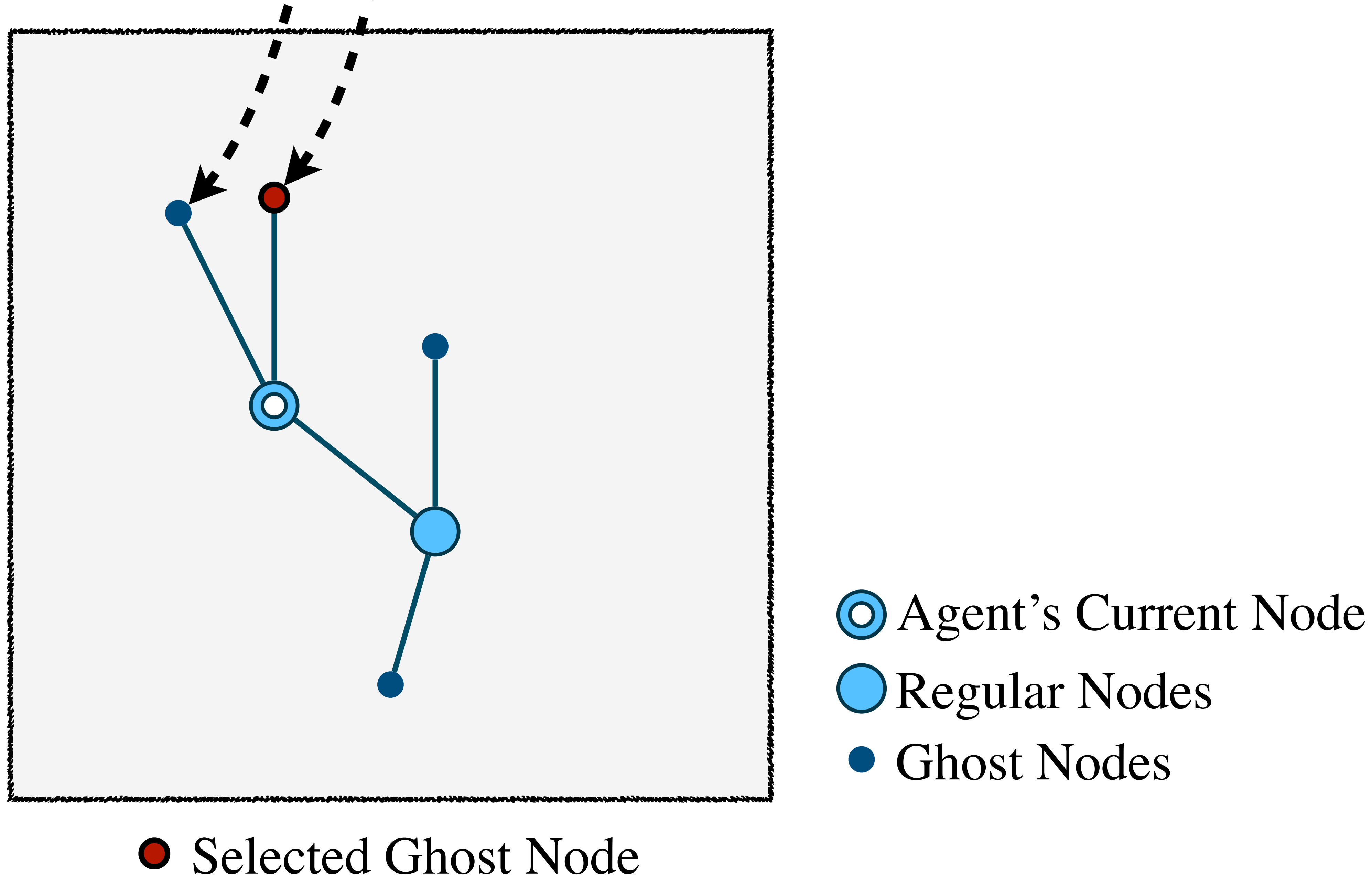
- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas

Topological Graph Representation



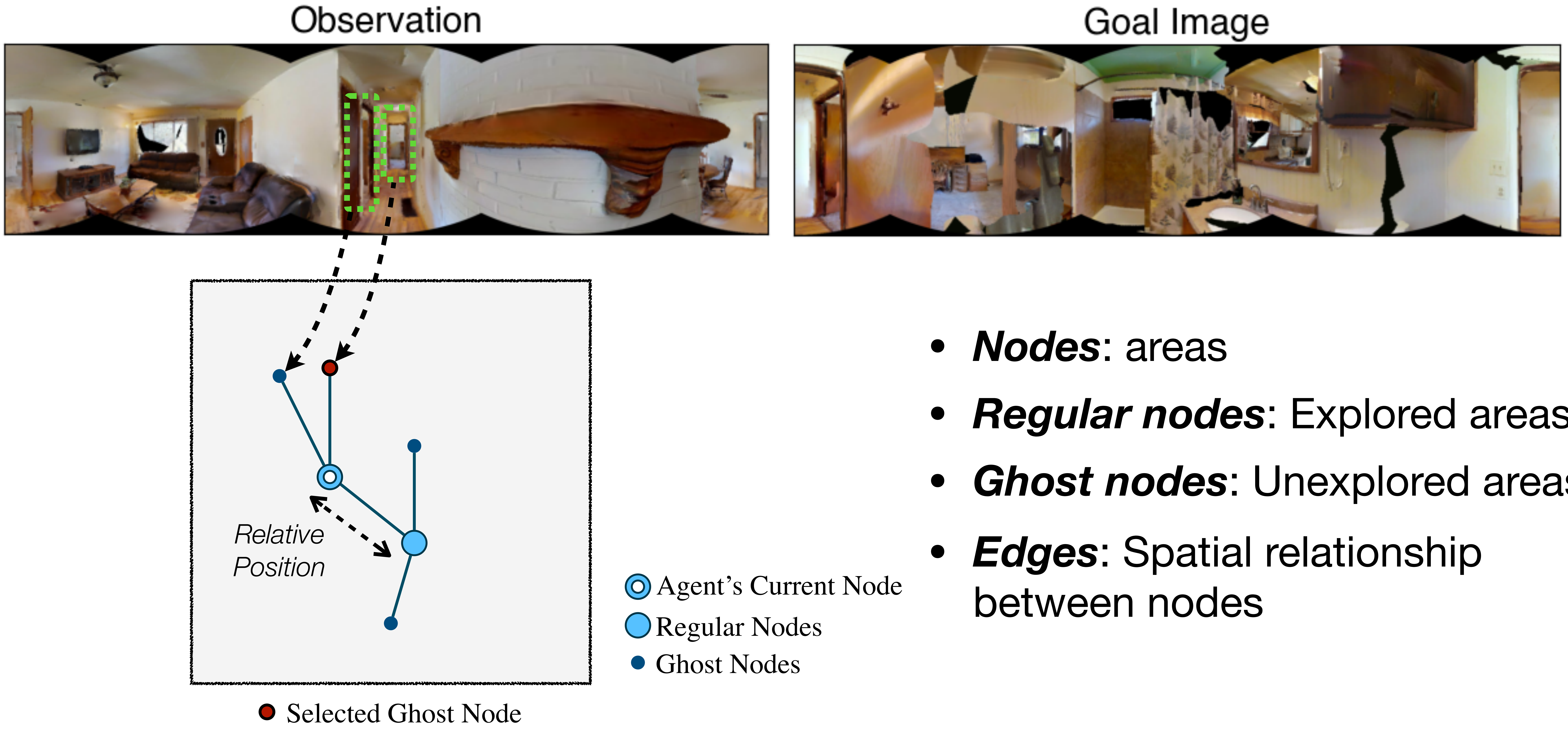
- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas

Topological Graph Representation



- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas

Topological Graph Representation



- ***Nodes***: areas
- ***Regular nodes***: Explored areas
- ***Ghost nodes***: Unexplored areas
- ***Edges***: Spatial relationship between nodes

Four learnable functions

Four learnable functions

$\mathcal{F}_G(I_1)$ = Geometric Prediction: Free directions

$\mathcal{F}_S(I_1, I_2)$ = Semantic Prediction: Closeness to target

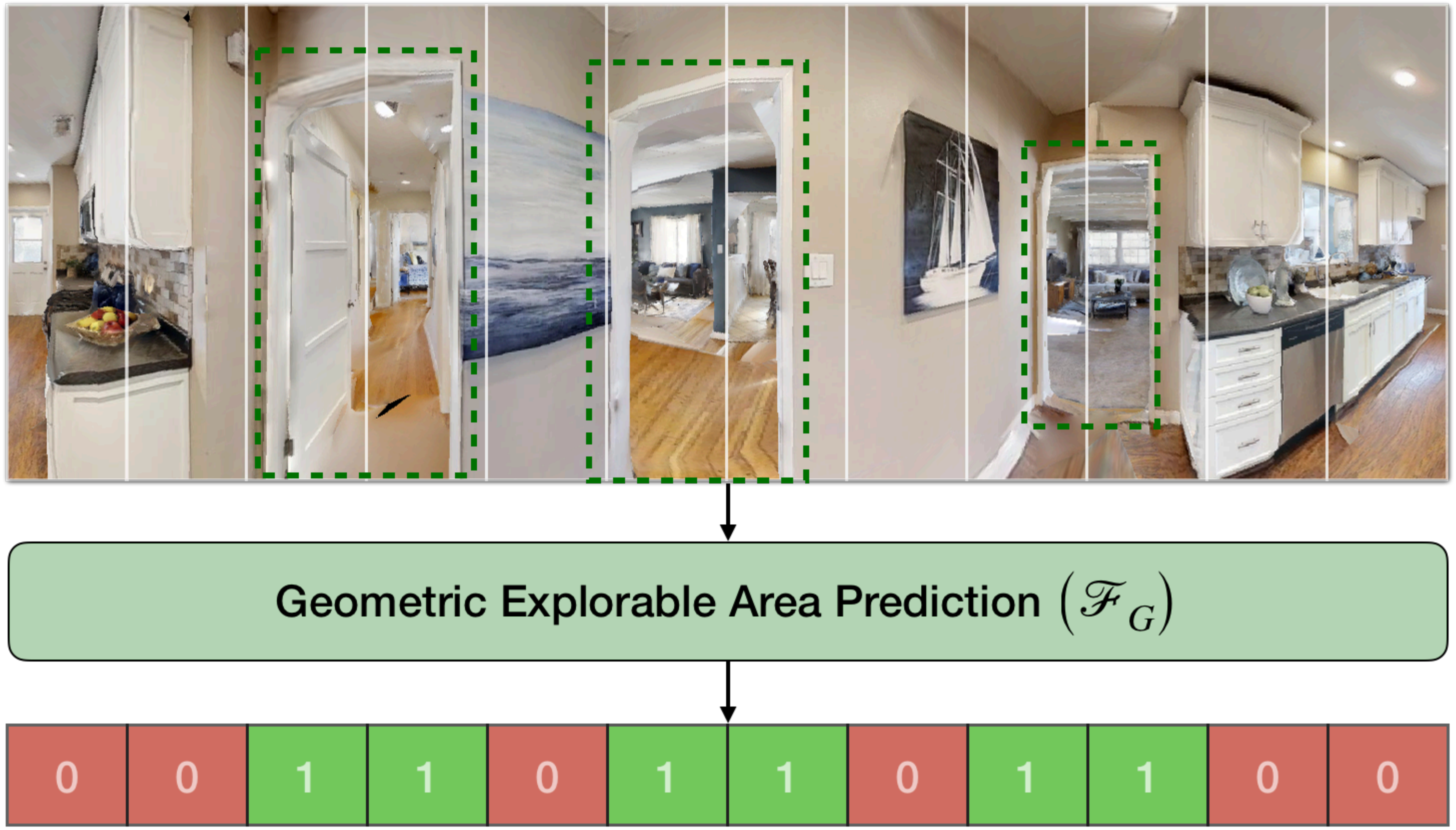
$\mathcal{F}_L(I_1, I_2)$ = Localization

$\mathcal{F}_R(I_1, I_2)$ = Relative Pose Prediction

Geometric Prediction

Geometric Prediction

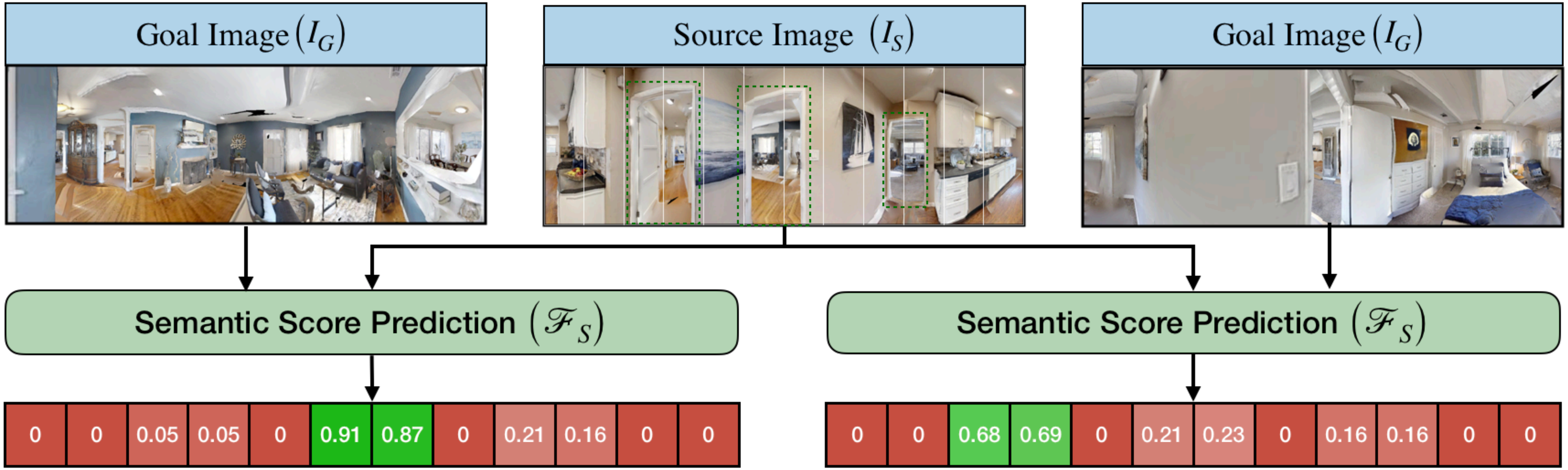
$\mathcal{F}_G(I_1)$ = Geometric Prediction: Free directions



Semantic Prediction

Semantic Prediction

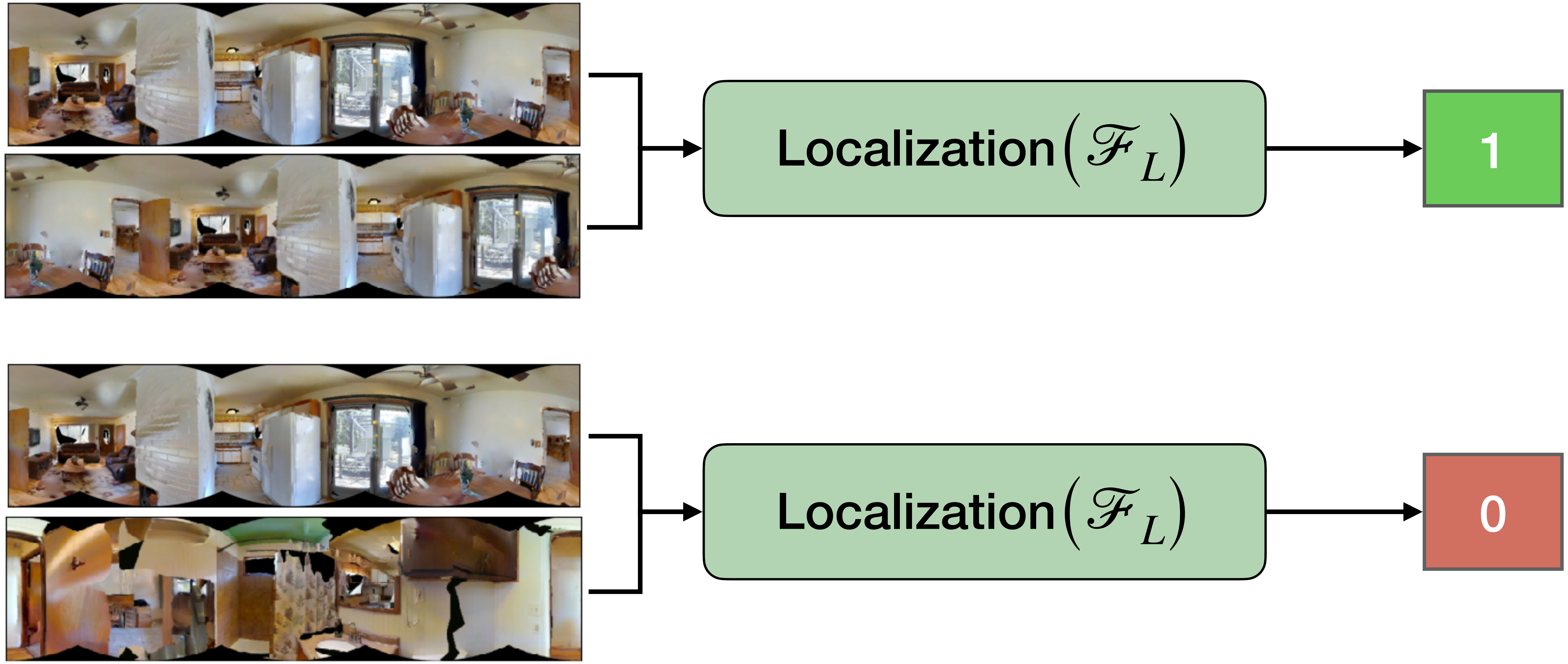
$\mathscr{F}_S(I_1, I_2)$ = Semantic Prediction: Closeness to target



Localization

Localization

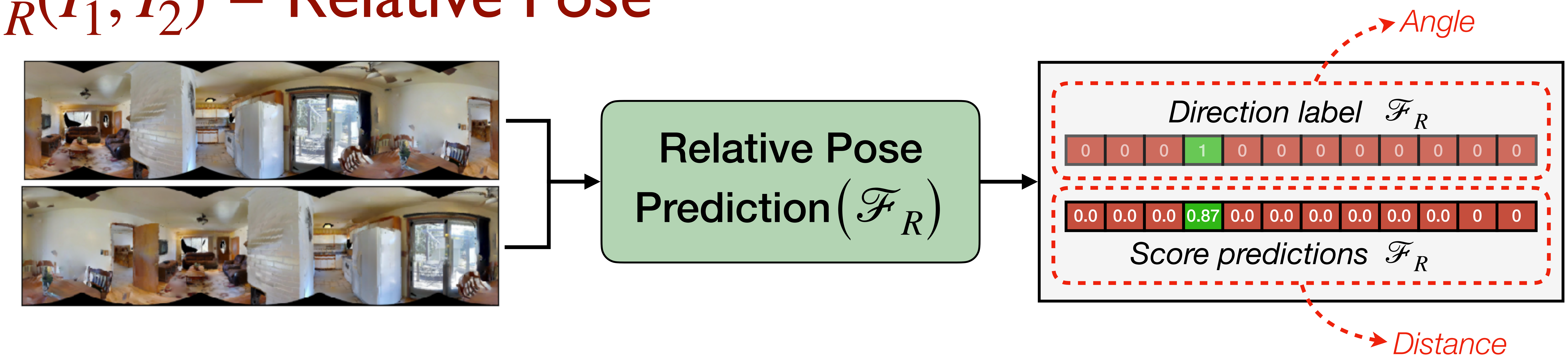
$\mathcal{F}_L(I_1, I_2)$ = Localization



Relative Pose Prediction

Relative Pose Prediction

$\mathscr{F}_R(I_1, I_2)$ = Relative Pose



Neural Topological SLAM

Neural Topological SLAM

$\mathcal{F}_G(I_1)$ = Geometric Prediction: Free directions

$\mathcal{F}_S(I_1, I_2)$ = Semantic Prediction: Closeness to target

$\mathcal{F}_L(I_1, I_2)$ = Localization

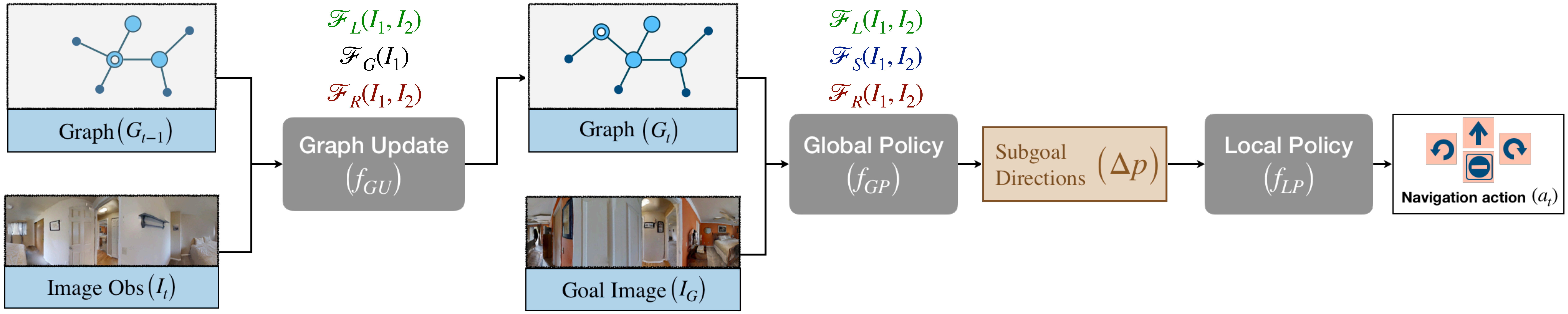
$\mathcal{F}_R(I_1, I_2)$ = Relative Pose Prediction

Neural Topological SLAM

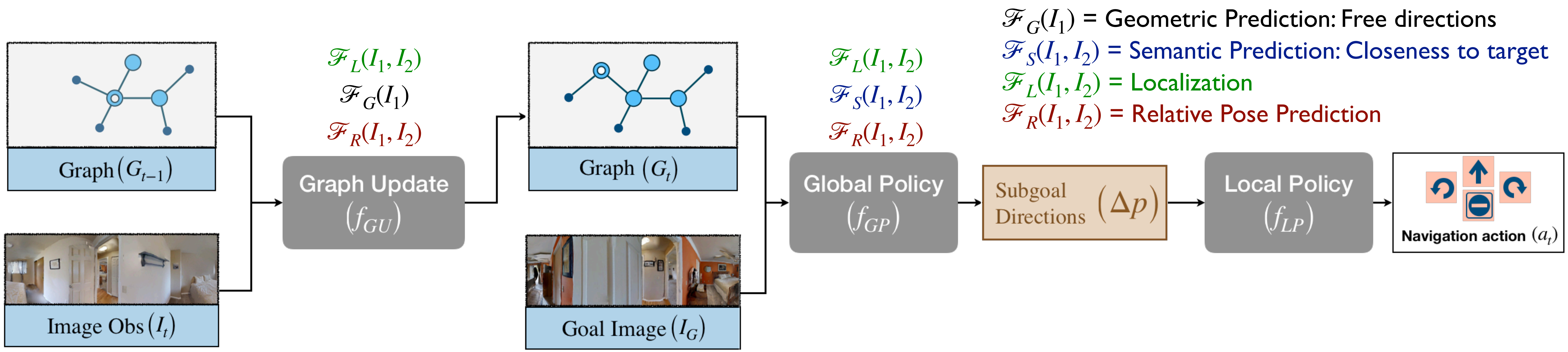
$\mathcal{F}_G(I_1)$ = Geometric Prediction: Free directions
$\mathcal{F}_S(I_1, I_2)$ = Semantic Prediction: Closeness to target
$\mathcal{F}_L(I_1, I_2)$ = Localization
$\mathcal{F}_R(I_1, I_2)$ = Relative Pose Prediction

Neural Topological SLAM

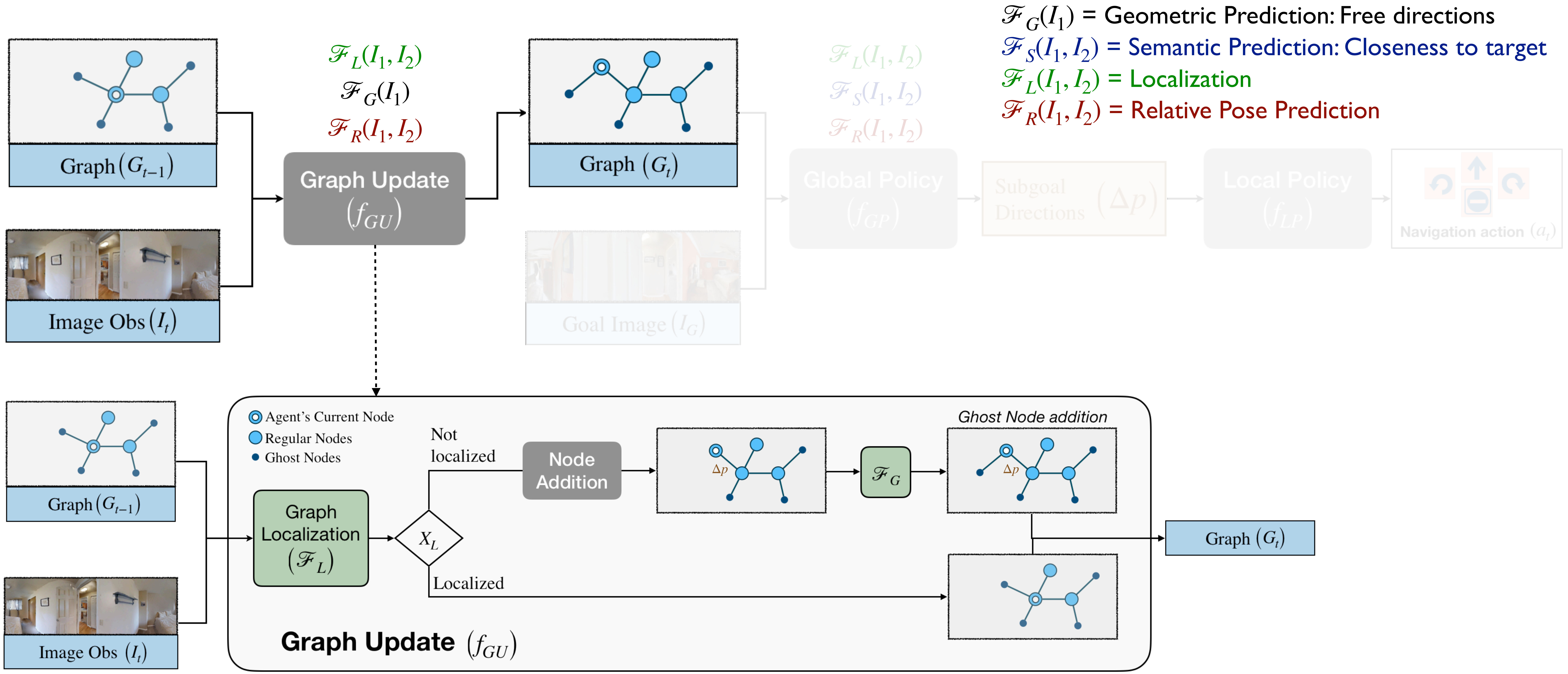
$\mathcal{F}_G(I_1)$ = Geometric Prediction: Free directions
$\mathcal{F}_S(I_1, I_2)$ = Semantic Prediction: Closeness to target
$\mathcal{F}_L(I_1, I_2)$ = Localization
$\mathcal{F}_R(I_1, I_2)$ = Relative Pose Prediction



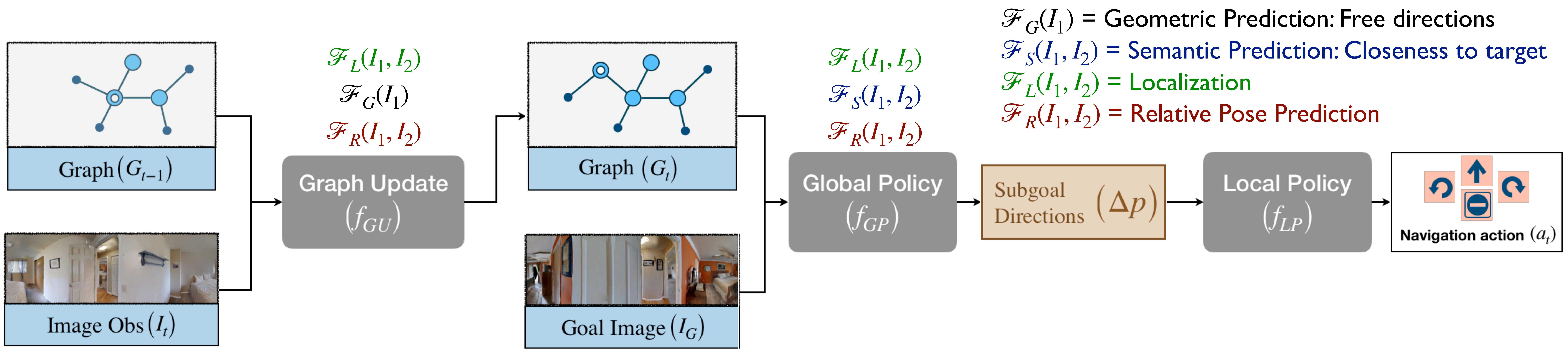
Neural Topological SLAM



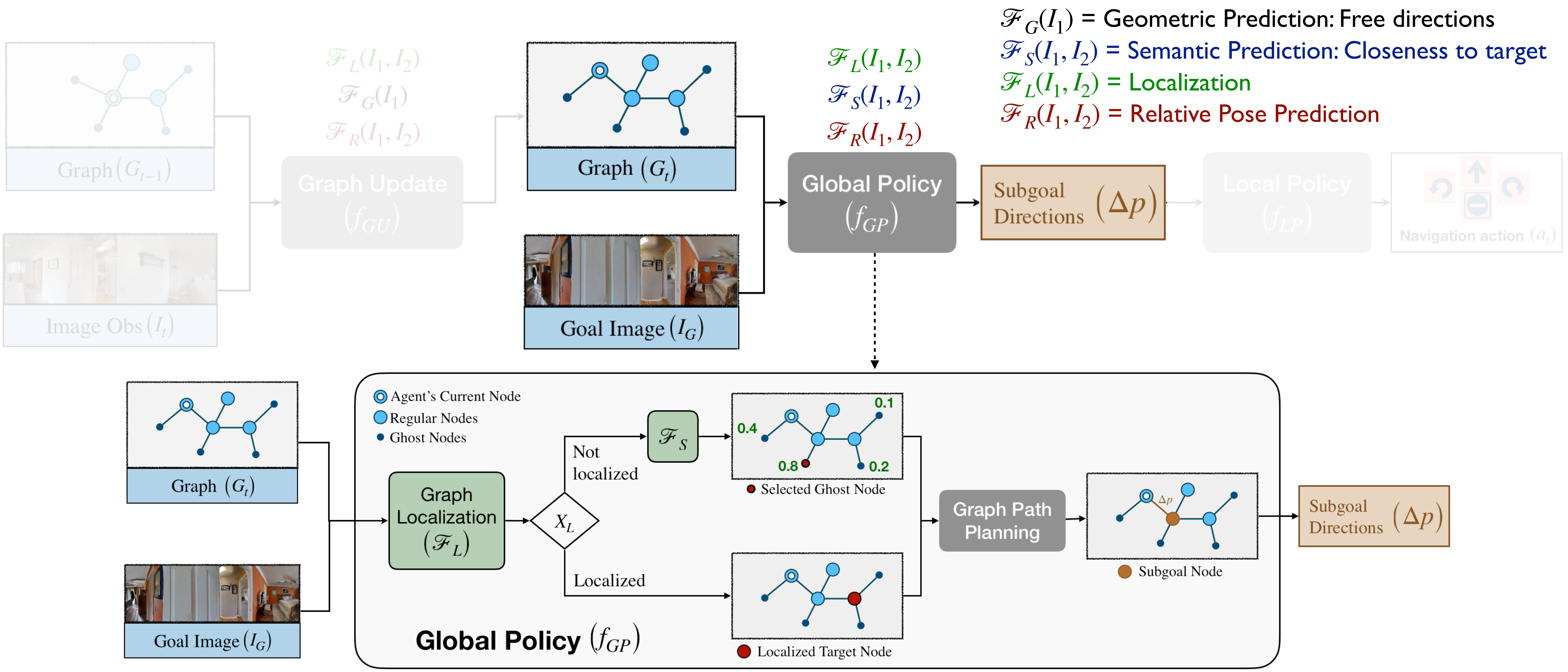
Neural Topological SLAM



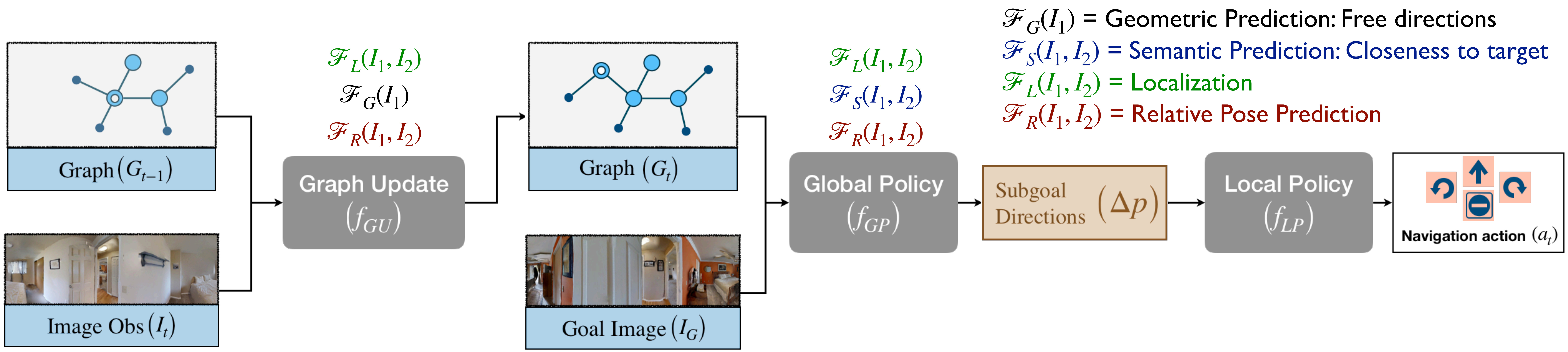
Neural Topological SLAM



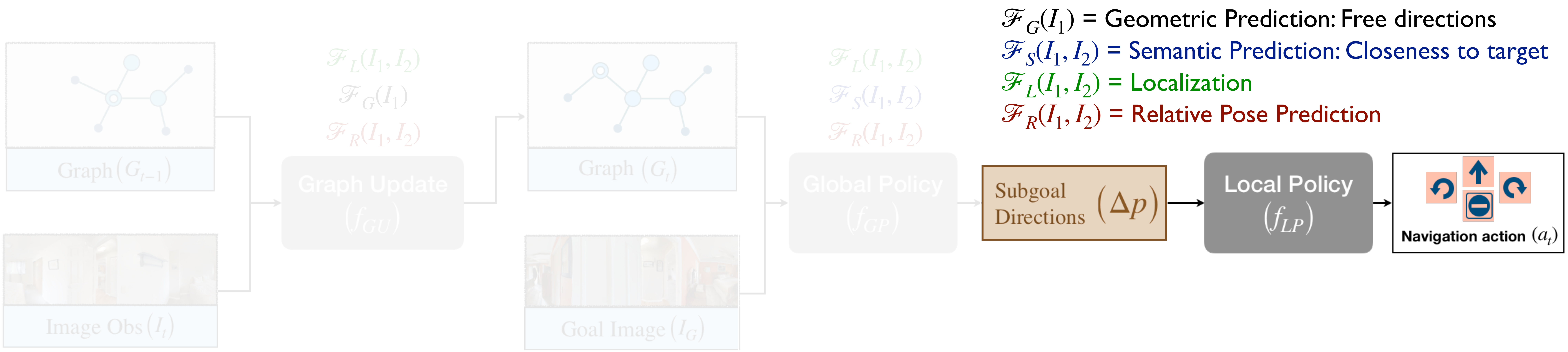
Neural Topological SLAM



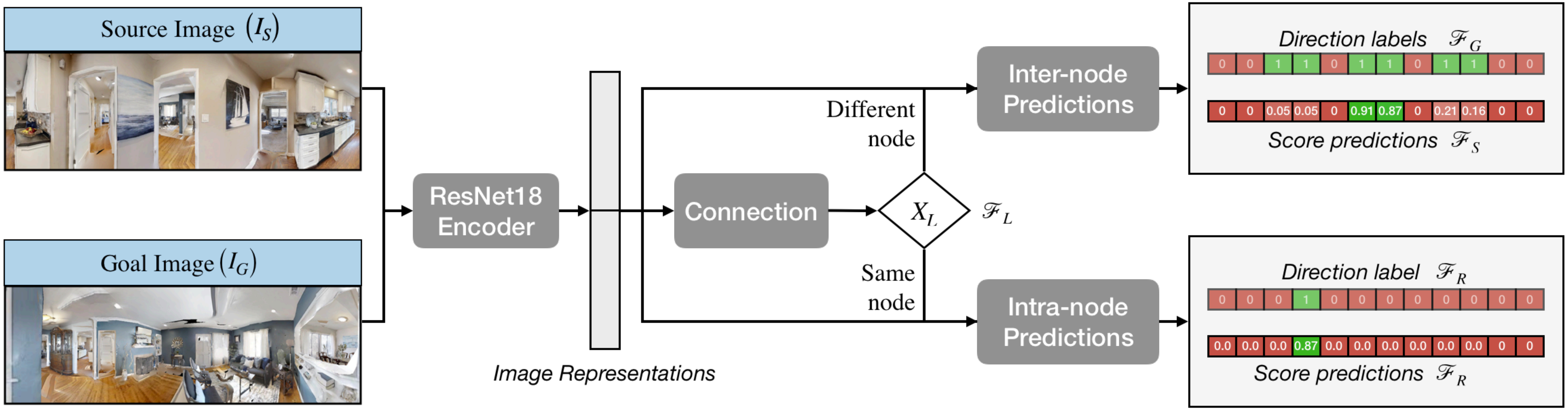
Neural Topological SLAM



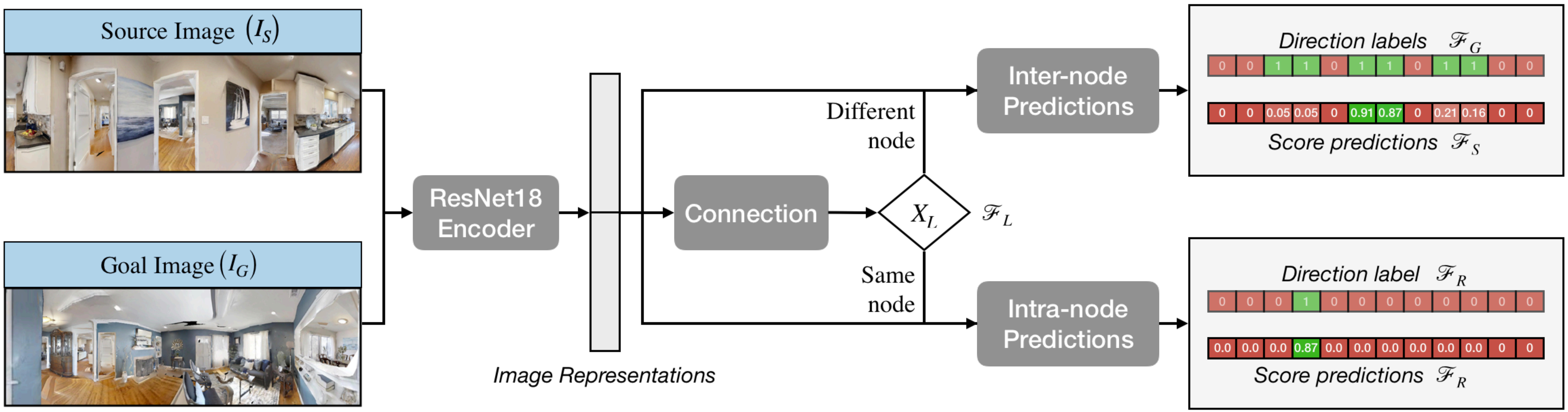
Neural Topological SLAM



Single supervised learning model



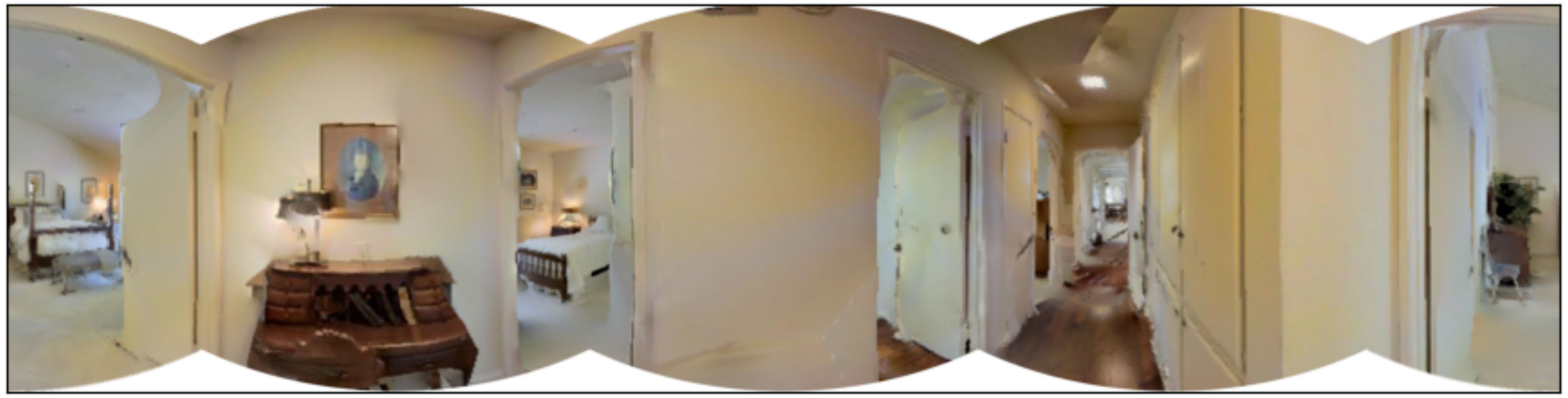
Single supervised learning model



- No reinforcement learning, no interaction needed
- Can be trained completely with static data

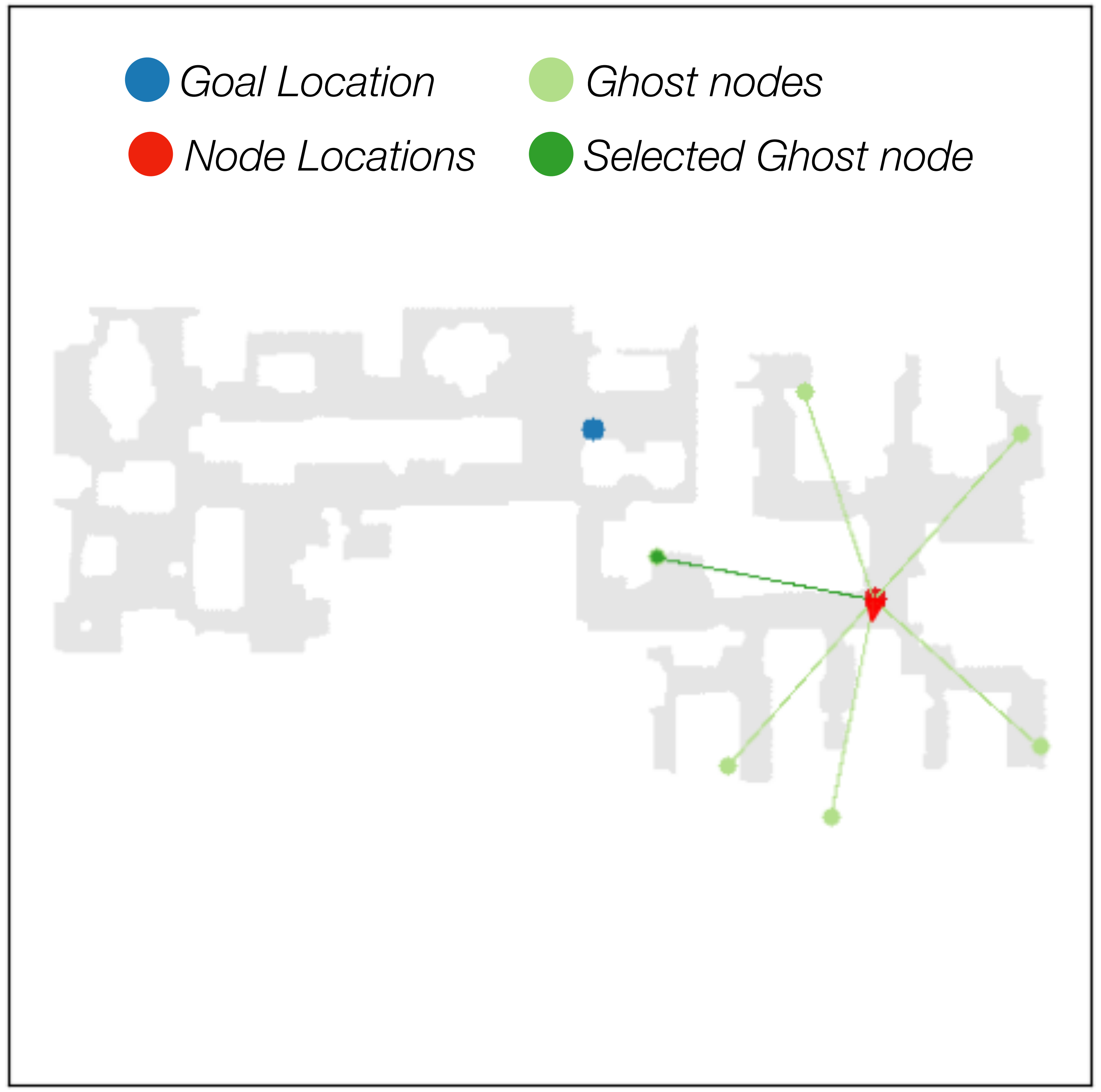
Demo video

Observation

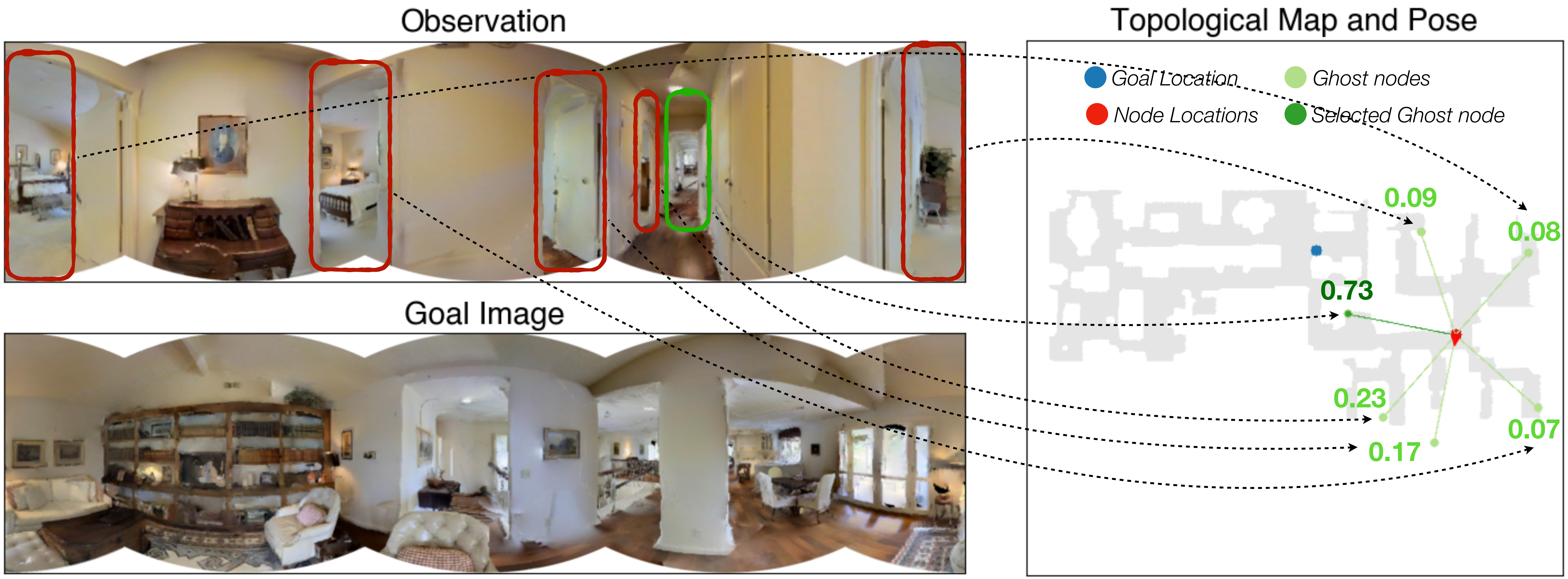


Goal Image

Topological Map and Pose



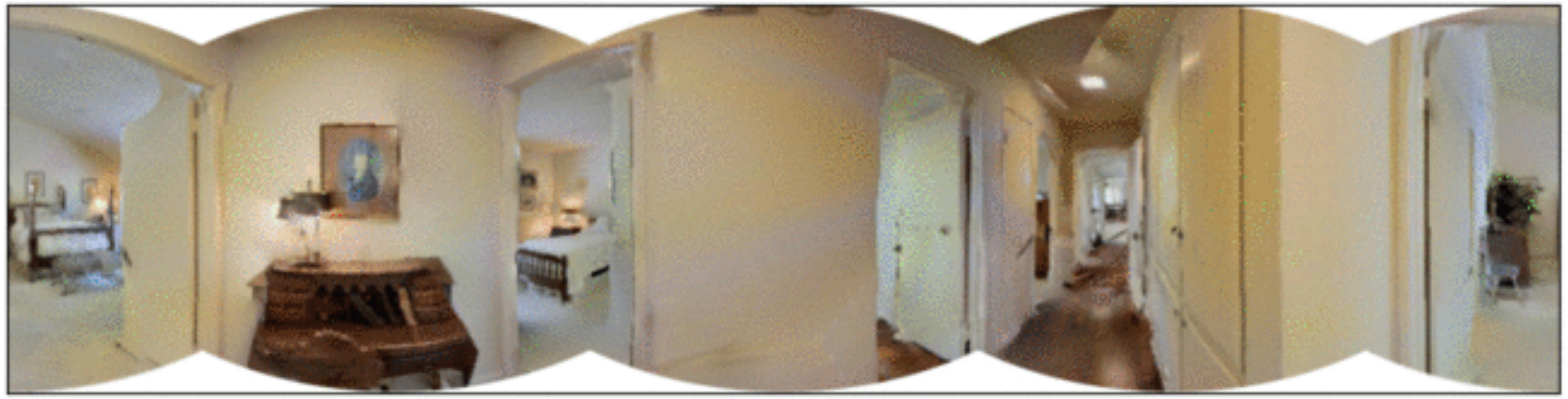
Demo video



Observation
Goal Image
Topological Map and Pose
Goal Location
Ghost nodes
Node Locations
Selected Ghost node
0.09
0.08
0.73
0.23
0.07
0.17

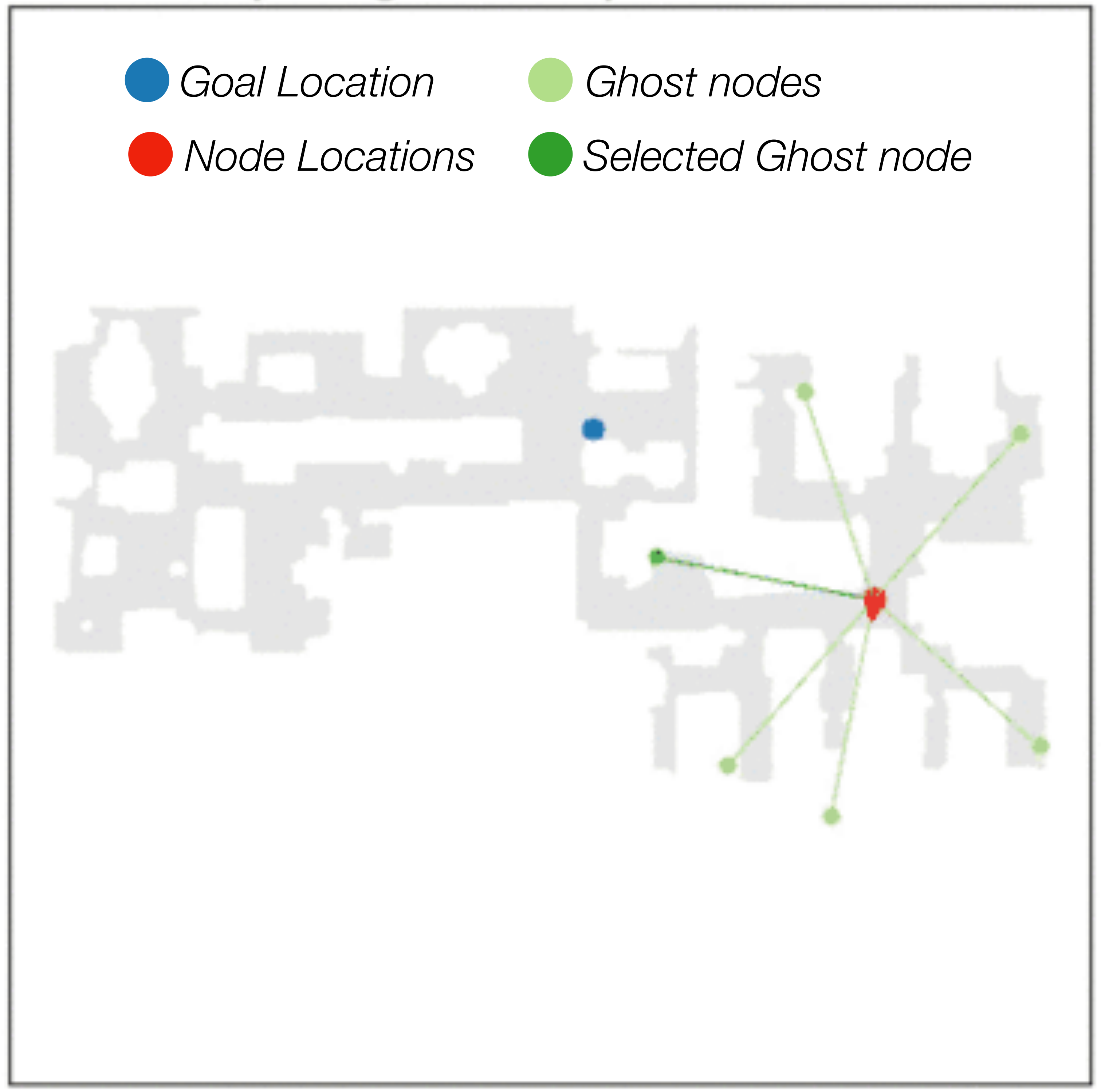
Demo video

Observation



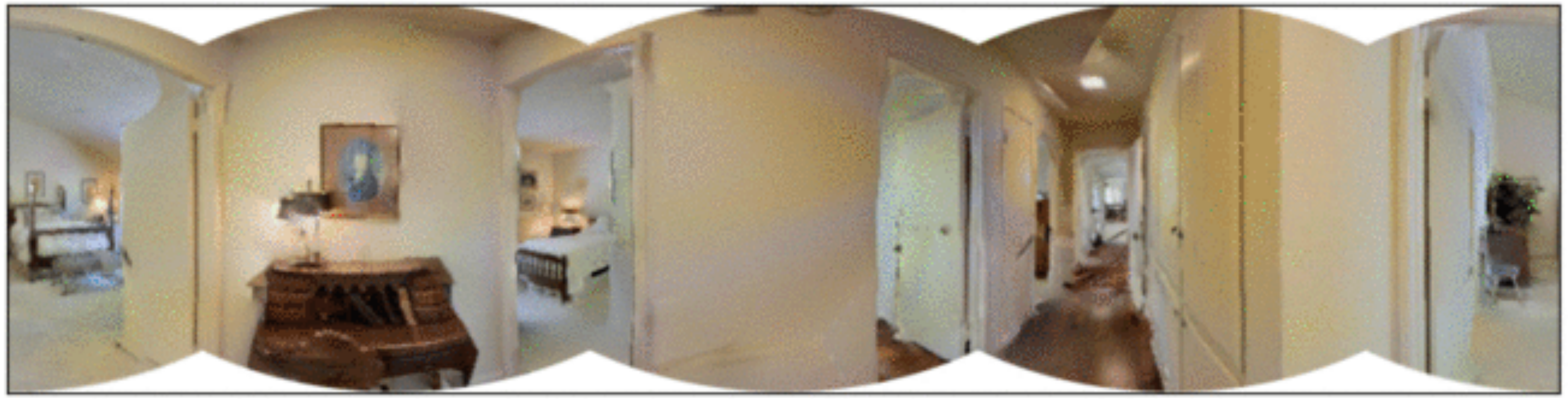
Goal Image

Topological Map and Pose



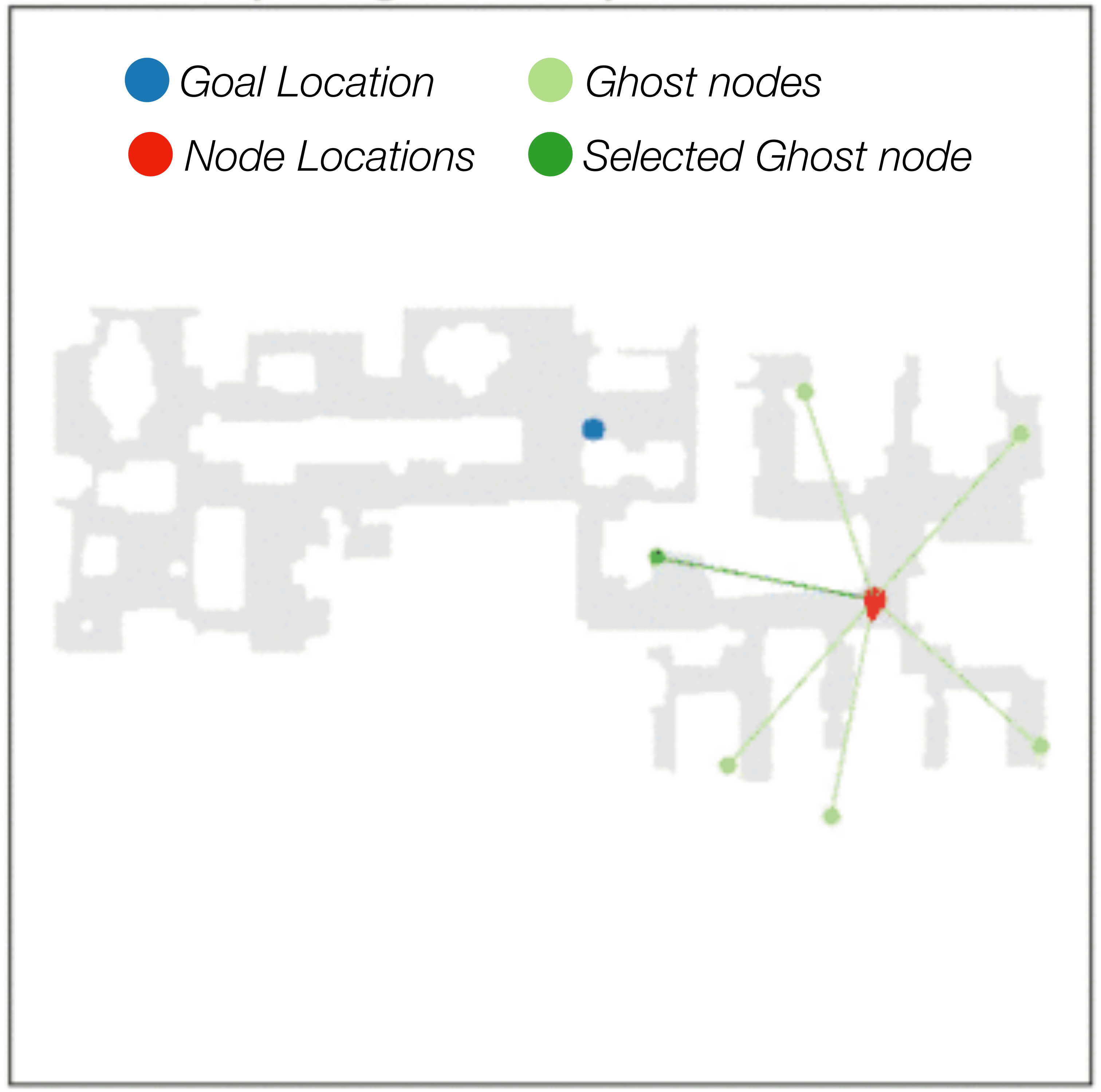
Demo video

Observation

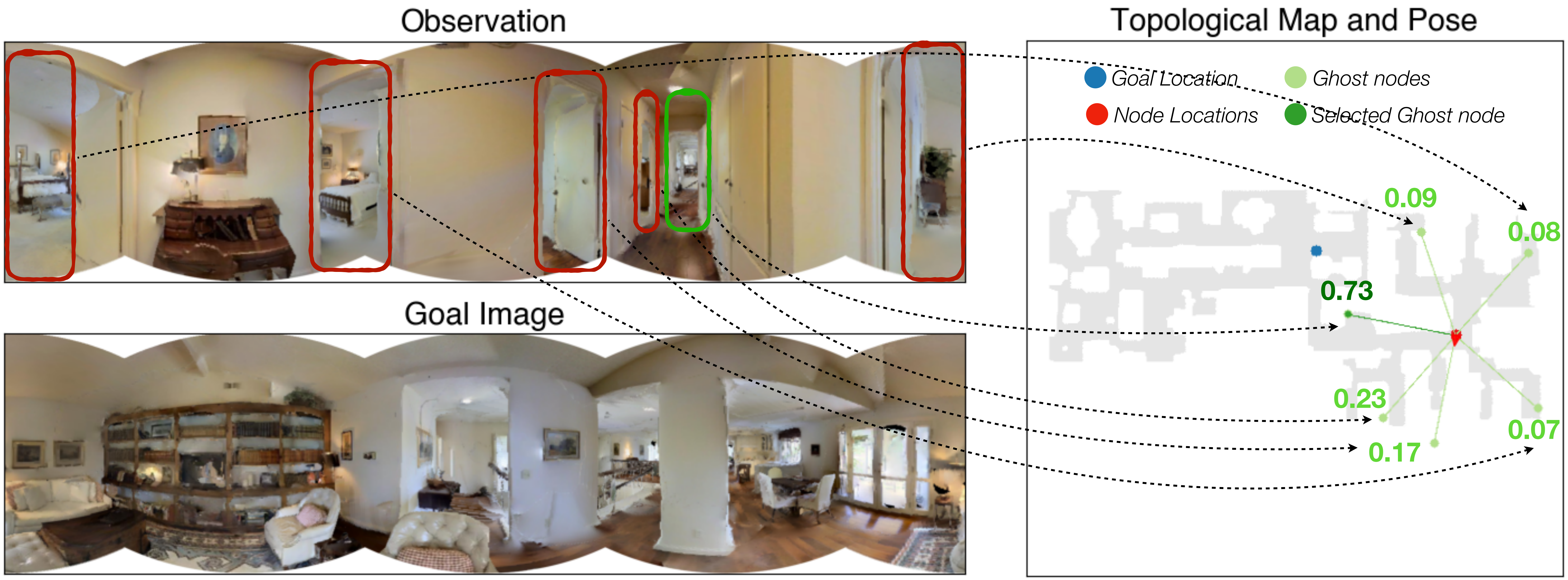


Goal Image

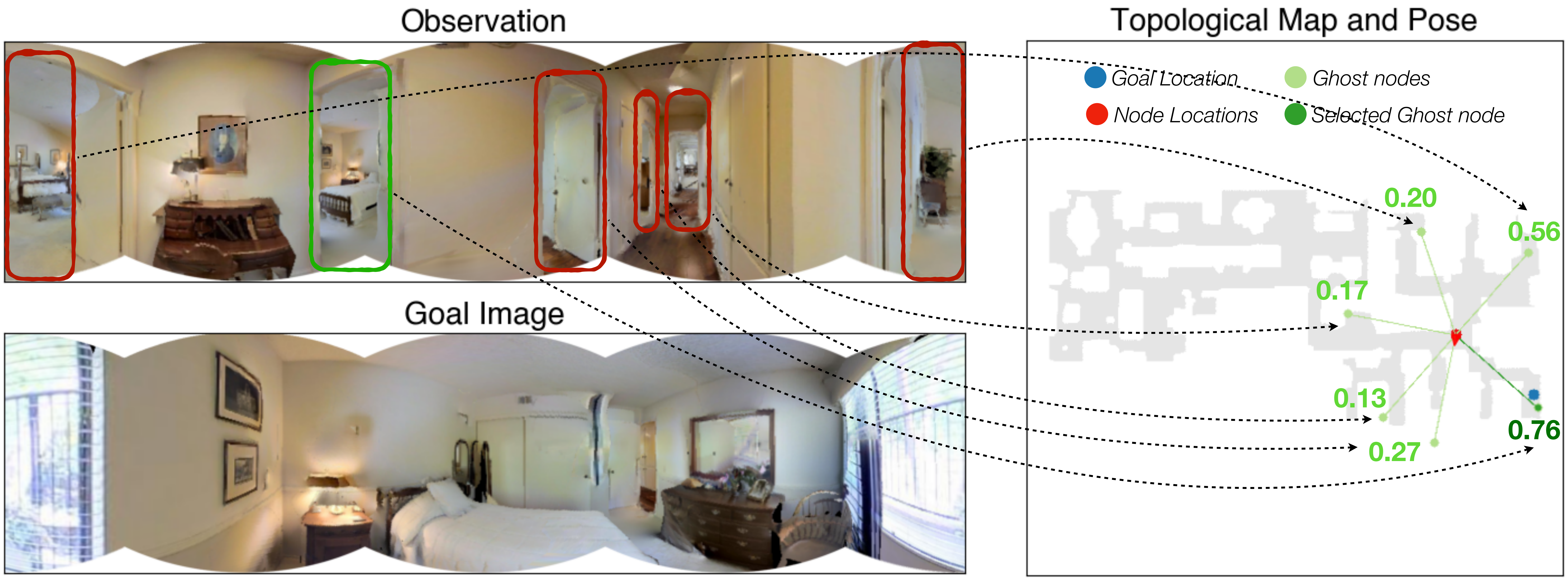
Topological Map and Pose



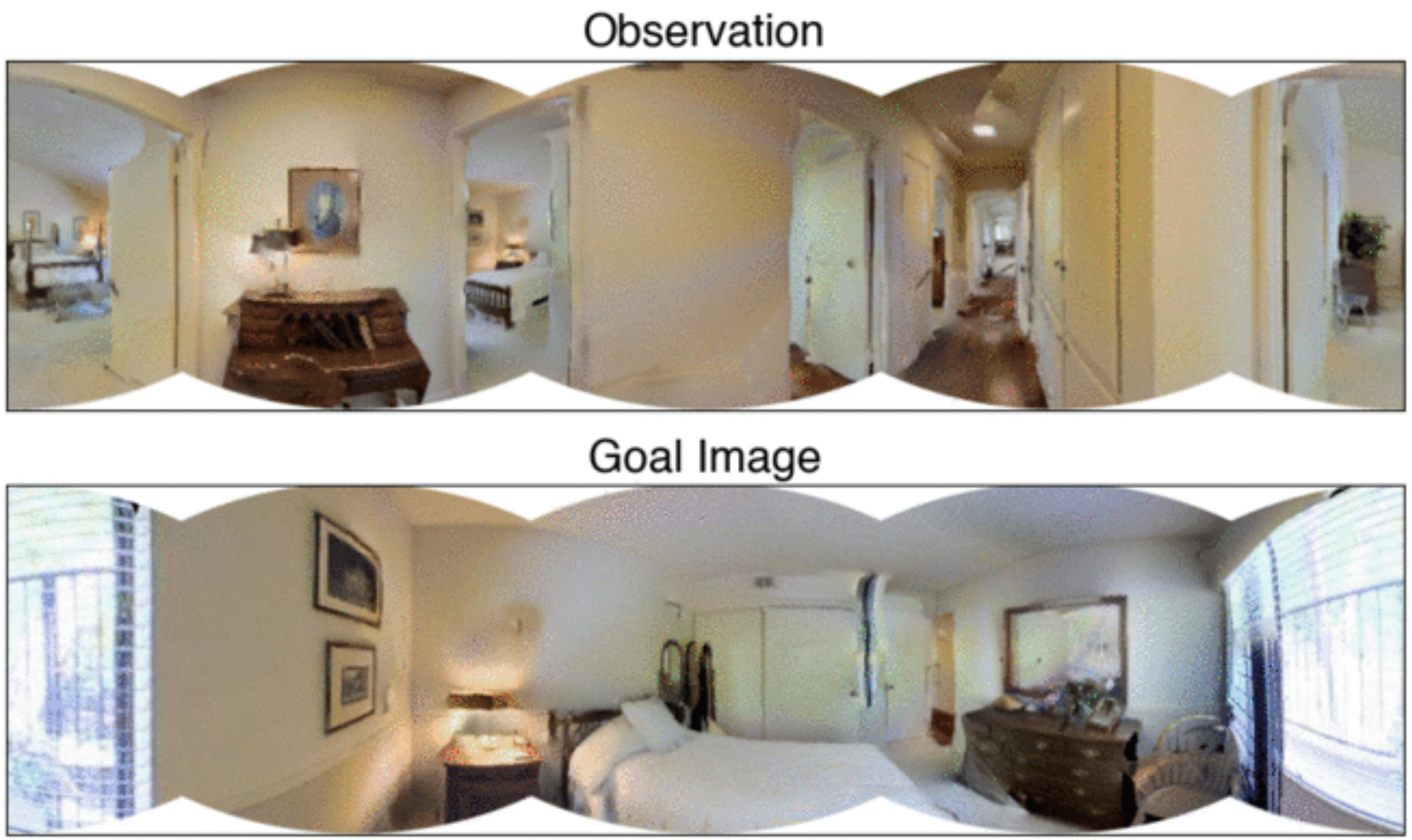
Learning Semantic Priors

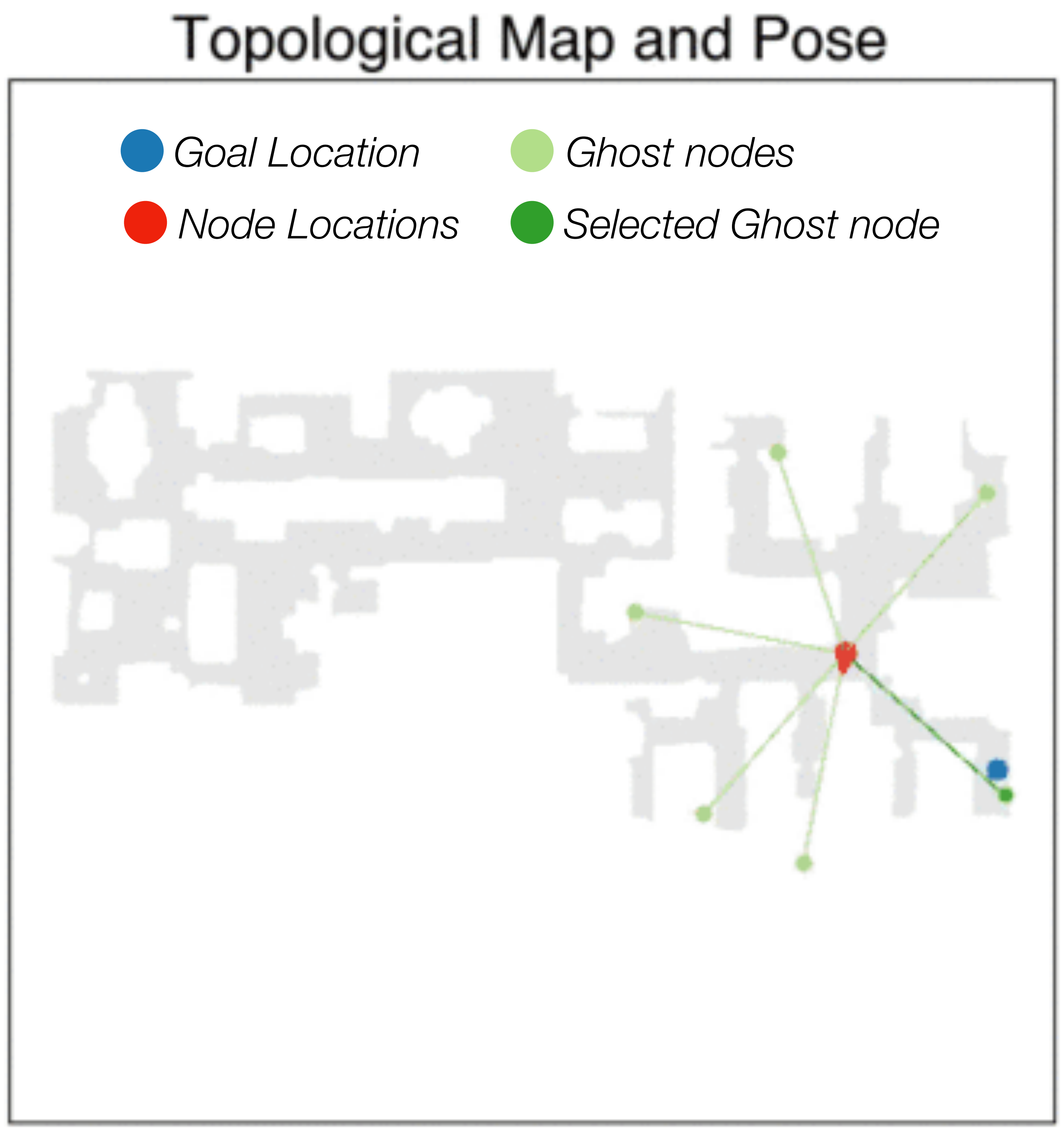


Learning Semantic Priors



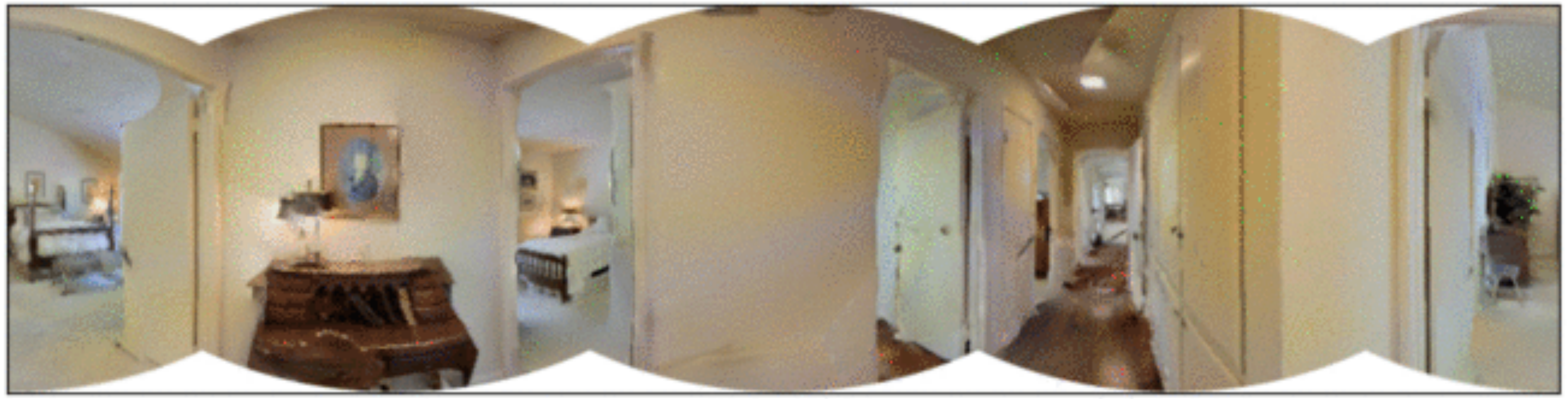
Learning Semantic Priors





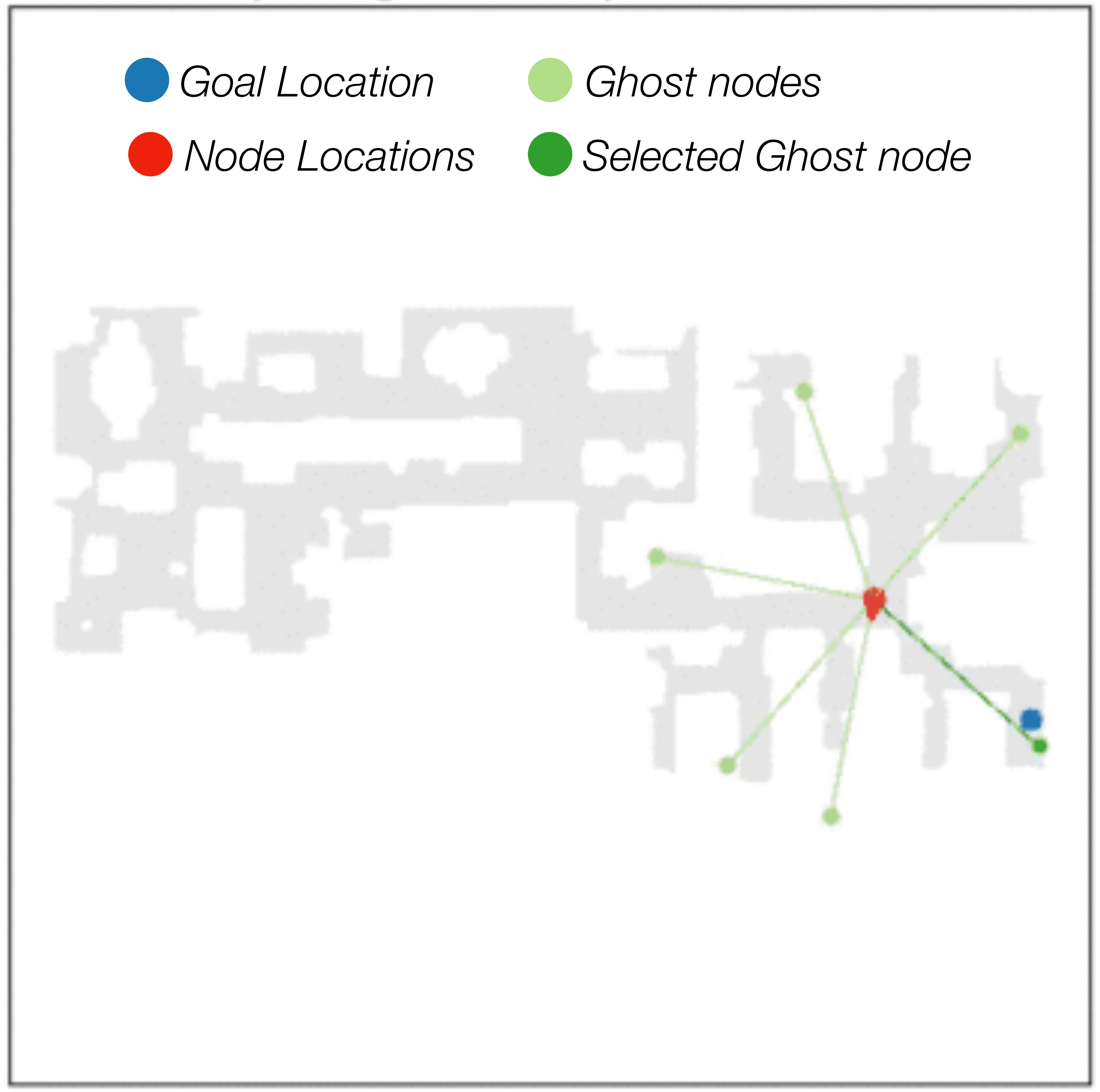
Learning Semantic Priors

Observation



Goal Image

Topological Map and Pose



Results

		RGB	RGBD	RGBD (No Noise)	RGBD (No Stop)
End-to-end Learning	LSTM + Imitation	0.10	0.14	0.15	0.18
	LSTM + RL	0.10	0.13	0.14	0.17
Modular Metric Maps	Occupancy Maps + FBE + RL	N/A	0.26	0.31	0.24
	Active Neural SLAM	0.23	0.29	0.35	0.39
Topological Maps	**Neural Topological SLAM**	**0.38**	**0.43**	**0.45**	**0.60**

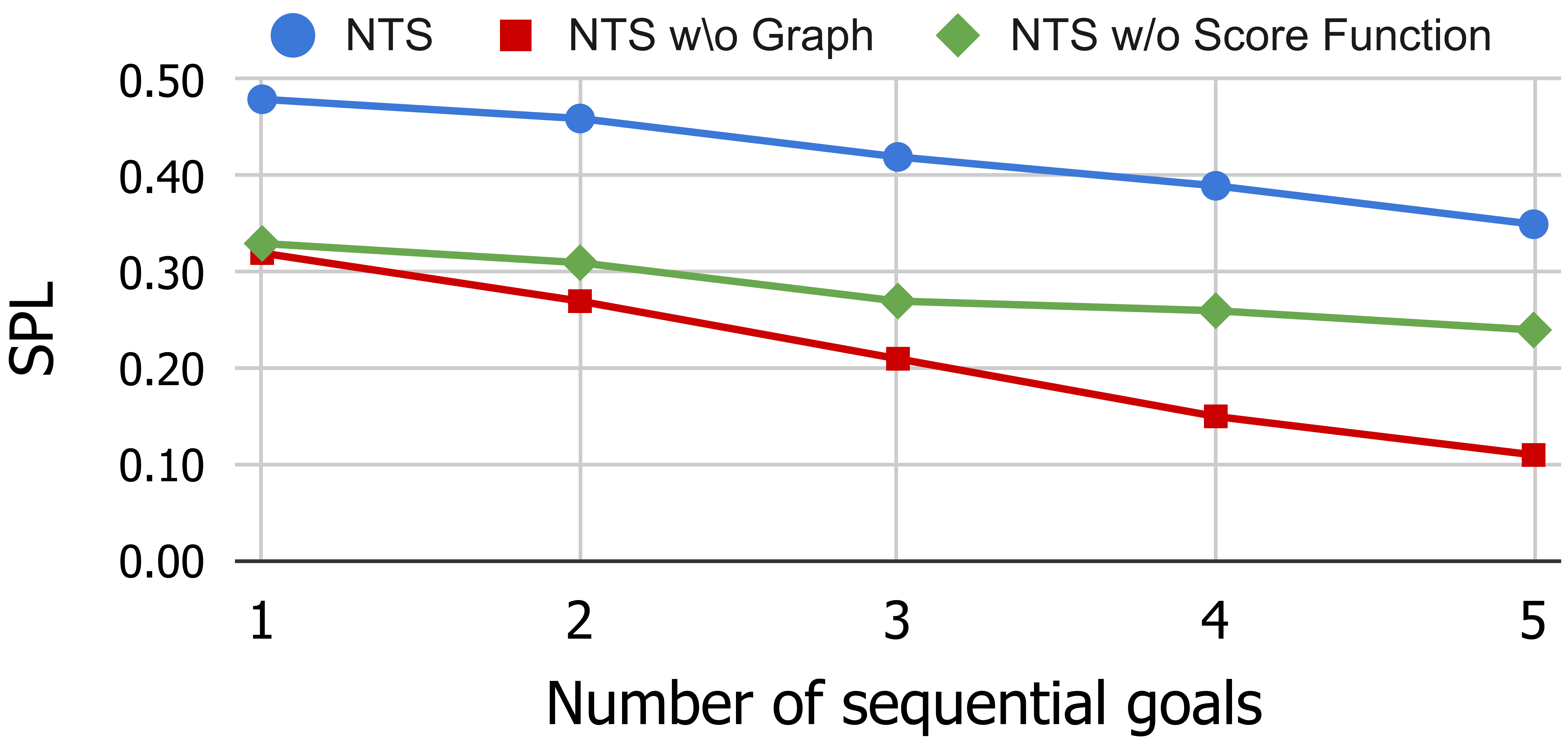
Results

Robustness to Pose Noise

		RGB	RGBD	RGBD (No Noise)	RGBD (No Stop)
End-to-end Learning	LSTM + Imitation	0.10	0.14	0.15	0.18
	LSTM + RL	0.10	0.13	0.14	0.17
Modular Metric Maps	Occupancy Maps + FBE + RL	N/A	0.26	0.31	0.24
	Active Neural SLAM	0.23	0.29	0.35	0.39
Topological Maps	**Neural Topological SLAM**	**0.38**	**0.43**	**0.45**	**0.60**

NTS is better than occupancy map models, captures and uses semantic priors.

Sequential Goals and Ablations



Sequential Goals and Ablations

Neural Topological SLAM for Visual Navigation

Devendra Singh Chaplot, Ruslan Salakhutdinov, Abhinav Gupta, Saurabh Gupta

CVPR 2020

Webpage: https://devendrachaplot.github.io/projects/Neural-Topological-SLAM

Thank you

Devendra Singh Chaplot

Webpage: http://devendrachaplot.github.io/
Email: chaplot@cs.cmu.edu
Twitter: @dchaplot


Neural Topological SLAM for Visual Navigation

Neural Topological SLAM for Visual Navigation

Devendra Singh Chaplot, Ruslan Salakhutdinov, Abhinav Gupta, Saurabh Gupta

CVPR 2020

Webpage: https://devendrachaplot.github.io/projects/Neural-Topological-SLAM

Thank you

Devendra Singh Chaplot

Webpage: http://devendrachaplot.github.io/

Email: chaplot@cs.cmu.edu

Twitter: @dchaplot