Topological Semantic Graph Memory for Image-Goal Navigation

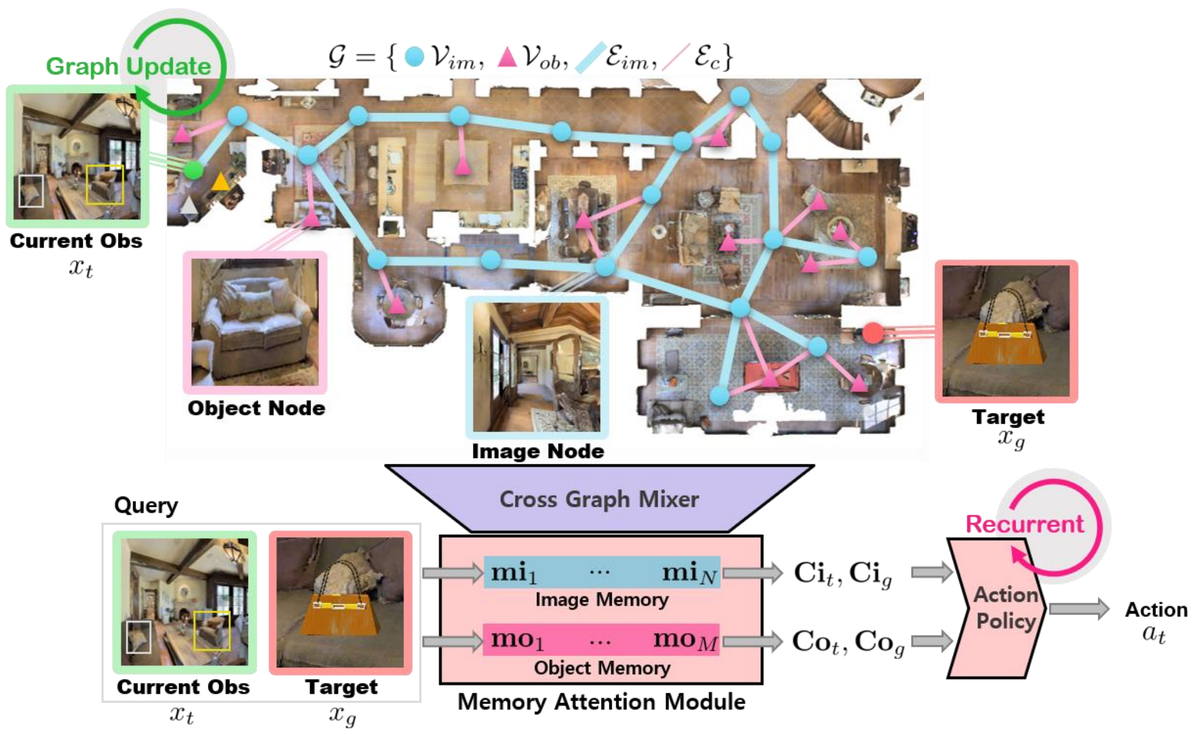
CoRL 2022 (accepted as **oral** presentation)

Nuri Kim Obin Kwon Hwiyeon Yoo Yunho Choi Jeongho Park Songhwai Oh

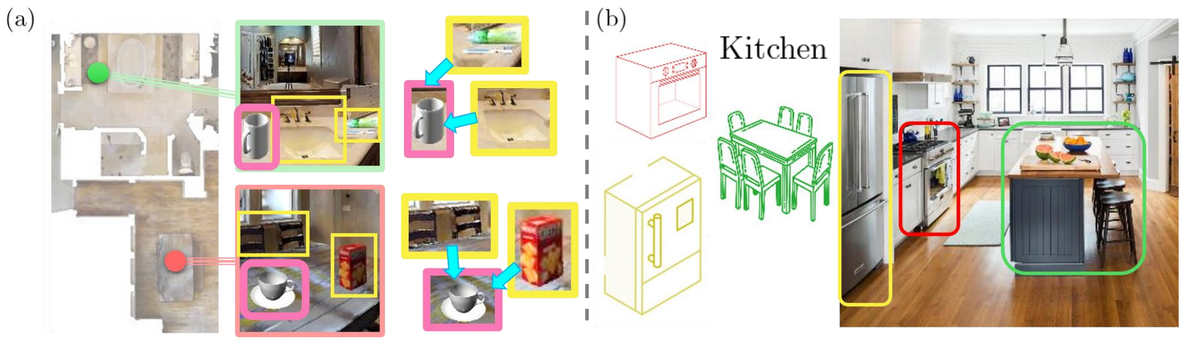
RLLAB http://rllab.snu.ac.kr

Abstract

This work proposes an approach to incrementally collect a landmark-based semantic graph memory and use the collected memory for image goal navigation. Given a target image to search, an embodied robot utilizes the semantic memory to find the target in an unknown environment. We present a method for incorporating object graphs into topological graphs, called ***Topological Semantic Graph Memory (TSGM)***. Although TSGM does not use position information, it can estimate 3D spatial topological information about objects.



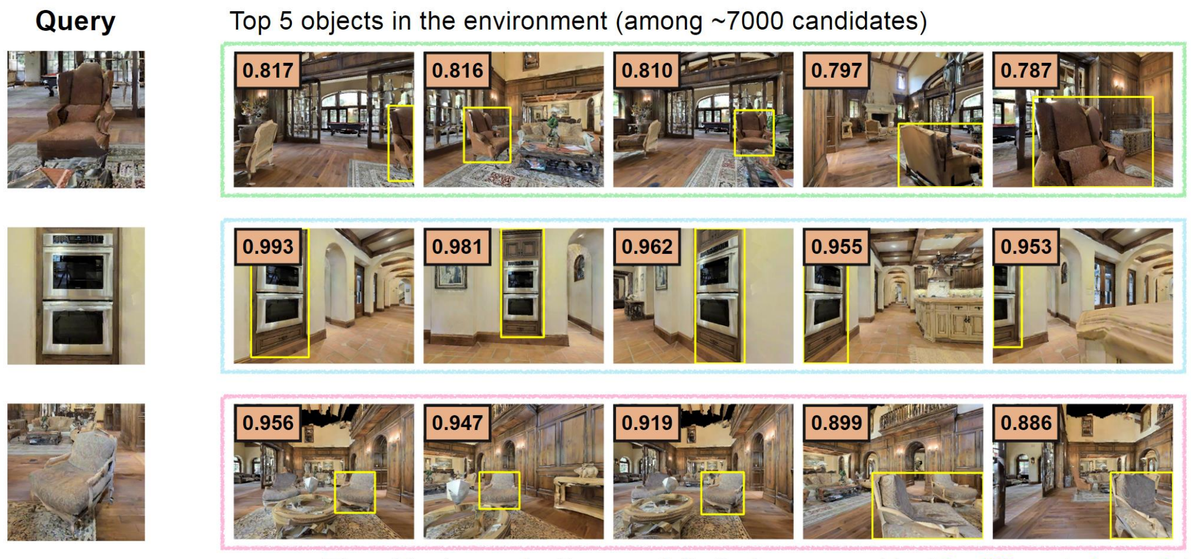
Why a robot needs a topological semantic memory?



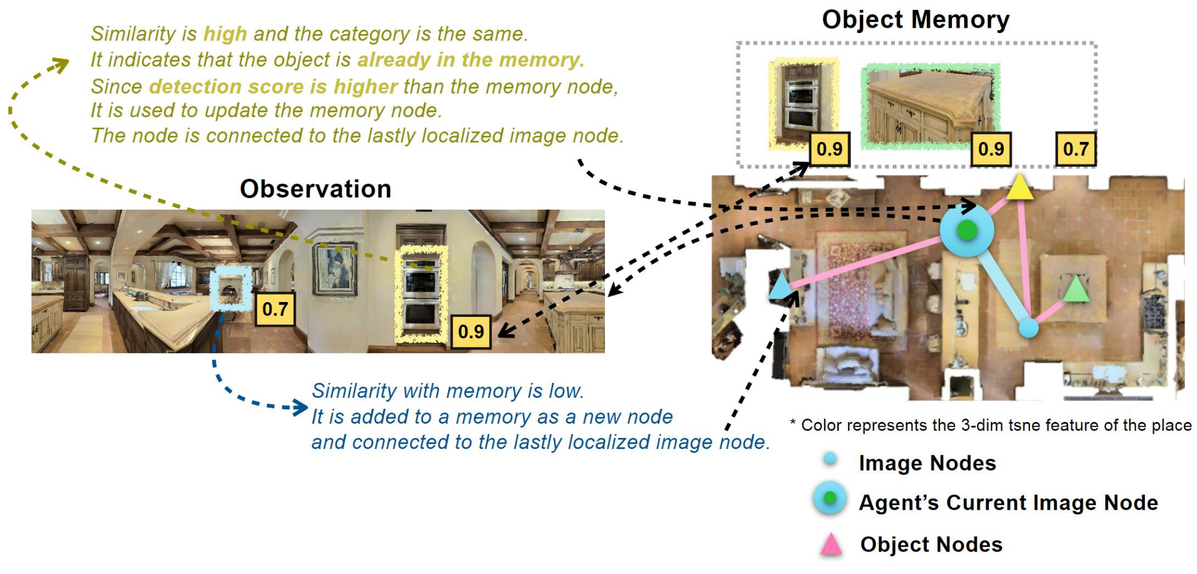
(a) The contextual representation, defining an object through neighboring objects, helps to eliminate the ambiguity of similar but different objects. For example, a cup in the kitchen can be perceived as one next to a chair and snack box, while a cup in the bathroom can be shown as one that is near to a toothbrush and washstand.

(b) A place can be better described through objects. A kitchen, for instance, can be defined by the presence of a refrigerator, oven, and dining table.

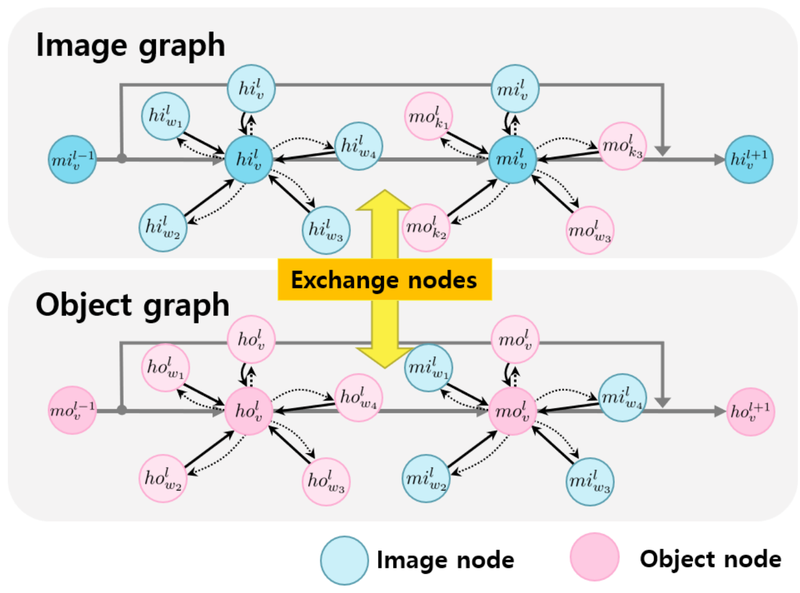
Graph Builder



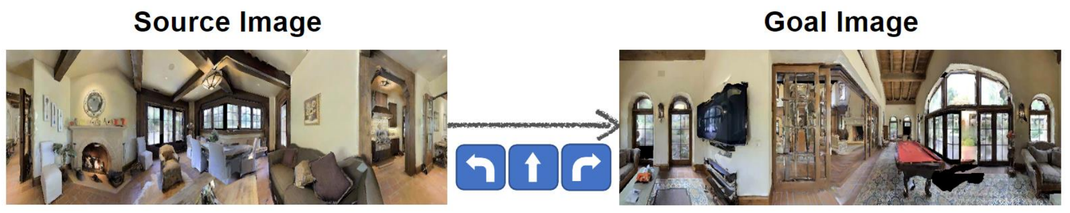
The object encoder successfully find a query object in different viewpoints



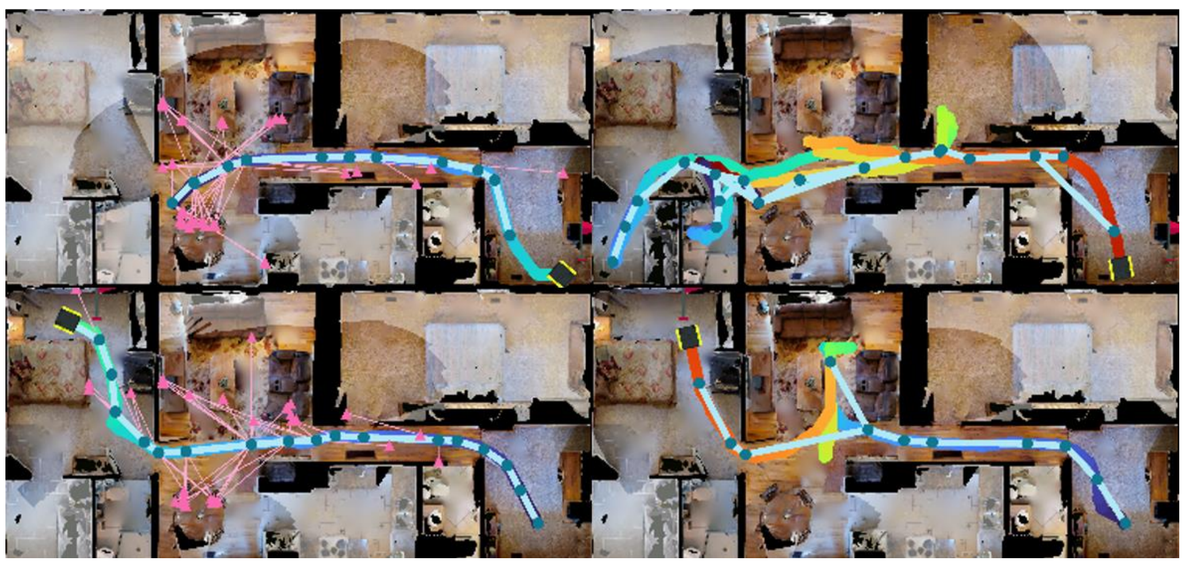
Cross Graph Mixer

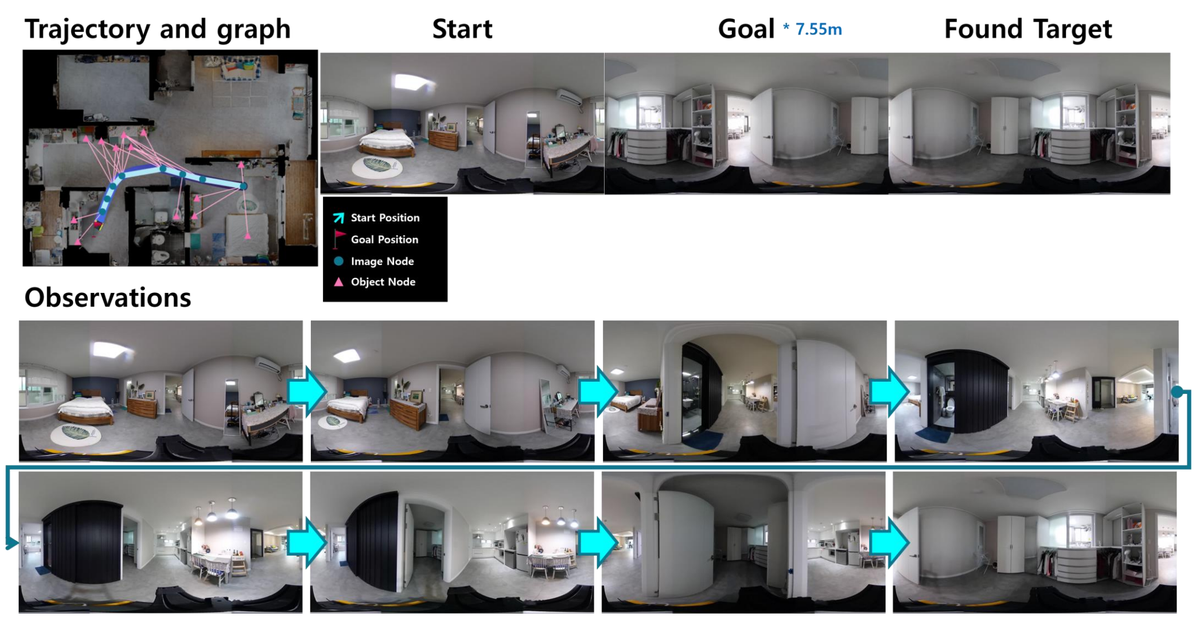


Task: Image Goal Navigation



Experiments





Results

Table 1: Comparison of TSGM with memory-based baselines on image goal navigation on Gibson.

Method	Memory	No Pose	Object	Easy		Medium		Hard		Overall	
				Success	SPL	Success	SPL	Success	SPL	Success	SPL
RGBD + RL [26]	implicit	✗	✗	72.5	69.5	53.1	48.6	22.3	17.7	49.3	45.3
Active Neural SLAM [17]	metric	✗	✗	74.2	20.5	68.4	22.9	29.9	11.0	57.5	18.1
Exp4nav [5]	metric	✗	✗	70.2	61.8	60.6	52.4	46.9	38.5	59.2	50.9
SMT [8]	graph	✗	✗	81.9	77.4	65.6	52.2	55.6	39.7	67.7	56.4
Neural Planner [20]	graph	✗	✗	71.7	41.3	64.7	38.5	42.0	27.0	59.5	35.6
SPTM [9]	graph	✔	✗	66.5	40.6	64.2	38.5	42.1	25.4	57.6	34.8
VGM [18]	graph	✔	✗	86.1	79.6	81.2	**68.2**	60.9	45.6	76.1	64.5
TSGM (Ours)	graph	✔	✔	**91.1**	**83.5**	**82.0**	68.1	**70.3**	**50.0**	**81.1**	**67.2**

Table 2: Comparison of TSGM with image goal navigation baselines on straight/curved episodes on Gibson.

Path Type	Method	Easy		Medium		Hard		Overall	
		Success	SPL	Success	SPL	Success	SPL	Success	SPL
Straight	NRNS [27]	67.1	57.8	52.4	41.2	32.6	22.4	50.7	40.5
	VGM [18]	81.0	54.4	82.0	69.9	67.3	54.4	76.7	59.6
	TSGM (Ours)	**94.4**	**92.1**	**92.6**	**84.3**	**70.3**	**62.8**	**85.7**	**79.7**
Curved	NRNS [27]	31.7	13.0	29.0	13.6	19.2	10.4	26.6	12.3
	VGM [18]	81.0	45.5	78.8	59.5	62.2	46.9	74.0	50.6
	TSGM (Ours)	**93.6**	**91.0**	**89.7**	**77.8**	**64.2**	**55.0**	**82.5**	**74.1**