

CSE 598
Action and Perception
Oct 23 - Lab 9

Drew

Overview

- Changing simulators... to **Habitat**
 - Colab Pro free for students?
 - Go over habitat structure and datasets
 - High-level overview of some visual navigation tasks
 - Quiz
 - **Note: no notebook today**

Overview of Habitat Sim / Lab!



> 10,000 FPS with a GPU!

How to Think about Embodied Simulators

Tasks



Simulators

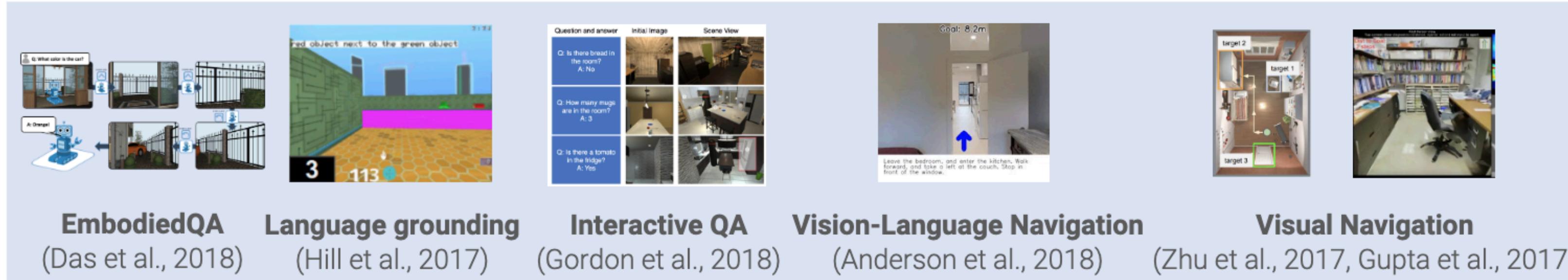


Datasets



How to Think about Habitat

Tasks



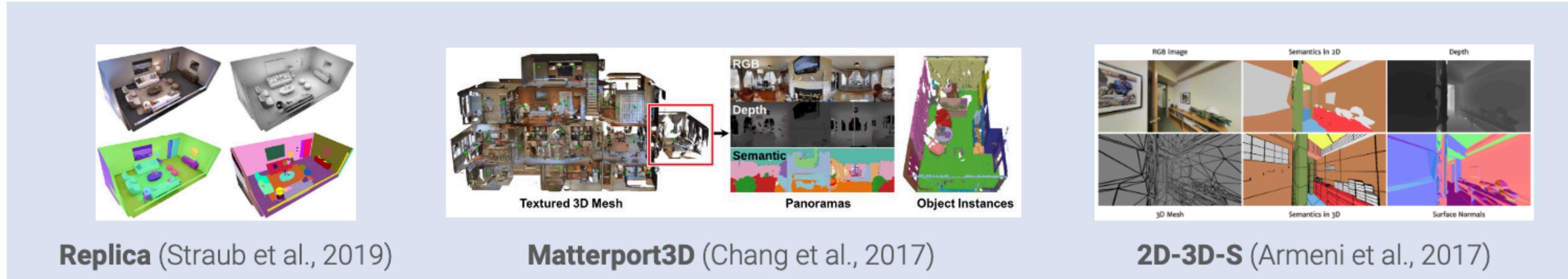
Habitat

Simulators



Habitat Lab

Datasets



Habitat Sim

Generic Dataset API

Design Philosophy: simply physics to speed up rendering

Overview of Habitat Sim / Lab!

Habitat Sim

Flexible, high-performance 3D simulator

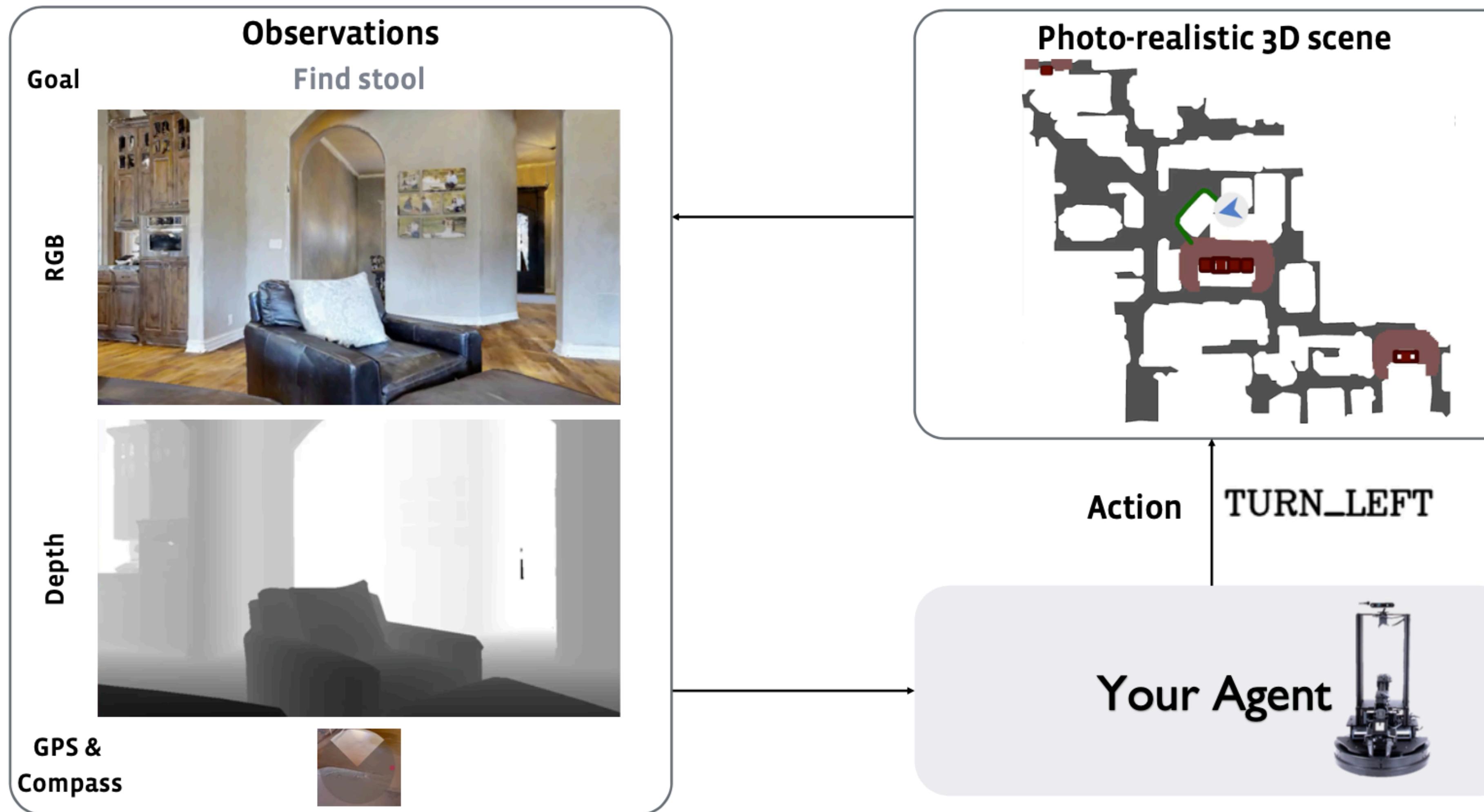
Defines: configurable agents, multiple sensors, high throughput rendering, and generic 3D dataset handling (with built-in support for Matterport3D, Gibson, and Replica datasets).

Habitat Lab

Modular high-level library for end-to-end development of embodied AI algorithms

Defines: embodied AI tasks (e.g. navigation, instruction following, question answering), configuring and training embodied agents (via imitation or reinforcement learning, or via classic SLAM), and benchmarking using standard metric

Overview of Habitat Sim / Lab!



Dataset Scale

Dataset	Replica	RoboTHOR	MP3D	Gibson (4+ only)	ScanNet	HM3D (ours)
Number of scenes	18	75	90	571 (106)	1613	1000
Navigable space (m ²)	0.56k	0.75k	30.22k	81.84k (7.18k)	10.52k	112.50k
Floor space (m ²)	2.19k	3.17k	101.82k	217.99k (17.74k)	39.98k	365.42k
Navigation complexity	5.99	2.06	17.09	14.25 (11.90)	3.78	13.31
Scene clutter	3.4	8.2	2.99	3.14 (3.04)	3.15	3.90

Table 1. Comparing Habitat-Matterport (HM3D) dataset with prior datasets.



00000-kfPV7w3FaU5.glb

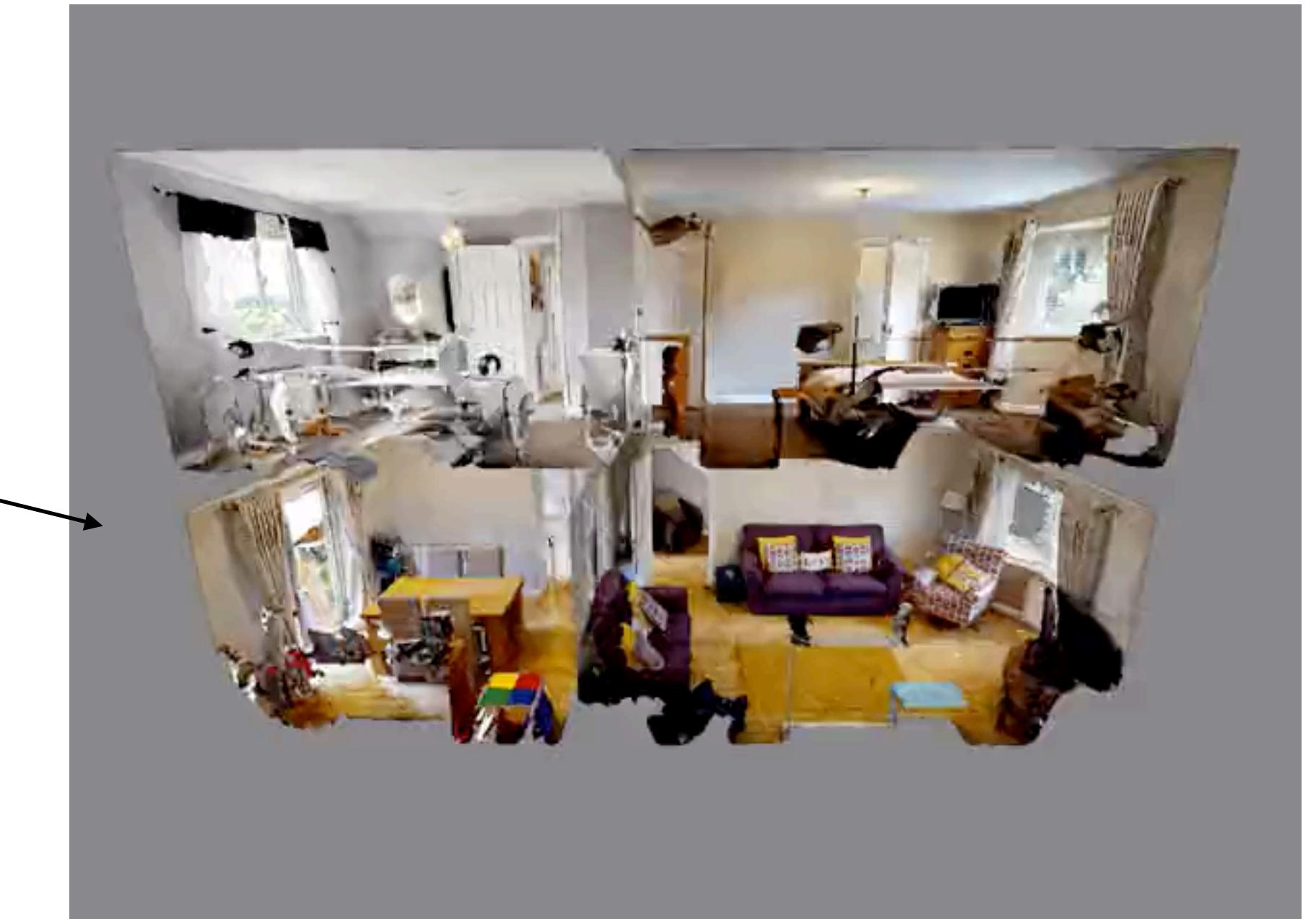
Dataset Scale

Dataset	Replica	RoboTHOR	MP3D	Gibson (4+ only)	ScanNet	HM3D (ours)
Number of scenes	18	75	90	571 (106)	1613	1000
Navigable space (m ²)	0.56k	0.75k	30.22k	81.84k (7.18k)	10.52k	112.50k
Floor space (m ²)	2.19k	3.17k	101.82k	217.99k (17.74k)	39.98k	365.42k
Navigation complexity	5.99	2.06	17.09	14.25 (11.90)	3.78	13.31
Scene clutter	3.4	8.2	2.99	3.14 (3.04)	3.15	3.90

Table 1. Comparing Habitat-Matterport (HM3D) dataset with prior datasets.



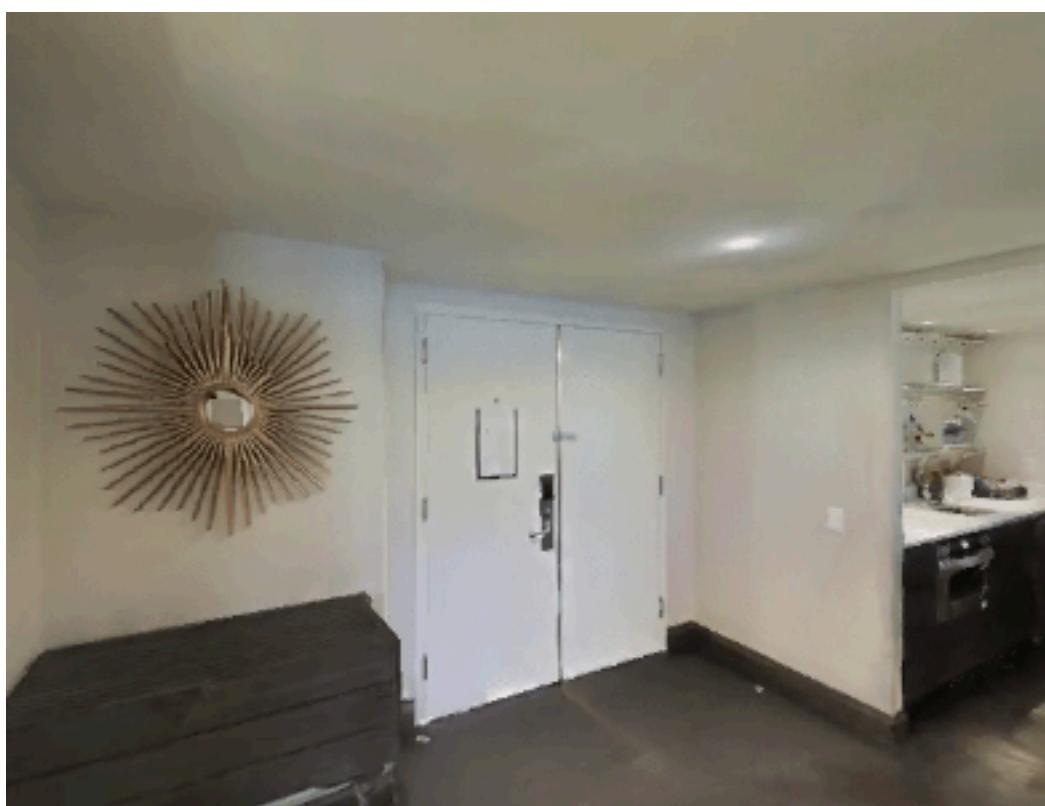
Both used for rendering and
constructing the **Navmesh**



00000-kfPV7w3FaU5.glb

Rendering with HabitatSim

mp3d



RGB



Depth



Semantic

Replica



A Flavor of Sim2Real Issues



Sliding Allowed



Sliding Not Allowed

What are some tasks?

Task: Visual Navigation

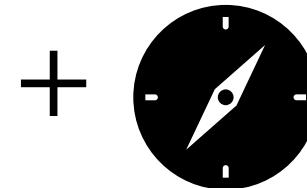
Visual Navigation Tasks

Navigate to **the target** in an unseen environment

PointNav: Navigate to (x, y)

TargetNav: Find green sphere

Observations



RGB + Compass + GPS

Environment



Matterport 3D Dataset

Task: Visual Navigation

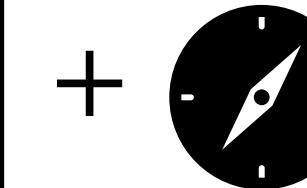
Visual Navigation Tasks

Navigate to **the target** in an unseen environment

PointNav: Navigate to (x, y)

TargetNav: Find green sphere

Observations



RGB + Compass + GPS

Environment



Matterport 3D Dataset

Evaluation:

Success weighted by **P**ath **L**ength:

$$SPL = \frac{1}{N} \sum_{i=1}^N S_i \frac{l_i}{\max(l_i, p_i)}$$

Task: Visual Navigation

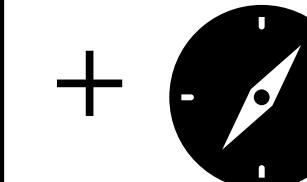
Visual Navigation Tasks

Navigate to **the target** in an unseen environment

PointNav: Navigate to (x, y)

TargetNav: Find green sphere

Observations



RGB + Compass + GPS

Environment



Matterport 3D Dataset

Evaluation:

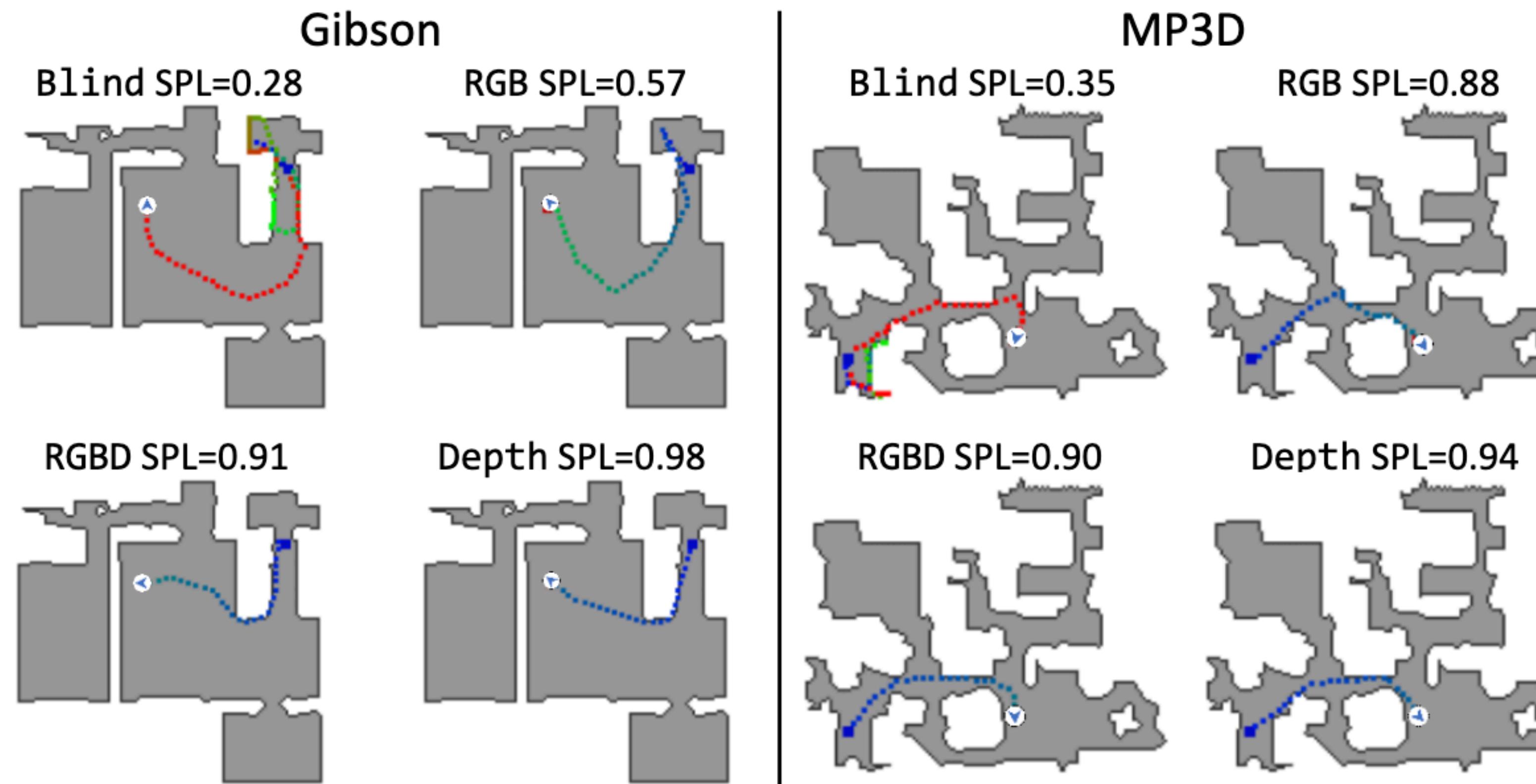
Success weighted by Path Length:

$$SPL = \frac{1}{N} \sum_{i=1}^N S_i \frac{l_i}{\max(l_i, p_i)}$$

Annotations for the evaluation formula:

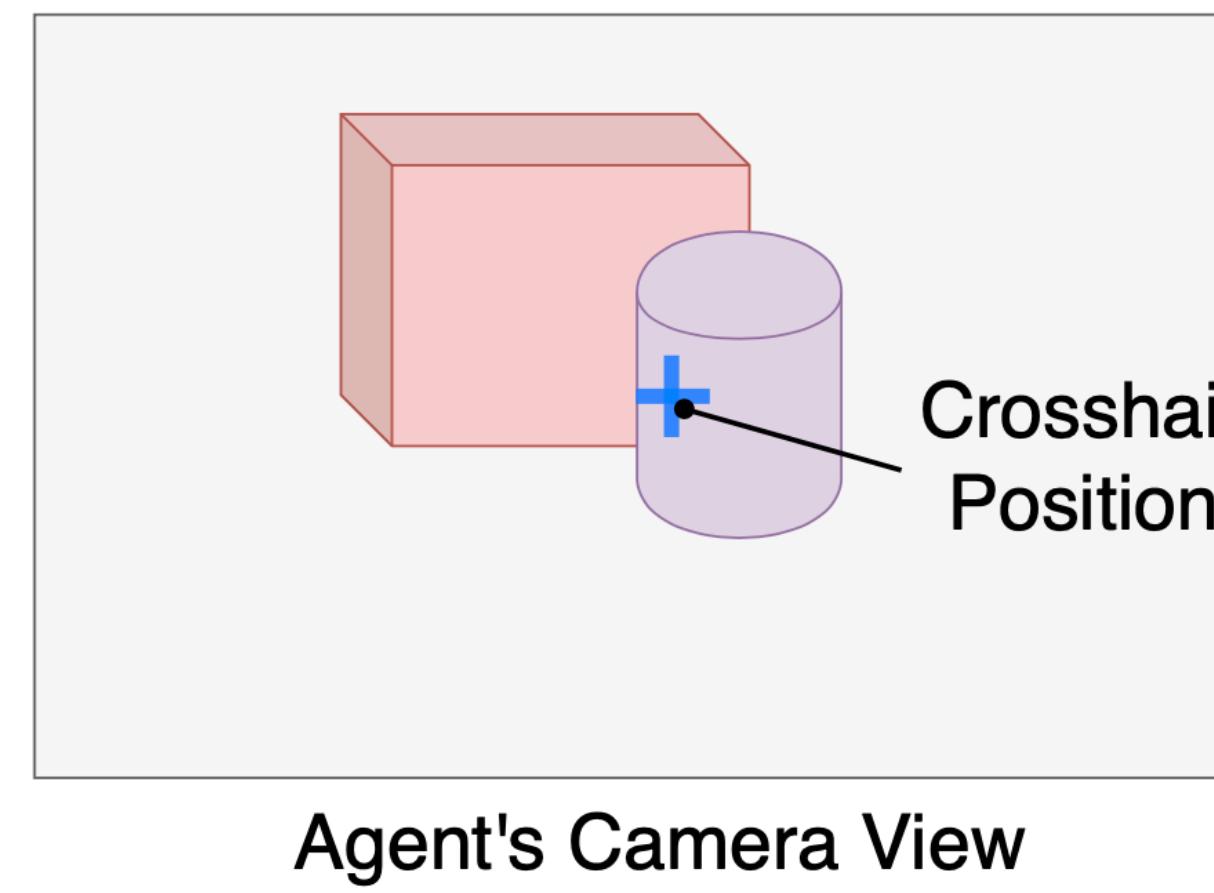
- Success Bool**: Points to the S_i term.
- Shortest geodesic path Len**: Points to the l_i term.
- Agent's path length**: Points to the p_i term.
- Num Eval Episodes**: Points to the N term.

Habitat SPL Examples

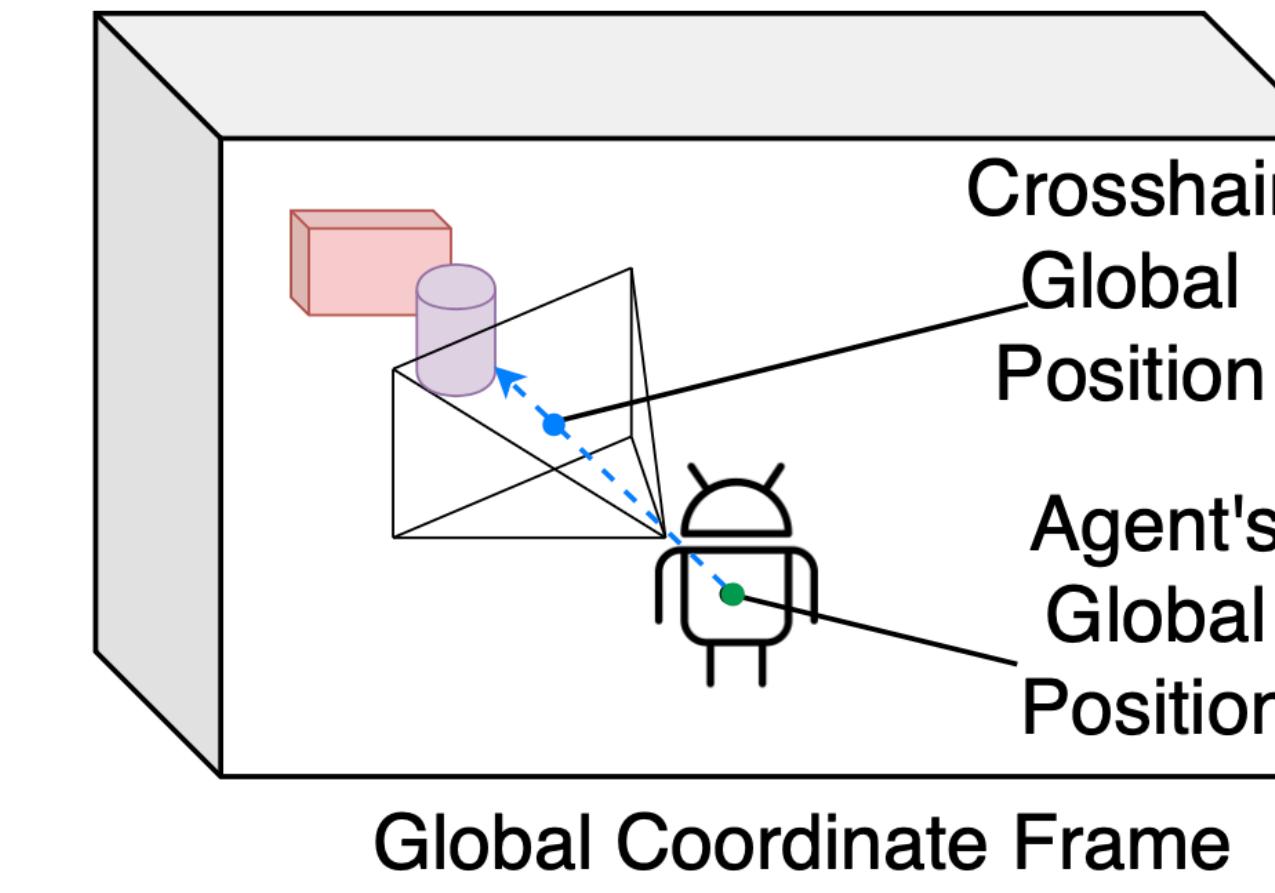


PointNav Task Trained Via HabitatLab PPO Impl

Task: Object Interaction & Rearrangement

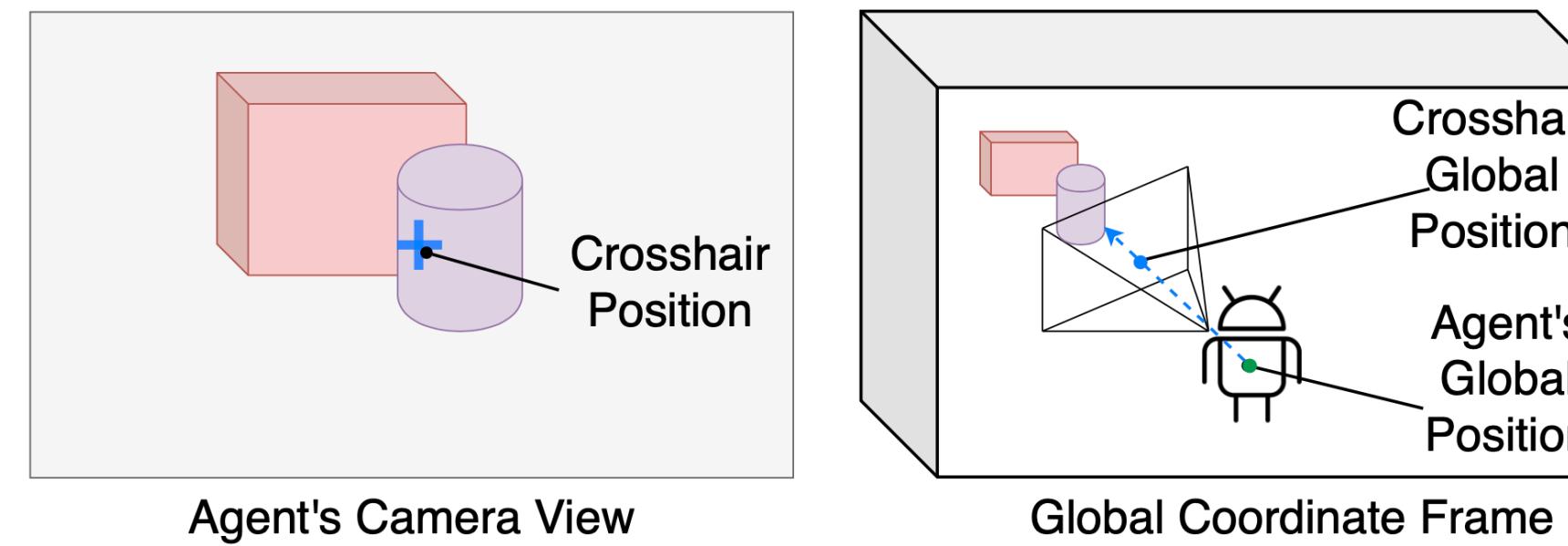


Agent's Camera View



Global Coordinate Frame

Task: Object Interaction & Rearrangement



HabitatLab2.0 Focused on different Rearrangement Tasks

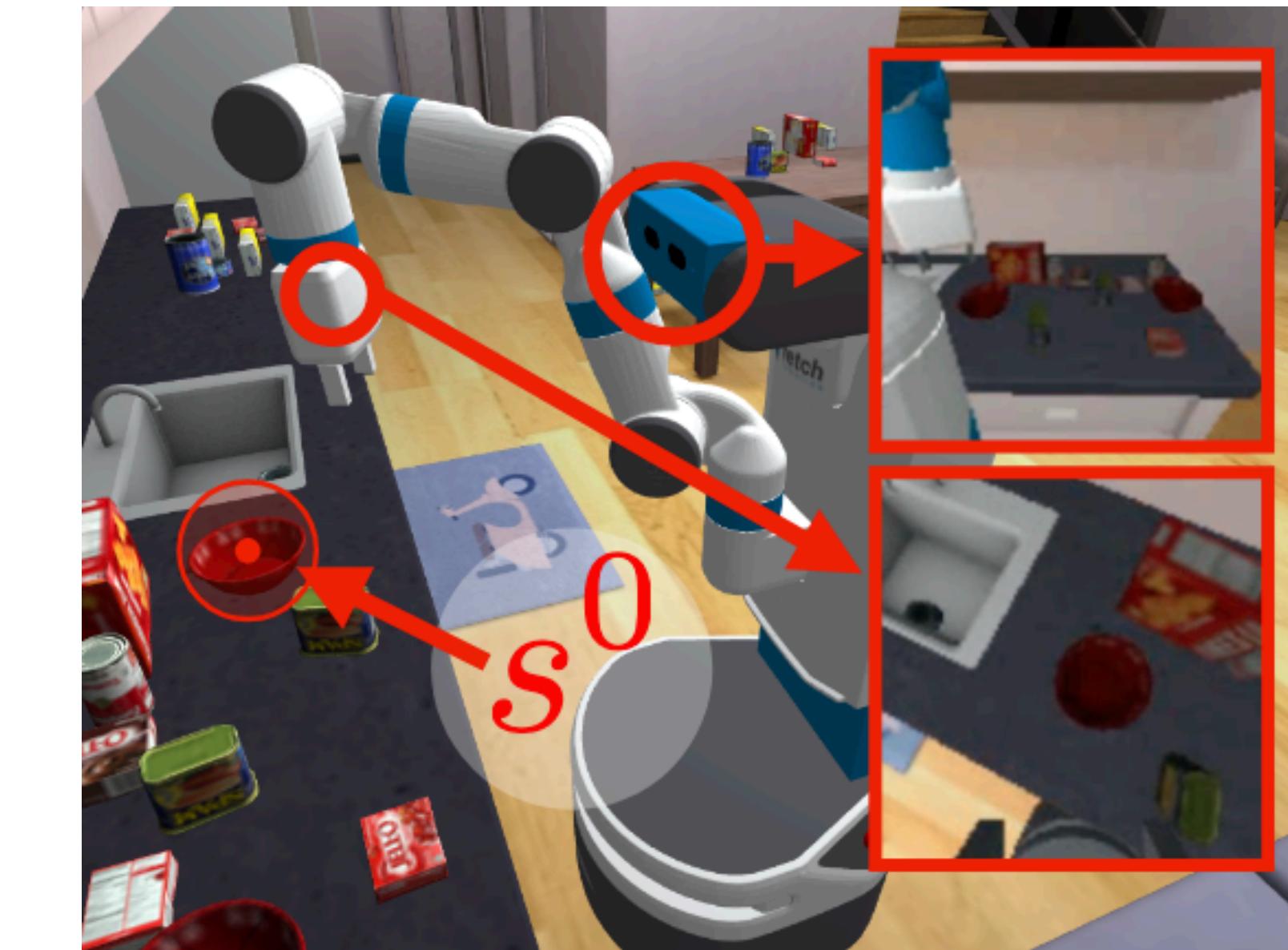


(a) Tidy House



(b) Prepare Groceries

Replica → ReplicaCAD



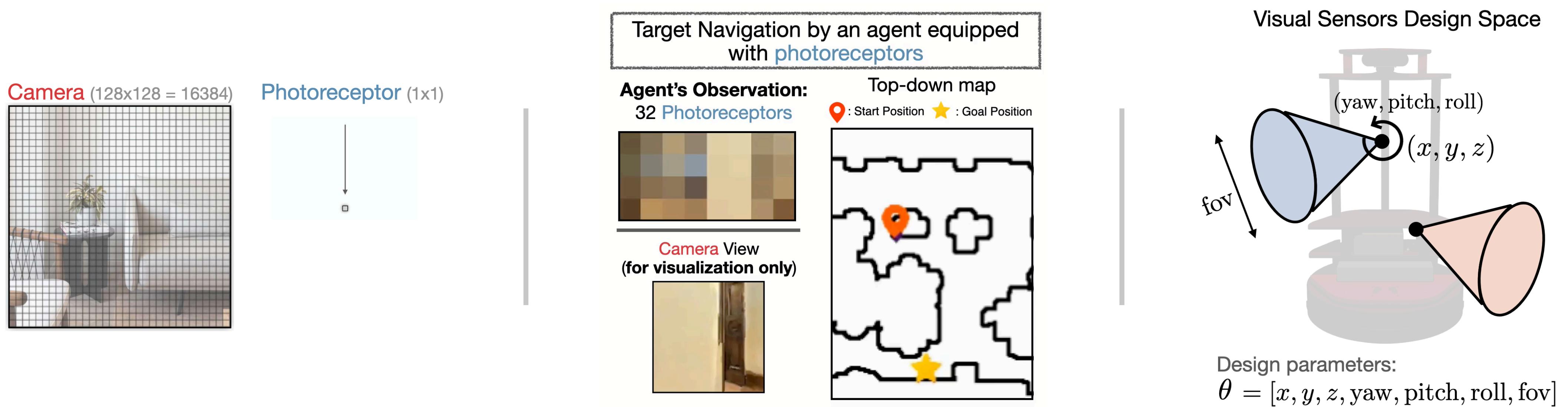
Pick-n-Place Task

Task: Multimodal (Audio-Visual) Navigation

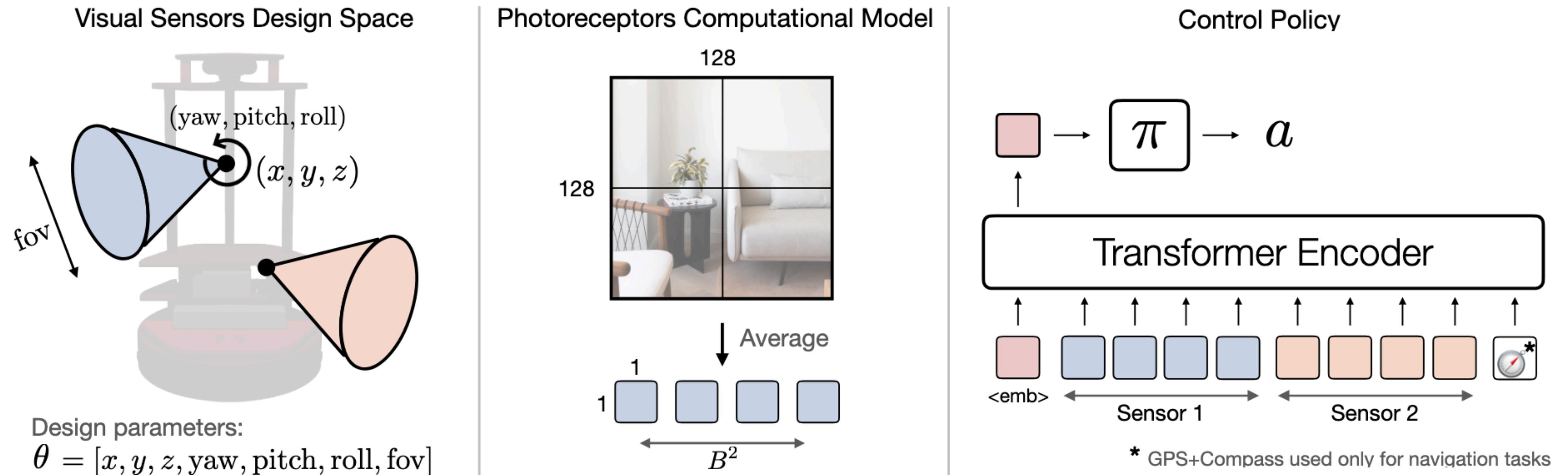


Task: Navigation from Constrained Vision

1. Simple photoreceptors can be used to solve non-trivial tasks
2. Their design (e.g. 6DOF Placement, FOV) is crucial for their effectiveness
3. Computational optimization methods can automate this design

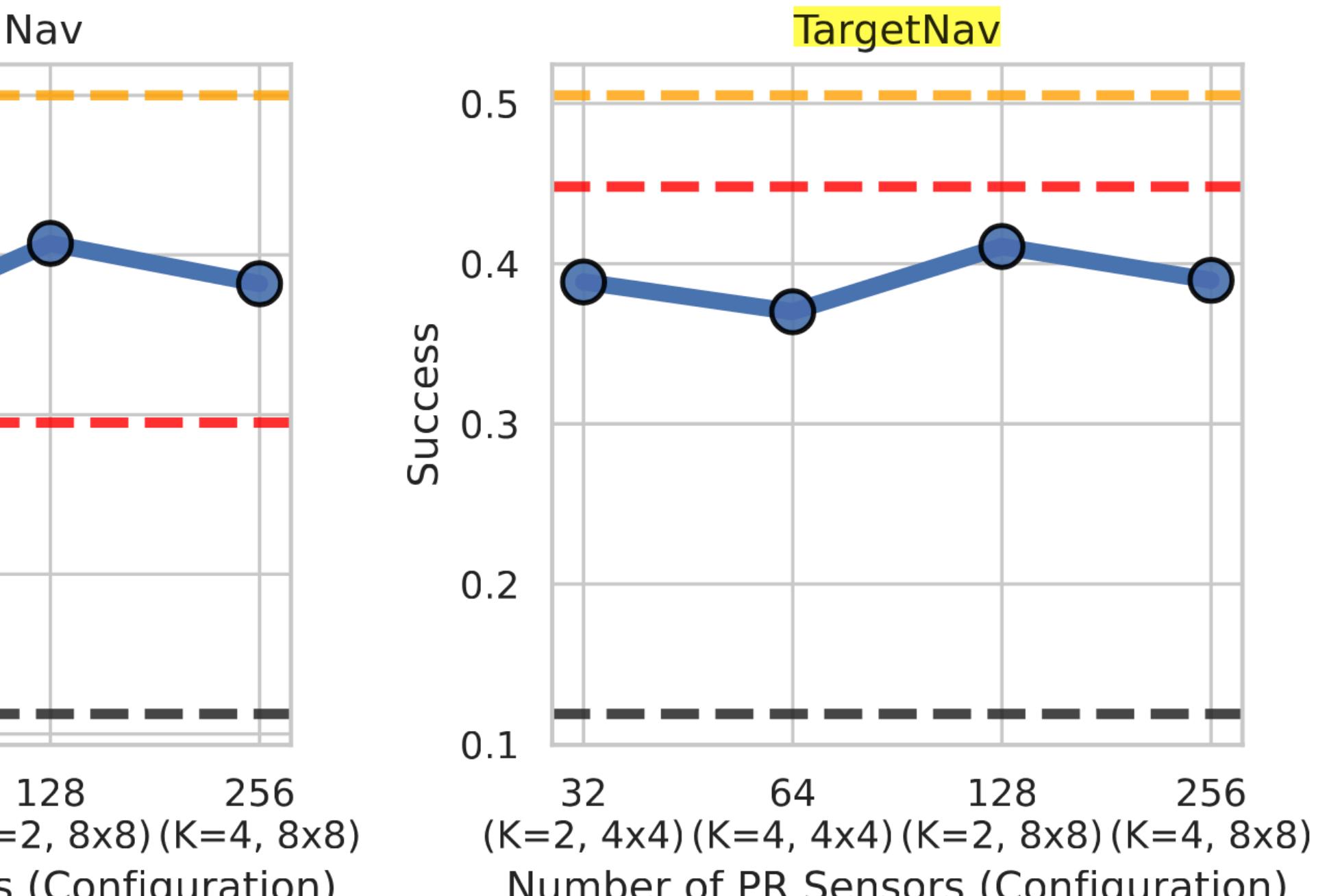
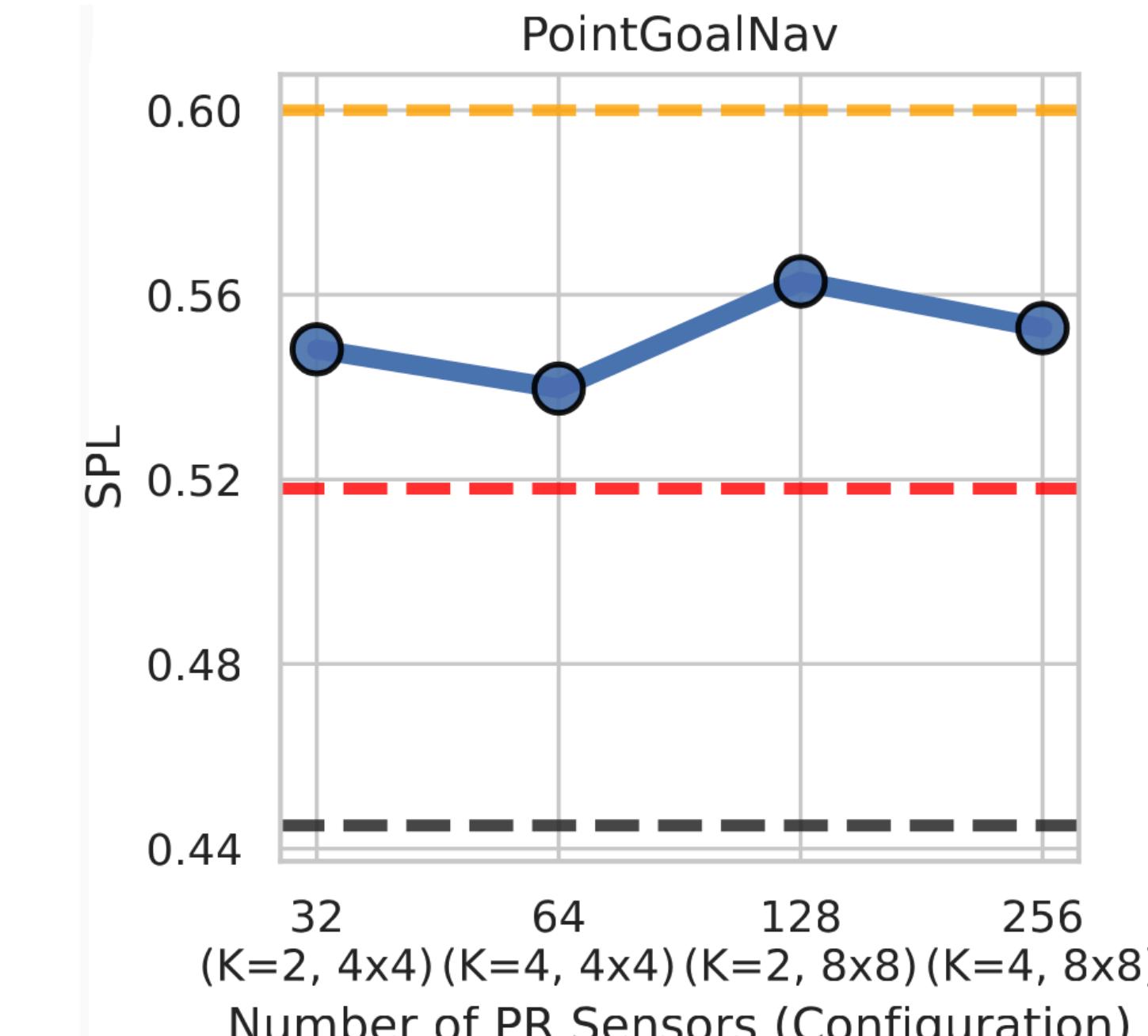
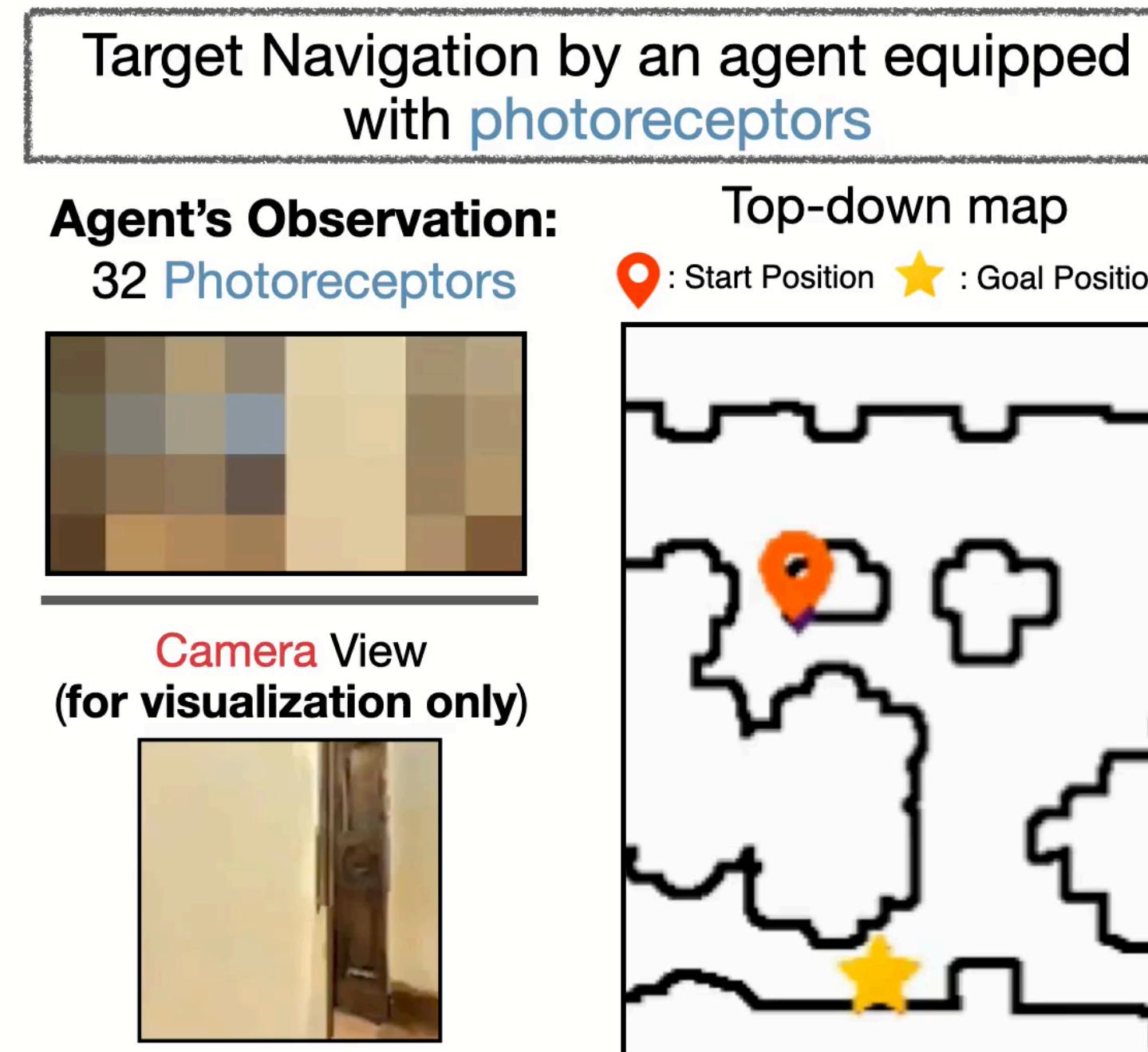


Task: Navigation from Constrained Vision

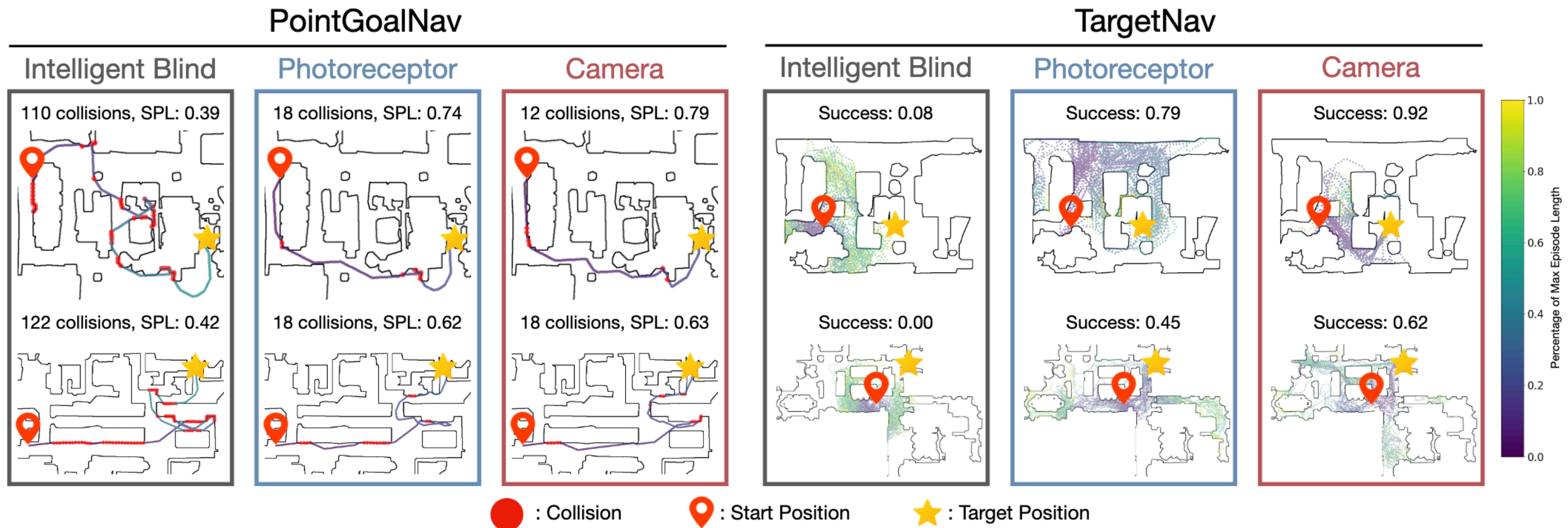


Task: Navigation from Constrained Vision

A handful of photoreceptors is enough

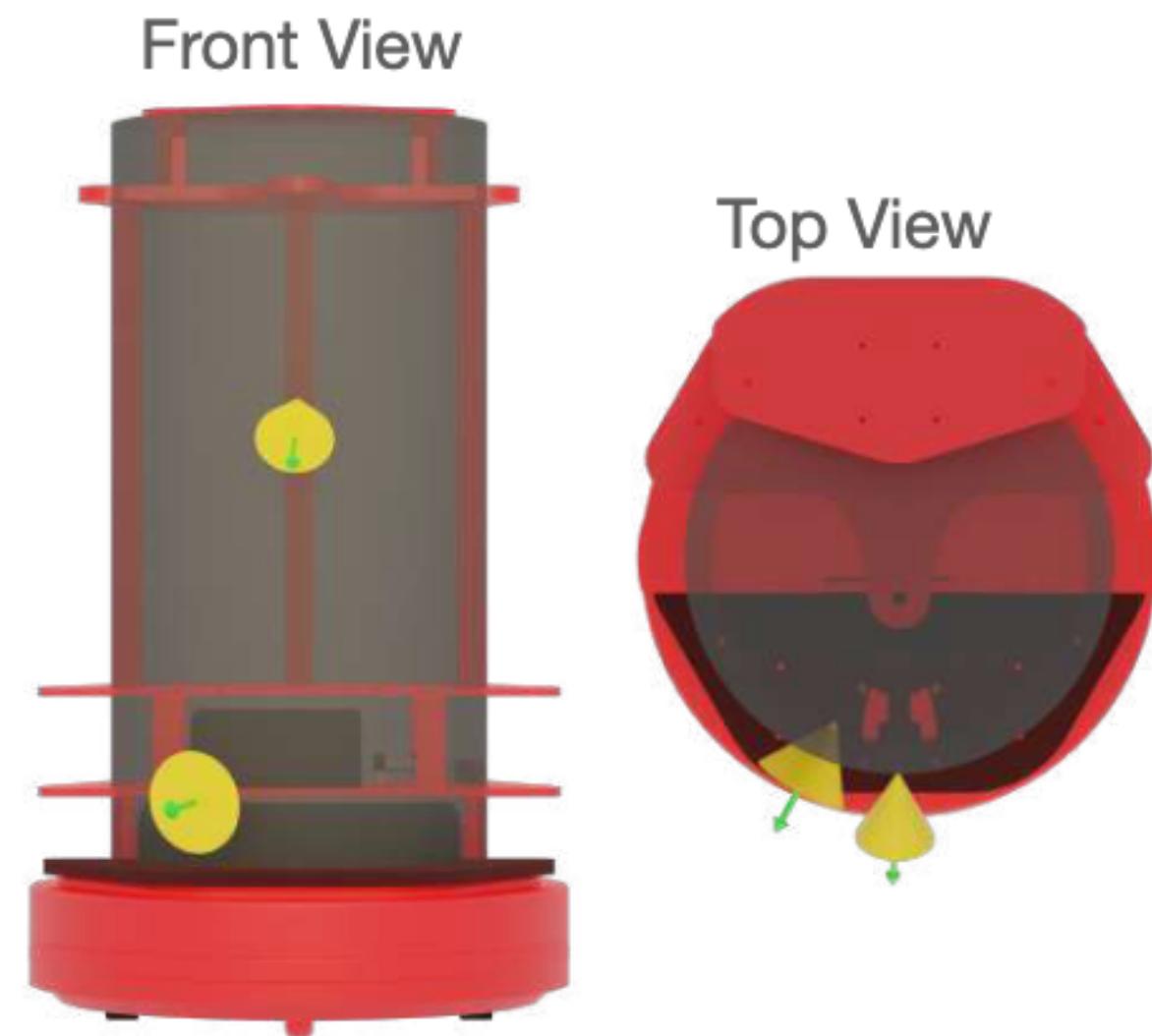


Task: Navigation from Constrained Vision

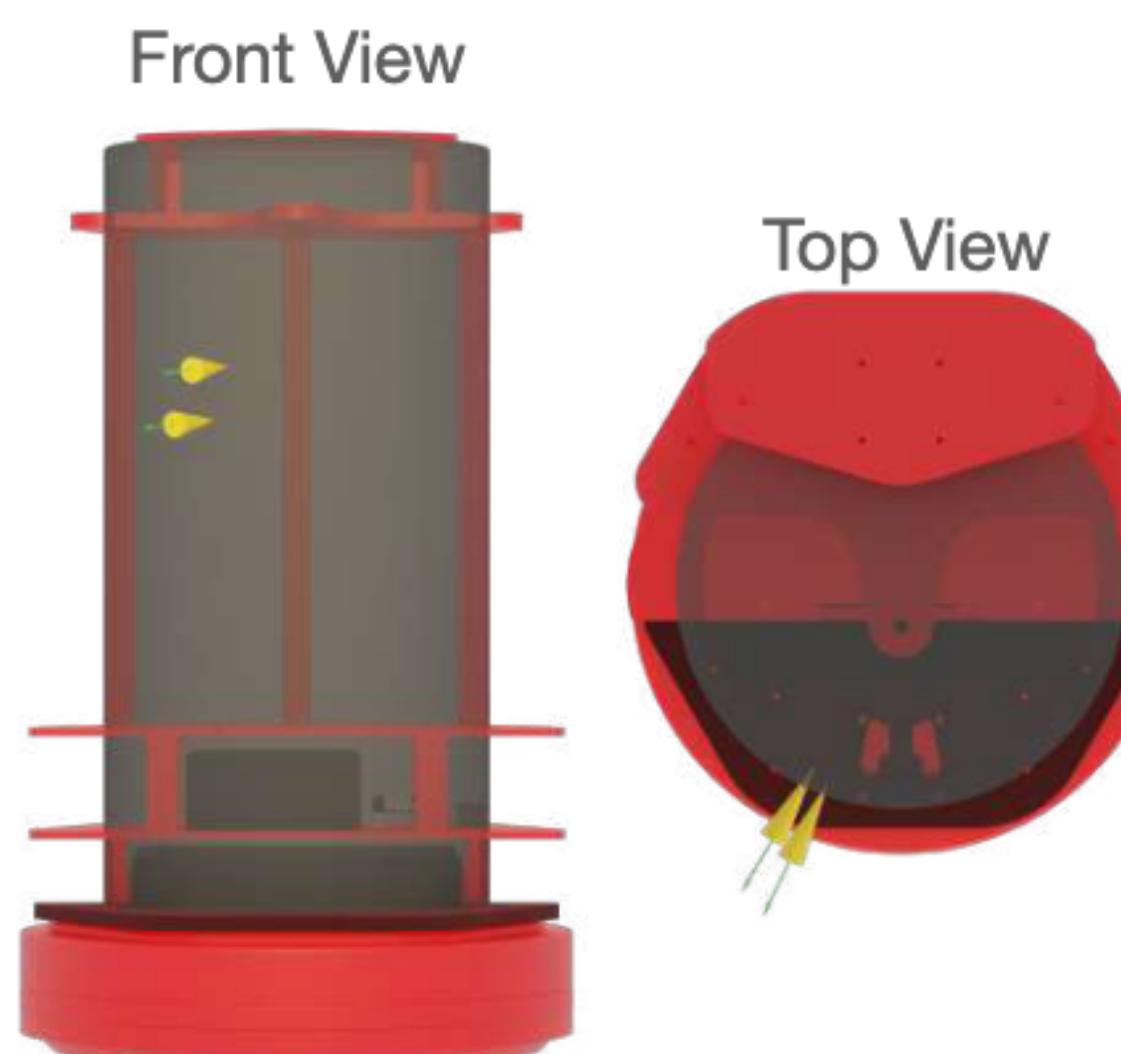


Task: Navigation from Constrained Vision

PointNav



TargetNav



Task: Language Conditioned Navigation



Quiz Time!

Go to Gradescope and take the **Lab09 Quiz**

Entry Code is **42W5EJ**

Thanks!