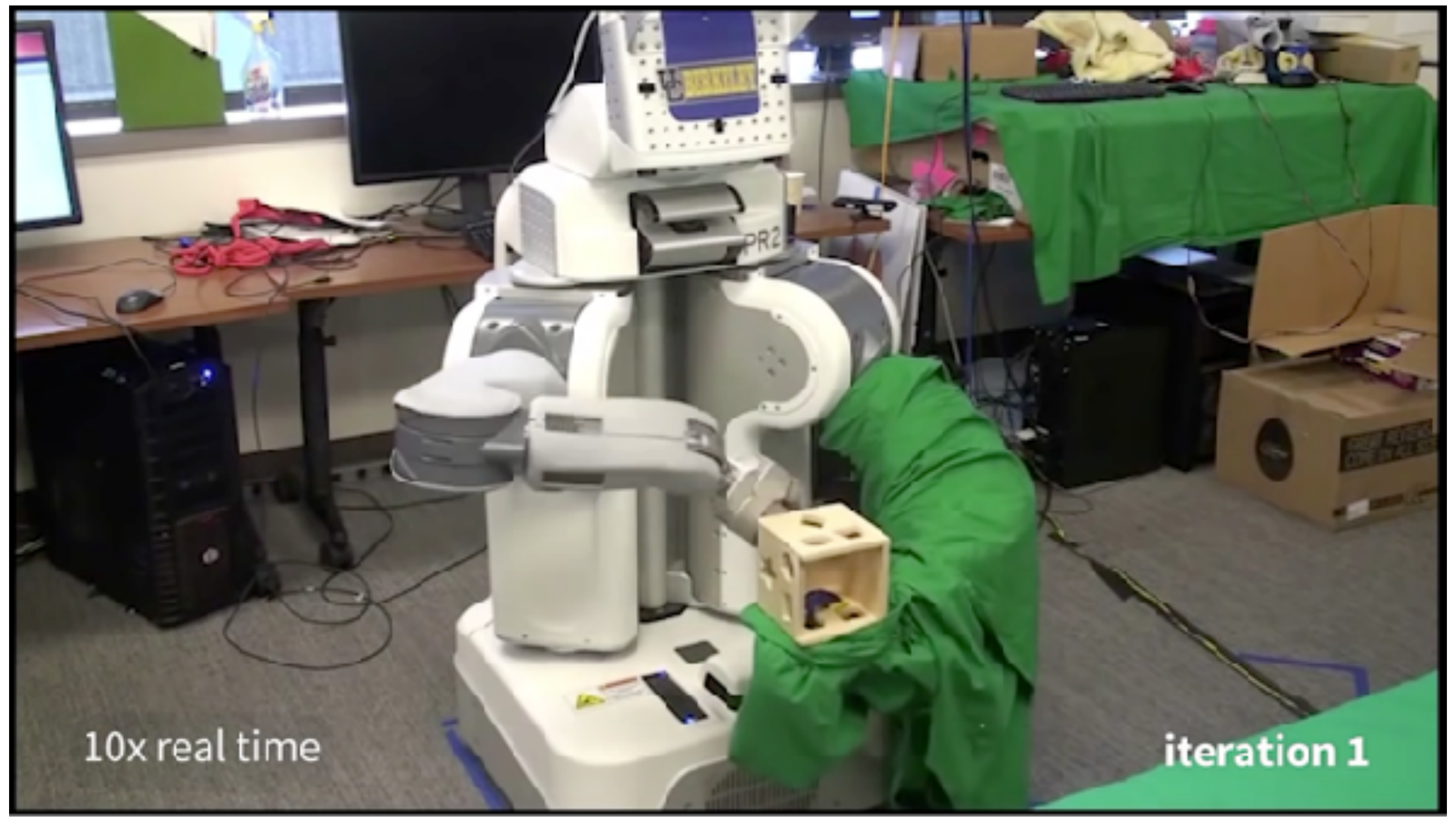
Social Learning

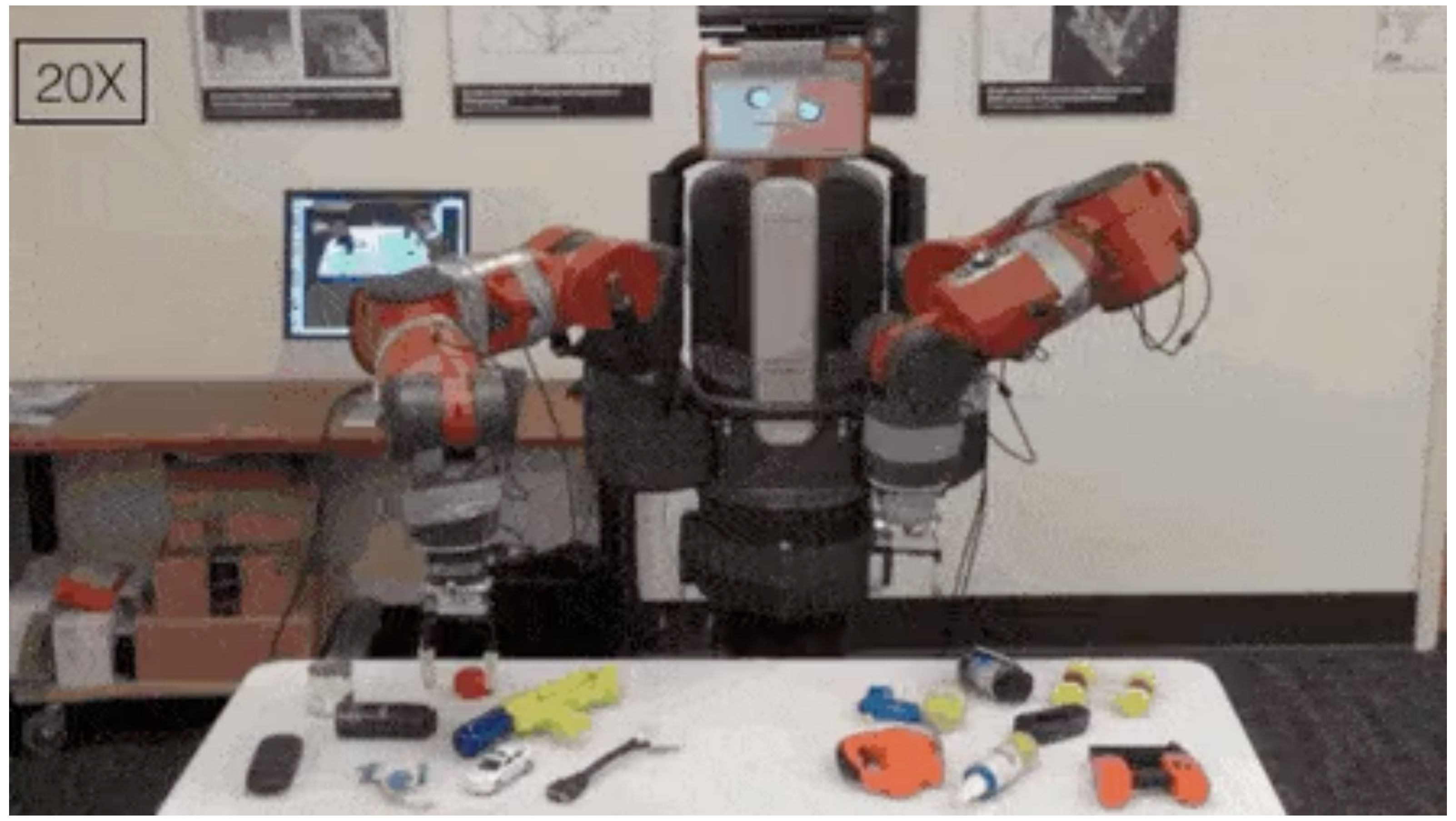
Saurabh Gupta
UIUC

Motivation

Learning in Computer Vision

Policy Learning in Robotics





How do we scale up learning for robotics?

Many different answers, but today, scaling up robot learning through egocentric videos.

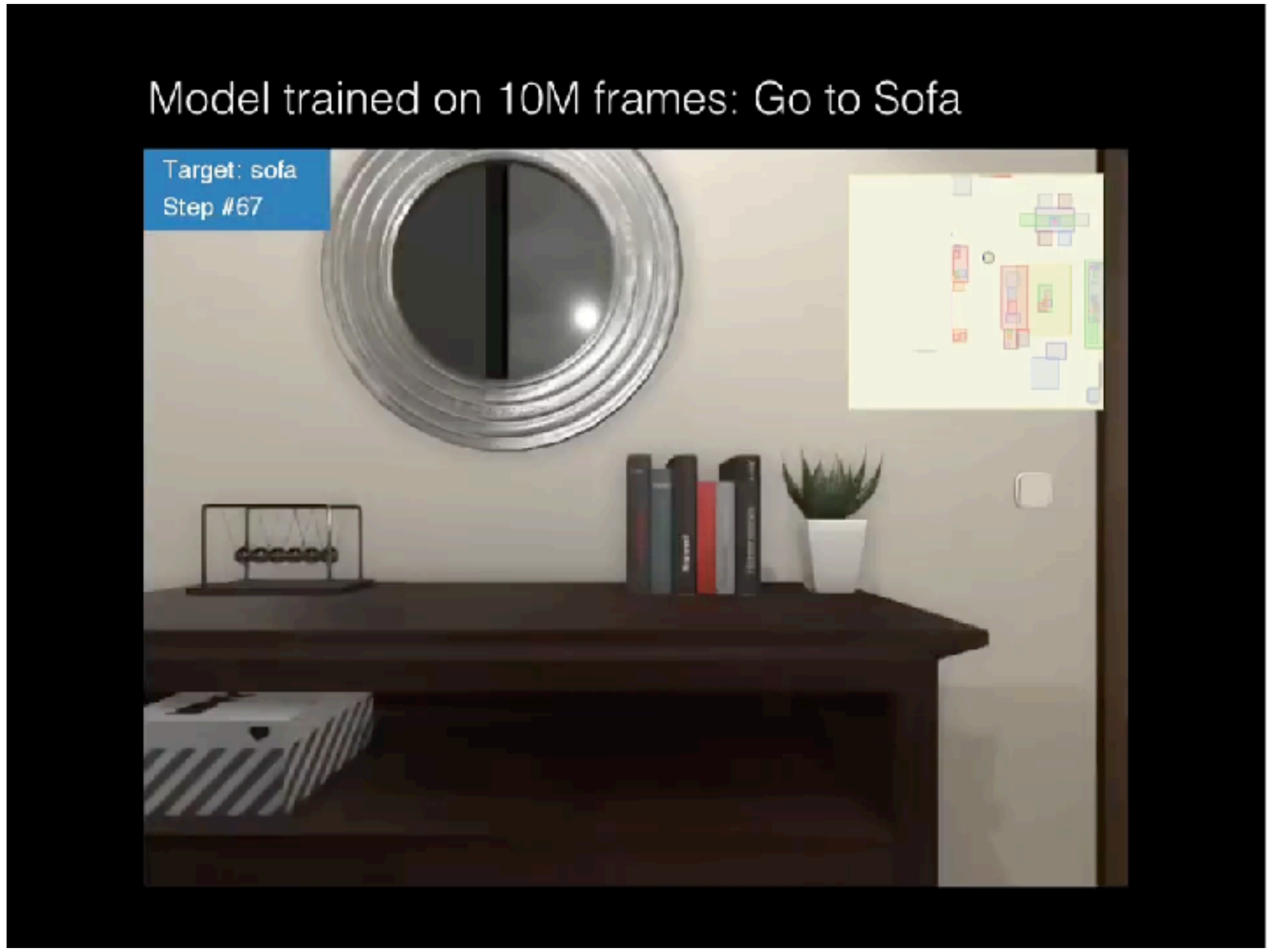
Egocentric Videos

How can we use such videos to learn policies for robots?

- We do it as adults
- Children:
  - Early imitation in children, as young as a few hours / days
  - Proto-referential imitation
- New Caledonian crows

Motivation

Policy Learning from Interaction



- Challenging to specify reward functions
- Impractically large sample complexity
- Learning signal derived solely from interaction
- Poor generalization due to lack of visual diversity in training, sim2real transfer

How can egocentric videos aid?

- Large diversity may provide good generalization.
- Demonstrations may directly show how to solve long horizon tasks.
- Depict what the world is like, and how it works.

Motivation

How can egocentric videos aid?

- Large diversity may provide good generalization.
- Demonstrations may directly show how to solve long horizon tasks.
- Depict what the world is like, and how it works.

Social Learning

What all can we learn?

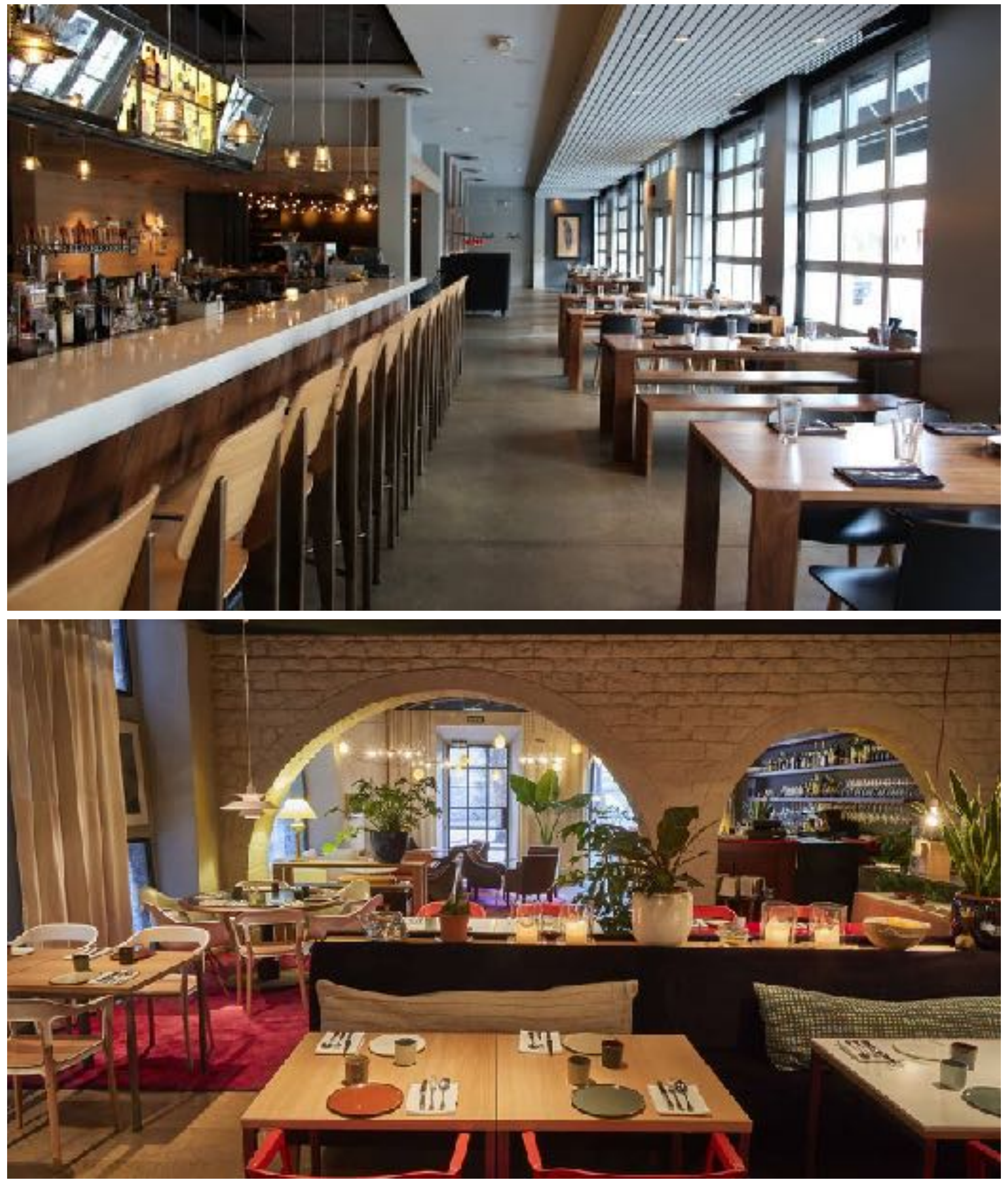
High-level plans

YouCook2 Dataset: http://youcook2.eecs.umich.edu/

Social Learning

What all can we learn?

High-level semantic priors



Finding a bathroom in a new restaurant

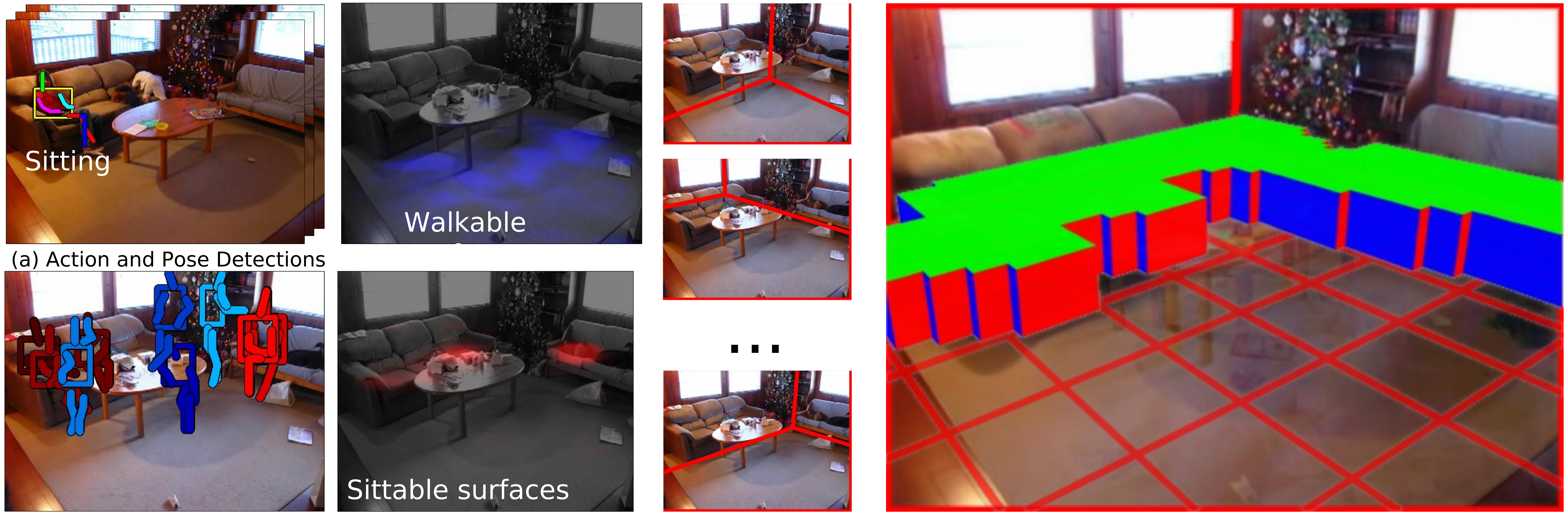
Learn by mining spatial co-occurrences from online videos

Semantic Visual Navigation by Watching YouTube Videos Matthew Chang, Arjun Gupta, Saurabh Gupta. NeurIPS, 2020

Social Learning

What all can we learn?

Environmental affordances (third-person time-lapse)



(a) Action and Pose Detections

(b) Poses (potentially aggregated over time)

(c) Estimates of functional surfaces

(d) 3D room geometry hypotheses

(e) Final 3D scene understanding

People Watching: Human Actions as a Cue for Single View Geometry.
D. F. Fouhey, V. Delaitre, A. Gupta, A. Efros, I. Laptev, and J. Sivic. ECCV 2012.

Social Learning

What all can we learn?

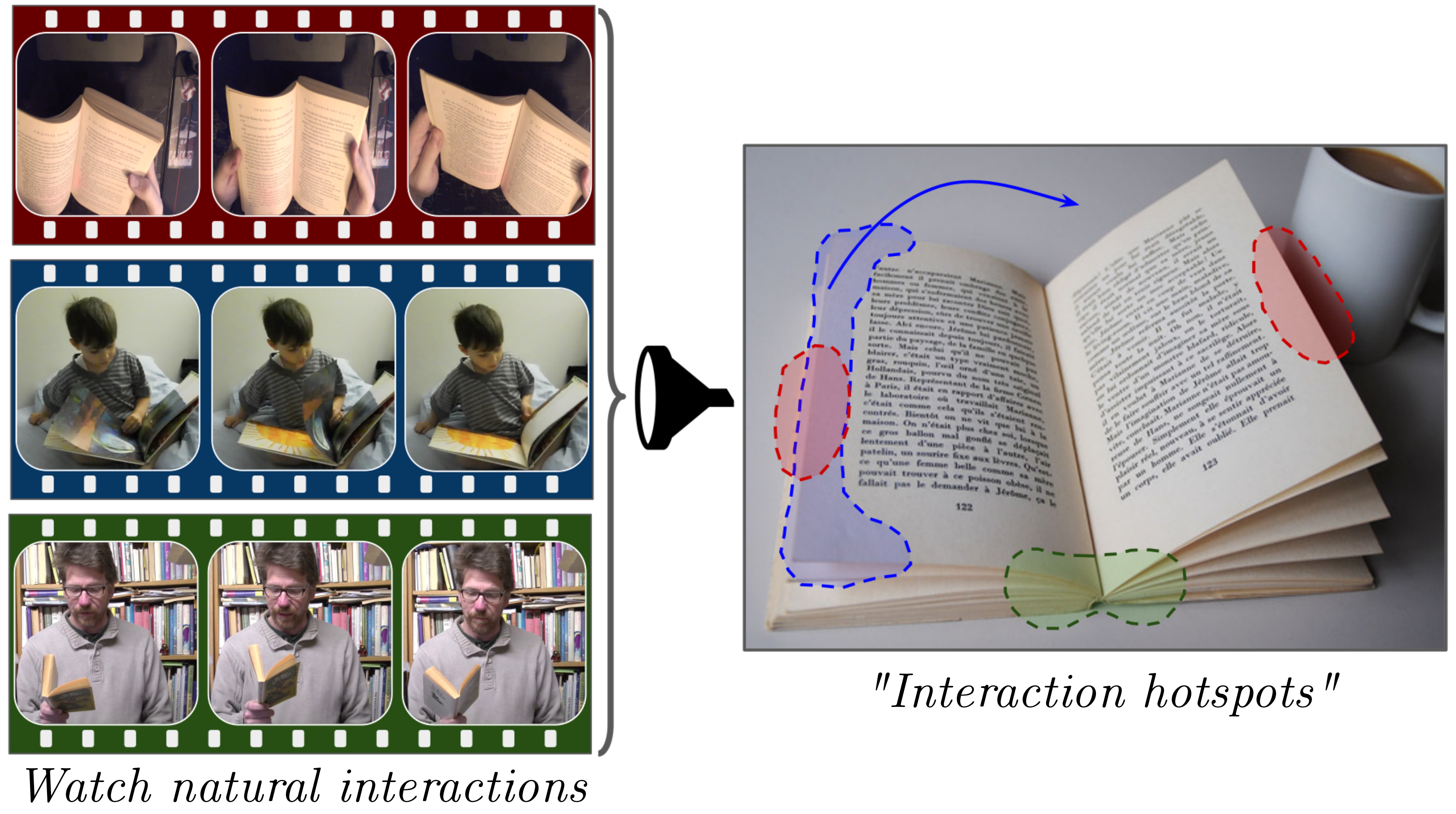
Environmental affordances
(first-person videos)

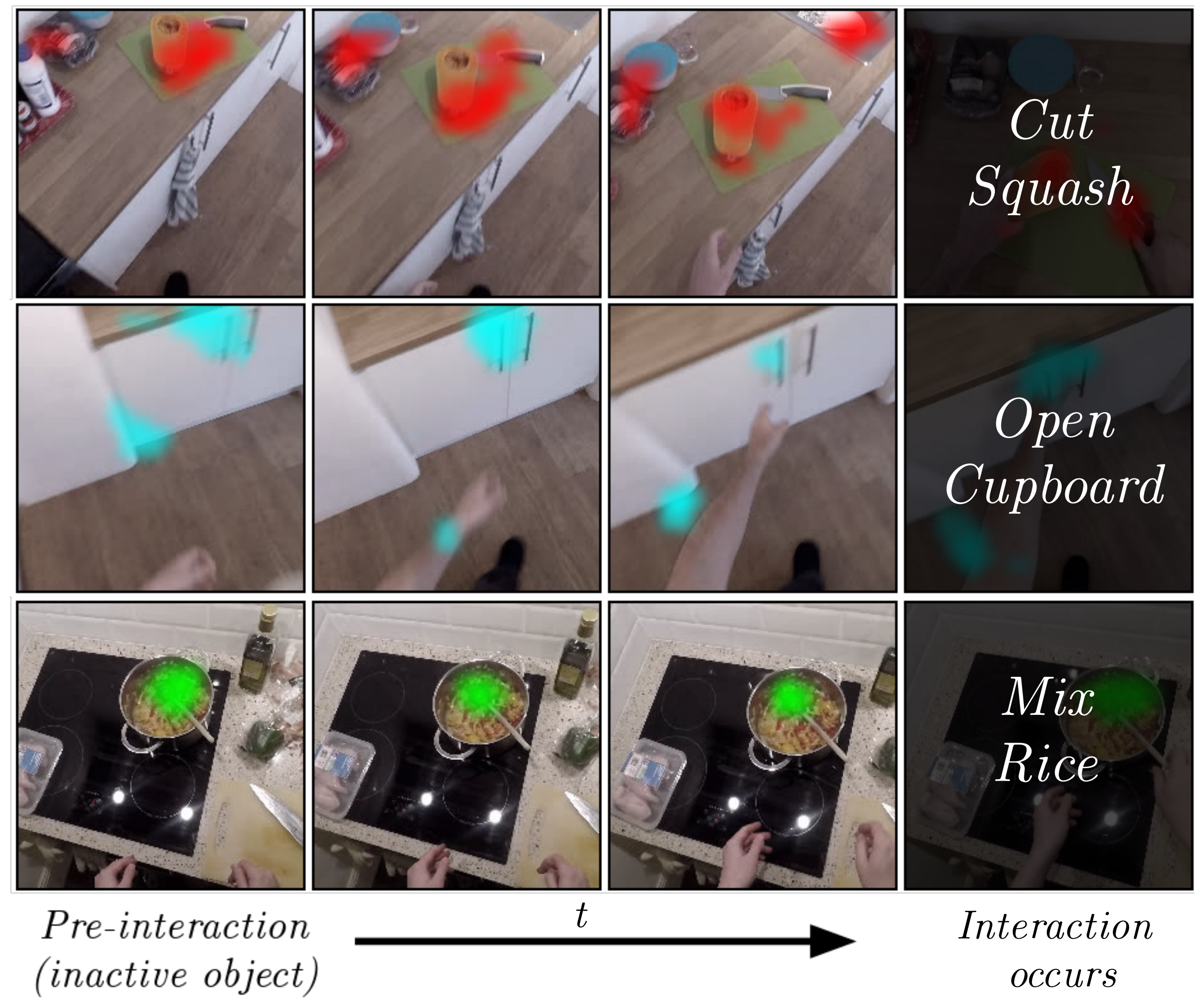
Learning Navigation Subroutines by Watching Egocentric Videos. A. Kumar, S. Gupta, J. Malik. CoRL 2019.

Social Learning

What all can we learn?

Priors for where to interact



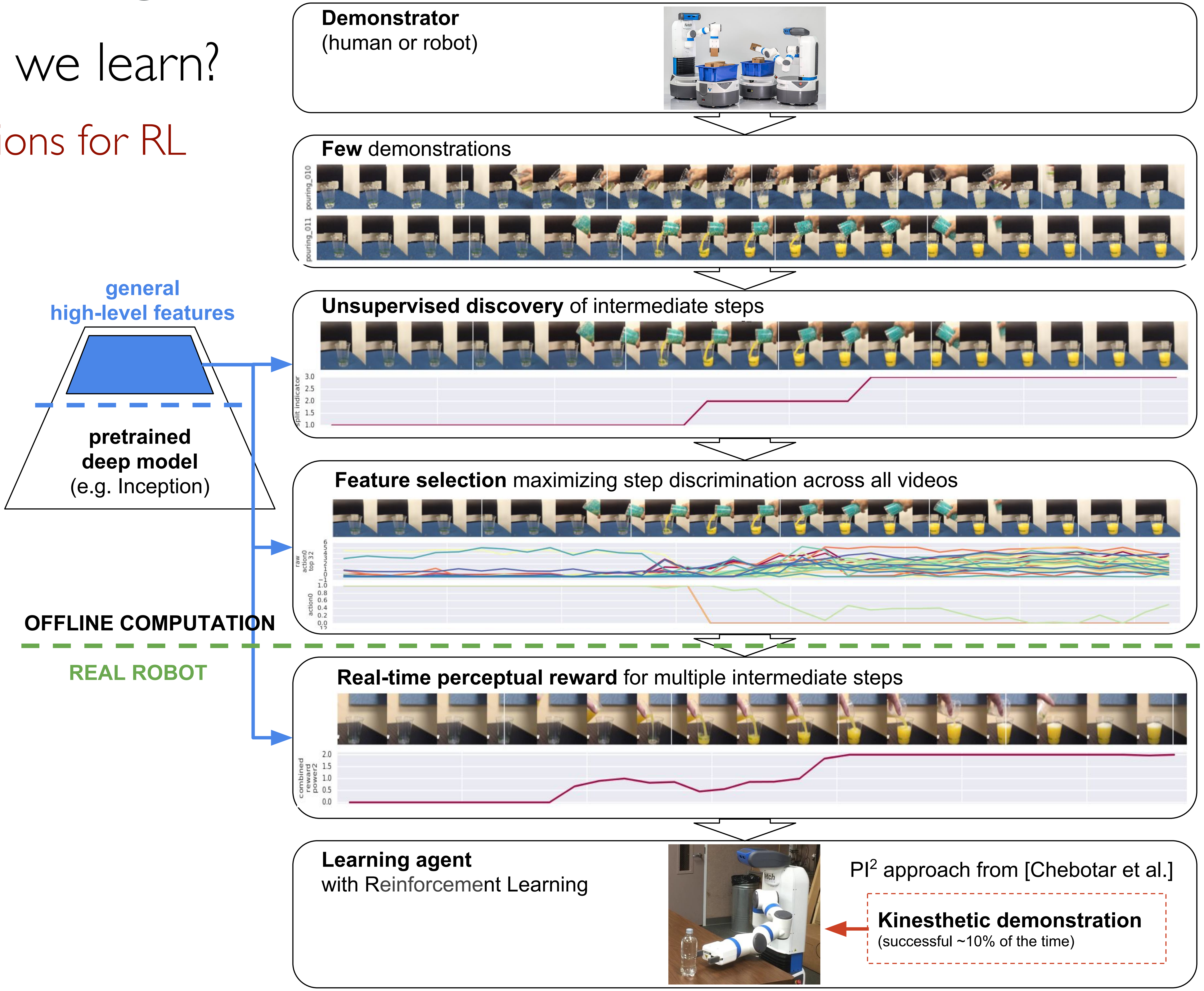


Grounded Human-Object Interaction Hotspots from Video
Tushar Nagarajan, Christoph Feichtenhofer, Kristen Grauman ICCV 2019.

Social Learning

What all can we learn?

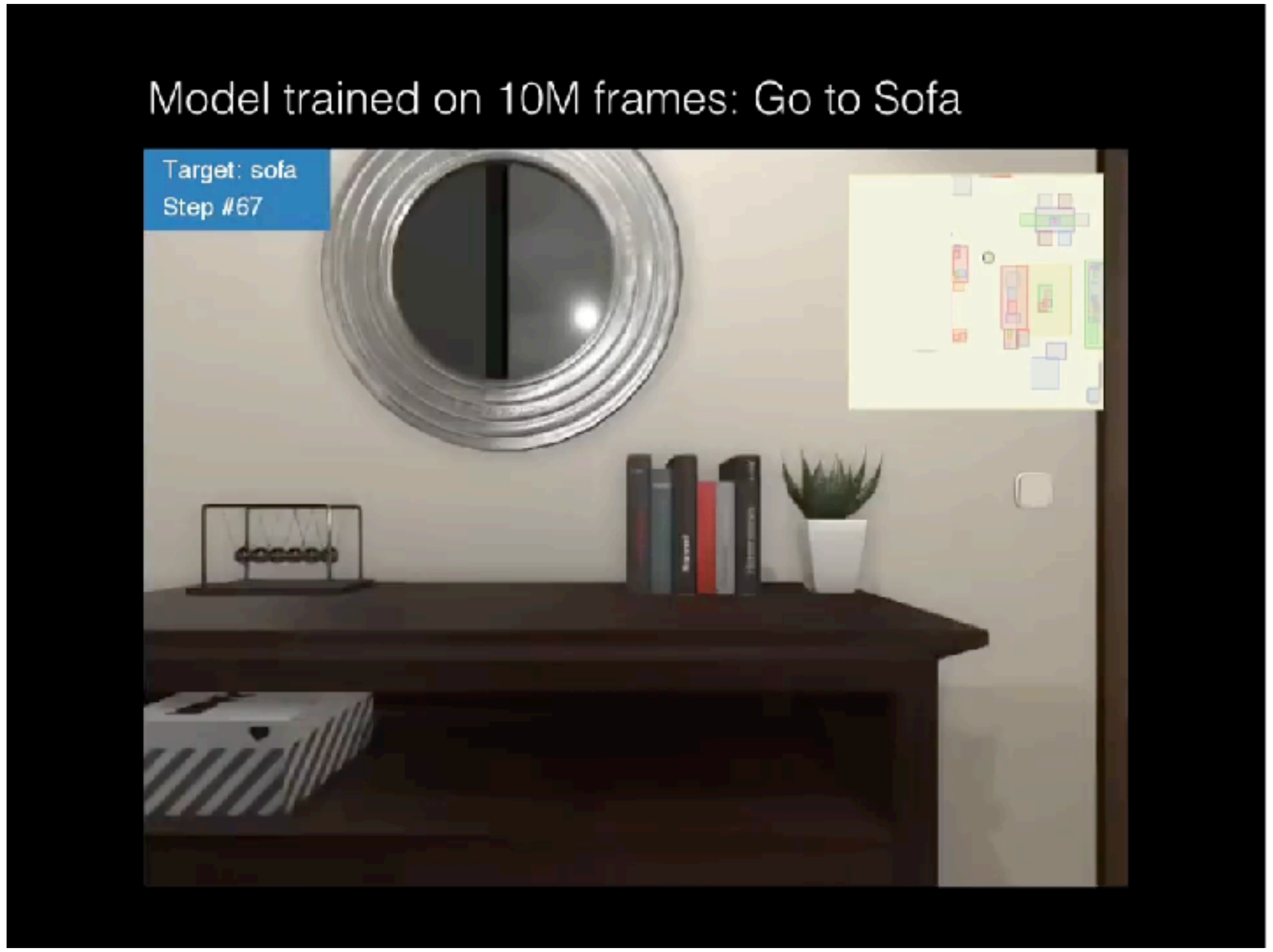
Reward functions for RL



Unsupervised Perceptual Rewards for Imitation Learning
Pierre Sermanet, Kelvin Xu, Sergey Levine. R:SS 2017

Motivation

Policy Learning from Interaction



- Challenging to specify reward functions
- Impractically large sample complexity
- Learning signal derived solely from interaction
- Poor generalization due to lack of visual diversity in training, sim2real transfer

How can egocentric videos aid?

- Large diversity may provide good generalization.
- Demonstrations may directly show how to solve long horizon tasks.
- Depict what the world is like, and how it works.

Discussion

What are ways in which social learning could be hard?

Motivation

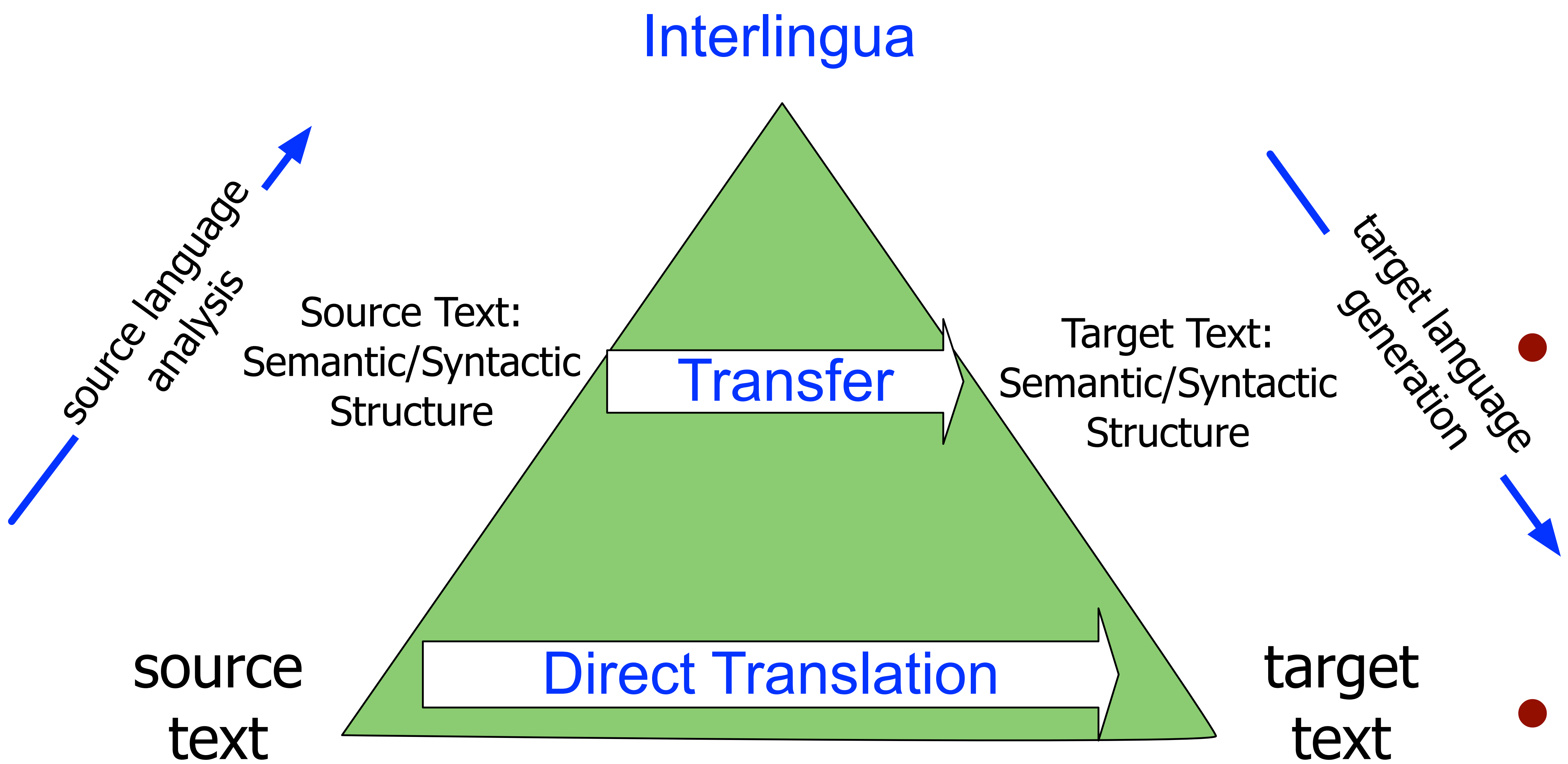
How can egocentric videos aid?

- Large diversity may provide good generalization.
- Demonstrations may directly show how to solve long horizon tasks.
- Depict what the world is like, and how it works.

However,

- Videos don't come with action labels
- Goals and intents are not known
- Depicted trajectories may be sub-optimal
- Embodiment gap (sensors / actions / capabilities)
- Only showcase positive data

"Opportunistic" Learning



The Vauquois triangle for machine translation

- *Interlingua* (text to abstract meaning space, back to text)
- *Transfer* (text to parse tree, transform tree, generate text)
- *Direct Translation* (translate word-by-word or phrase-by-phrase)

- *Increasing depth of understanding for analysis and generation*
- *Decreasing amount of transfer knowledge needed*

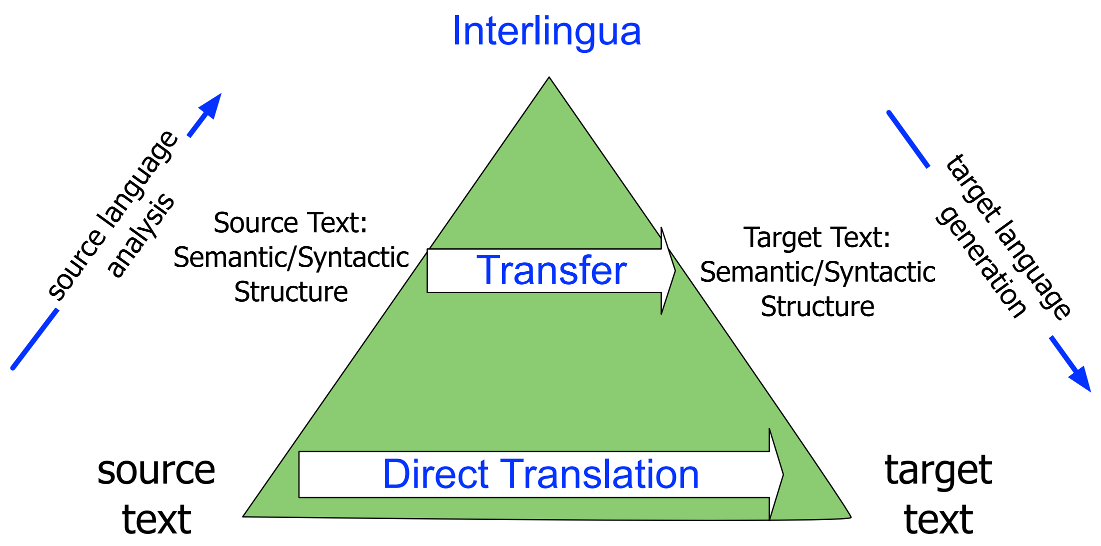
Figure from Jurafsky and Martin

“Opportunistic” Learning

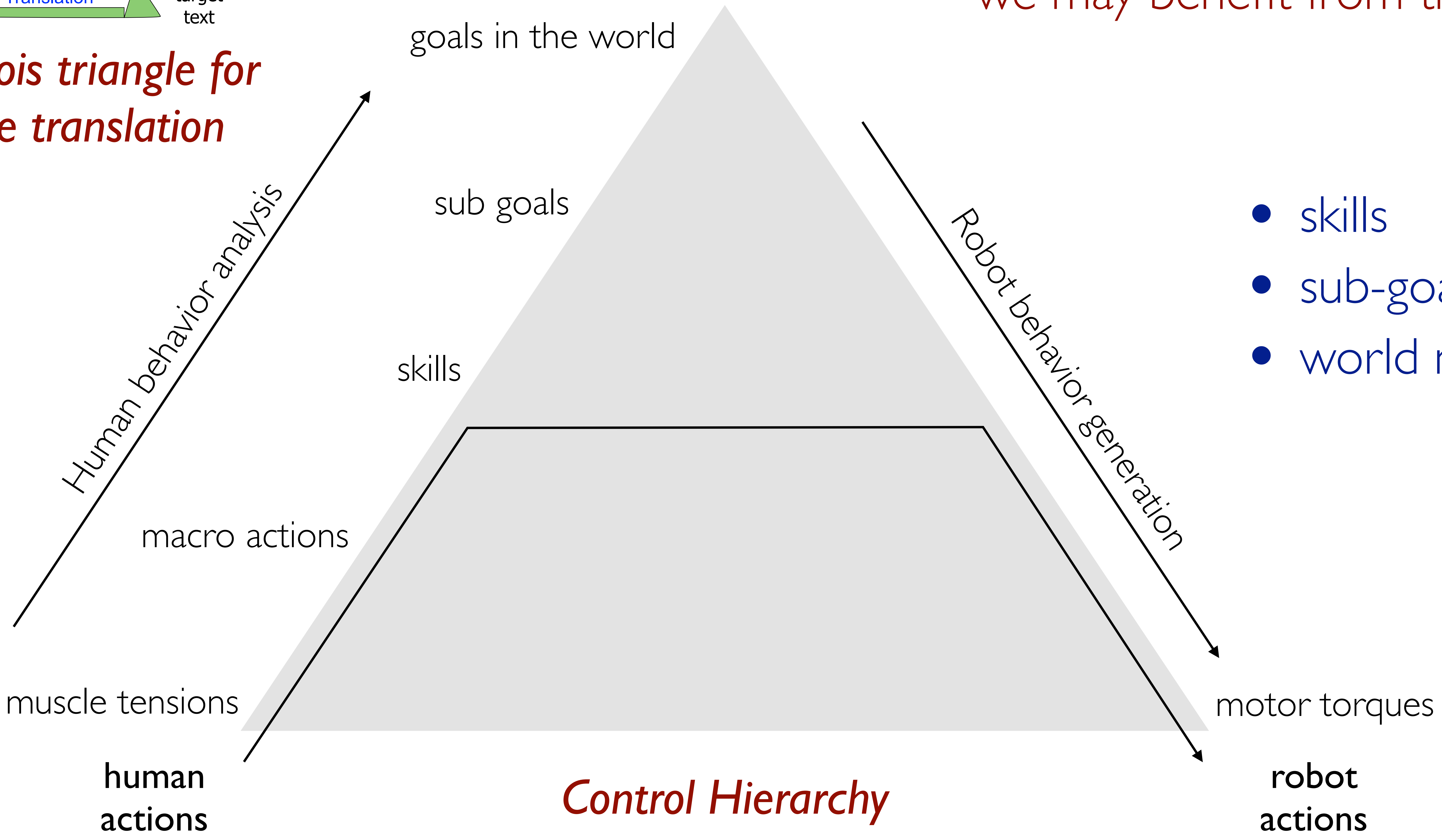
Depending on the amount of gap between:

- goals,
- embodiment,
- what we can observe in videos

we may benefit from transfer at different levels.



The Vauquois triangle for machine translation



Control Hierarchy

- skills
- sub-goals
- world models

Learning Navigation Subroutines From Egocentric Videos*

Ashish Kumar Saurabh Gupta Jitendra Malik

CoRL 2019

** Note that this was an earlier work, where we used rendered videos as opposed to real videos.*

Problem Statement

Input: Egocentric Video

Output: Low-level Skills

Robot w/camera

Skill: Go down hallway

Skill: Go around obstacles

Skill: Go through door

Challenges

- Videos don't come with action labels

  $\Longrightarrow$ Action Grounding via an Inverse Model

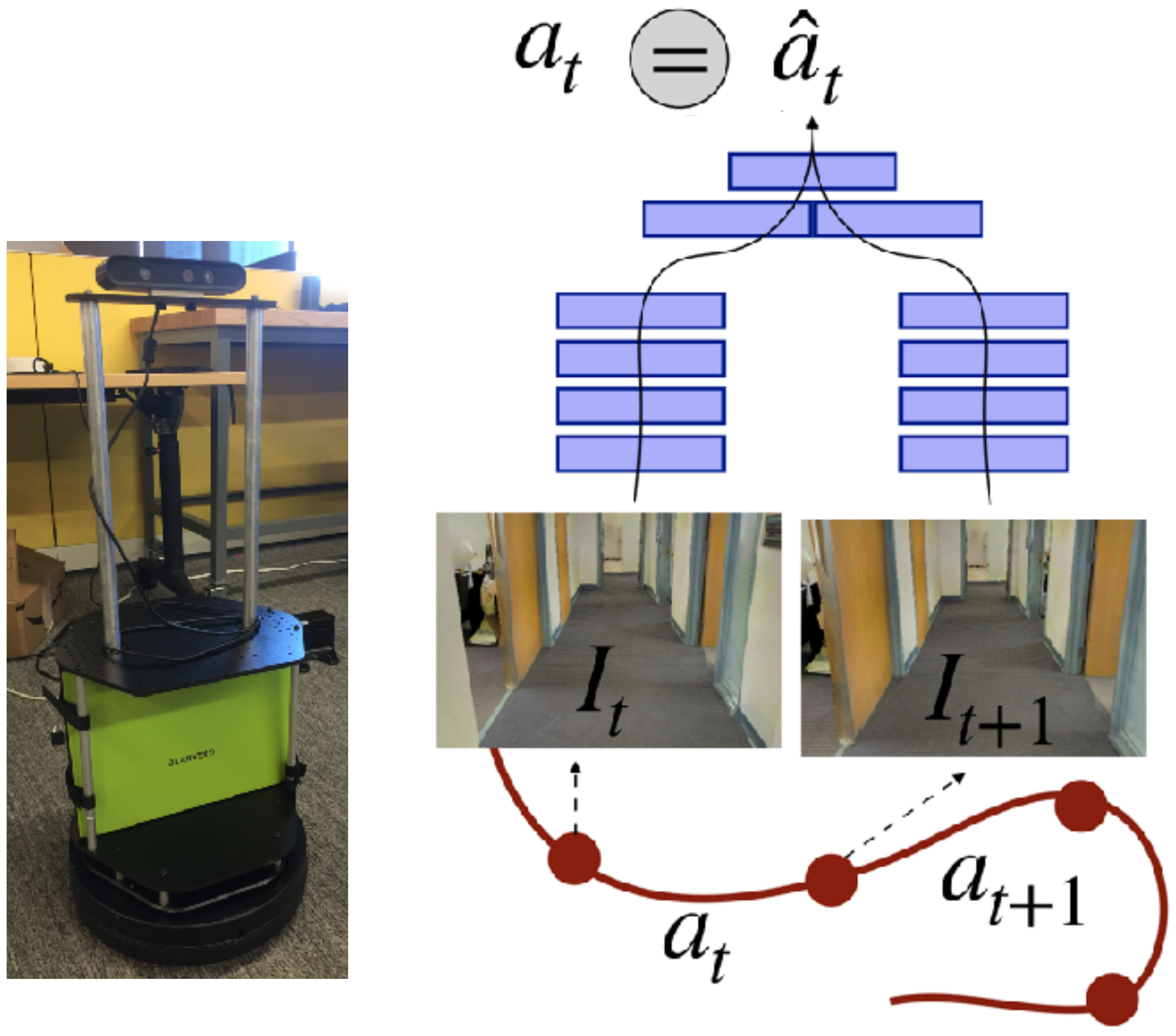
- Possible to do multiple different things, intents are not apriori known

  $\Longrightarrow$ Jointly mine for intents using a latent variable model

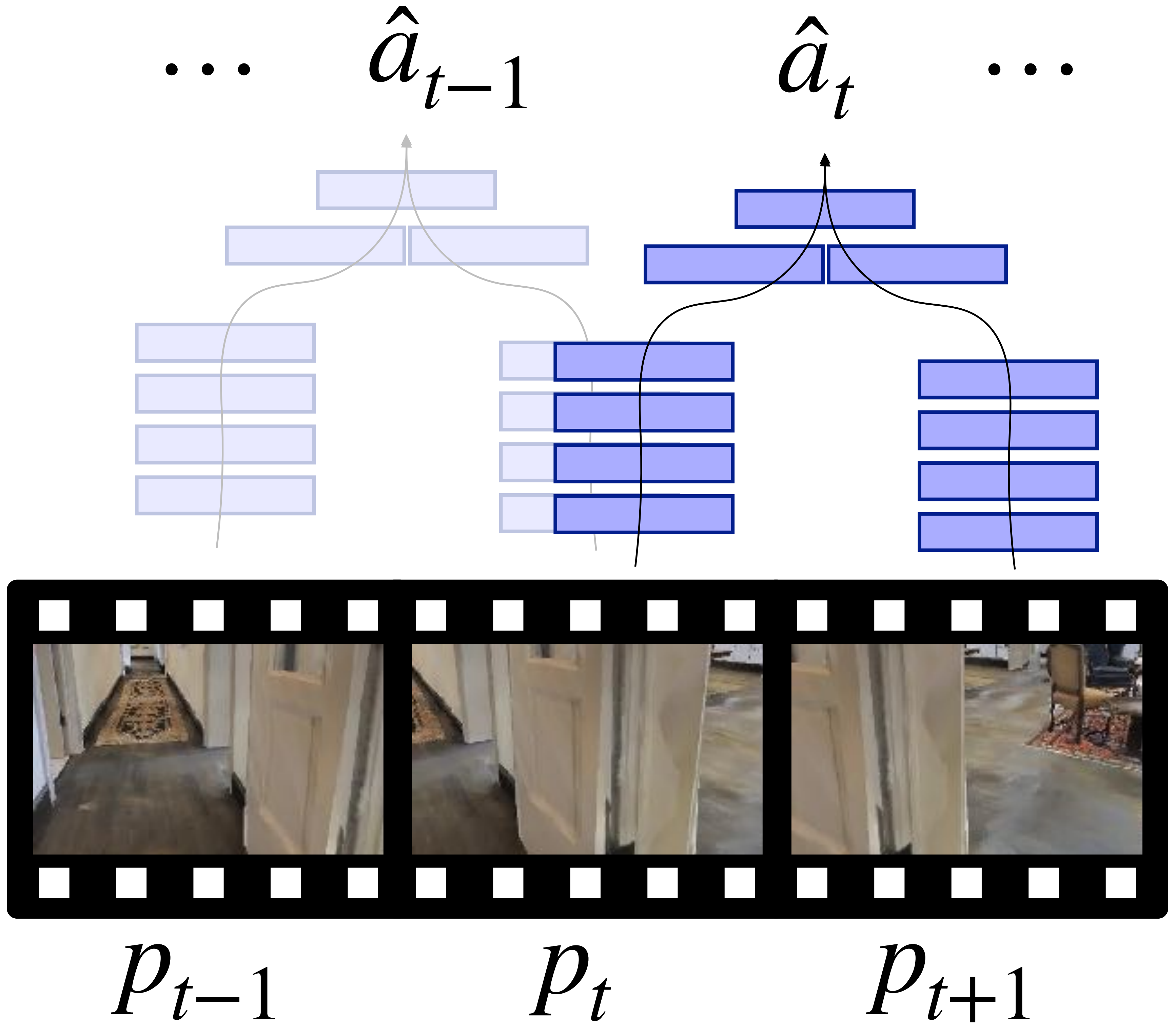
More Details

⟹ Action Grounding via an Inverse Model

a) Learning an inverse model

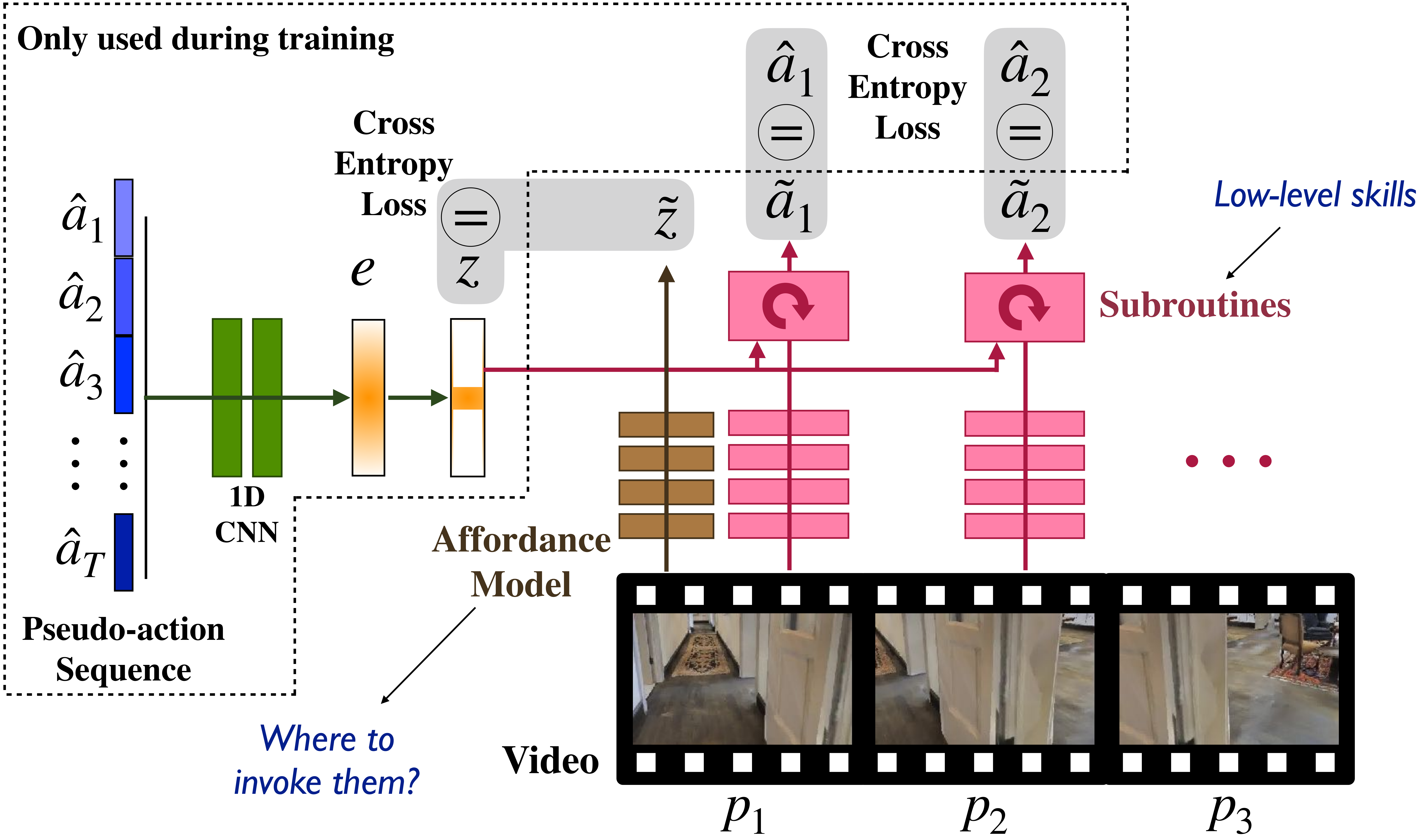


b) Pseudo-labeling Egocentric Videos

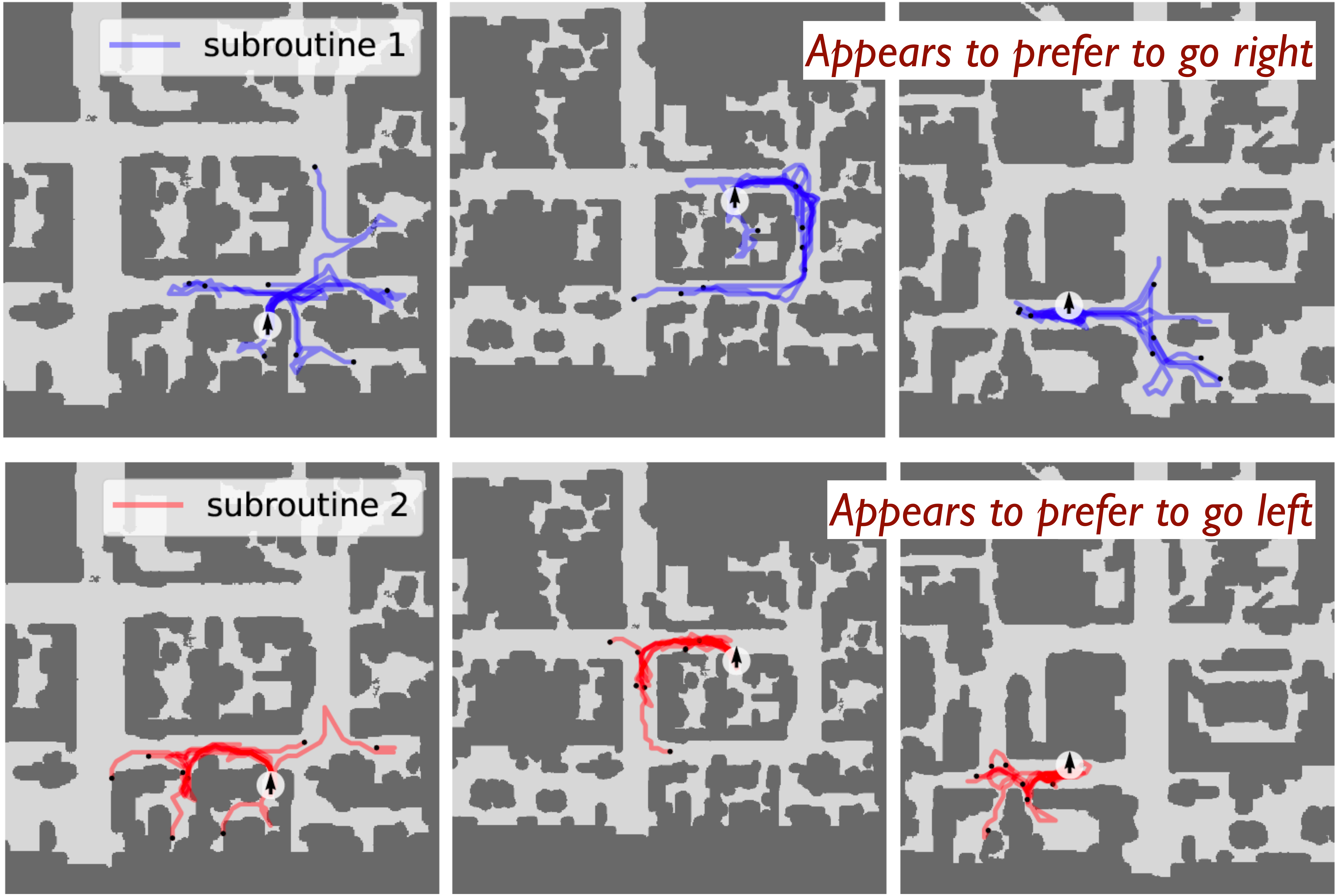


Alternative: Structure from motion

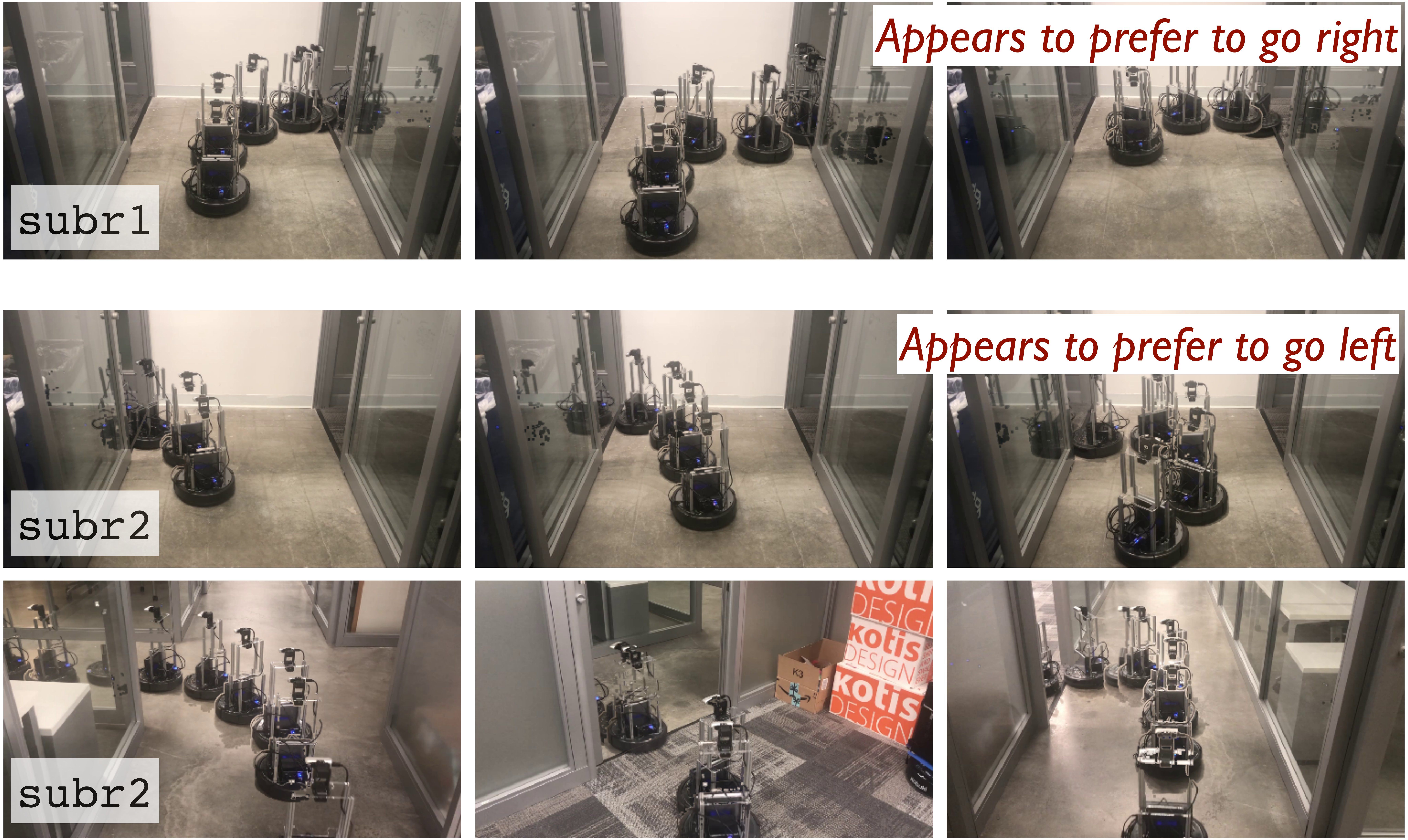
c) Learning subroutines by mining intents



Results (Subroutines)



Results (Subroutines)



Results (Affordance model)

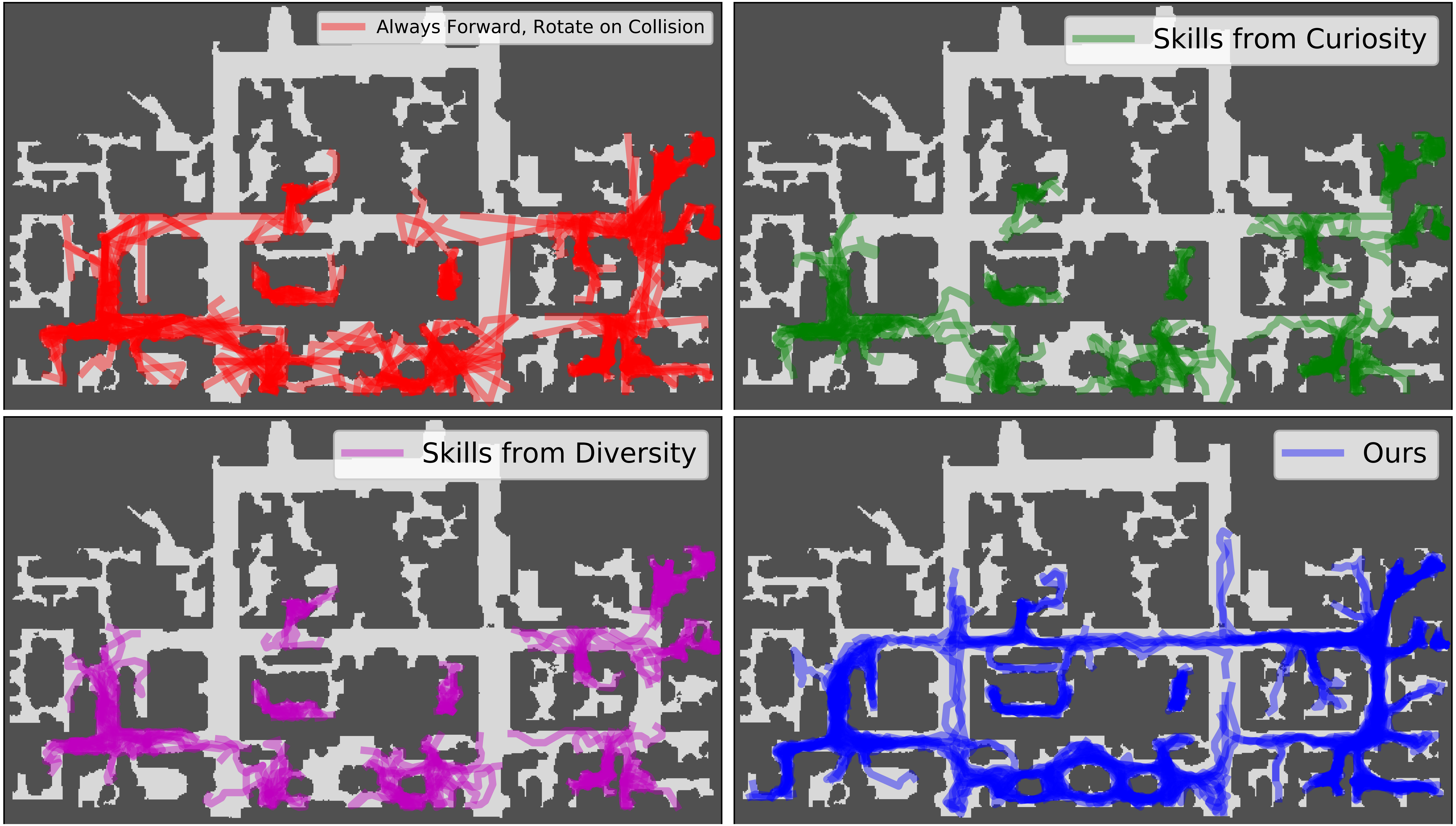
Using Subroutines and Affordances

A. As is for exploration

Method	**# Samples**	**ADT**	**Max. Dist.**	**Collision Rate (%)**
Random	0	0.96	4.34	62.5
Forward Bias Policy	0	0.66	7.19	80.2
Always Forward, Rotate on Collision	0	0.62	8.20	66.3
Skills from Diversity [13]	$10M$	0.79	4.90	64.0
Skills from Curiosity [27]	$10M$	0.83	4.36	61.3
Our (Exploration via Subroutines)	$45K$	**0.34**	**11.06**	**12.0**

B. Eysenbach, A. Gupta, J. Ibarz, and S. Levine. ICL 2019. **Diversity is all you need: Learning skills without a reward function.**
D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell. ICML 2017. **Curiosity-driven exploration by self-supervised prediction.**

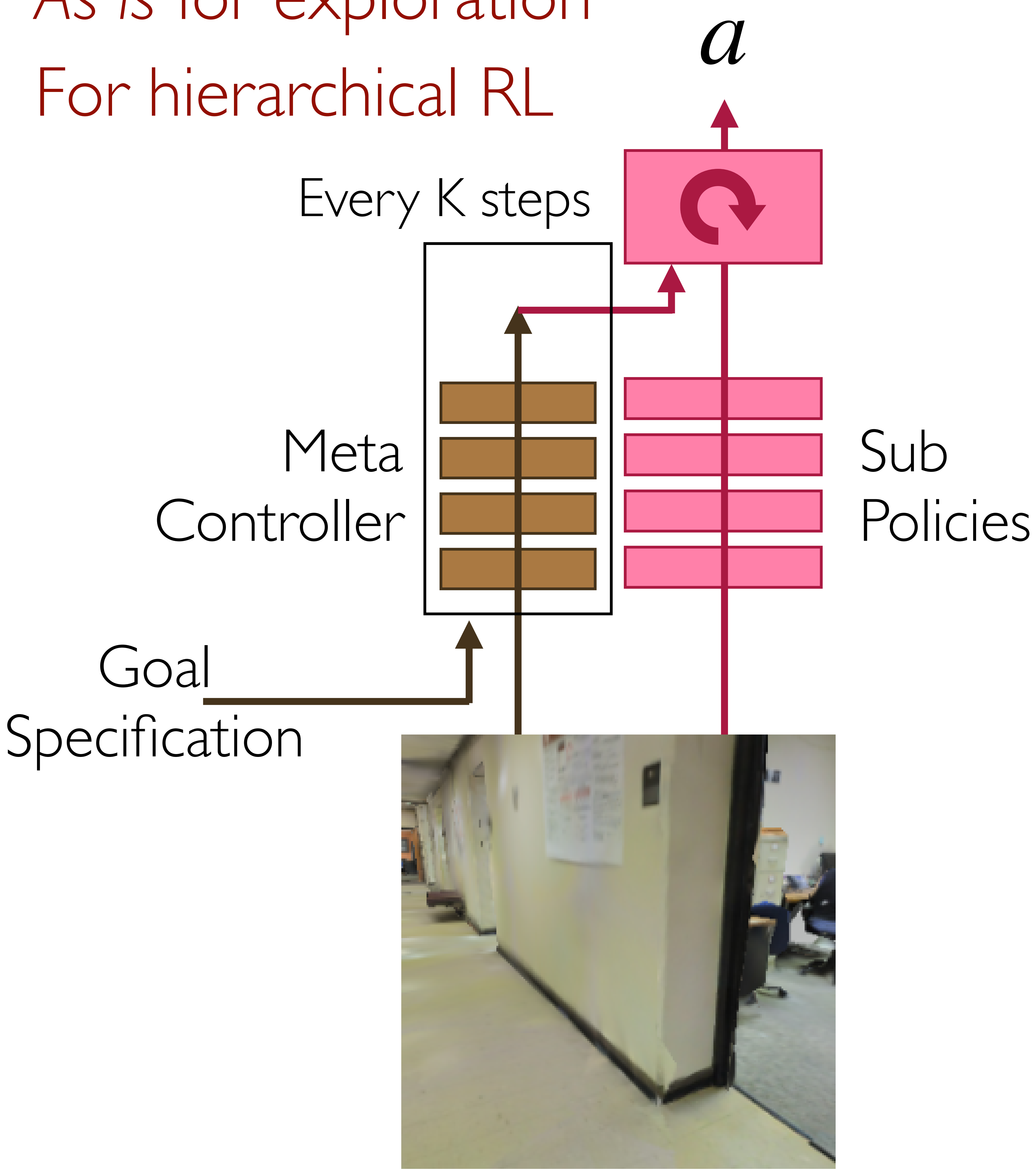
Exploration Comparisons



Using Subroutines and Affordances

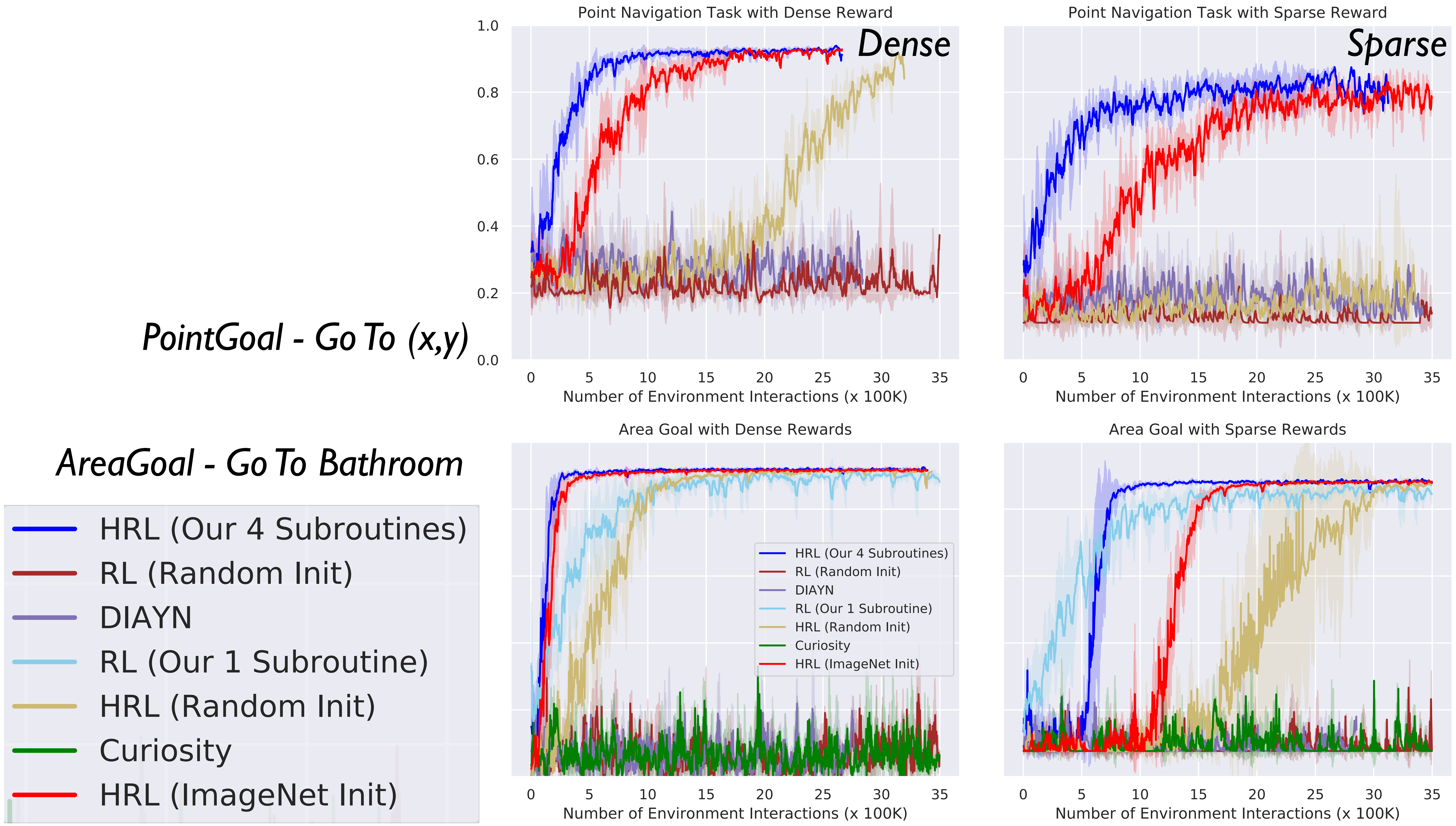
A. As is for exploration

B. For hierarchical RL



1. Use Subroutines as sub-policies.

2. Use Affordance Model to initialize meta-controller to guide meta-controller towards feasible sub-policies.



PointGoal - Go To (x,y)
AreaGoal - Go To Bathroom
Point Navigation Task with Dense Reward
Dense
Point Navigation Task with Sparse Reward
Sparse
Area Goal with Dense Rewards
Area Goal with Sparse Rewards
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Number of Environment Interactions (x 100K)
HRL (Our 4 Subroutines)
RL (Random Init)
DIAYN
RL (Our 1 Subroutine)
HRL (Random Init)
Curiosity
HRL (ImageNet Init)