Occupancy Anticipation for Efficient Exploration and Navigation

ECCV 2020 (Spotlight)

Project page: http://vision.cs.utexas.edu/projects/occupancy_anticipation

Santhosh Kumar Ramakrishnan[1,2]

Ziad Al-Halah[1]

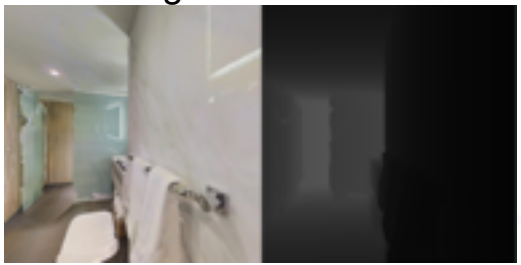
Kristen Grauman[1,2]

[1] The University of Texas at Austin

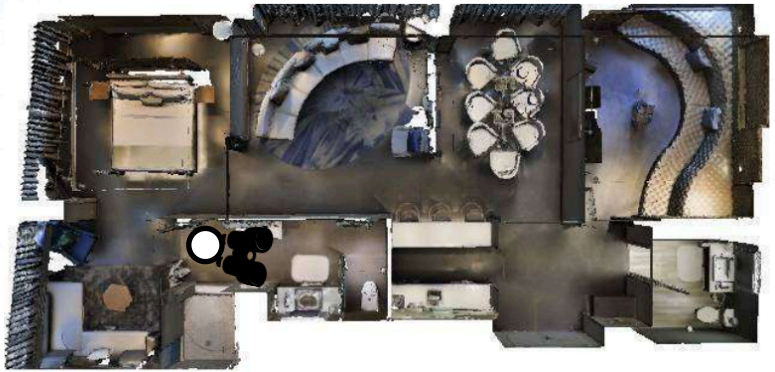
[2] Facebook AI Research

Embodied visual exploration

Egocentric view



3D environment



Top-down occupancy map

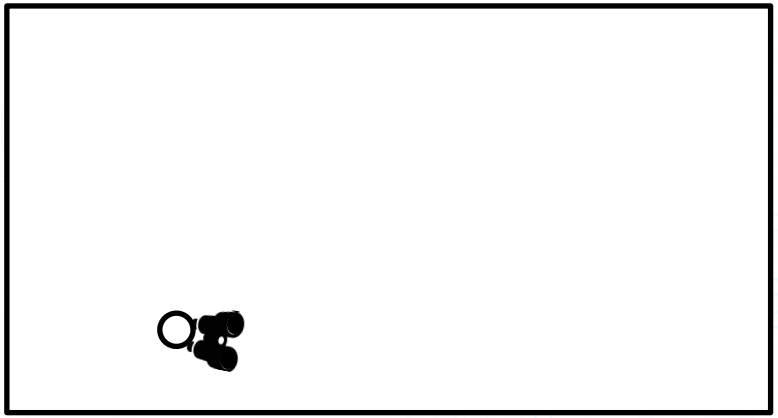
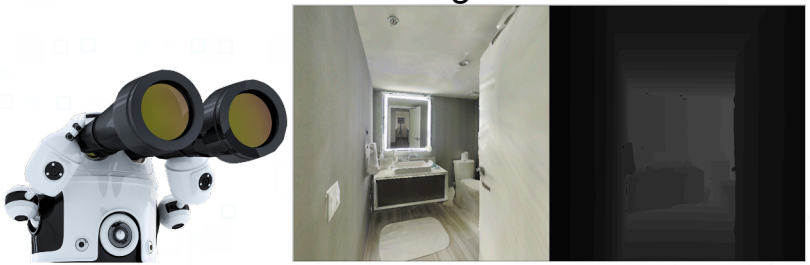


Image credits: Matterport3D dataset, Chang et al., 3DV 2017

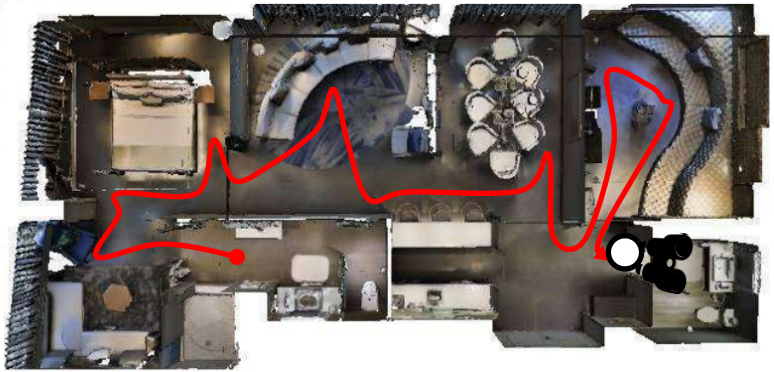
Speaker: Santhosh K. Ramakrishnan, UT Austin & Facebook AI Research

Embodied exploration

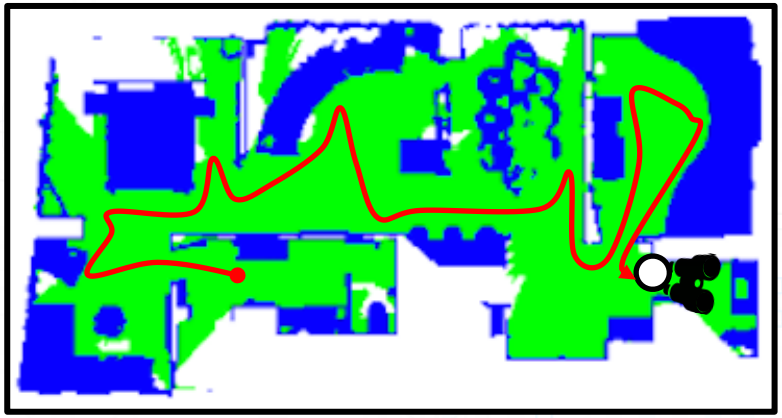
Egocentric view



3D environment



Top-down occupancy map

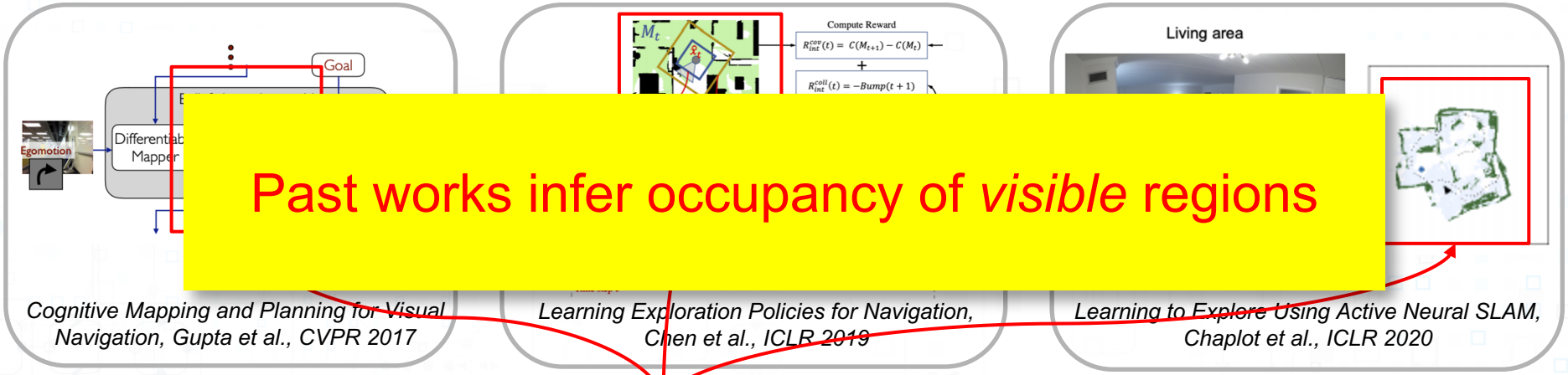


Occupied Free Unknown

Image credits: Matterport3D dataset, Chang et al., 3DV 2017

Speaker: Santhosh K. Ramakrishnan, UT Austin & Facebook AI Research

Spatial occupancy for exploration / navigation



Cognitive Mapping and Planning for Visual Navigation, Gupta et al., CVPR 2017

Learning Exploration Policies for Navigation, Chen et al., ICLR 2019

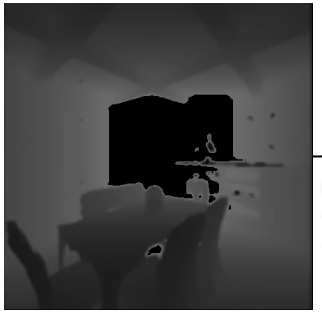
Learning to Explore Using Active Neural SLAM, Chaplot et al., ICLR 2020

Past works infer occupancy of *visible* regions

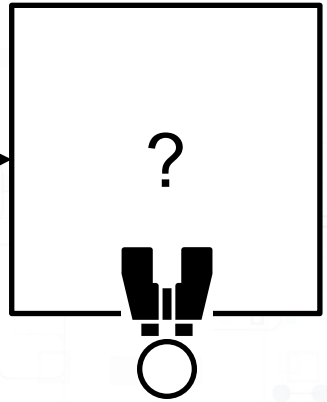
Spatial occupancy maps

IQA: Visual Question Answering in Interactive Environments, Gordon et al., CVPR 2018
Look, Listen, and Act: Towards Audio-Visual Embodied Navigation, Gan et al., ICRA 2020
Object Goal Navigation using Goal-oriented Semantic Exploration, Chaplot et al., CVPR 2020 workshop

Idea - anticipate occupancy for unseen regions

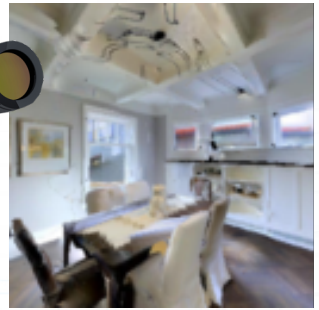


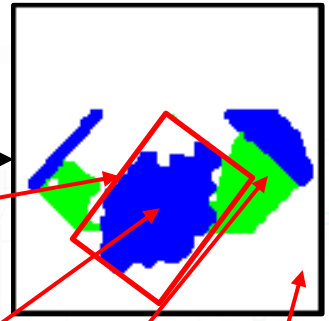
Top-down occupancy map



Occupied Free Unknown

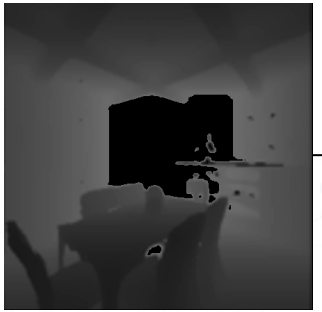
Idea - anticipate occupancy for unseen regions



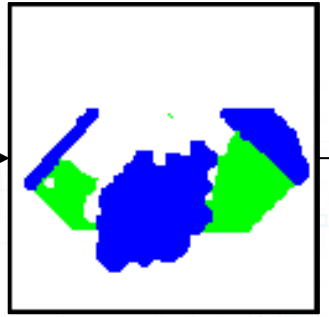


Occupied Free Unknown

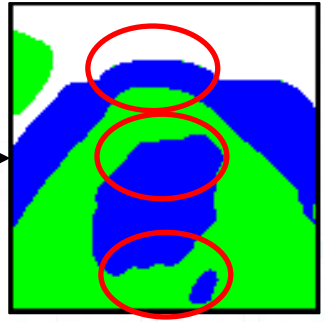
Idea - anticipate occupancy for unseen regions



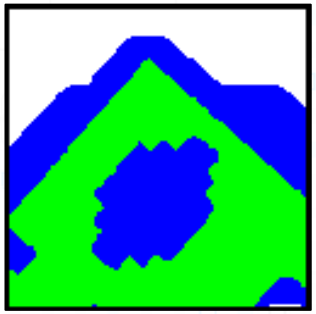
Visible occupancy
(standard)



Anticipated occupancy
(proposed)

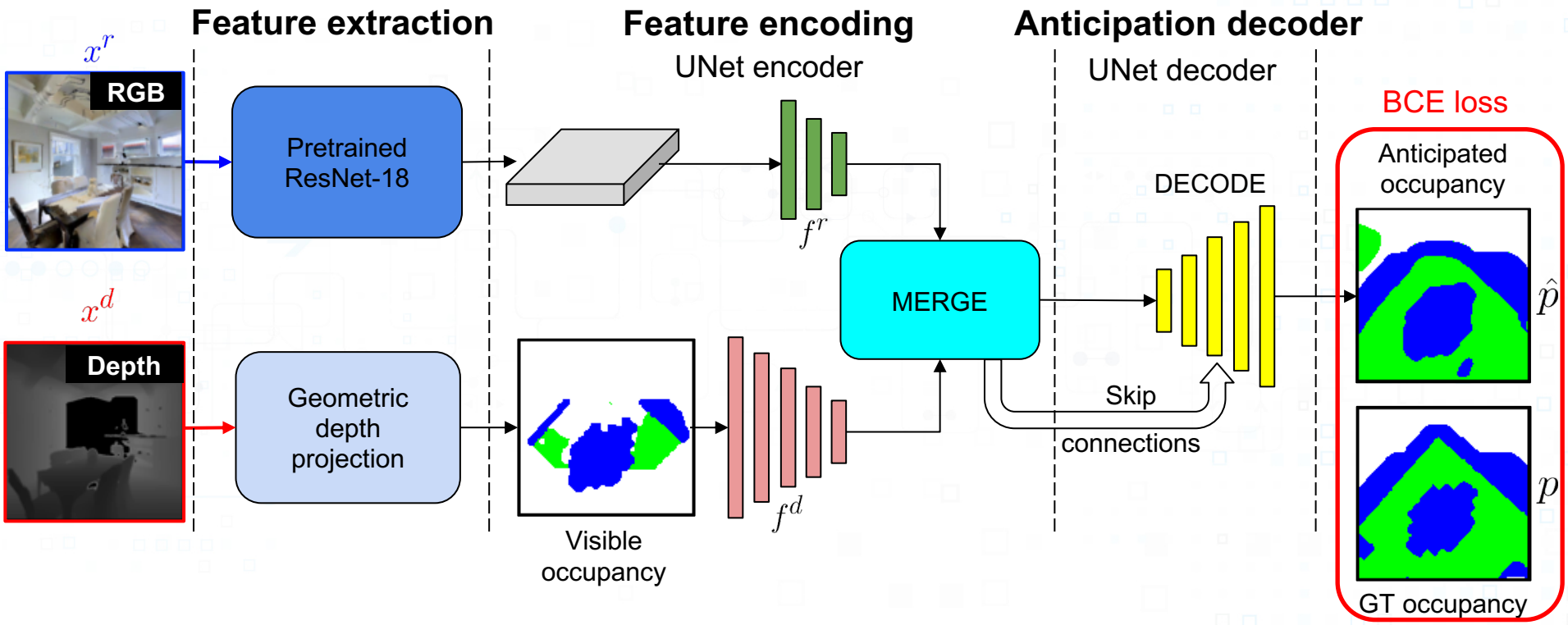


Occupied Free Unknown

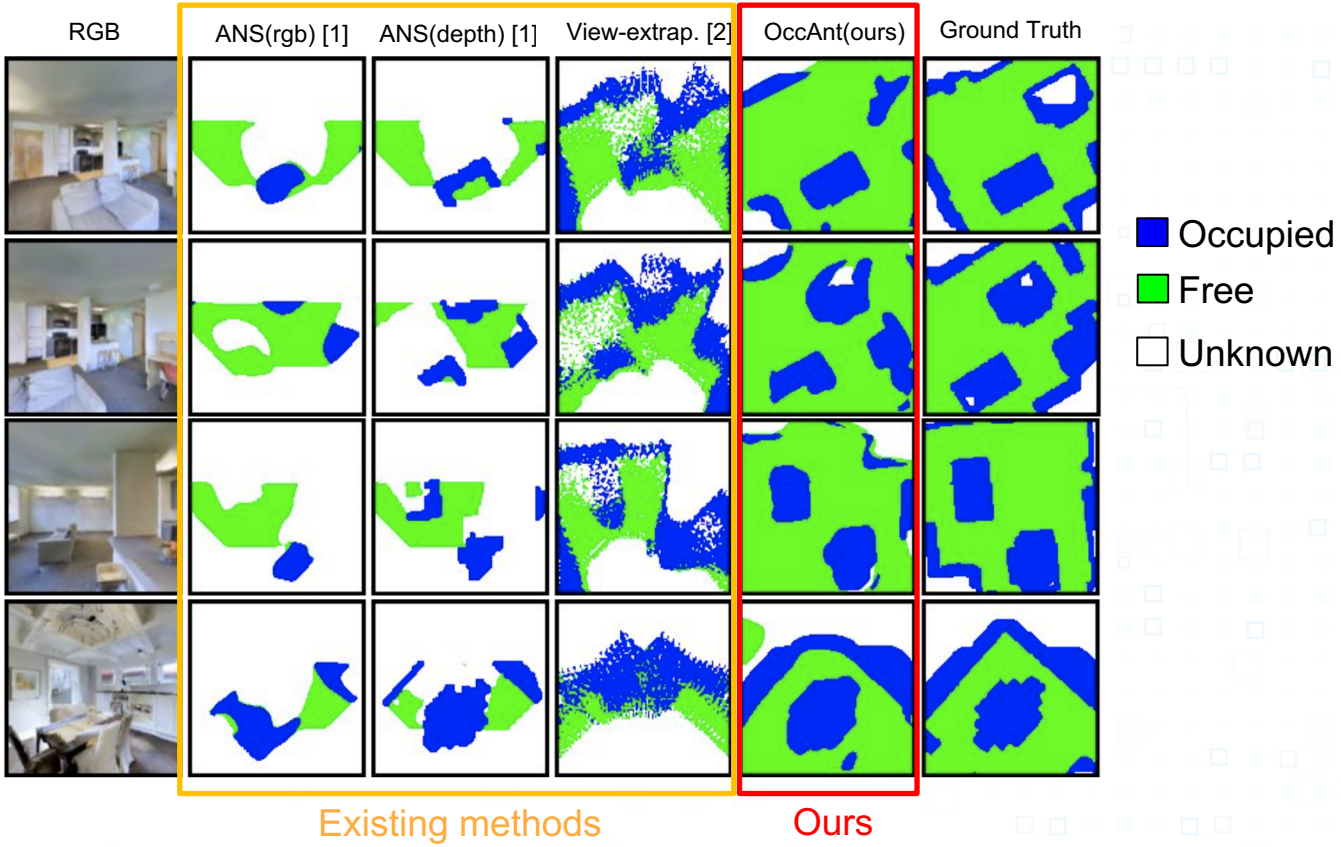


Ground-truth occupancy

Occupancy anticipation architecture



Map anticipation results



[1] Chaplot et al., ICLR 2020
[2] Song et al., CVPR 2018

Map anticipation results

Semantic and geometric context permits intelligent anticipation

Evaluation on Gibson 3D dataset

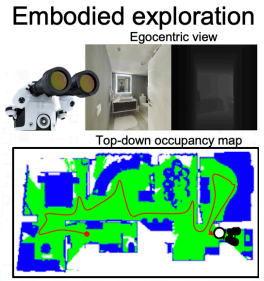
Method	IoU			F1 score		
	free	occ.	mean	free	occ.	mean
all-free	0.301	0.0	0.1505	0.436	0.0	0.218
all-occupied	0.0	0.251	0.126	0.0	0.392	0.196
ANS(rgb) [1]	0.121	0.149	0.135	0.196	0.249	0.225
ANS(depth) [1]	0.145	0.241	0.193	0.231	0.376	0.304
View-extrap. [2]	0.155	0.264	0.210	0.250	0.404	0.327
OccAnt(depth) [3]	0.504	**0.619**	0.561	0.638	**0.750**	0.694
OccAnt(rgbd) [3]	**0.515**	0.615	**0.565**	**0.649**	0.748	**0.698**

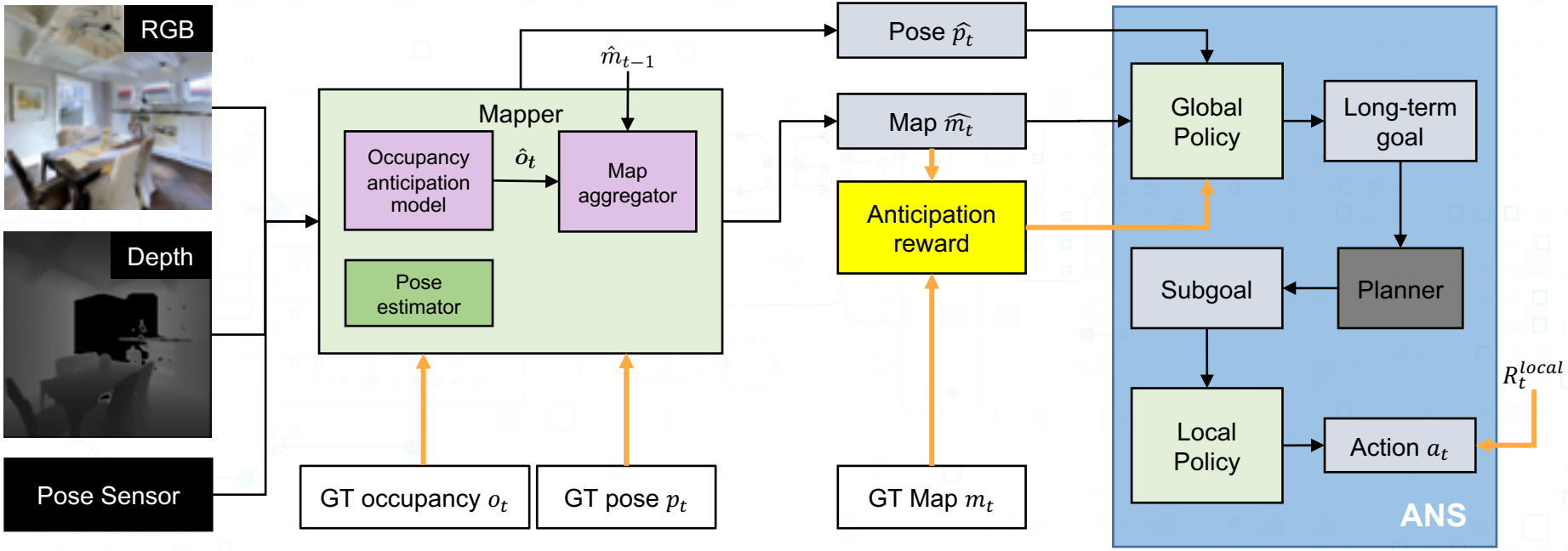
Existing methods

Ours

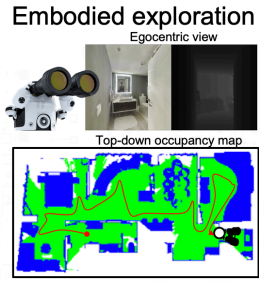
[1] Learning to Explore Using Active Neural SLAM, Chaplot et al., ICLR 2020
[2] Im2Pano3D: Extrapolating 360 Structure and Semantics Beyond the Field of View, Song et al., CVPR 2018
[3] OccAnt = Occupancy Anticipation for Efficient Exploration and Navigation, Ramakrishnan et al., ECCV 2020 (ours)

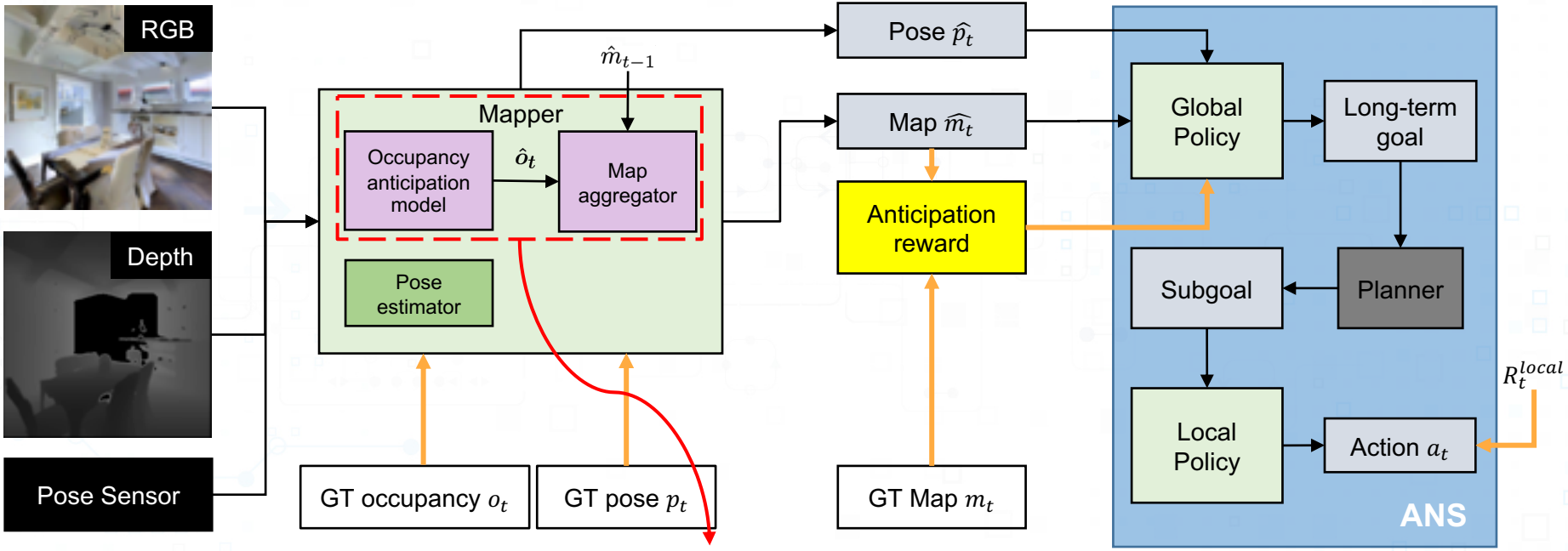
Exploration with occupancy anticipation





Exploration with occupancy anticipation



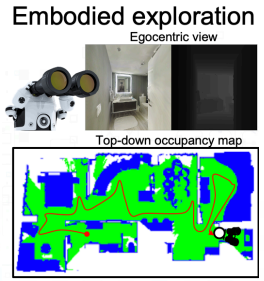


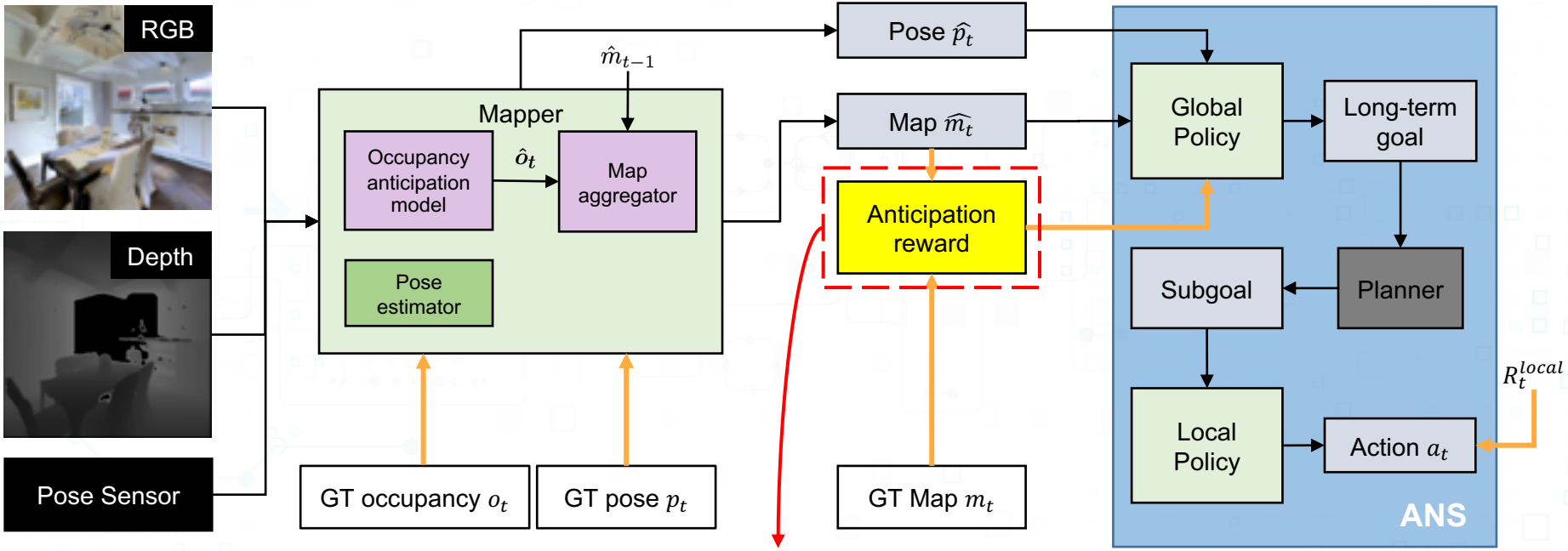
Proposed mapping + aggregation:
Reasoning about unseen regions from context

ANS = Chaplot et al., ICLR 2020

training signals

Exploration with occupancy anticipation



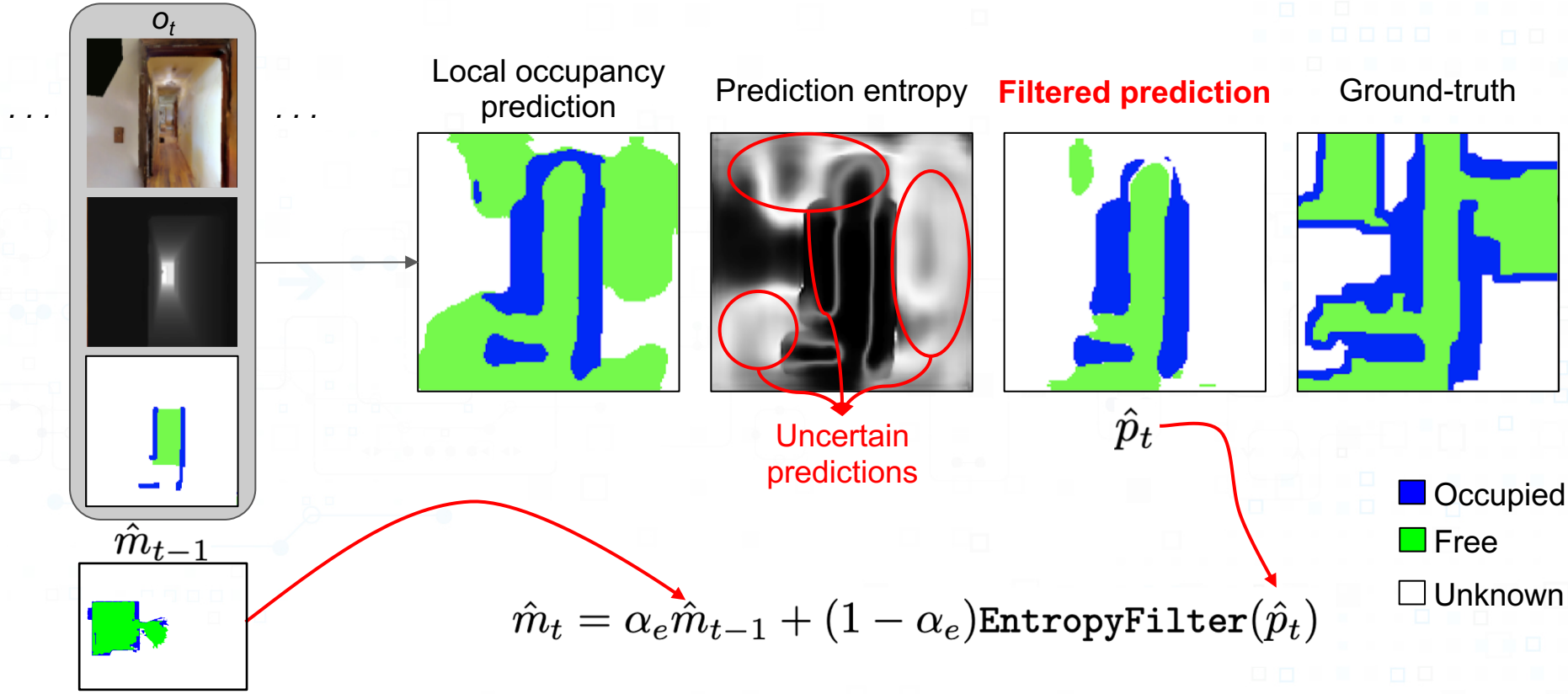


Proposed anticipation-based reward:
Learns intelligent behaviors for efficient exploration

ANS = Chaplot et al., ICLR 2020

training signals

Novel map aggregation strategy

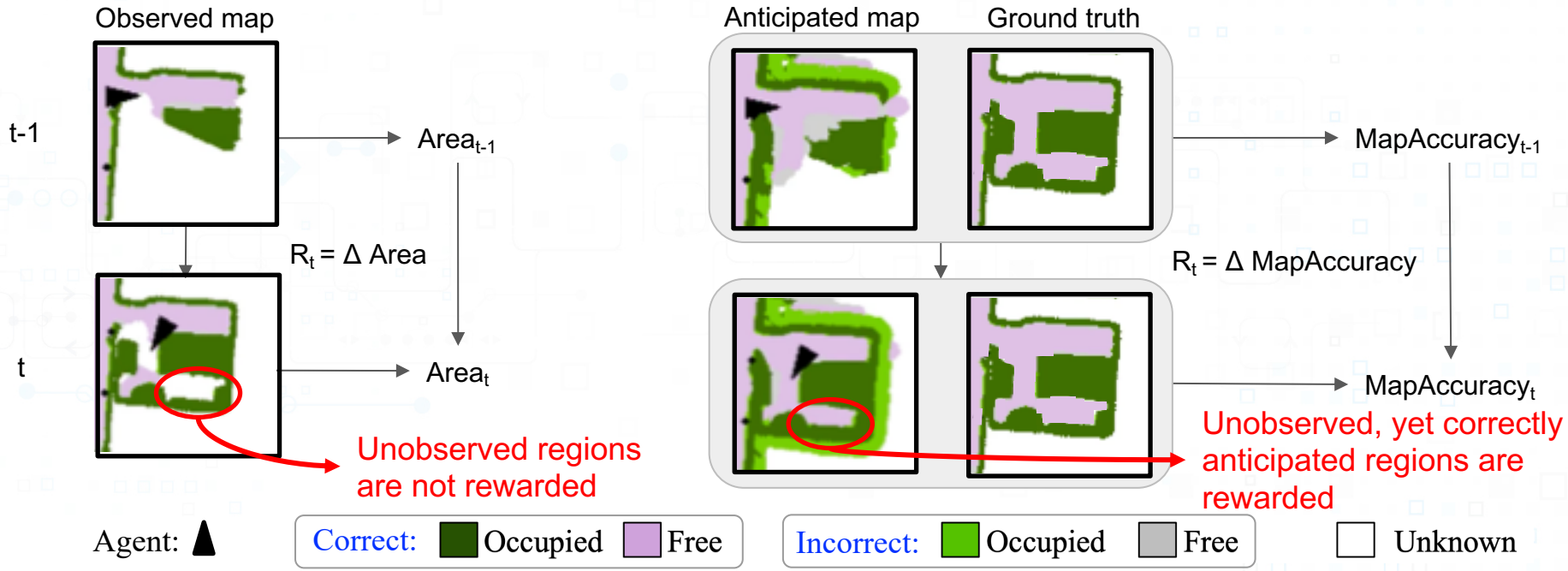


$$\hat{m}_t = \alpha_e \hat{m}_{t-1} + (1 - \alpha_e)\texttt{EntropyFilter}(\hat{p}_t)$$

Anticipation-based reward function

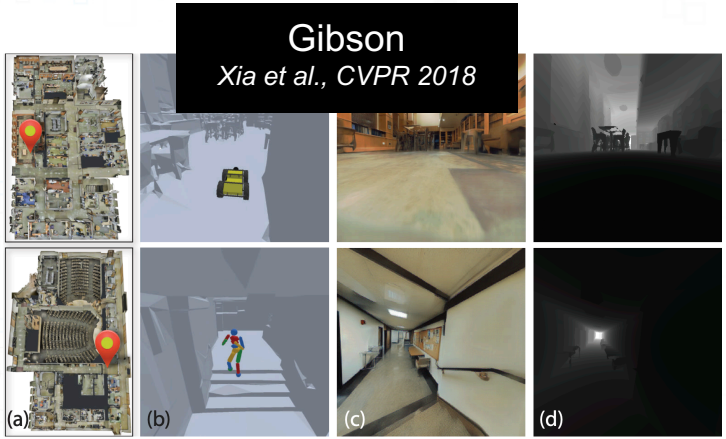
Traditional coverage reward:
Rewards increase in *area seen*

Anticipation-based reward function:
Rewards increase in *area accurately mapped*



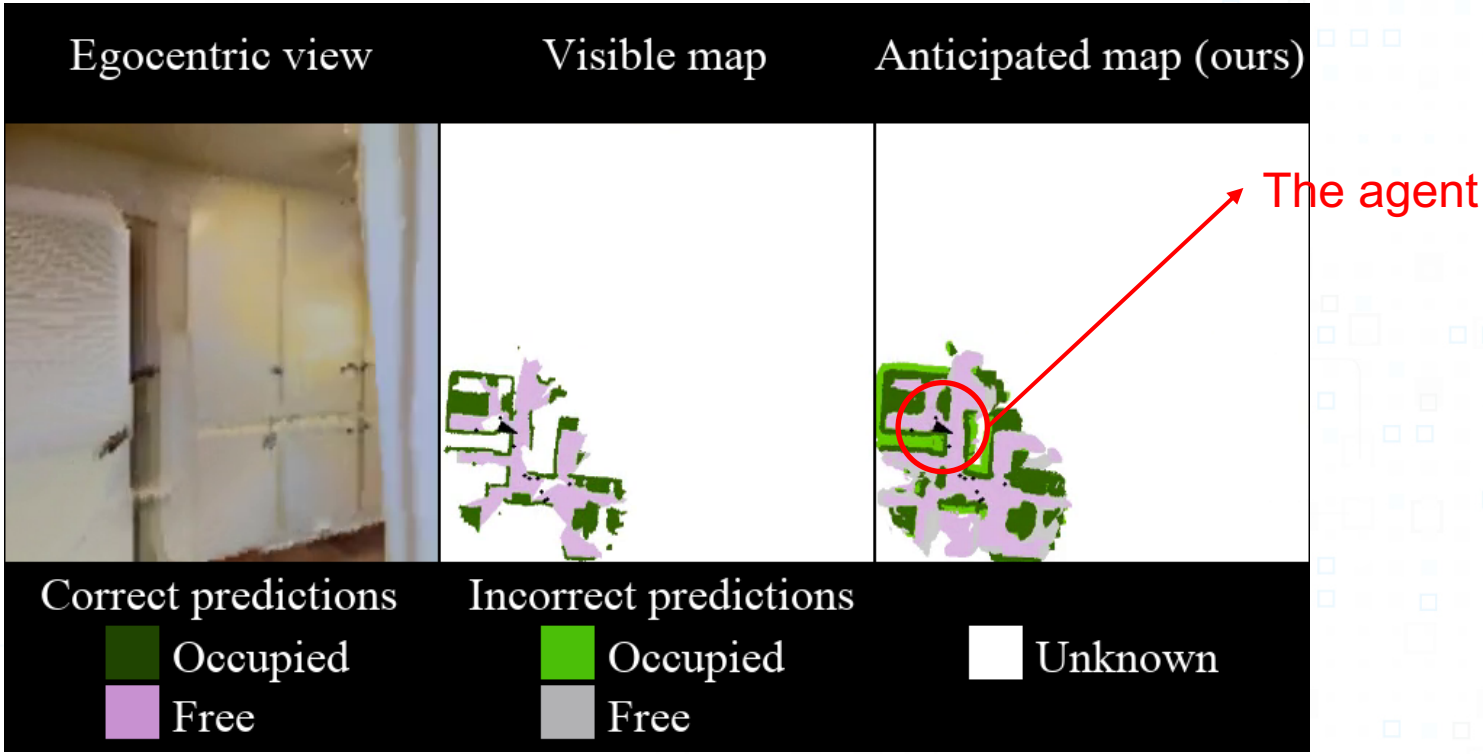
Experimental setup

- Noisy action + sensing
- Photorealistic 3D environments



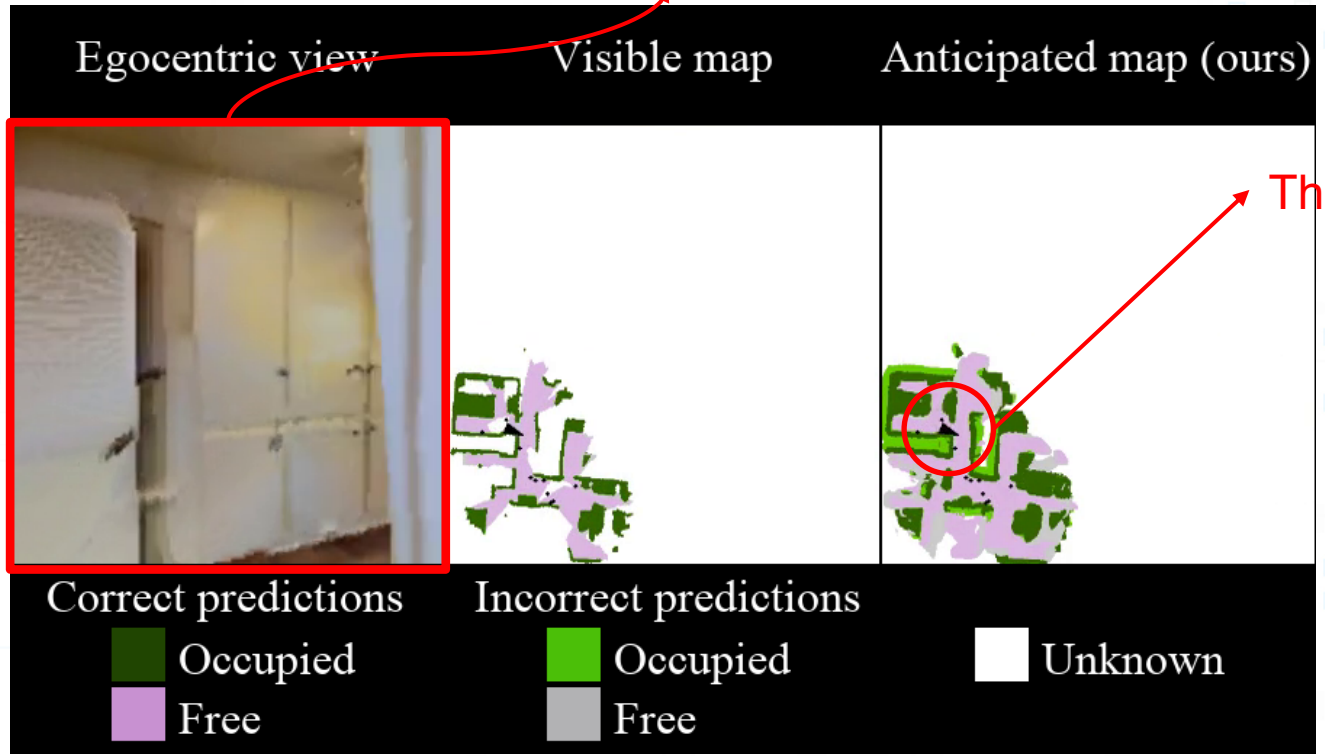
Evaluation metrics:
- Map accuracy
- Mean IoU (occupied and free classes)

Exploration results



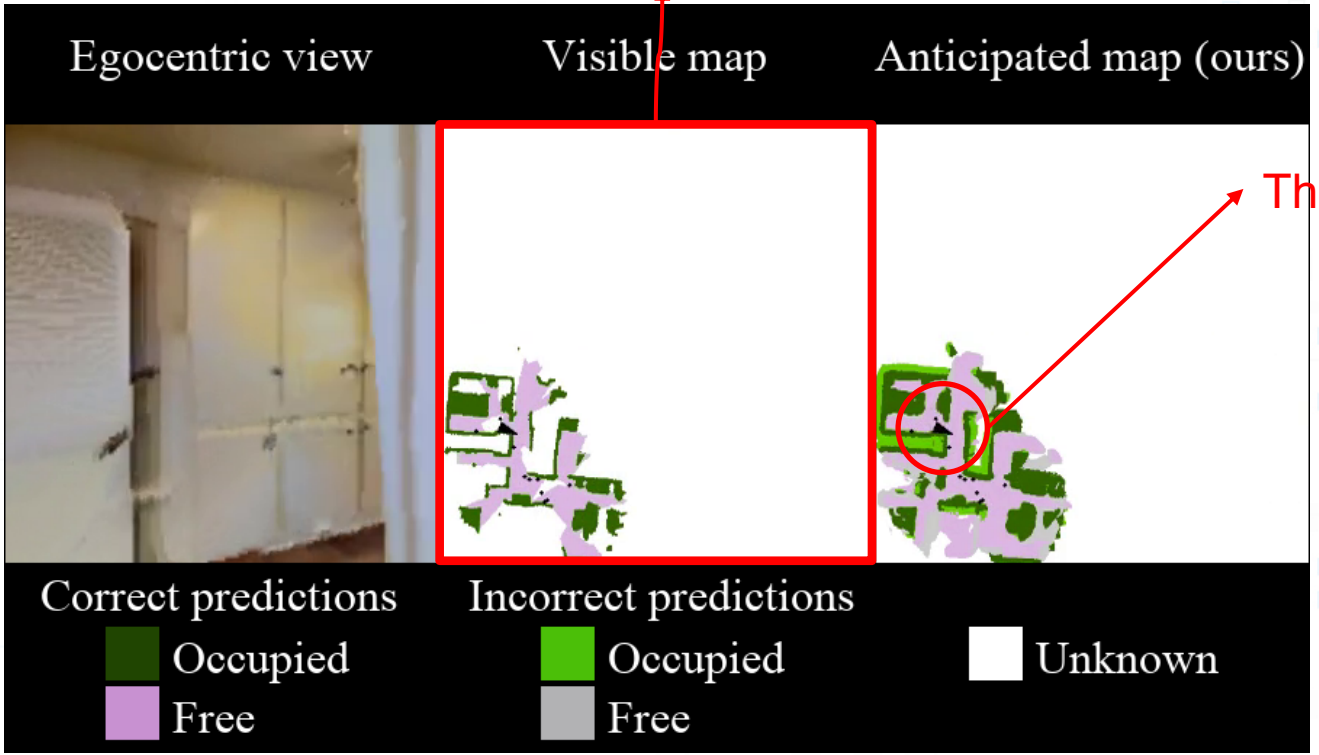
Explora

Agent's egocentric RGB view



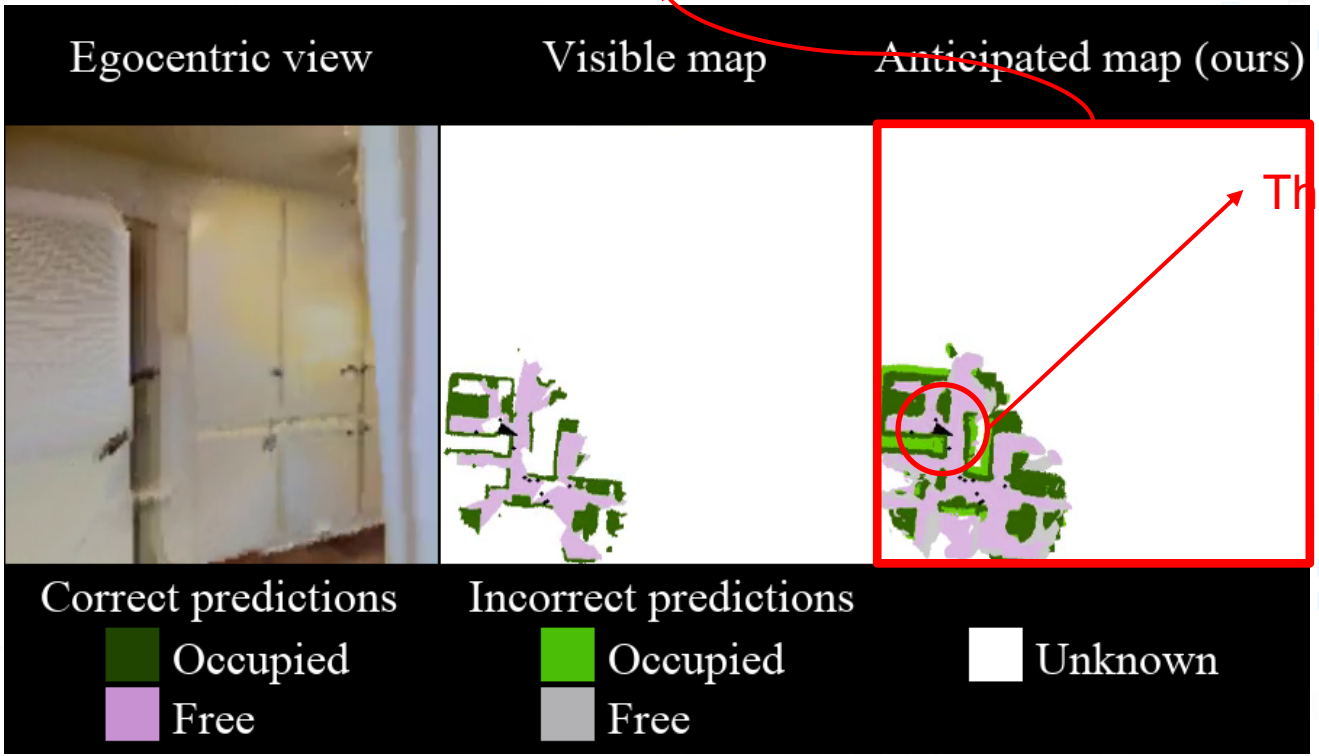
Explora

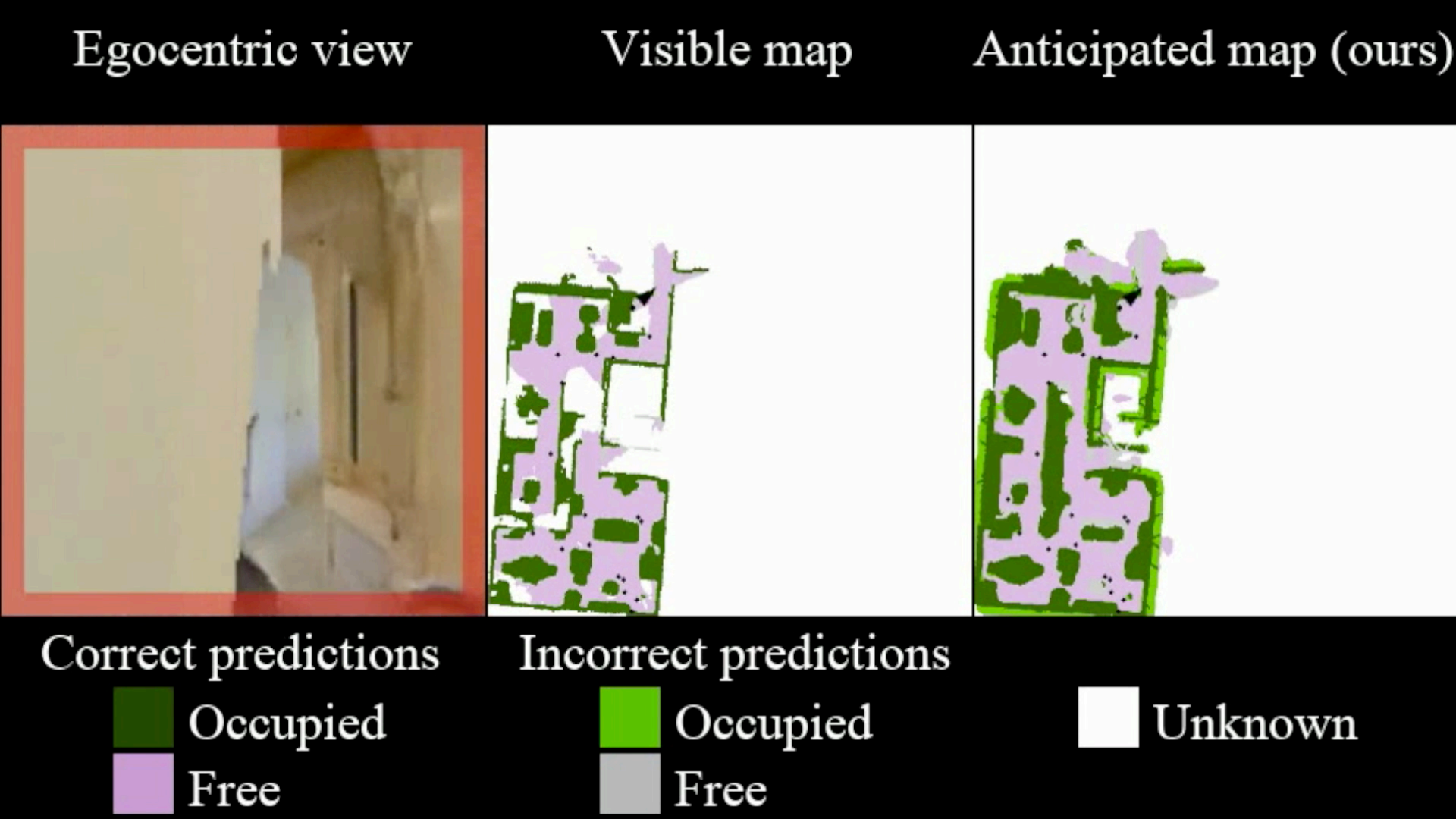
Top-down map built using *observed* regions (standard approaches today)

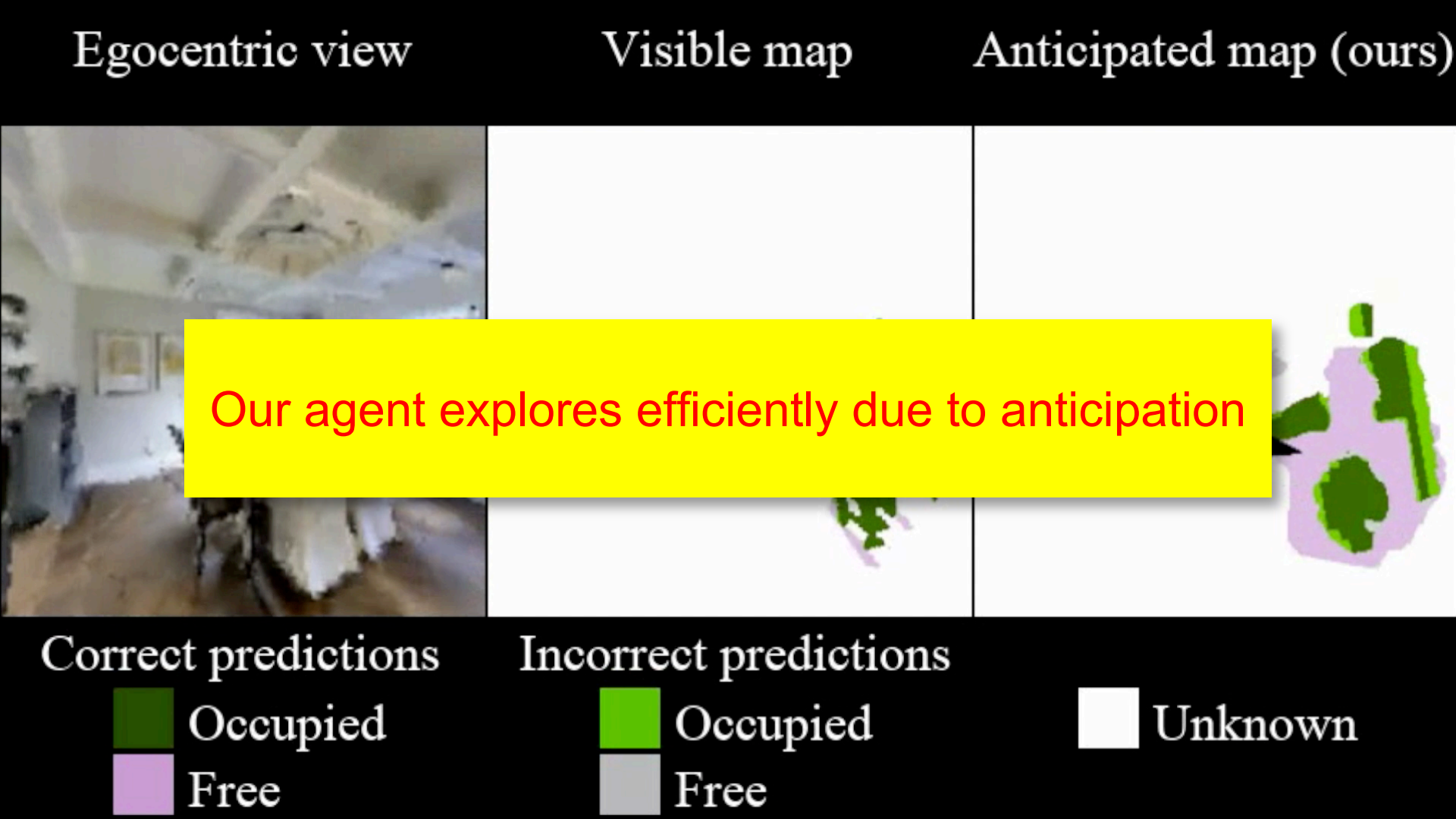


Explor

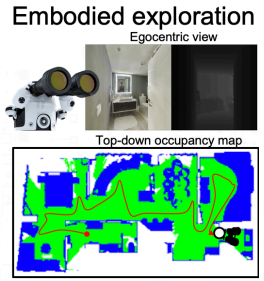
Top-down map built by *anticipating* occupancy of *unobserved* regions (ours)



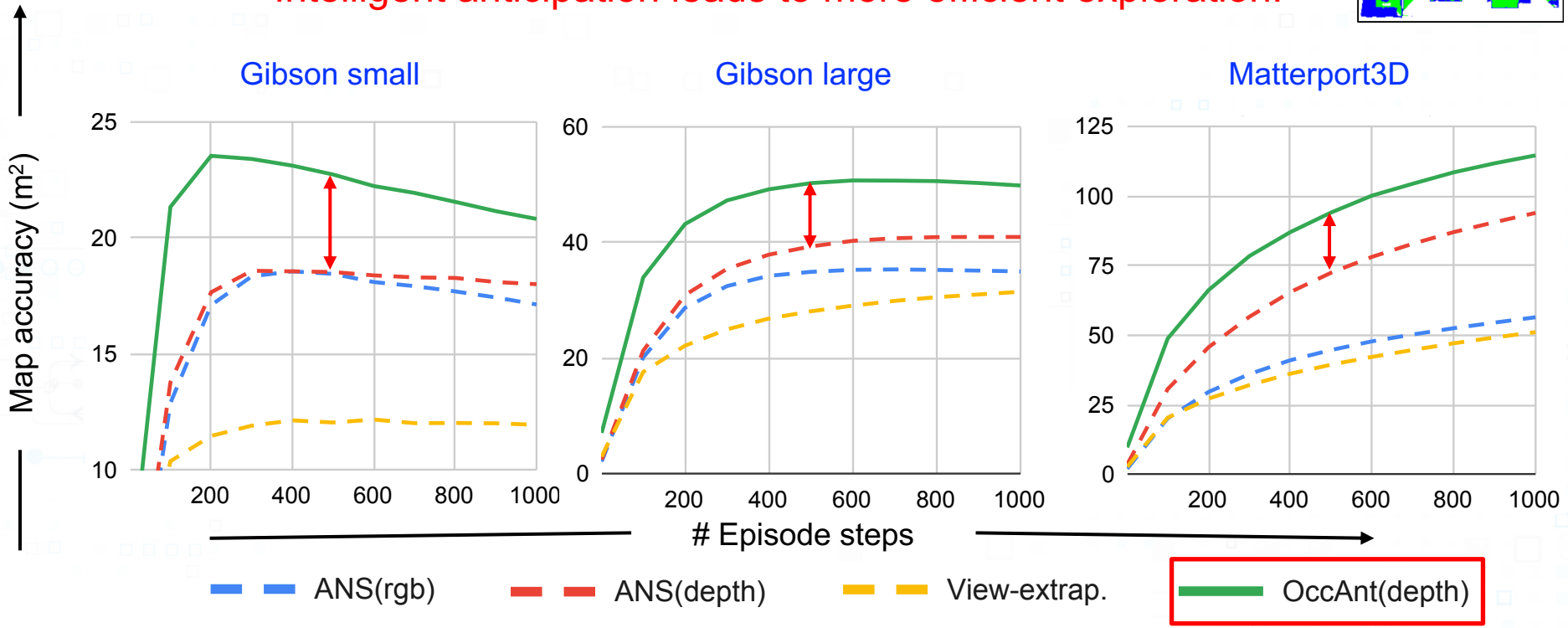

Egocentric view
Visible map
Anticipated map (ours)
Correct predictions
Occupied
Free
Incorrect predictions
Occupied
Free
Unknown


Egocentric view
Visible map
Anticipated map (ours)
Our agent explores efficiently due to anticipation
Correct predictions
Occupied
Free
Incorrect predictions
Occupied
Free
Unknown

Exploration results

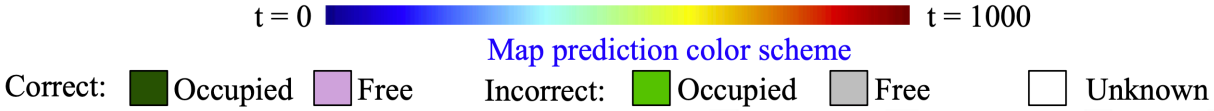


Intelligent anticipation leads to more efficient exploration.

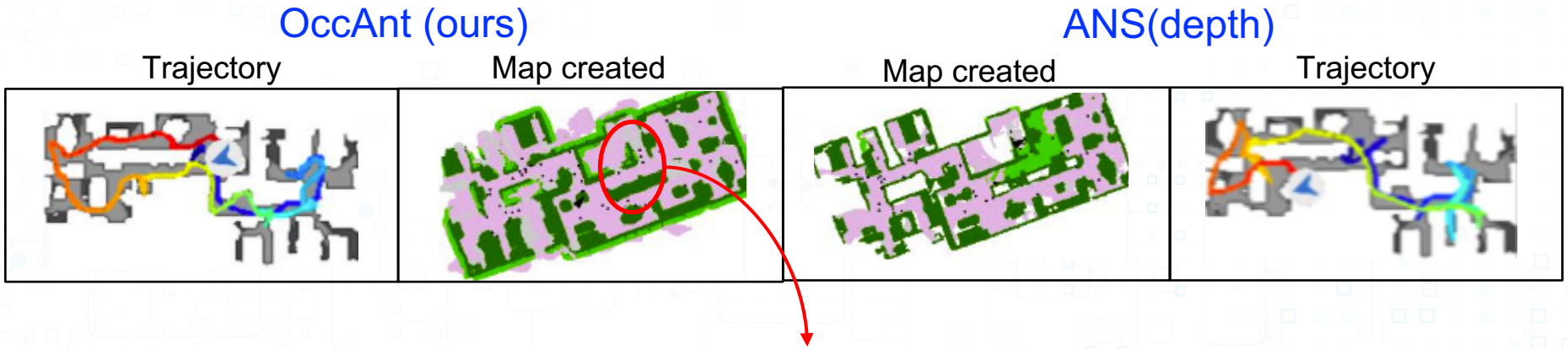


Exploration results: strengths

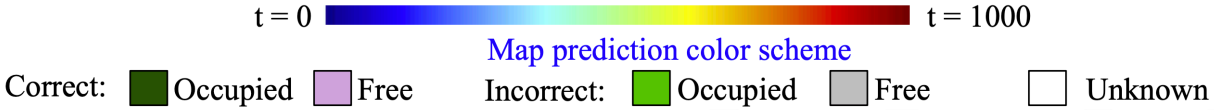
Baseline agent stuck exploring narrow spaces



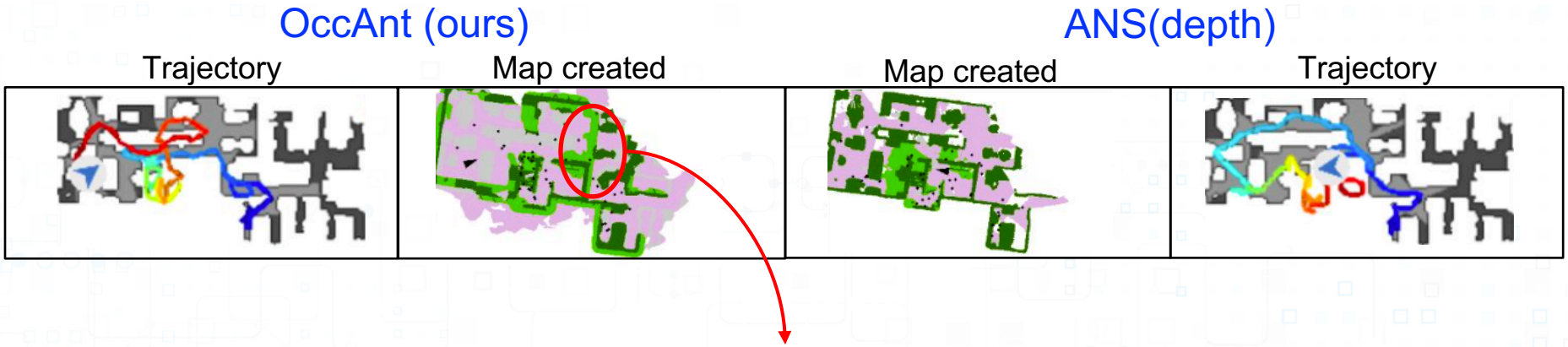
Exploration results: strengths



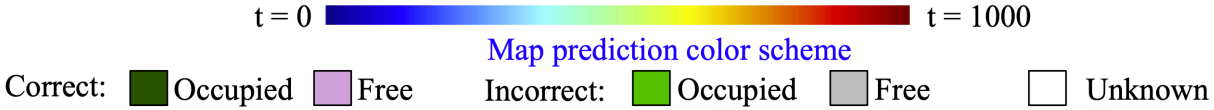
Our approach has better map quality for similar area seen



Exploration results: weakness

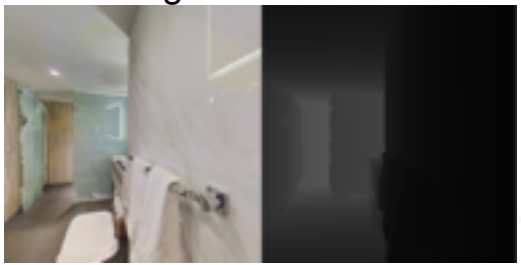


In some cases, incorrect anticipation can reduce map quality for our approach

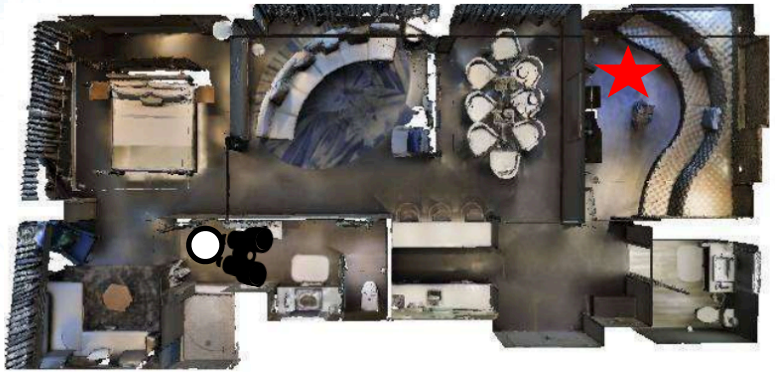


PointGoal navigation

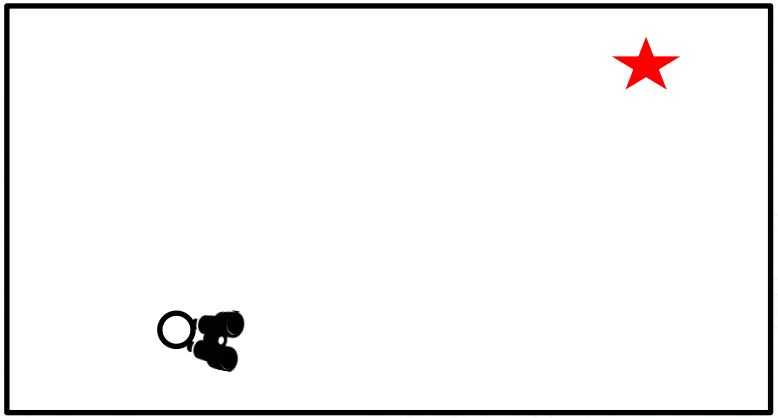
Egocentric view



3D environment



Top-down occupancy map



Navigation goal

Occupied Free Unknown

PointGoal navigation results

Intelligent anticipation leads to more efficient navigation.

	Gibson Validation		
Method	SPL	Success Rate	Time taken
ANS(rgb) [9]	0.668	0.879	254.109
ANS(depth)	0.768	0.866	226.161
View-extrap.	0.104	0.333	835.556
OccAnt(depth)	0.778	0.913	194.751
OccAnt(rgbd)	**0.800**	**0.930**	**171.874**

View-extrap. = Im2Pano3D: Extrapolating 360 Structure and Semantics Beyond the Field of View, Song et al., CVPR 2018
ANS = Learning to Explore Using Active Neural SLAM, Chaplot et al., ICLR 2020
OccAnt = Occupancy Anticipation for Efficient Exploration and Navigation, Ramakrishnan et al., ECCV 2020 (ours)

Habitat PointNav challenge 2020

Winner of the Habitat 2020 PointNav challenge

	Test standard			Test challenge		
Rank	Team	SPL %	Success %	Team	SPL %	Success %
1	**OccupancyAnticipation**	19.2	24.8	**OccupancyAnticipation**	20.9	27.5
2	ego-localization [13]	10.4	13.6	ego-localization [13]	14.6	19.2
3	Information Bottleneck	5.0	7.5	DAN [28]	13.2	25.3
4	cogmodel_team	0.8	1.3	Information Bottleneck	6.0	8.8
5	UCULab	0.5	0.8	cogmodel_team	0.7	1.2
6	Habitat Team (DD-PPO) [65]	0.0	0.2	UCULab	0.1	0.2

[13] Datta et al., Integrating egocentric localization for more realistic pointgoal navigation agents, CVPR 2020 workshop
[28] Karkus et al., Differentiable algorithm networks for composable robot learning, RSS 2019
[65] Wijmans et al., Learning near-perfect pointgoal navigators from 2.5 billion frames, ICLR 2020

Occupancy Anticipation for Efficient Exploration and Navigation

ECCV 2020 (Spotlight)

Project page: http://vision.cs.utexas.edu/projects/occupancy_anticipation

Santhosh Kumar Ramakrishnan[1,2]

Ziad Al-Halah[1]

Kristen Grauman[1,2]

[1] The University of Texas at Austin

[2] Facebook AI Research