OCTOBER 2023

TEXAS
The University of Texas at Austin

From Social- To Everything- Navigation

JOYDEEP BISWAS
Associate Professor, Computer Science Department, The University of Texas at Austin

Thanks to my collaborators!


Haresh Karnan


Zichao Hu


Luisa Mao


Rohan Chandra


Amir Hossain Raj


Amir Payandeh


Arthur Zhang

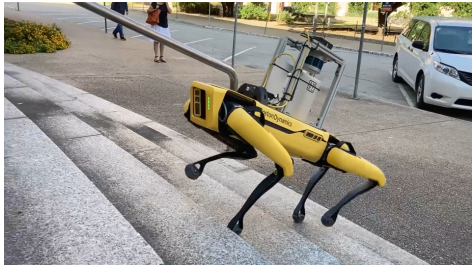

Xueso Xiao


Garrett Warnell

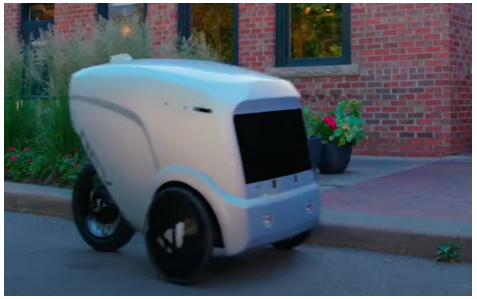

Peter Stone

Anirudh Nair, Sören Pirk, Alexander Toshev, Justin Hart, Elvin Yang, Daniel Farkash, Chaitanya Eranki, Christina Zhang, Ji-Hwan Park, Raymond Hong, Pranav Kalyani, Lochana Kalyanaraman, Arsh Gamare, Arnav Bagad, Maria Esteva, Vrushabh Zinage, Efstathios Bakolas, Zayne Sprague, Jarrett Holtz

Context: Urban-Scale Service Mobile Robots



https://www.amazon.com/Introducing-Amazon-Astro/dp/B078NSDFSB



https://refraction.ai/

https://www.diligentrobots.com/moxi

https://www.starship.xyz/the-starship-robot/

What does socially compliant robot navigation look like?

Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations for Social Navigation

Haresh Karnan[1], Anirudh Nair[1], Xuesu Xiao[1], Garrett Warnell[1, 2], Sören Pirk[3], Alexander Toshev[3], Justin Hart[1], Joydeep Biswas[1], and Peter Stone[1, 4]

[1]The University of Texas at Austin

[2]Army Research Laboratory

[3]Google Robotics

[4]Sony AI

DEVCOM ARMY RESEARCH LABORATORY

AMRL

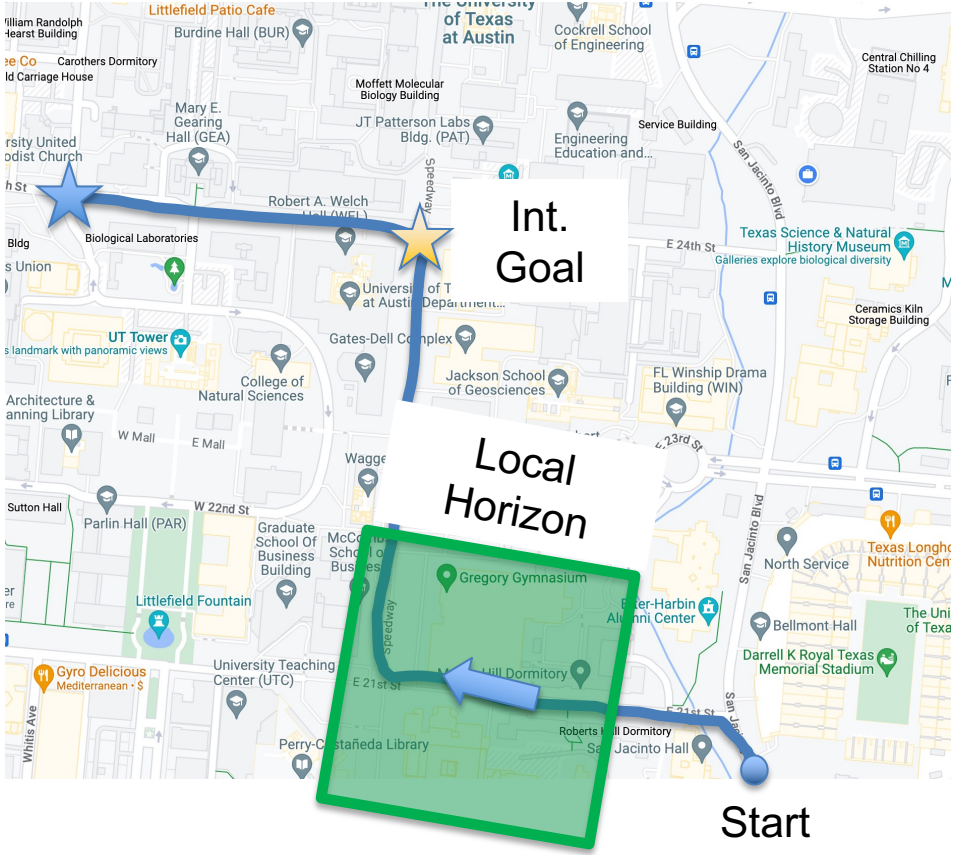
Google Research

LARG Learning Agents Research Group The University of Texas at Austin

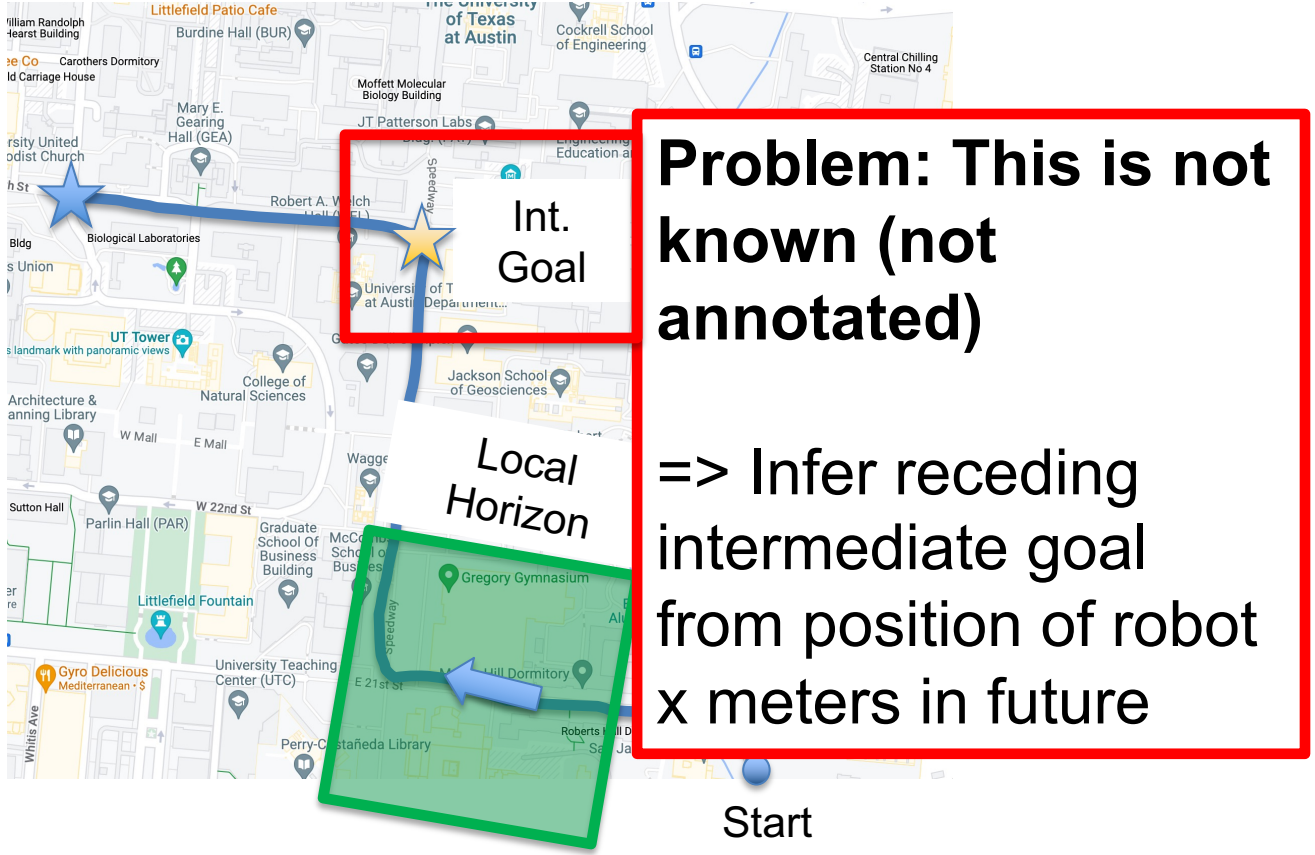
Geometric Navigation: How bad is it?

Local Geometric Navigation Formulation

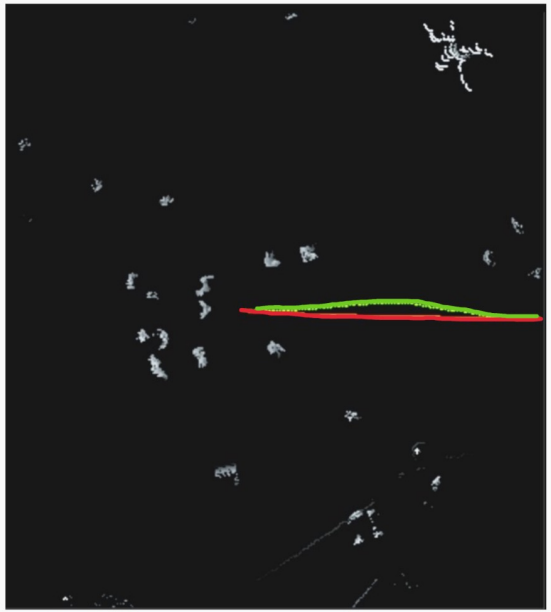
Nav. Goal



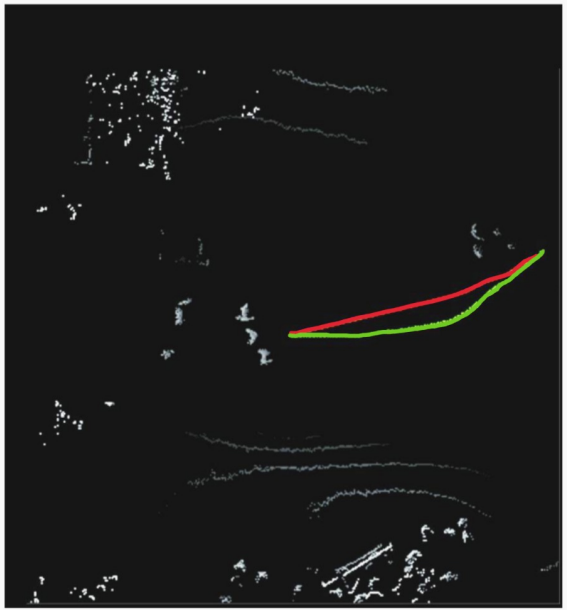
Local Geometric Navigation Formulation



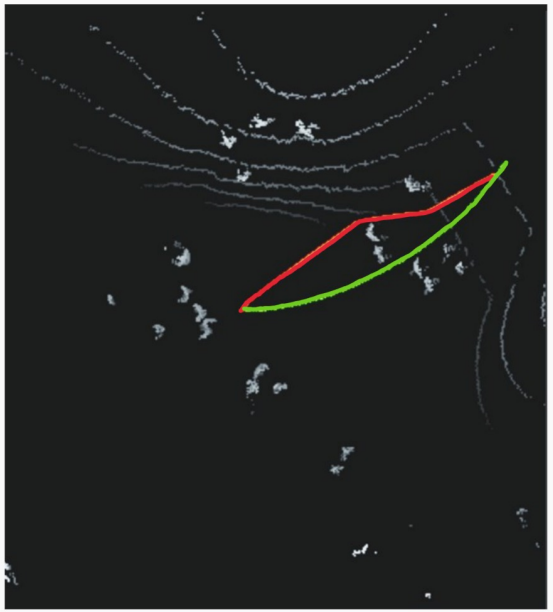
Hausdorff Distance: Geom. Nav. vs. Human



Hausdorff Distance ≈ 1.0



Hausdorff Distance ≈ 2.0



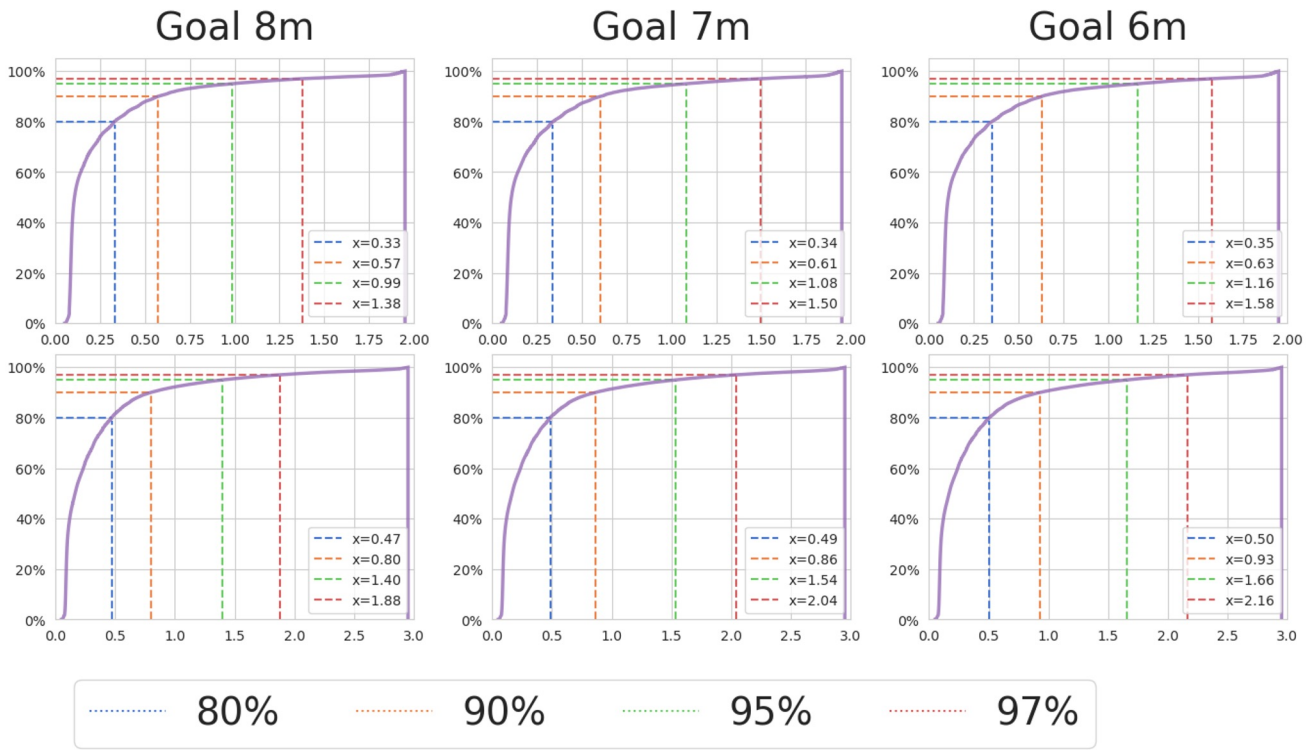
Hausdorff Distance ≈ 3.0

Amir Hossain Raj, Zichao Hu, Haresh Karnan, Rohan Chandra, Amirreza Payandeh, Luisa Mao, Peter Stone, Joydeep Biswas, Xuesu Xiao (2023). Targeted Learning: A Hybrid Approach to Social Robot Navigation. Arxiv preprint: 2309.13466

Geometric Navigation vs. Human Demonstrations

Local Horizon: 2m

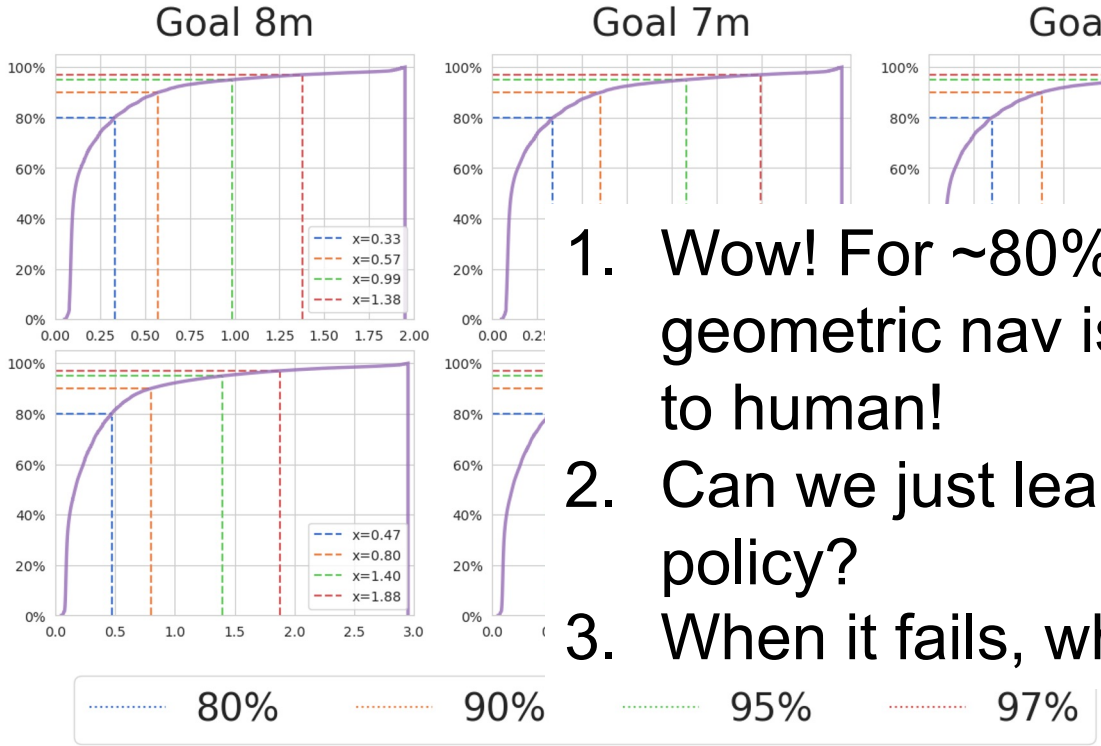
Local Horizon: 3m



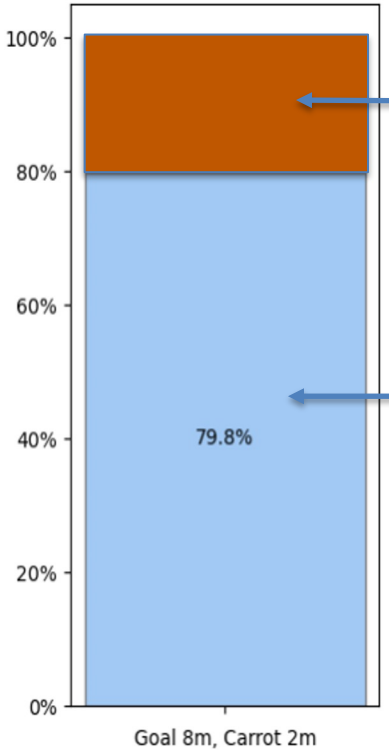
Geometric Navigation vs. Human Demonstrations

Local Horizon: 2m

Local Horizon: 3m



1. Wow! For ~80% of cases, geometric nav is very similar to human!
2. Can we just learn a better policy?
3. When it fails, why does it fail?



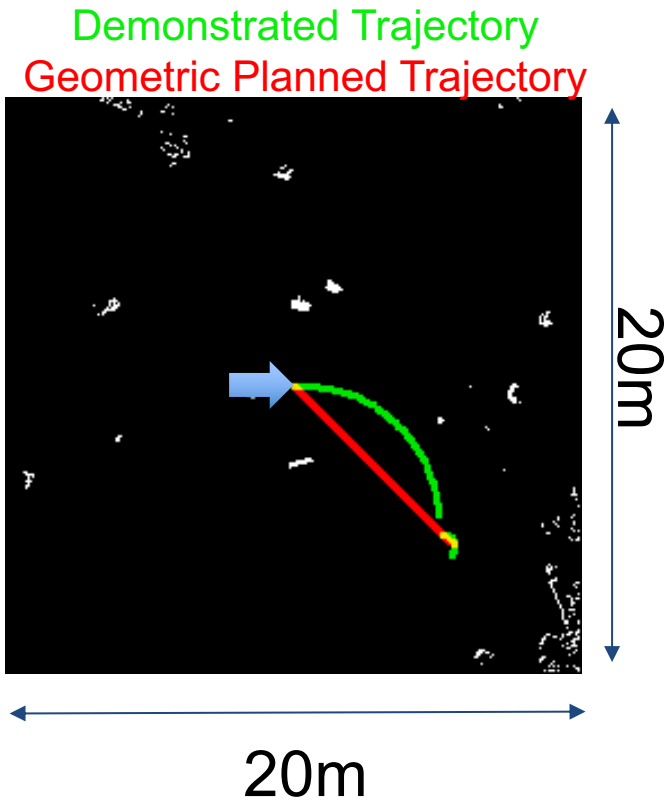
Social Navigation (?)

Geometric nav. Plan agrees with human

Plan agreement:
Hausdorff distance <= 0.5m

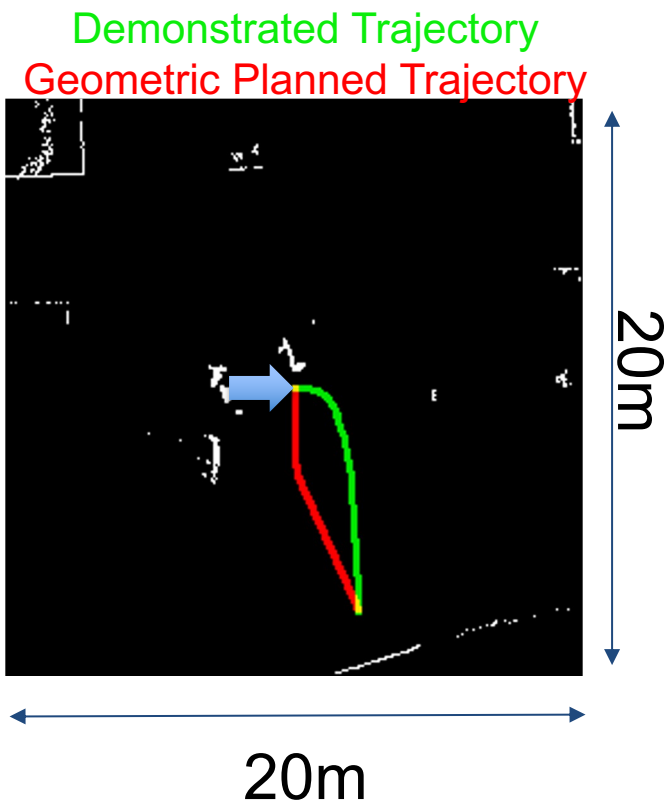
Manual Inspection:
Randomly sample, and classify

Avoiding Gravel



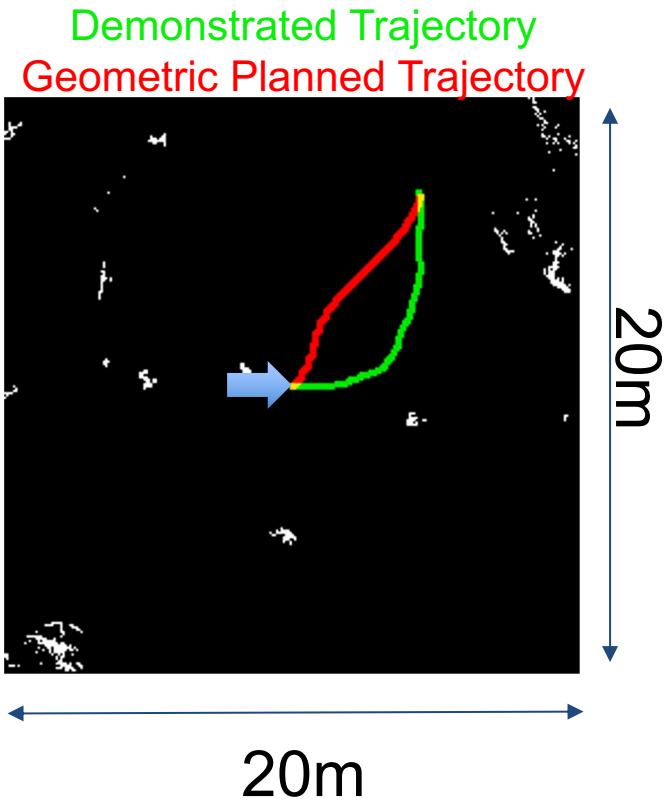
Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*

Crossing At Marked Crosswalks



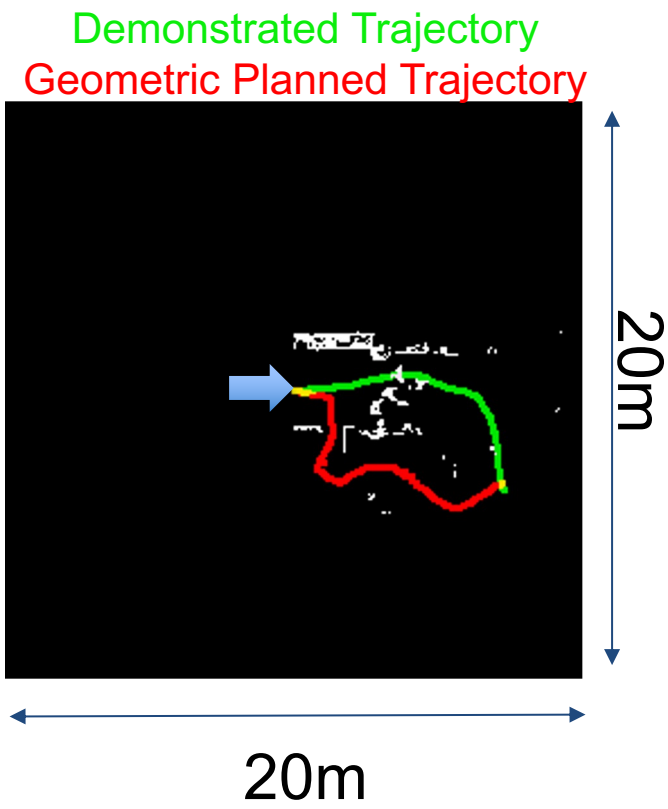
Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*

Crossing At Unmarked Crosswalks



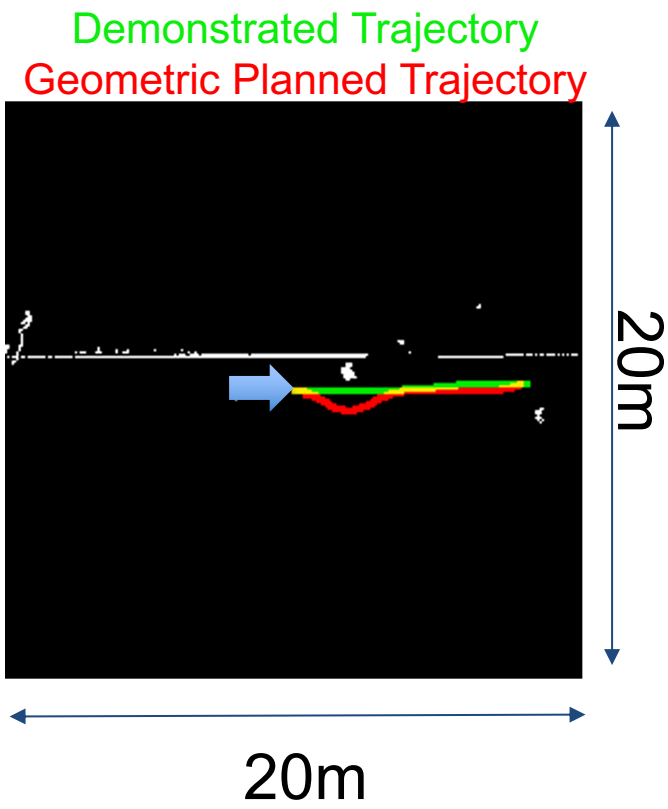
Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*

Avoiding Non-Public Areas



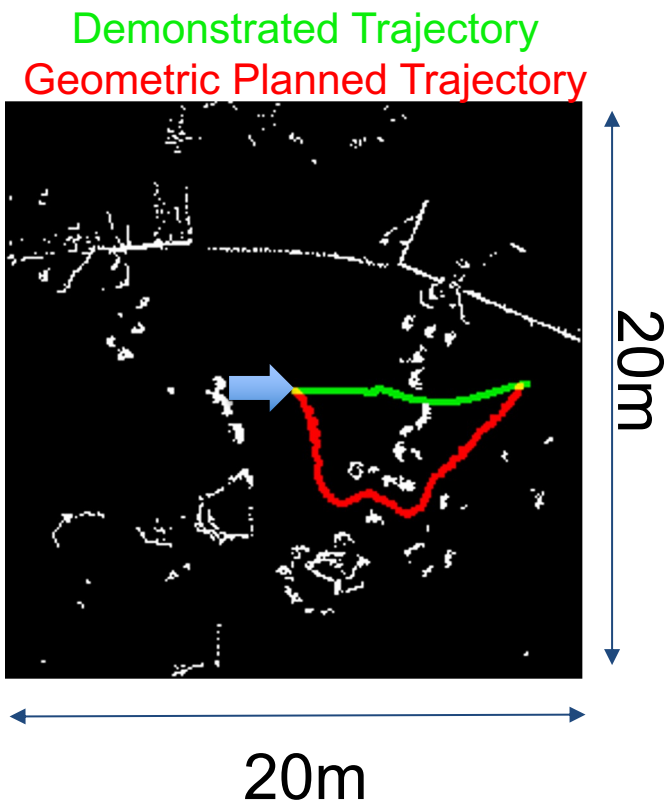
Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*

“Narrow Hallway” sans hallway

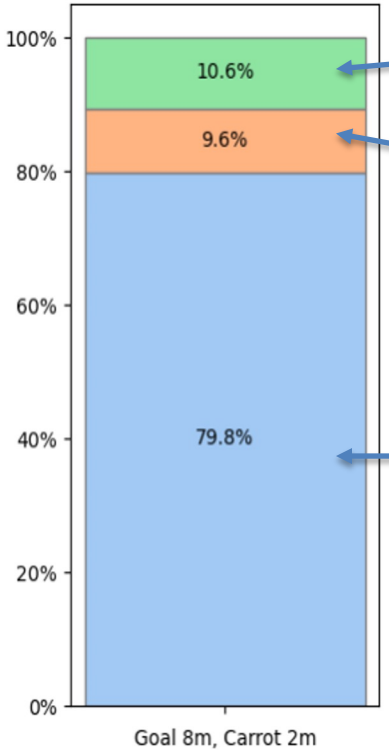


Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*

Cutting Across A Queue



Haresh Karnan, Anirudh Nair, Xuesu Xiao, Garrett Warnell, Soeren Pirk, Alexander Toshev, Justin Hart, Joydeep Biswas, Peter Stone (2022). Socially CompliAnt Navigation Dataset (SCAND): A Large-Scale Dataset Of Demonstrations For Social Navigation. *IEEE Robotics and Automation Letters, 7(4), pp. 11807-11814*



Social Factors

Other Factors

Geometric nav. Plan agrees with human

Quote from a SCAND driver [unedited, emphasis mine]

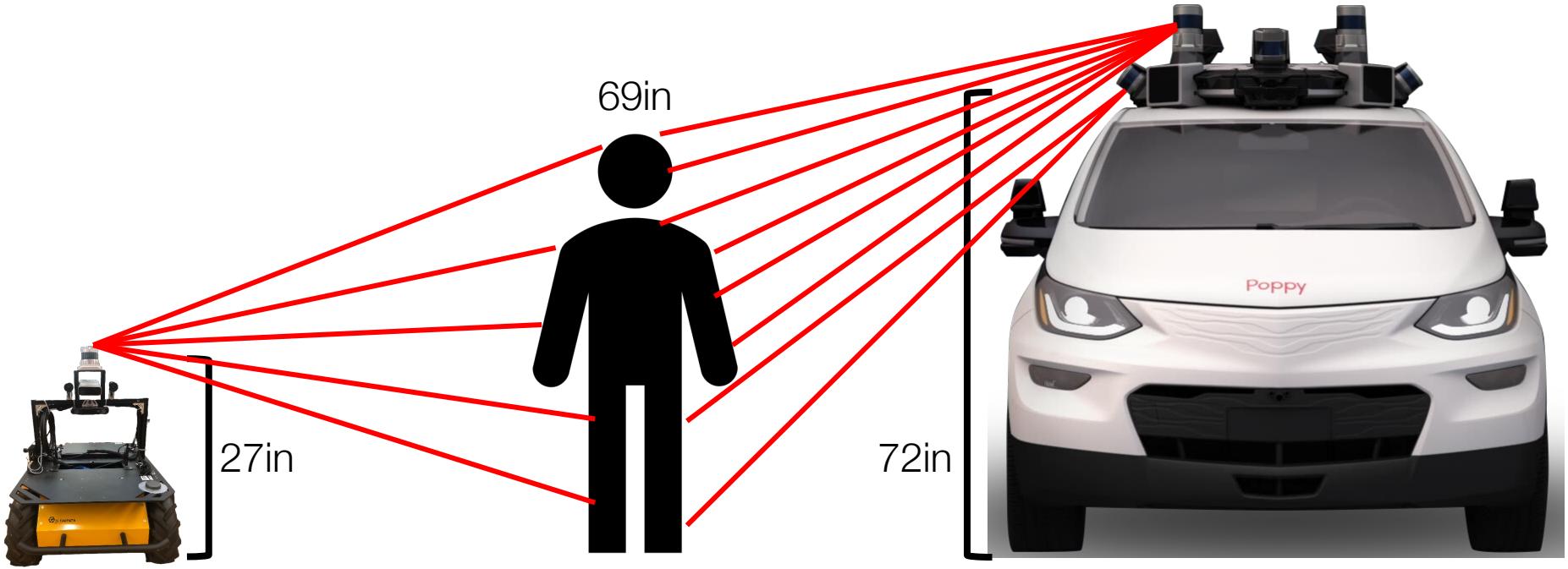
1. **right side of the sidewalk**
2. keeping away from **unpreferred terrain**, such as the **metal cover terrain / bushes** on the right in speedway
3. **keeping away from groups of people and not cut through them (social compliance)**
4. Keeping away from **"under construction" zones**
5. being **terrain-aware** - taking the stairs instead of an incline with bushes (near UT tower)
6. maybe sunlight. operators could have preferred shaded sidewalks over sunny spaces

Maybe we need to detect more than just humans…

Question: Why Not Leverage AV Datasets?

Geiger et al, “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”, CVPR 2012
Sun et al, “Scalability in Perception for Autonomous Driving: Waymo Open Dataset”, CVPR 2020
Caesar et al, “nuScenes: A multimodal dataset for autonomous driving”, CVPR 2020

Sensor and Viewpoint Differences Between Robots and AVs



Urban Robot

Figure Dimensions are to Scale

Autonomous Vehicle[1]

[1]https://getcruise.com/ride/

The UT Campus Object Dataset

The UT Campus Object Dataset is a **multiclass, multimodal** urban robotics **dataset**, with **1.3 million 3D bounding box annotations** for **53 object classes**, **204 million annotated points** for **24 terrain classes**, and **globally consistent** pseudo-ground truth **poses**.

Zhang, A., Eranki, C., Zhang, C., Park, J.H., Hong, R., Kalyani, P., Kalyanaraman, L., Gamare, A., Bagad, A., Esteva, M. and Biswas, J., 2023. Towards Robust Robot 3D Perception in Urban Environments: The UT Campus Object Dataset. *arXiv preprint arXiv:2309.13549*.

https://amrl.cs.utexas.edu/coda/

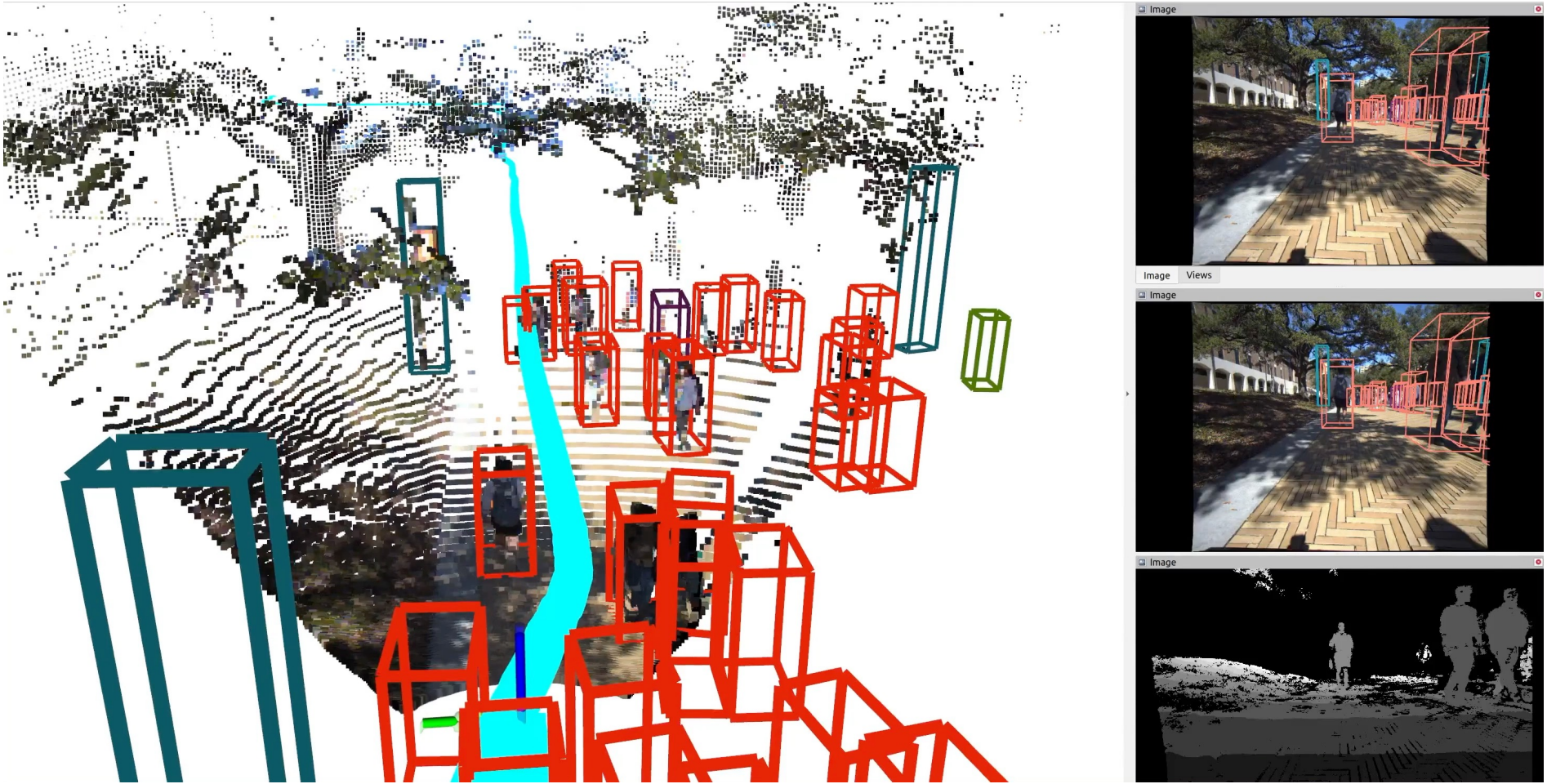
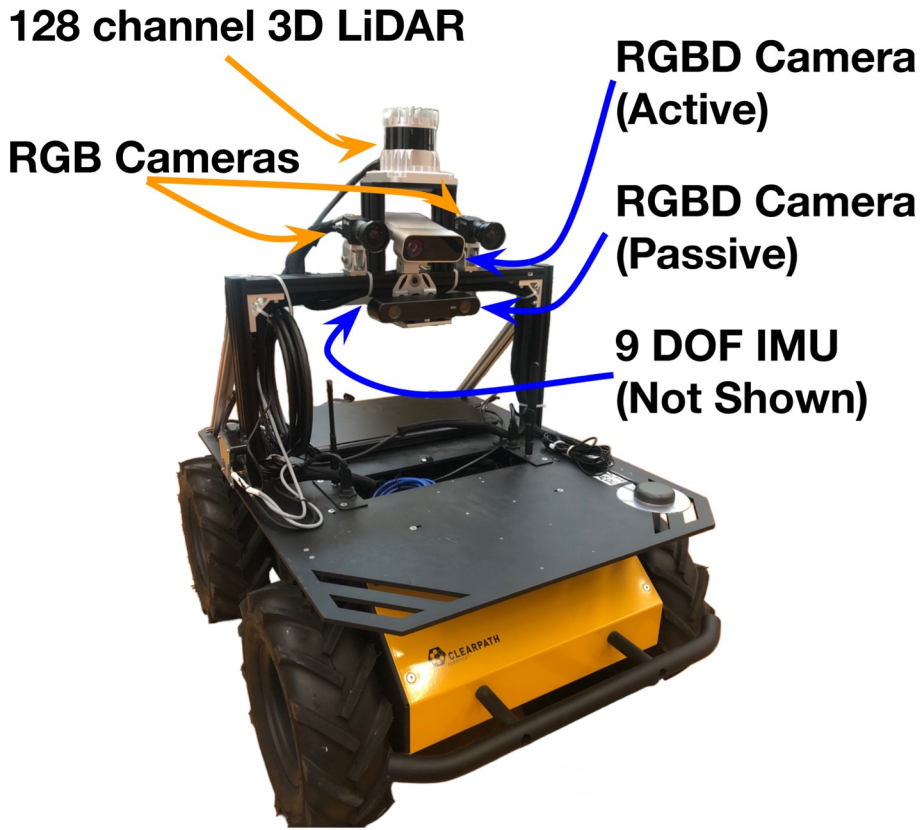


Image
Image
Views
Image
Image

AM
RL

Hardware Synchronization between LiDAR and Stereo RGB Cameras



https://amrl.cs.utexas.edu/coda/

UT Campus Object Dataset | Scenes in Geographic Areas

AMRL

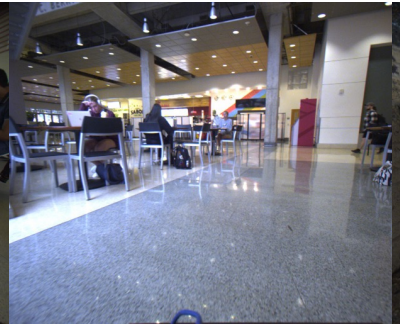
5 Geographic Areas with Distinct Features

SWY – Blue
GDC – Green
WCP – Brown
GUAD – Red
UNB - Purple

SWY
Hybrid Traffic Area

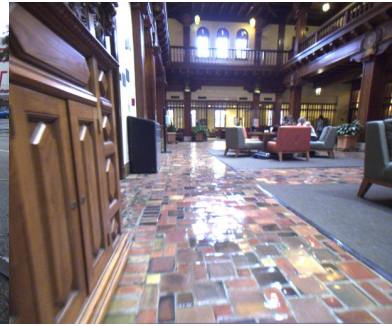
GDC
Open Class Area

WCP
Cafeteria Dining Area



GUAD
Road/Street Area

UNB
Library Study Area



https://amrl.cs.utexas.edu/coda/

UT Campus Object Dataset | Weather/Lighting Diversity (Outdoors)

AMRL

SWY
(Left to Right)
Sunny
Cloudy
Rainy

GUAD
(Left to Right)
Sunny
Cloudy
Rainy
Dark

UNB
(Left to Right)
Sunny
Cloudy
Rainy
Dark

UT Campus Object Dataset | **Viewpoint Diversity**

AMRL

Red bounding box indicates common objects between views

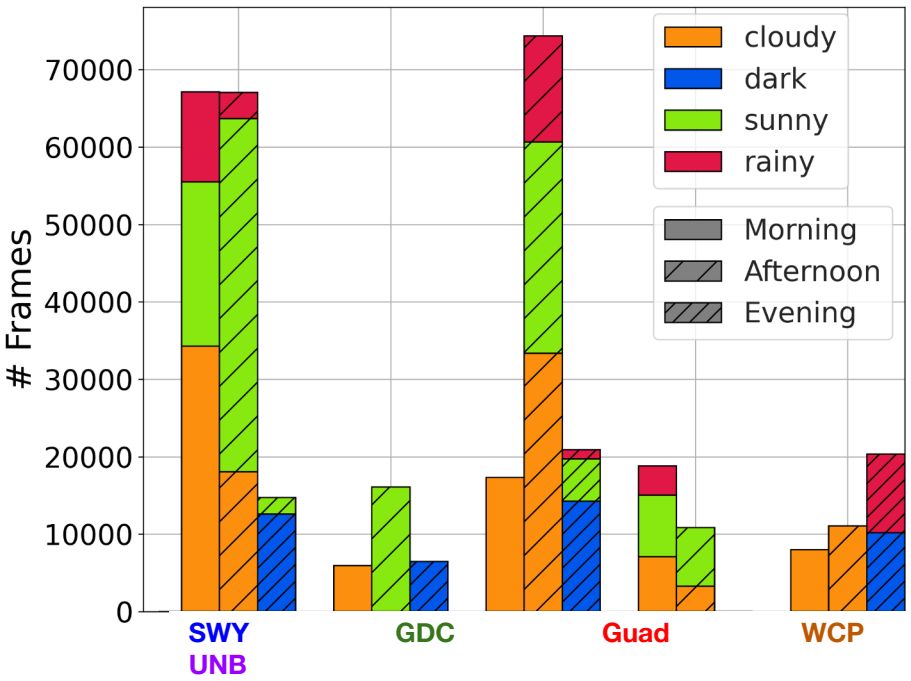
GDC
Circular Statue

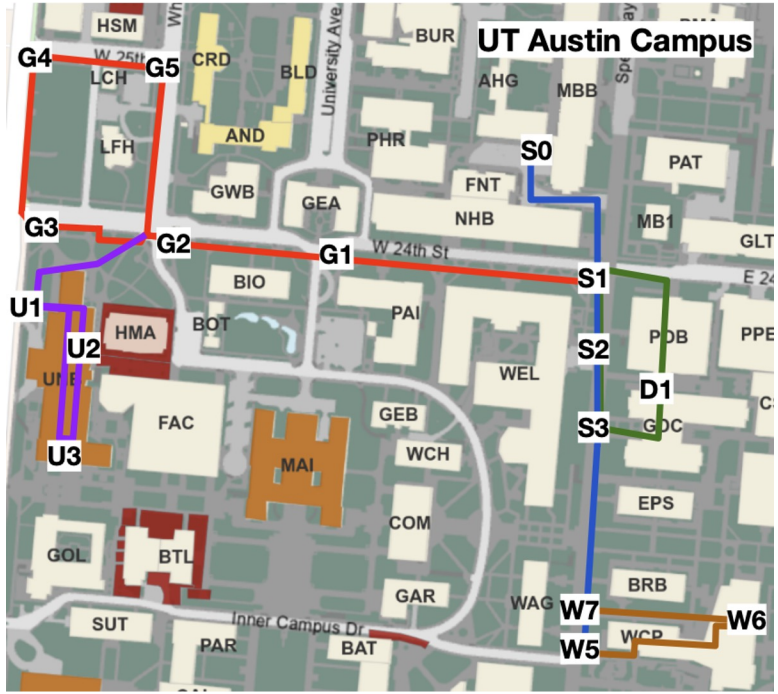
GUAD
Shoe Store by Crosswalk

UNB
Crosswalk by Library

WCP
Cafeteria Entrance

UT Campus Object Dataset | **Scene Diversity**



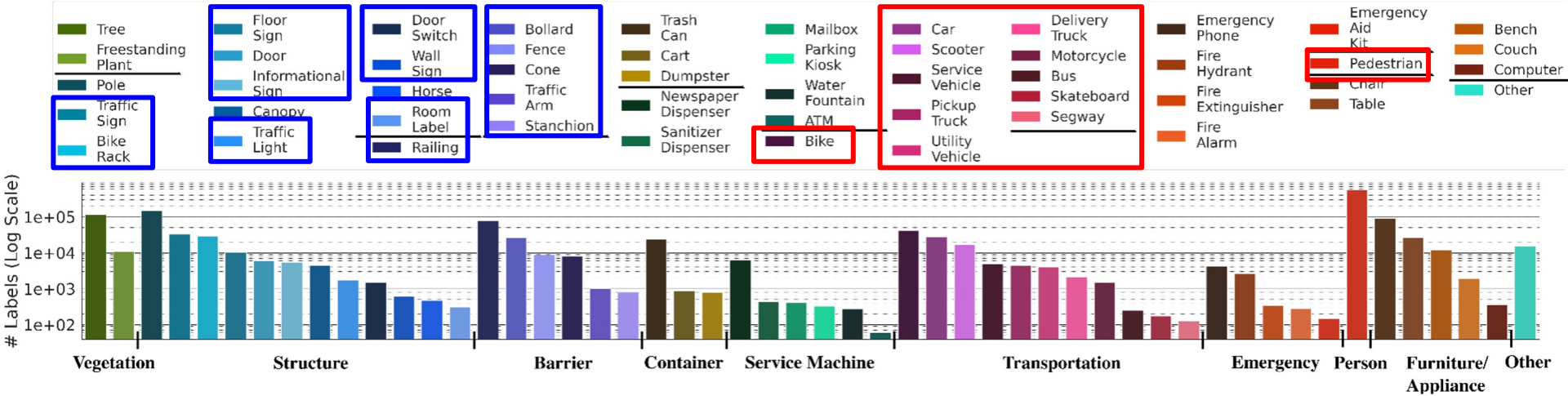


https://amrl.cs.utexas.edu/coda/

UT Campus Object Dataset | **Object Class Diversity**

Situational Awareness for Static Objects

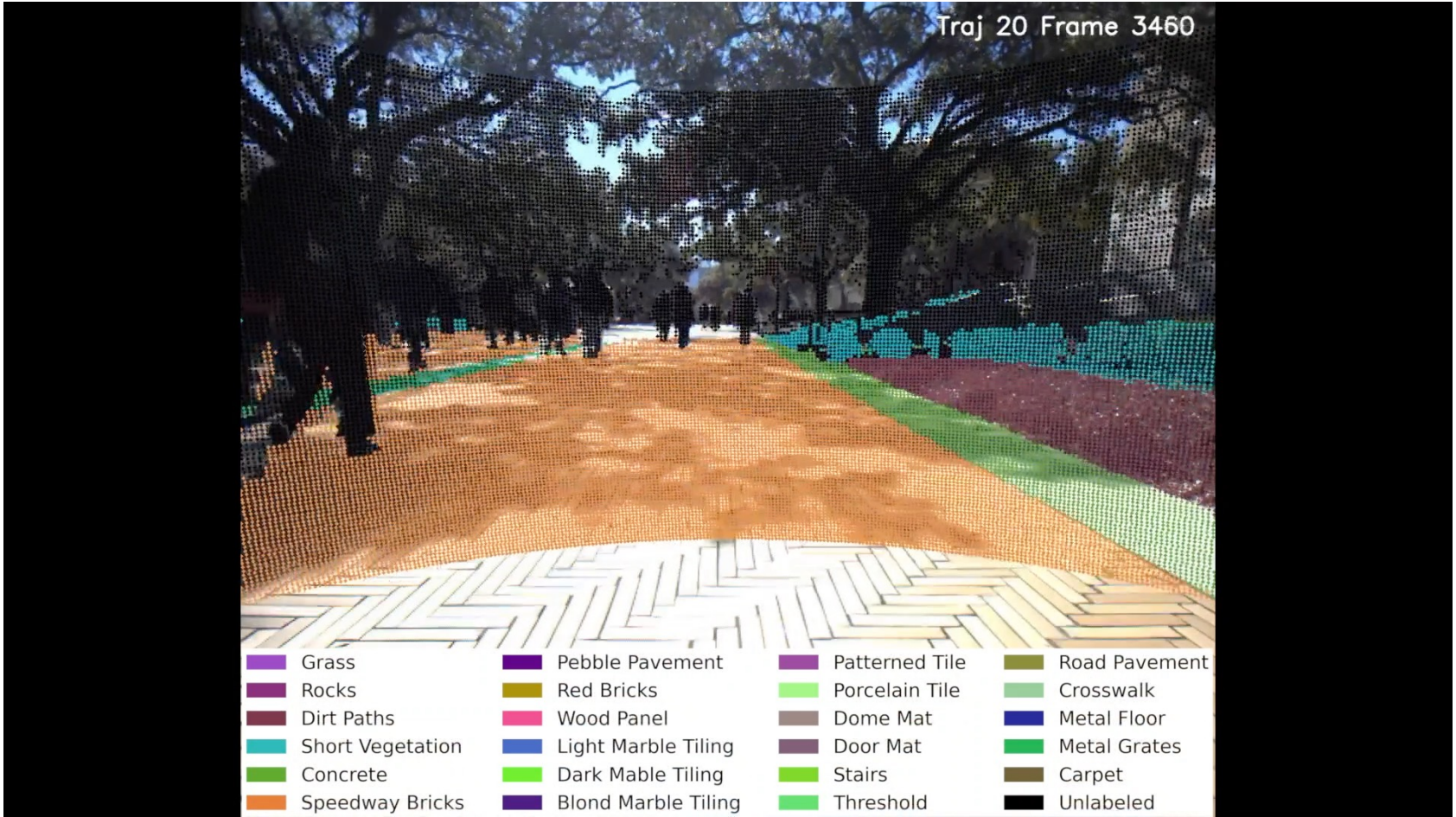
Situational Awareness for Dynamic Objects



Number of bounding box annotations per object class

UT Campus Object Dataset | Terrain Class Visualization

Terrain Annotations from 3D LiDAR Projected on RGB Image



AM RL

Evaluation on Car, Pedestrian, Cyclist Detections

Pretrain \ Dom. Adapt	Direct	Finetune	Direct + Unsup. Domain Adapt.
nuScenes[1]	21.30	91.39	14.07
Waymo[2]	46.20	93.12	38.27
CODa	92.08		

Baseline: Train/Test on CODa

With CODa Labels

Without CODa Labels

[1]Caesar et al, "nuScenes: A multimodal dataset for autonomous driving", CVPR 2020
[2]Sun et al, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset", CVPR 2020
[3]Shaoshuai et al, "PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection", CVPR 2020

AM RL

Evaluation on Pedestrian-only Detection

Train \ Test	Precision	Recall	F1	# Pedestrian Labels
nuScenes	42.12	11.83	18.48	2.10 million
Waymo	52.76	17.19	25.94	2.03 million
CODa	57.38	25.31	35.13	0.57 million
JRDB	64.15	27.15	38.15	1.8 million

Worst  Best

Geiger et al, “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”, CVPR 2012
Sun et al, “Scalability in Perception for Autonomous Driving: Waymo Open Dataset”, CVPR 2020
Caesar et al, “nuScenes: A multimodal dataset for autonomous driving”, CVPR 2020
Yin et al, “Center-based 3D Object Detection and Tracking”, CVPR 2021
Martin Martin et al, “JRDB: A Dataset and Benchmark of Egocentric Robot Visual Perception of Humans in Built Environments”, TPAMI 2021

An aside: How far will terrain-aware navigation take you?

Haresh Karnan, Elvin Yang, Daniel Farkash, Garrett Warnell, Joydeep Biswas, Peter Stone (2023). Self-Supervised Terrain Representation Learning from Unconstrained Robot Experience. *Conference on Robot Learning (CoRL) 2023*

STERLING Extended Deployment

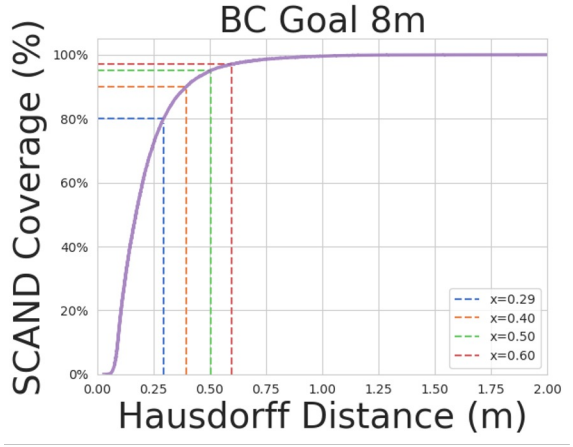
Haresh Karnan, Elvin Yang, Daniel Farkash, Garrett Warnell, Joydeep Biswas, Peter Stone (2023). Self-Supervised Terrain Representation Learning from Unconstrained Robot Experience. *Conference on Robot Learning (CoRL) 2023*

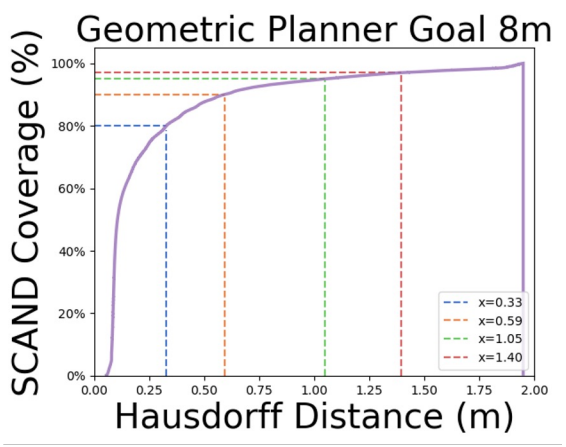
STERLING Extended Deployment

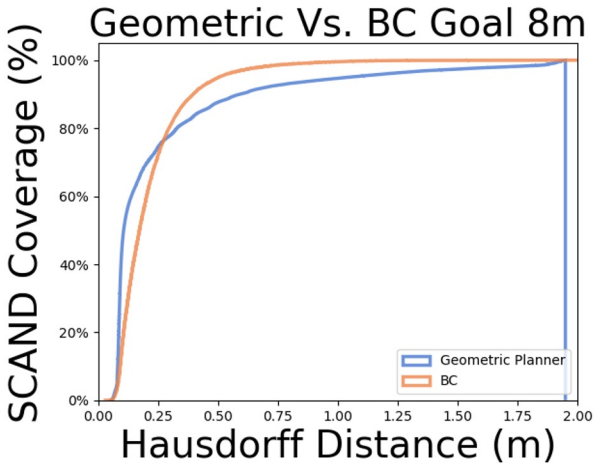
How about direct behavior cloning (no object detection)? + A hybrid approach

Amir Hossain Raj, Zichao Hu, Haresh Karnan, Rohan Chandra, Amirreza Payandeh, Luisa Mao, Peter Stone, Joydeep Biswas, Xuesu Xiao (2023). Targeted Learning: A Hybrid Approach to Social Robot Navigation. Arxiv preprint: 2309.13466

End-to-end Behavior Cloning







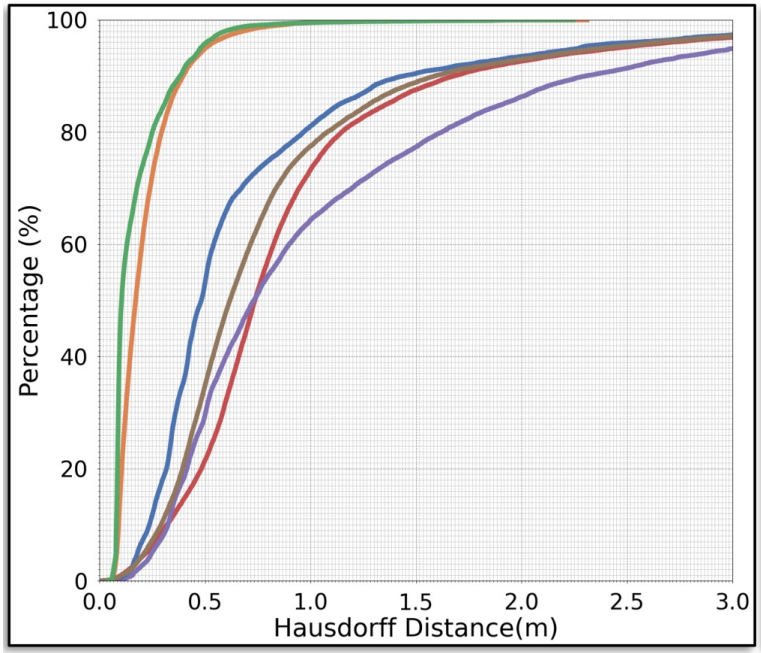
80% 90% 95% 97%

Hybrid Geometric-Learned Navigation

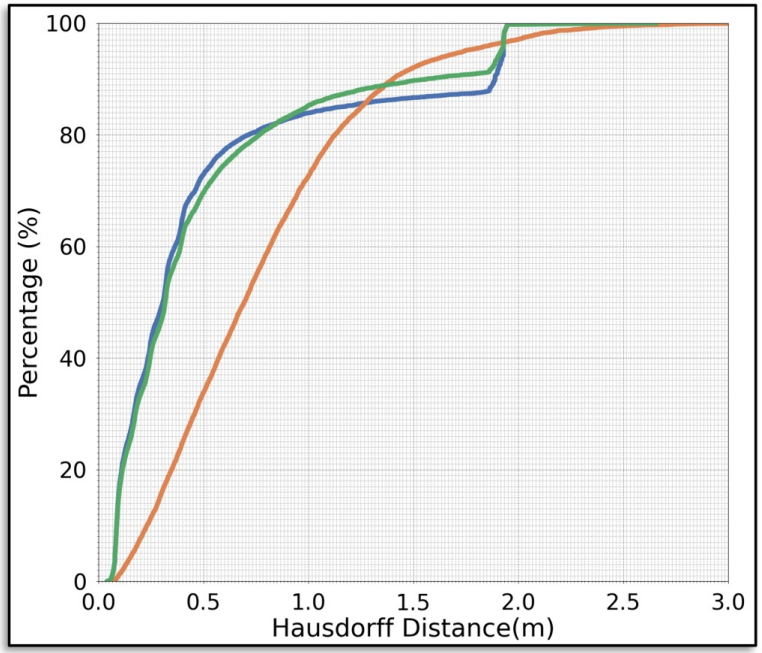
```
if (ExpectGeometryGood(s))
    return GeometricNavigation(s)
else
    return BehaviorCloning(s)
```

Amir Hossain Raj, Zichao Hu, Haresh Karnan, Rohan Chandra, Amirreza Payandeh, Luisa Mao, Peter Stone, Joydeep Biswas, Xuesu Xiao (2023). Targeted Learning: A Hybrid Approach to Social Robot Navigation. Arxiv preprint: 2309.13466

Hybrid Planner: Geometric + Behavior Cloning



In-distribution



Out-of-distribution

Hybrid | move_base | BC | Human-Aware | move_base(s) | CoHAN

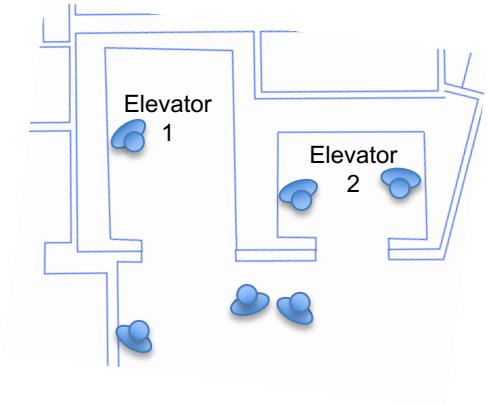
How about scenarios that depend only on multiagent factors? Enter: Social Minigames

Rohan Chandra, Vrushabh Zinage, Efstathios Bakolas, Joydeep Biswas, Peter Stone (2023). Decentralized Multi-Robot Social Navigation in Constrained Environments via Game-Theoretic Control Barrier Functions. Arxiv preprint: 2308.10966

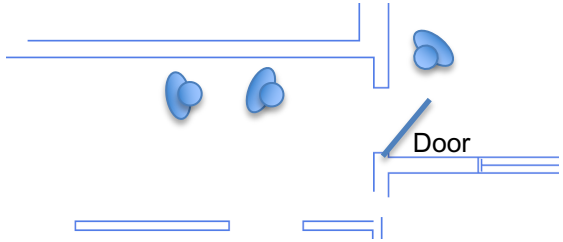
Zayne Sprague, Rohan Chandra, Jarrett Holtz, Joydeep Biswas (2023). SOCIALGYM 2.0: Simulator for Multi-Agent Social Robot Navigation in Shared Human Spaces. Arxiv preprint: 2303.05584

Jarrett Holtz, Joydeep Biswas (2022). SOCIALGYM: A Framework for Benchmarking Social Robot Navigation. In Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on, pp. 11246-11252.

Social Mini-Games

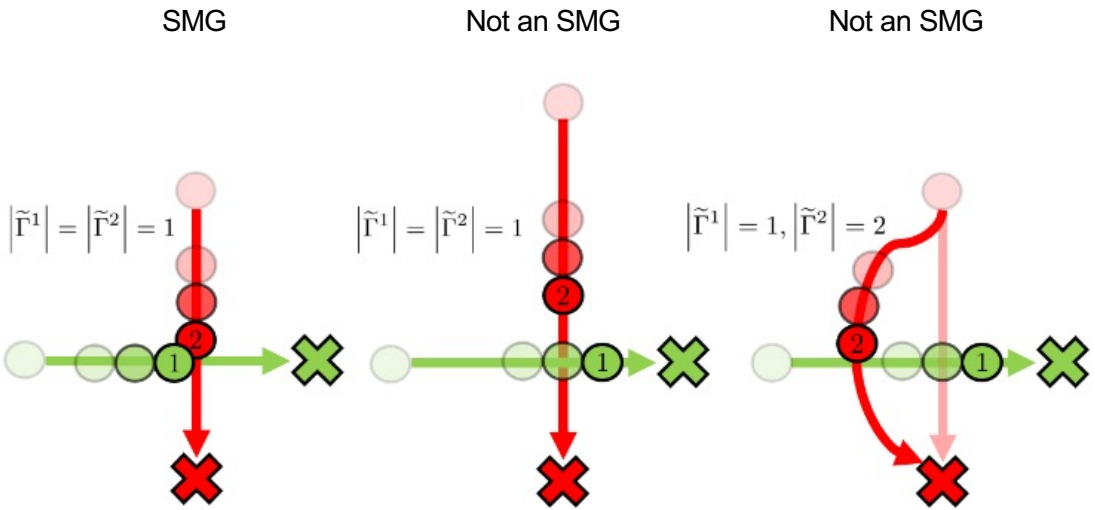


The Elevator Mini-Game



The Doorway Mini-Game

Social mini-games (SMGs)



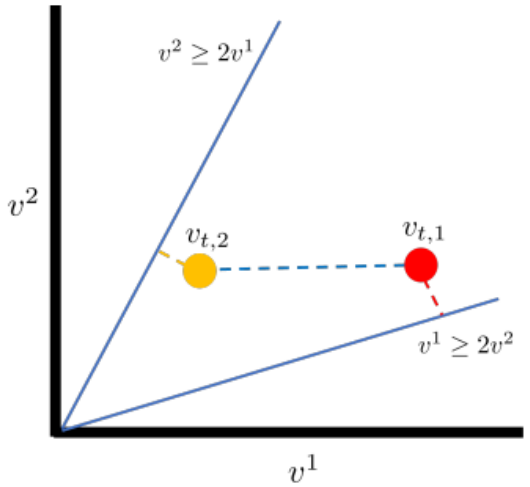
- Each agent has a set of preferred timed trajectories.
- If all the preferred trajectories intersect over some time interval, the scenario is an SMG.

Rohan Chandra, Vrushabh Zinage, Efstathios Bakolas, Joydeep Biswas, Peter Stone (2023). Decentralized Multi-Robot Social Navigation in Constrained Environments via Game-Theoretic Control Barrier Functions. Arxiv preprint: 2308.10966

Key idea

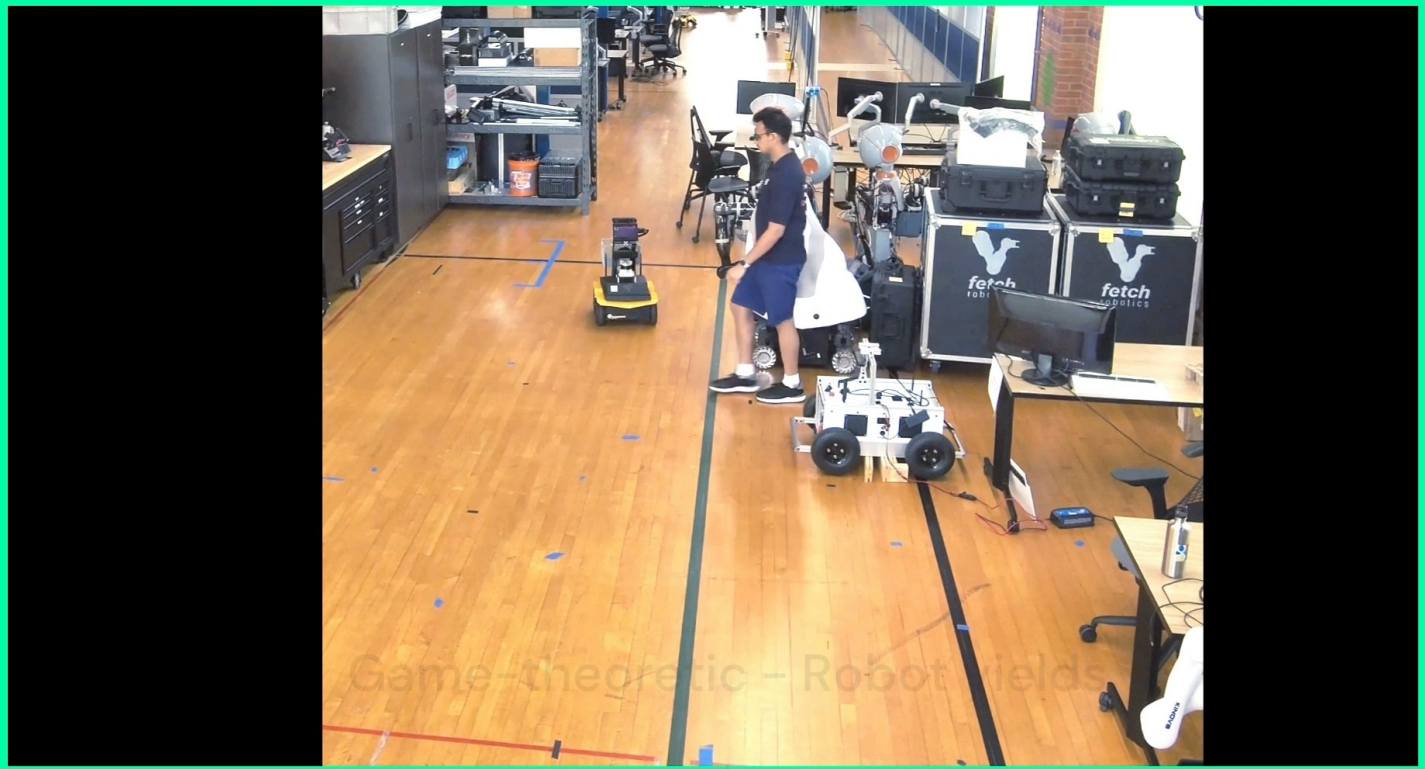
Each agent projects their velocity onto liveness and safety sets.

$$\widetilde{v}_t = \arg \min_{\mu \in \mathscr{C}_\ell(t)} \|v_t - \mu\|_2$$

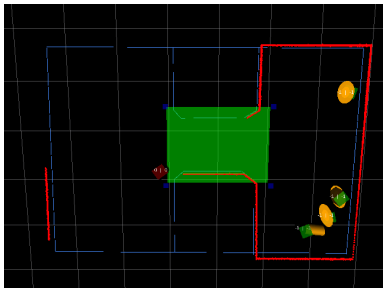


Rohan Chandra, Vrushabh Zinage, Efstathios Bakolas, Joydeep Biswas, Peter Stone (2023). Decentralized Multi-Robot Social Navigation in Constrained Environments via Game-Theoretic Control Barrier Functions. Arxiv preprint: 2308.10966

Real Robot Experiments



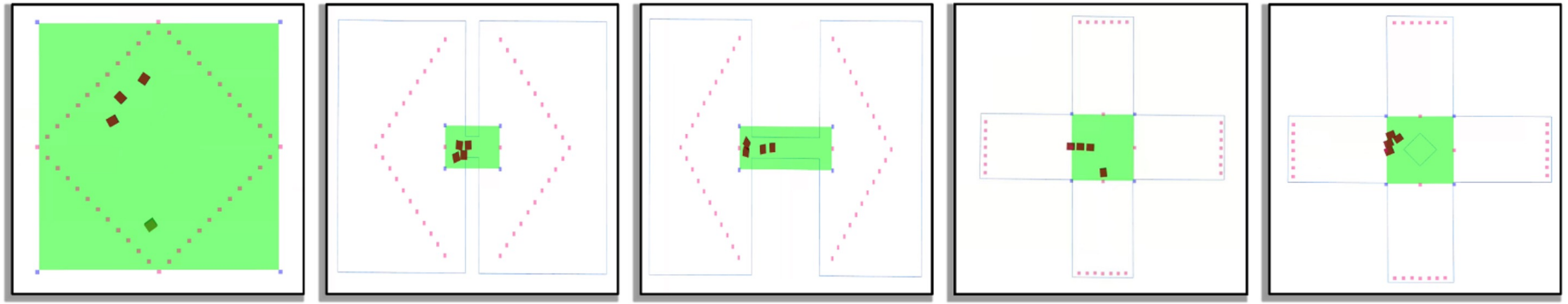
SocialGym 2.0: Simulator for Multi-Agent Social Robot Navigation in Shared Human Spaces



https://github.com/ut-amrl/SocialGym2

Sprague, Z., **Chandra, R**., Holtz, J., & Biswas, J. SOCIALGYM 2.0: Simulator for Multi-Agent Social Robot Navigation in Shared Human Spaces. (submitted to) RSS (2023).

Social mini-games



(a) Open Scenario (b) Doorway Scenario (c) Hallway Scenario (d) Intersection Scenario (e) Roundabout Scenario

Concluding Questions

- How often, and under what circumstances is geometric navigation not sufficient?
- Which entities are of relevance for context-aware social navigation, and how can we detect them?
- Rethinking Social Navigation: what will be the winning hybrid between geom. and learned policies?
- How do we account for all contextual factors cohesively in ~~social-~~ ***everything-*** navigation?

https://amrl.cs.utexas.edu/coda/