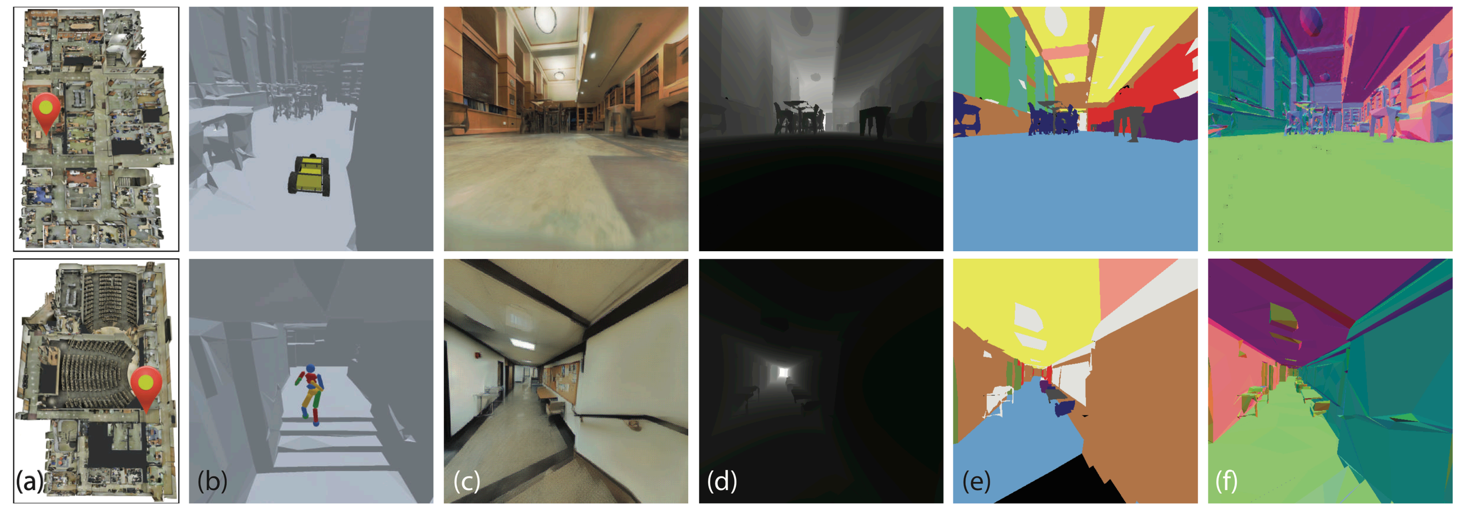
Gibson Env: Real-World Perception for Embodied Agents

Fei Xia*[1], Amir R. Zamir*[1,2], Zhiyang He*[1], Alexander Sax[1], Jitendra Malik[2], Silvio Savarese[1]

http://gibsonenv.stanford.edu

[1] Stanford University [2] Berkeley University of California

Introduction



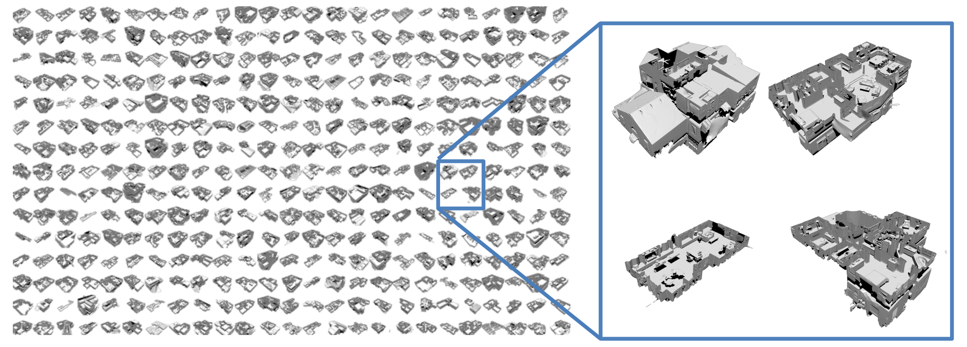
From datasets to environments

- Learning in **real world**: slow, fragile
- Learning in **simulation**: generalization difficulties: (1) photorealism (2) semantic distribution mismatch

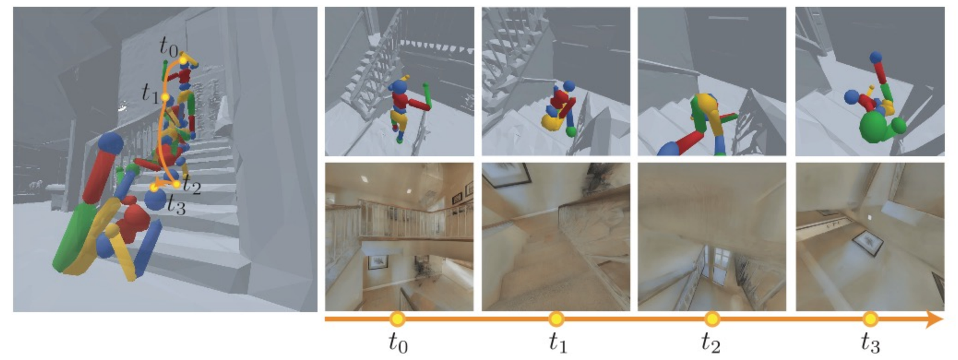
we propose

Gibson Environment

1. Database: real-world RGBD panoramas

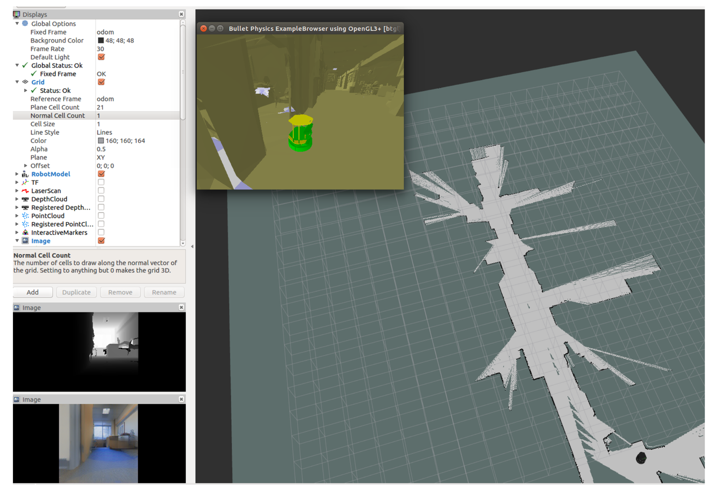


2. Physics engine: PyBullet

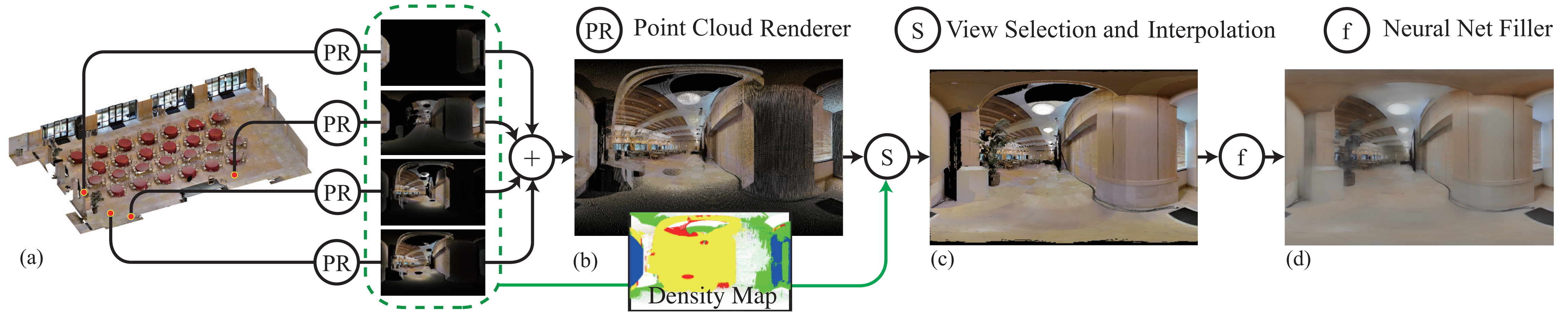


3. View synthesis: Neural Network filler and Goggle mechanism (Method Section)

4. ROS integration & Gym integration



Rendering Engine



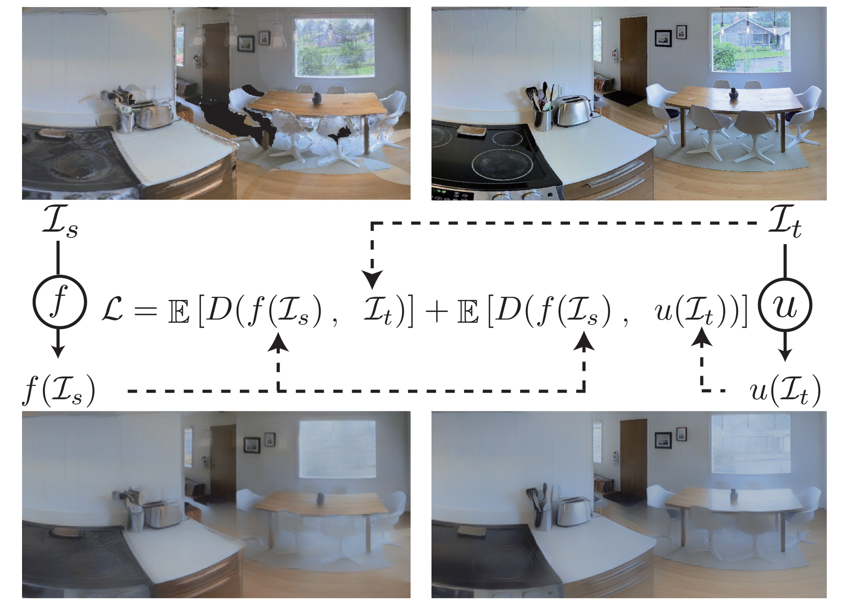
Our rendering engine has 3 stages:

1. (a)->(b) Point cloud rendering: **reproject** points to new view
2. (b)->(c) View selection: adaptively **select** view to retrieve points from; Interpolation: **interpolate** rendered points to image
3. (c)->(d) Neural network: **fill** holes and fix geometry issues

Training techniques:

- Color matching loss
- Stochastic identity initialization
- Perceptual loss for $D(I_1, I_2)$

Goggle Mechanism



Close domain gap

Point Cloud Rendering → Image (f, u)

f: input point cloud output rendering — filler

u: input real images output rendering — goggles

$\mathcal{I}_s$: point cloud rendering $\mathcal{I}_t$: real images

$f(\mathcal{I}_s)$: NN rendering $u(\mathcal{I}_t)$: real images through goggle

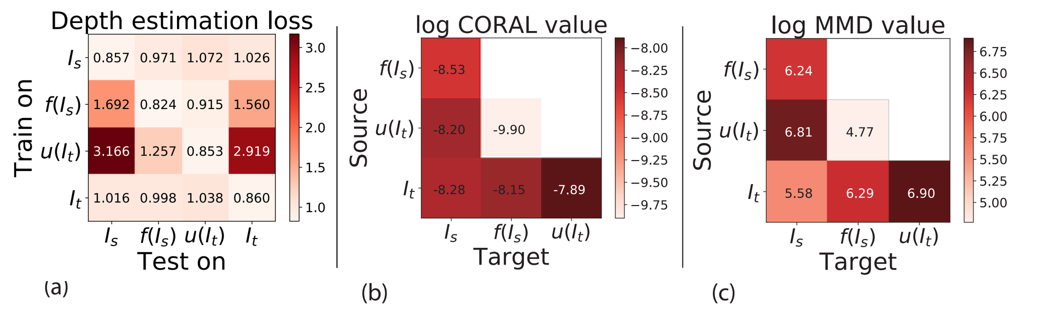
Experimental Results

Dataset comparison

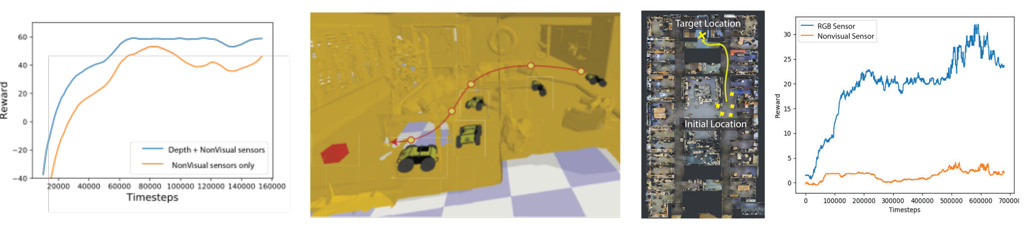
Dataset	Gibson	SUNCG	Matterport3D
Number of Spaces	572	45622	90
Total Coverage m^2	211k	5.8M	46.6K
SSA	1.38	0.74	0.92
Nav. Complexity	5.98	2.29	7.80
Real-World Transfer Err	$0.92^{\S}$	$2.89^{\dagger}$	$2.11^{\dagger}$

View Synthesis Qualitative Results

Transferring to Real World Results



Example tasks in Gibson



Visual obstacle avoidance Distant navigation

Conclusions and Limitations

We propose Gibson Environment for developing real world perception for active agents.

Limitations (future work):

- Dynamic contents
- Manipulation

Source code https://github.com/StanfordVL/GibsonEnv

Download Gibson dataset: http://gibsonenv.stanford.edu

Browser dataset online: http://gibsonenv.stanford.edu/database/

(* Equal Contribution)