RL Applications

Kunal Pratap Singh

27/03/2025

École polytechnique fédérale de Lausanne

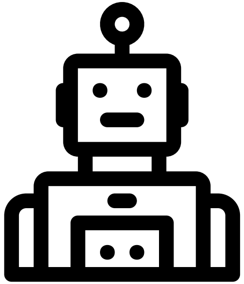
Embodied AI

1. Dietke et al., Retrospectives on the Embodied AI workshop, 2022

Embodied AI

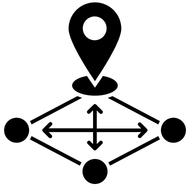
Navigation

Navigation



Goal

“Find a bed”



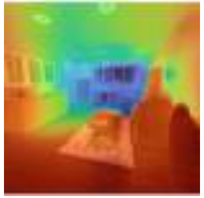
GPS

Navigation

PointNav

RGB

Depth

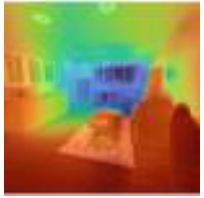


GPS

ObjectNav

RGB

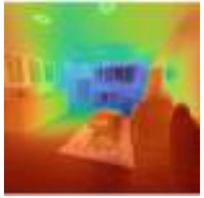
Depth



ImageNav

RGB

Depth



Sensors

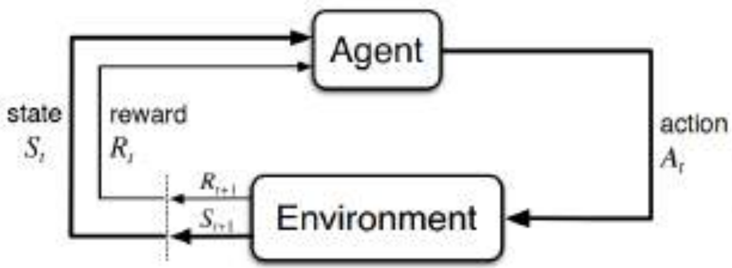
Target

GPS

"Find me a bed"

Navigation

Reinforcement learning in Navigation



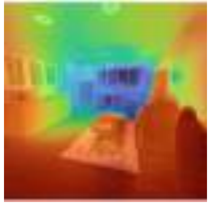
Agent

Environment

State, S_t

RGB

Depth



Action, A_t

Forward, Backward
RotateLeft, RotateRight

Reward, R_t

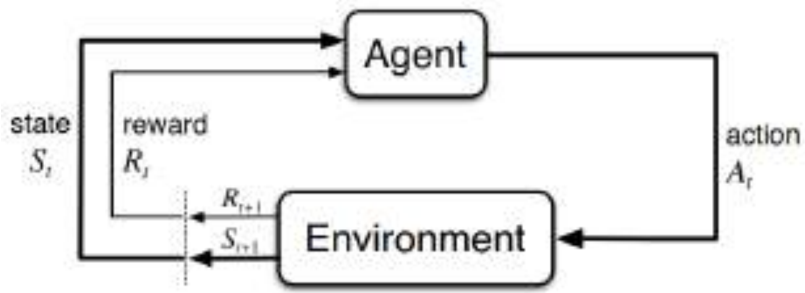
?

Navigation

Reinforcement learning in Navigation

Navigation

Reinforcement learning in Navigation



Reward, R_t

$$R_{success} = \begin{cases} 2.5, & \text{if reach goal} \\ 0, & \text{otherwise} \end{cases}$$

Terminal Reward

$$R_{slack} = -0.01$$

$$R_{progress} = -distance(pos_t, pos_{goal})$$

Reward Shaping

Agent

Environment

Target

Navigation

Reinforcement learning in Navigation

Reward, R_t

$$T = 1$$

$$R_{success} = 0 \quad R_{slack} = -0.01$$

$$R_{progress} = -5$$

$$R_{T=1} = 0 + -0.01 + -5$$

Target : Desk

Navigation

Reinforcement learning in Navigation

Reward, R_t

$$T = 2$$

$$R_{success} = 0 \quad R_{slack} = -0.01$$

$$R_{progress} = -3$$

$$R_{T=2} = 0 + -0.01 + -3$$

Target : Desk

Navigation

Reinforcement learning in Navigation

Reward, R_t

$$T = 3$$

$$R_{success} = 2.5 \quad R_{slack} = -0.01$$

$$R_{progress} = -1$$

$$R_{T=3} = 0 + -0.01 + -1$$

Navigation

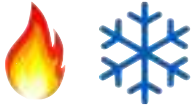
Reinforcement learning in Navigation

Reward, R_t

In Summary,

$$R_{T=1} = -5.01$$

$$R_{T=2} = -3.01$$



$$R_{T=3} = 1.49$$

Target : Desk

Navigation

Reinforcement learning in Navigation

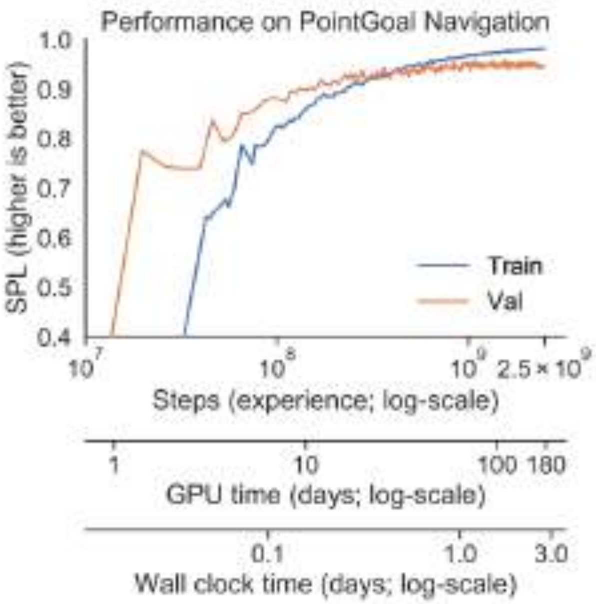
Habitat

ProcTHOR

Gibson

Navigation

DD-PPO: Distributed Decentralised PPO[1]



1. Wijmans et al. DD-PPO: Learning Near-Perfect PointGoal Navigators from 2.5 Billion frames

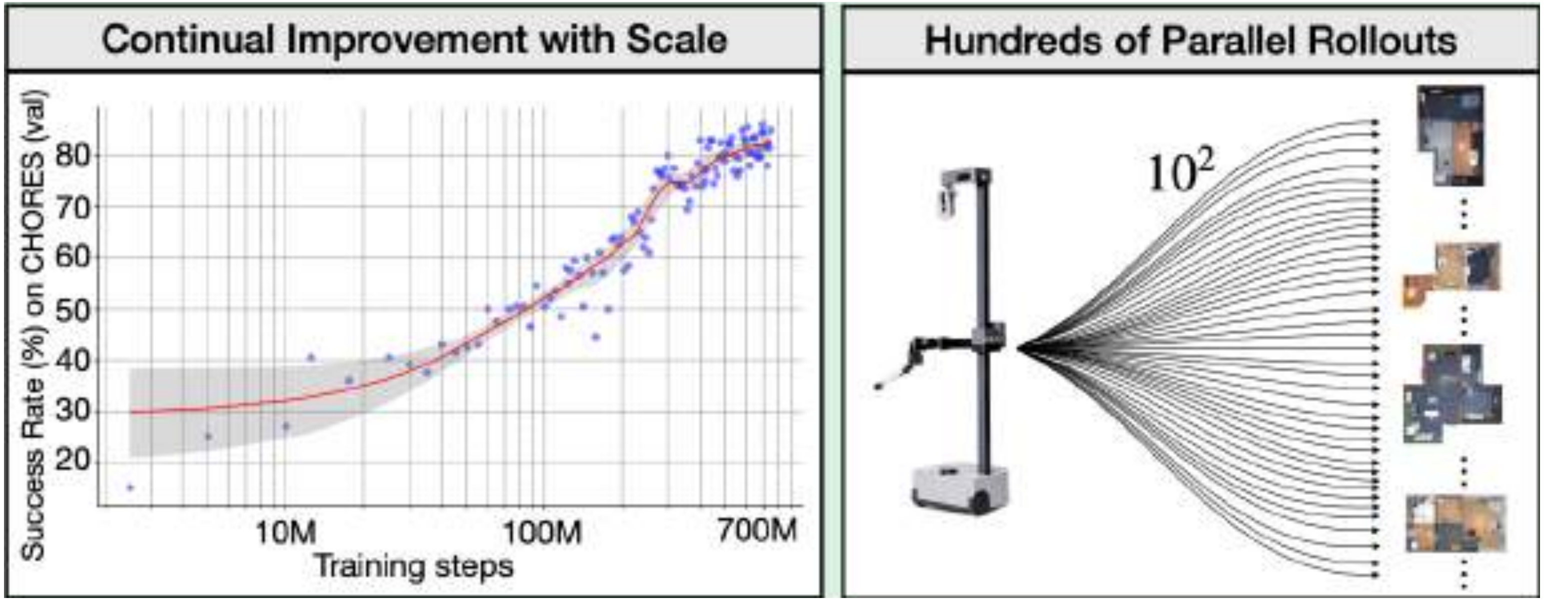
Navigation

DD-PPO: Decentralised Distributed PPO[1]

1. Wijmans et al. DD-PPO: Learning Near-Perfect PointGoal Navigators from 2.5 Billion frames

Navigation

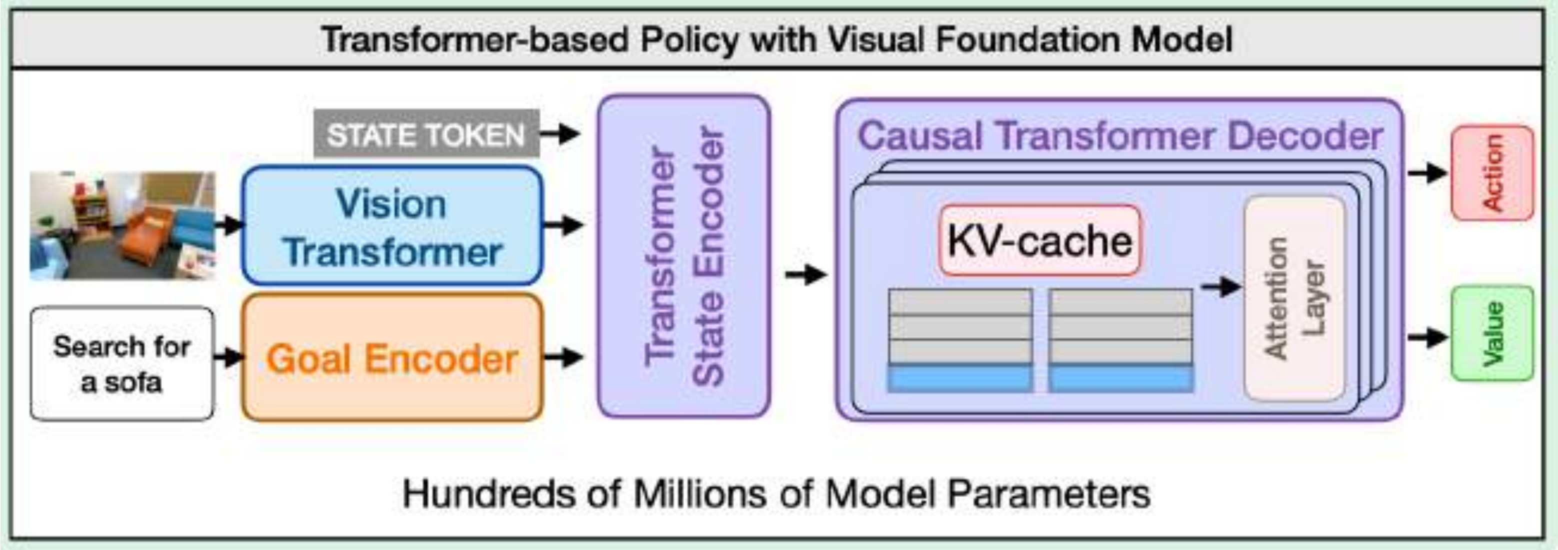
Poliformer: Scaling On-policy RL with Transformers[1]



1. Zeng et al. PoliFormer: Scaling On-Policy RL with Transformers Results in Masterful Navigators.

Navigation

Poliformer: Scaling On-policy RL with Transformers[1]



1. Zeng et al. PoliFormer: Scaling On-Policy RL with Transformers Results in Masterful Navigators.

Navigation

Poliformer: Scaling On-policy RL with Transformers[1]

1. Zeng et al. PoliFormer: Scaling On-Policy RL with Transformers Results in Masterful Navigators.

Blind Navigation Agents

Emergence of Maps in the memories Blind Navigation Agents[1]

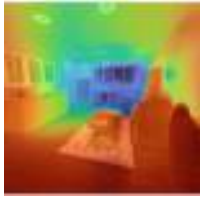
PointNav

RGB

Depth

GPS

Sensors



Target

GPS

Blind PointNav

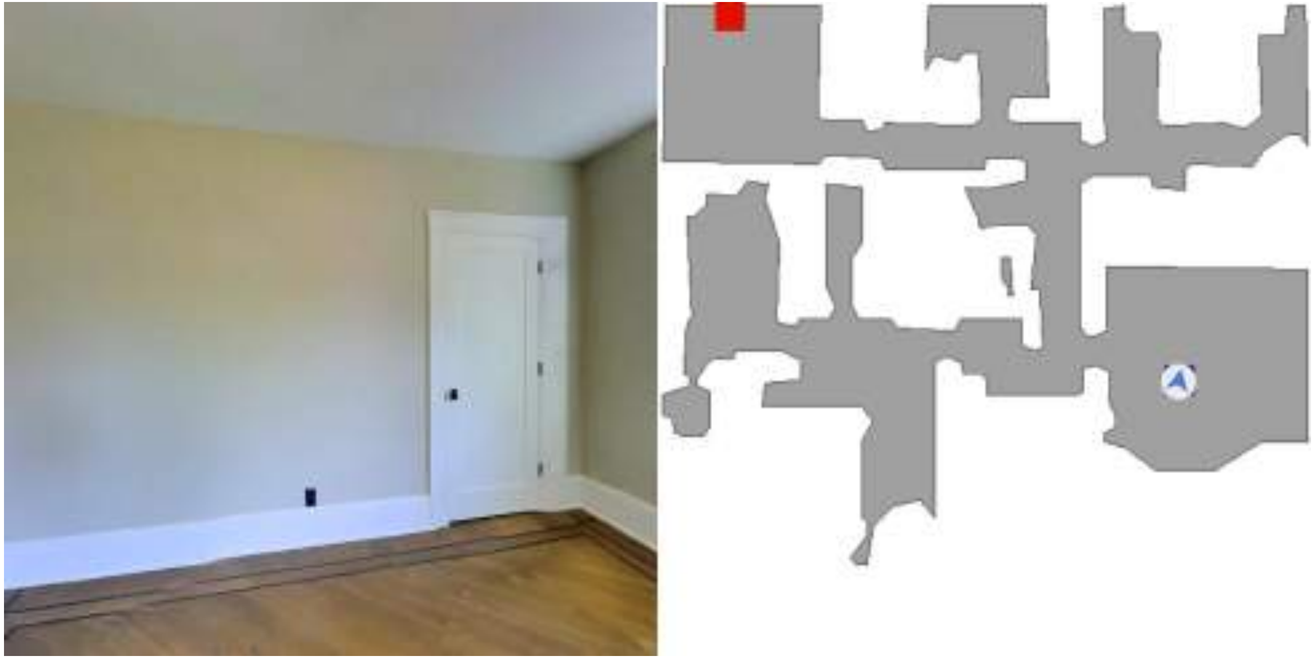
GPS

GPS

1. Wijmans et al. Emergence of Maps in the memory of blind agents.

Blind Navigation Agents

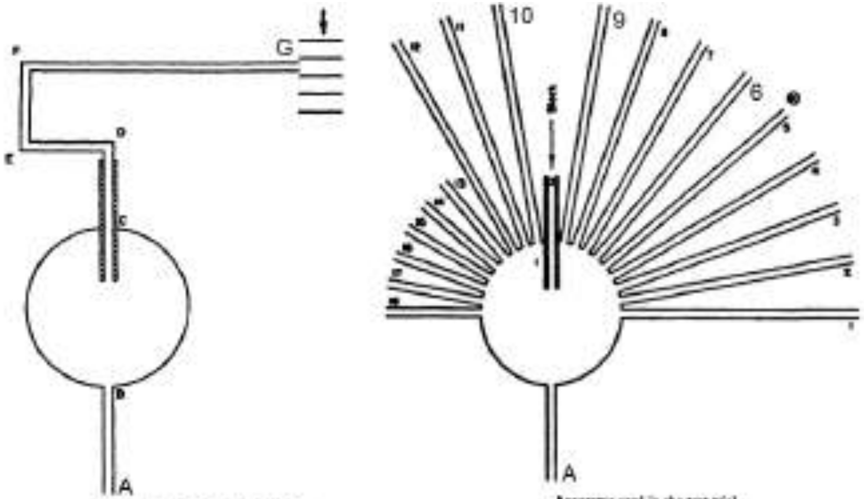
Emergence of Maps in the memories Blind Navigation Agents[1]



1. Wijmans et al. Emergence of Maps in the memory of blind agents.

Blind Navigation Agents

Cognitive Maps (1948)



Apparatus used in preliminary training

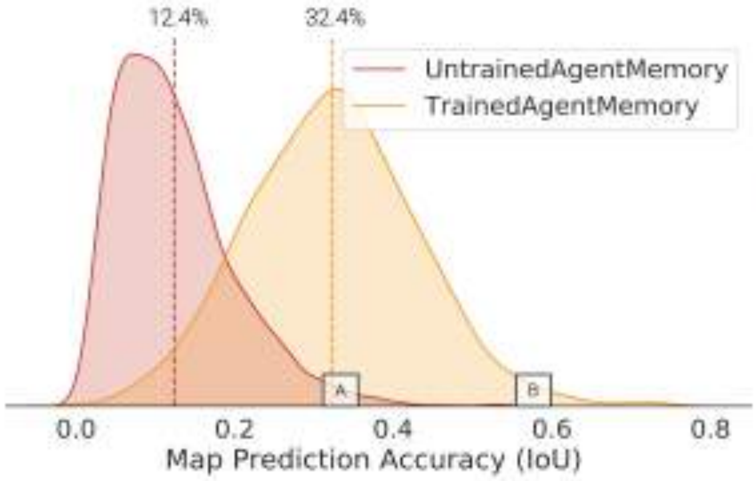
Apparatus used in the test trial

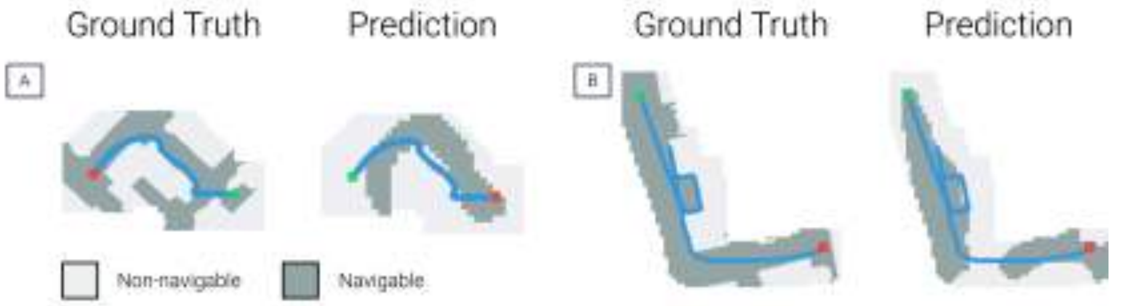
(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and short-cut. *J. exp. Psychol.*, 1946, 36, p. 17.)

Blind Navigation Agents

Emergence of Maps in the memories Blind Navigation Agents[1]

- Trained Blind Agent have better memory of free space in the environment.

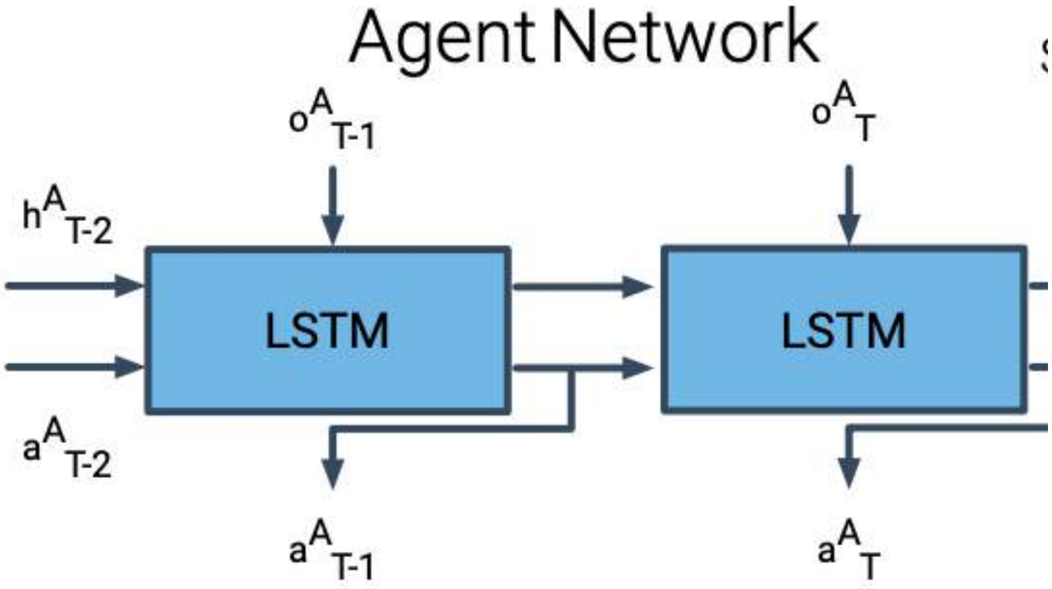




1. Wijmans et al. Emergence of Maps in the memory of blind agents.

Blind Navigation Agents

Emergence of Maps in the memories Blind Navigation Agents[1]

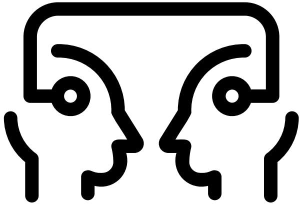


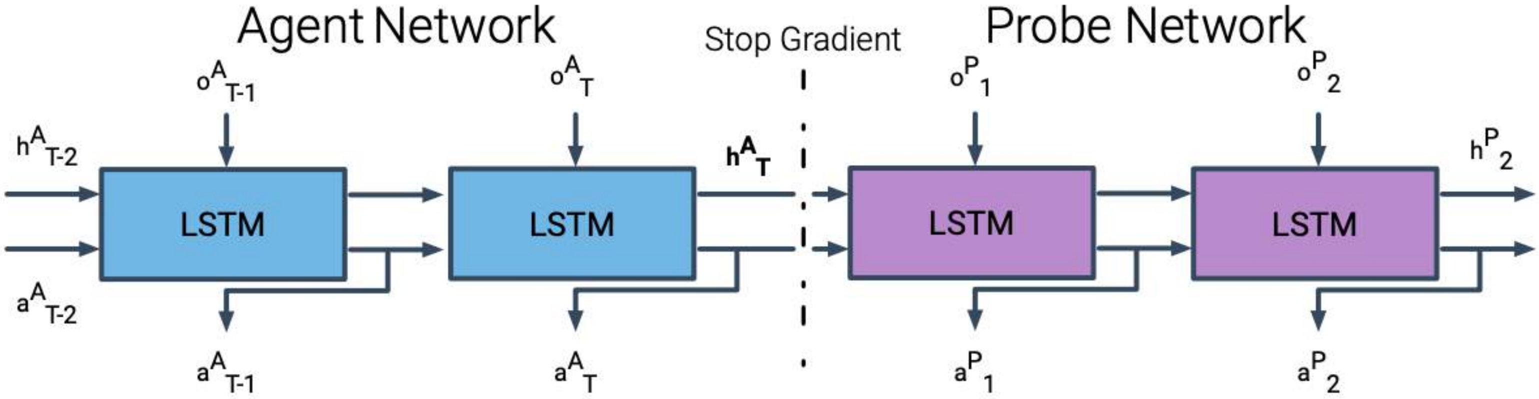
1. Wijmans et al. Emergence of Maps in the memory of blind agents.

Blind Navigation Agents

Emergence of Maps in the memories Blind Navigation Agents[1]

- Transfer memory from an trained agent, to an untrained one.





1. Wijmans et al. Emergence of Maps in the memory of blind agents.

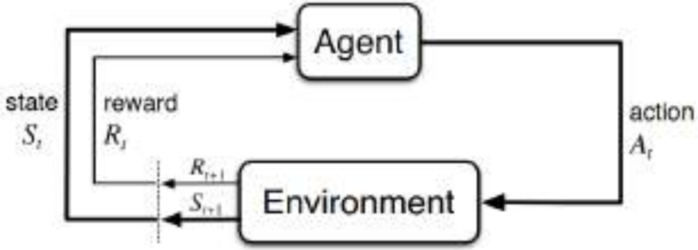
Blind Navigation Agents

Emergence of Maps in the memories Blind Navigation Agents[1]

1. Wijmans et al. Emergence of Maps in the memory of blind agents.

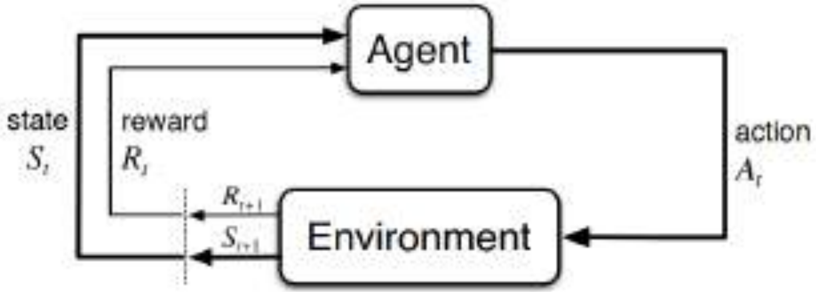
RL is not all you need

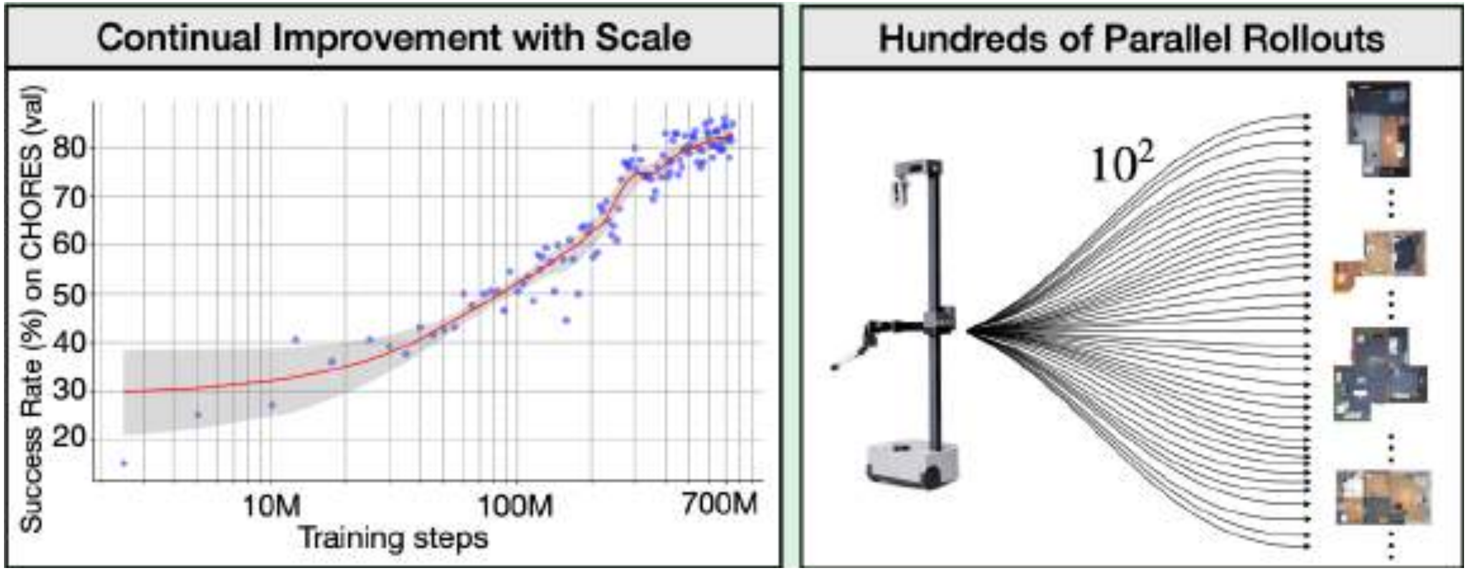
- Reward engineering is hard for complex tasks.



RL is not all you need

- RL is sample inefficient.
- Millions of interactions for a task.





Recent success in NLP / Vision

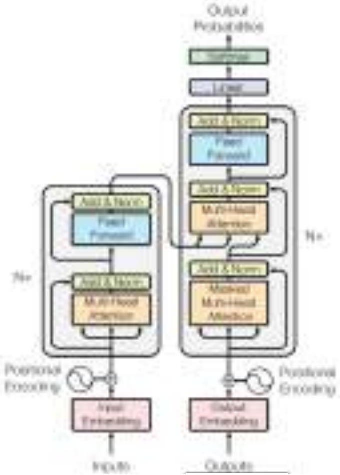
Foundation Models

Recent success in NLP / Vision

Foundation Models - Two key components

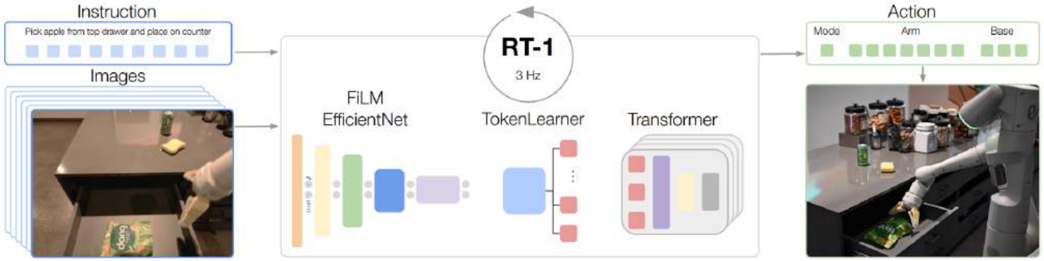
Large scale data

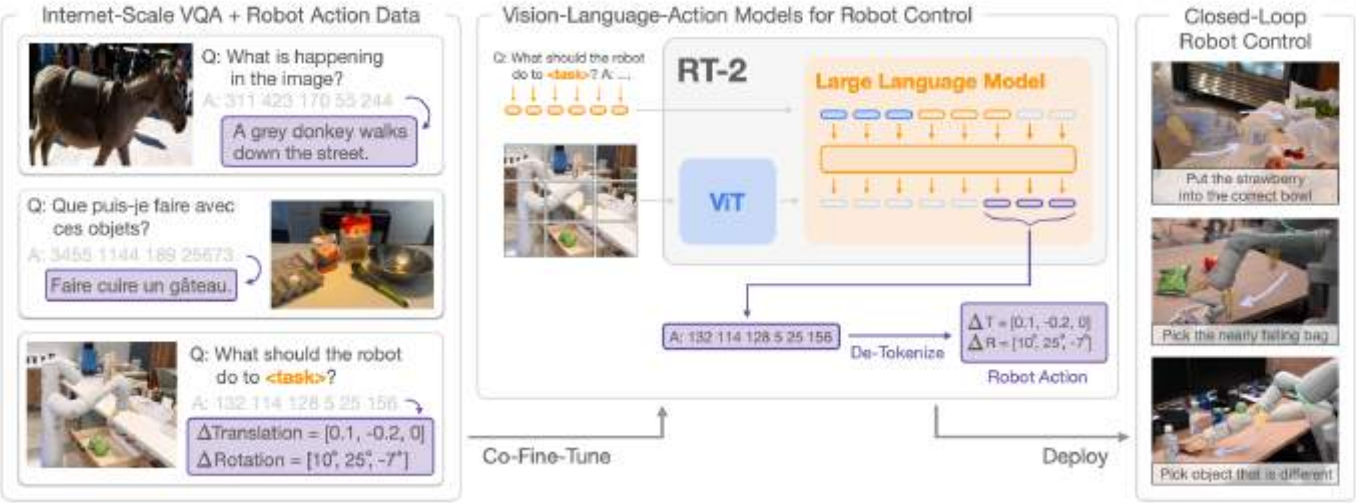
Architecture



ChatGPT for Robotics

Foundation Model for Robotics





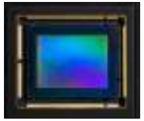
Ecological Approach

- Gibson: “Ask not what’s inside your head, but what is your head inside.”

Visual Sensors in Computer Vision and Nature

Computer Vision *status quo*

- High-resolution Camera
- Intuitive design (by humans)



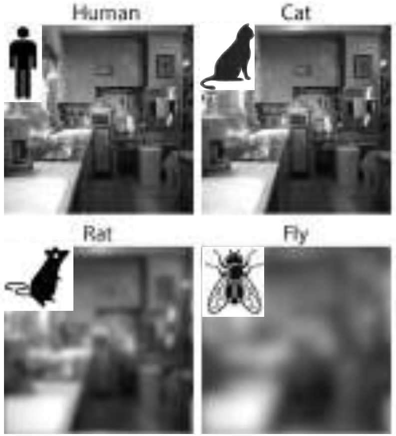
Camera

(e.g., 128x128 = 16,384)

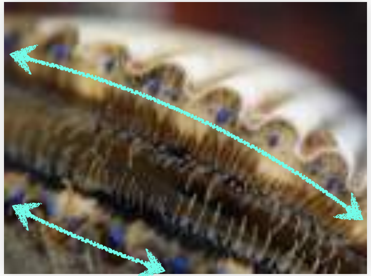
Nature (simple eyes)

- Simple low-resolution eyes
- Optimized design (via evolution)

Qualitative comparison of eyes resolutions:



Distributed Design of Primitive Eyes of a Scallop

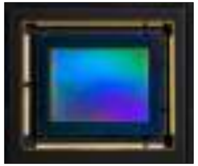


Visual Acuity and the Evolution of Signals, Caves et al. 2018.

Slide Credit: Visual Morphology

How far can a 1-pixel camera go?

- We use simple photoreceptor sensors (**<1%** of a camera resolution).



Camera

(e.g., 128x128 = 16,384)

Photoreceptor

(1x1 = 1)

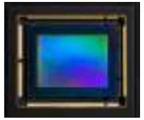
Camera

Photoreceptors

- Slide Credit: Visual Morphology

How far can a 1-pixel camera go?

- Can simple photoreceptor sensors solve vision tasks?
- What is the role of their design?
- Develop a computational design optimization method.



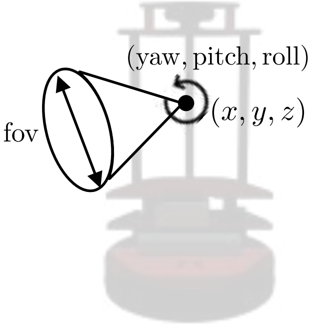
Camera

(e.g., 128x128 = 16,384)

Photoreceptor

(1x1 = 1)

Vision tasks:
e.g., visual navigation.

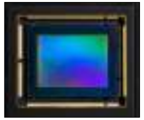


Visual sensors **design.**

- Slide Credit: Visual Morphology

How far can a 1-pixel camera go?

- **Can simple photoreceptor sensors solve vision tasks?**
- What is the role of their design?
- Develop a computational design optimization method.



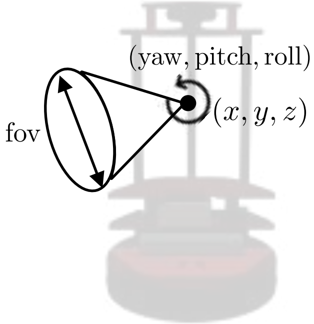
Camera

(e.g., 128x128 = 16,384)

Photoreceptor

(1x1 = 1)

Vision tasks:
e.g., visual navigation.



Visual sensors **design.**

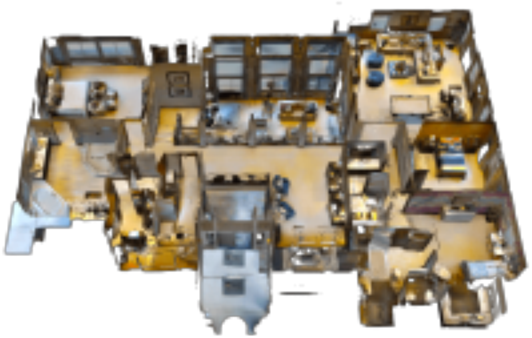
Slide Credit: Visual Morphology

Visual Navigation using Photoreceptors

Visual Navigation Task:

Task: navigate to **the target** in an unseen environment

MatterPort3D Scans of Real-World Apartments



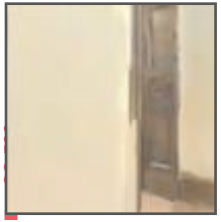
- Slide Credit: [Visual Morphology]

Visual Navigation using Photoreceptors

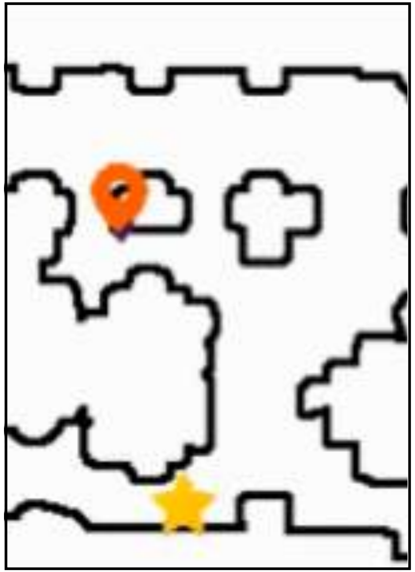
- A <u>handful of PRs can be enough</u> to solve visual navigation meaningfully well.

Observation:
32 **Photoreceptors**

Camera View
(**for visualization only**)



Top-down map:
: Start Position : Goal Position

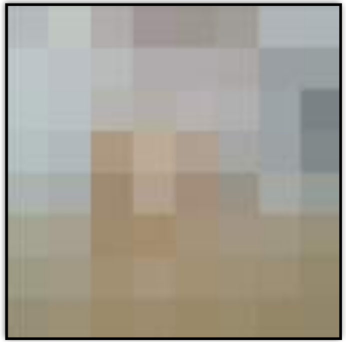


Success Rate on Unseen Scenes

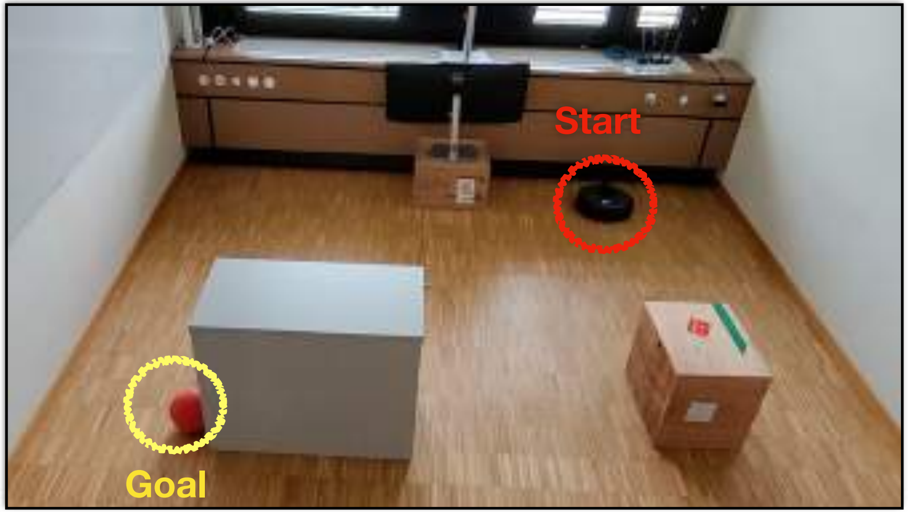
Simple Photoreceptors in the Real World

- Simple photoreceptors show non-trivial generalization to the real world.

Visual Observation:
64 **Photoreceptors**



Third-Person View:
(for visualization only)



Robot Camera View
(**for visualization only**)

Slide Credit: Visual Morphology

Credit: Jason Toskov for setting up and running this experiment.

Continuous Control with Photoreceptors

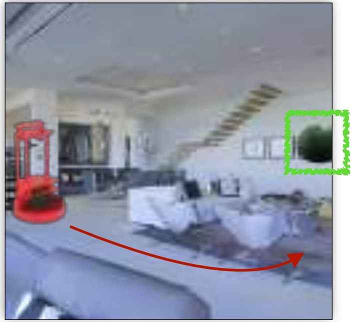
- Example of the Walker: Walk task solved with 4 photoreceptors.

Placement of **Photoreceptors:**

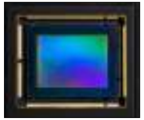
Slide Credit: Visual Morphology

How far can a 1-pixel camera go?

- Can <u>simple photoreceptor</u> sensors solve vision tasks?
- **What is the role of their <u>design</u>?**
- **Develop a <u>computational design optimization</u> method.**

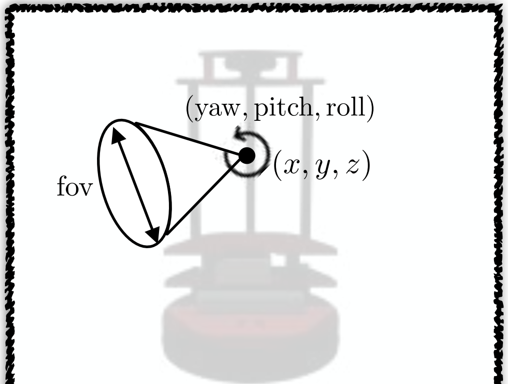


Solve **vision tasks**,
e.g., visual navigation.



Camera
(e.g., 128x128 = 16,384)

Photoreceptor
(1x1 = 1)



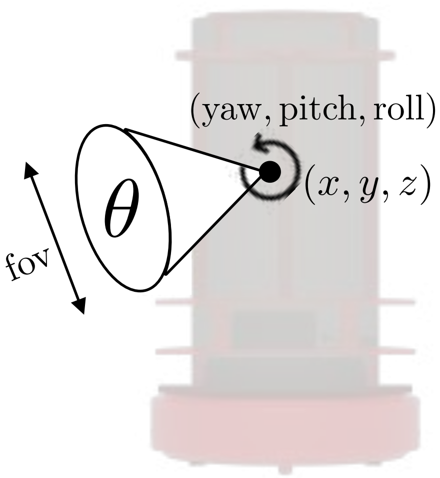
Visual sensors **design.**

Slide Credit: [Visual Morphology]

Designing Visual Sensors

- **Design parameters:** position (on the agent's body), orientation, FoV

Visual Sensor Design:



$$\theta = [x, y, z, \text{yaw}, \text{pitch}, \text{roll}, \text{fov}]$$

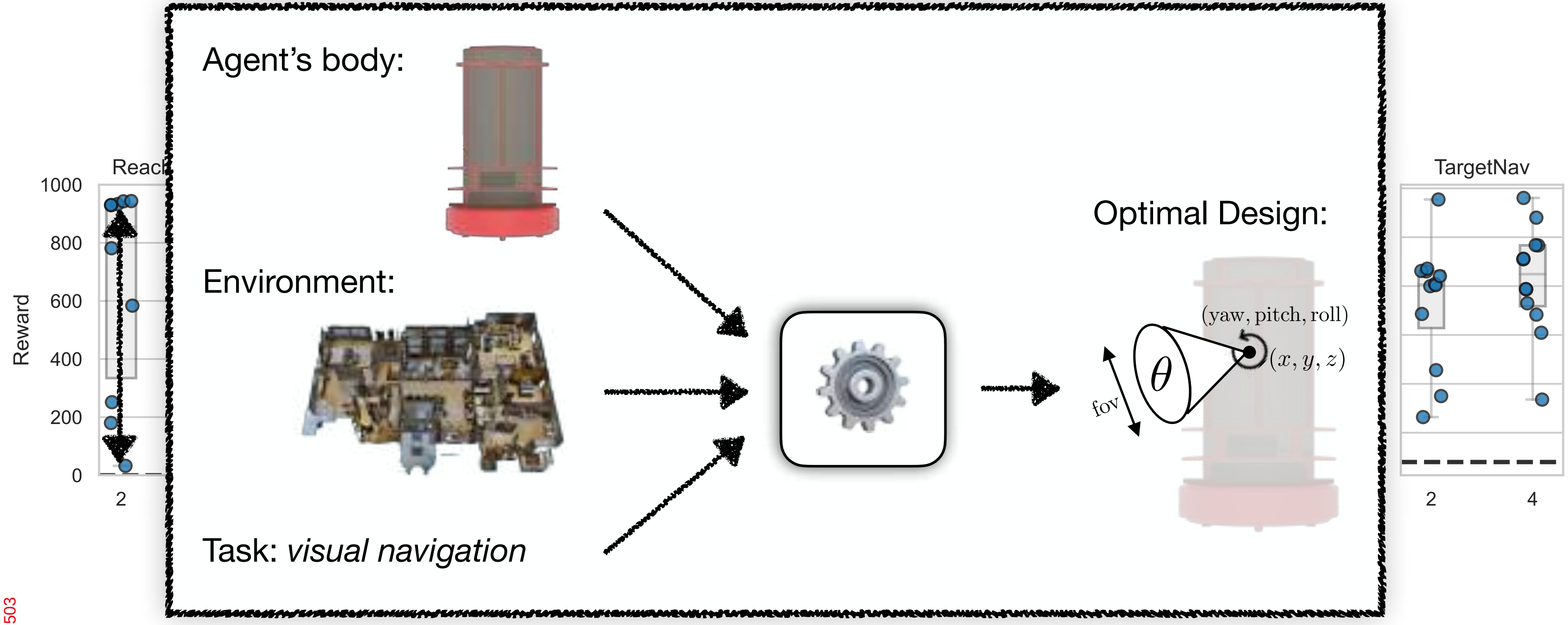
Design Types:

- 🎲 Random design sampled uniformly from the design space.
- 🤖 Computational design optimized for a specific body, environment, and task.
- 🤔 Intuitive design engineered by humans (via a human survey).

- Slide Credit: Visual Morphology

Does the design matter?

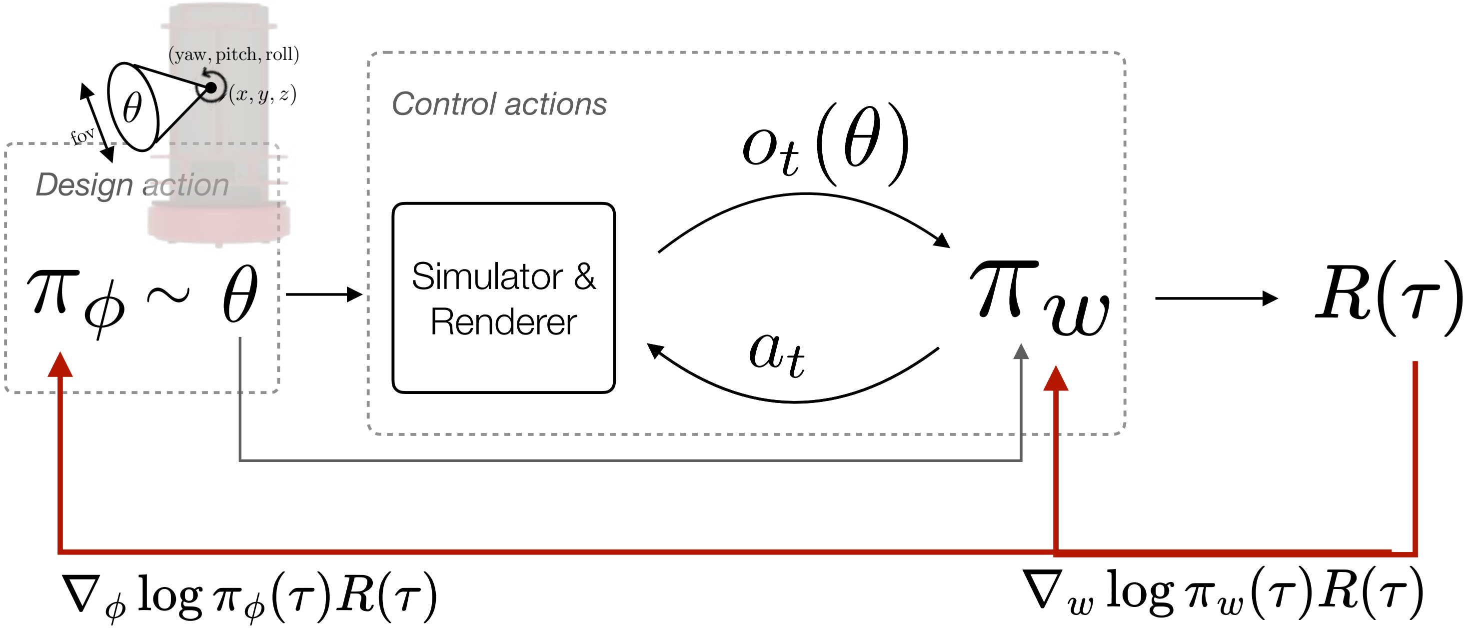
- The design plays a crucial role in the effectiveness of simple visual sensors.



- Slide Credit: Visual Morphology

Computation Design Optimisation

- We jointly optimize both the control and design policy.



- Slide Credit: Visual Morphology

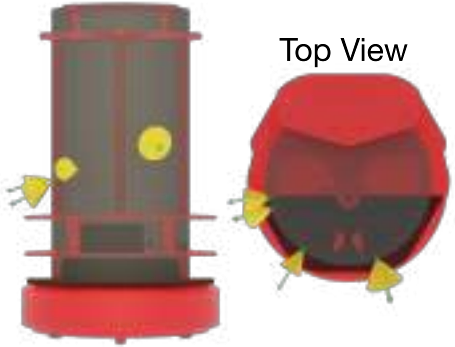
Transform2Act: Learning a Transform-and-Control Policy for Efficient Agent Design, Ye et al., 2022.

Computation Design Optimisation

- Examples designs for navigation tasks.

Front View

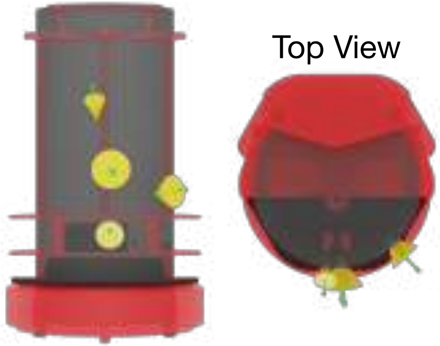
Top View



Initial Design (Random)
PointGoalNav (4x4 grid) SPL = 0.39
TargetNav (8x8 grid) Success = 0.29

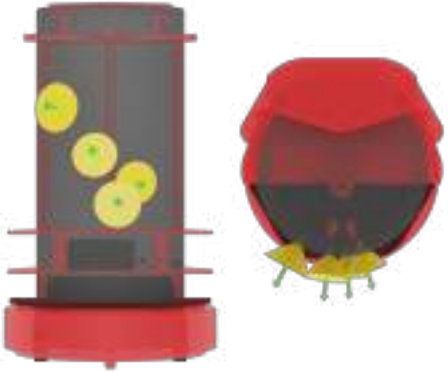
Design Optimization
for *PointGoalNav*

Top View



Design After Optimization
PointGoalNav (4x4 grid) SPL = 0.51

Design Optimization
for *TargetNav*



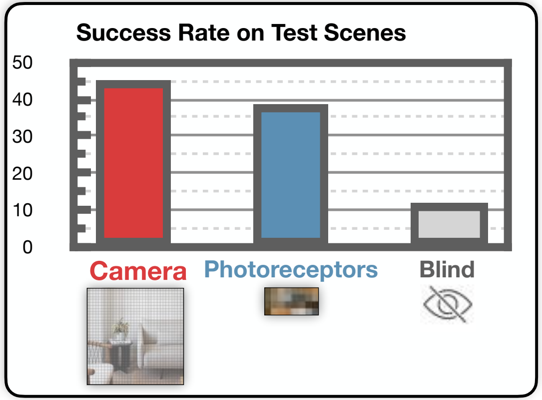
Design After Optimization
TargetNav (8x8 grid) Success = 0.32

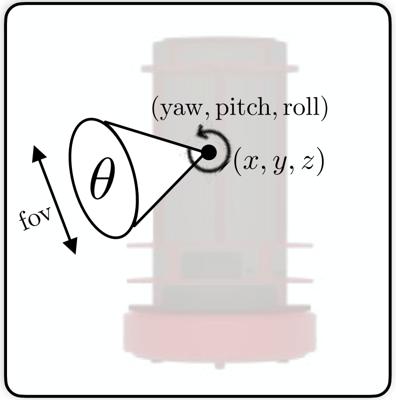
- Slide Credit: [Visual Morphology]

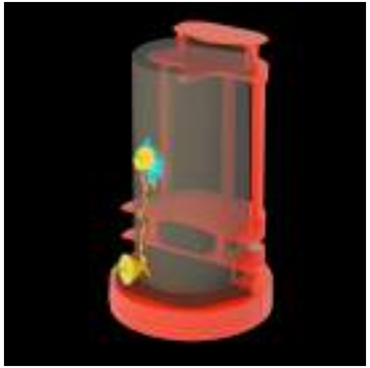
How far can a 1-pixel camera go?

- Simple photoreceptors are effective visual sensors.
- Design is essential for the effectiveness of photoreceptors.
- Computational Design Optimisation using Reinforcement learning.

Project Page:







- Slide Credit: [Visual Morphology]

Questions?

https://vilab.epfl.ch/