CS-503
Visual Intelligence: Machines and Minds

Amir Zamir

16.05.2024

École
polytechnique
fédérale
de Lausanne

Logistics

Logistics

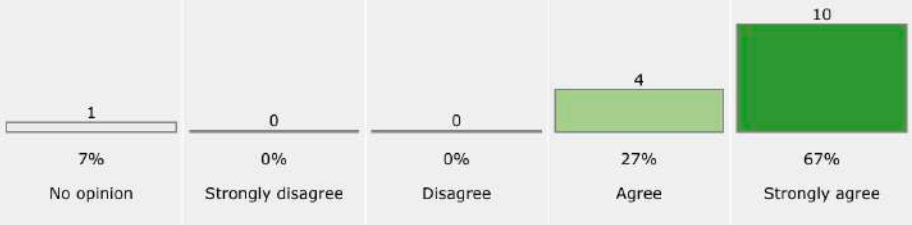
- Next assignment notebook due 01/04/2025 23:59 CET.

Week Num.	Date	Item
1	20.02	- lecture 1
2a	25.02	- lecture 2
2b	27.02	- lecture 3
3a	04.03	- lecture 4
3b	06.03	- lecture 5
4a	11.03	- lecture 6 (+ Q&A)
	11.03	- Transformers notebook assignment due
4b	13.03	- lecture 7
5a	18.03	- lecture 8
5b	20.03	- lecture 9
6a	25.03	- lecture 10
6b	27.03	- lecture 11 (+ Q&A)
	01.04	- Active agents notebook assignment due
7a	01.04	- lecture 12
7b	03.04	- lecture 13
8a	08.04	- lecture 14
8b	10.04	- lecture 15 (+ Matchmaking session)
	13.04	- Project proposals due
	15.04	- all subsequent sessions from 15.04 onwards are for Q&A
	18.04	- Project proposals due, when revision is needed.
	22.04	- MidSem break - No classes
	25.04	- MidSem break - No classes
	29.04	- Foundation Models assignment due
	01.05	- lecture 16
	09.05	- Project progress report due
	13.05	- Robustness assignment due (extra credit)
	20.05	- Moodle homework due
	26.05	- Final project presentation video due
	27.05	- Final project presentation Part I
	29.05	- Final project presentation Part II
	30.05	- Project report due

Logistics

Year	2024-2025
Course	Visual intelligence : machines and minds
Questionnaire	Indicative feedback of teaching (since 2022-2023)
Nb Registered	120
Nb Answered	15

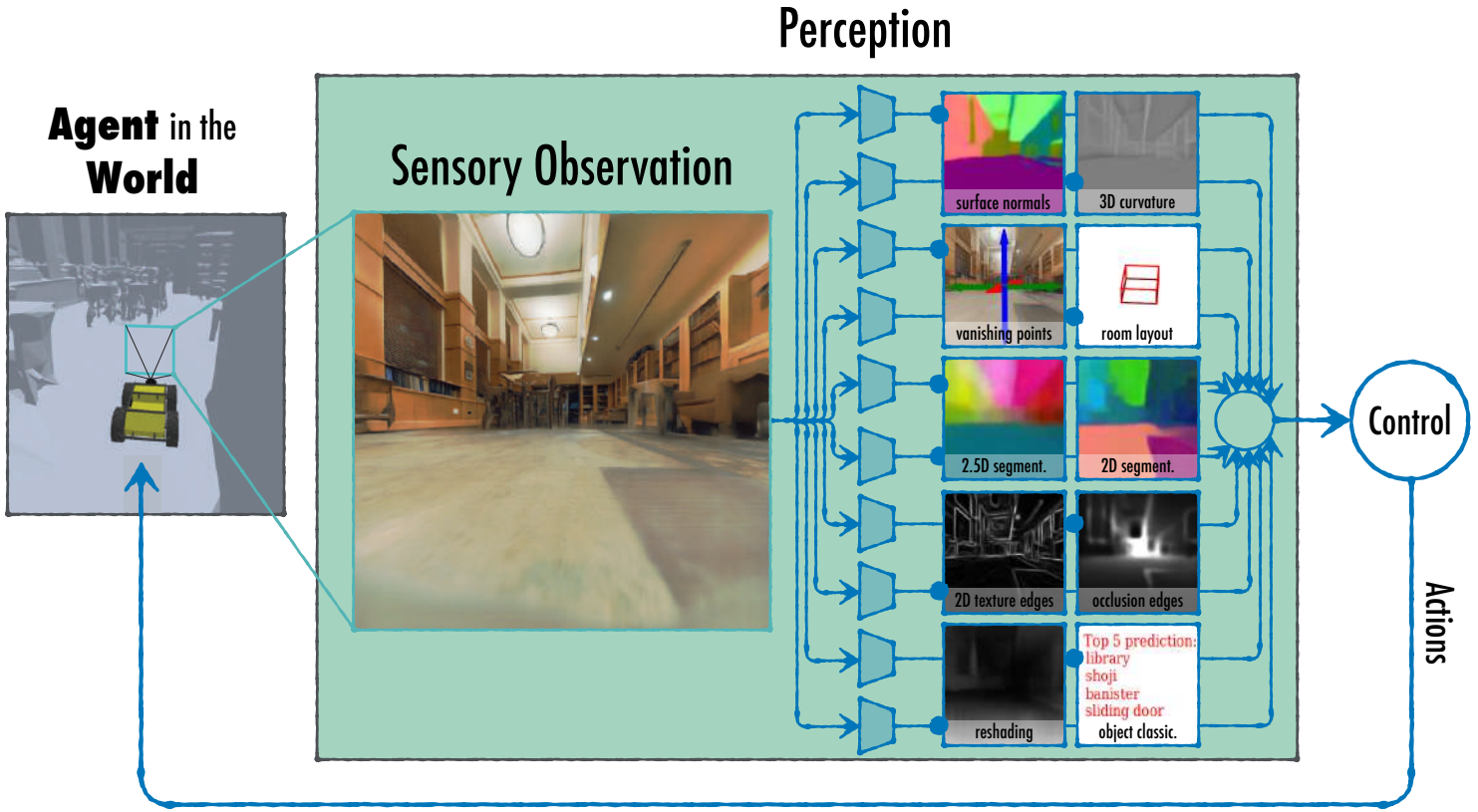
The running of the course enables my learning and an appropriate class climate



- I would recommend to use Edstem for the class forum instead of moodle since it is way better than moodle and "Laughably simple" to use
- One of the best courses at EPFL <3
- The homeworks could have a bit more details or hints about what we should do, because it's sometimes really unclear (like the current one, I spent more time trying to find which function we where supposed to use/ complete than understanding the concepts…
- Very interesting course, thank you very much!…give us access to the recordings of the lecture would help…if for some reason we miss one lecture there is not a lot of information in the slides to catch up the missed lecture.
    - Also the communication of the organization of the course was more confusing than helping.…e.g. not really specifying what the "moodle assignment" will be is a bit uncomfortable.

Recap

Vision In-the-loop




Frequency
Rank order of items

Gibson Environment

Large Real Space

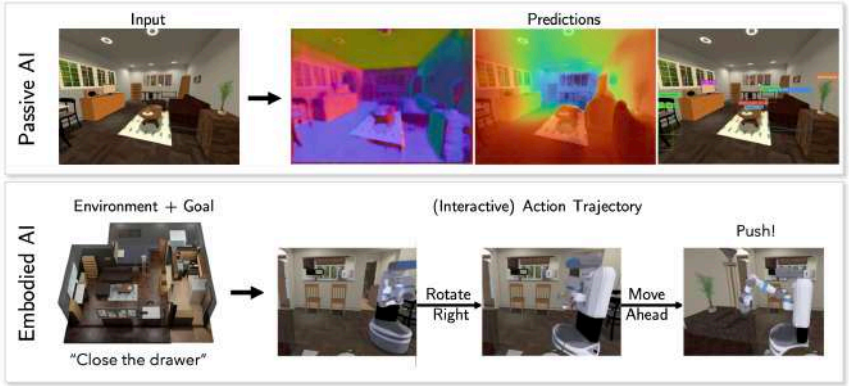
Active Agent

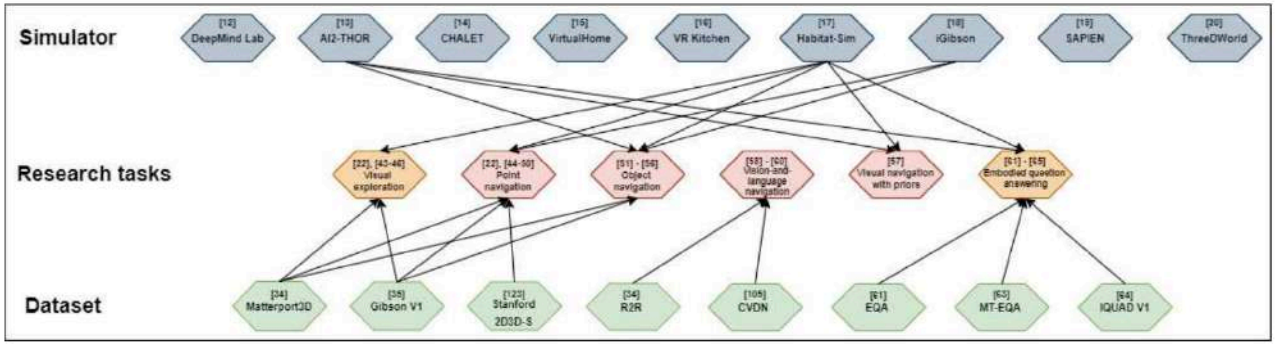
RGB Frame Stream

Additional Modalities

Common Tasks (~so far)

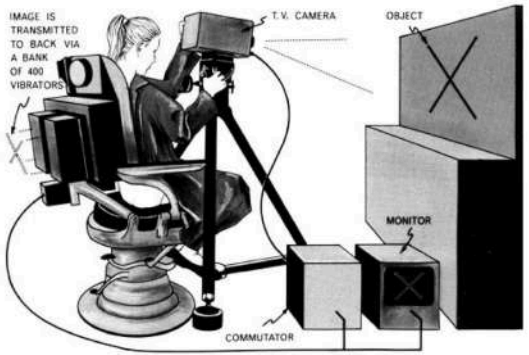
- (1) visual navigation
- (2) rearrangement
- (3) embodied vision-and-language





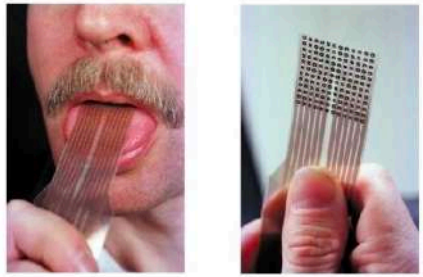
Retrospectives on the Embodied AI Workshop, Deitke et al., 2022

Sensorimotor Contingency



Bach-y-Rita et al., Vision substitution by tactile image projection, Nature (1969)

Tongue Display Unit



Sampaio, E., S. Maris , and P. Bach-y-Rita. 2001
Brain plasticity: 'Visual' acuity of blind persons via the tongue. *Brain Research* 908(July 13):204.

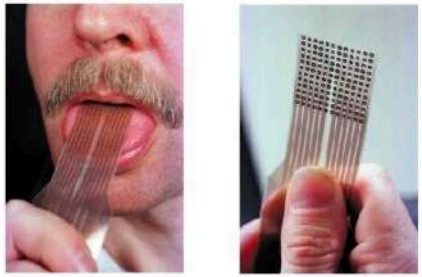
David Ha 2022. Erismann & Kohler 1931. Stratton 1897.
Paul Bach-y-rita (1934-2006) (the father of sensory substitution.

Summary

- The architectural connection between vision and action may be denser than what we think.
- What is “correct”?
  - What appears to matter is the close connection with downstream utility/action ⇒ things get

    (constantly) calibrated vs. being hard-wired to be the “correct” way
  - **Engineering implication**: close the connection with downstream utility of vision (and learning continually) vs. hard-engineering a known configuration
    - **Inductive bis:** still, some structure and selective relearning is still in play.

Tongue Display Unit



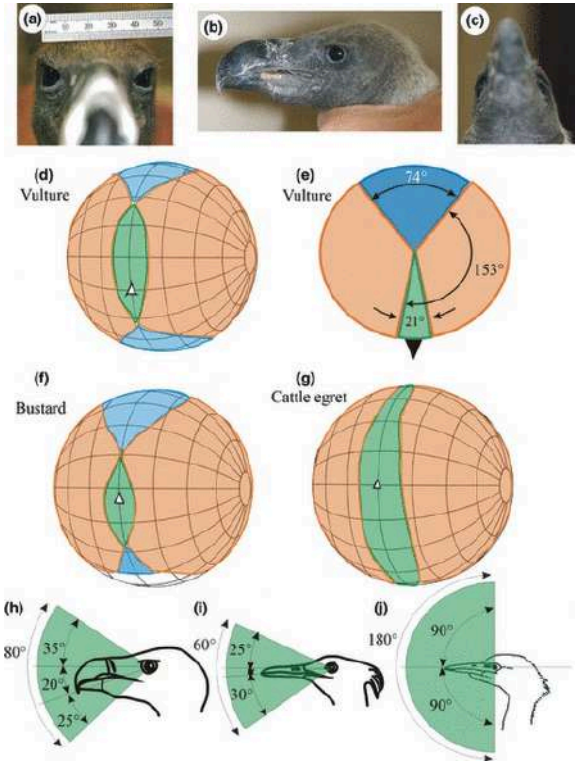
Visual Morphologies

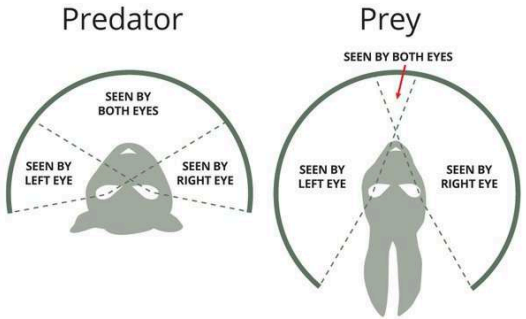
- **Visual perception** is closely linked to agent's actions, body, and surrounding environment.

- M. Bank et al., 2015. M. Land, 2002
- Refs in evolutionary & computational biology, psychology, neuroscience. J. Gibson 1979, P. Churchland et al.,1994, A. Parker 2003.

Clams

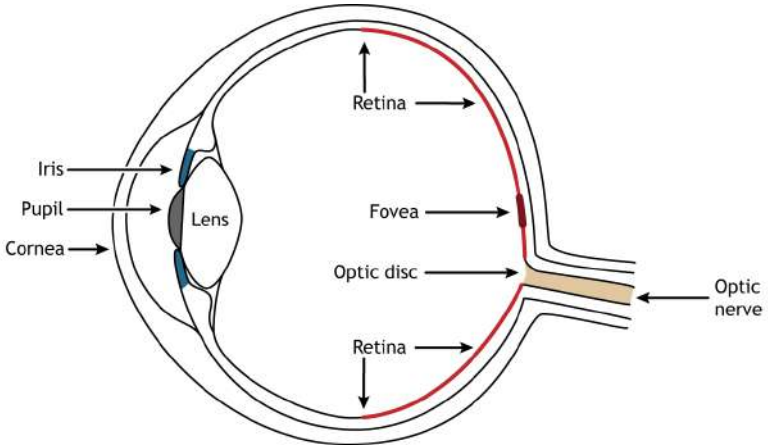
Filed of view



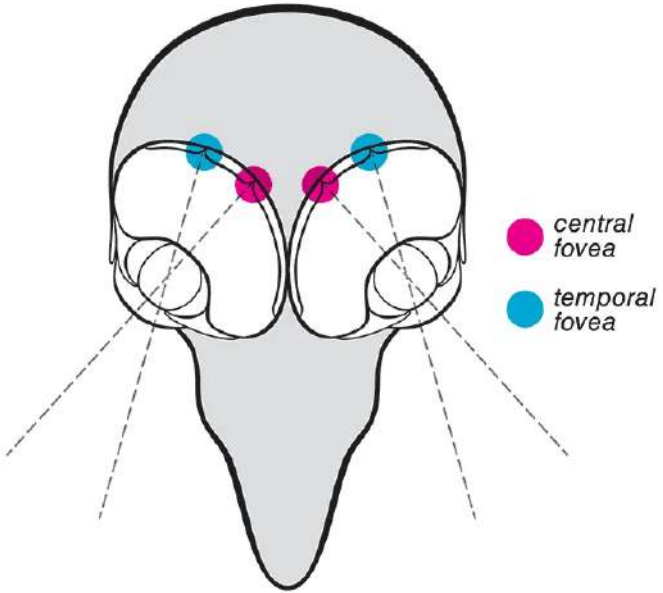


Acuity zones

Acuity zones/Fovea

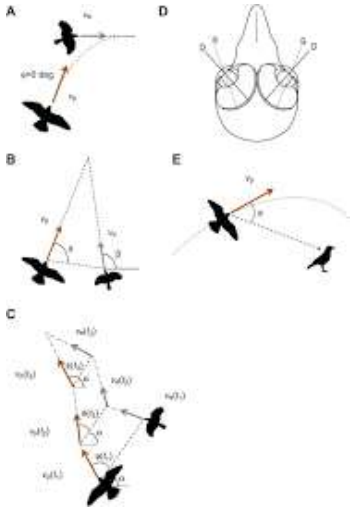


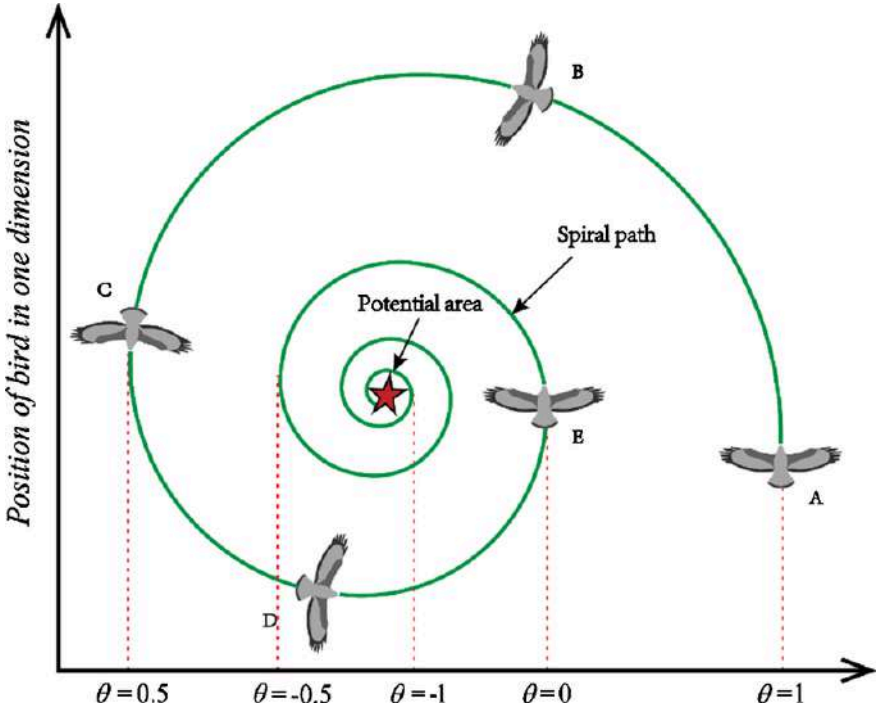
Human



Hawks, falcons, etc.

Acuity zones





Darkness Adaptation

- Obvious trick:

Darkness Adaptation

- Obvious trick: going larger

Darkness Adaptation

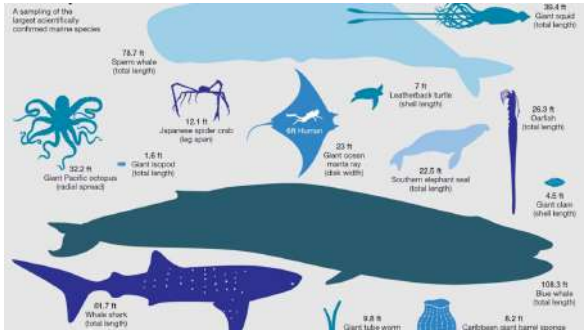
- Integrating receptors: Sweat Bee
- Recycling photons: reflective eyes

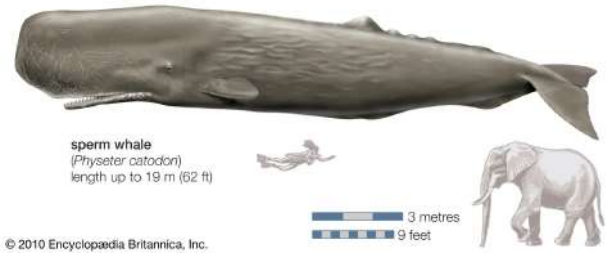
Darkness Adaptation

- Giant Squid

Darkness Adaptation

- Giant Squid



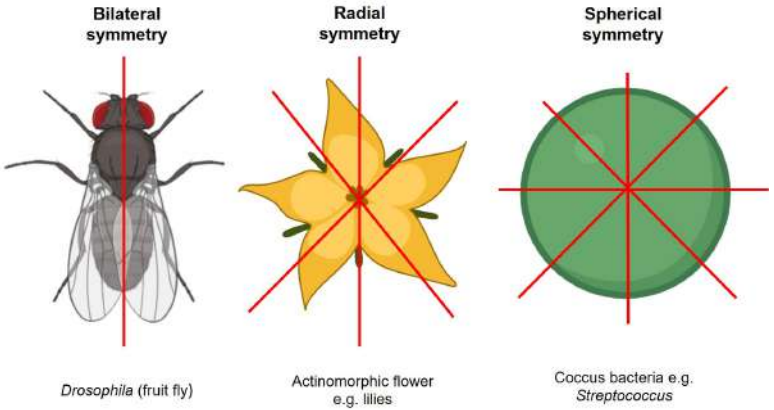


Anthropomorphism

Anthropomorphism

Some weird ones

- Symmetry



Some weird ones

- Brittle Star

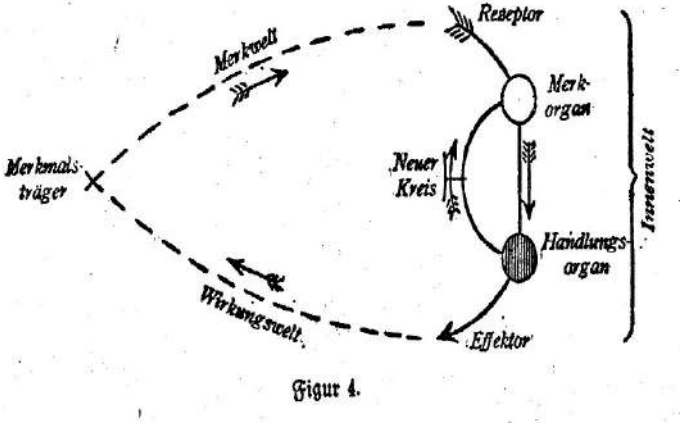
Ecological Vision

- Gibson: “Ask not what's inside your head, but what your head's inside” (Mace, 1977)

James J Gibson

Umwelt

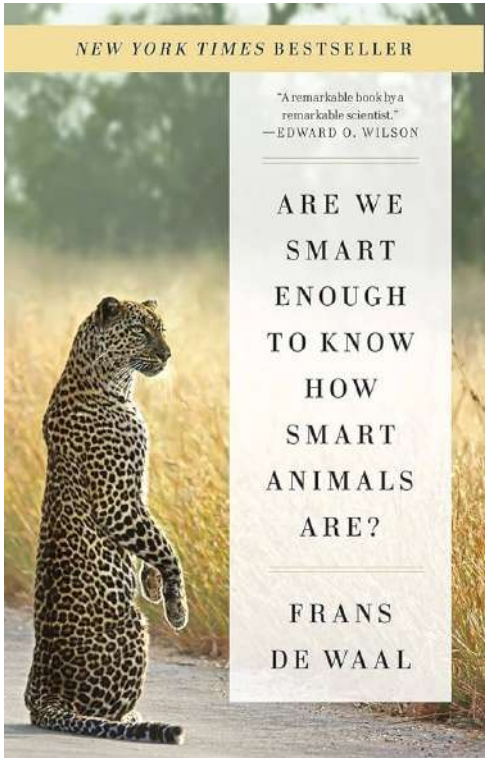
- The Sensory bubble of an animal
- Jakob Johann von Uexküll (1864-1944)

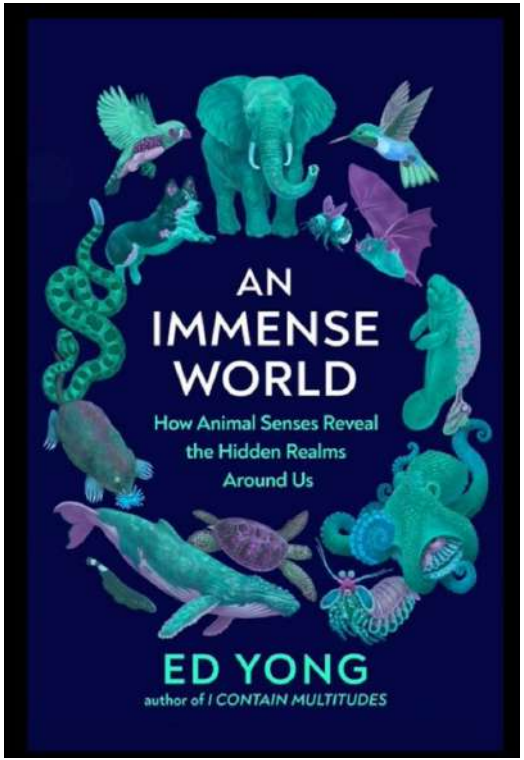


Figur 4.

On physically-grounded & perceptual intelligence

Frans de Waal (1948-2024)





Designing the Visual Morphology

A simple eye.

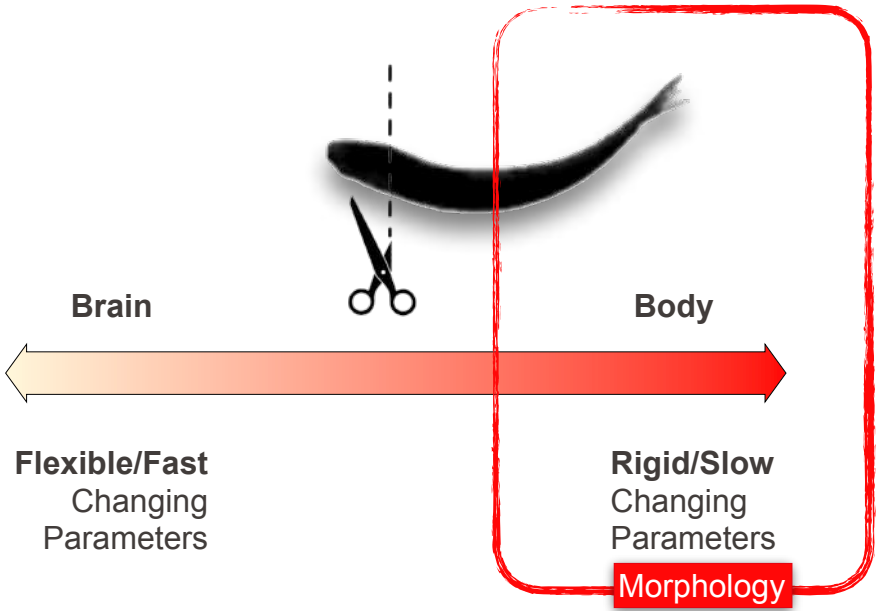
Fish Swimming

Beal 2006

Zamir

Dead Fish Swimming

- Is the fish intelligent?
- Where is the intelligence?



Specialization

Mechanical Morphology

Perceptual Morphology

Specialization

Mechanical Morphology

Perceptual Morphology

- Adaptions to the special ecological context
- Ecological Theory
- No Free Lunch theorem: if you want to succeed everywhere (i.e., not specialize), you won't be efficient anywhere.

Automatic Design of Morphologies

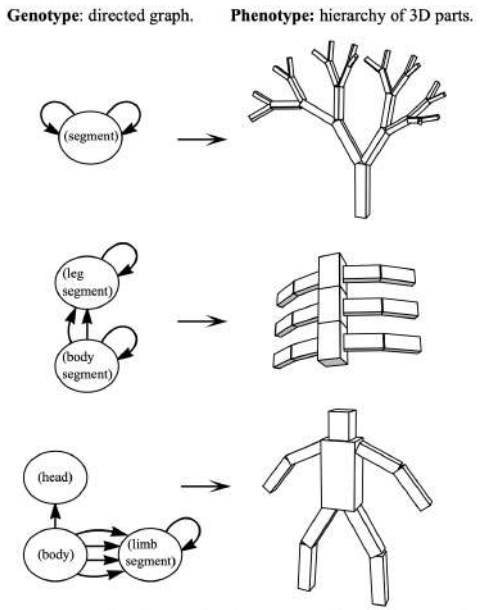


Figure 1: Designed examples of genotype graphs and corresponding creature morphologies.

- Evolving Virtual Creatures, Sims 1994

Automatic Design of Morphologies

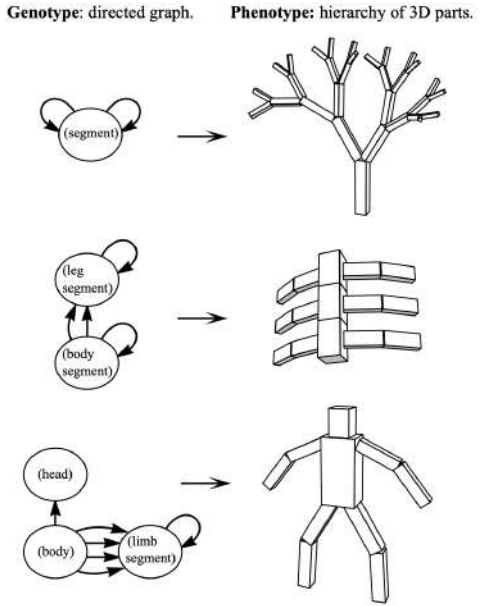
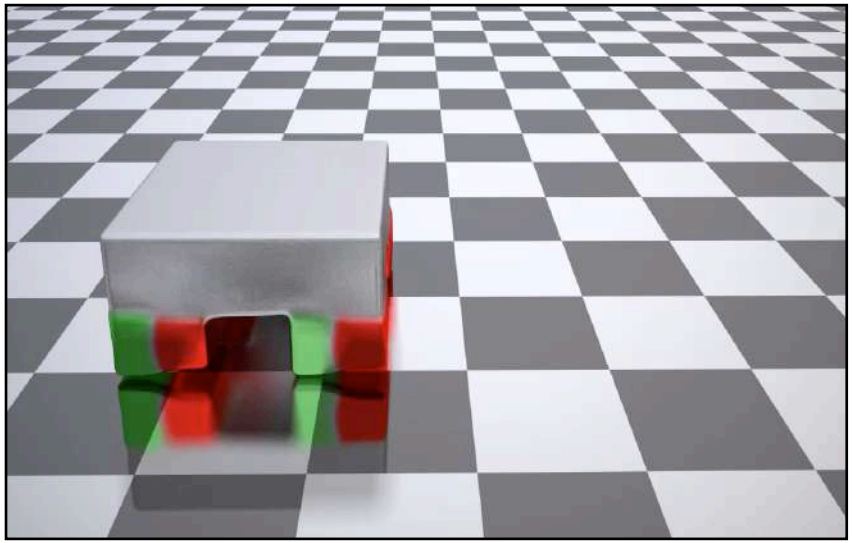


Figure 1: Designed examples of genotype graphs and corresponding creature morphologies.

- Evolving Virtual Creatures, Sims 1994

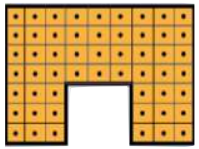


- **2024**: Advances in computing, optimization, simulators, mathematical models, learning.

Methodology of Automated Design

Design Parameter Space

(e.g. voxel location & elasticity)

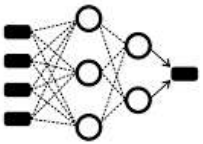


Model

Physics-based

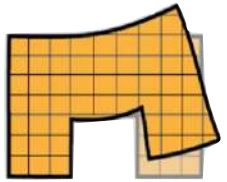
$$\Delta t \approx \Delta x \sqrt{\frac{\rho}{E}}$$

Learning-based



Downstream Goal

(e.g. run faster)



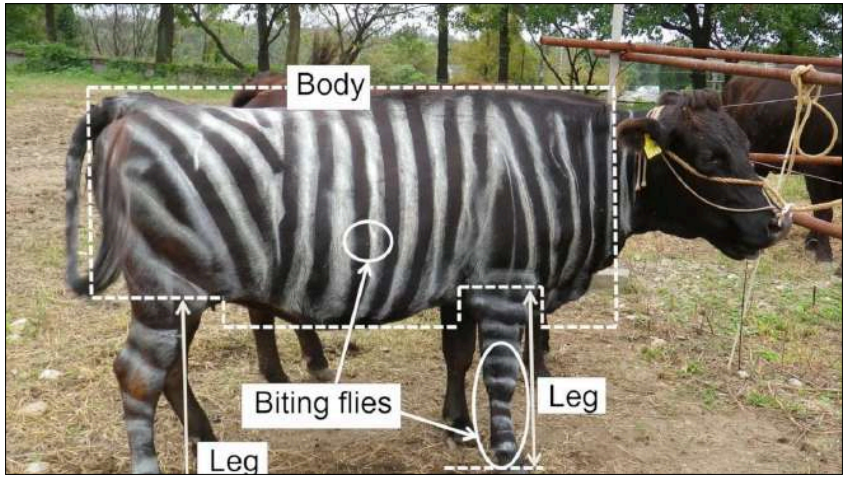
- **Related problems**: X Discovery (material, drug, etc.), AlphaGo, computational imaging design, etc.

When/Why not manually design?

- Unintuitive domains ⇒ Poor world model
  ⇒ automated design.

Pillcam

Zamir



Sasha Sax. Kojima et al,. Ed Yong 2022.

Designing Morphologies

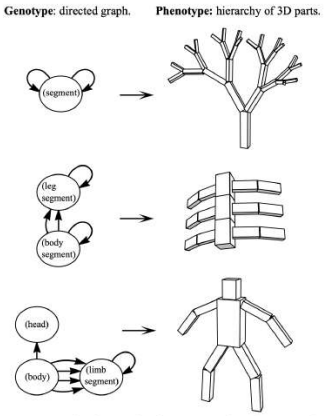
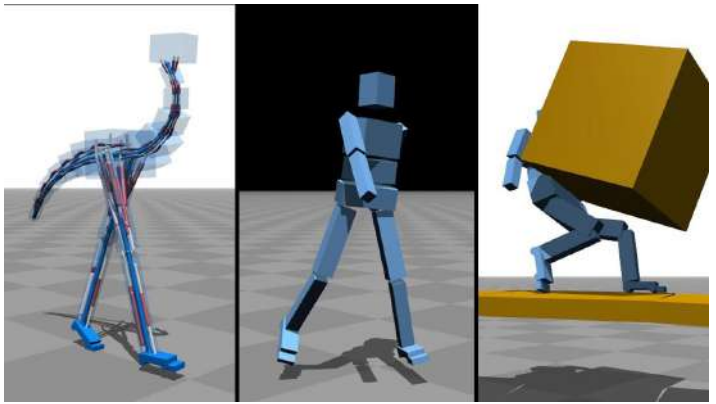


Figure 1: Designed examples of genotype graphs and corresponding creature morphologies.



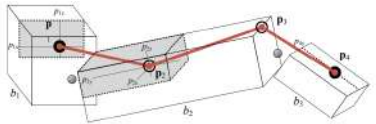
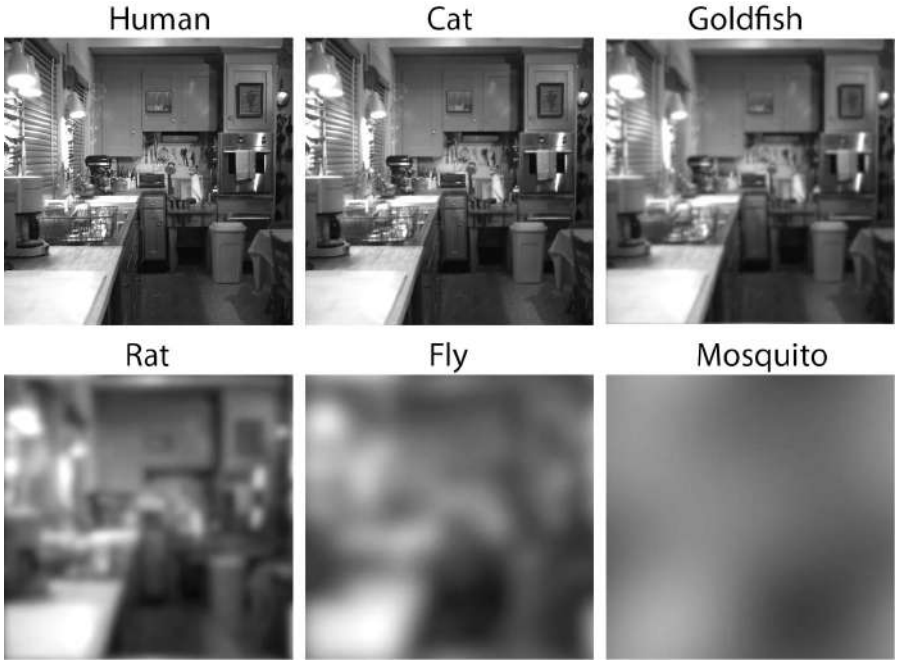


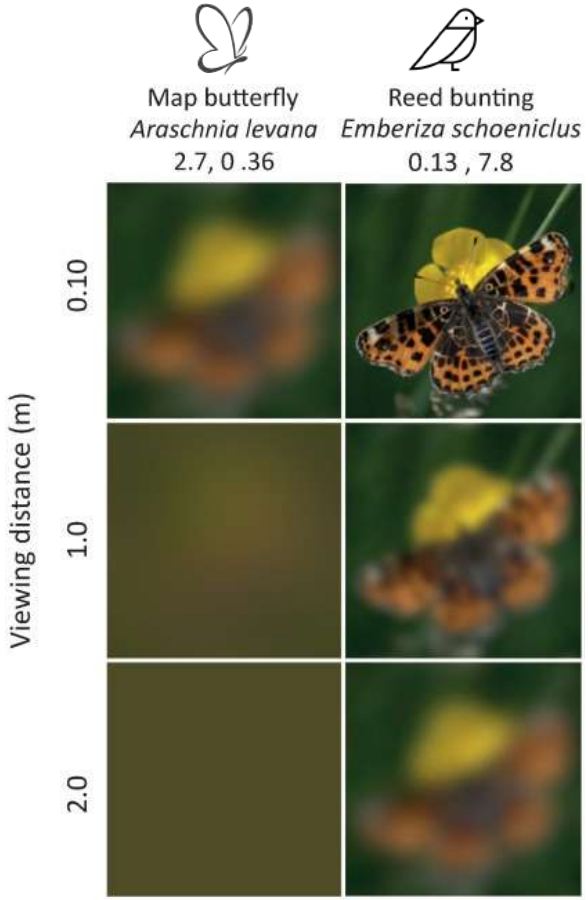
Figure 3: *Muscle attachment points that will be optimized within a constrained region. In this example, muscle point* $\mathbf{p}_1$ *is constrained to a 2D surface, muscle point* $\mathbf{p}_2$ *is constrained to a 3D volume,* $\mathbf{p}_3$ *is fixed, and* $\mathbf{p}_4$ *is constrained to a line. The actual areas used in our experiments are shown in Figure 8.*

Evolving Virtual Creatures, Sims 1994

Flexible Muscle-Based Locomotion for Bipedal Creatures, Geijtenbeek 2013

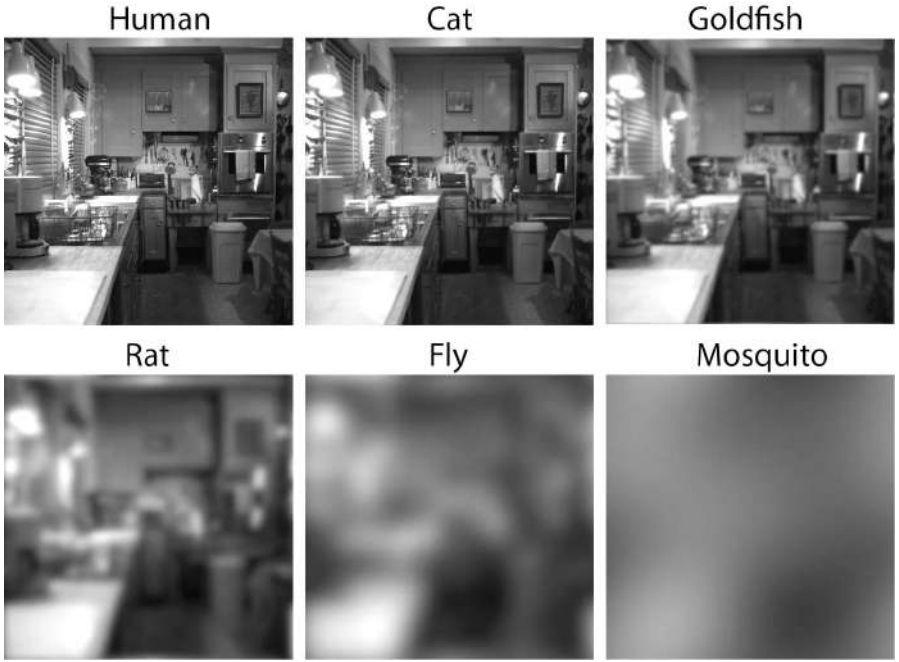
How far can simple eyes go?

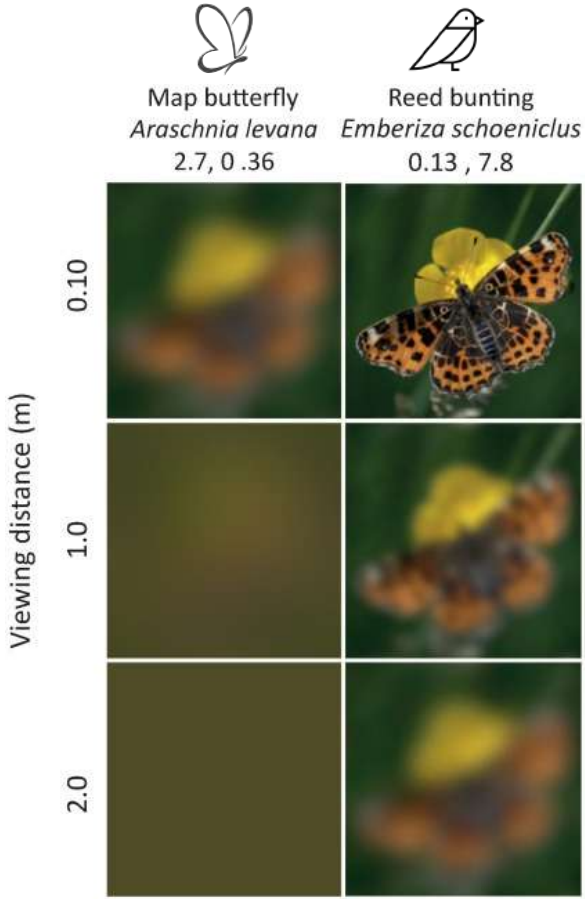




Visual Acuity and the Evolution of Signals, Caves et al. 2018.
An Immense World: How Animal Senses Reveal the Hidden Realms Around Us, Ed Yong. 2022.

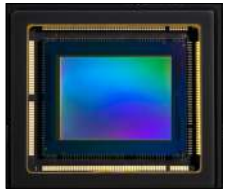
How far can simple eyes go?



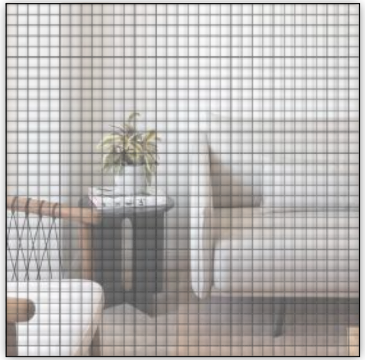


Visual Acuity and the Evolution of Signals, Caves et al. 2018.
An Immense World: How Animal Senses Reveal the Hidden Realms Around Us, Ed Yong. 2022.

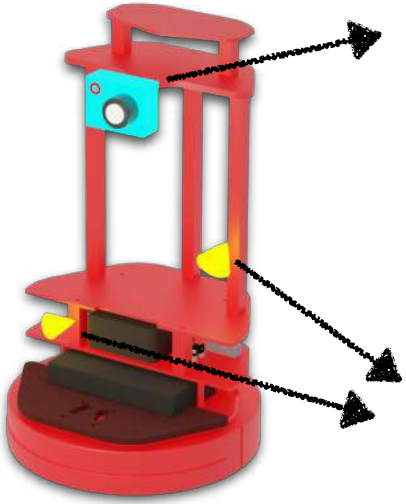
How far can simple eyes go?



Camera *(128x128 = 16,384)*



Photoreceptor *(1x1)*



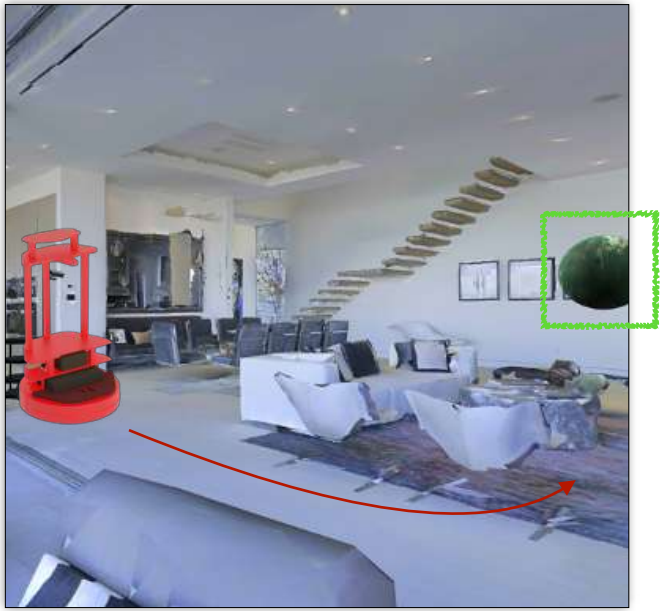
Camera

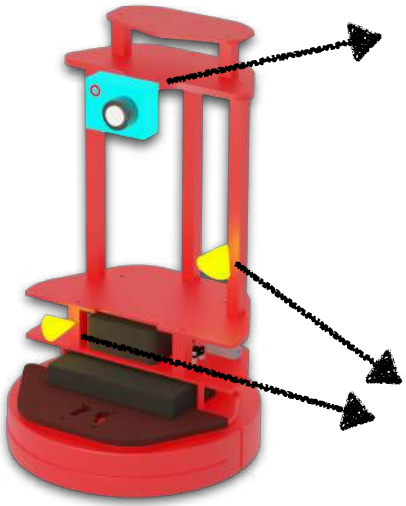
vs.

Photoreceptors

Solving Vision Tasks with Simple Photoreceptors Instead of Cameras. Atanov, Singh, Fu, Yu, Spielberg, Zamir. 2024.

How far can simple eyes go?





Camera

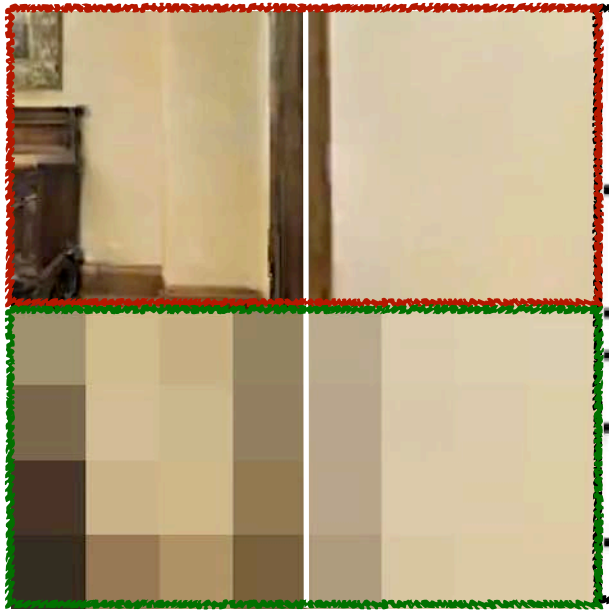
vs.

Photoreceptors

Solving Vision Tasks with Simple Photoreceptors Instead of Cameras. Atanov, Singh, Fu, Yu, Spielberg, Zamir. 2024.

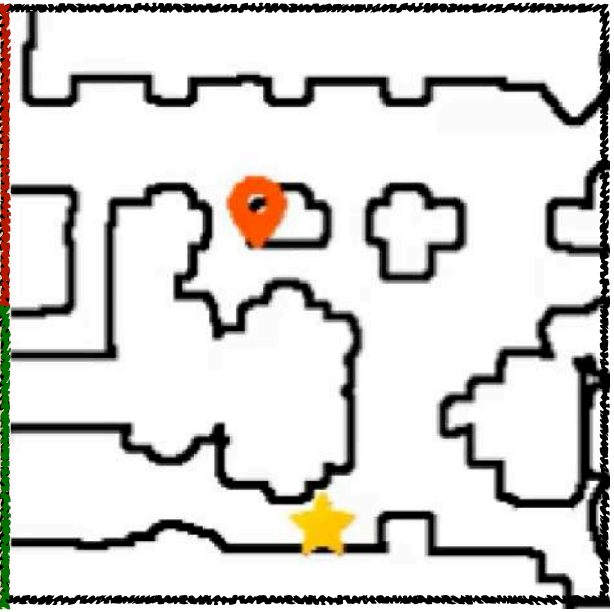
How far can simple eyes go?

Camera
(for visualization only)



Photoreceptors
(agent's Observation)

2 x (4x4)

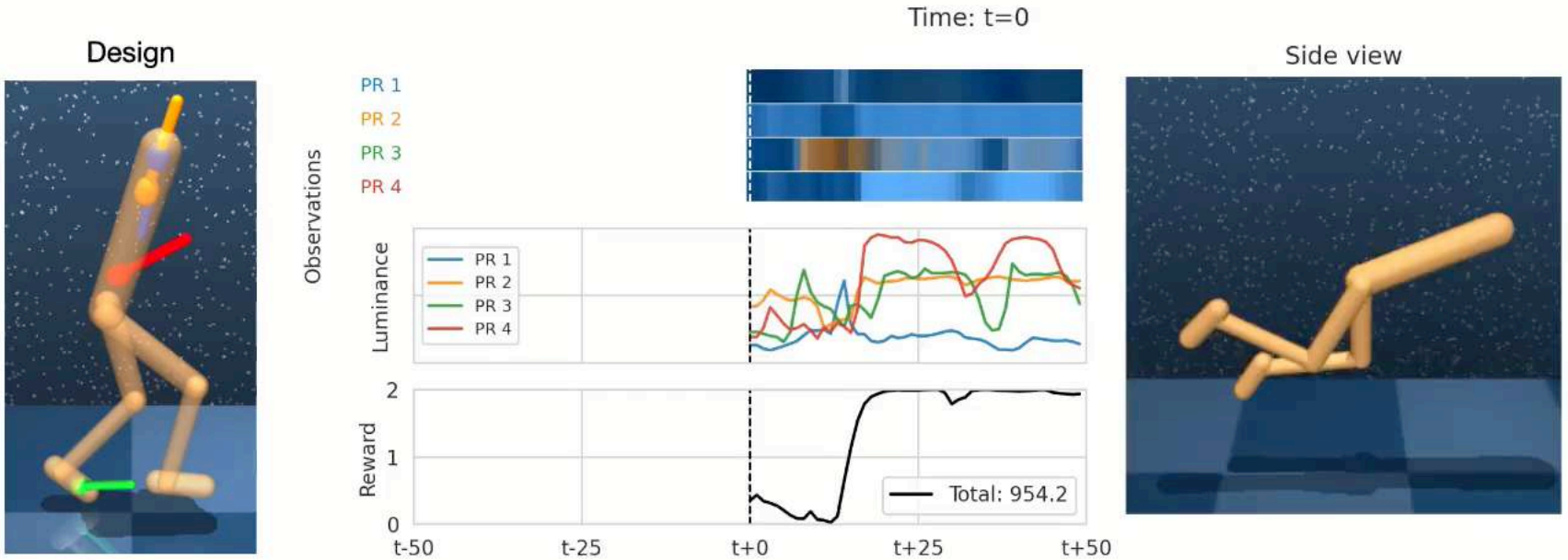


Start

Target

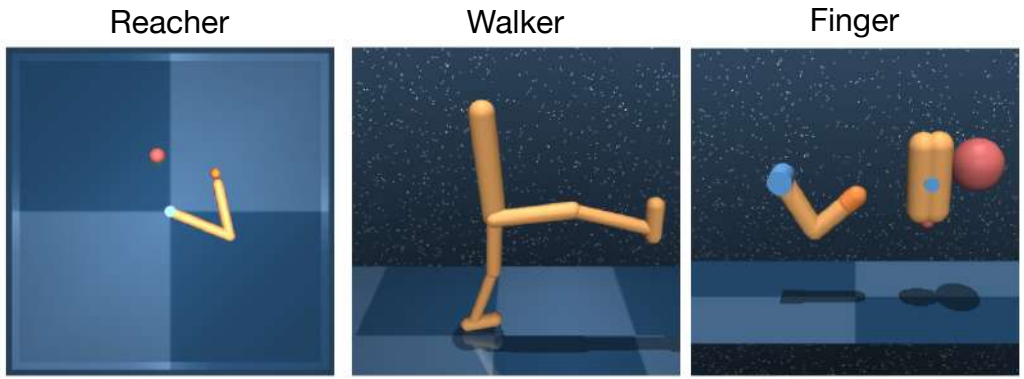
Map
(for visualization only)

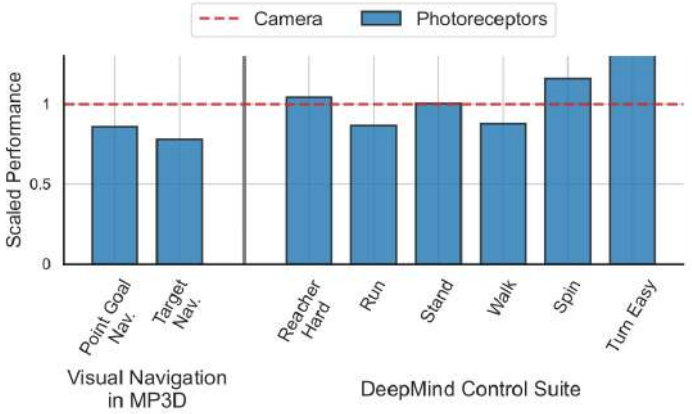
How far can simple eyes go?



Solving Vision Tasks with Simple Photoreceptors Instead of Cameras. Atanov, Singh, Fu, Yu, Spielberg, Zamir. 2024.

How far can simple eyes go?

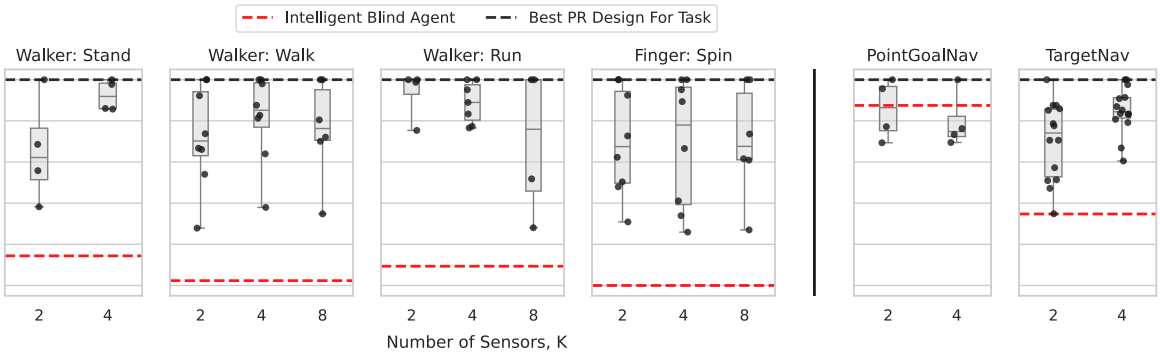


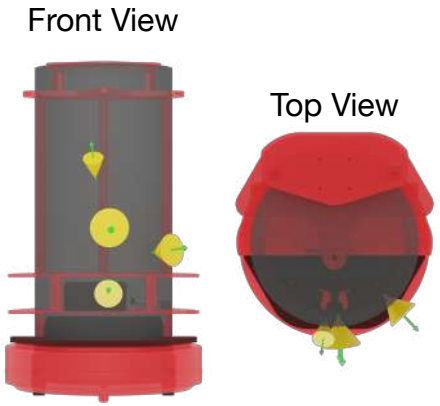


i. **Simple photoreceptors can be competitive** (w/ cameras) for visually-guided behavior for many tasks.

Solving Vision Tasks with Simple Photoreceptors Instead of Cameras. Atanov, Singh, Fu, Yu, Spielberg, Zamir. 2024.

How far can simple eyes go?



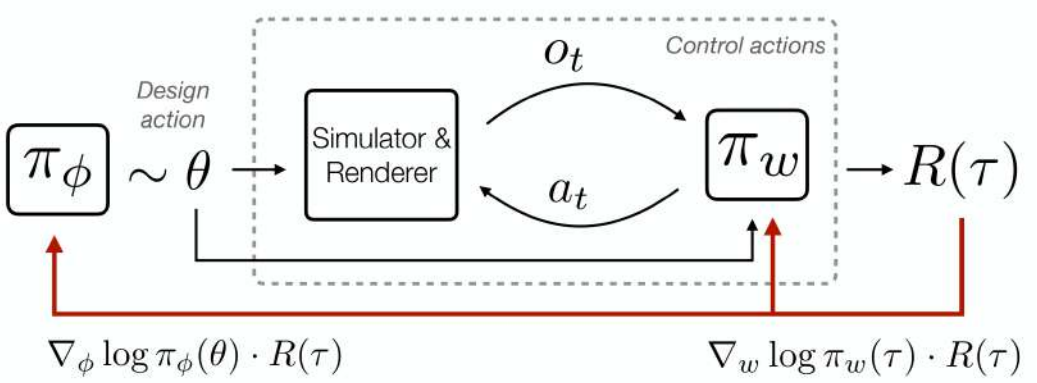


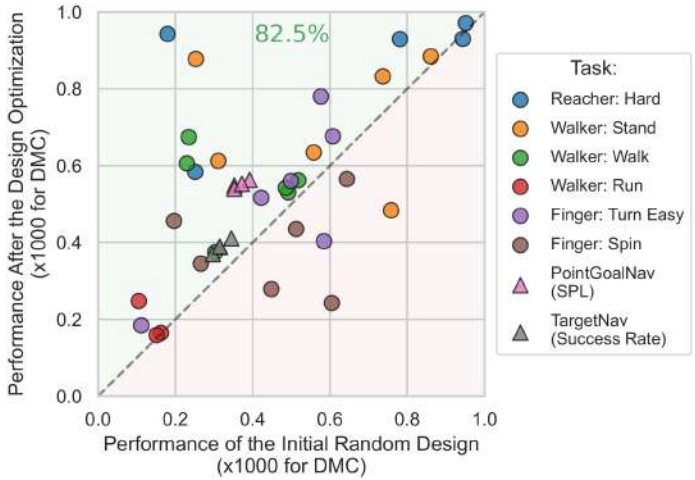
i. **Simple photoreceptors can be competitive** (w/ cameras) for visually-guided behavior for many tasks.

ii. The strategic placement of photoreceptors matter ⇒ An **optimization-based automatic design** method.

Solving Vision Tasks with Simple Photoreceptors Instead of Cameras. Atanov, Singh, Fu, Yu, Spielberg, Zamir. 2024.

Designing the morphology



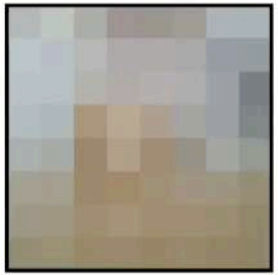


Designing the morphology

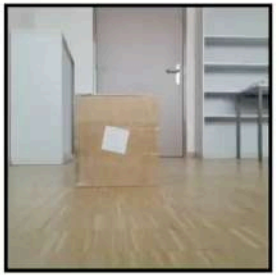
Target Navigation by Photoreceptor equipped robot

External (Third-person) View

Visual Observation:
64 Photoreceptors



Robot Camera View
(for visualization only)



Questions?

Amir Zamir (amir.zamir@epfl.ch)

Oguzhan Kar (head TA)
Kunal Pratap Singh
Rishubh Singh
vimm-ta@groupes.epfl.ch

https://vilab.epfl.ch/