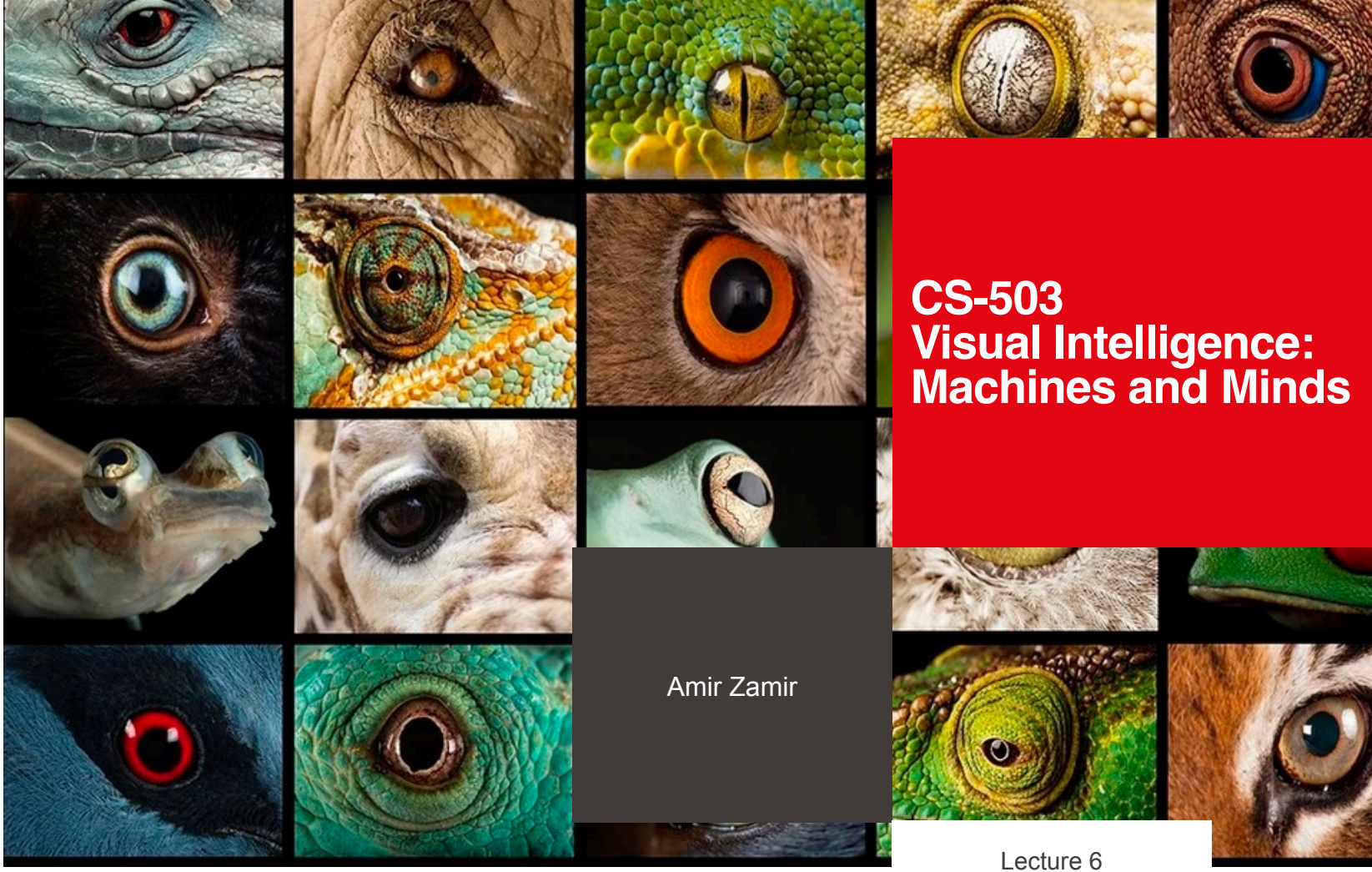


CS-503
Visual Intelligence: Machines and Minds

Amir Zamir

Lecture 6

École polytechnique fédérale de Lausanne

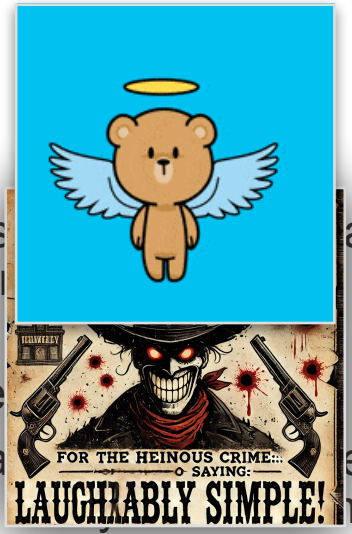
Logistics

Logistics

- First assignment notebook due 11/03/2025 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

Anonymous feedback entry



Hi !

This arly in th ut I just wanted to thank you for
you kindness and your insights.

You incredibl
rare

I wa emester ons but this course
has mpact.

It's laughably simple, but it ha
usefulness of what I do whilst maintaining a critical eye. And most importantly,
that we're here to learn, not to perform, which you easily forget in this
environment.

So from the bottom of my hea t, thank you

🚨 **WANTED!** 🚨

💰 **BOUNTY: ICCV Conference Trip to Hawaii** 🌴🏝️

FOR THE CRIME OF SAYING:
"*Laughably simple*"

This outlaw has made things **sound too easy**, and we can't let that slide. If you have **seen, heard, or been a victim** of this phrase, you may be eligible for the bounty!

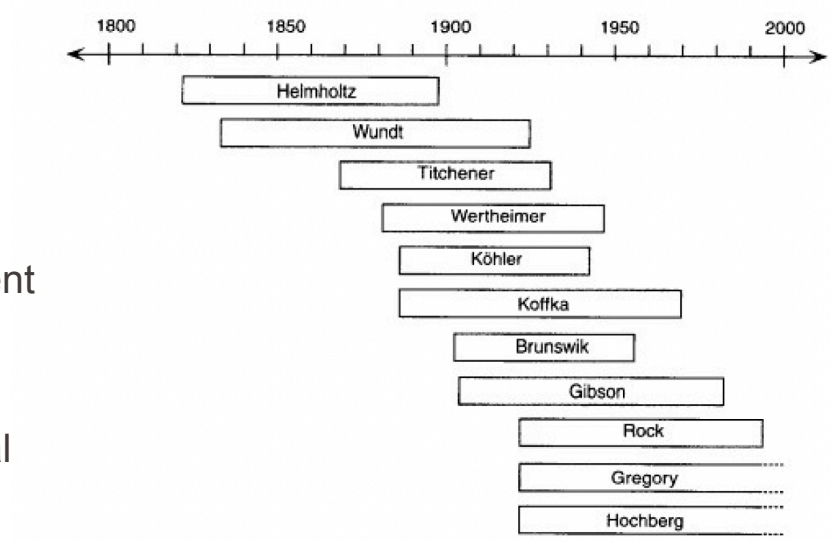
🔍 **LAST KNOWN WHEREABOUTS:**

- Any place where confidence runs high
- Tech forums, puzzle competitions, or trivia nights
- Probably sipping coffee while casually solving hard problems

📣 **REWARD:**
A luxurious **trip to Hawaii** 🌊, where things are **genuinely laughably simple.**

Recap

Vision theories



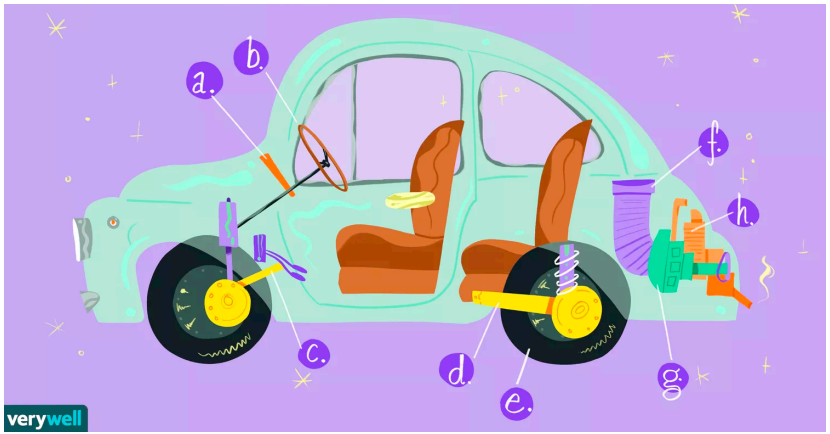
- What's a "theory"? An integrated/consistent set of statements/hypotheses about underlying principles of something.
  - That not only organizes and explains known facts (eg existing experimental results), but also makes predictions about new ones.

THEORY	NATIVISM vs. EMPIRICISM	ATOMISM vs. HOLISM	ORGANISM vs. ENVIRONMENT	PRINCIPAL ANALOGY	METHOD
Structuralism	Empiricism	Atomism	Organism	Chemistry	Trained Introspection
Gestaltism	Nativism	Holism	Organism	Physical Field Theory	Naive Introspection
Ecological Optics	Nativism	Holism	Environment	Mechanical Resonance	Stimulus Analysis

Stephen Palmer

Structuralism

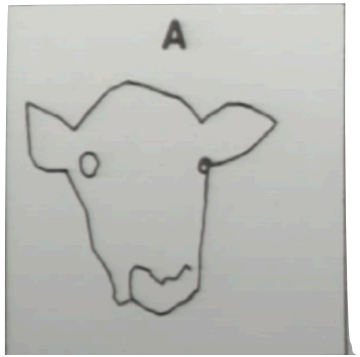
- Progressive “concatenation" of “sensory atoms”

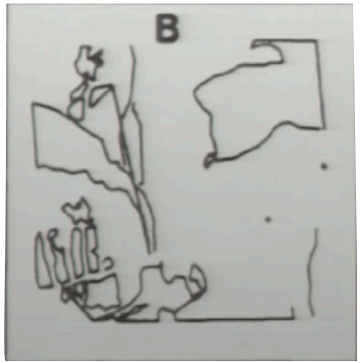


Wilhelm Wundt

Gestaltism

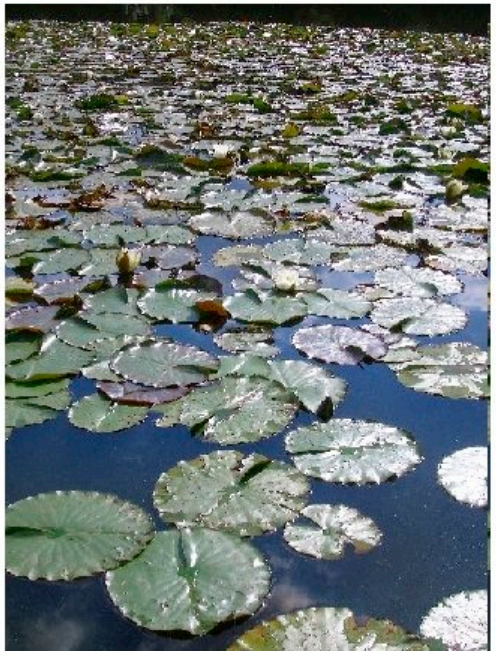
- Whole is more than the sum of parts.

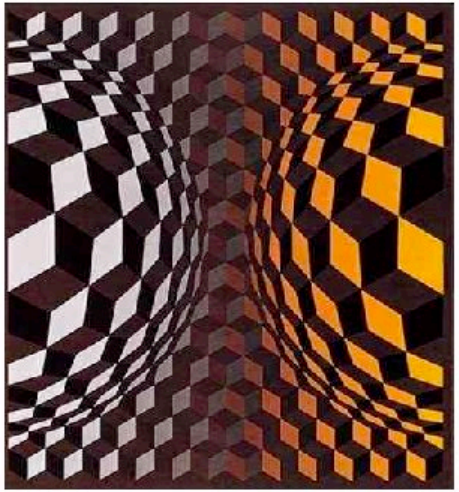




Ecological Approach

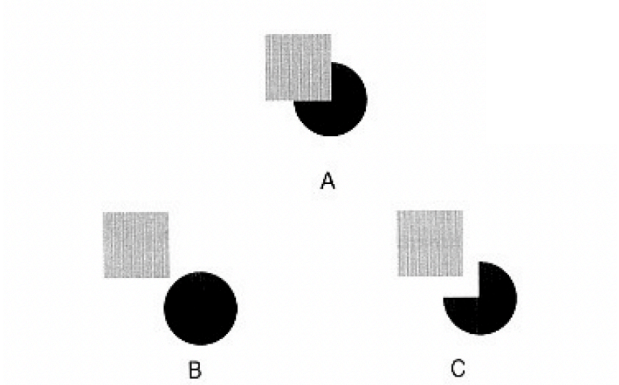
- Gibson: “Ask not what's inside your head, but what your head's inside” (Mace, 1977)

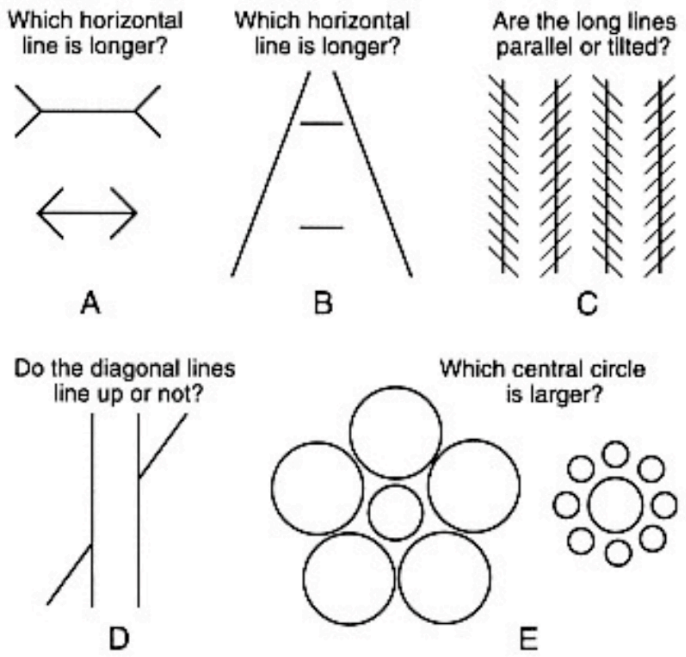




Constructivism

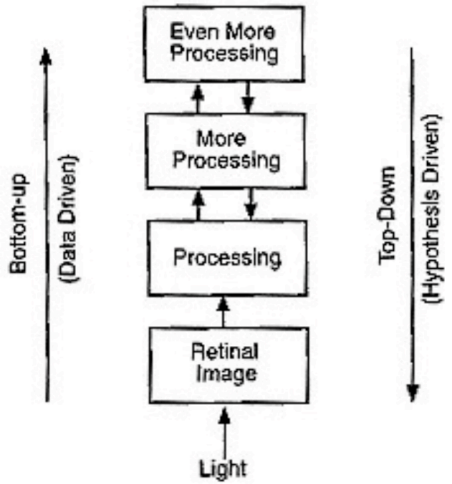
- Vision:
  - an indeterminate inverse problem from retinal images.
  - a “reconstruction” of the reality.
- Something besides the retinal image is needed.
- Likelihood Principle





Feedback

- Vision appears to be more than bottom-up association.

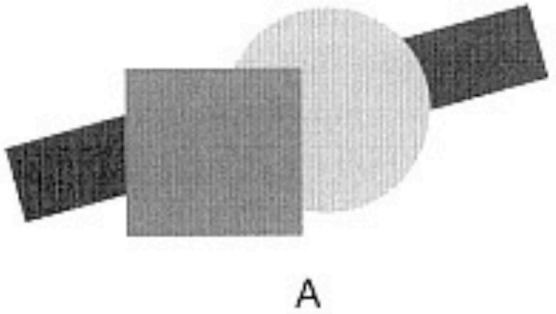


Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

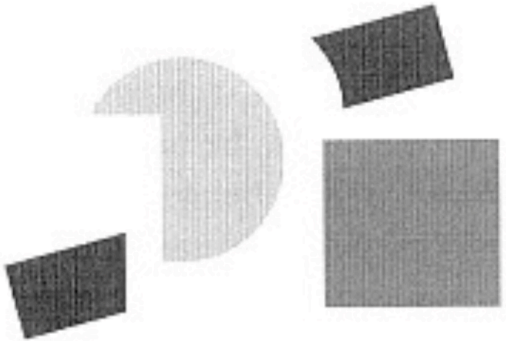
- Vision: A *useful* *reconstruction* of the world in a bottom-up and *top-down* way.

Perception as modeling the environment

- *The observer is constructing a **model** of what environment situation might have produced the observed pattern of sensory stimulation*



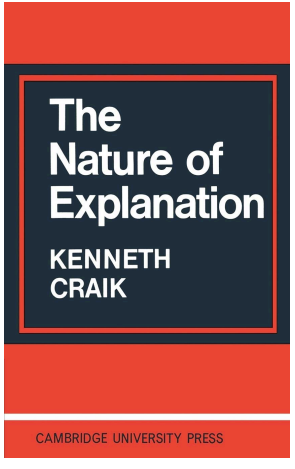
A



B

Mental Models

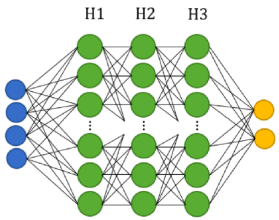
Kenneth Craik (1943)



- *"If the organism carries a "small-scale model" of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilise the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it."*
- Model:
  - something that parallels a reality.
  - enables prediction, planning, and counterfactual reasoning/ imagination.
    - (vs reactive)
  - keeps the relevant aspects and simplifies others.

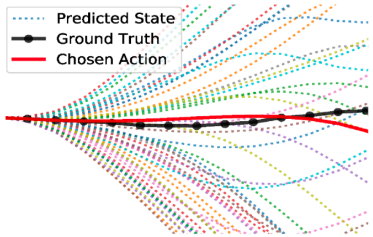
Models

- Model-Based Reinforcement Learning



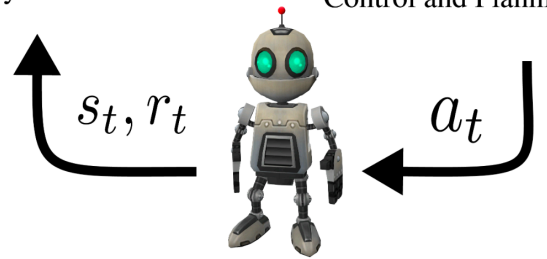
$$s_{t+1} = f_\theta(s_t, a_t)$$

Dynamics Model



$$\arg\max_{u(t)} \sum_{t=0}^{T} r(s_t, a_t)$$

Control and Planning



While improving:

1. Agent acts in environment
2. Learn model of dynamics

$$p_\theta = \arg\max_\theta \sum_{i=1}^{N} \log\, p_\theta(s_{t+1}|s_t, a_t)$$

3. Plan actions to maximize reward

$$a^* = \arg\max_a \sum_{t=0}^{T} \gamma^t r(s_t, a_t)$$

$$s.t.\ s_{t+1} \sim p_\theta(s_{t+1}|s_t, a_t)$$

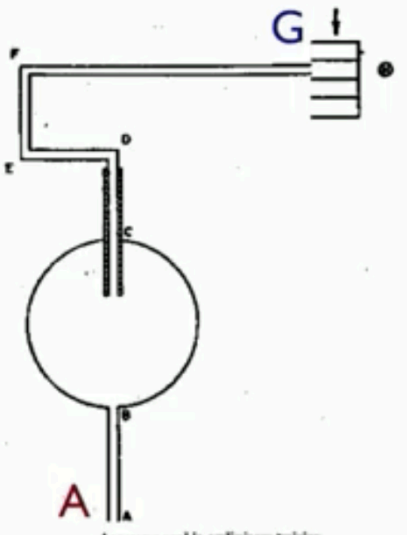
Cognitive Maps (1948)

Zamir

COGNITIVE MAPS IN RATS AND MEN [1]

BY EDWARD C. TOLMAN

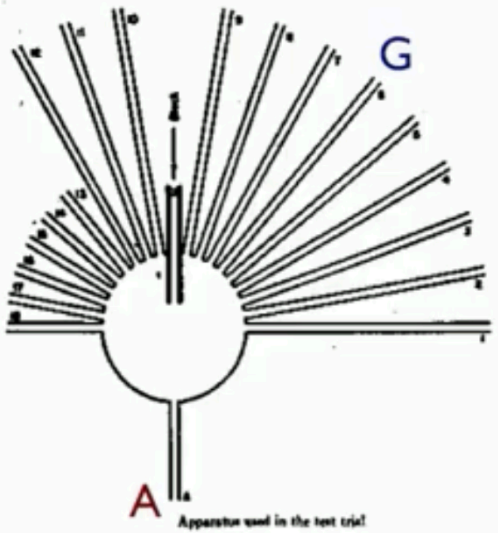
University of California



Apparatus used in preliminary training

FIG. 15

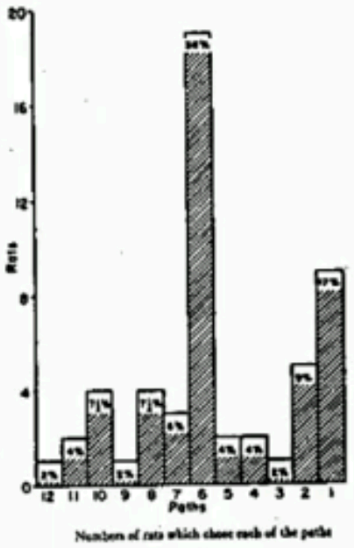
(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 16.)



Apparatus used in the test trial

FIG. 16

(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.)



Numbers of rats which chose each of the paths

FIG. 17

(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 19.)

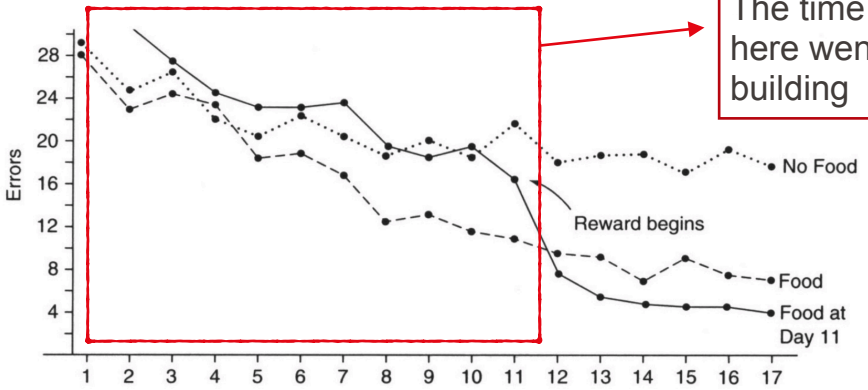
Latent Learning
Tolman & Honzik 1930

Group 1: Rewarded: Day 1 – 17: Every time they got to end, given food (i.e. reinforced).

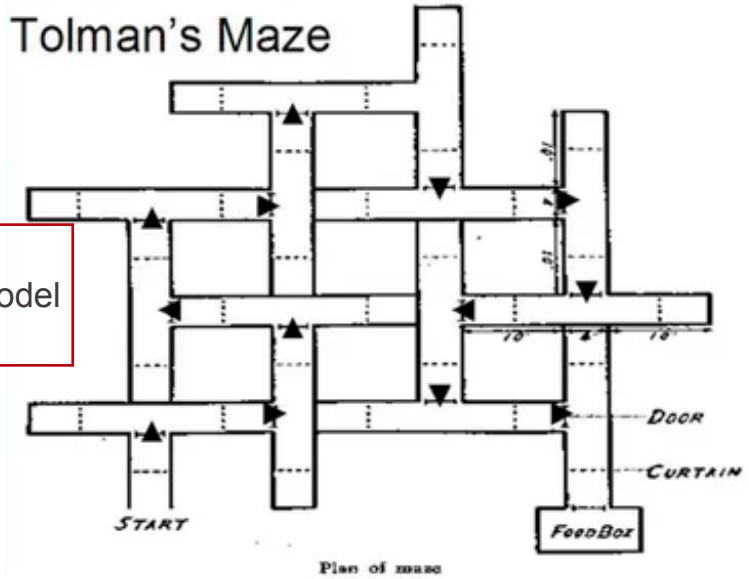
Group 2: Delayed Reward

- Day 1 - 10: Every time they got to end, taken out.
- Day 11 -17: Every time they got to end, given food (i.e. reinforced).

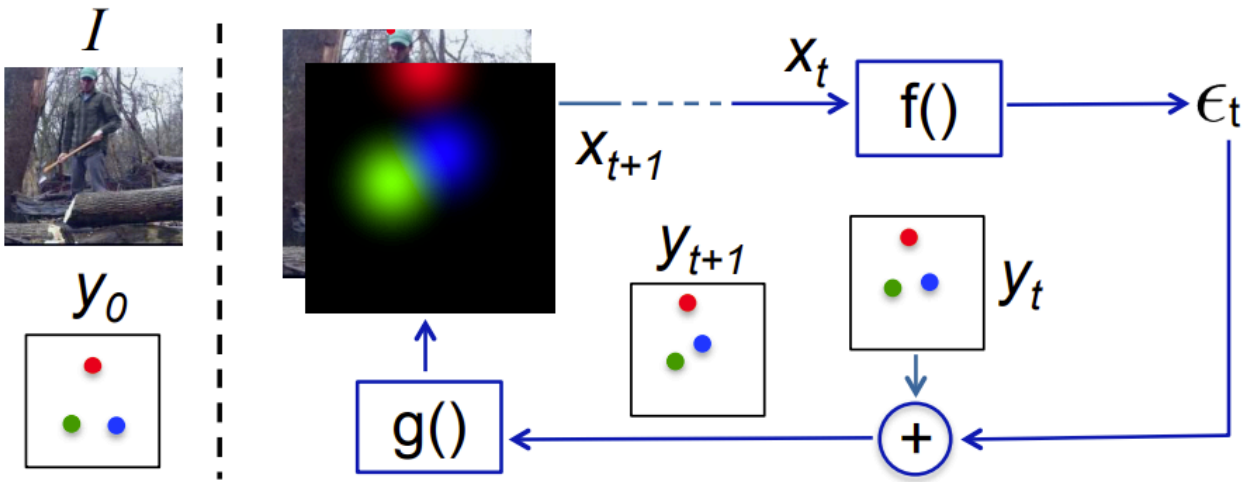
Group 3: No reward: Day 1 – 17: Every time they got to end, taken out.



The time spent here went to model building



Feedback

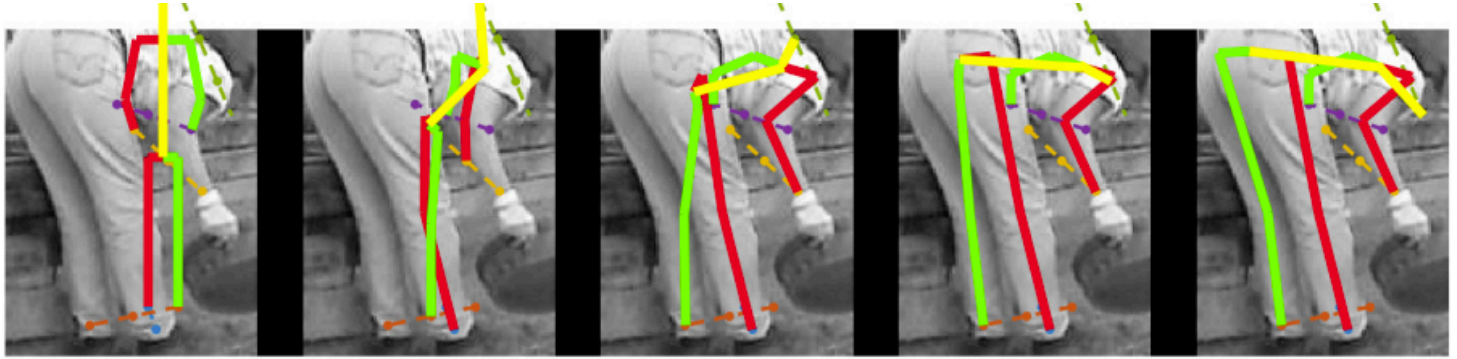


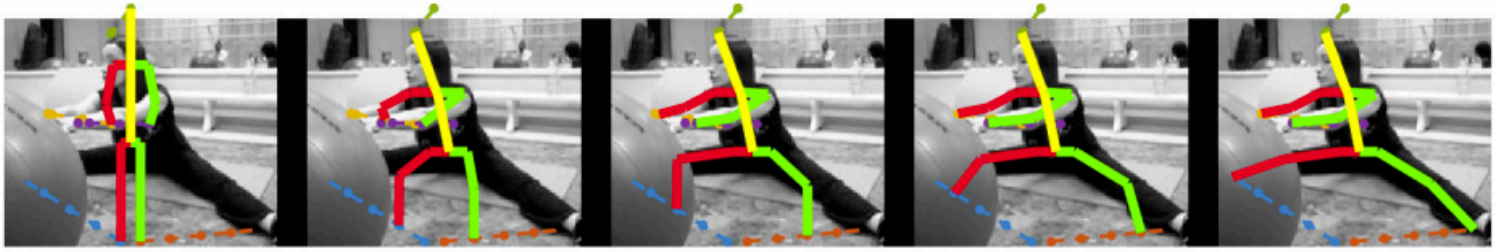
- $\epsilon_t = f(x_t)$
- $y_{t+1} = y_t + \epsilon_t$
- $x_{t+1} = I \oplus g(y_{t+1})$

$$\min_{\Theta_f, \Theta_g} \sum_{t=1}^{T} h(\epsilon_t, e(y, y_t))$$

Human Pose Estimation with Iterative Error Feedback, J Carreira, P Agrawal, K Fragkiadaki, J Malik, CVPR 2016

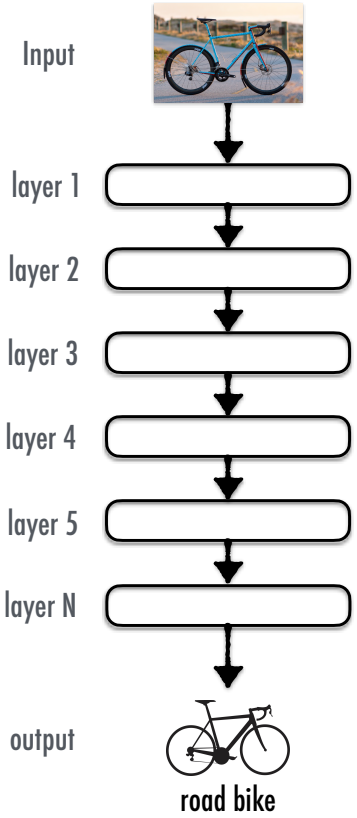
Feedback

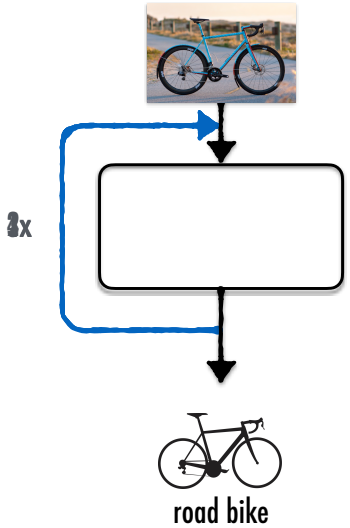




	Head	Shoulder	Elbow	Wrist	Hip	Knee	Ankle	UBody	FBody
Iterative Error Feedback (IEF)	95.2	91.8	80.8	71.5	82.3	73.7	66.4	81.4	81.0
Direct Prediction	92.9	89.4	74.1	61.7	79.3	64.0	53.3	75.1	74.8
Iterative Direct Prediction	91.9	88.5	73.3	59.9	77.5	61.2	51.8	74.0	73.4

Human Pose Estimation with Iterative Error Feedback, J Carreira, P Agrawal, K Fragkiadaki, J Malik, CVPR 2016



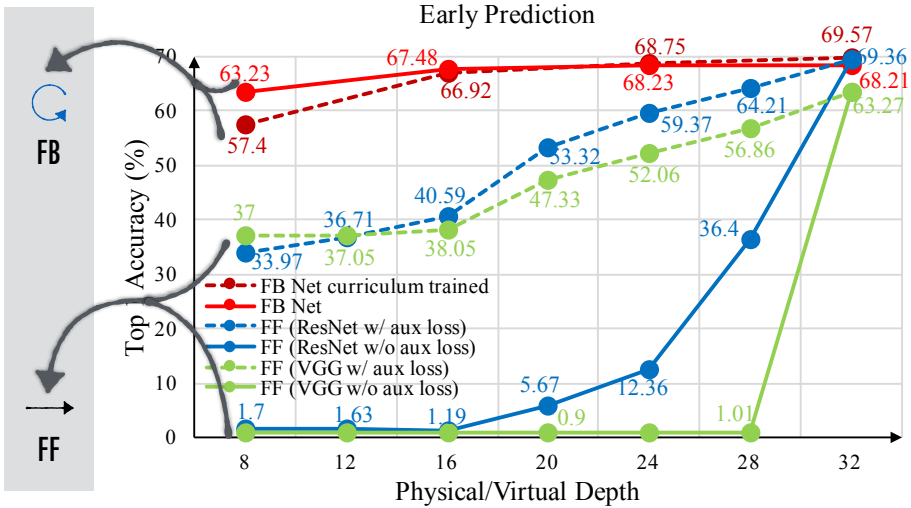


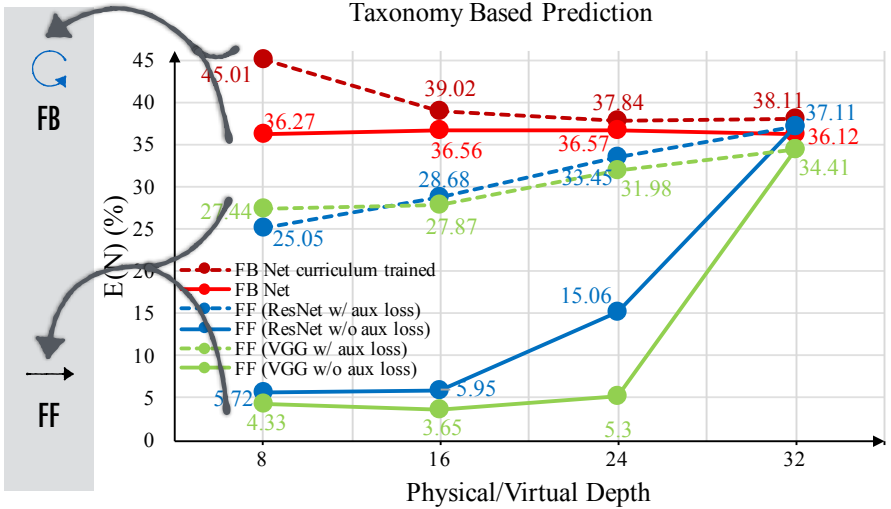
Feedforward model.

Feedback model.

Feedback Networks, CVPR 2017

Early Prediction & Taxonomic Prediction

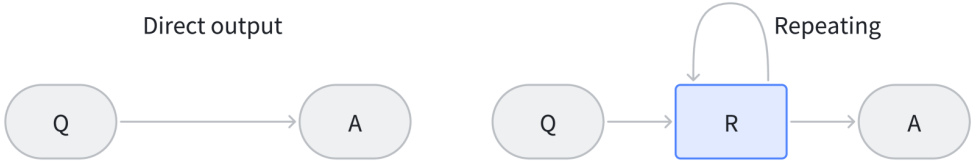


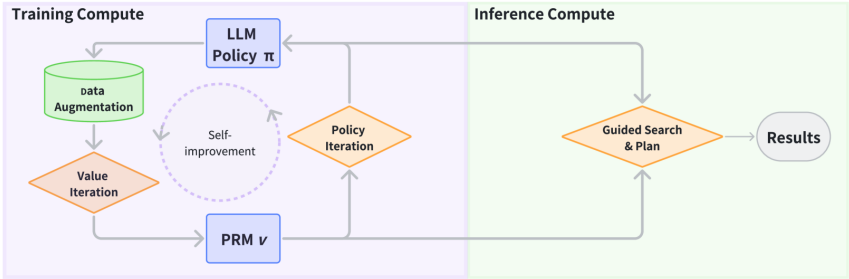


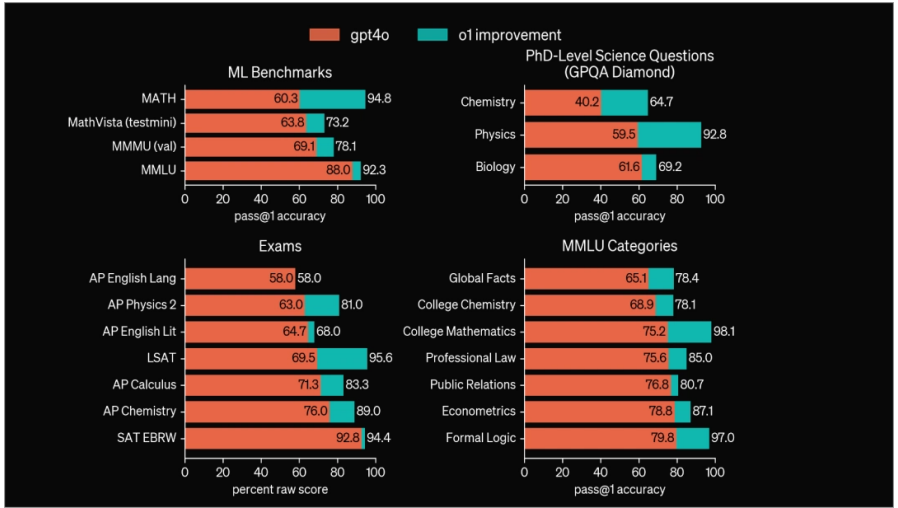
(details in the main paper)

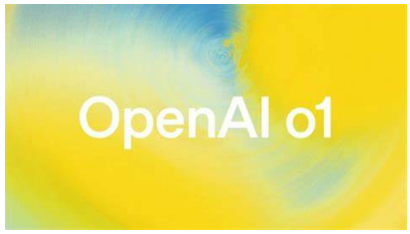
Feedback

- LLM “reasoning”
- TBD in the FM lectures



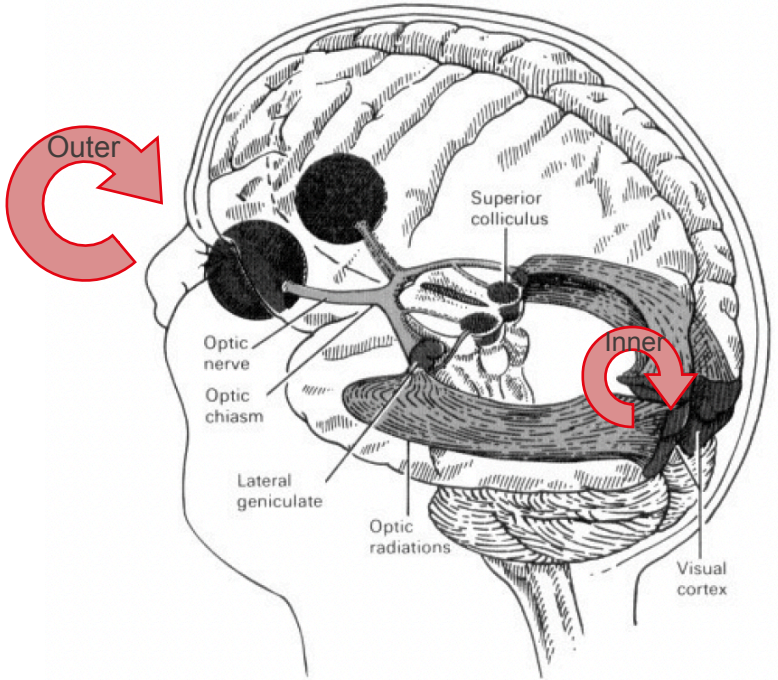




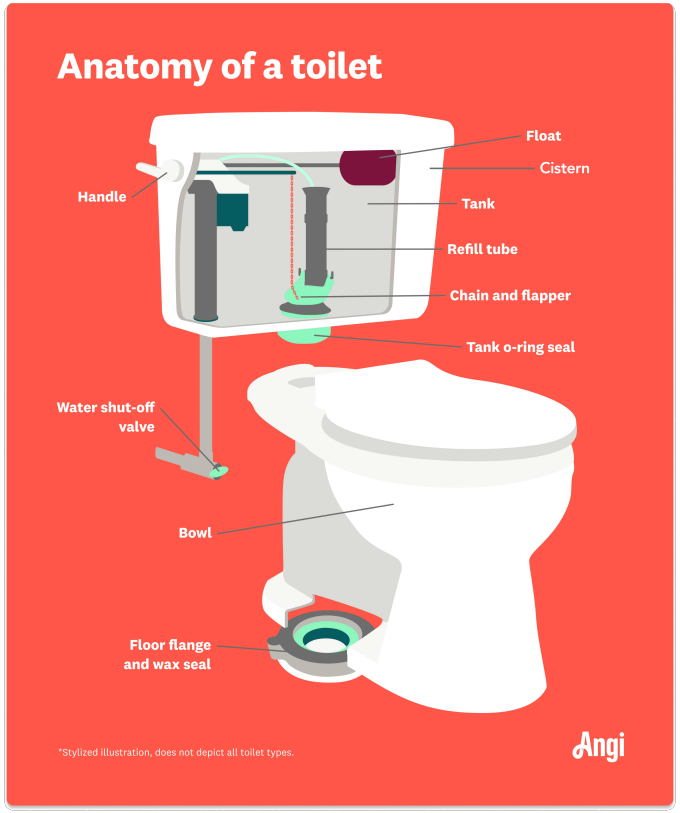


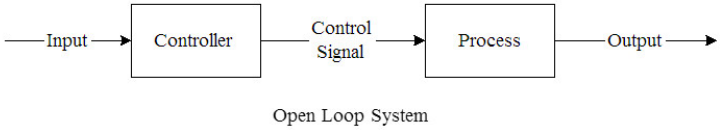
Feedback

- Inner loop
  - top-down processing without external feedback from the world.
  - e.g. IEF (iterative error feedback, 2016), Attention, Feedback Networks (2017), diffusion.

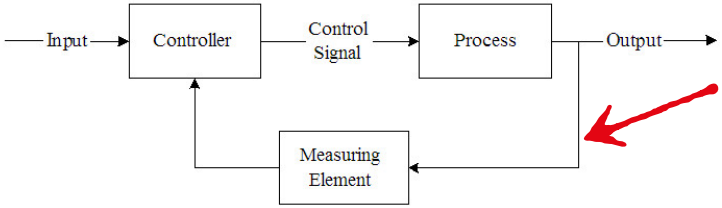


Outer loop Feedback





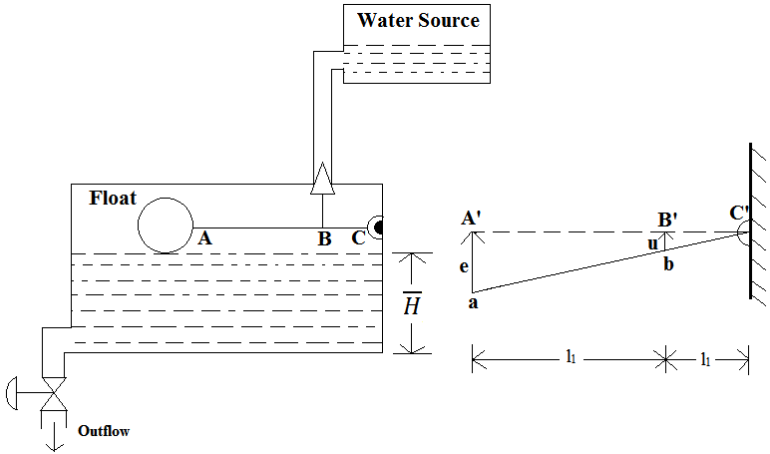
Open Loop System

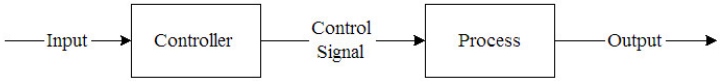


Closed Loop System

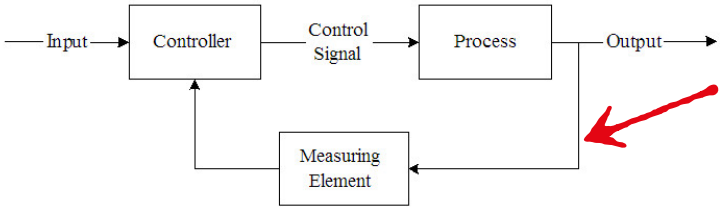
Realpars

Outer loop Feedback





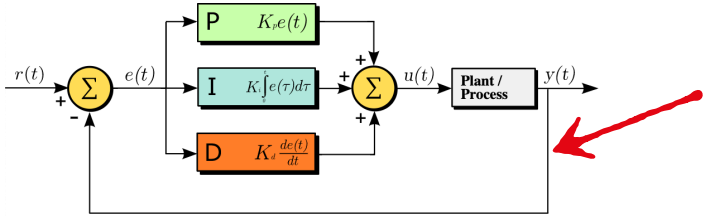
Open Loop System

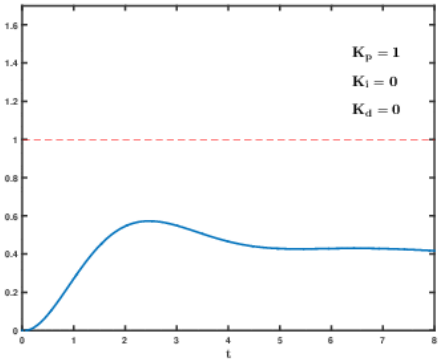


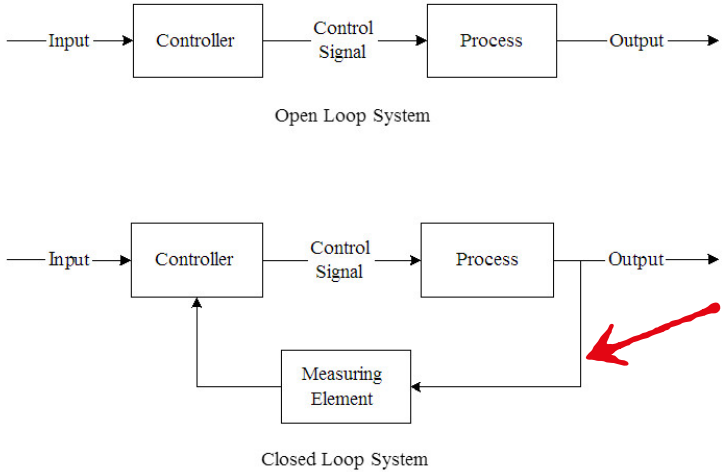
Closed Loop System

Outer loop Feedback

- E.g. PID controller



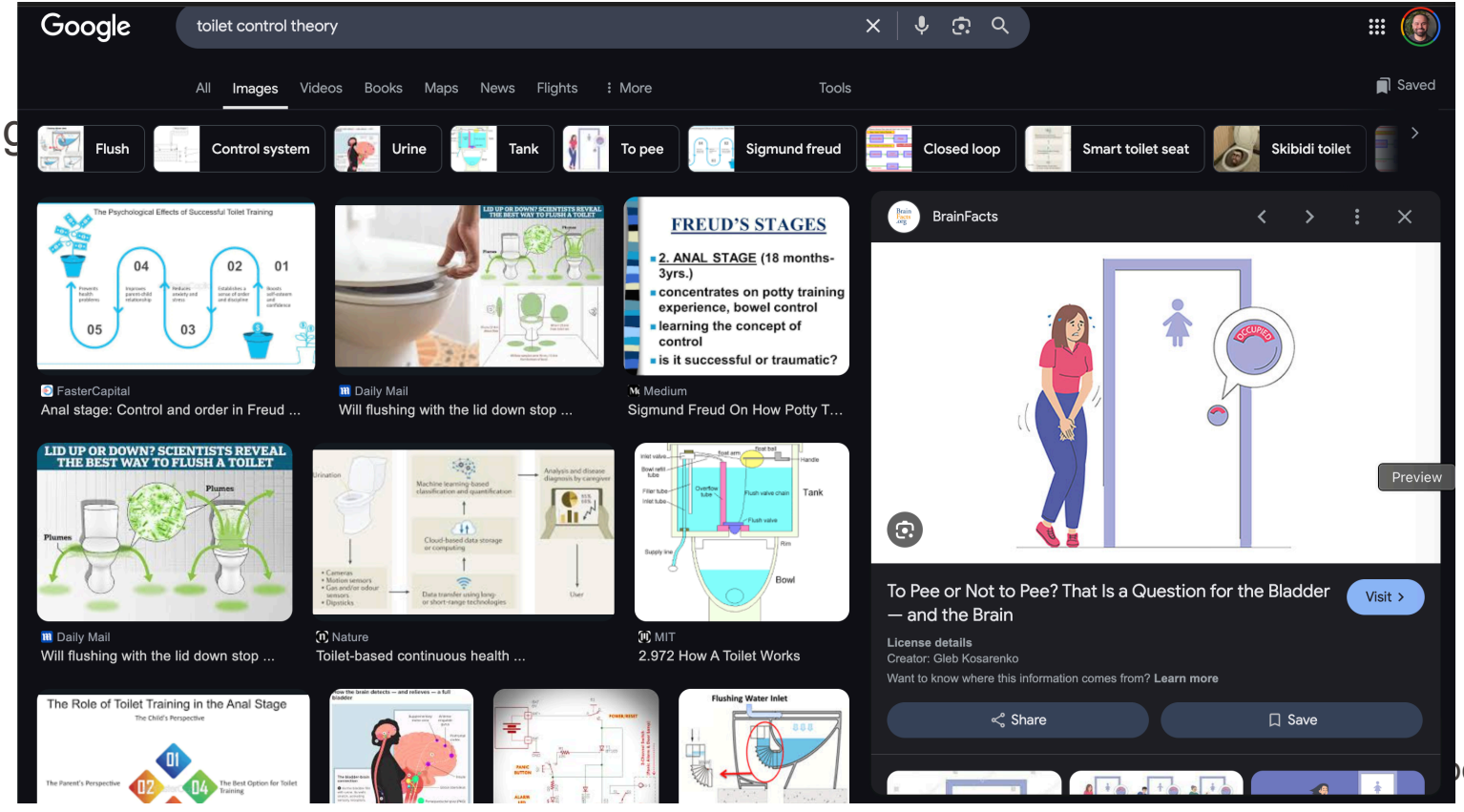




Wikipedia

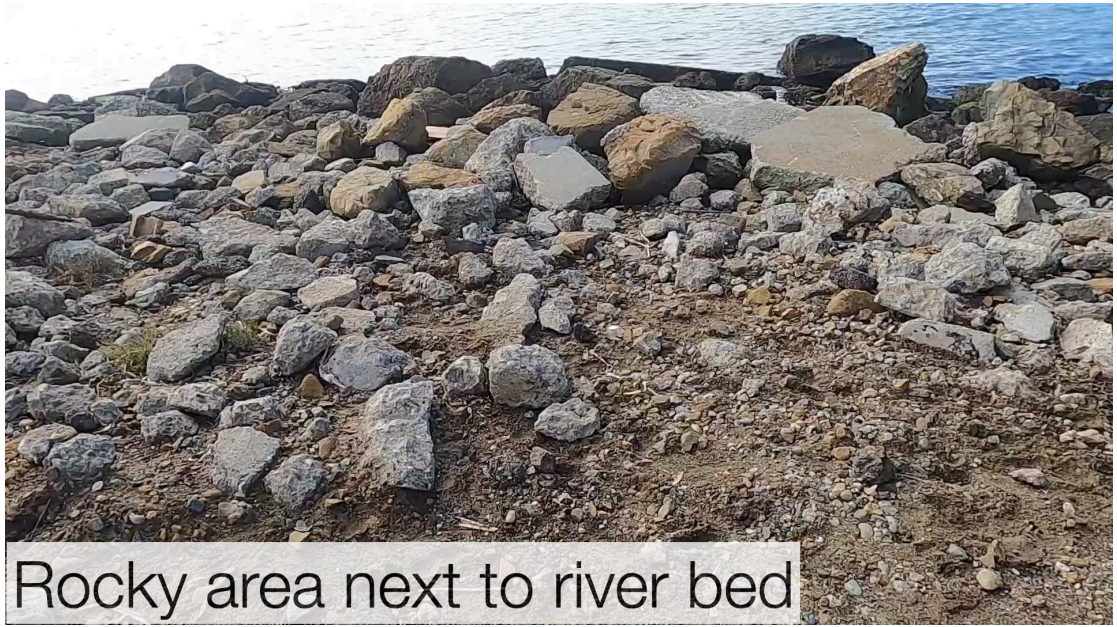
Outer loop Feedback

- E.g



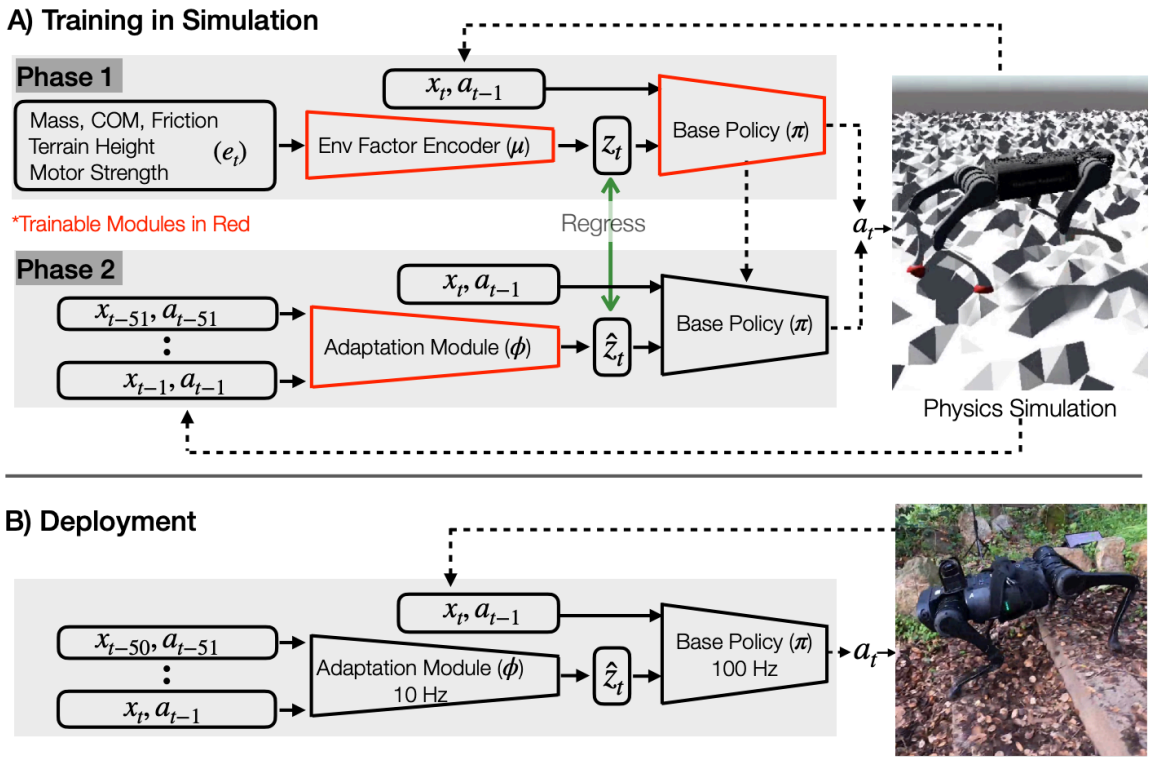
Outer loop Feedback

- RMA
  - Motor Adaptation



RMA: Rapid Motor Adaptation for Legged Robots, Kumar et al., RSS 2021

RMA
Motor Adaptation



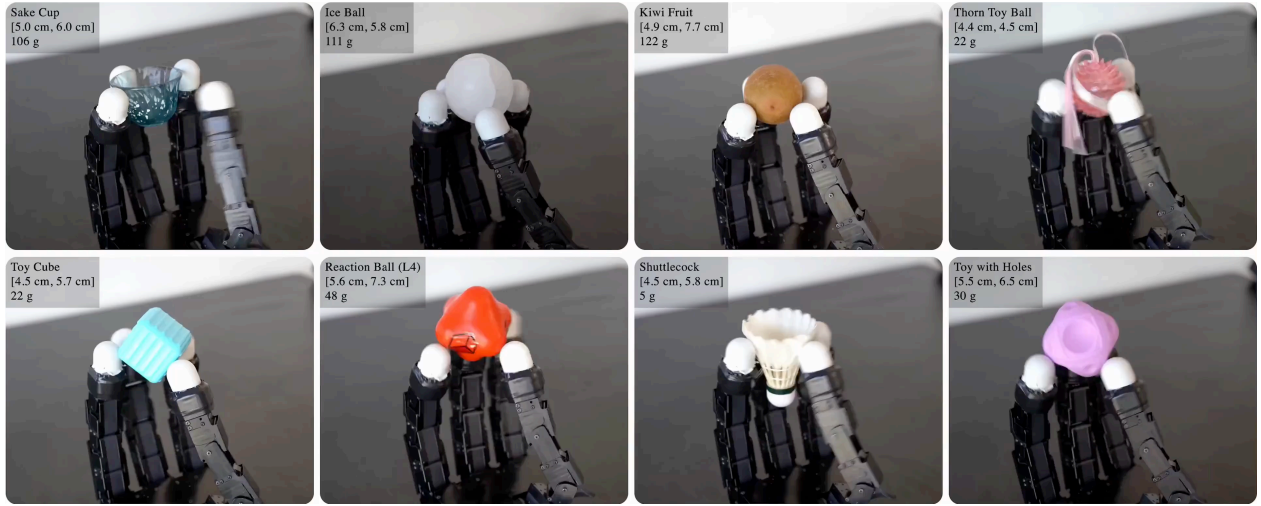
RMA: Rapid Motor Adaptation for Legged Robots, Kumar et al., RSS 2021

RMA
Motor Adaptation

RMA: Rapid Motor Adaptation for Legged Robots, Kumar et al., RSS 2021

RMA
Motor Adaptation

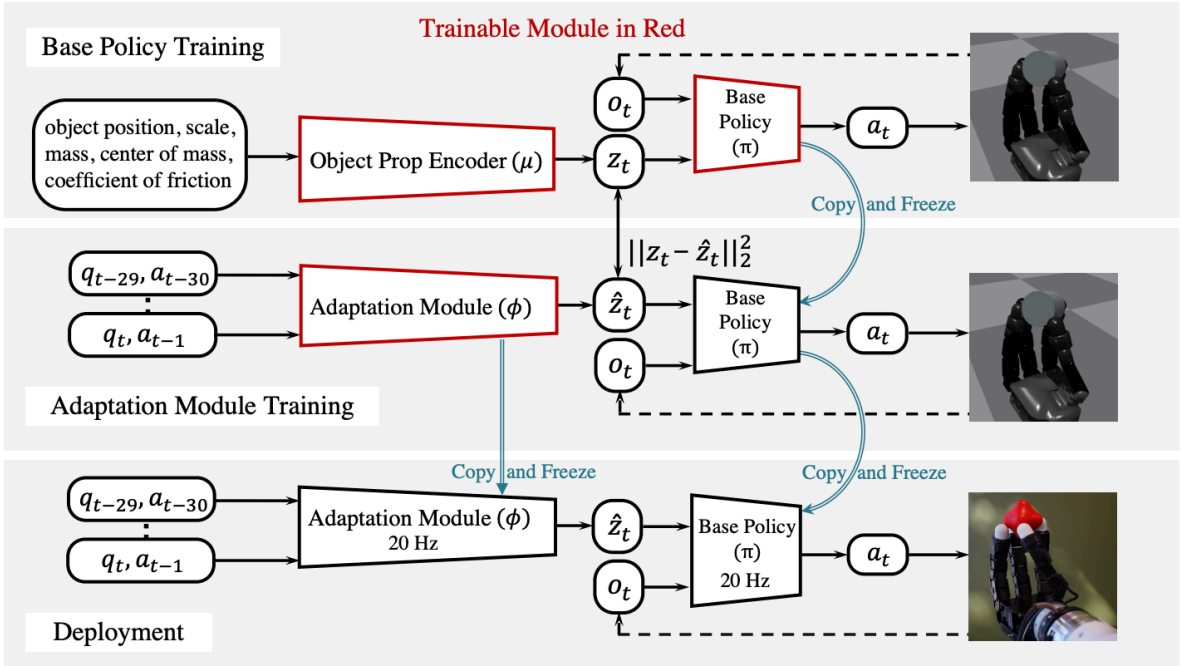
- RMA for in-hand object rotation



In-Hand Object Rotation via Rapid Motor Adaptation, Qi et al., CoRL 2022

RMA
Motor Adaptation

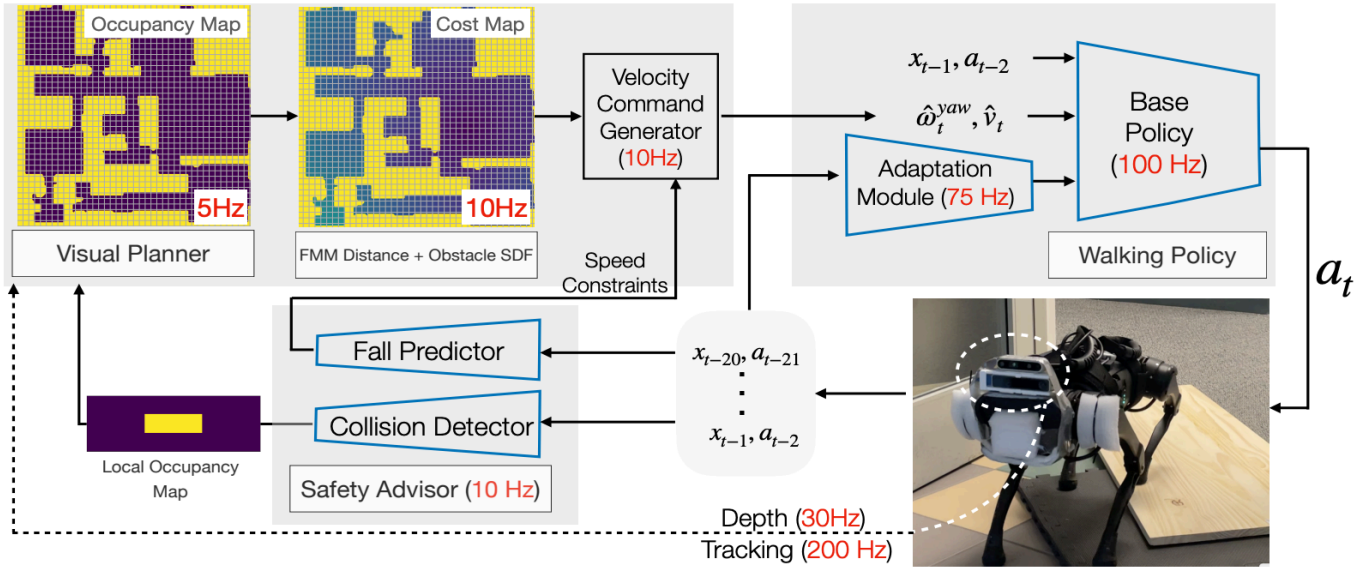
- RMA for in-hand object rotation



In-Hand Object Rotation via Rapid Motor Adaptation, Qi et al., CoRL 2022

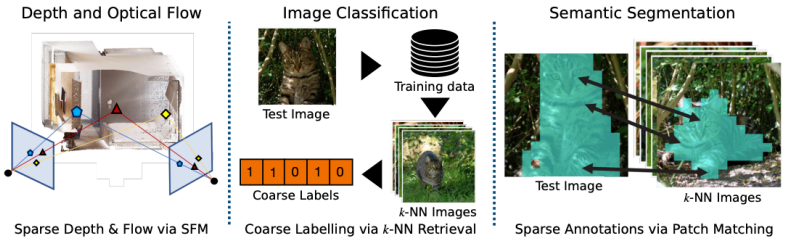
RMA
Motor Adaptation

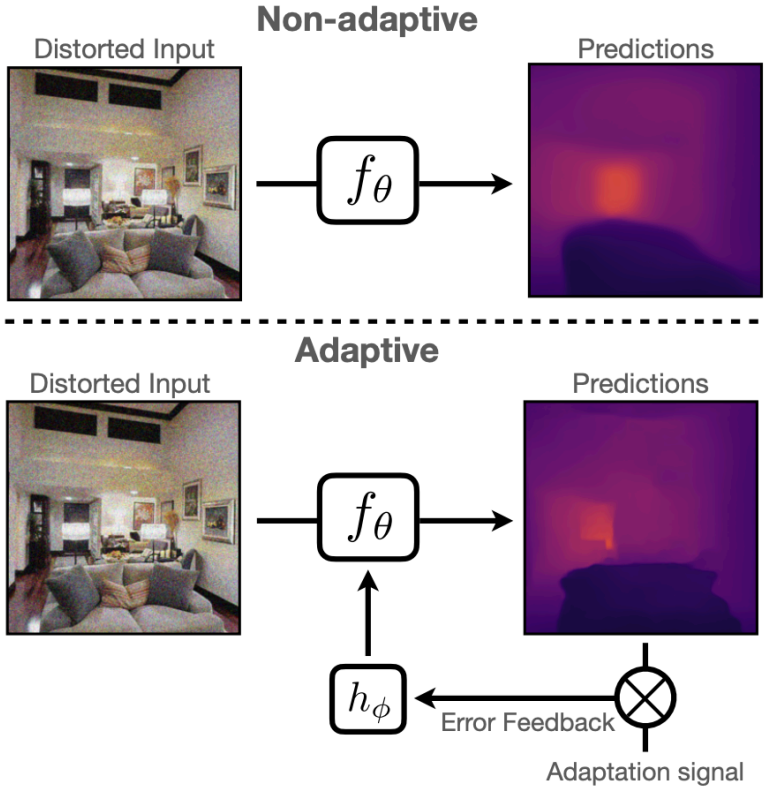
- +visual sensing



Rapid Network Adaptation, Yeo et al., ICCV 2023. https://rapid-network-adaptation.epfl.ch/

RNA
Network Adaptation

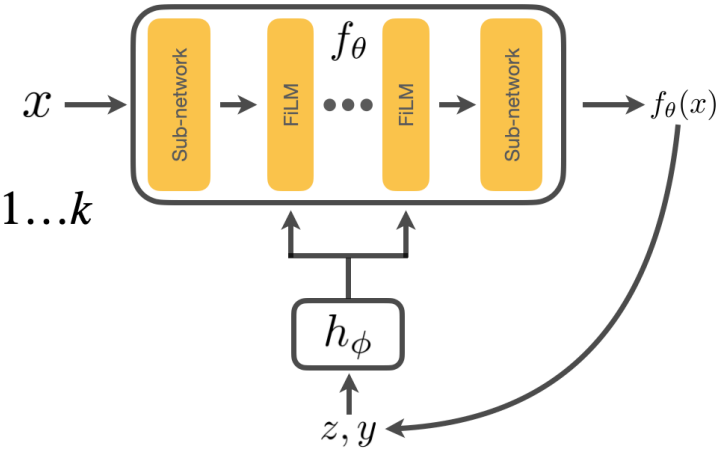




Rapid Network Adaptation, Yeo et al., ICCV 2023. https://rapid-network-adaptation.epfl.ch/

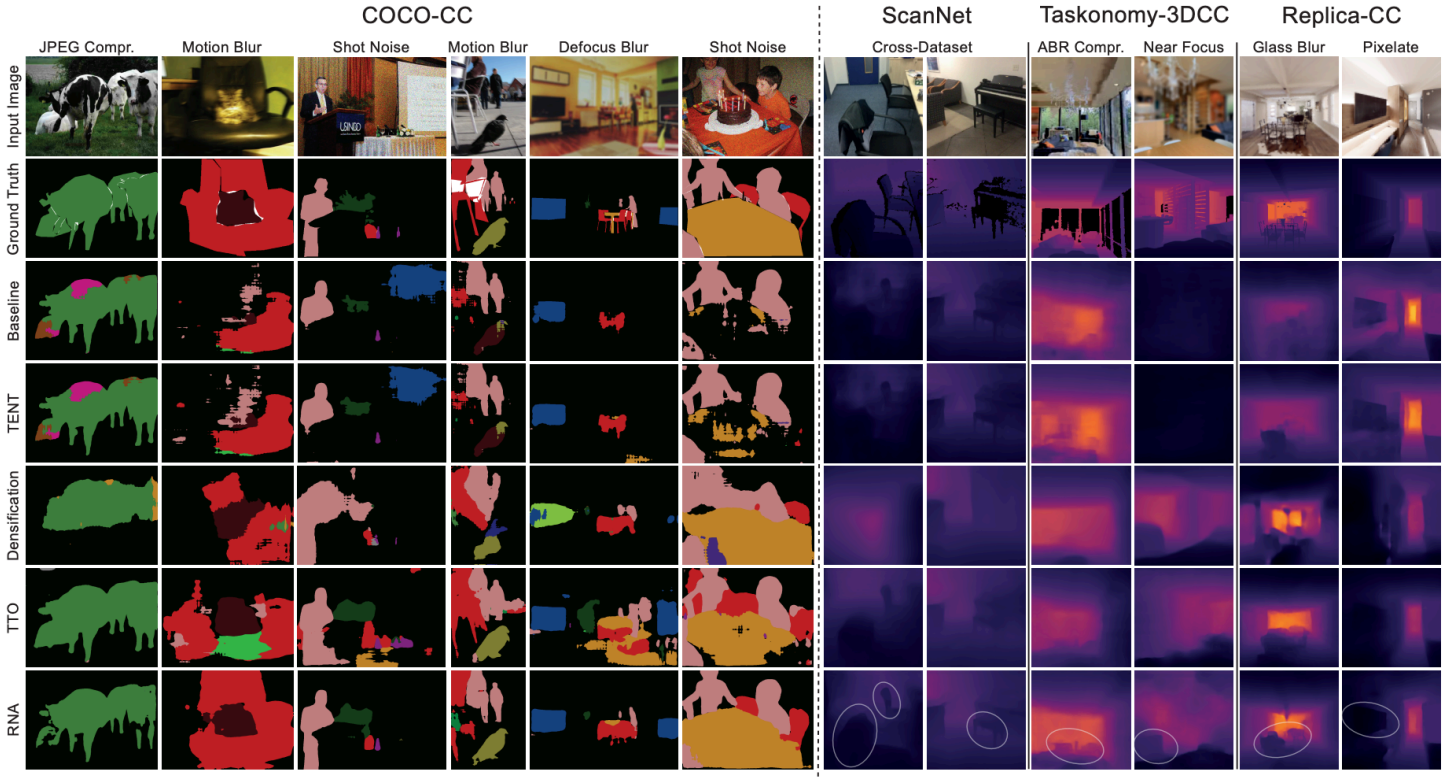
RNA Network Adaptation

- Implementation:
  - Insert k FiLM layers into f_θ.
    - $\text{FiLM}(x_i; \gamma_i, \beta_i) = \gamma_i \odot x_i + \beta_i, \quad \forall i \in 1 \ldots k$
  - h_ϕ takes the adaptation signal, z, and prediction, y, as inputs and predicts $\{\gamma_i, \beta_i\}_{i=1}^k$
  - h_ϕ is able to generalize to unseen shifts and is flexible



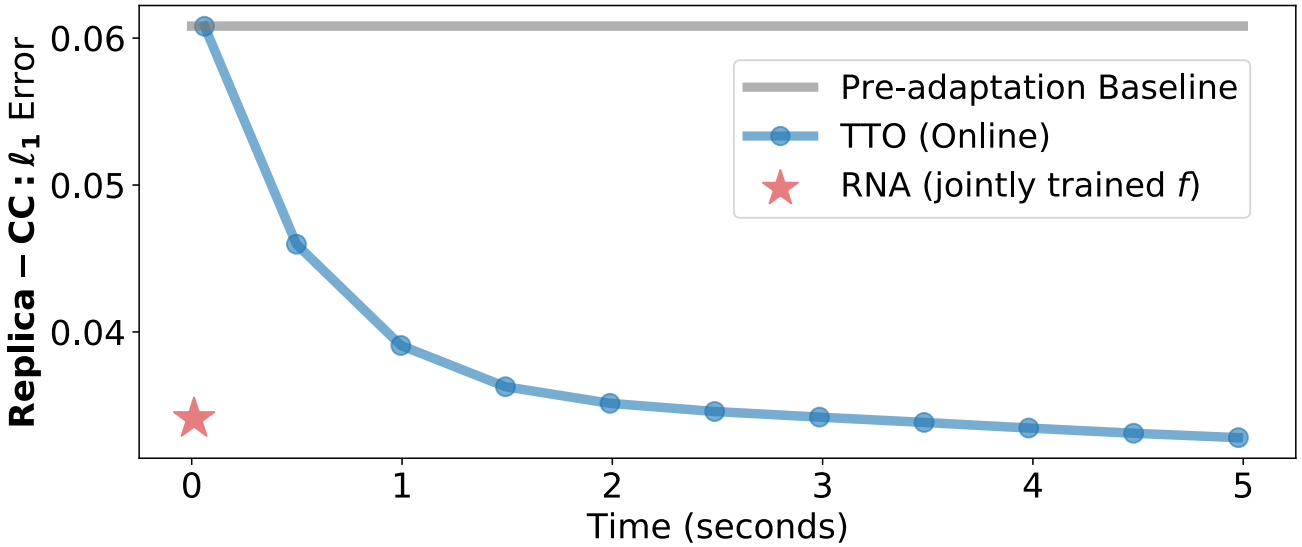
Rapid Network Adaptation, Yeo et al., ICCV 2023. https://rapid-network-adaptation.epfl.ch/

RNA
Network Adaptation



Rapid Network Adaptation, Yeo et al., ICCV 2023. https://rapid-network-adaptation.epfl.ch/

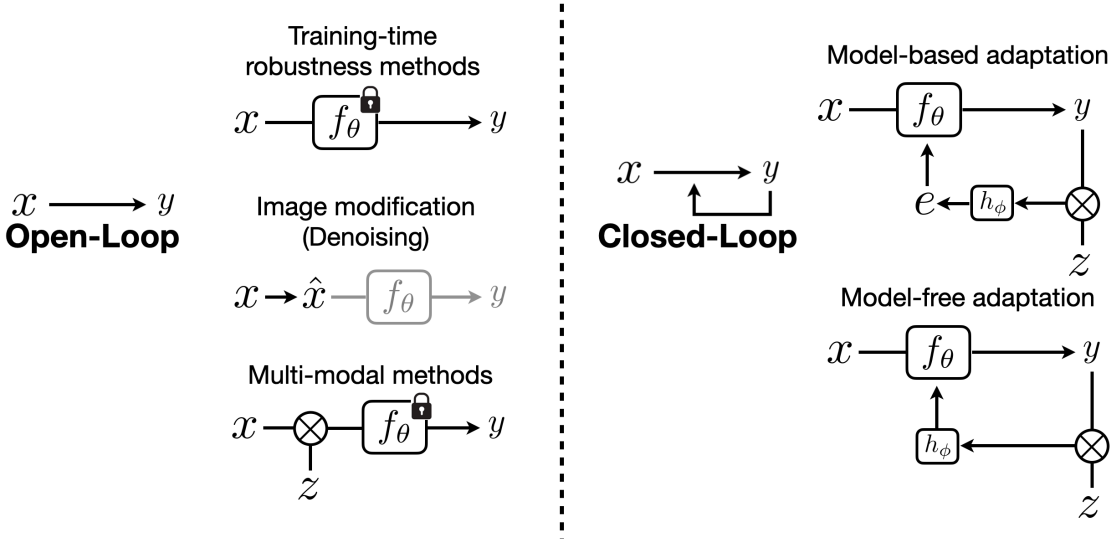
RNA
Network Adaptation



Rapid Network Adaptation, Yeo et al., ICCV 2023. https://rapid-network-adaptation.epfl.ch/

RNA
Network Adaptation

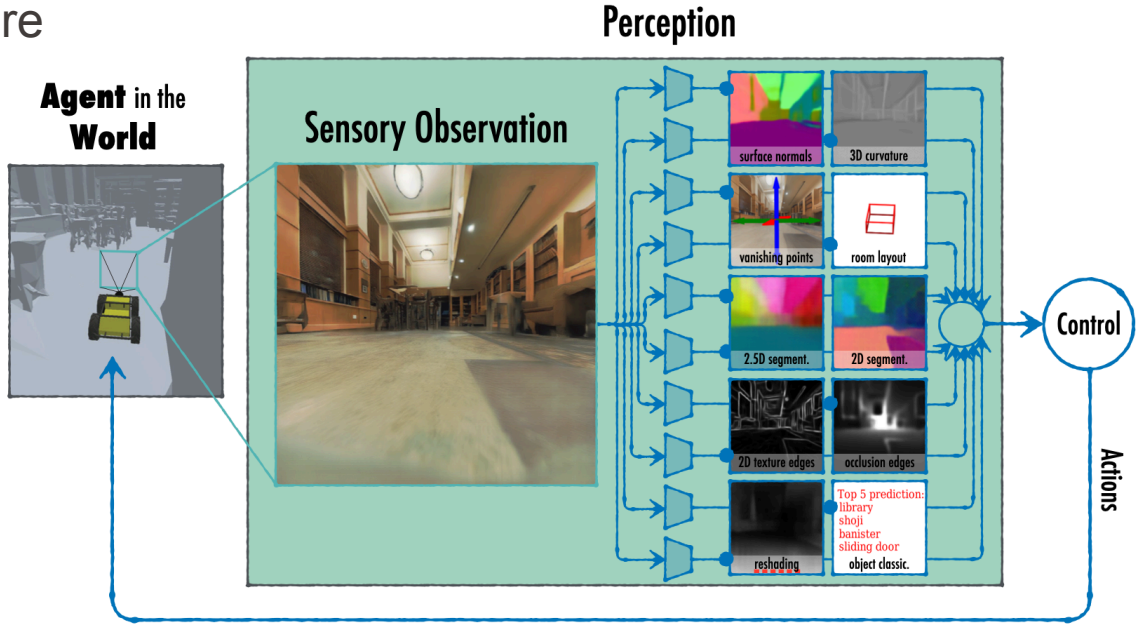
- Methods for handling distribution shifts



Fast Adaptation of Neural Networks Using Test-time Feedback, Yeo et al., 2023

Outer loop Feedback

- Most vision-action systems
  - In active vision lecture



Questions?

https://vilab.epfl.ch/