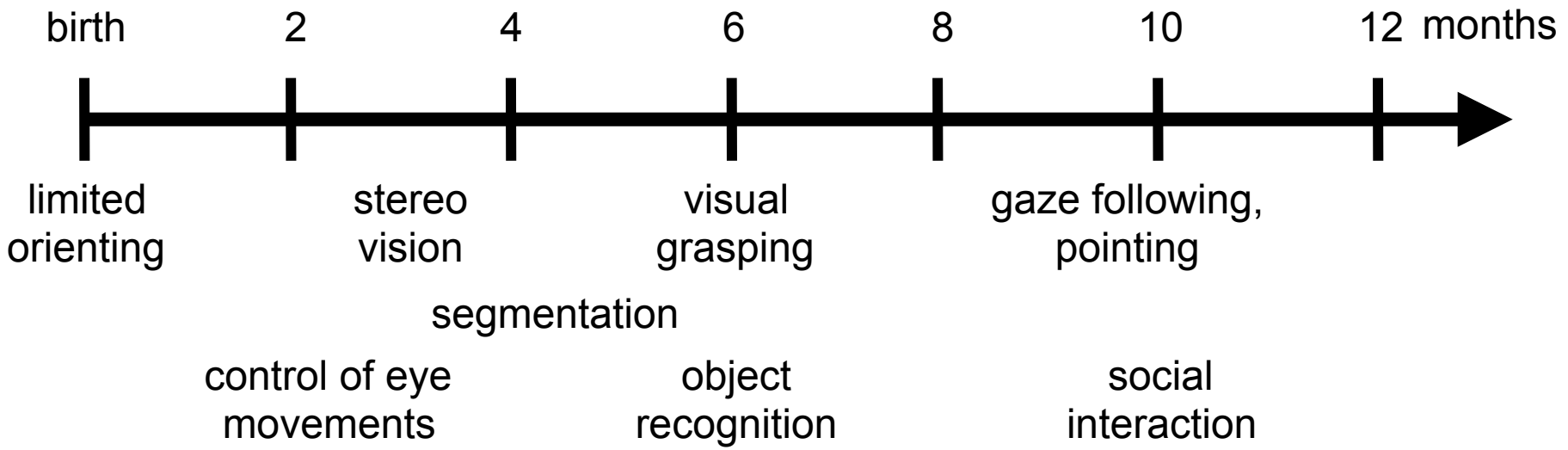
Active Efficient Coding

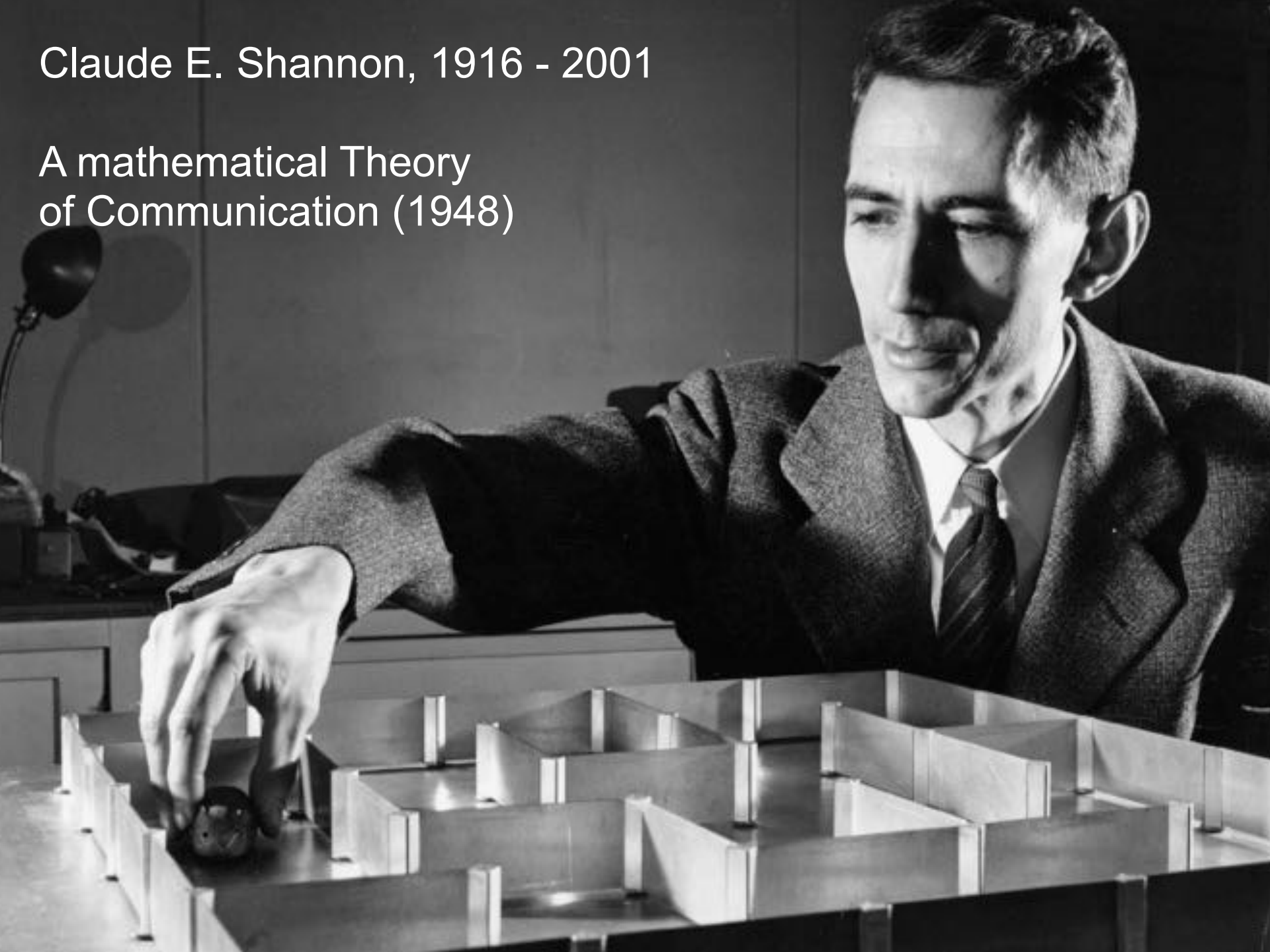
for building self-calibrating vision systems

Jochen Triesch
Frankfurt Institute
for Advanced Studies &
Goethe Univ. Frankfurt

Infant Development

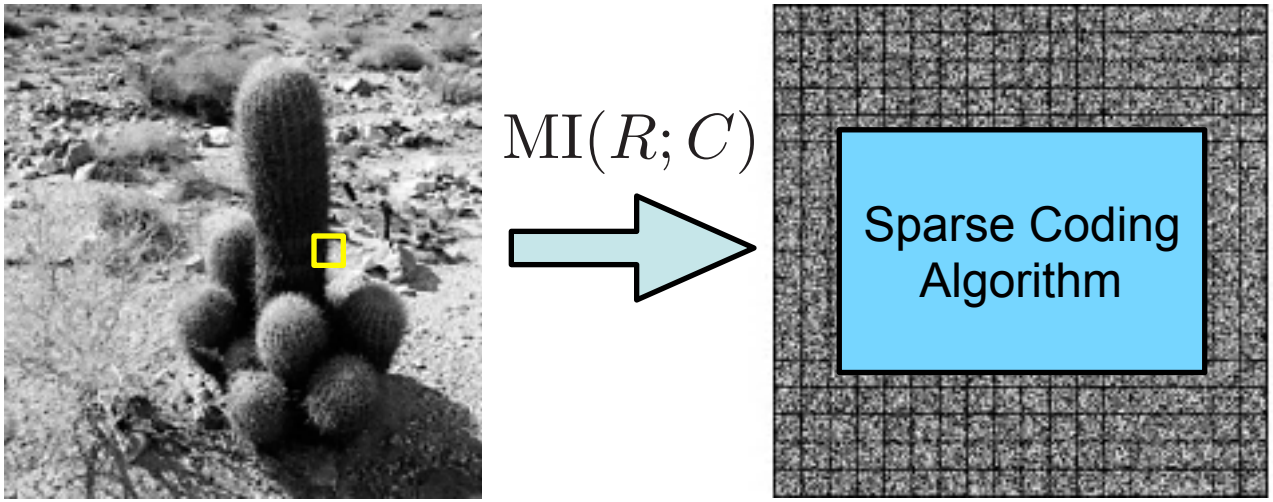


How can a developing agent autonomously learn a good model of the world?

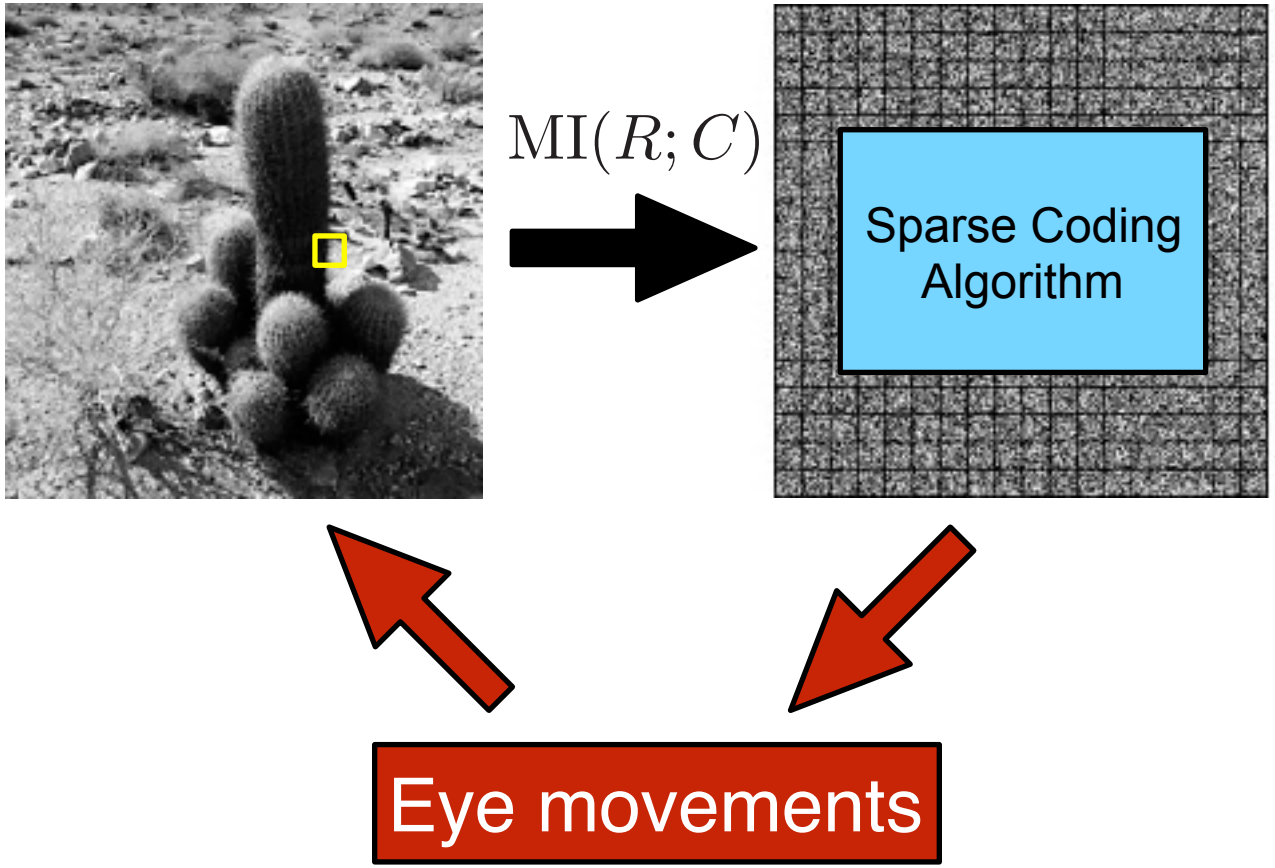

Claude E. Shannon, 1916 - 2001
A mathematical Theory
of Communication (1948)

- **Efficient Coding Hypothesis:** biological sensory systems exploit redundancies in the sensory signals to encode information from the environment more efficiently (Attneave 1954, Barlow 1961, Laughlin 1981, Linsker 1988, Attick 1992, ...)

- **Sparse Coding:** sensory systems employ representations with few active neurons for any given stimulus (Földiák 1990, Olshausen&Field 1996, Bell&Sejnowski 1997, Rao&Ballard 1999, ...)



Active Efficient Coding




C. Rothkopf


C. Teulière


B. Shi

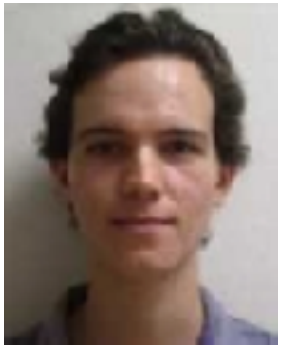

M. Fronius


S. Jeong


L. Lonini


V. Narayan


T. Chandrapala


S. Forestier


A. Priamikov


S. Eckmann

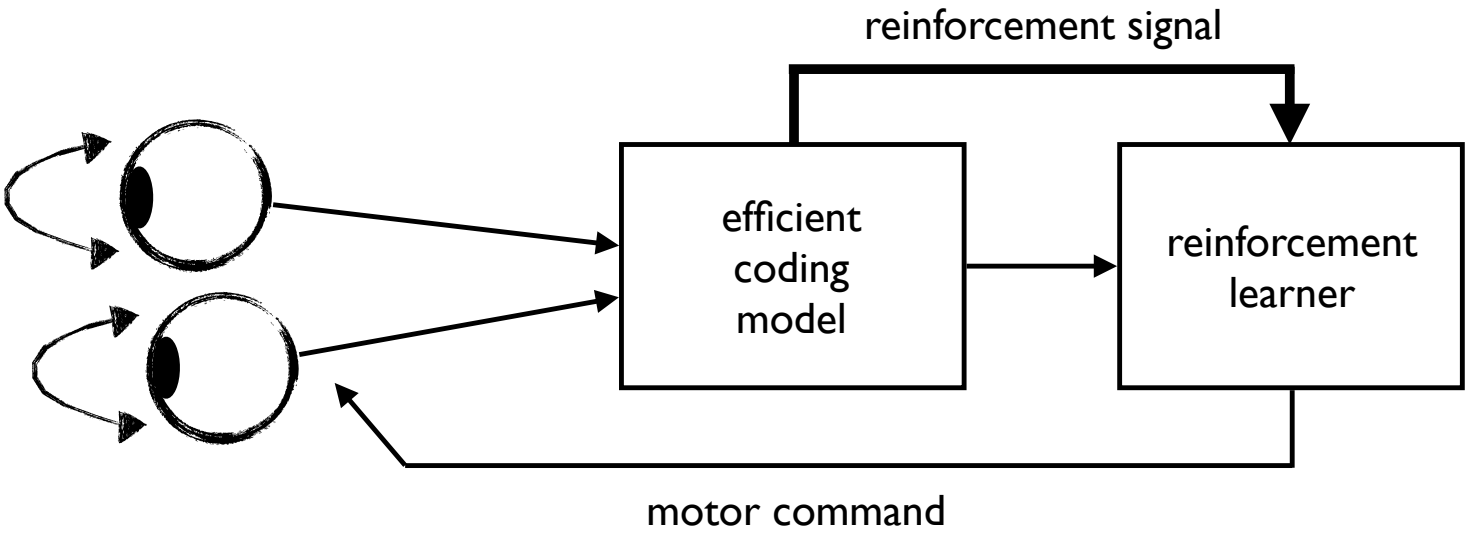

I. Schneider


L. Klimmasch


A. Lelais


C. Wilmot

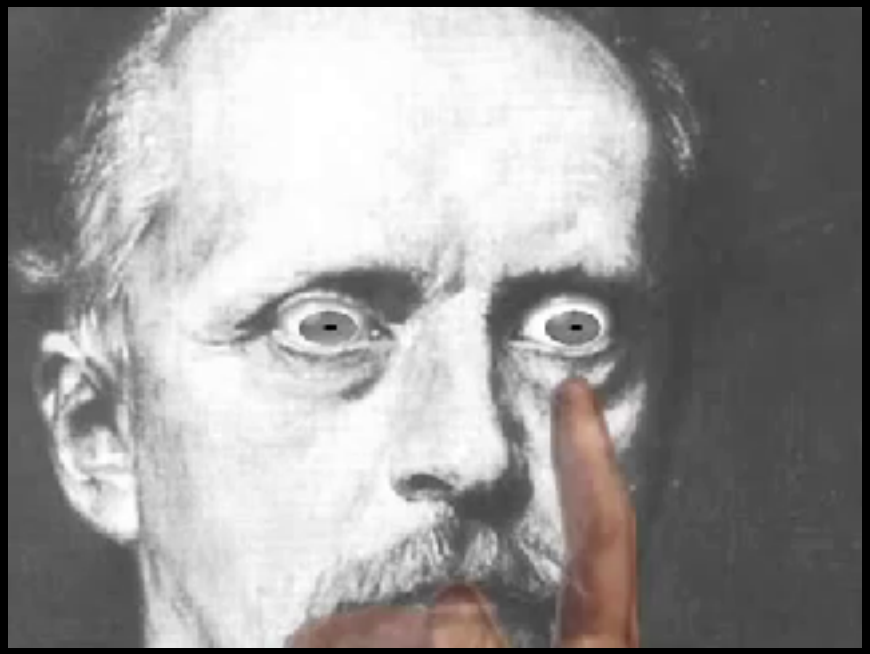
Active Efficient Coding



- reinforcement signal encodes a measure of coding efficiency (intrinsically motivated learning, cf. Schmidhuber, 1991, 2009; Oudeyer 2007; Gottlieb et al., 2013)
- allows fully self-calibrating active perception systems

Zhao et al., 2012; Teulière et al., 2014

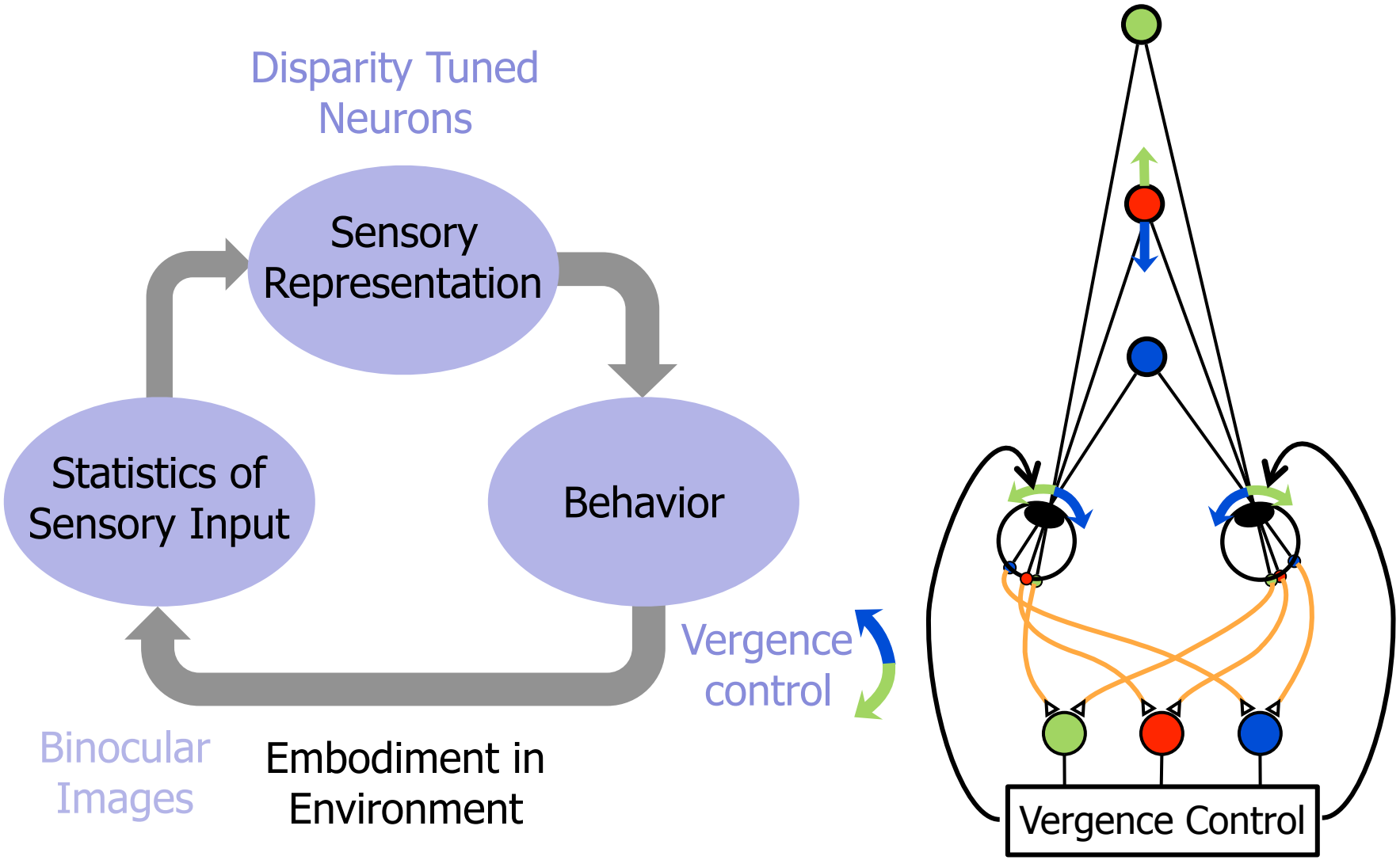
vergence



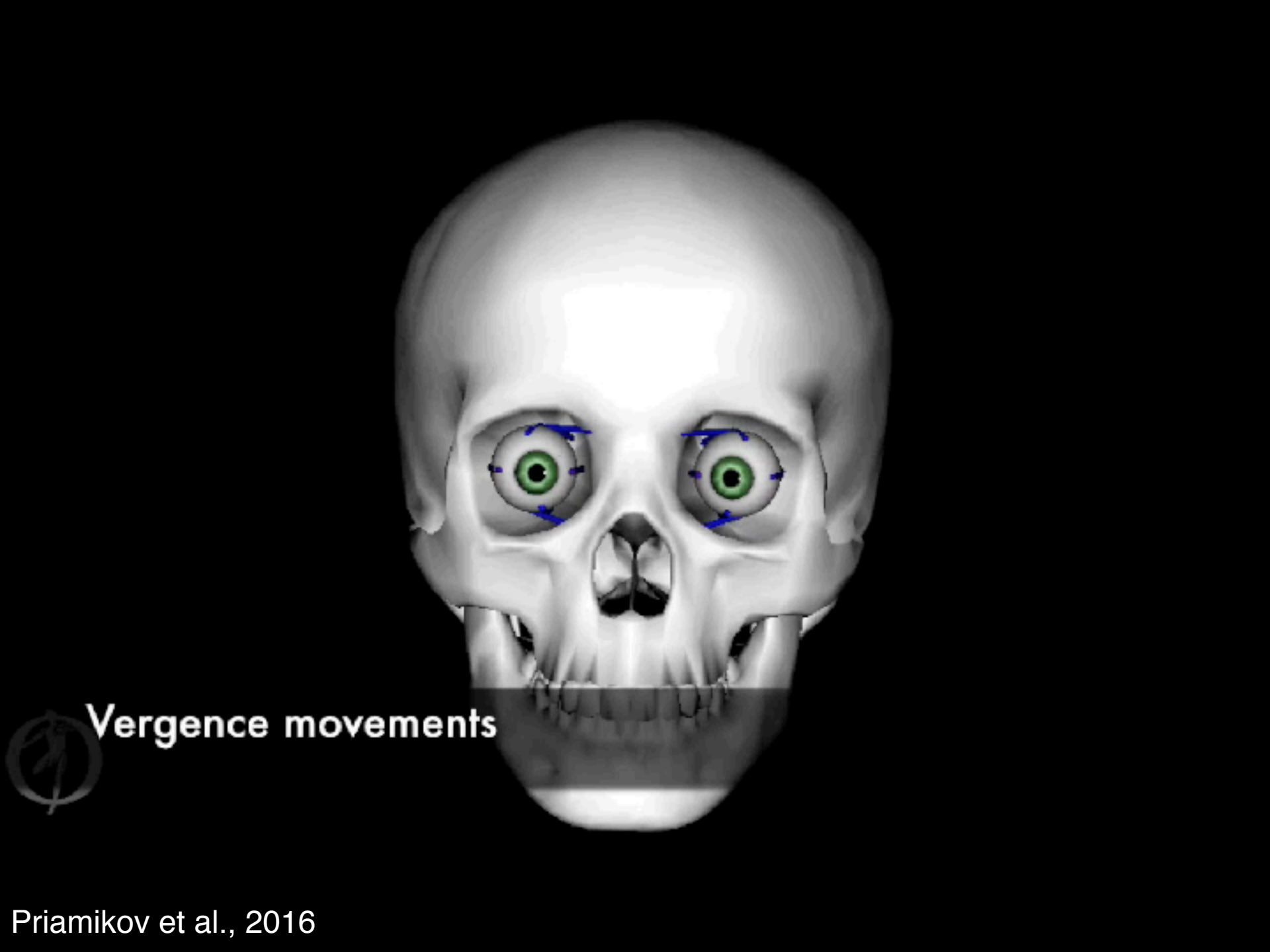
pursuit

animations by Mark Bolding

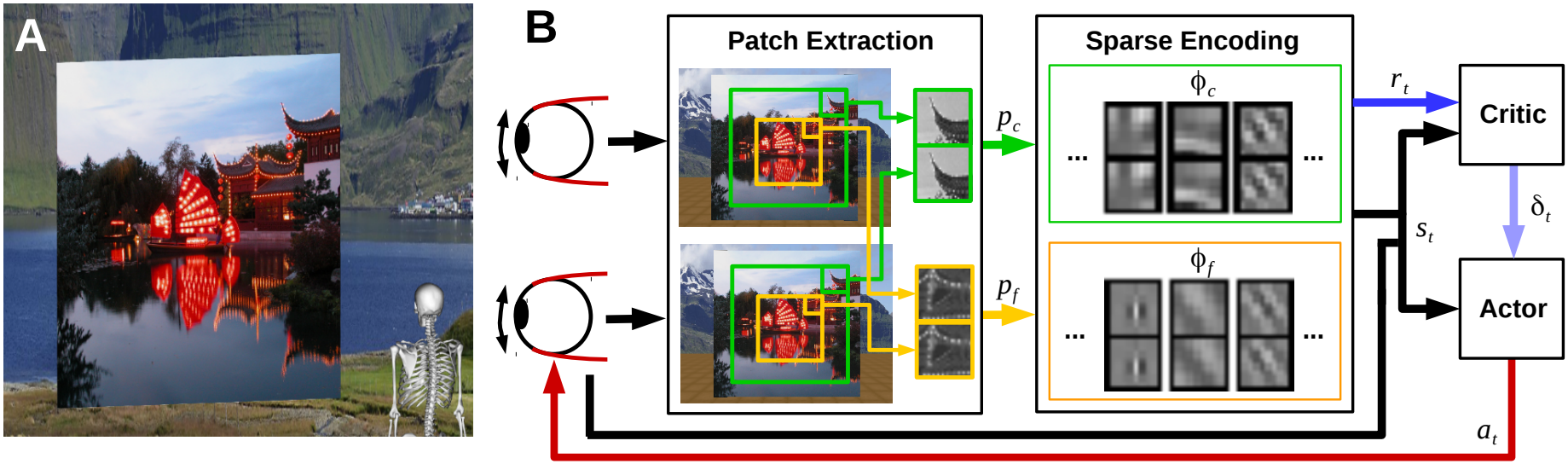
The case of binocular vision



Priamikov et al., 2016



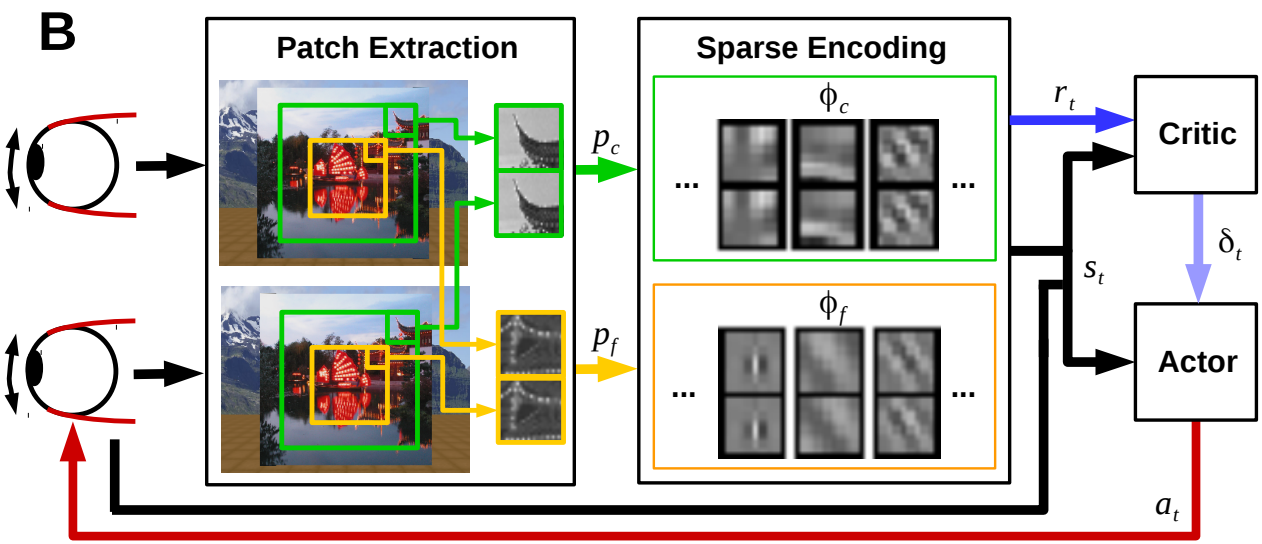
Priamikov et al., 2016

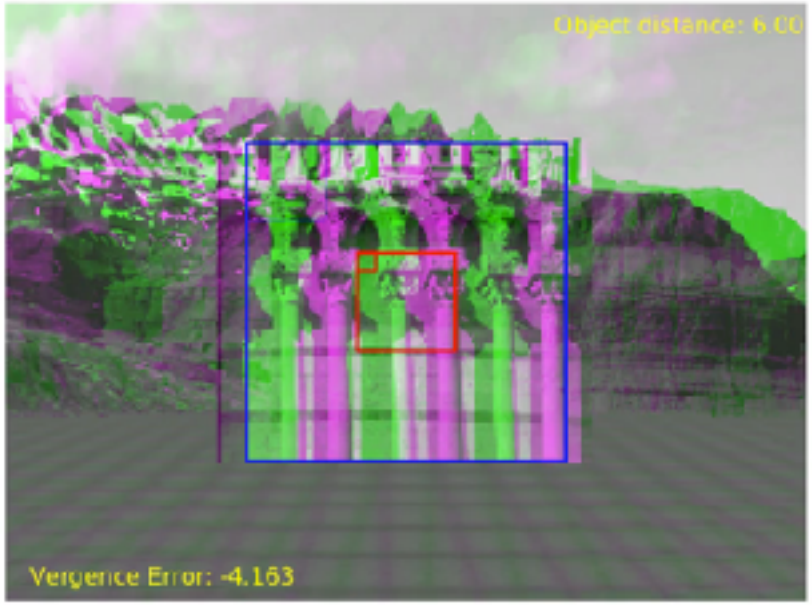


- Sparse Coding: matching pursuit algorithm with gradient descent learning (Mallat & Zhang, 1993)

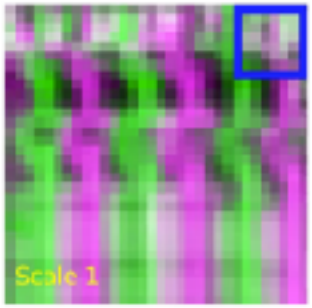
- Reinforcement Learning: CACLA+VAR algorithm (Van Hasselt & Wiering, 2007)

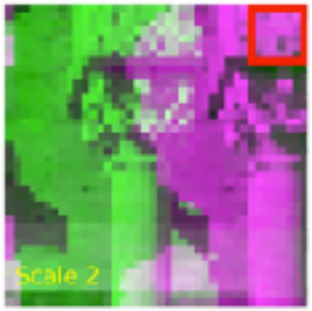

A


B
Patch Extraction
Sparse Encoding
p_c
p_f
ϕ_c
ϕ_f
r_t
Critic
δ_t
s_t
Actor
a_t


Object distance: 6.00
Vergence Error: -4.163

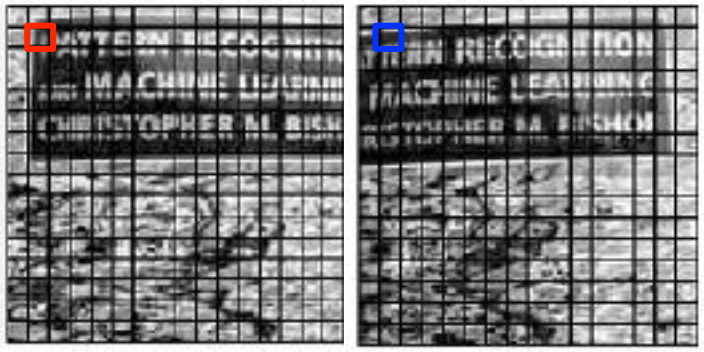
1 pixel = 0.22°


Scale 1

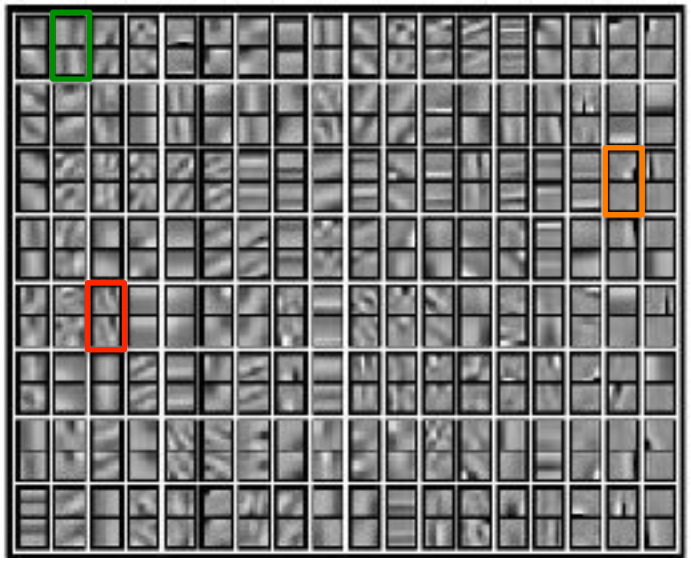

Scale 2

Learned encoding of stereo patches

original binocular image



learned stereo-basis functions



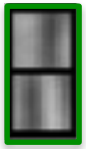
stereo patches:

L
R

sparse encoding:

$$\begin{pmatrix} I_{i,L} \\ I_{i,R} \end{pmatrix} \approx \sum_{n=1}^{N} a_{i,n}(t) \begin{pmatrix} \phi_{n,L}(t) \\ \phi_{n,R}(t) \end{pmatrix}$$

tuned to zero disparity



tuned to non-zero disparity

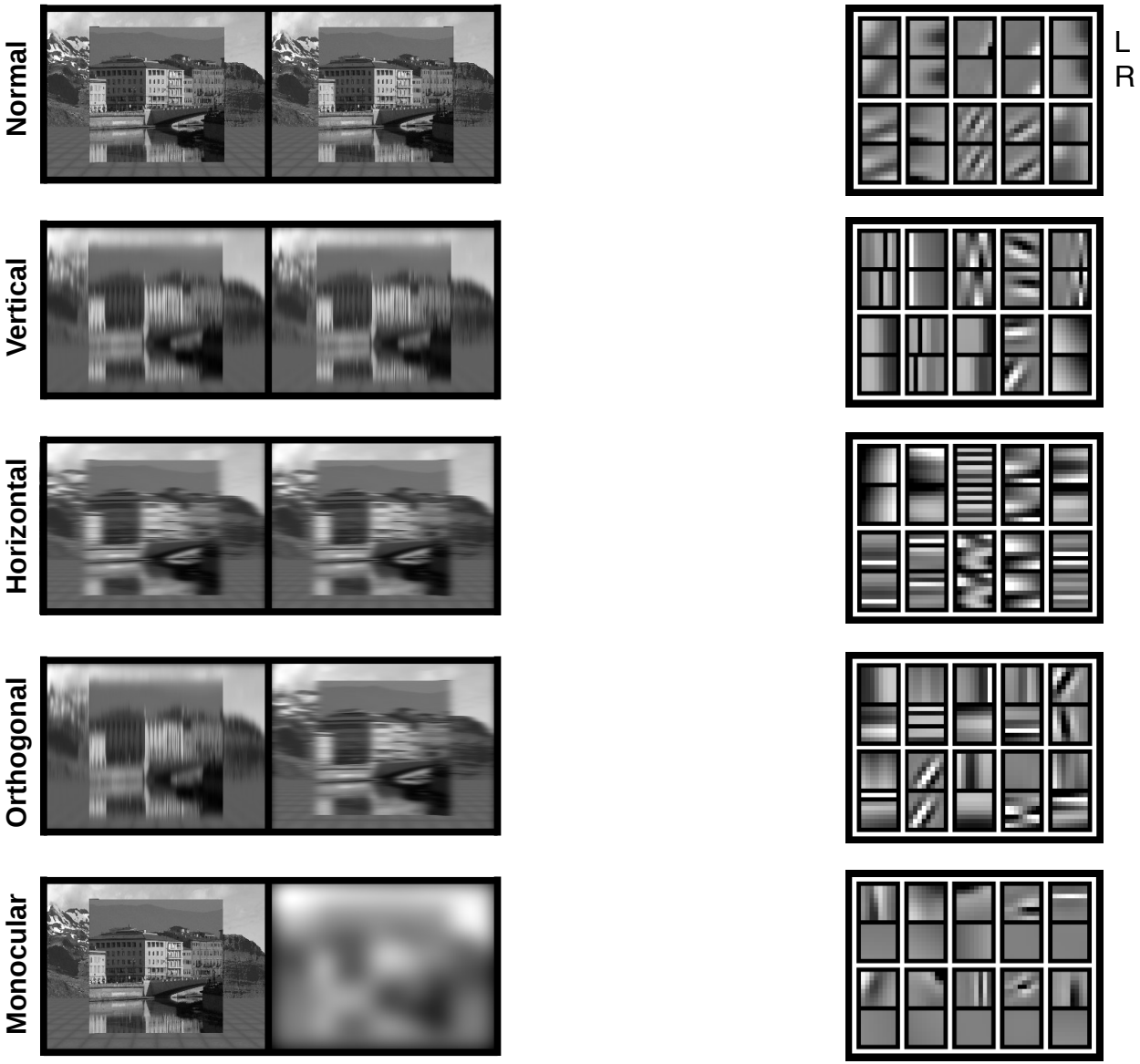
monocular

Alternate Rearing Conditions

Blurring the input along the
x and/or y-direction:

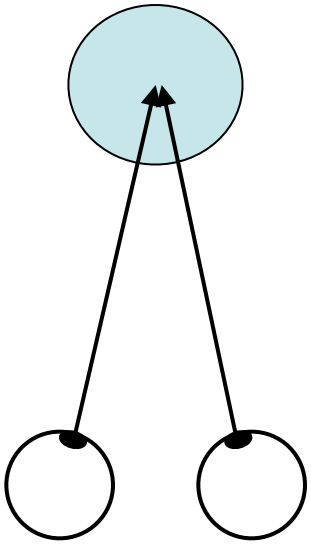
$$K_{\sigma_x,\sigma_y}(x,y) = \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right)$$

Klimmasch et al., SAB 2018 (best paper award); Klimmasch et al., in preparation

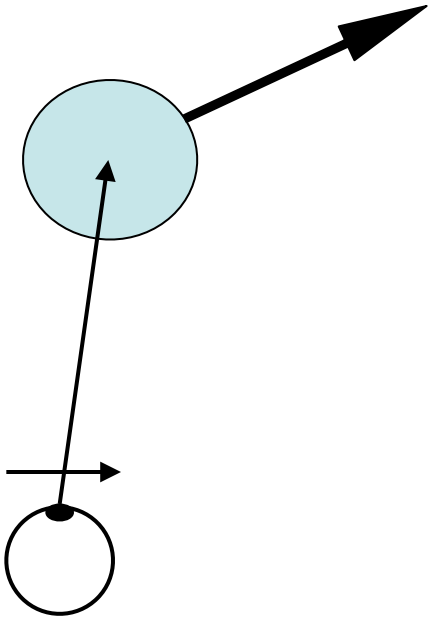


Klimmasch et al., SAB 2018 (best paper award); Klimmasch et al., in preparation

How general is Active Efficient Coding?



disparity tuning and vergence eye movements

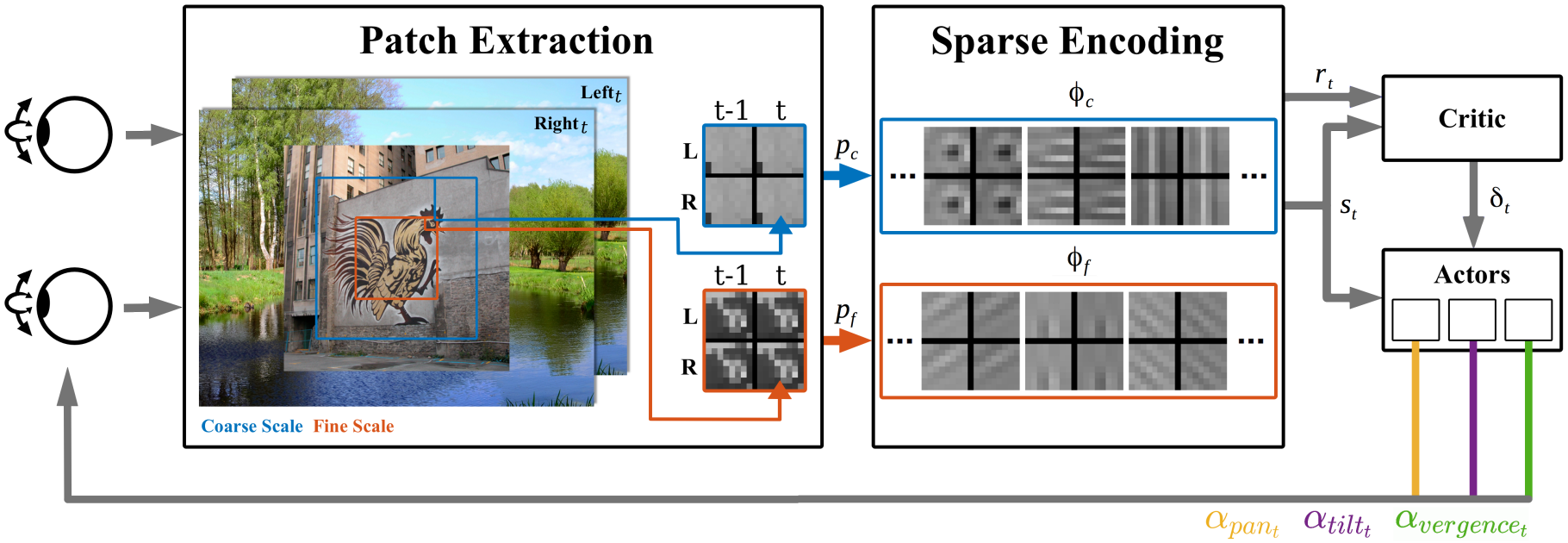


motion tuning and pursuit eye movements

Zhang et al. 2014, 2016; Teulière et al. 2014; Narayan et al. 2014

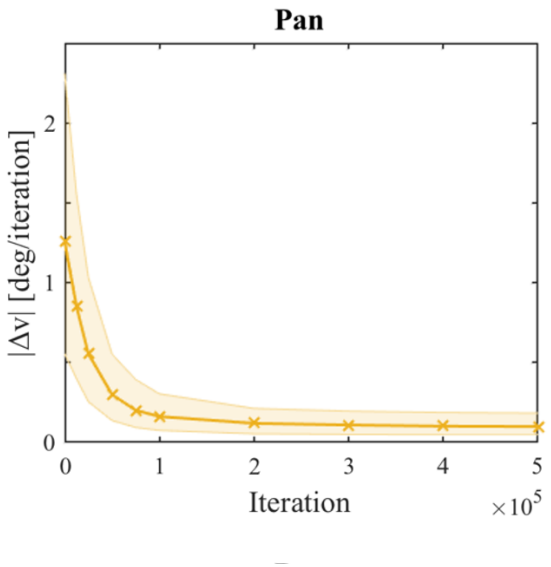
Learning Environment

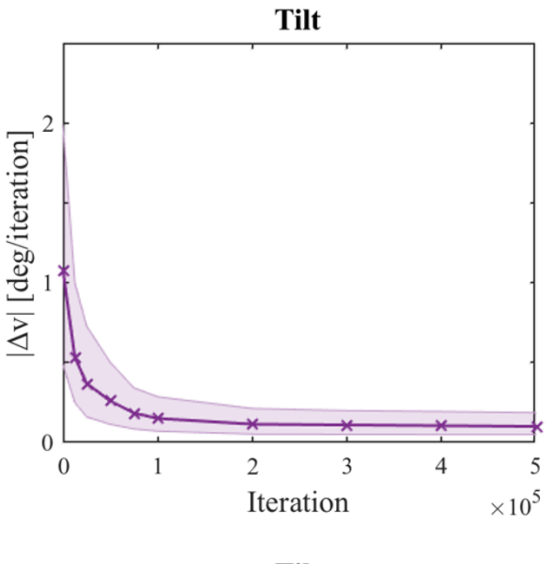
Lelais et al., submitted

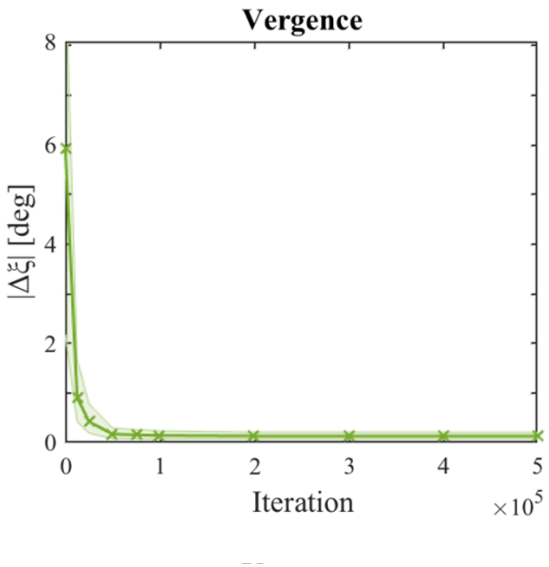


- Sparse Coding: matching pursuit algorithm with gradient descent learning (Mallat & Zhang, 1993)
- Reinforcement Learning: natural actor critic algorithm with discrete actions (Bhatnagar et al., 2009)

Learning Performance



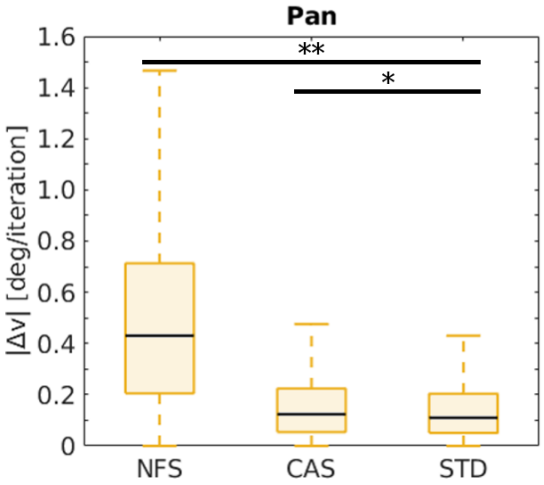


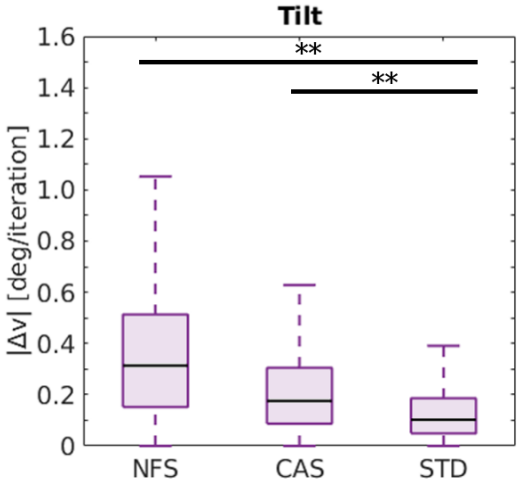


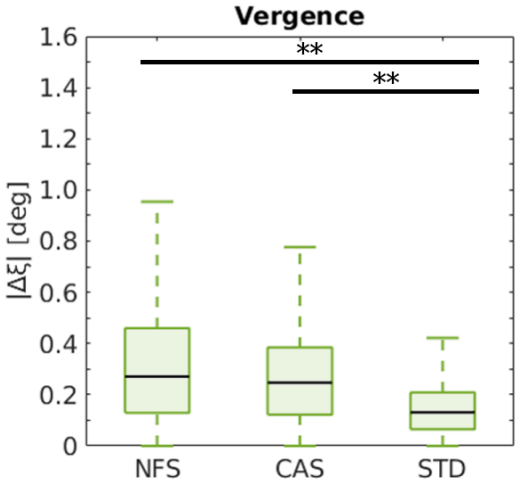
Lelais et al., submitted

speed: × 2.4

What limits the performance?



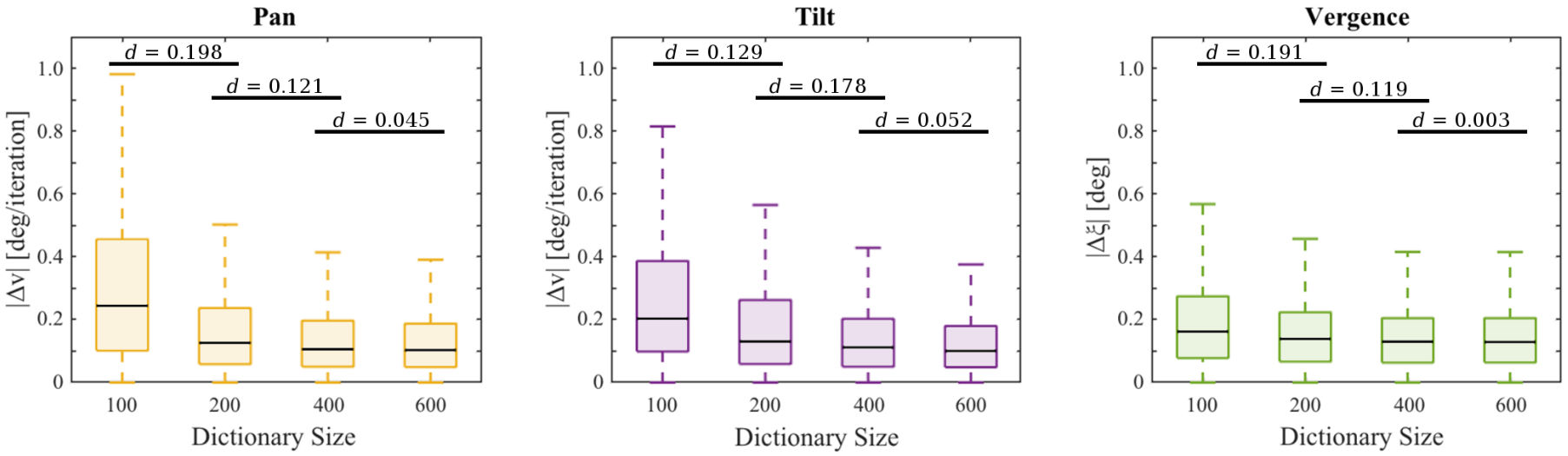




NFS: no fine scale (sensory limitation)
CAS: coarse action set (motor limitation)
STD: standard configuration

Lelais et al., submitted

Are there also internal limitations?

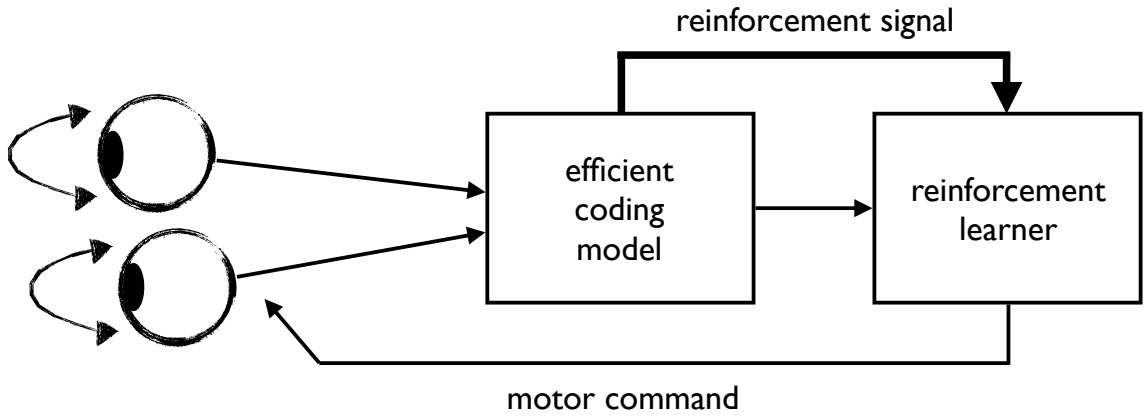


self-improvement until system hits a sensory, motor, or internal limitation

Lelais et al., submitted

Discussion

- **Active Efficient Coding:** „*Move your sense organs to make sensory encoding as efficient as possible!*“
- **self-calibration:** of a range of eye movements
  - vergence & pursuit
  - but also: optokinetic nystagmus, torsion, accommodation
- **general:** applicable to other sensory modalities



Outlook 1: medical applications

Amblyopia:
What causes derailed development?
Can we develop better treatment methods?

Eckmann et al., submitted; Schneider et al., in preparation

Outlook 2: event-based cameras

- **Advantages of new sensors:** low power (<10 mW), low latency (> 10,000 fps), high dynamic range (> 120 dB)
- **Funding:** French Tech Chair (Université Clermont Auvergne)

Prophesee

PhD position in Clermont-Ferrand available!

JOHANNA QUANDT STIFTUNG

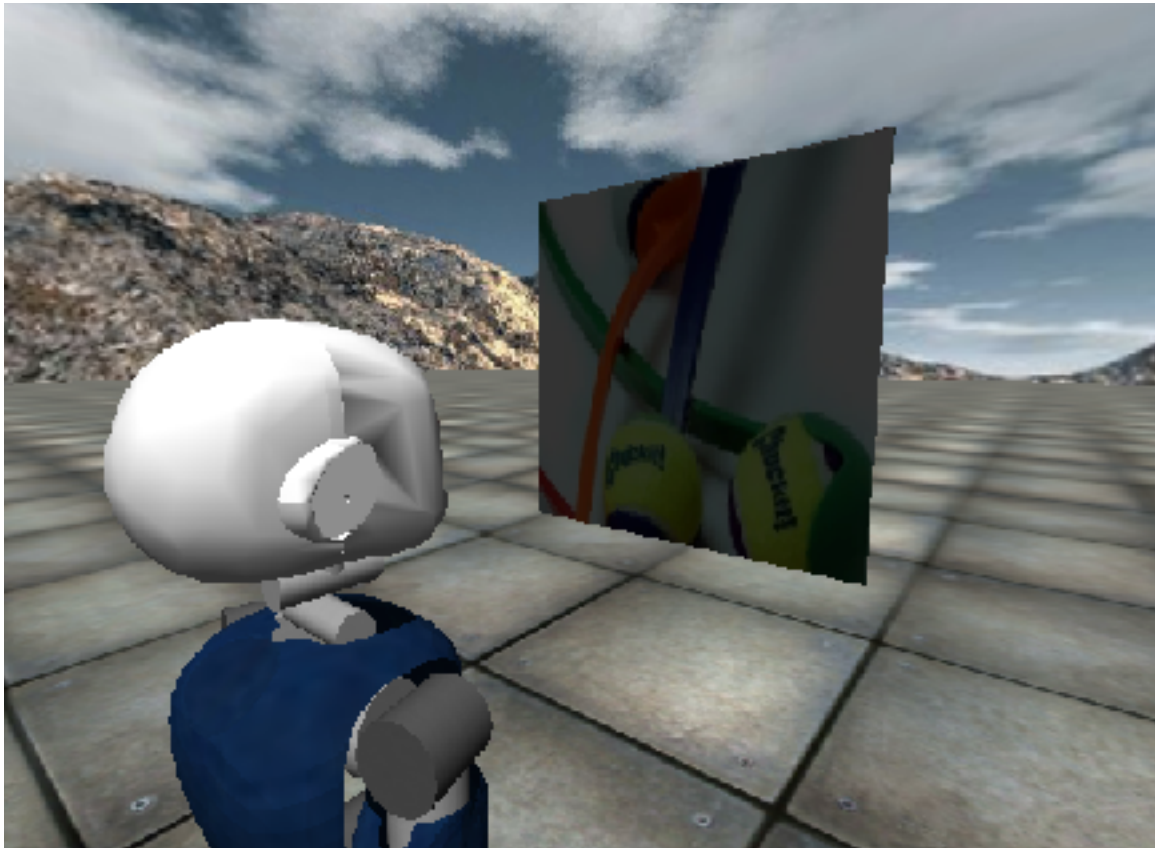
NEURDREAM

SPP 2041 Computational Connectomics

GOAL Robots

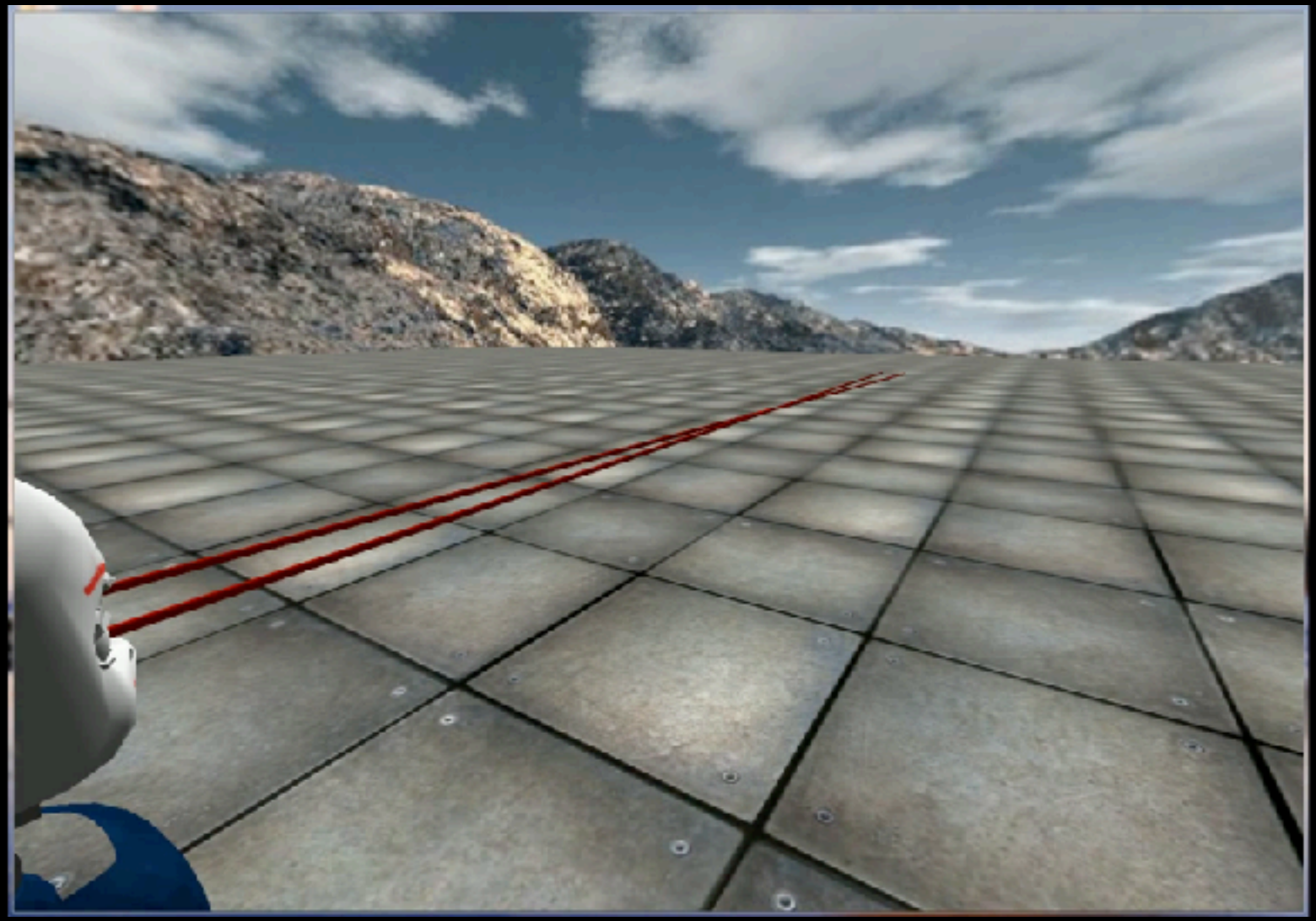
Thank You!

Adding Overt Attention



Binocular saliency mechanism: look at the most surprising location given your current „world“ model

Bruce & Tsotsos, 2009; Zhu et al., 2017



Zhu et al., ICDL 2017

Test on physical iCub