Pervasively neural cognitive architectures: What they entail, what they may deliver, how they remain elusive

Gregor Schöner
gregor.schoener@rub.de
Institute for Neural Computation
Ruhr-University Bochum, Germany

Neural inspiration in autonomous driving/robotics

- has returned center stage largely through Deep Learning, primarily in vision
- past discomfort about lack of theoretical penetration/proof of competence has faded as performance has become better
- rely on benchmarks with an eye for super-human performance

Neural inspiration

- acceptance is helped by fact that NN are used in a narrow way…
- essentially as intelligent filters that extract information for neural representations… on which action/decisions are based
- decision making, planning, coordination, and control remain to a large extent outside the neural metaphor

Neural inspiration

- this helps circumscribing the problem on limited verifiability
- but is also a limitation in itself… => underusing a potential revolution…
- => embodied cognition in closed loop rather than static representations and plans

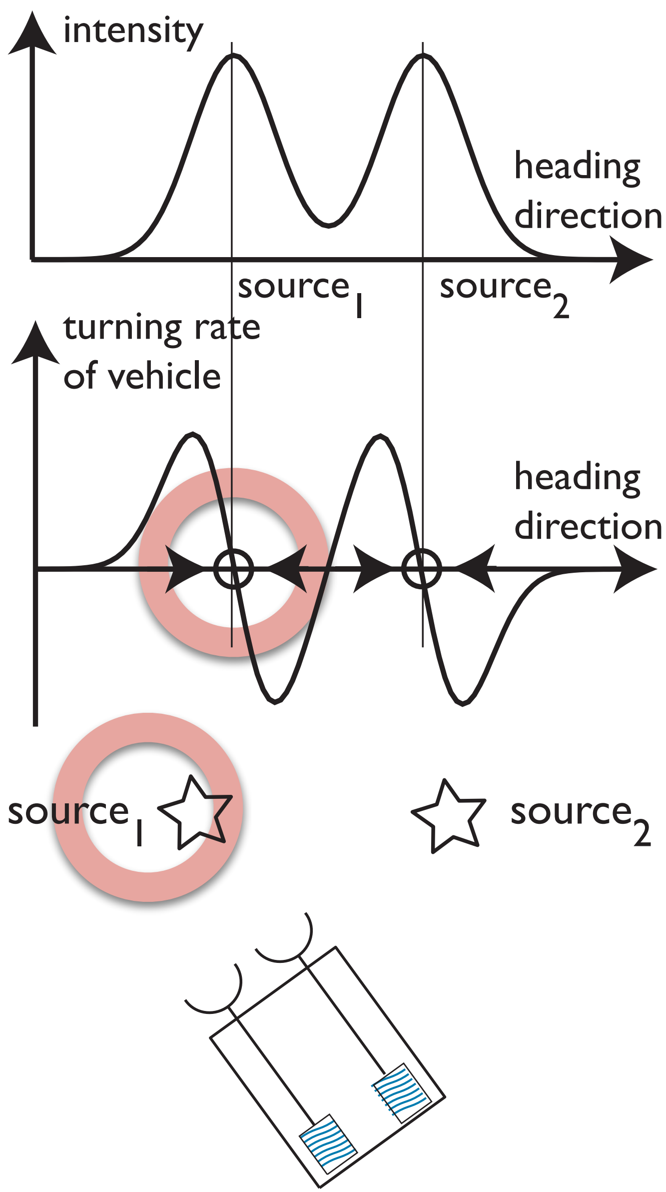
Cognition in the wild…

- attention/gaze
- active perception/working memory
- action plans/decisions/ sequences
- goal orientation
- motor control
- background knowledge
- learning from experience

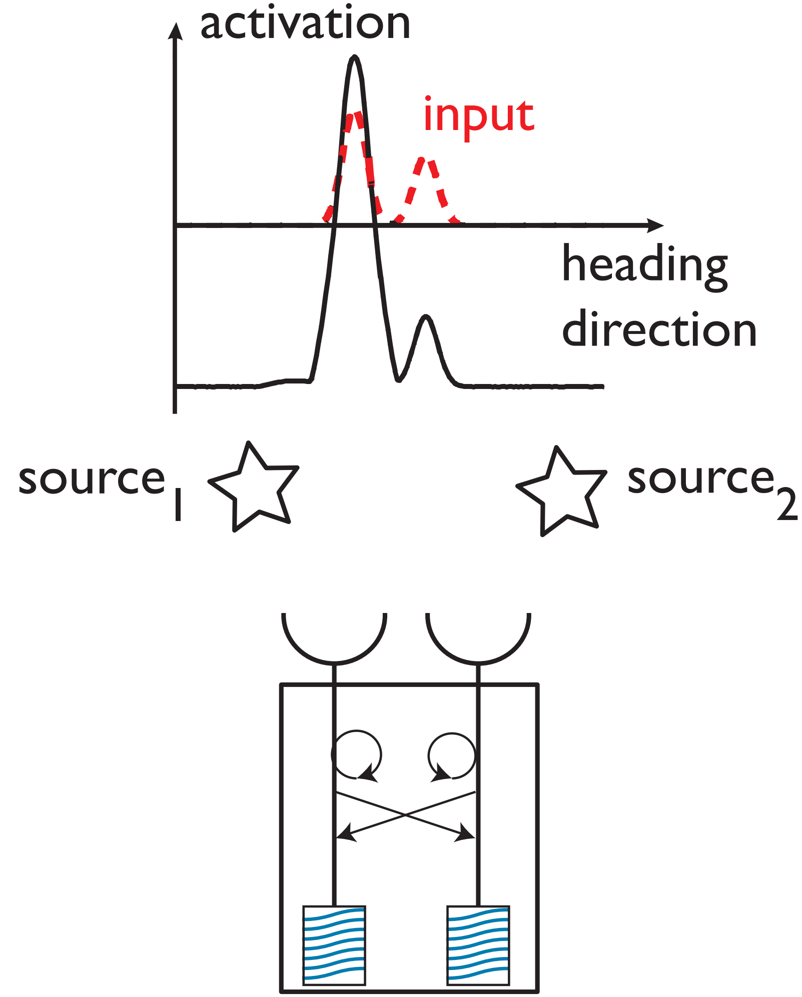
=> underlying neural processes have dynamic properties

- graded state
- continuous time
- from which discrete events and categorical behavior emerge
- continuous/intermittent link to the sensory and motor surfaces in closed loop
- => dynamics
- => stability

Two forms of dynamics



behavioral dynamics

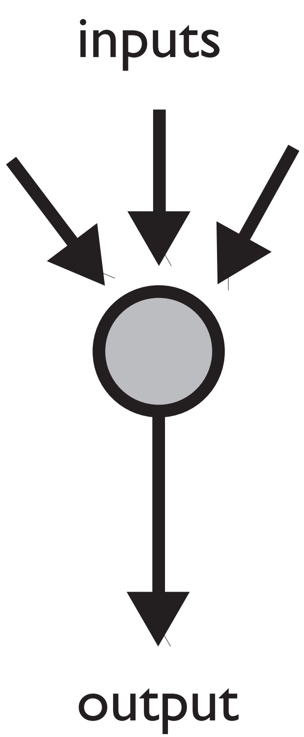


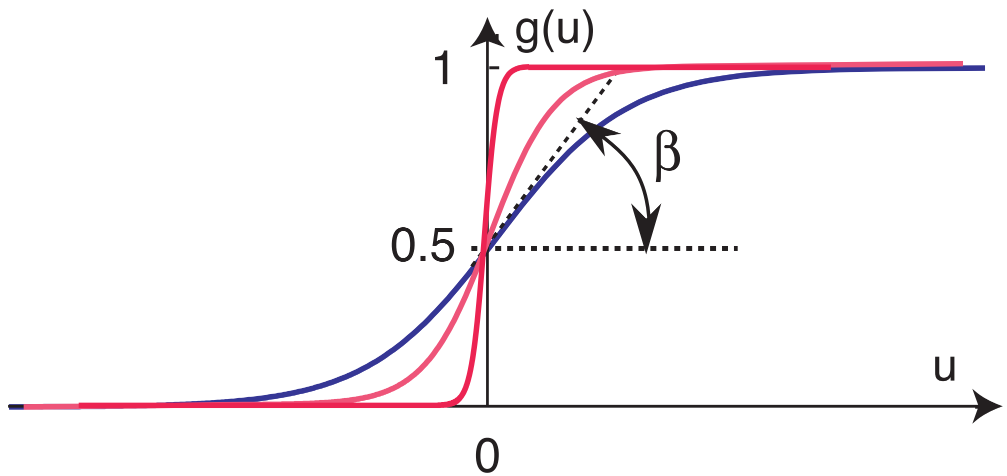
neural dynamics

Embodiment hypothesis

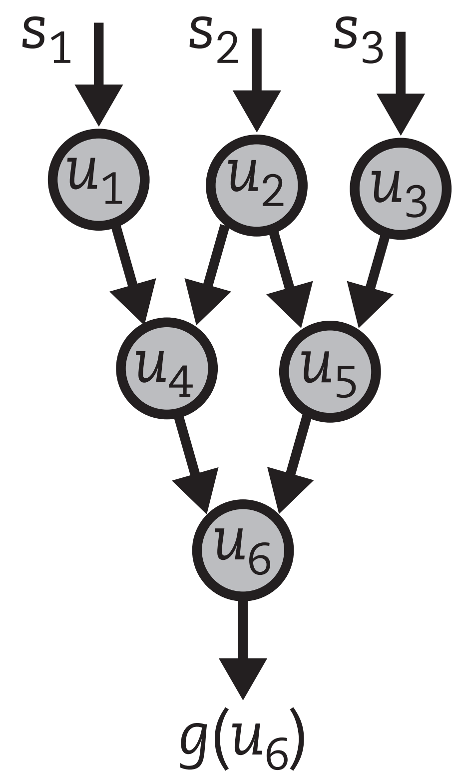
- all cognition is like soccer playing = has the properties of embodied cognition
- => embodied cognition reaches all forms of cognition, including "higher" symbolic reasoning

Forward neural networks form input-output mappings/functions

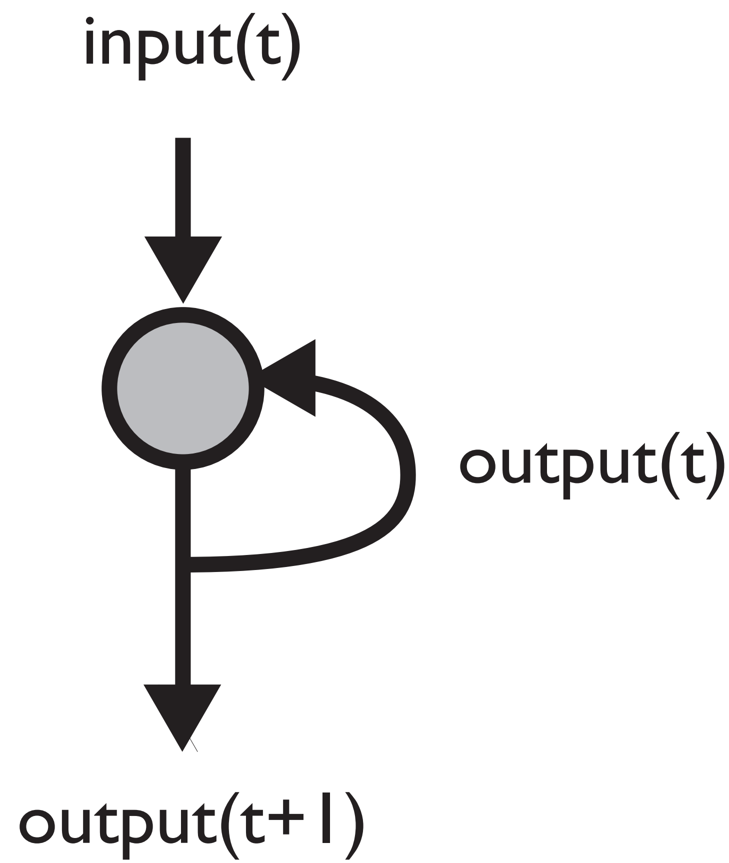




$$\text{output} = g\left(\sum(\text{inputs})\right)$$



Recurrent neural networks require concept of (continuous) time => neural dynamics



$$\dot{u}(t) = -u(t) + h + \text{input}(t) + g(u(t))$$

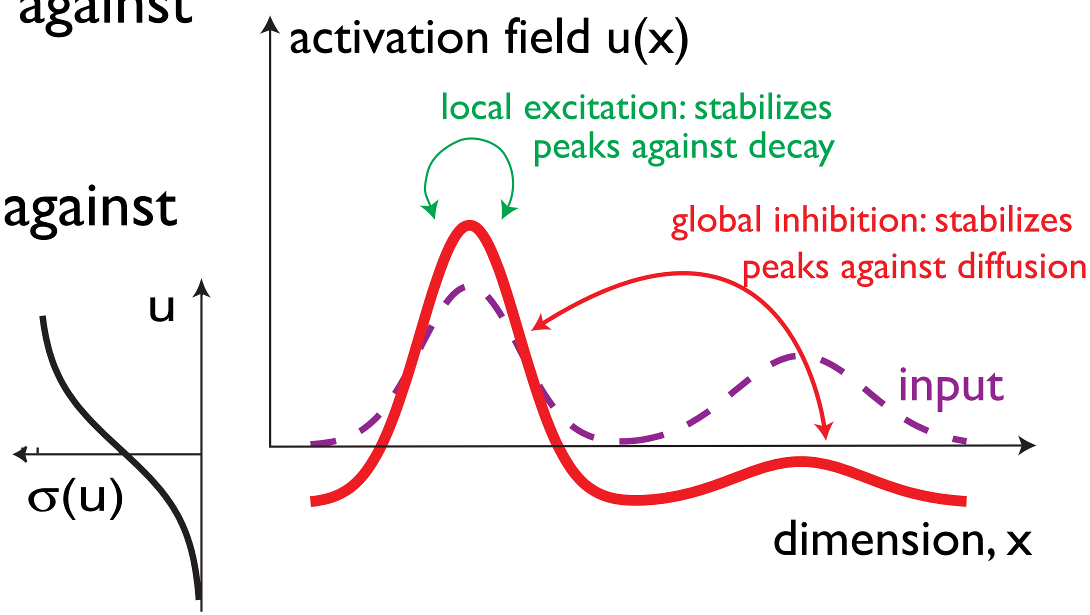
activation state, *u(t)*

Self-generated/self-stabilized activation patterns in neural dynamics

- key to cognition… !
- activation generated/sustained by recurrent connectivity (rather than by input)
- who recurrently excites who?
- => need for embedding in low-dimensional space

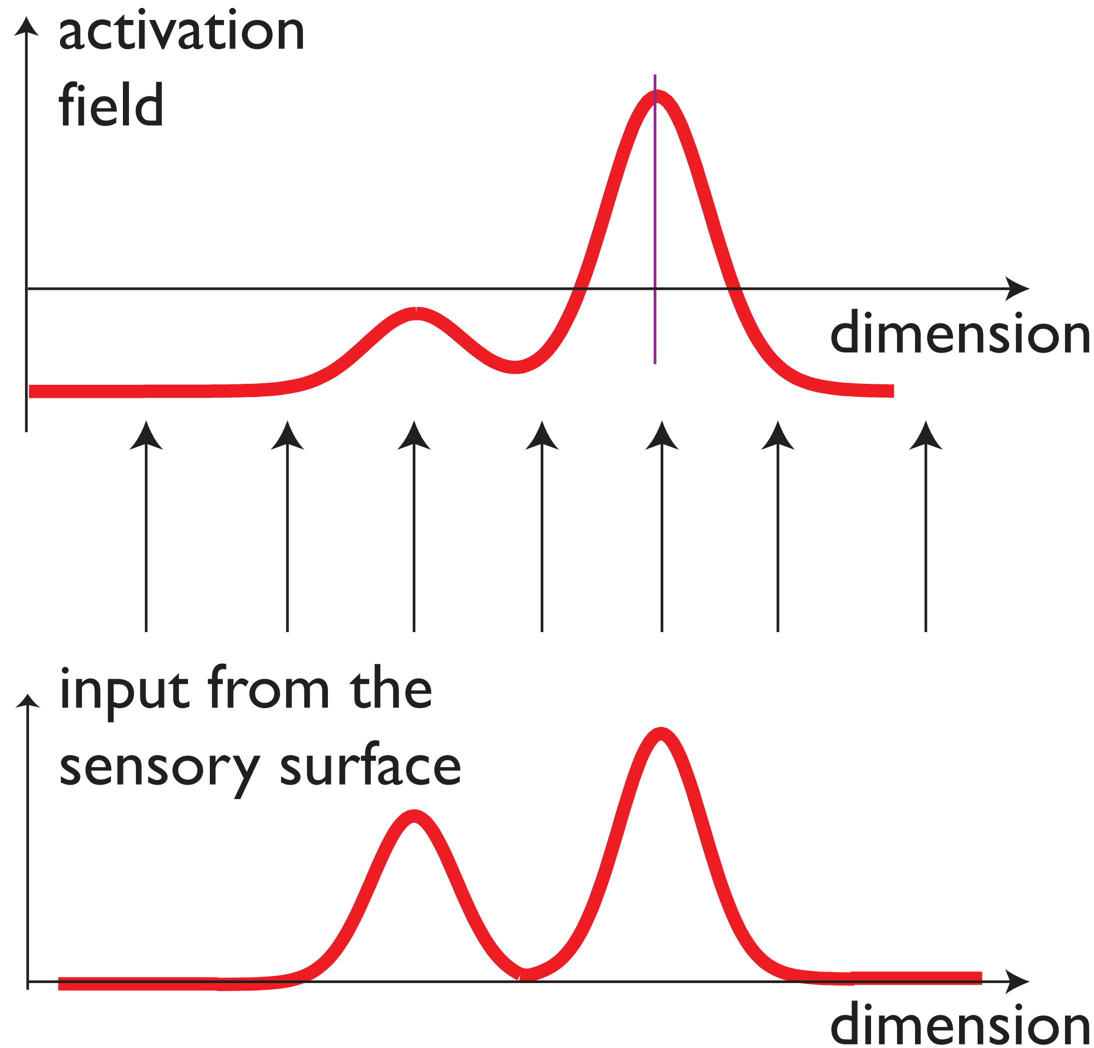
Dynamic neural fields

- localized activation patterns stabilized by regular pattern of recurrent connectivity
    - stabilized by excitatory coupling against decay
    - stabilized by inhibitory coupling against diffusive spread
- => attractors



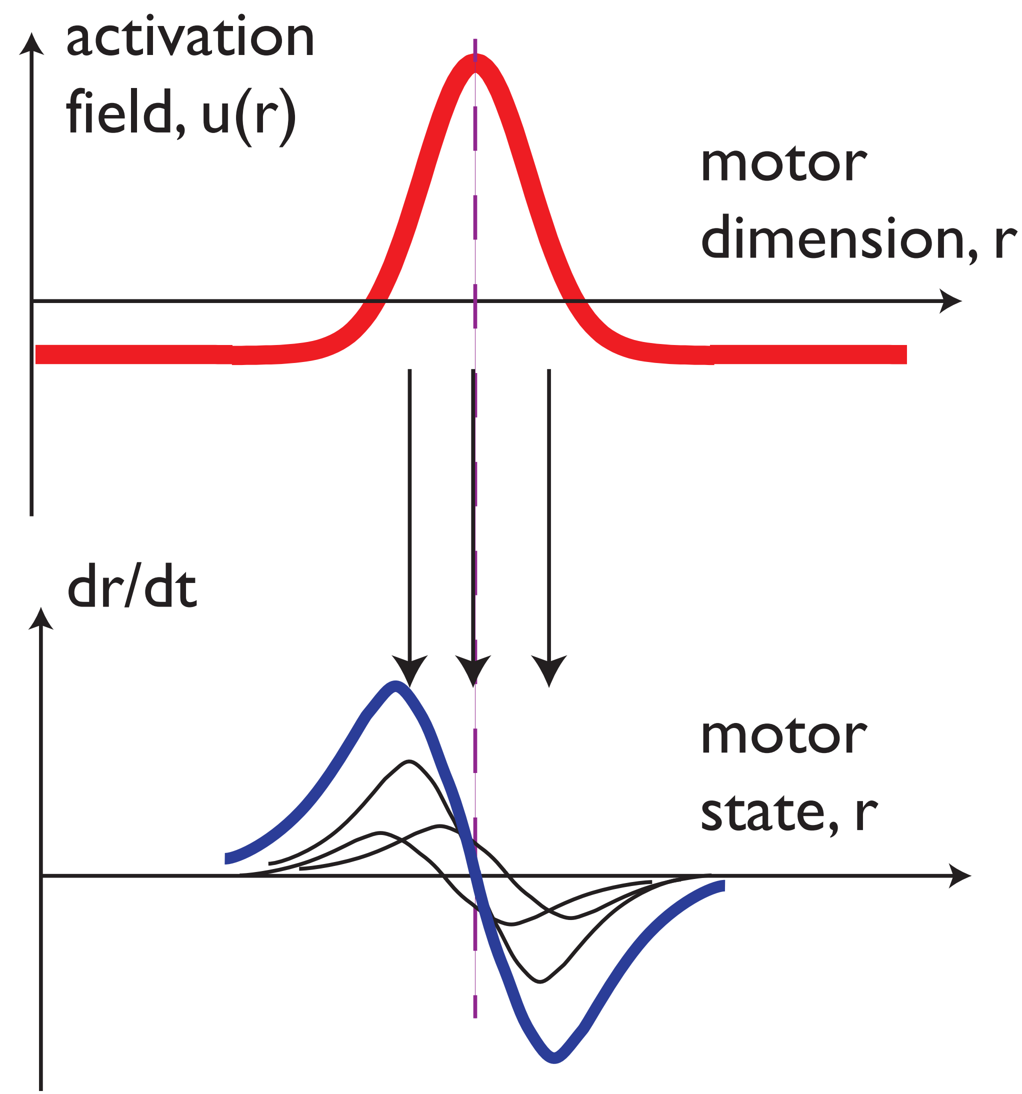
Embedding in low-dimensional space

- through forward connectivity from sensory surface
- e.g., feature maps…
- e.g. deep networks…



Embedding in low-dimensional space

- through forward projections onto motor surfaces…
- => behavioral dynamics
- e.g., through peripheral reflex loops



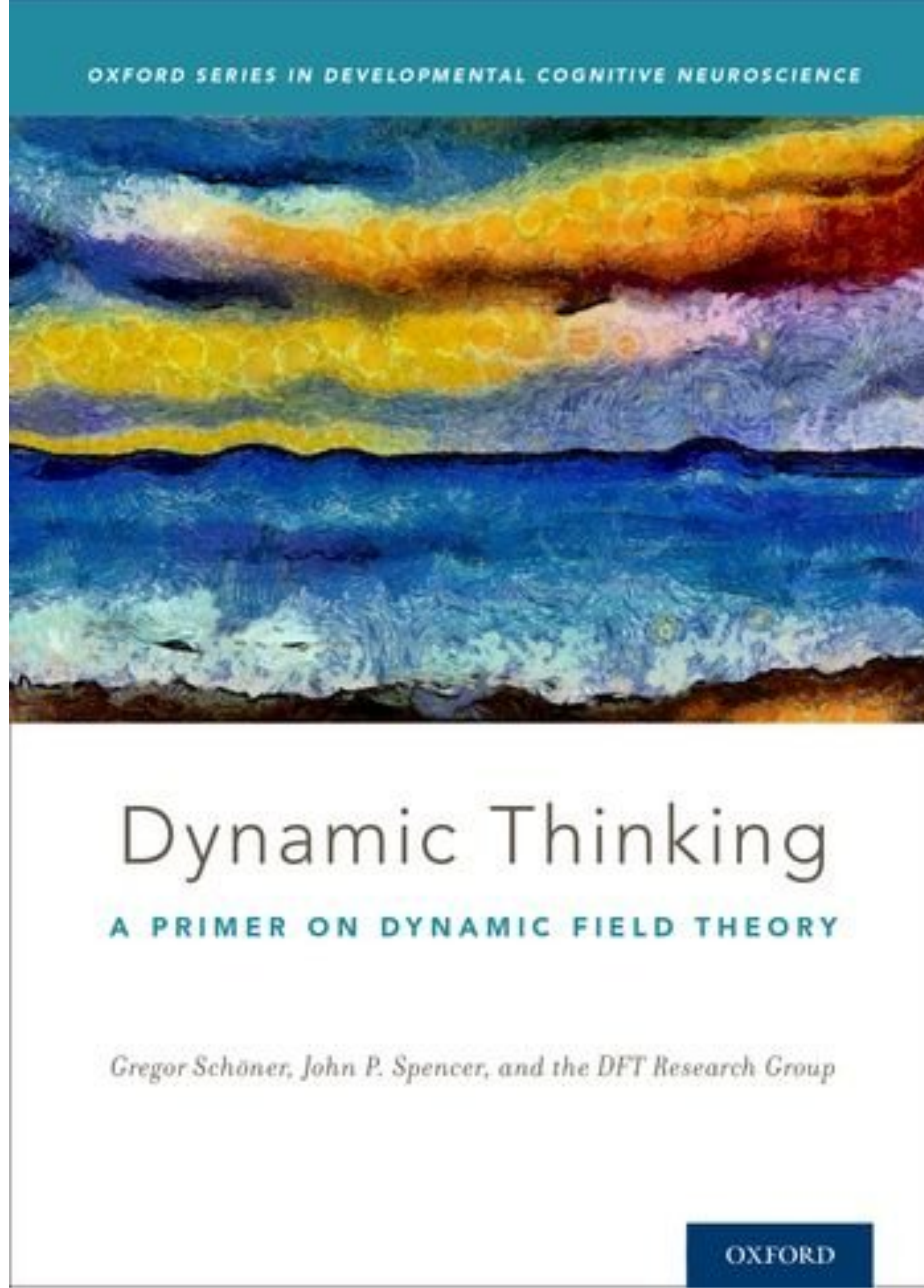
=> simulation

Dynamic Field Theory

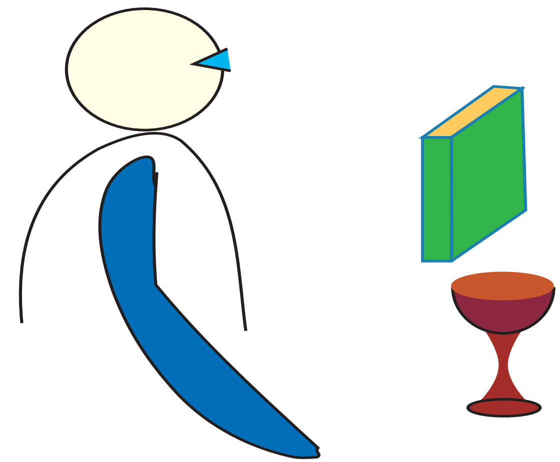
- attractor activation states
  - input driven solution (sub-threshold)
  - self-stabilized solution (peak, supra-threshold)
- instabilities
  - detection/reverse detection => events
  - selection => decision making
  - memory => WM
  - boost driven detection: => categories, switching

=> Dynamic Field Theory

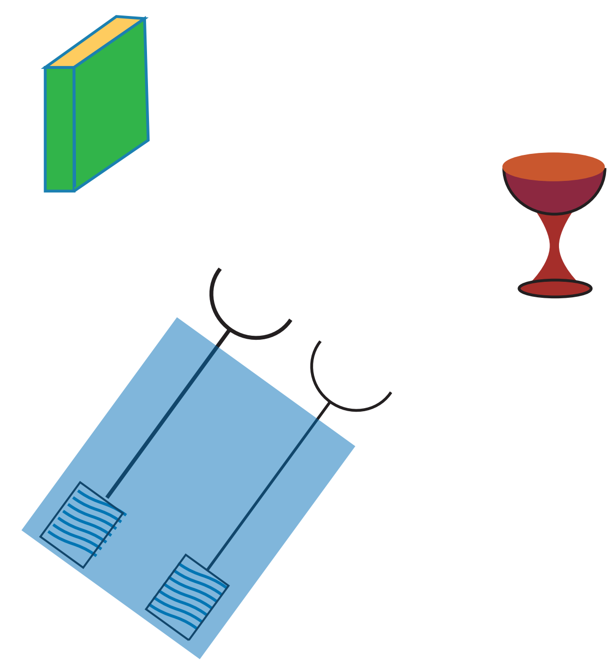
- “pervasively” neural processing accounts of behavior and cognition
- dynamicfieldtheory.org



Building embodied cognition

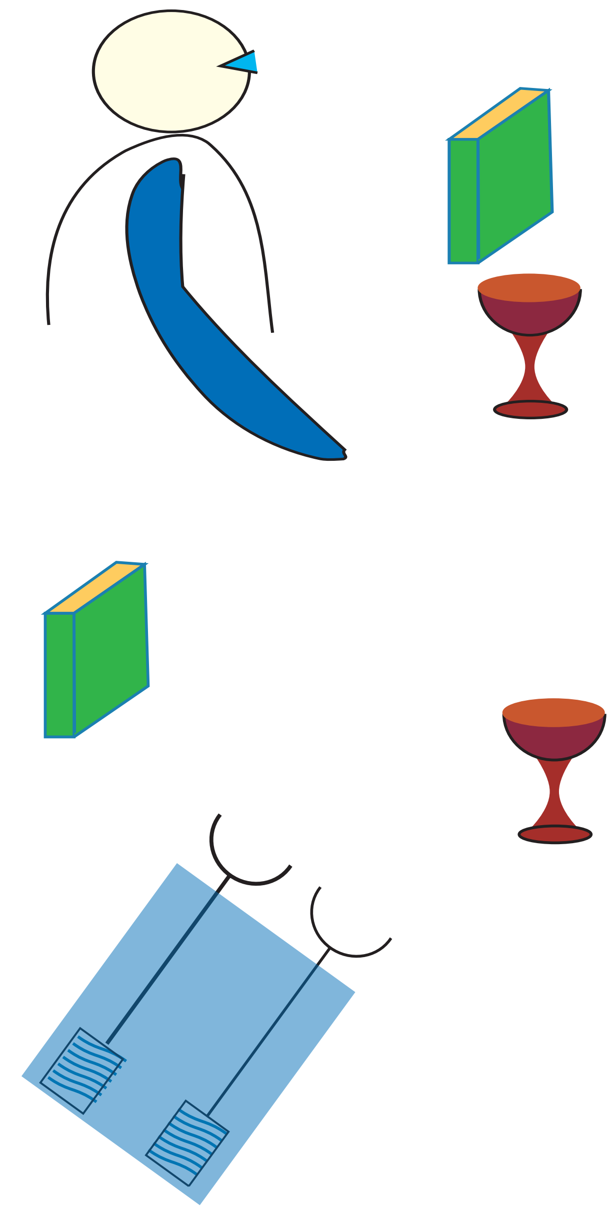


- two scenarios
  - “table-top”: directing action or thought at objects in a scene
  - “navigation”: generating actions while moving through a world



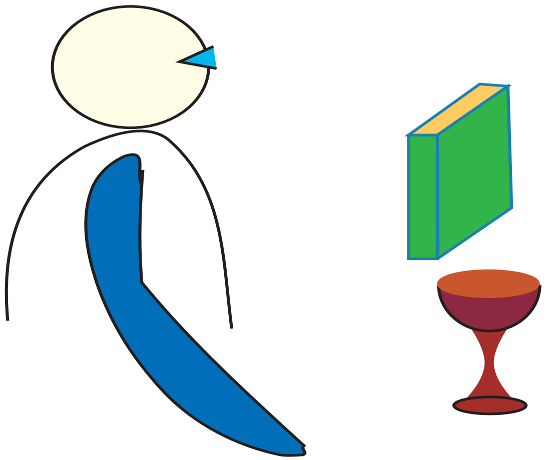
Building embodied cognition

- notions in both scenarios
  - scene representation
  - generating sequences
  - concepts
  - goals, knowledge, problem solving

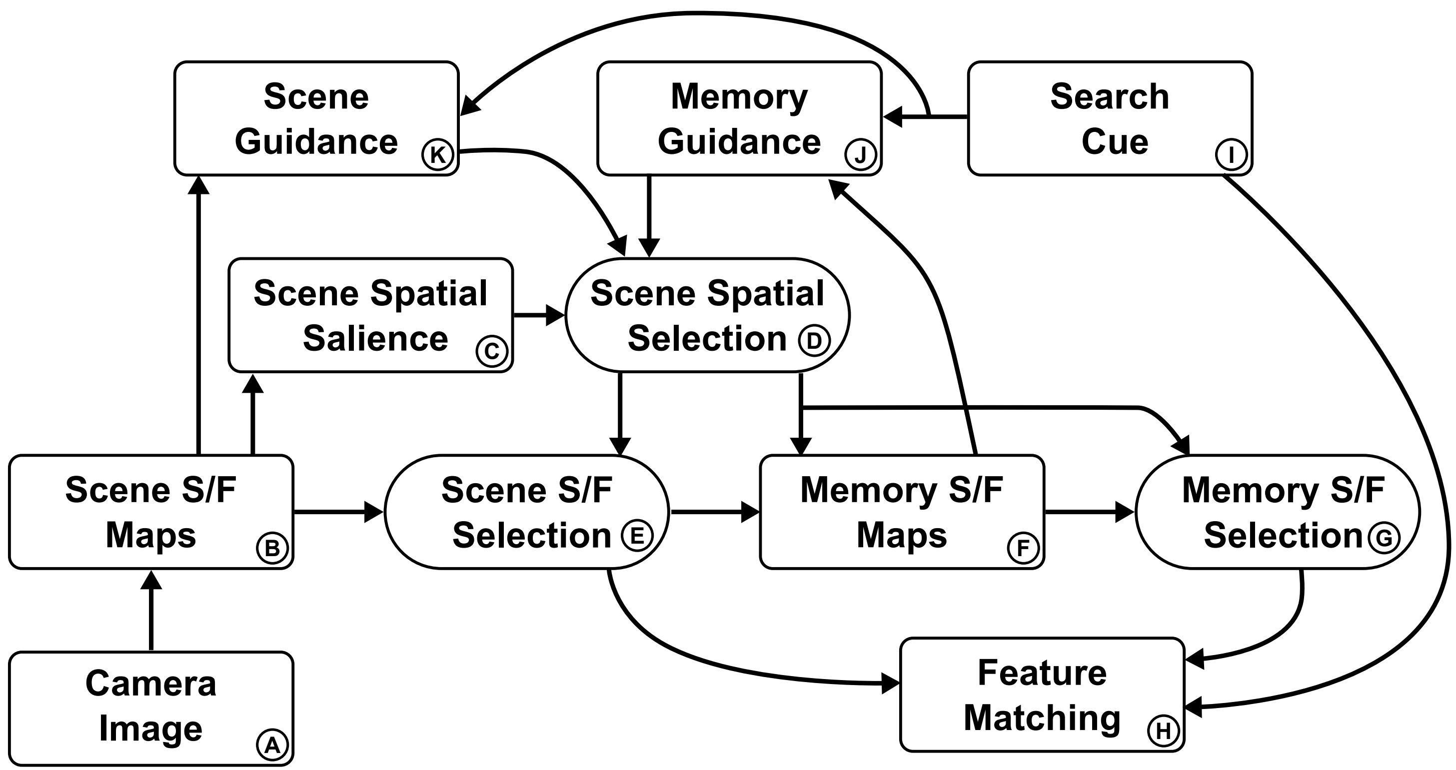


Scene representation

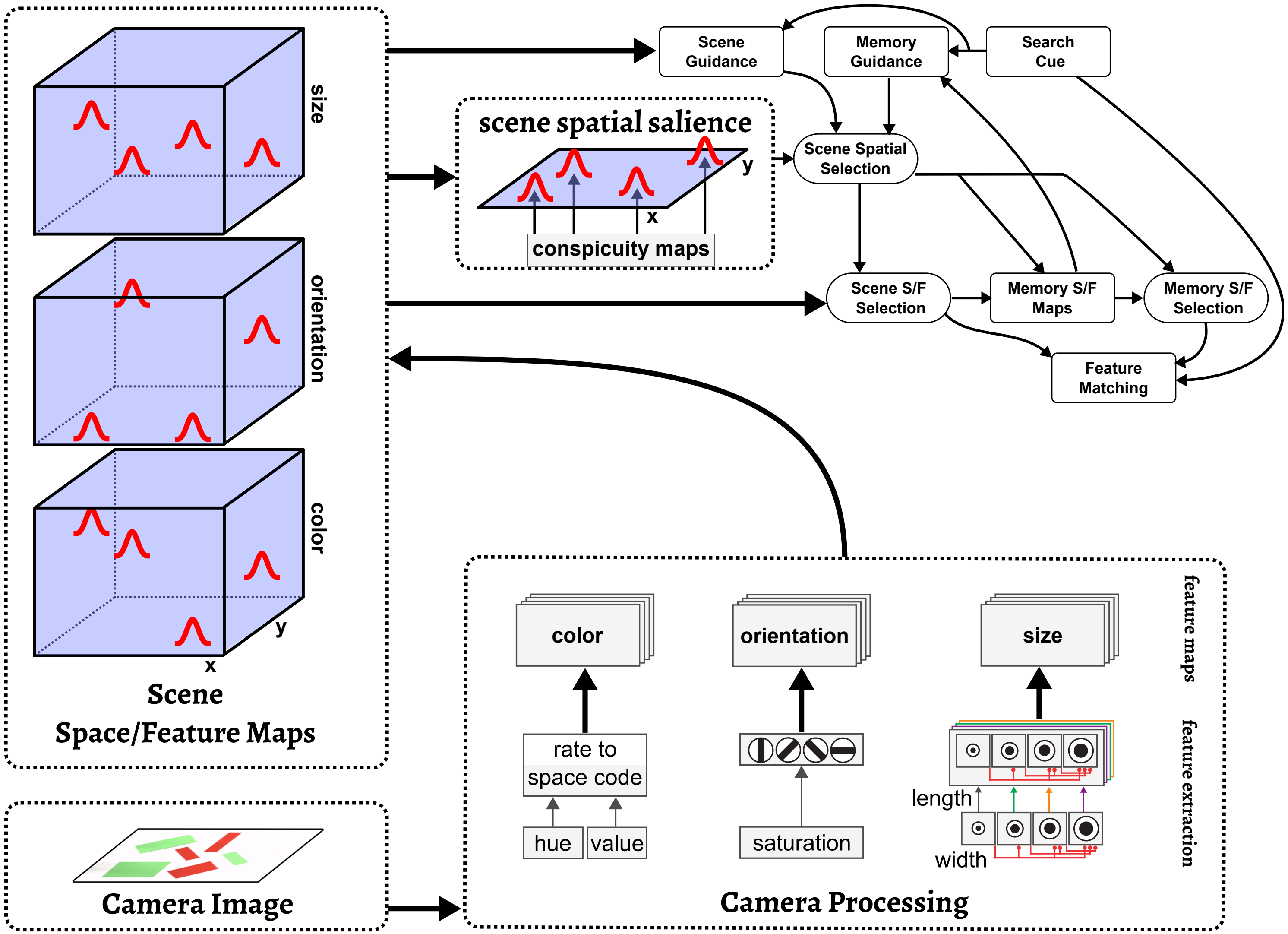
- any action directed at an object begins with bringing the object into the attentional foreground
  - visual search
  - attentional selection in scene memory
  - (selection within mental maps)



Neural dynamic architecture of scene memory and visual search



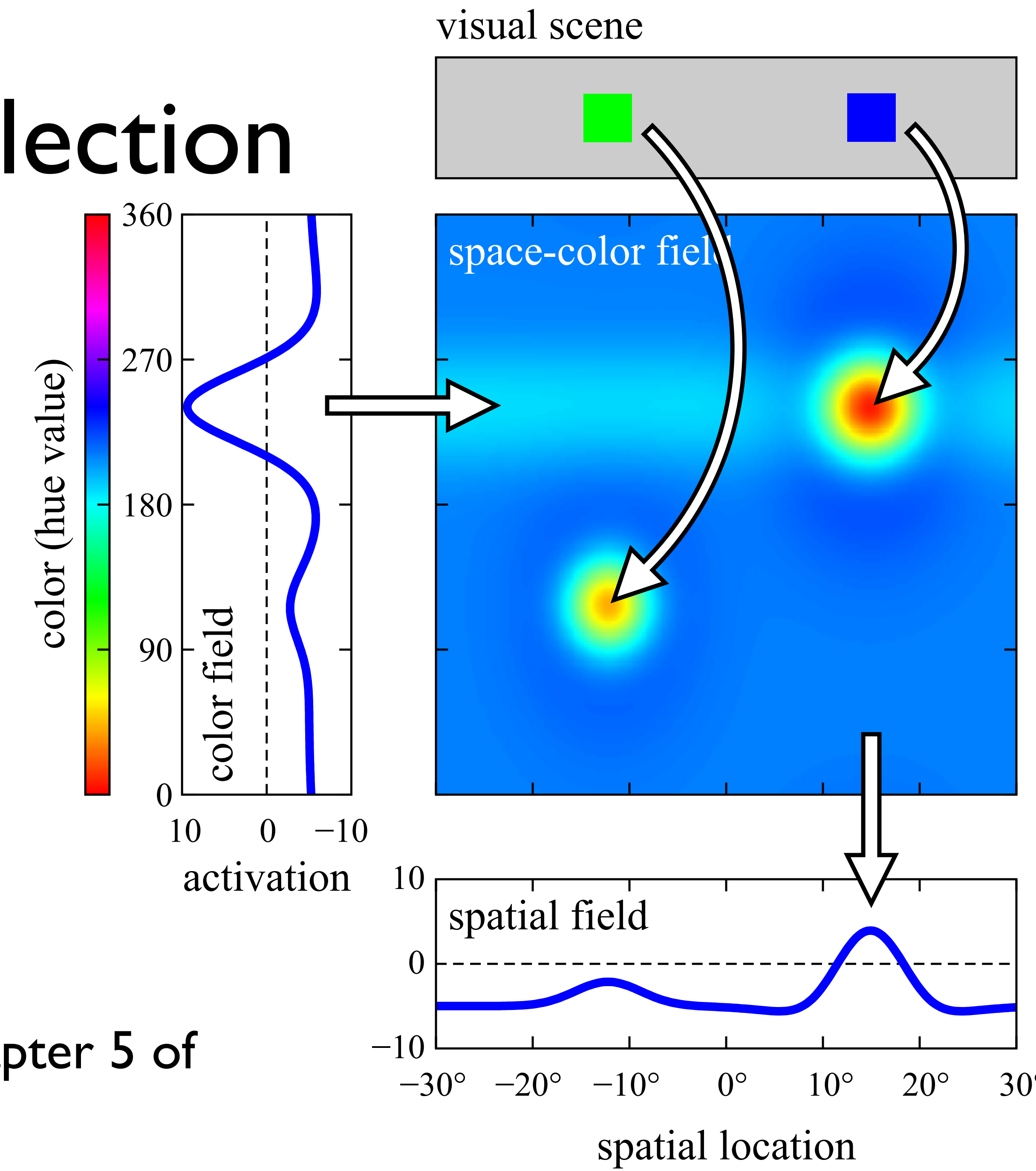
[Grieben et al, Attention, Perception & Psychophysics, in press]



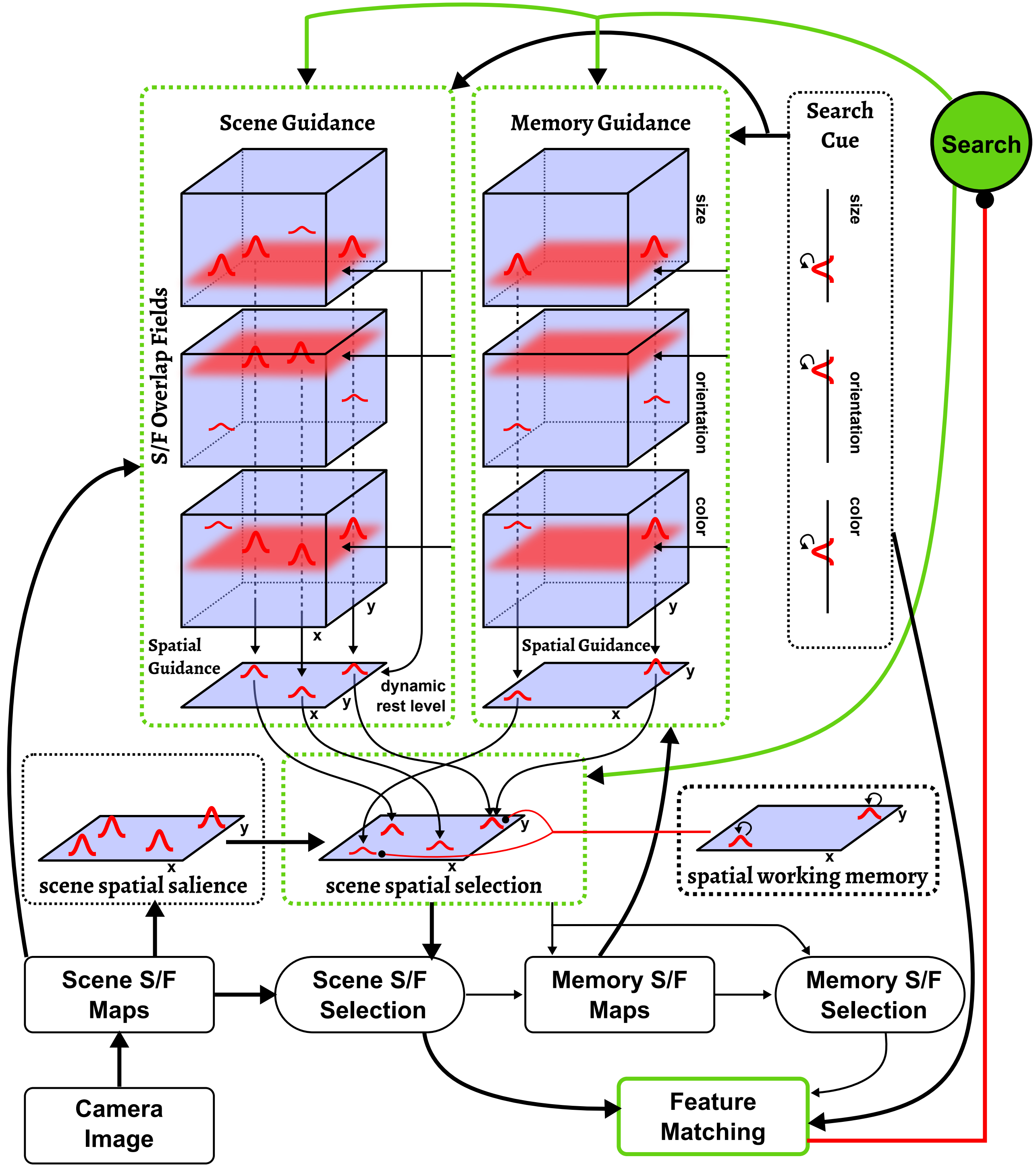
[Grieben et al, Attention, Perception & Psychophysics, in press]

Attentional selection

- boost driven detection instability

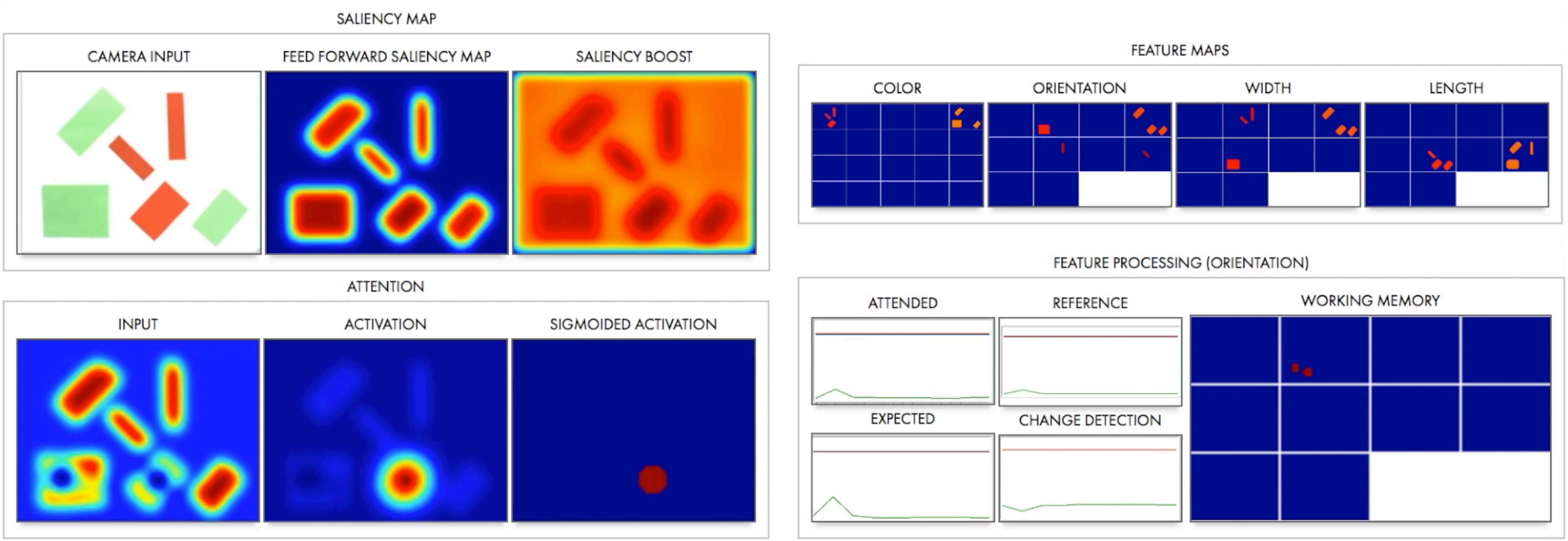


[Schneegans, Lins, Spencer, Chapter 5 of DFT Primer, 2016]

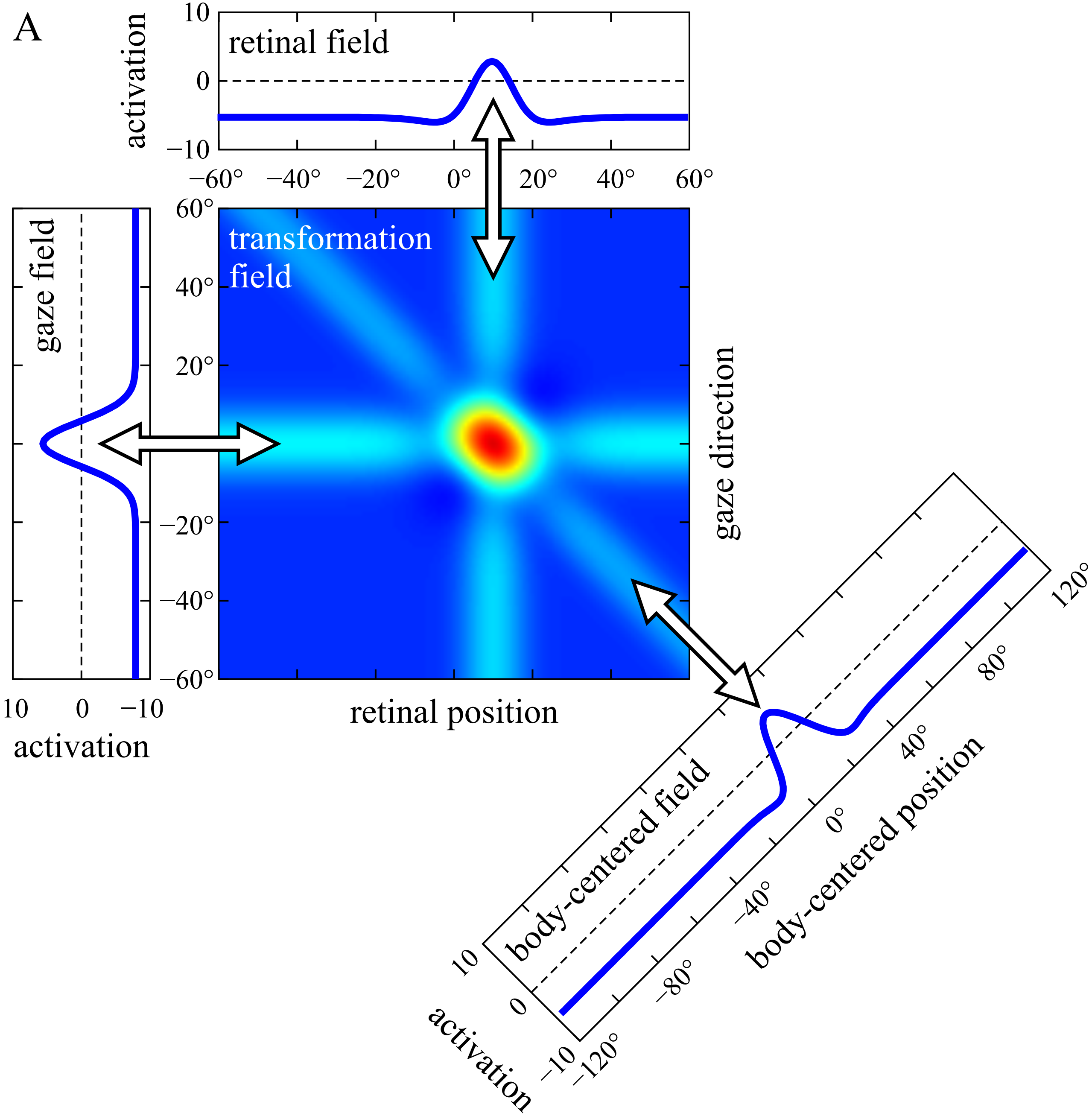


[Grieben et al, Attention, Perception & Psychophysics, in press]

Autonomous sequences of visual exploration and cued visual search



Coordinate transformatio



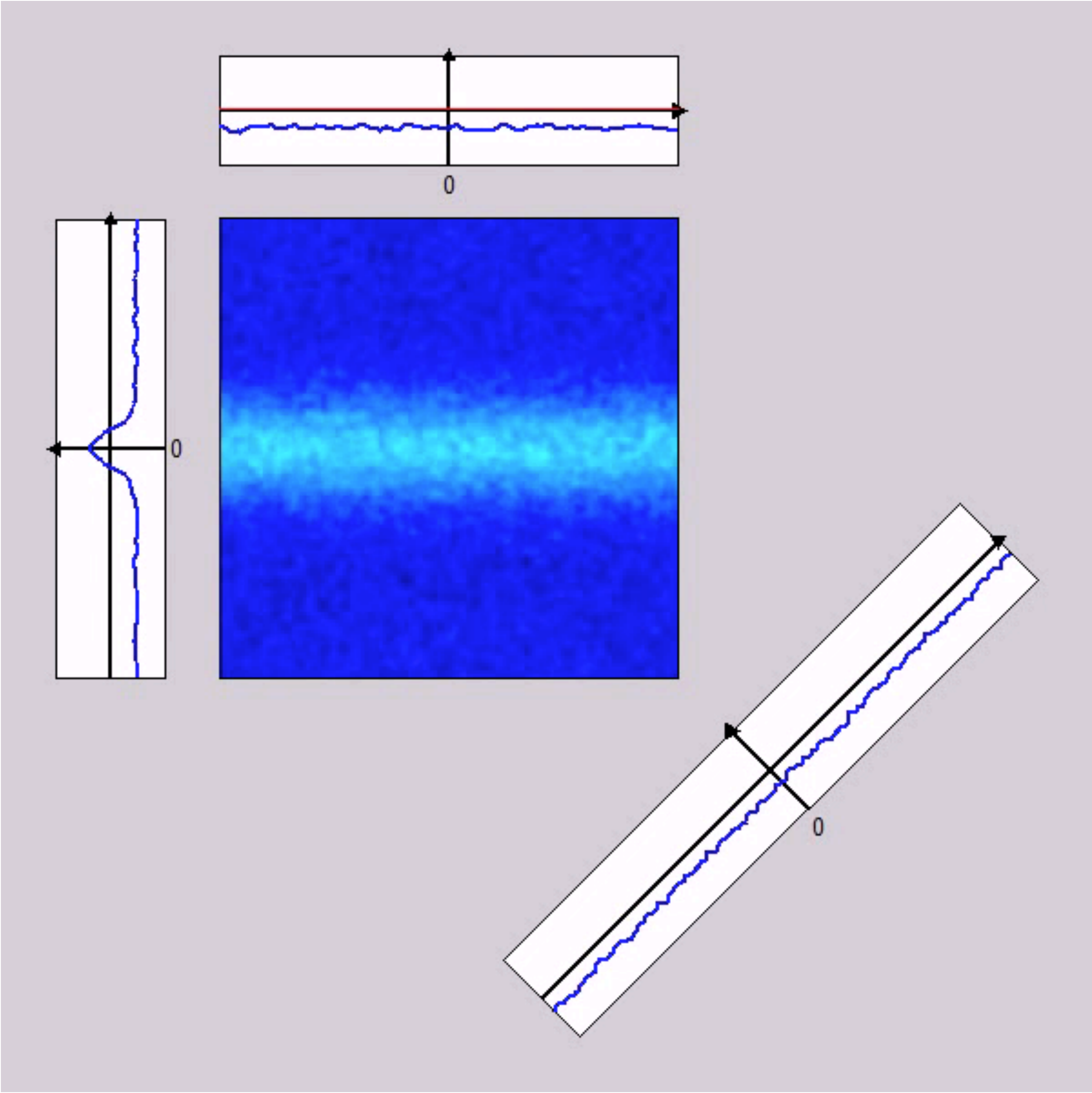
- are another function of multi-dimensional activation fields

[Schneegans, Chapter 7 of DFT Primer, 2016]

Coordinate transformations

- are another function of multi-dimensional activation fields

[Schneegans, Chapter 7 of DFT Primer, 2016]

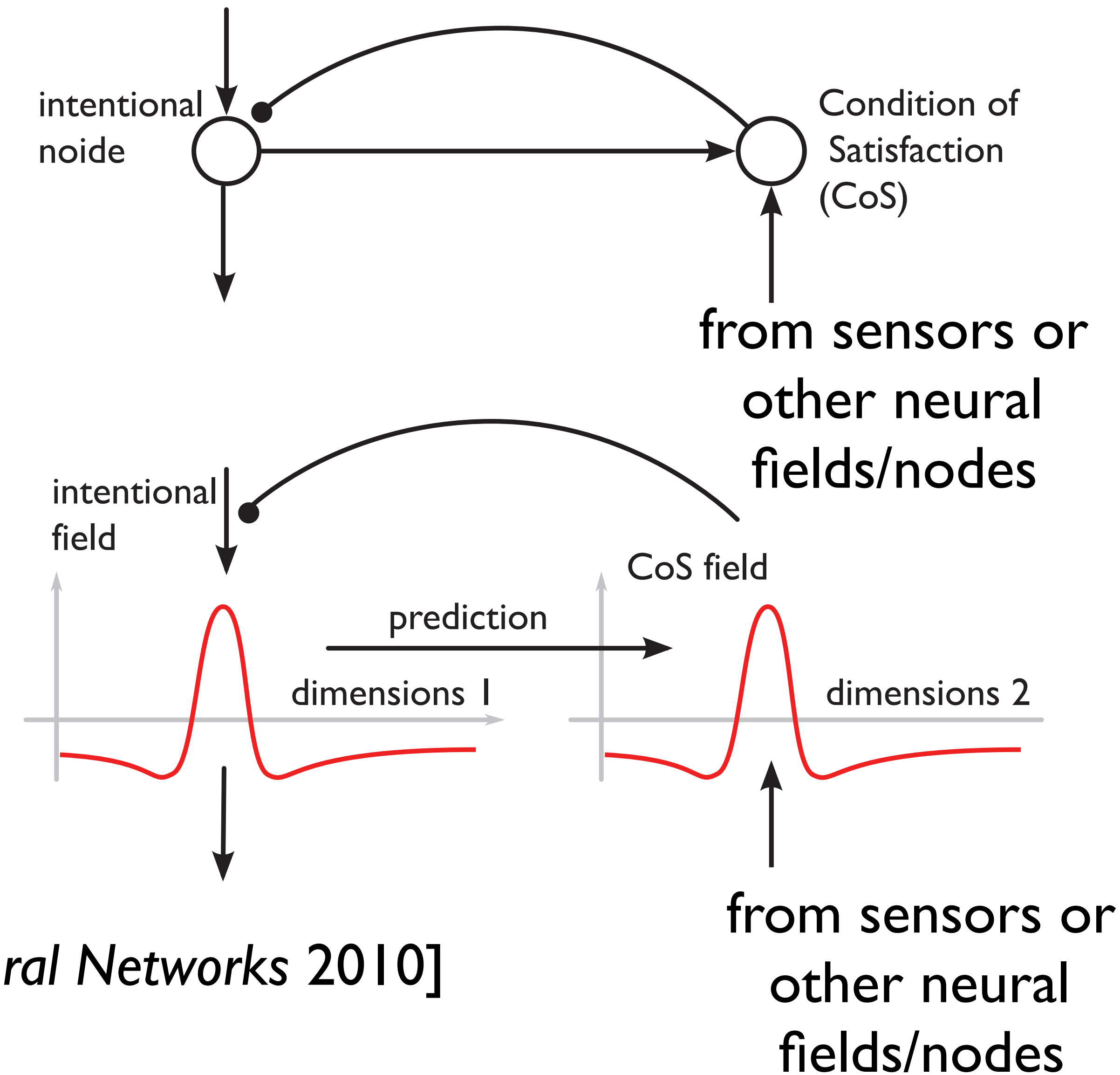


Sequence generation

- Actions and thoughts consist of sequences of mental states that unfold autonomously
  - not necessarily “triggered” by external inputs

Sequences in neural dynamics

- intentional states predict/pre-activate their CoS=Condition of satisfaction
- match of prediction activates CoS and inhibits the intentional state

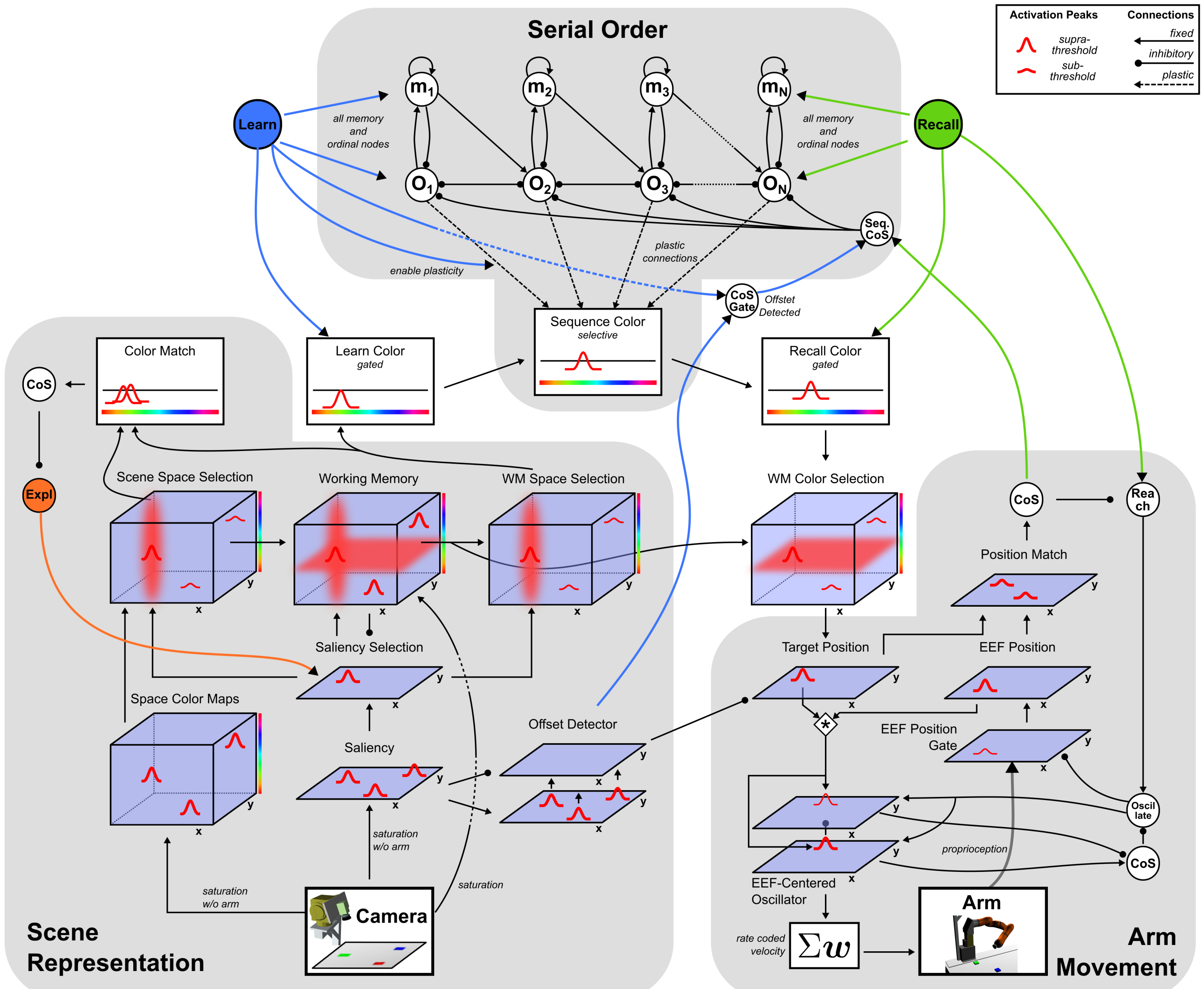


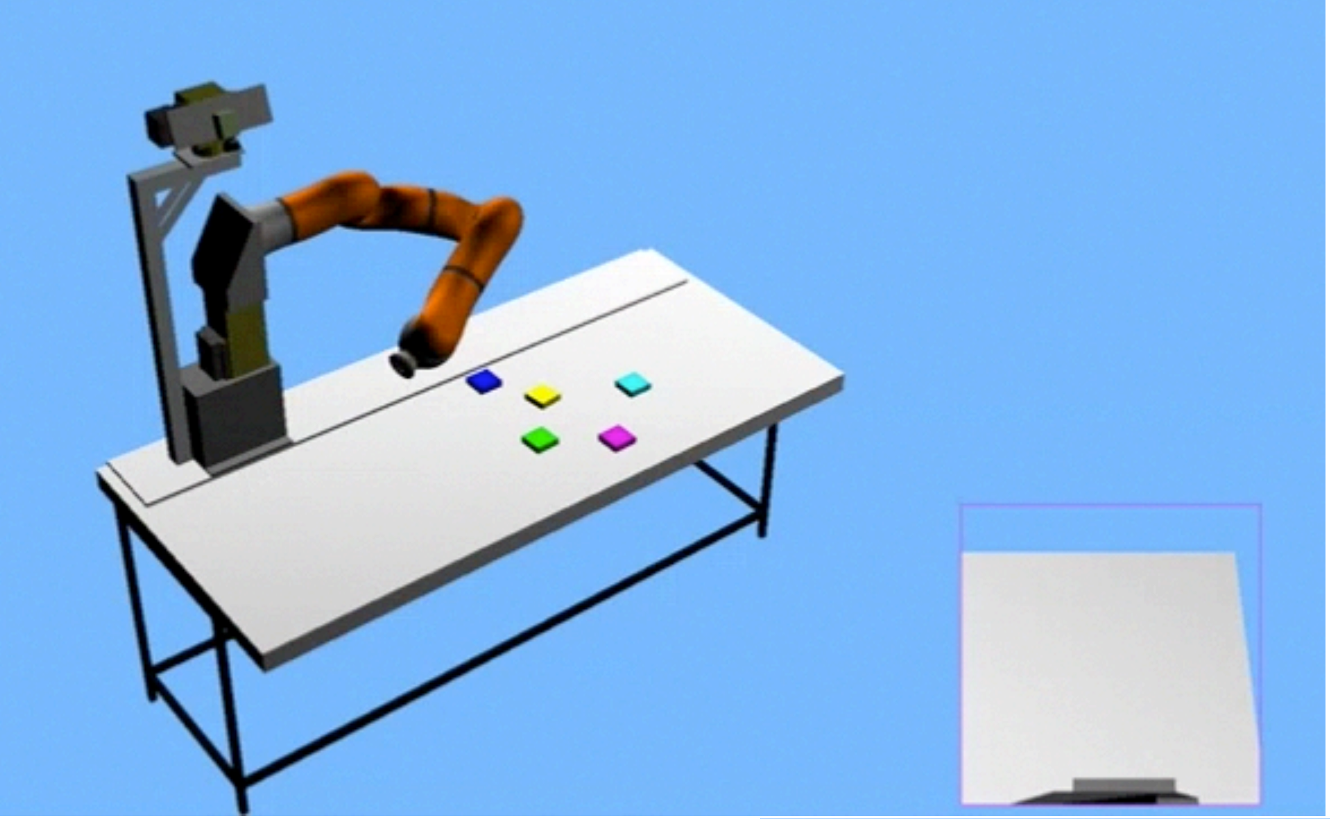
[Sandamirskaya, Schöner, *Neural Networks* 2010]

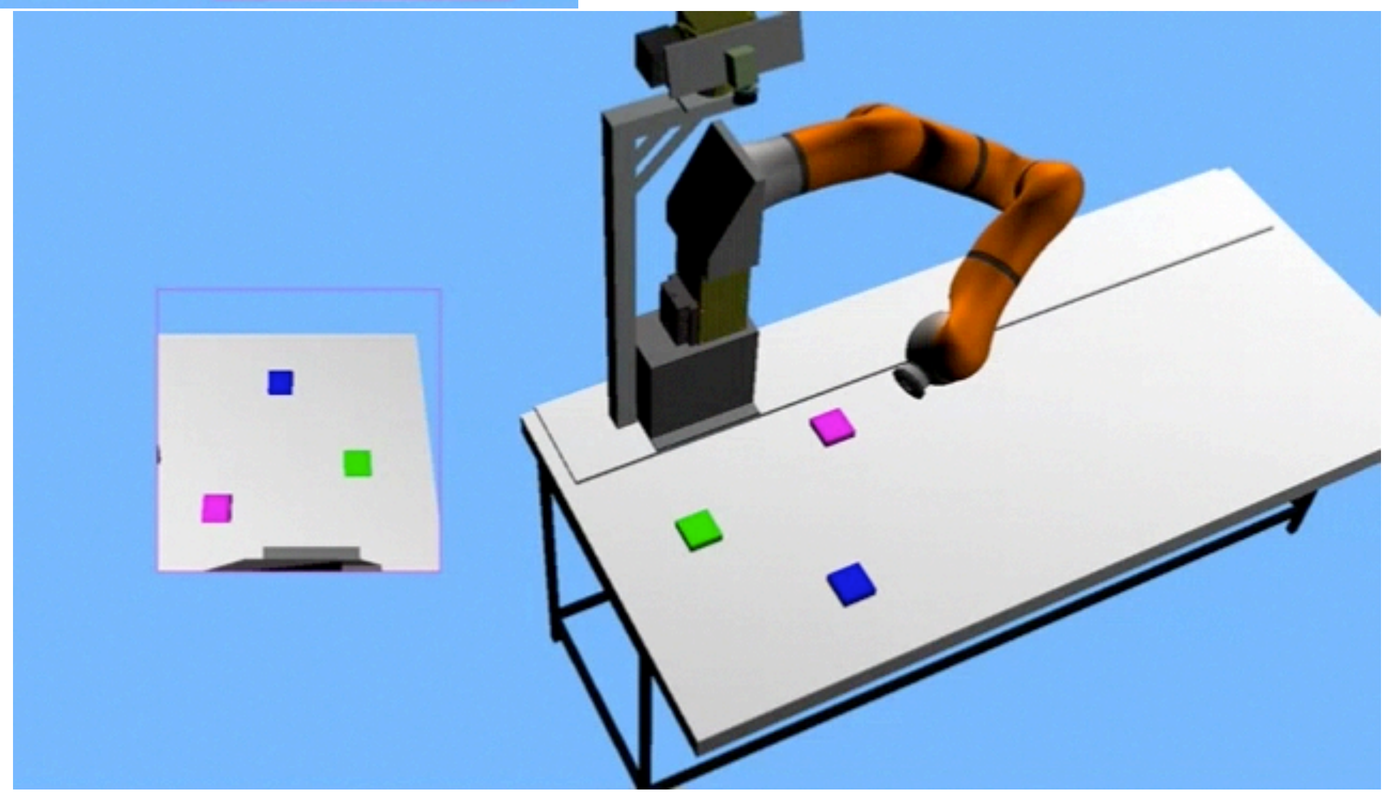
Sequence generation directed at objects

- serial order from demonstration
- => sequence of pointing movements

[Tekülve et al., *Frontiers in Neurorobotics* (in press)]





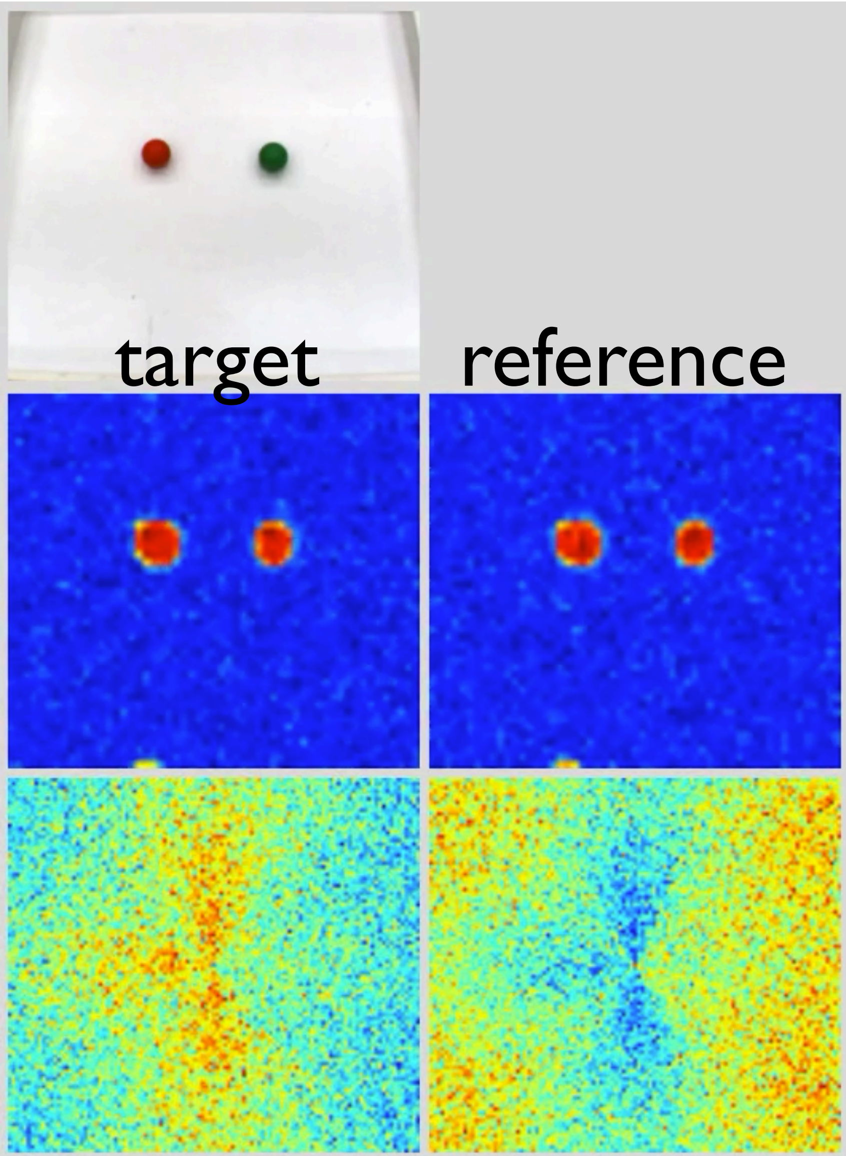


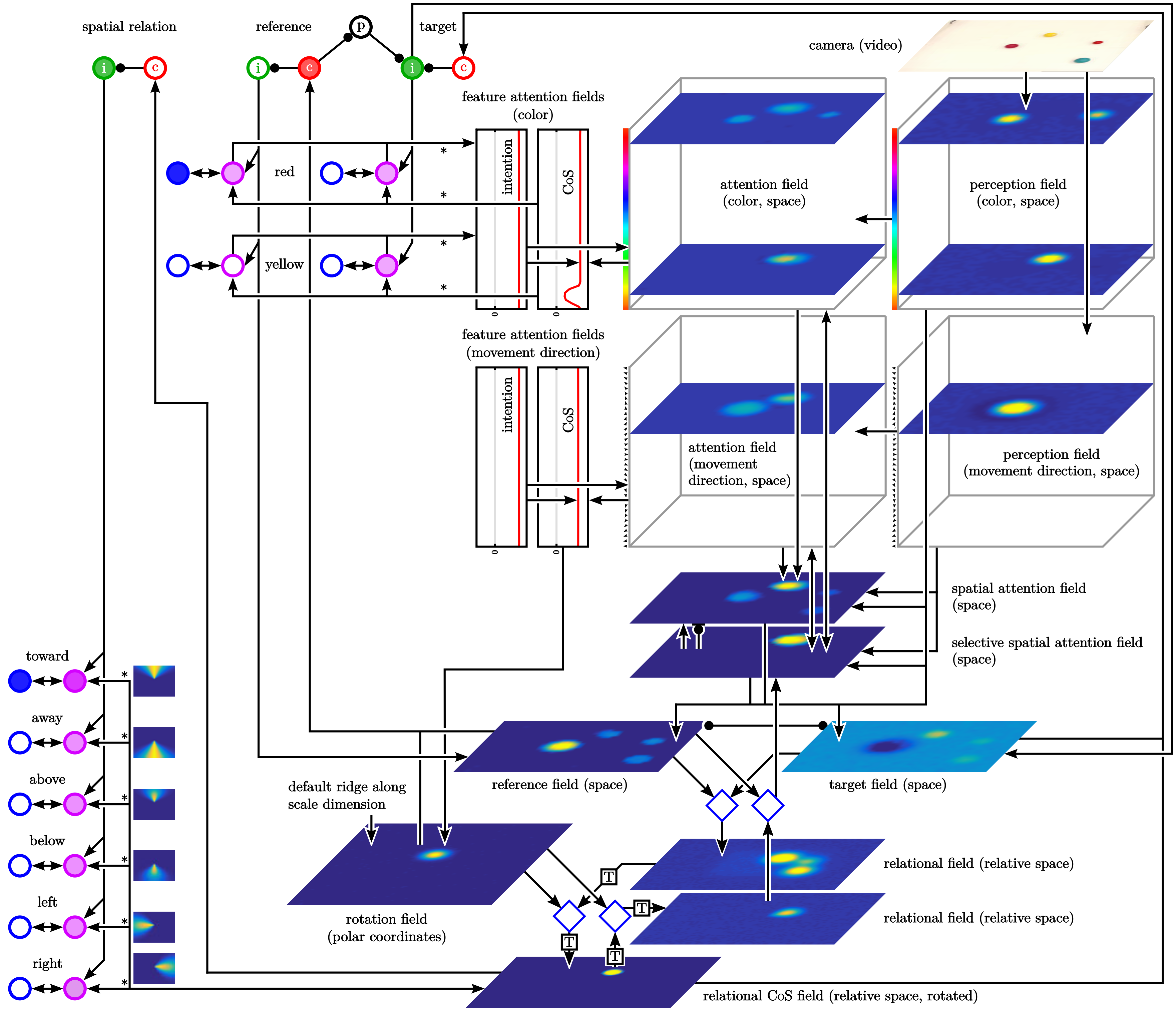
Concepts, relational thinking

- Talking about objects entails bringing the targeted object into the attentional foreground

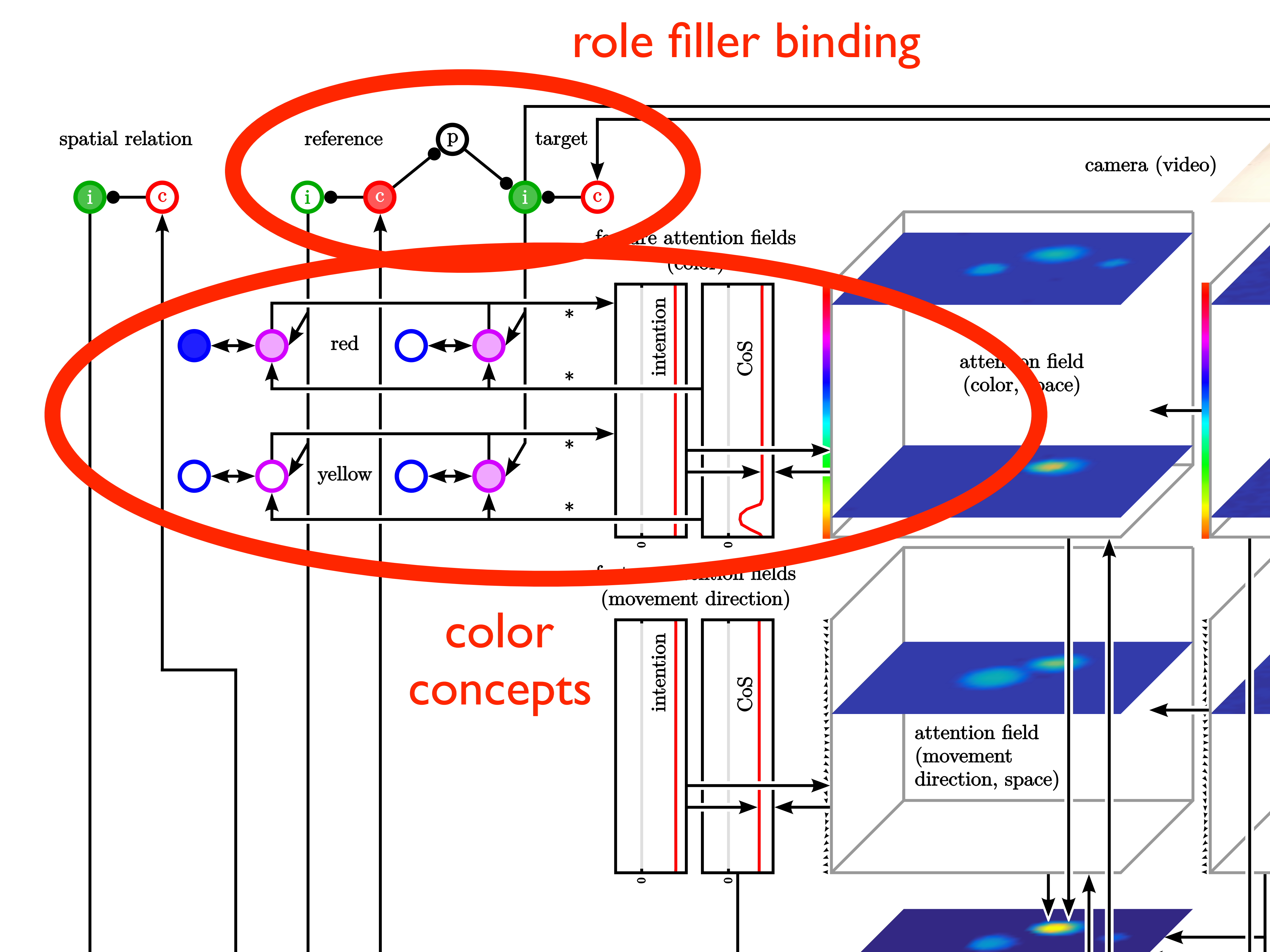
[Lipinski, Sandamirskaya, Schöner 2009
… Richter, Lins, Schöner, *Topics* 2017]

"red to the left of green"



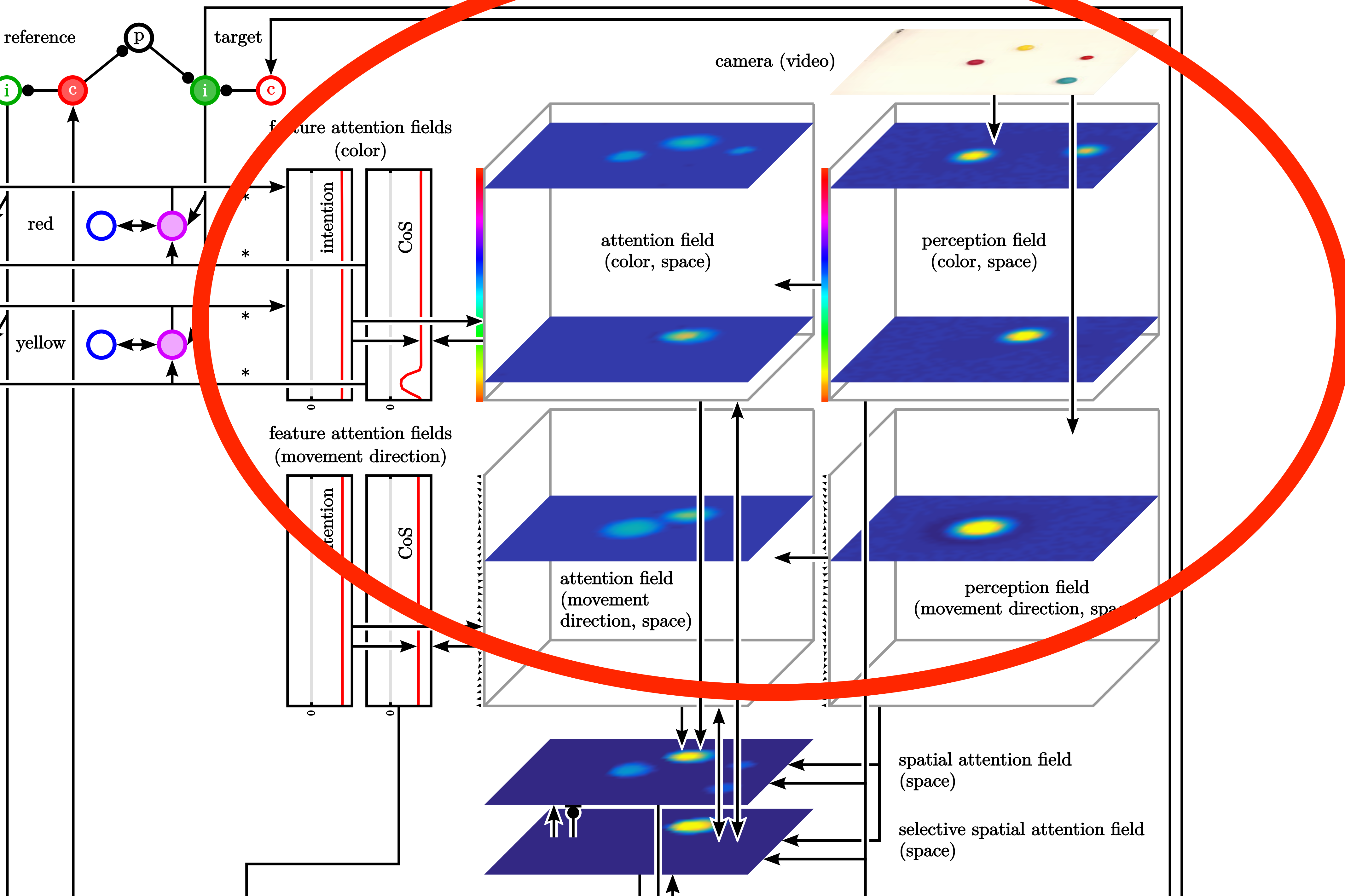


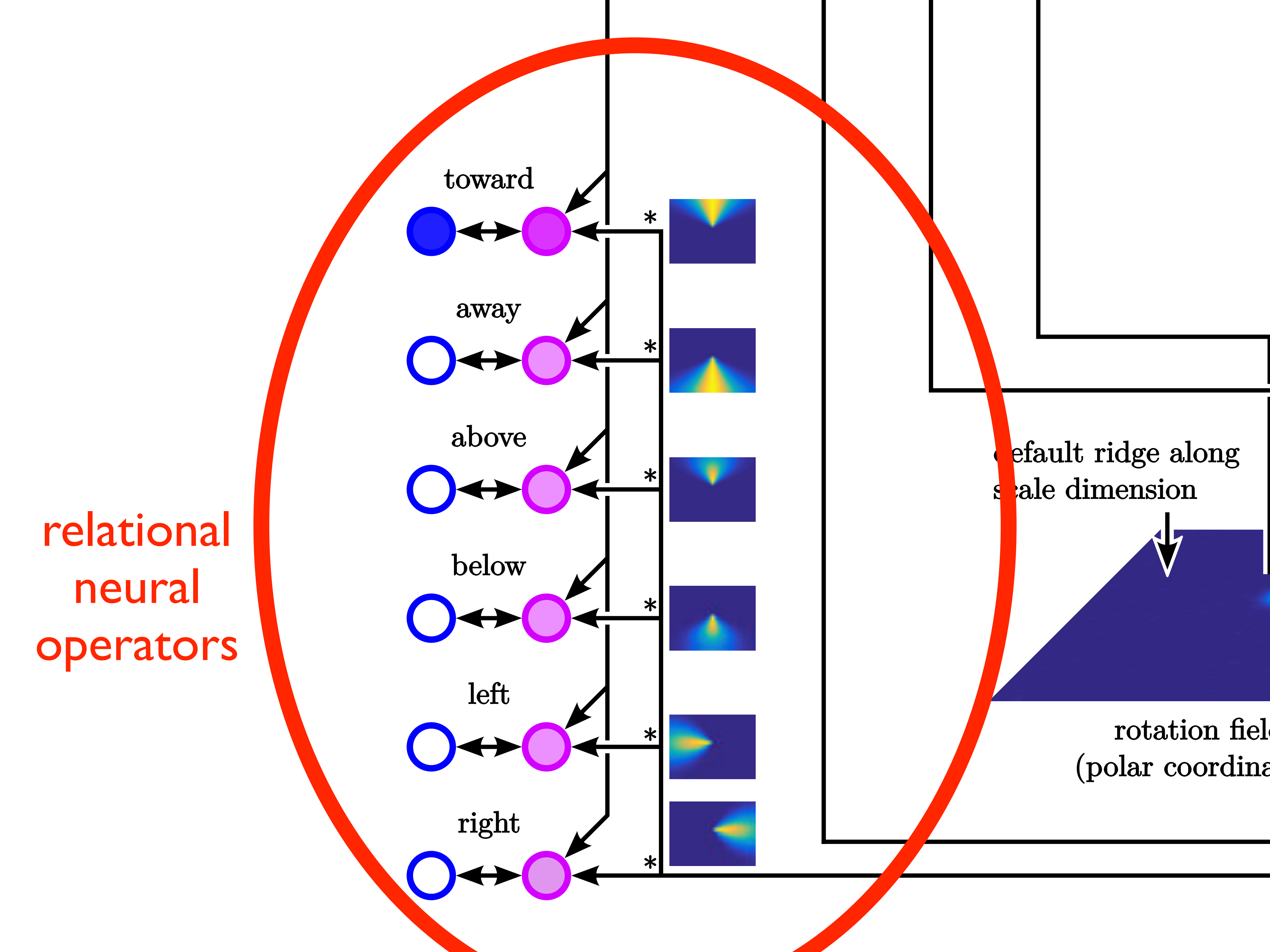
[Richter, Lins, Schöner, ToPiC (2017)]

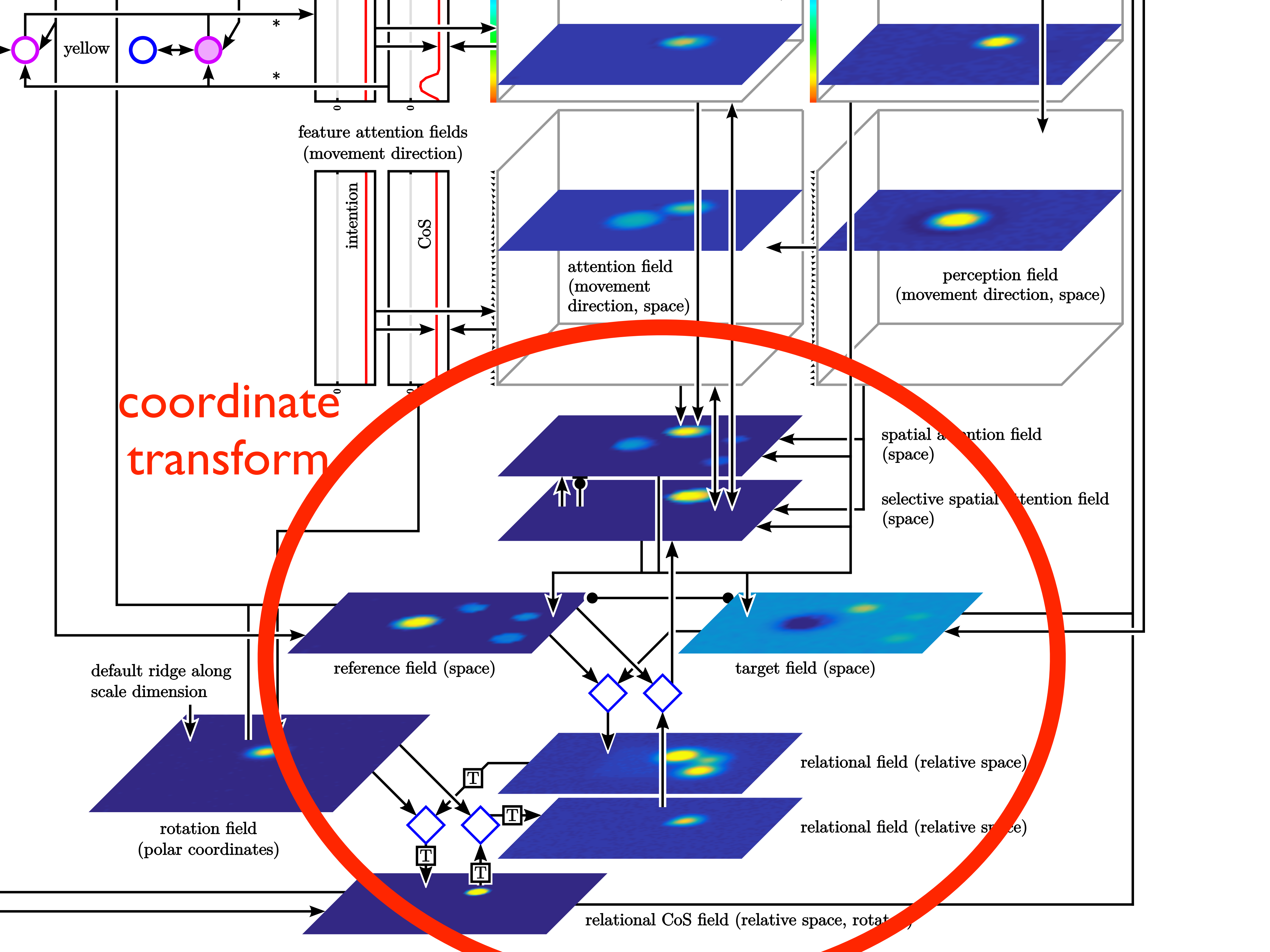


role filler binding
spatial relation
reference
target
p
i
c
camera (video)
red
yellow
intention
CoS
(movement direction)
color concepts
attention field (movement direction, space)

cued visual search

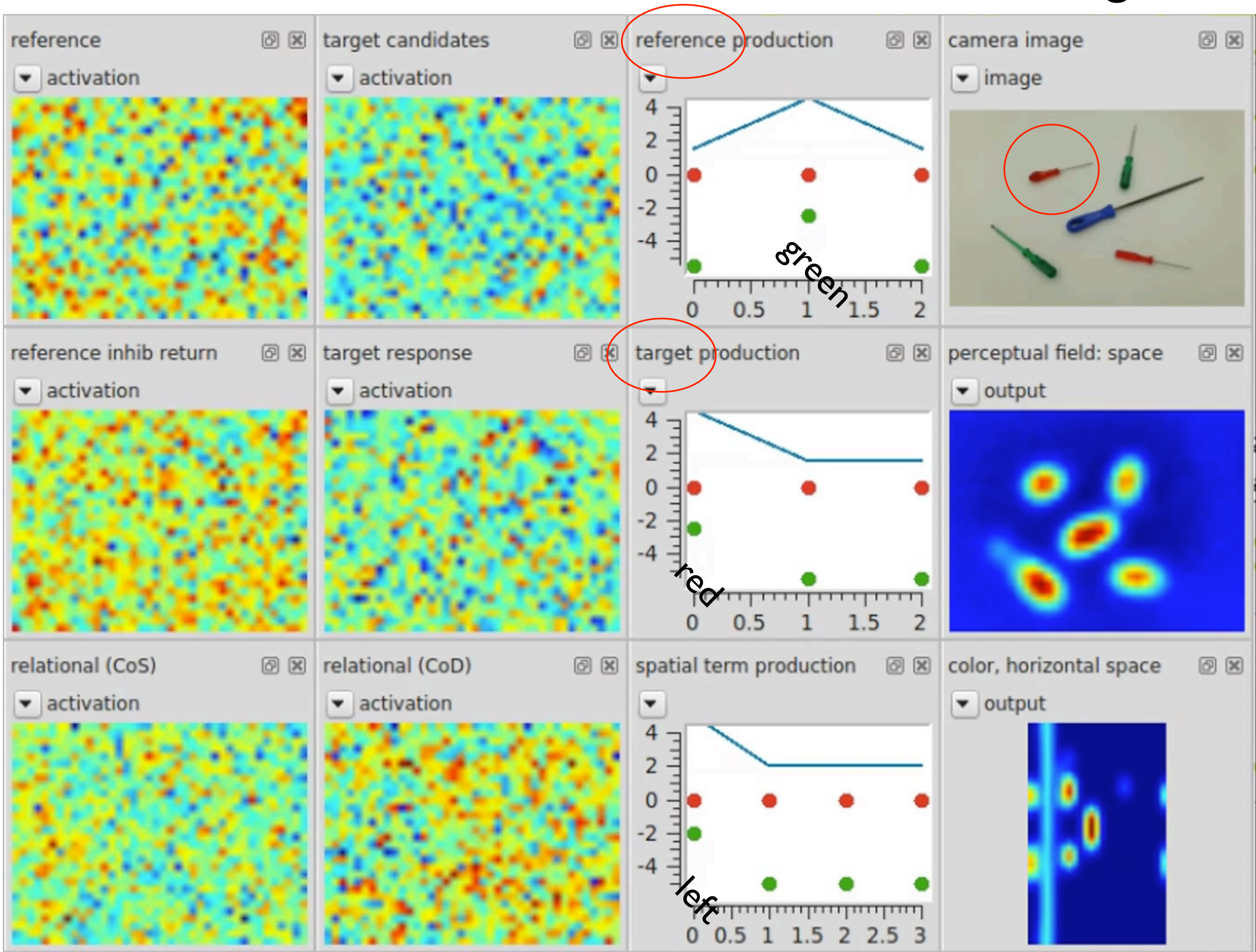



relational
neural
operators
toward
away
above
below
left
right
default ridge along
scale dimension
rotation field
(polar coordina



yellow
feature attention fields
(movement direction)
intention
CoS
attention field
(movement
direction, space)
perception field
(movement direction, space)
coordinate
transform
spatial a ntion field
(space)
selective spatial tention field
(space)
reference field (space)
target field (space)
default ridge along
scale dimension
relational field (relative space)
relational field (relative sp ce)
rotation field
(polar coordinates)
T
relational CoS field (relative space, rotat)

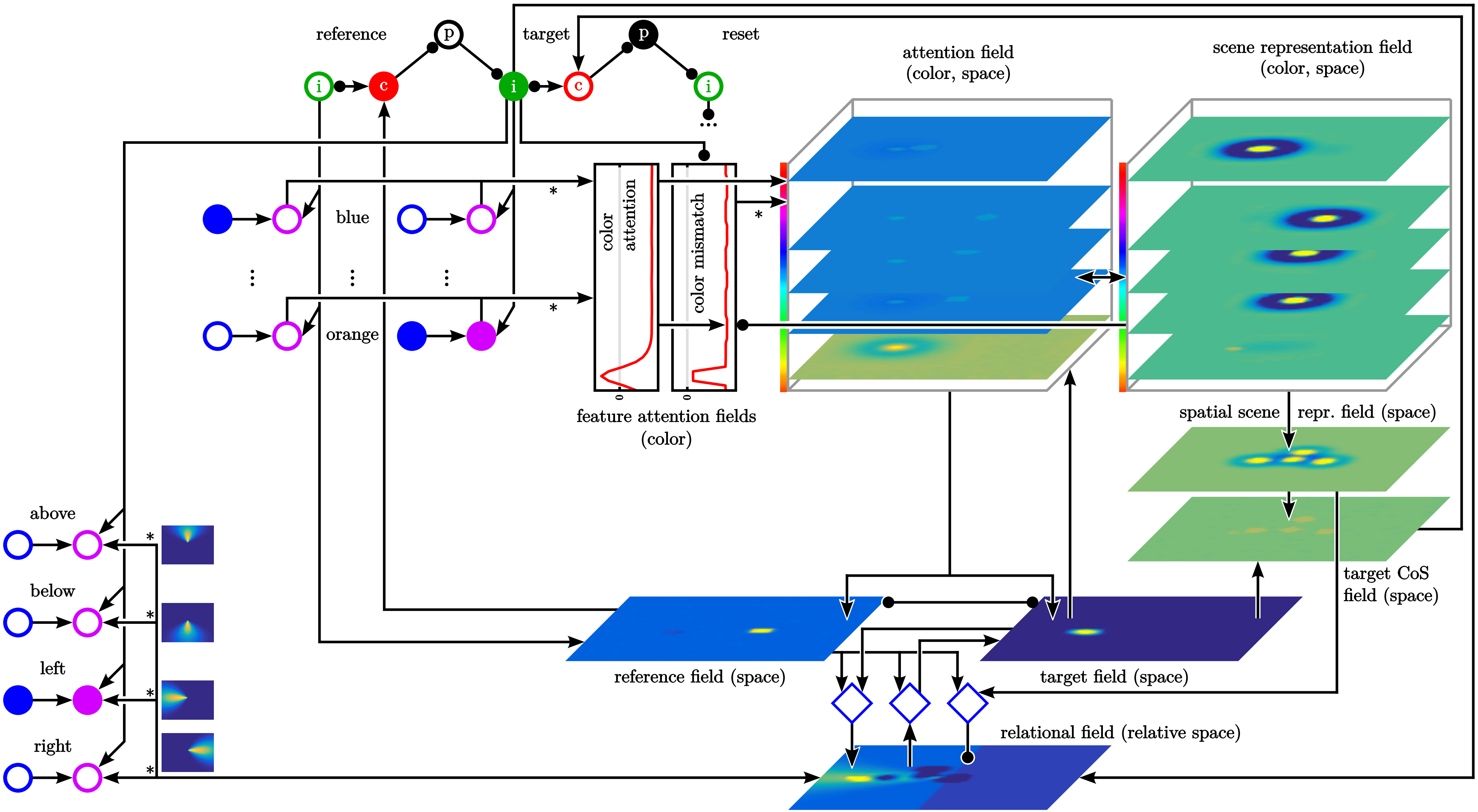
“red to the left of green”



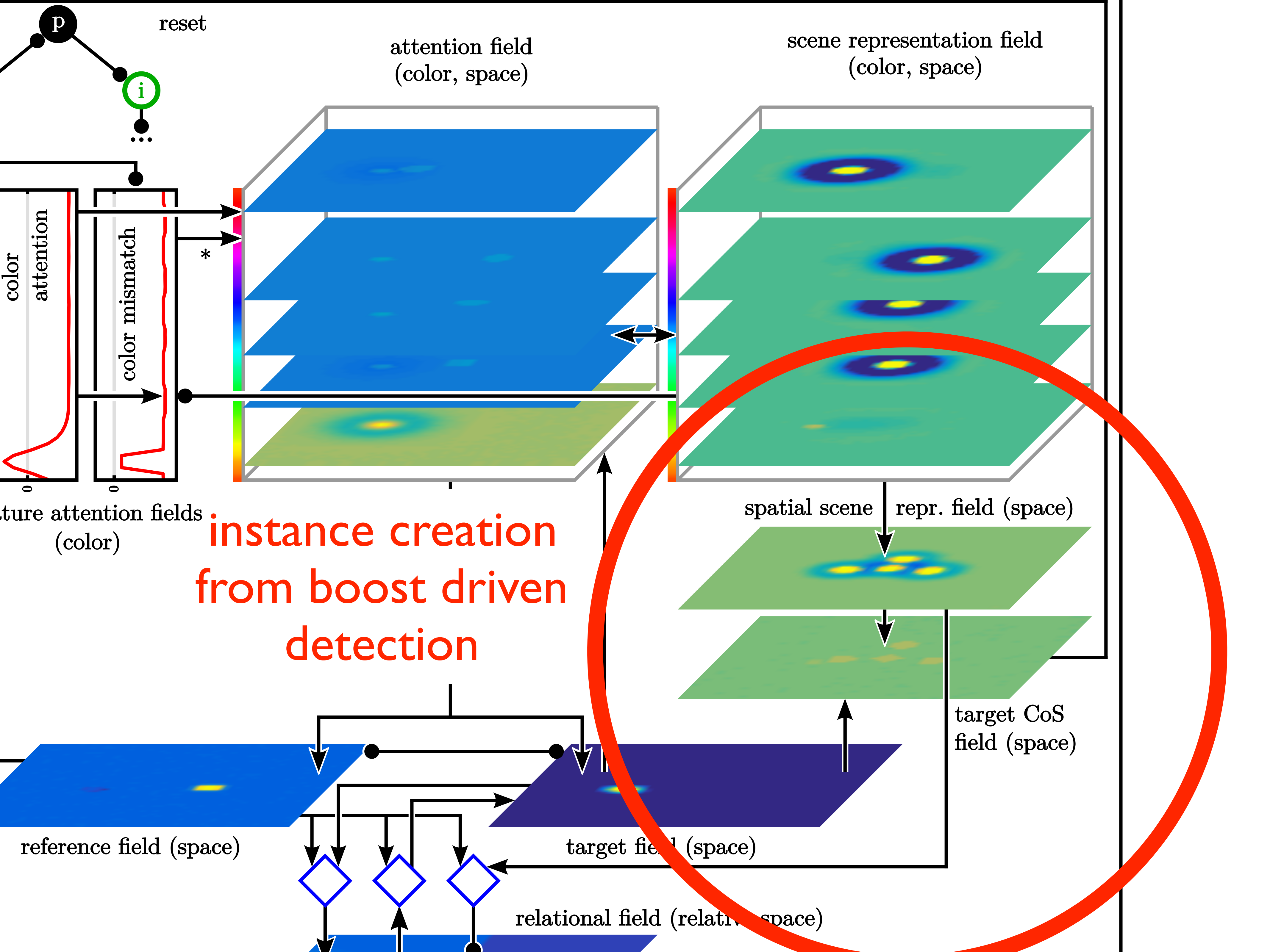
purely “mental” scene representation

- from propositions
  - “There is a cyan object above a green object.”
  - “There is a red object to the left of the green object.”
  - “There is a blue object to the right of the red object.”
  - “There is an orange object to the left of the blue object.”
- inference
  - “Where is the blue object relative to the red object?”

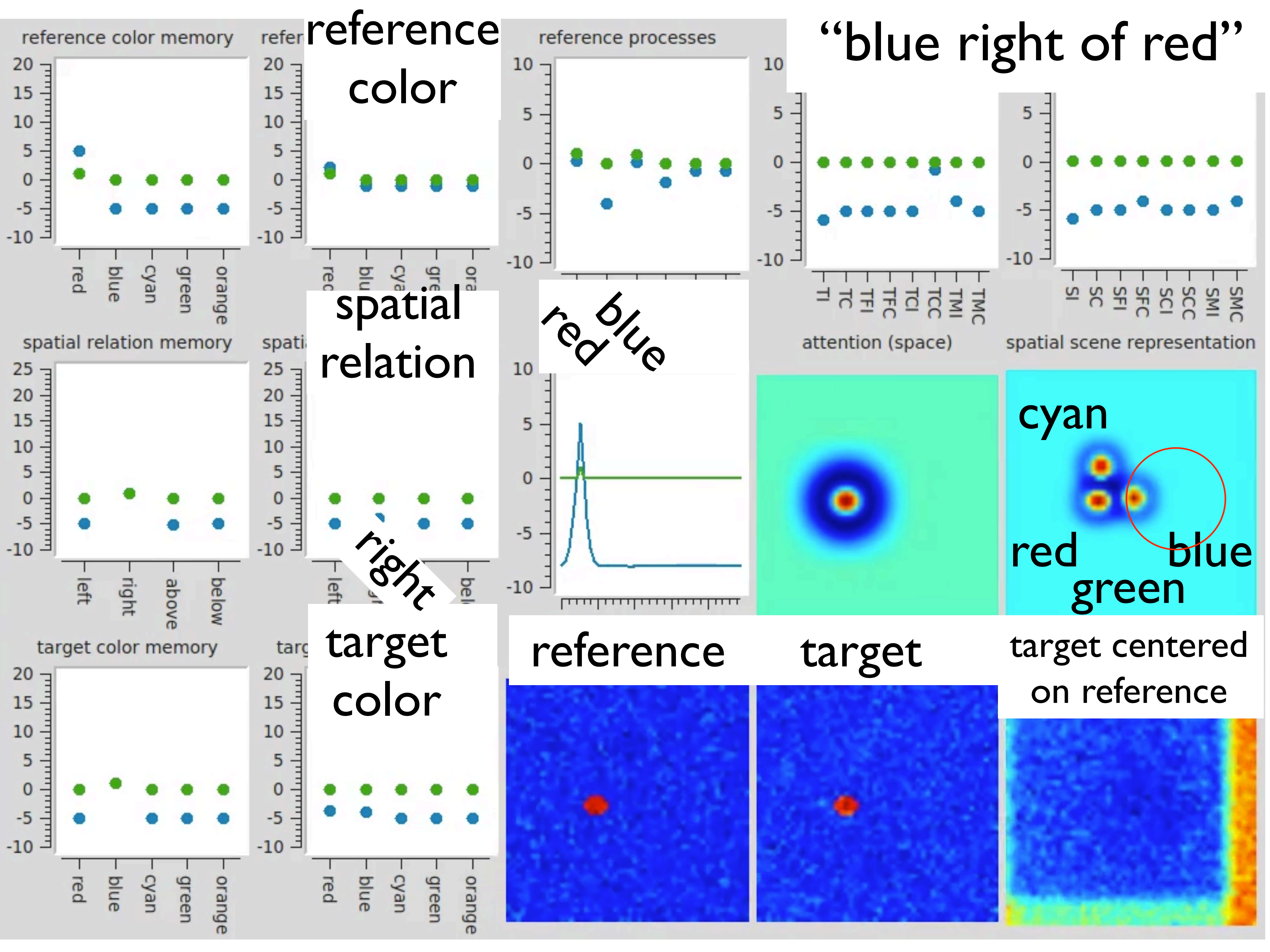
[Ragni, Knauff, *Psych Rev* 2013]



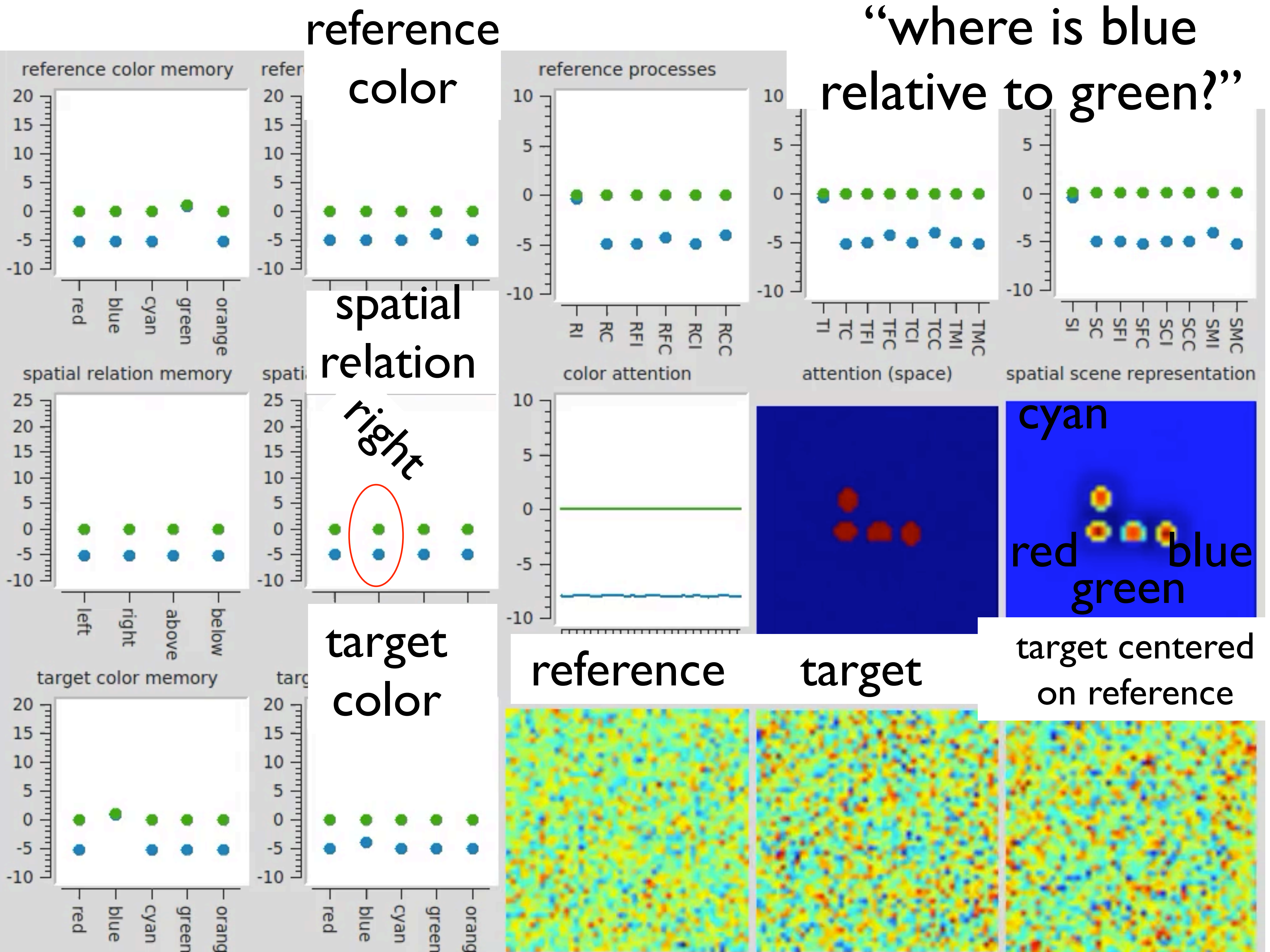
[Kounatidou, Richter, Schöner, CogSci 2018]



instance creation from boost driven detection

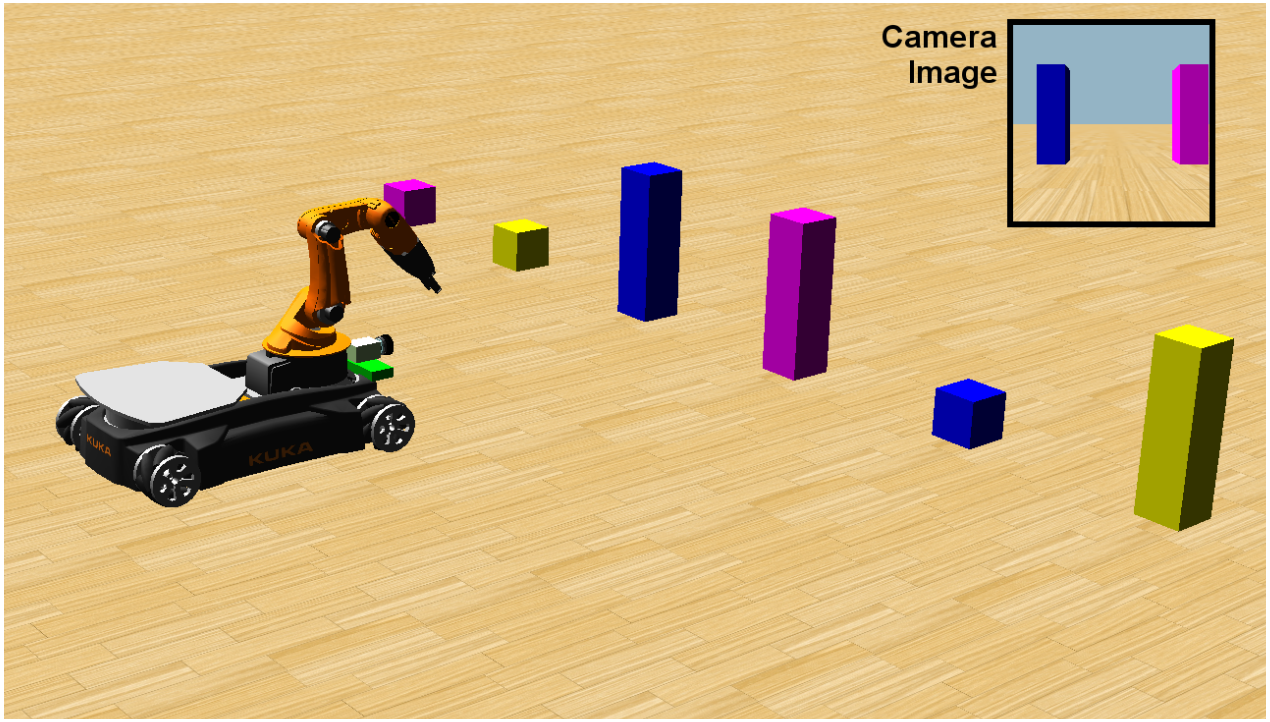


"blue right of red"
reference color memory
reference color
reference processes
spatial relation memory
spatial relation
red
blue
right
attention (space)
spatial scene representation
cyan
red
blue
green
target color memory
target color
reference
target
target centered on reference

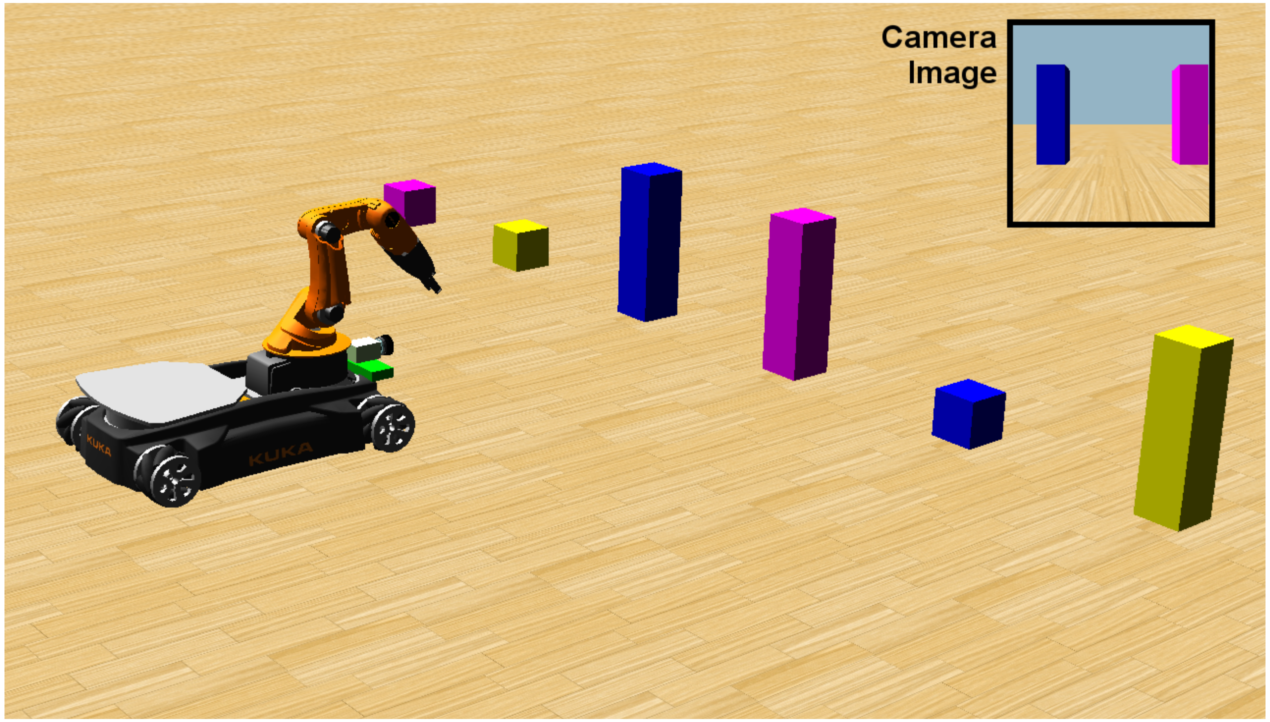


"where is blue relative to green?"
reference color
spatial relation
right
target color
reference color memory
reference processes
spatial relation memory
color attention
attention (space)
spatial scene representation
target color memory
cyan
red
blue
green
reference
target
target centered on reference
red
blue
cyan
green
orange
left
right
above
below
RI
RC
RFI
RFC
RCI
RCC
TI
TC
TFI
TFC
TCI
TCC
TMI
TMC
SI
SC
SFI
SFC
SCI
SCC
SMI
SMC

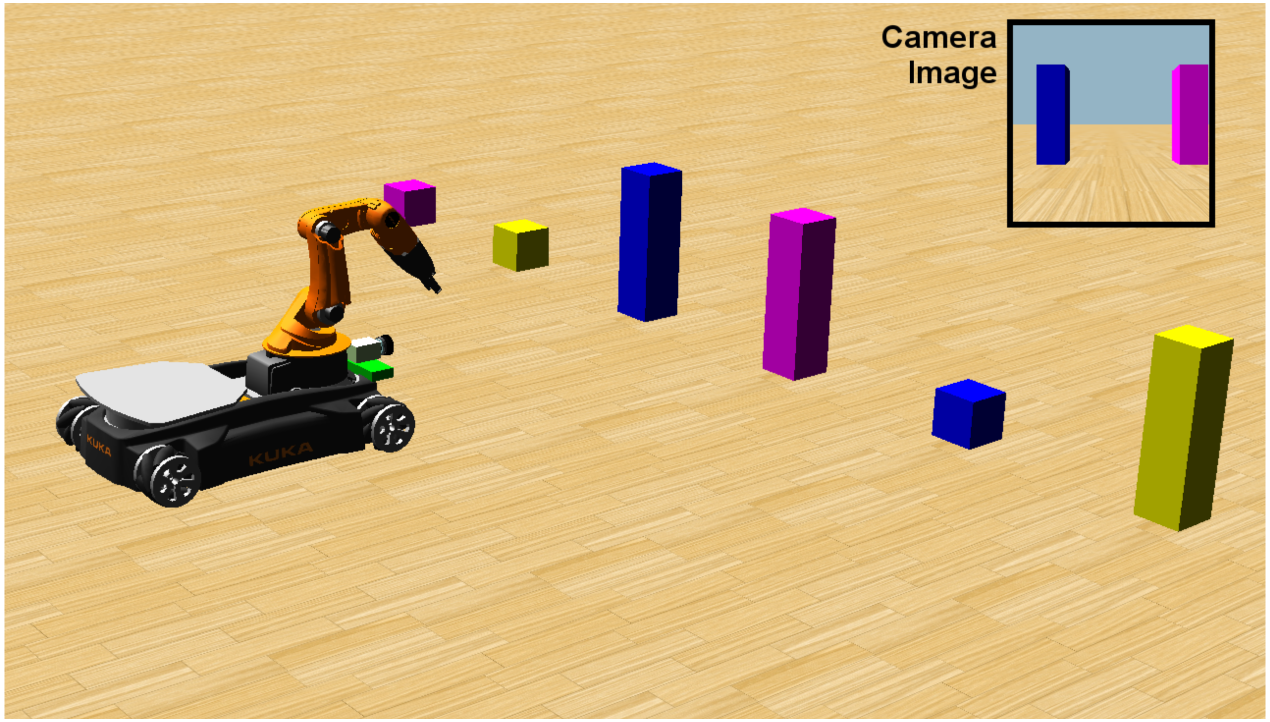
Goals, knowledge, problem solving



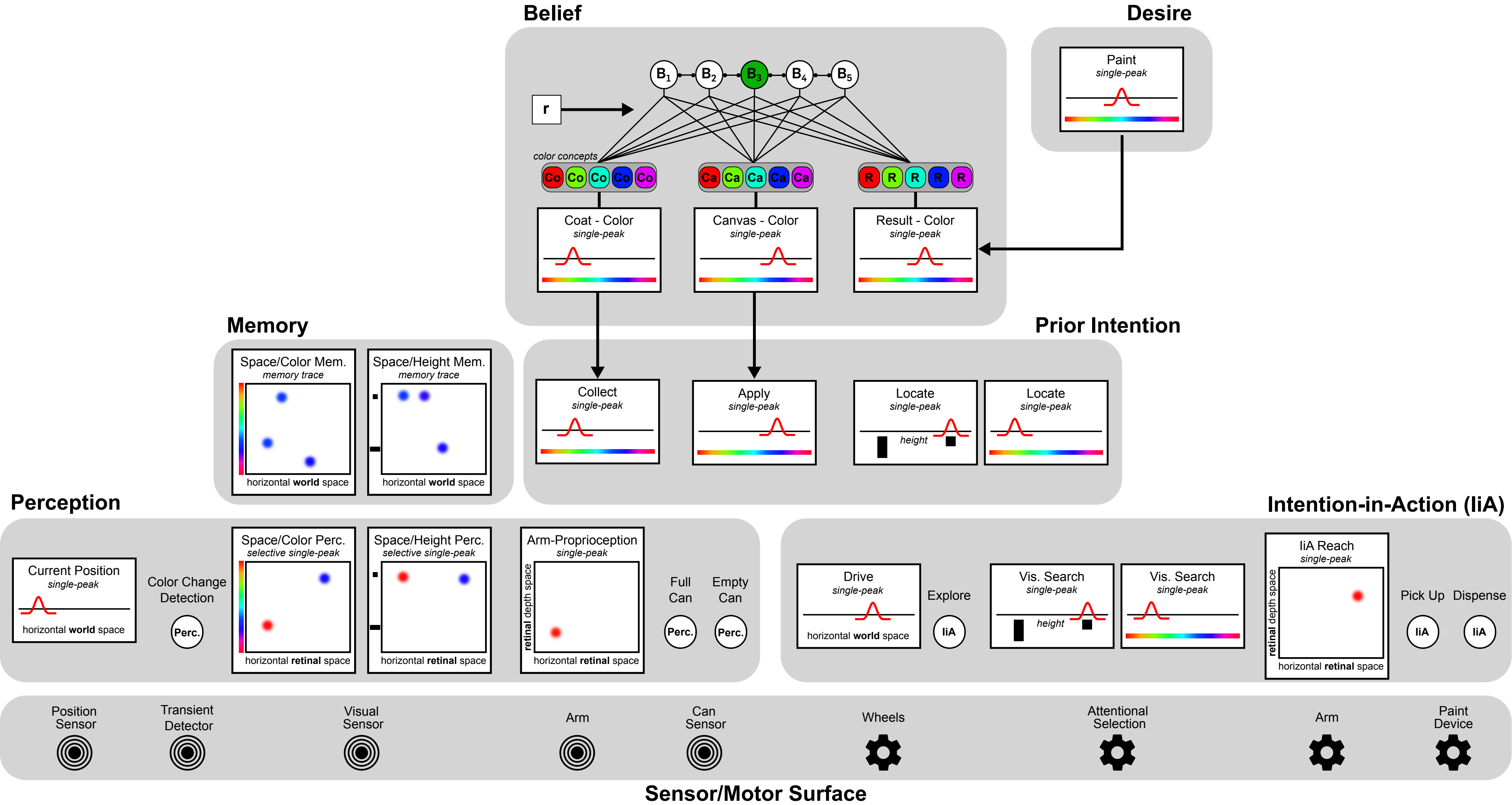
- world
  - colored objects (small)
  - paint buckets (tall)
  - vehicle with arm
- perception
  - see color/feature
  - sense position, arm, paint in gripper
- intention in action
  - move in 1D
  - reach to take up paint
  - reach to apply a coat of paint

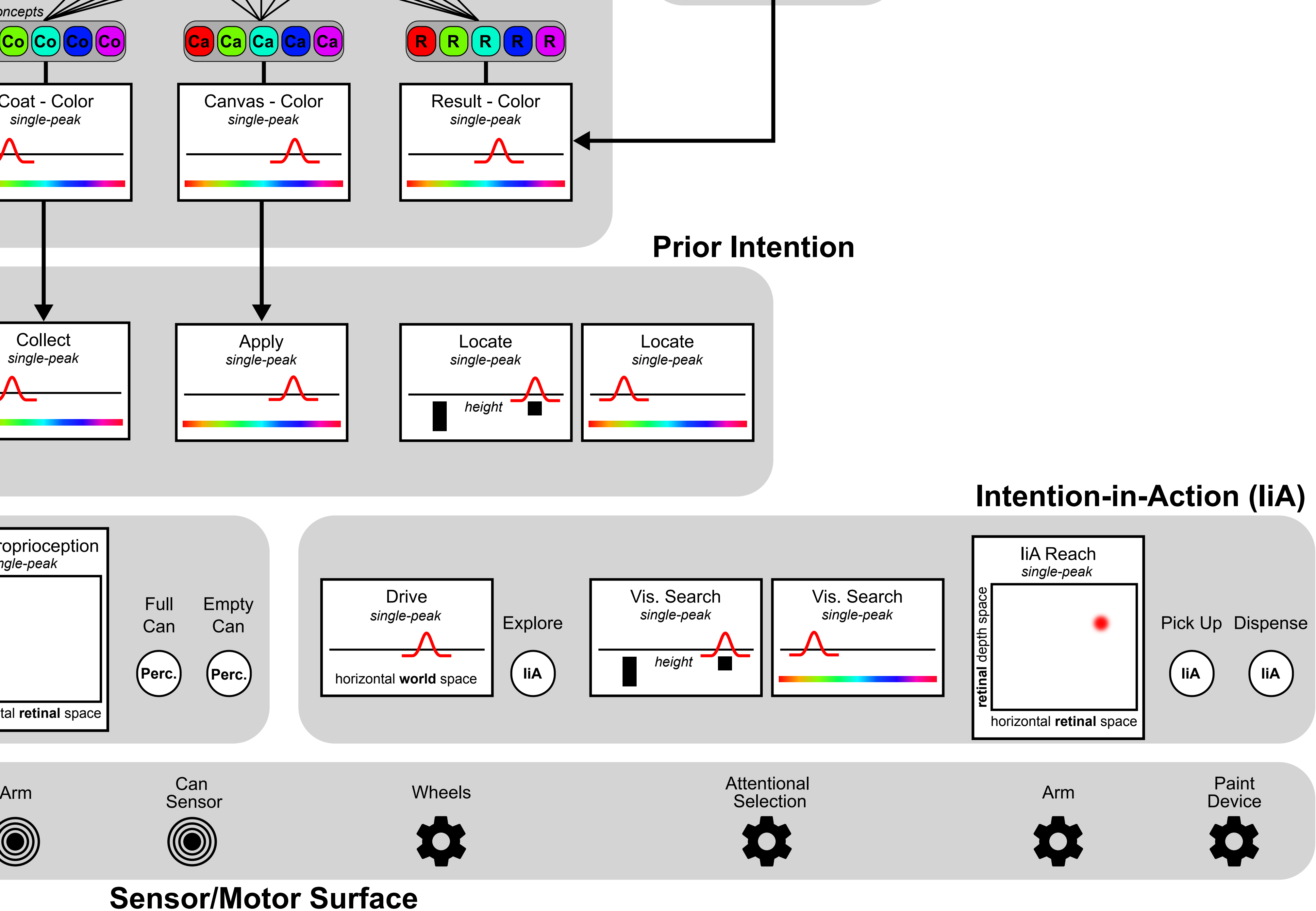


- memory
  - of visual scene
- prior intentions
  - search to paint
  - search to load paint
  - reach to apply paint
  - move to a recalled location
- beliefs
  - rules linking color concepts: which paint on which canvas generates which new color
- desires
  - for cubes of a particular color



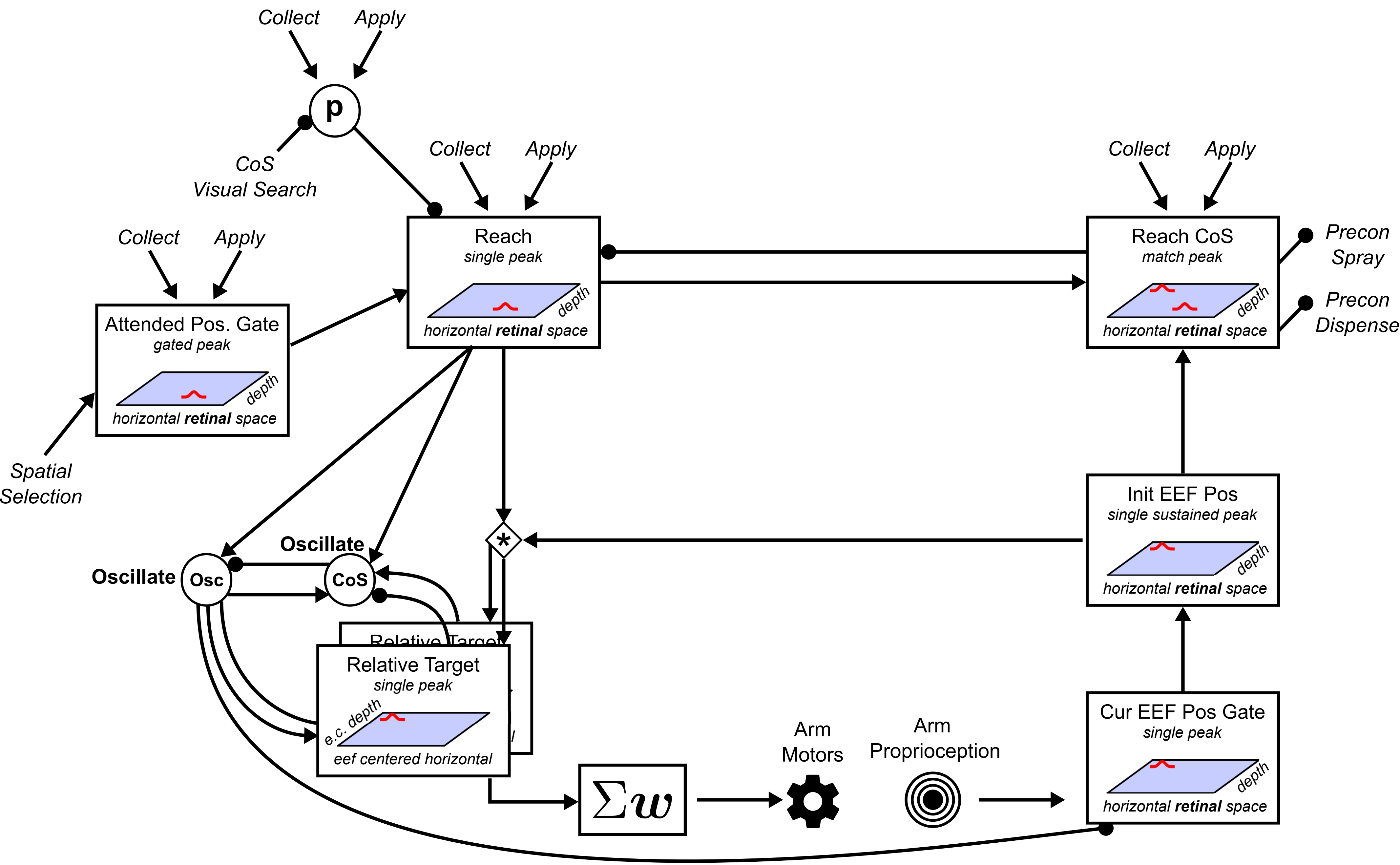
Neural dynamic architecture

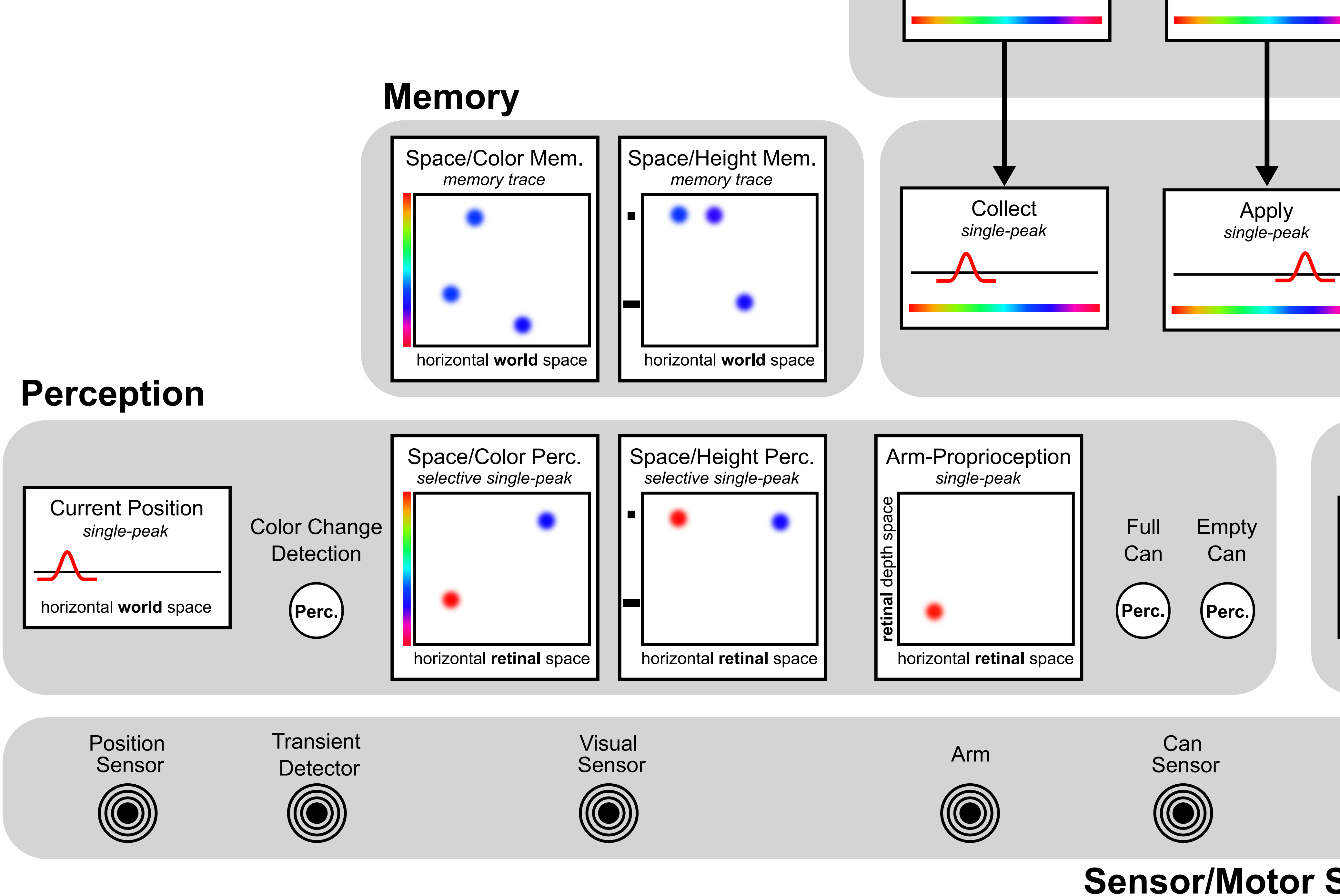




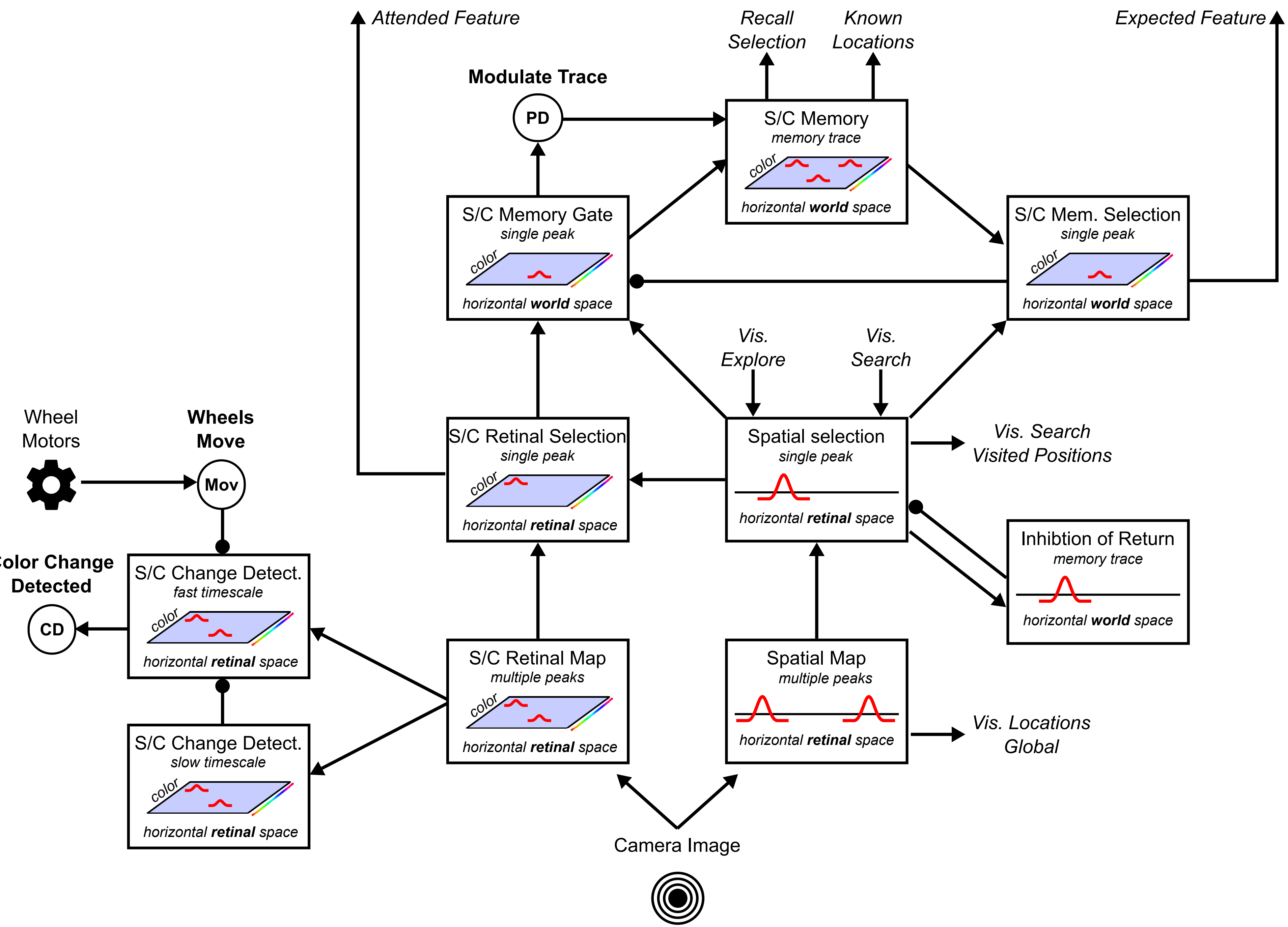
oncepts
Co Co Co Co
Ca Ca Ca Ca Ca
R R R R R
Coat - Color
single-peak
Canvas - Color
single-peak
Result - Color
single-peak
Prior Intention
Collect
single-peak
Apply
single-peak
Locate
single-peak
height
Locate
single-peak
Intention-in-Action (IiA)
roprioception
ngle-peak
tal retinal space
Full Can
Perc.
Empty Can
Perc.
Drive
single-peak
horizontal world space
Explore
IiA
Vis. Search
single-peak
height
Vis. Search
single-peak
IiA Reach
single-peak
retinal depth space
horizontal retinal space
Pick Up
IiA
Dispense
IiA
Arm
Can Sensor
Wheels
Attentional Selection
Arm
Paint Device
Sensor/Motor Surface

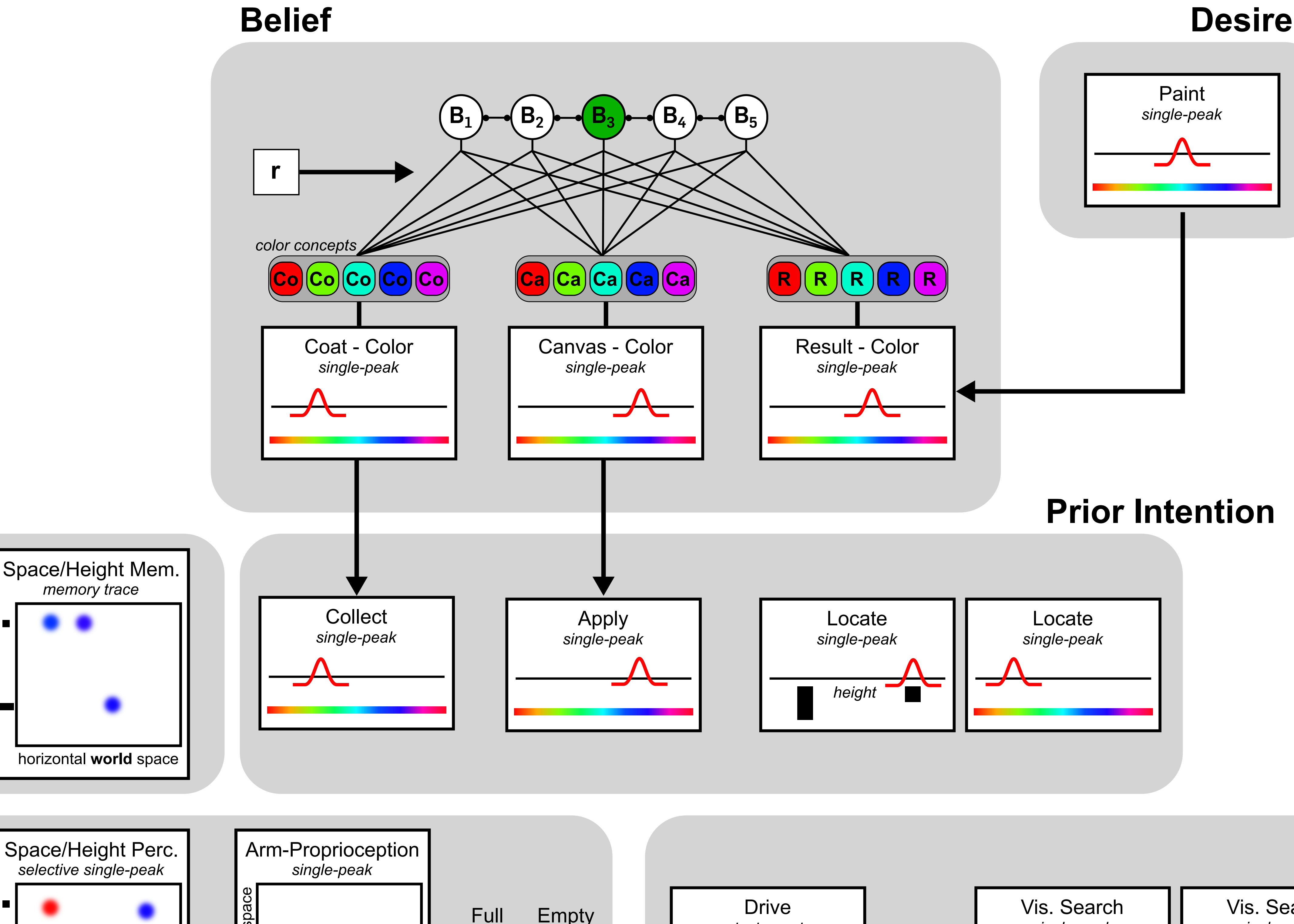
Intention in action: reach



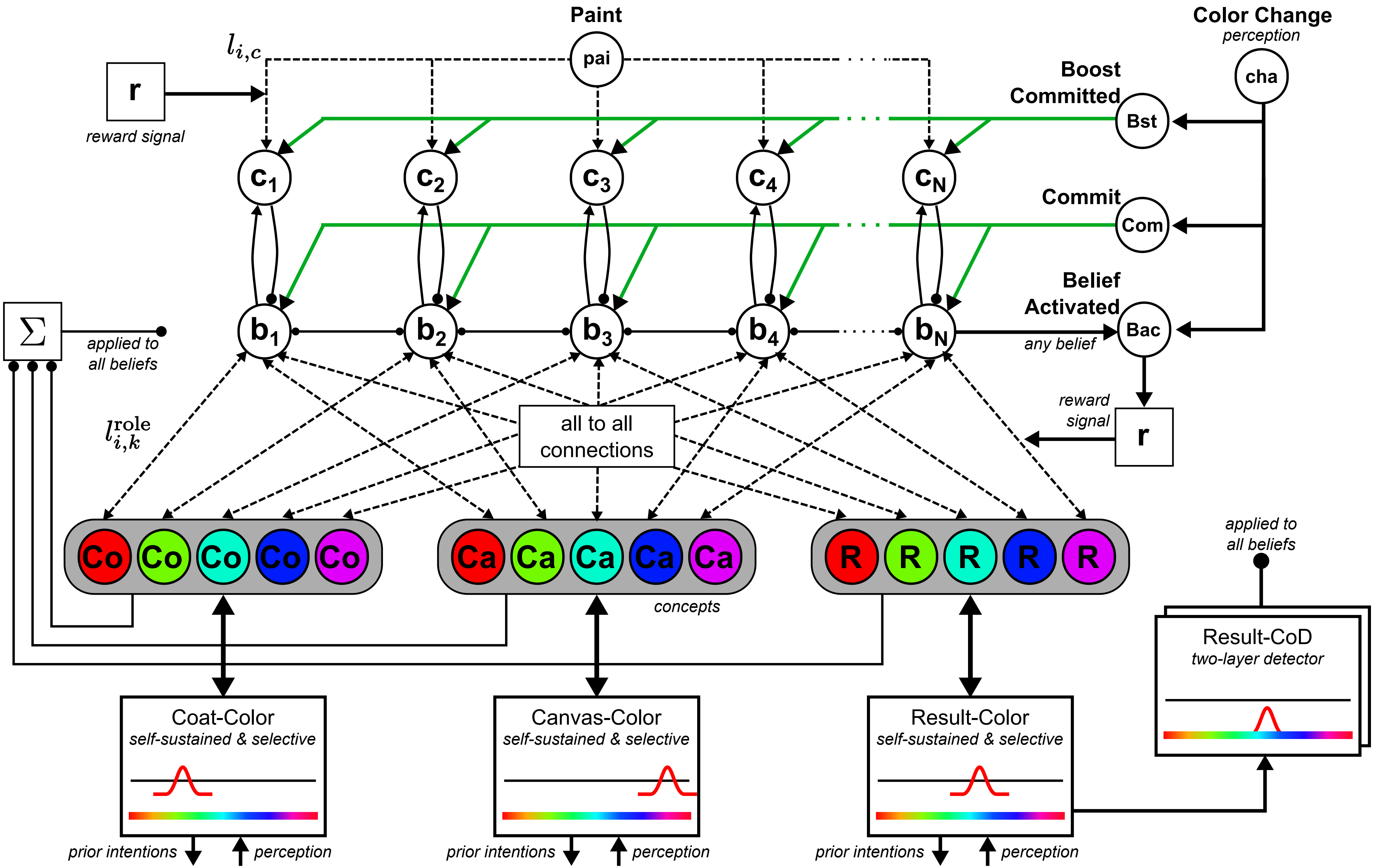

Memory
Space/Color Mem.
memory trace
horizontal **world** space
Space/Height Mem.
memory trace
horizontal **world** space
Collect
single-peak
Apply
single-peak
Perception
Current Position
single-peak
horizontal **world** space
Color Change Detection
Perc.
Space/Color Perc.
selective single-peak
horizontal **retinal** space
Space/Height Perc.
selective single-peak
horizontal **retinal** space
Arm-Proprioception
single-peak
retinal depth space
horizontal **retinal** space
Full Can
Perc.
Empty Can
Perc.
Position Sensor
Transient Detector
Visual Sensor
Arm
Can Sensor
Sensor/Motor S

Perception and memory



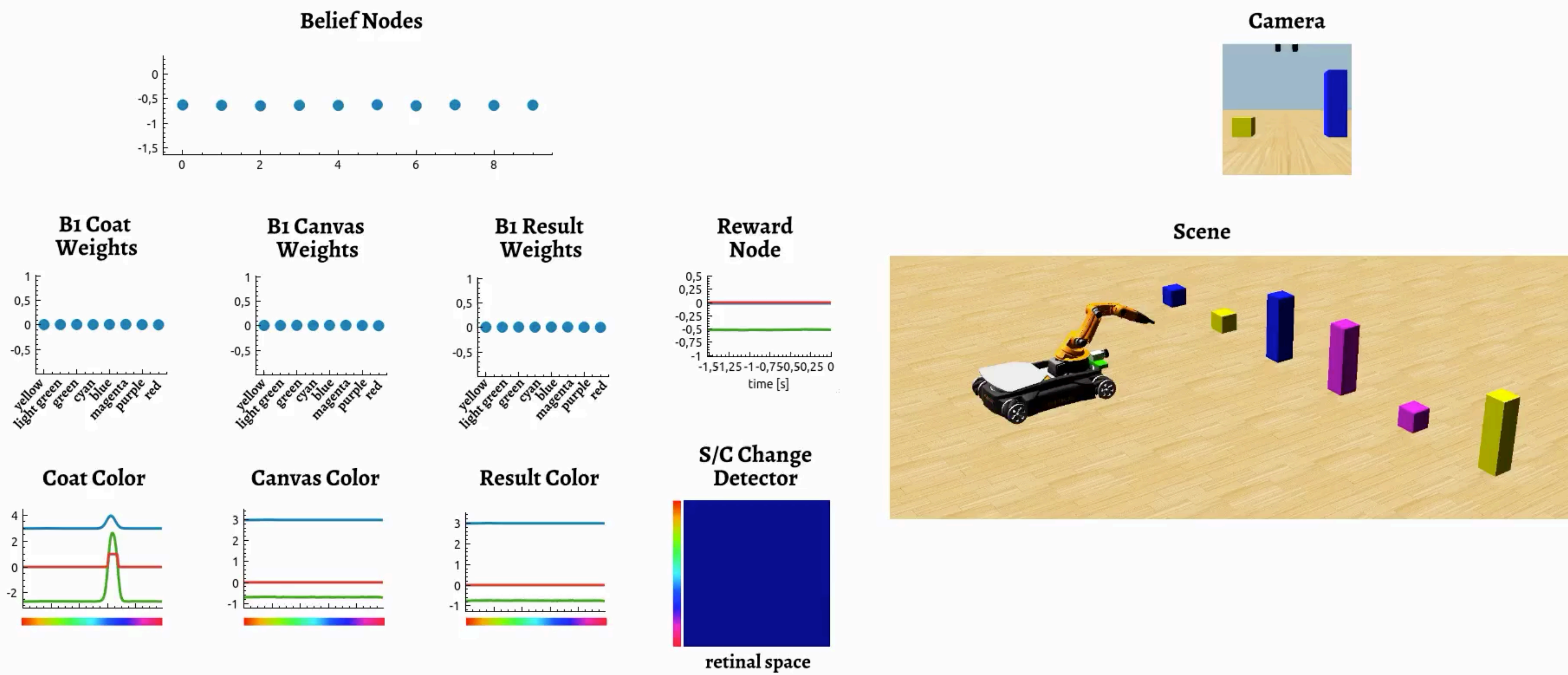

Belief
Desire
r
color concepts
Co Co Co Co Co
Ca Ca Ca Ca Ca
R R R R R
Coat - Color
single-peak
Canvas - Color
single-peak
Result - Color
single-peak
Paint
single-peak
Prior Intention
Space/Height Mem.
memory trace
horizontal world space
Collect
single-peak
Apply
single-peak
Locate
single-peak
height
Locate
single-peak
Space/Height Perc.
selective single-peak
Arm-Proprioception
single-peak
Full
Empty
Drive
Vis. Search

Learning a new belief



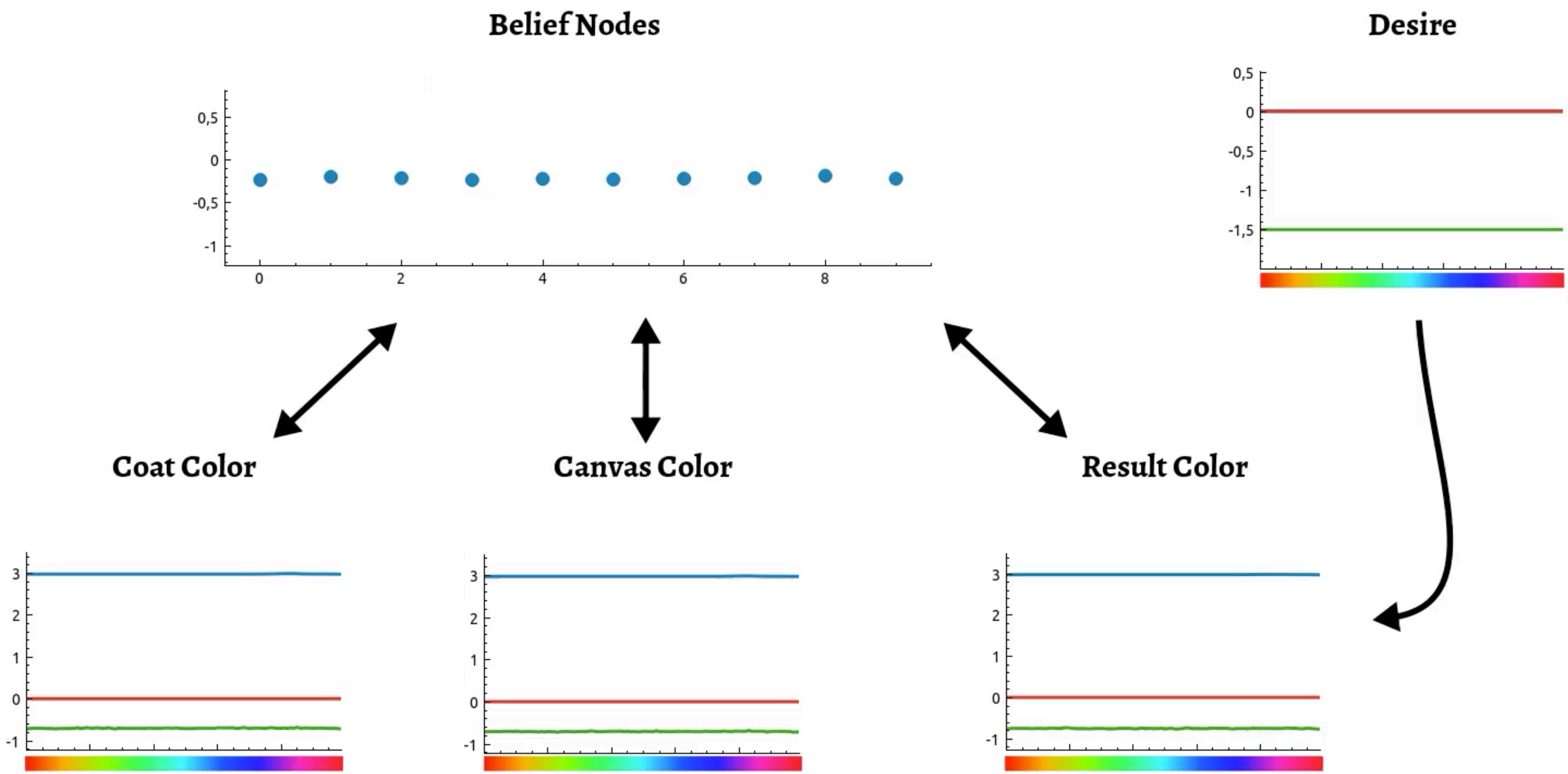
Learn a new belief

[while exploring: applying blue paint to yellow cube]



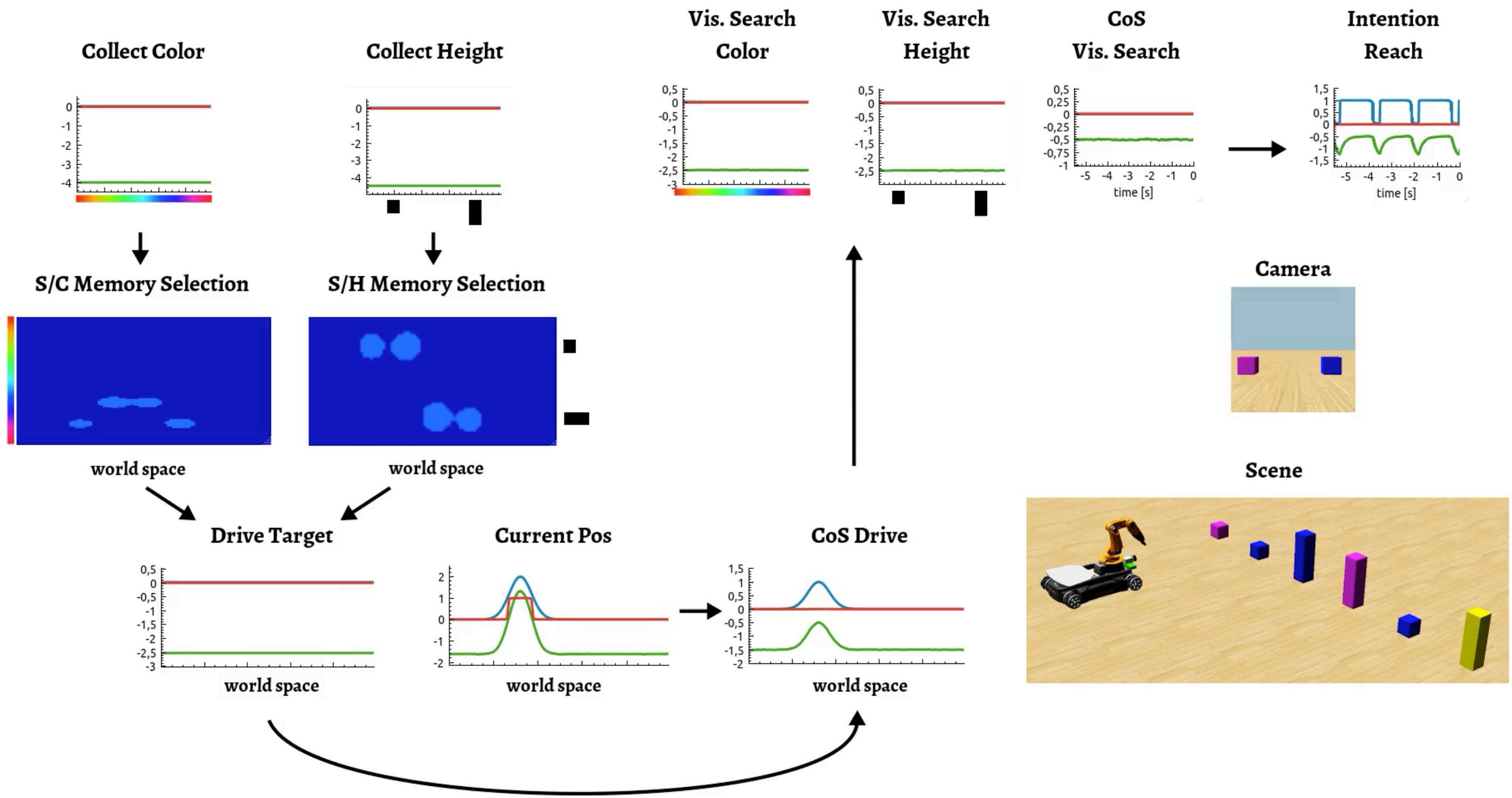
Recall a belief

[triggered by a desire and objects in scene memory]



Recall-drive-search

[based on a desire and an activated belief, looking for a tall pink object, which is in memory]



Summary

- low-dimensional activation fields enable the autonomous generation of sequences of mental states
  - events emerge from detection decisions
  - attentional foreground from selection decisions
  - low-dimensional activation fields as substrates for scene and mental maps
  - concepts, neural operators, and coordinate transforms enables generalization, inference
- stability => robustness => architectures
- enables autonomous learning

Conclusions

- a privileged level of description for "pervasively neural" process accounts for behavior and thinking

TOPICS

TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2019) 1–15

The Dynamics of Neural Populations Capture the Laws of the Mind

Gregor Schöner

What would pervasive neural processing “buy us” ?

- embodiment for “free”: updating, control, coupling
- coherent architectures: understandable
- learn back-ground knowledge (rather than program many special cases)
- autonomous learning
- low-energy implementations => Yulia Sandamirskaya

A long way off?

- not necessarily… framework is becoming visible
- scaling to realistic scenarios as a challenge
  - use learning (deep learning?) to extract the low-dimensional representations within which neural dynamic cognition may work
- autonomous learning within this vision… still a challenge
  - develop the process infrastructure for that…
  - but also: study implications for its use