Scene Representation for Robots: From Elementary Behaviors to Autonomous Grasping

Stephan Zibner

RUB

INI

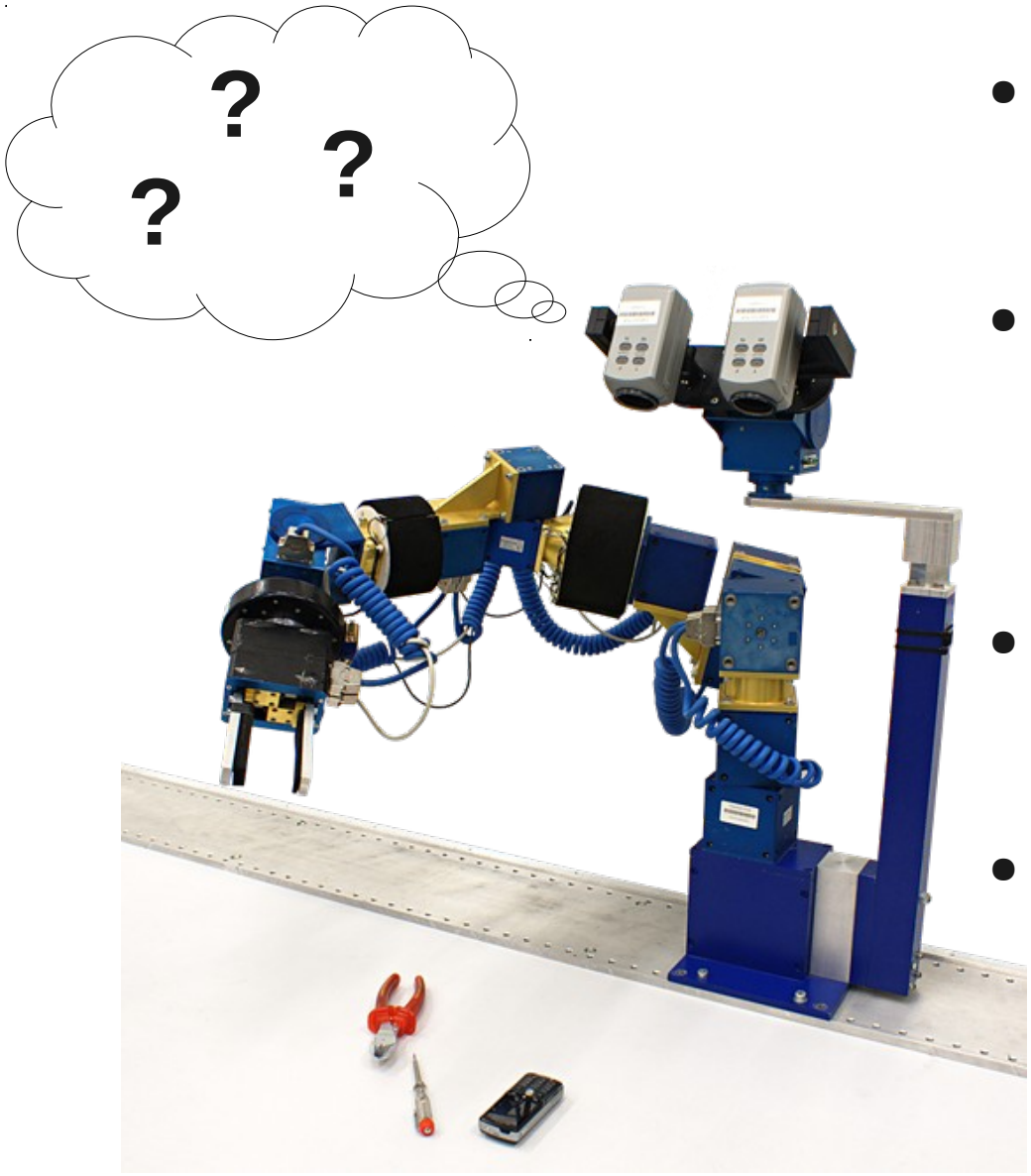
SEVENTH FRAMEWORK PROGRAMME

Outline

- Representing Scenes
- Building Blocks
  - Saliency
  - Elementary Cognitive Units
  - Change Detection
- Autonomous Behaviors
  - Exploration
  - Maintenance
  - Query
- Reaching and Grasping

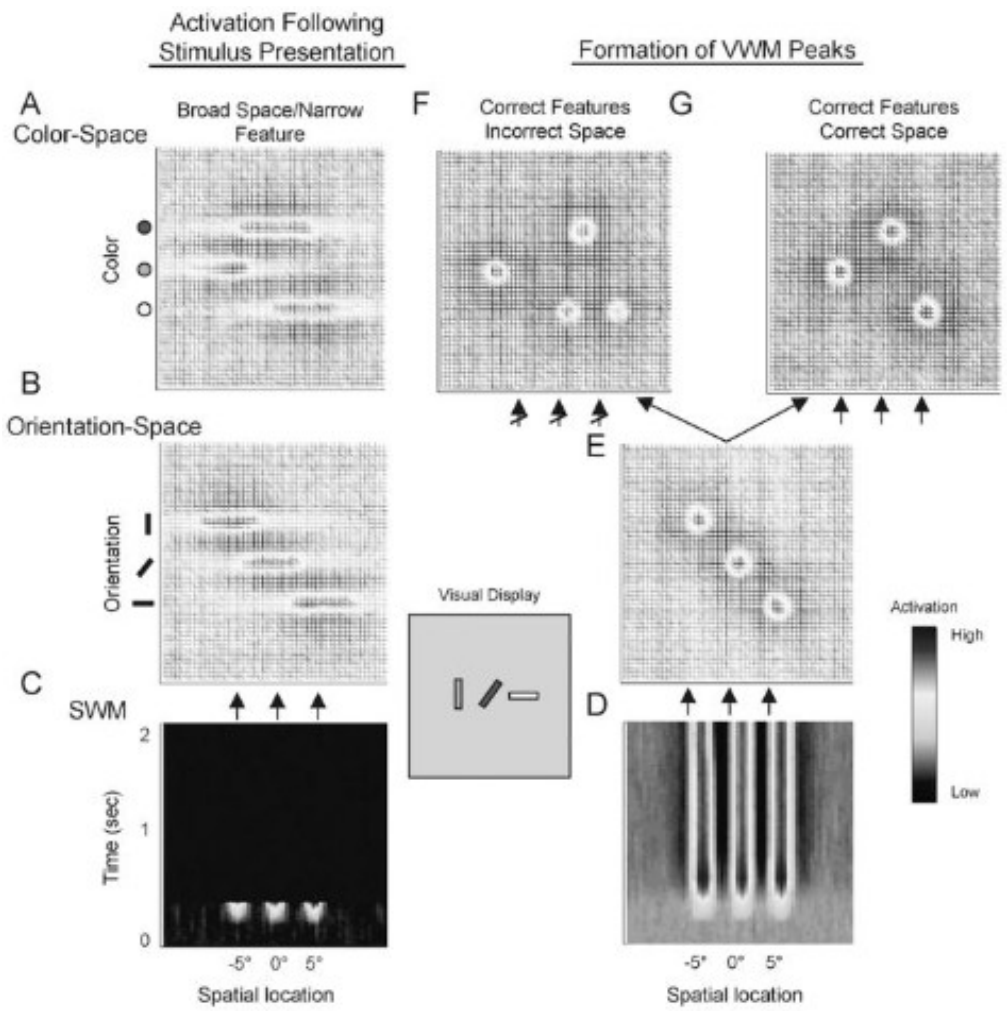
Representing Scenes

Scene Representation



- internal representation of environment
- foundation for every higher cognitive operation and action
- stable despite eye and body movements
- limited capacity, link to long-term memory

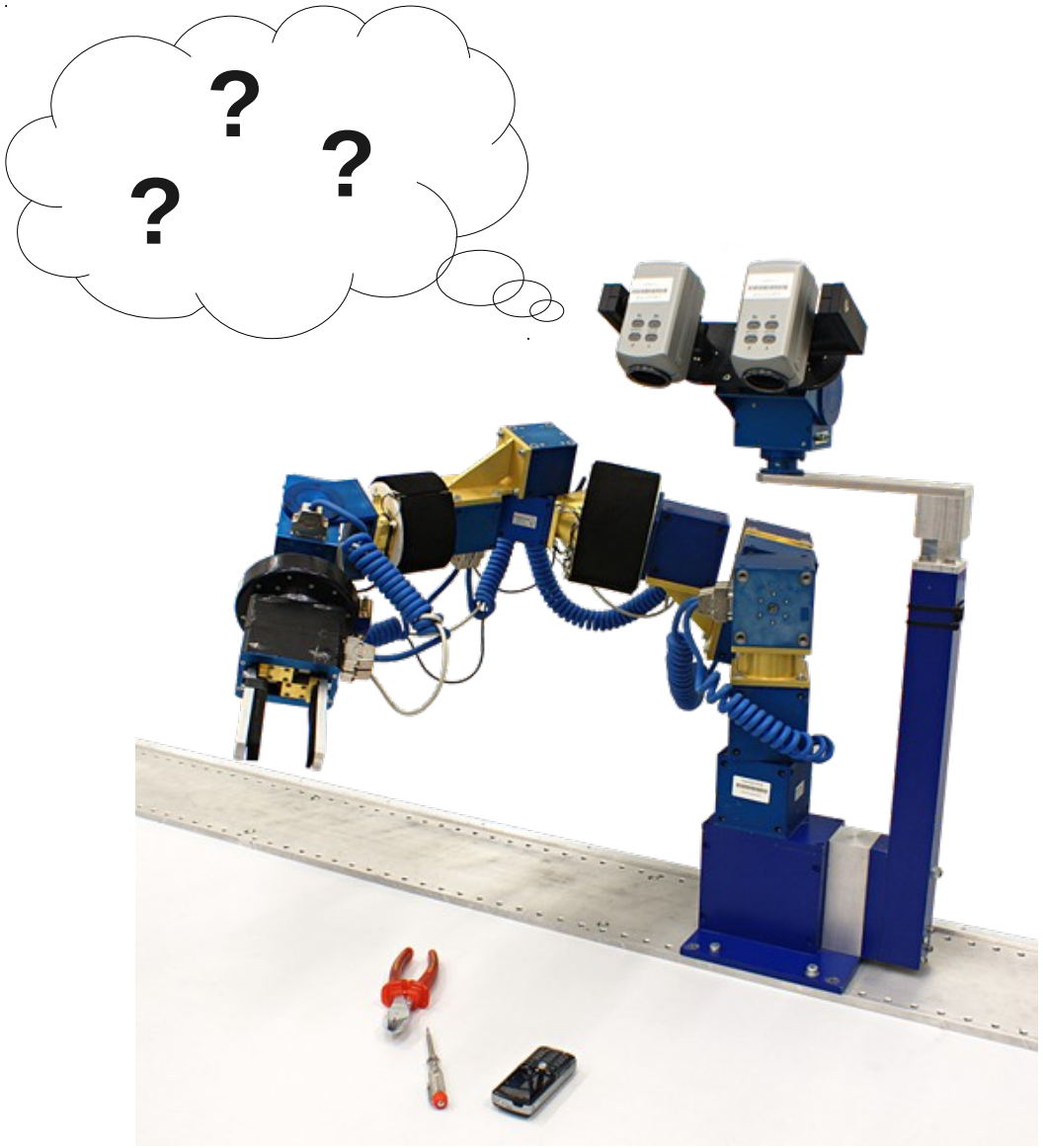
DFT Model



Johnson, Spencer, Schöner (2008)

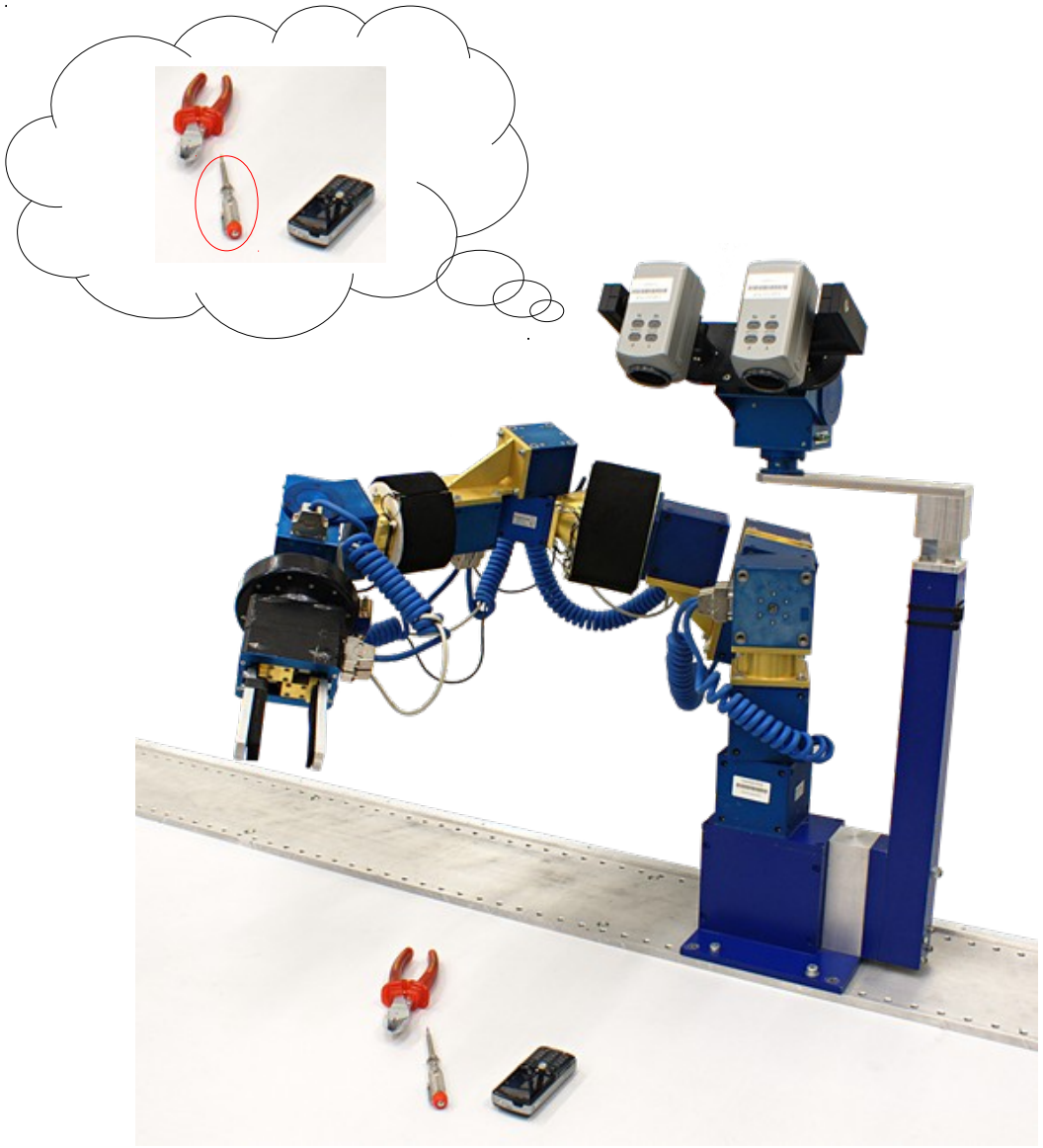
- humans solve tasks in scenes effortlessly (in normal life)
- model of human visual working memory based on dynamic neural fields
- decomposition of features, binding through shared space

Robotic Scenario



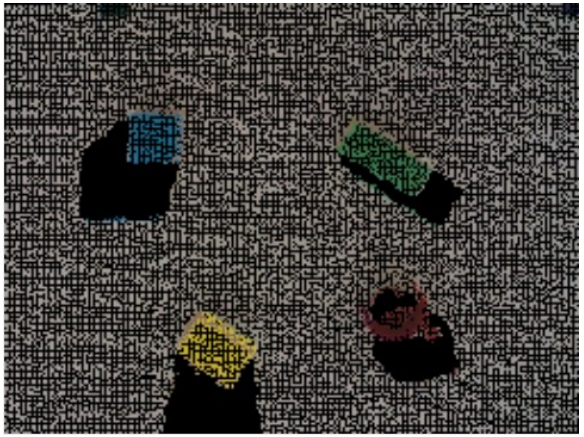
- apply to table-top scenario and human interaction
- use the internal representation for behavior generation (e.g., grasping)
- interact with humans (“hand me the red screwdriver”, “what's to the left of the pliers?”)

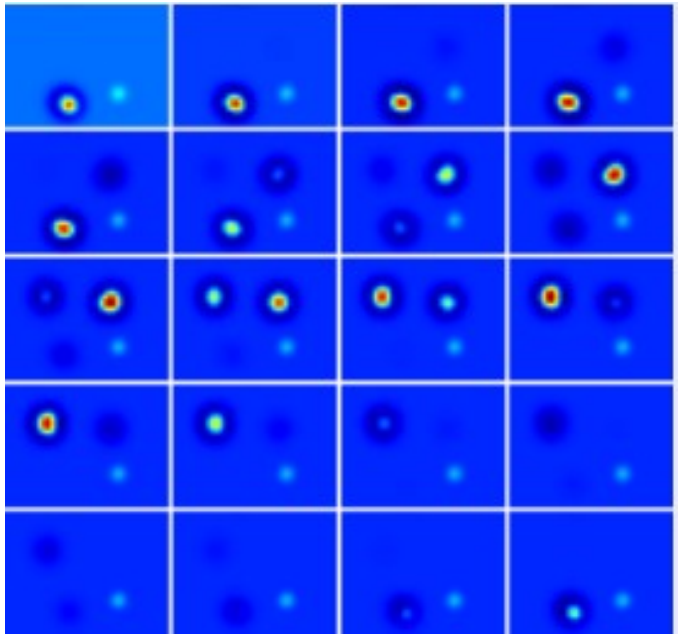
Robotic Scenario: Behaviors



- explore the environment and store objects and their features internally
- maintain the internal representation
- query the represen-tation to create autonomous action based on the representation

Robotic Scenario: Challenges

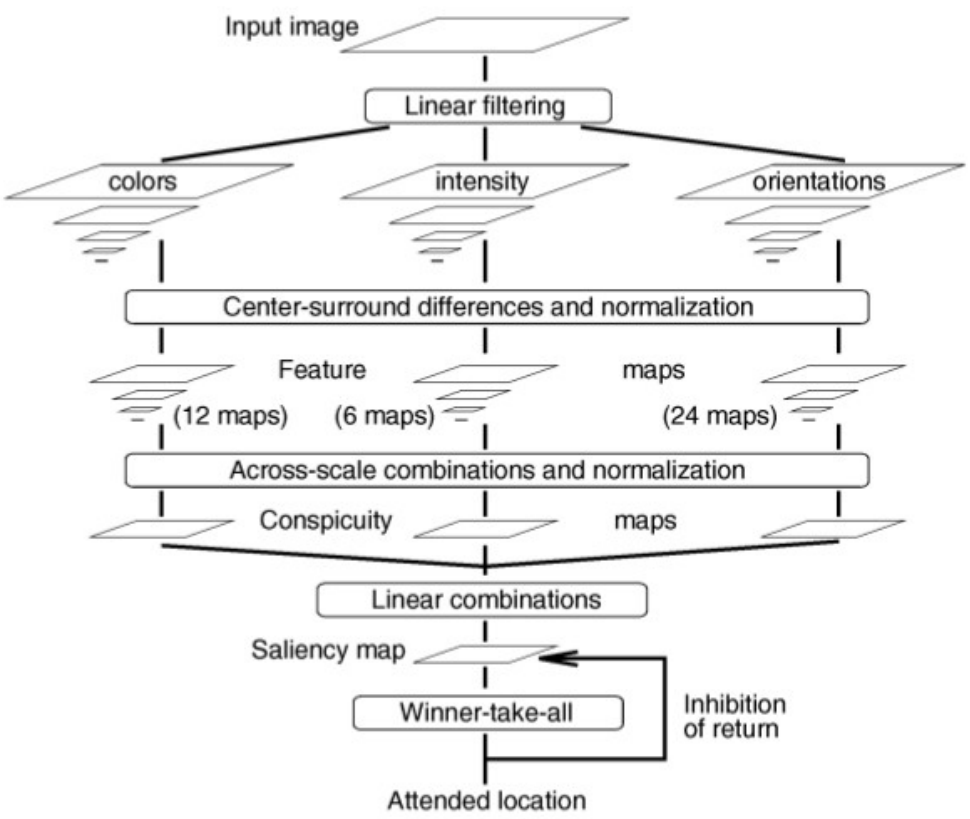


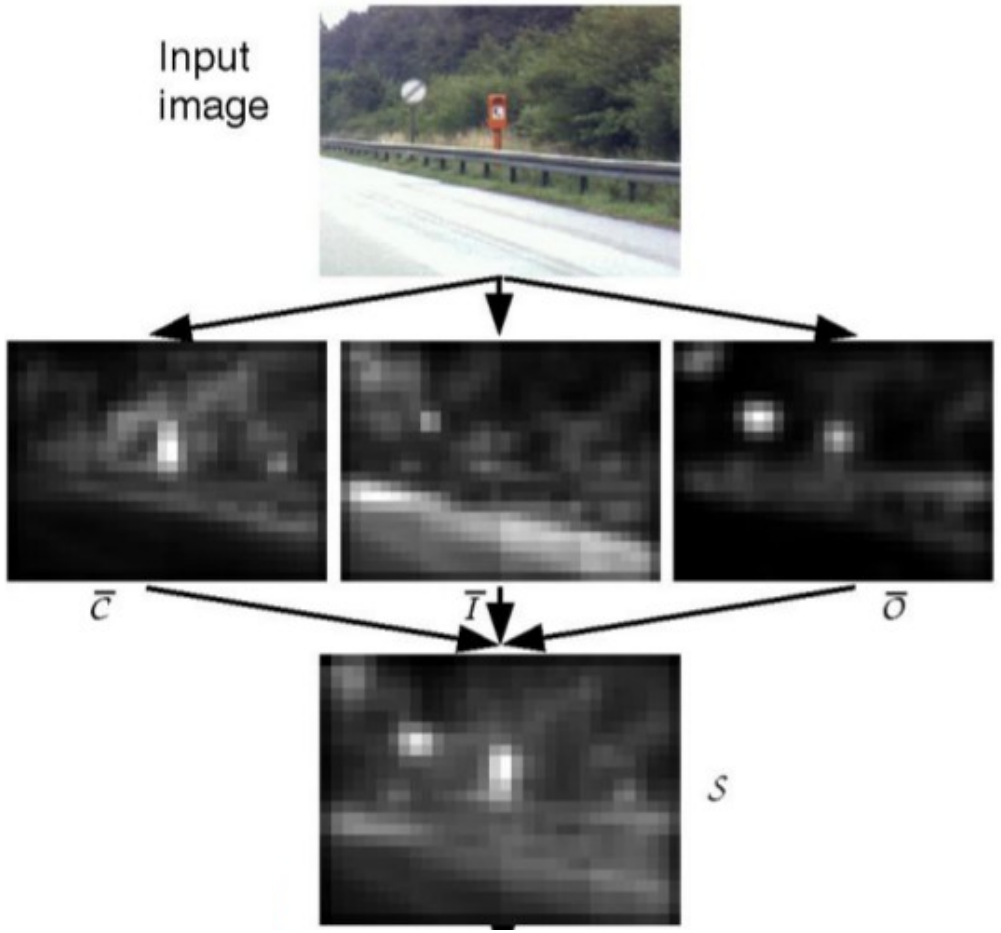


- real sensory input
- moving sensors
- limited field of view
- 3D space
- dynamic scenes
- multiple behaviors
- computational constraints
- and many more ...

Building Blocks

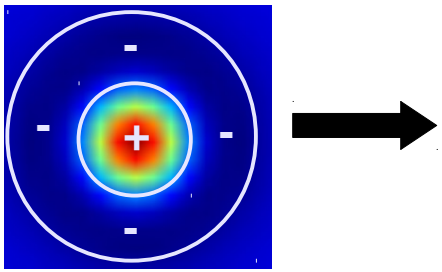
Saliency

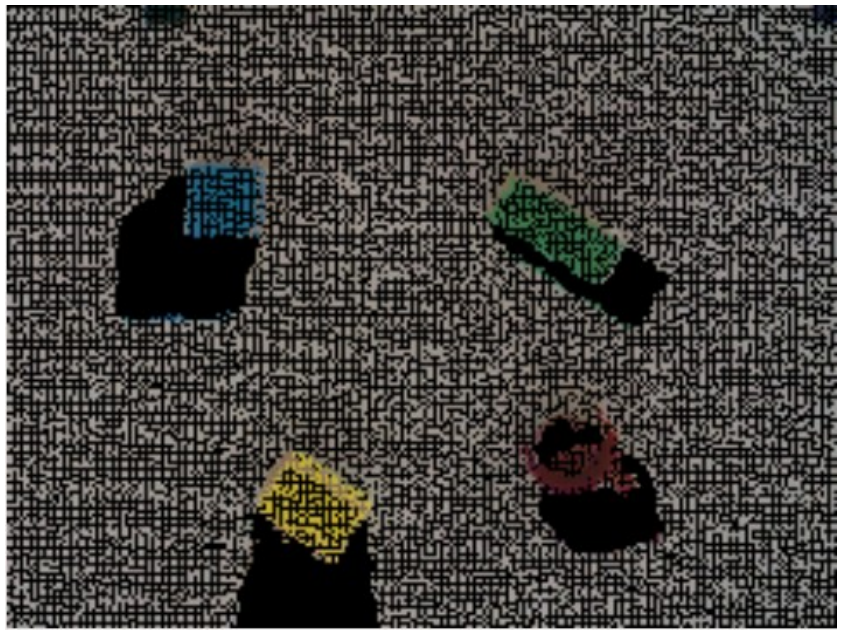




Itti, Koch, Niebur (1998)

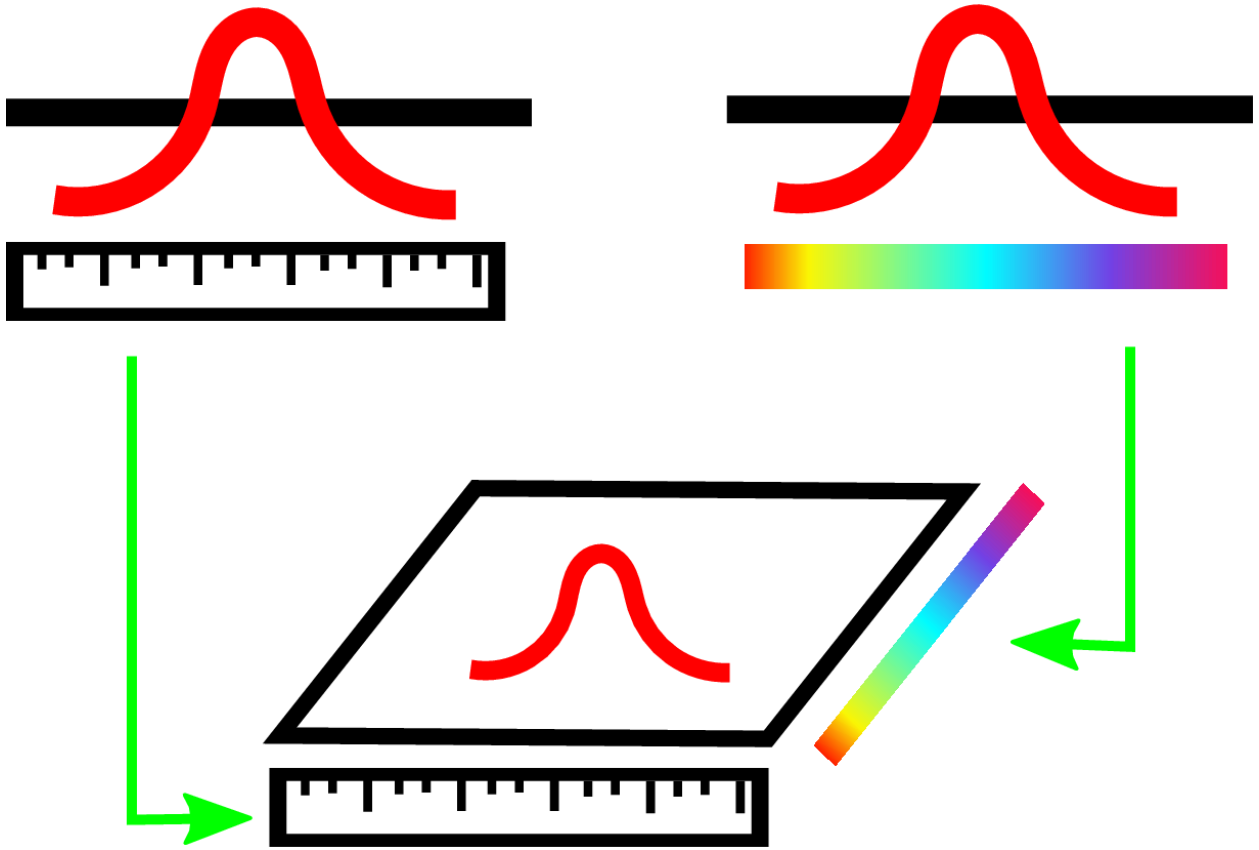
Saliency





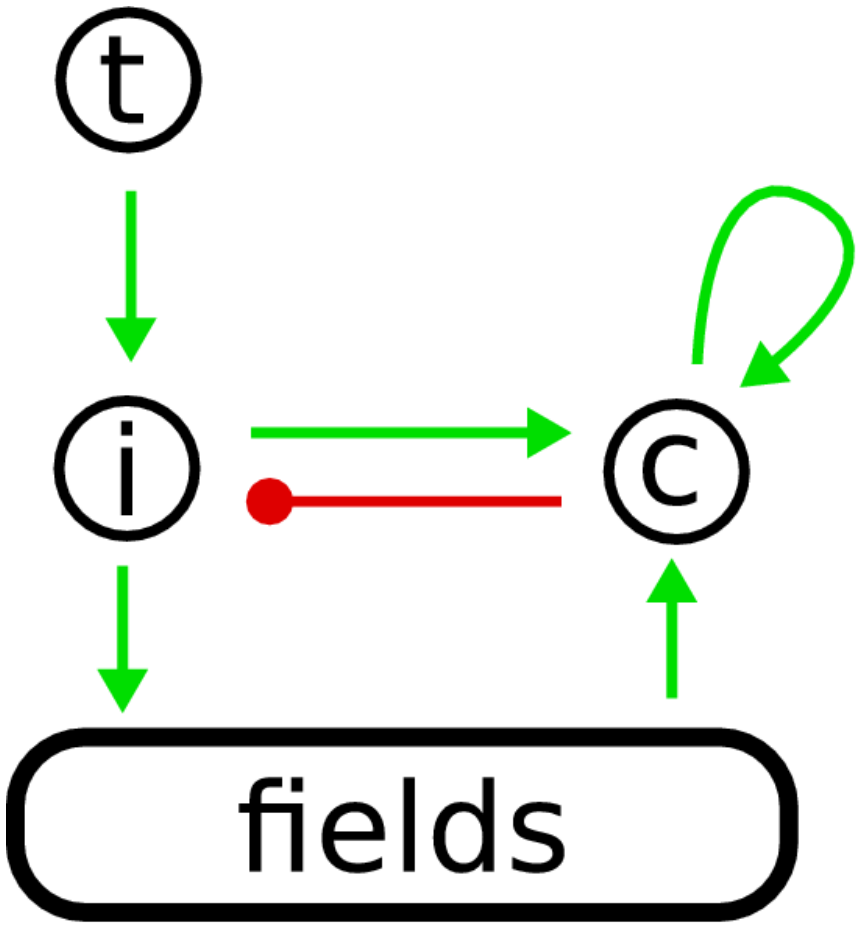
- on-/off-center responses
- uniform regions result in zero responses
- objects fitting into on-center region produce non-zero responses
- these lead to detection decision in fields

Space-Feature Links



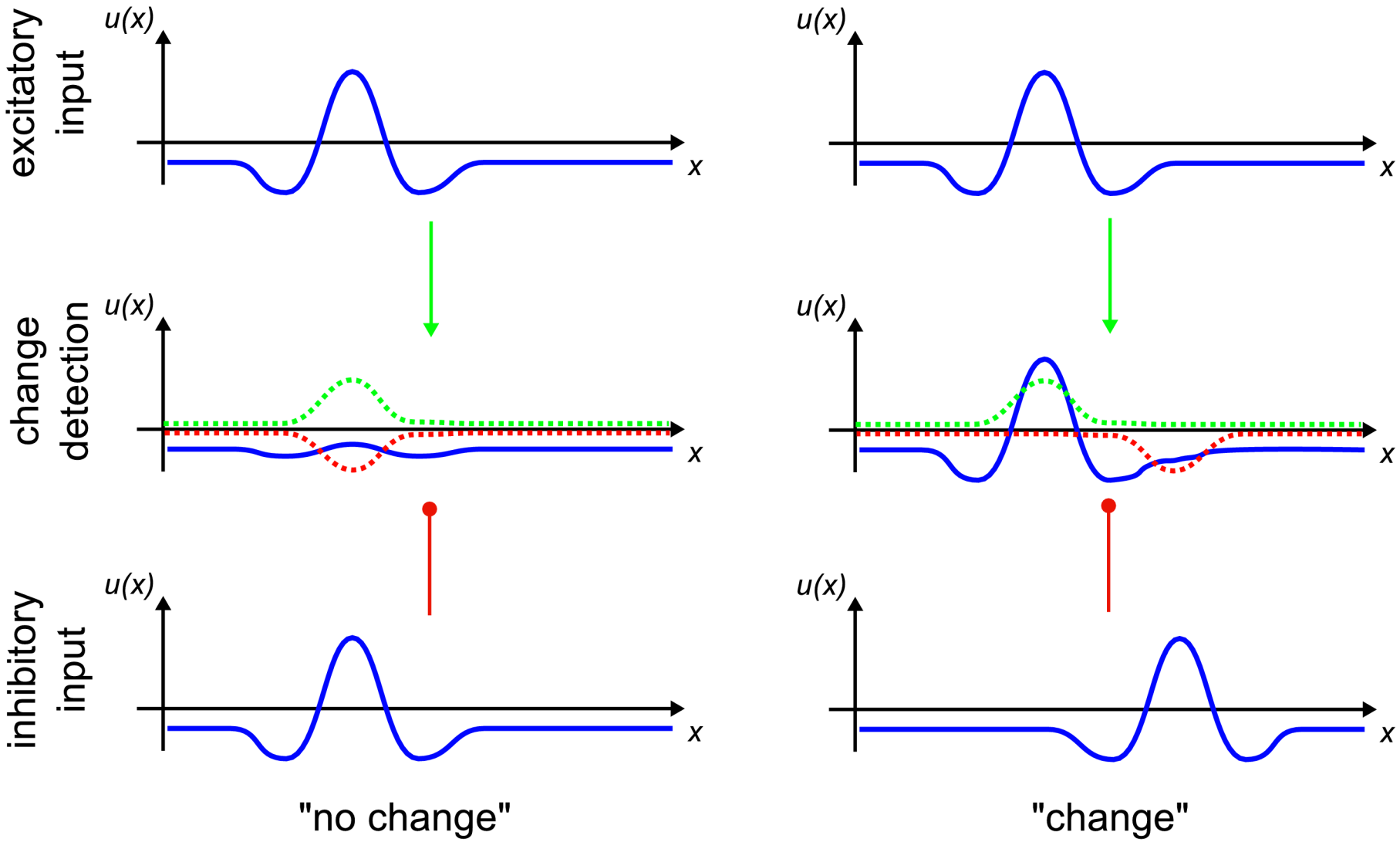
Feature estimates are linked to spatial positions.

Autonomy of Behaviors



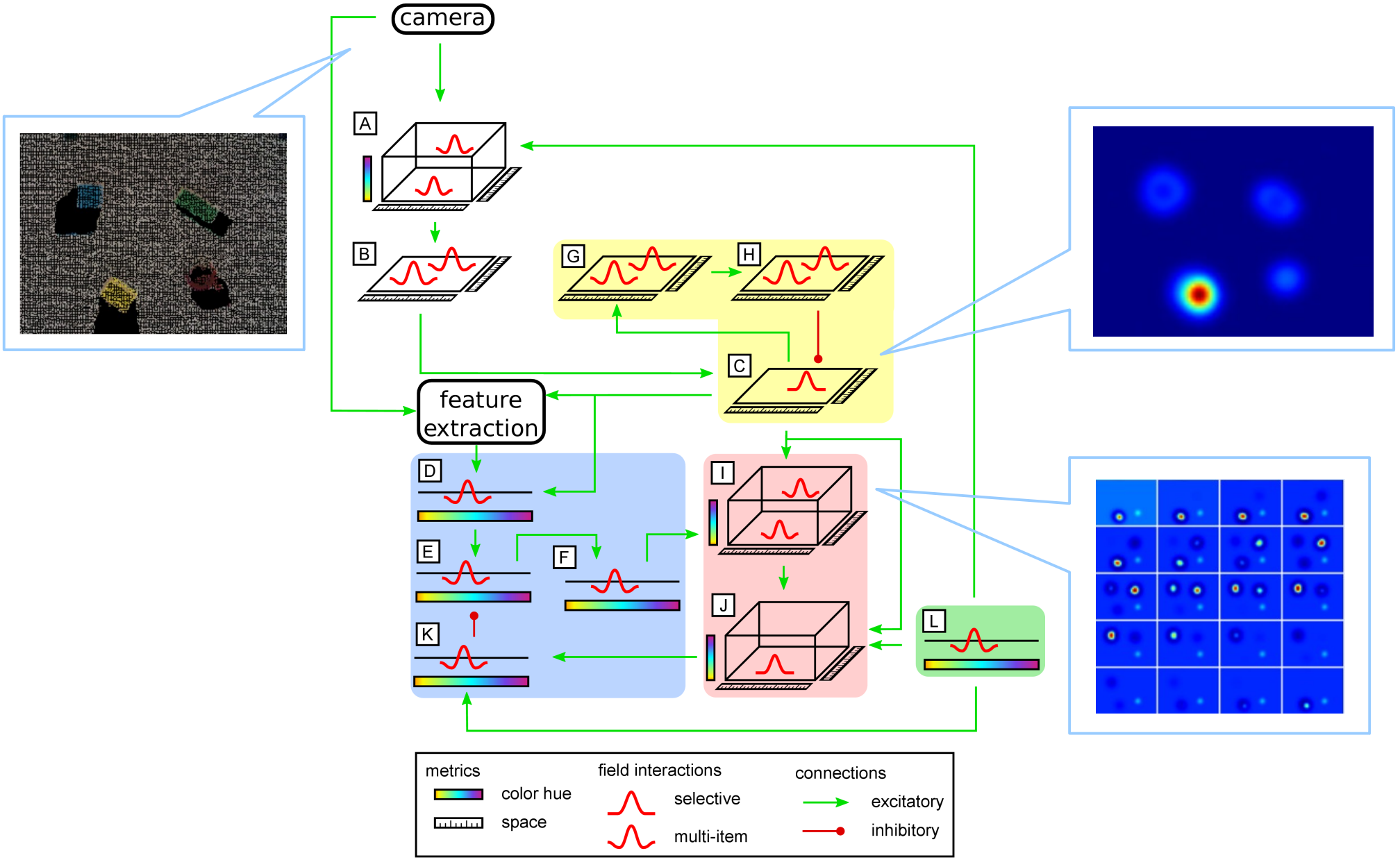
- elementary cognitive units (ECU)
- intention node boosts fields
- CoS node detects completion of behavior in fields
- sequences or exclusion through precondition nodes

Change Detection with Fields

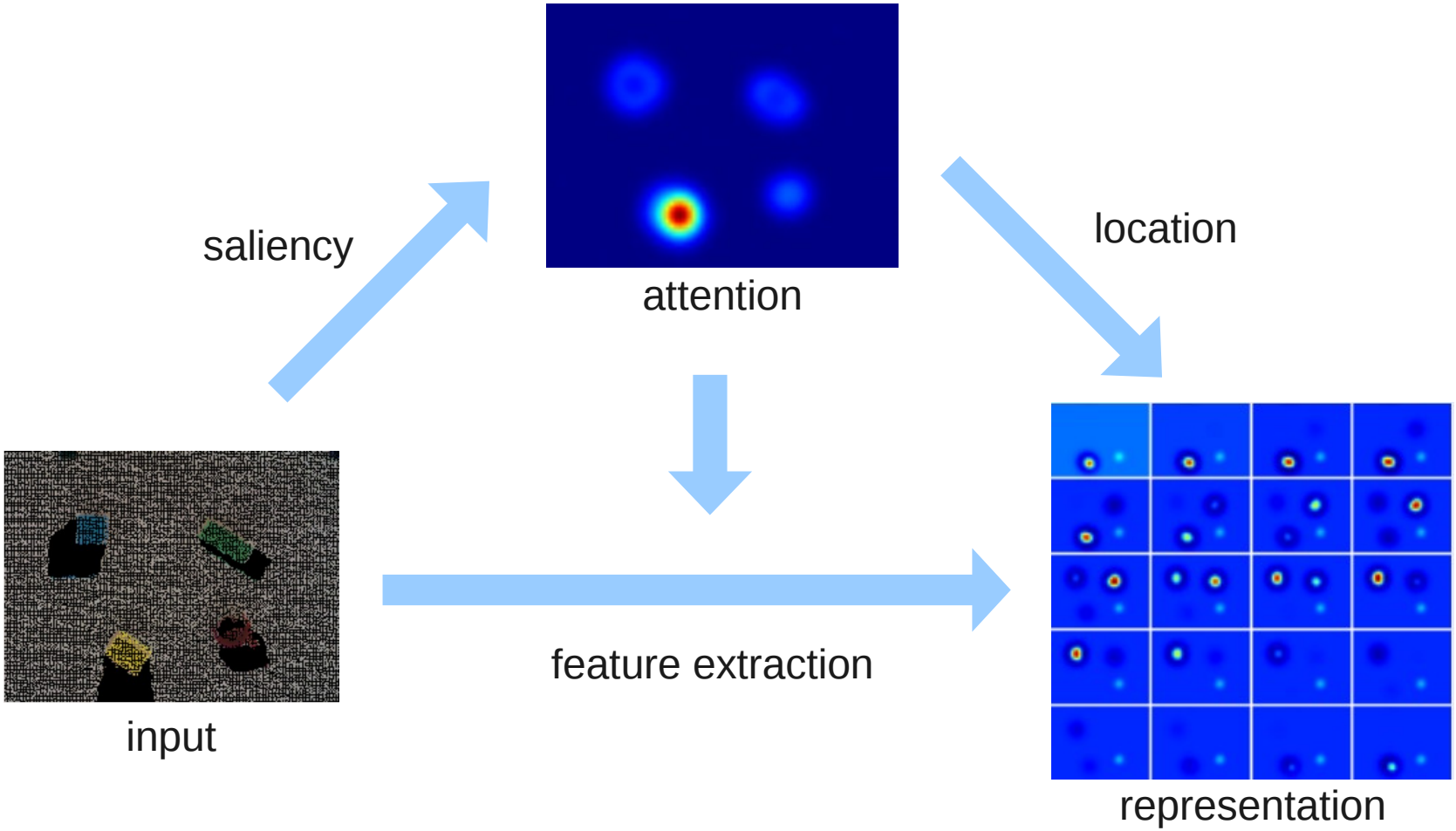


Autonomous Behaviors

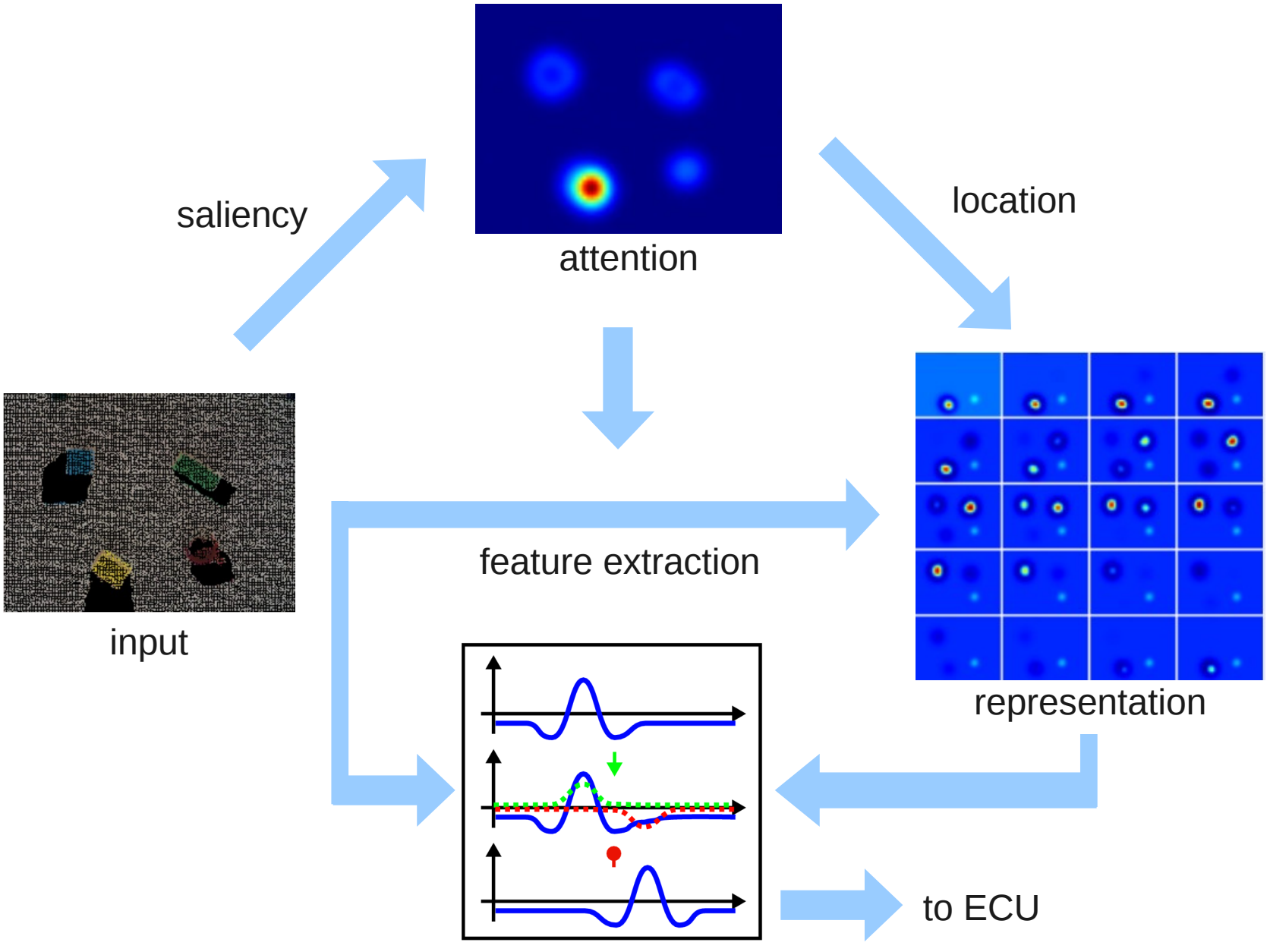
Architecture Overview

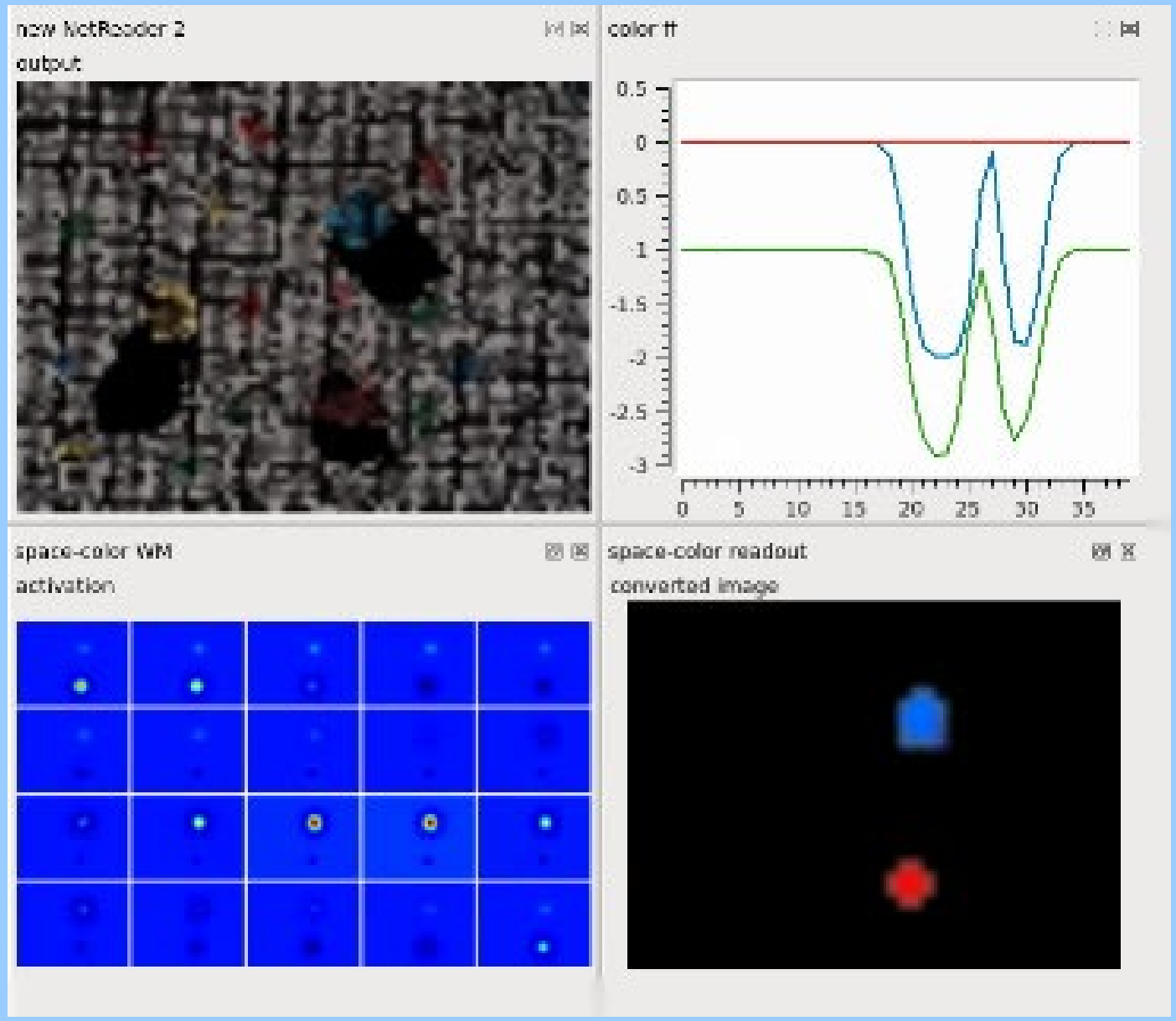


Exploration Behavior



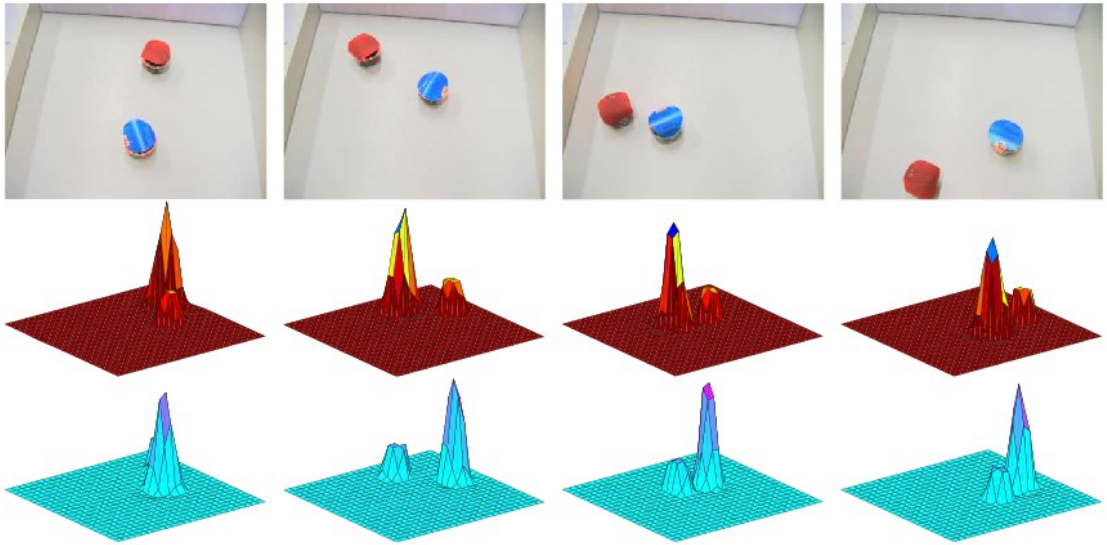
Autonomy of Exploration




new NetReader 2
output
color ff
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
0
5
10
15
20
25
30
35
space-color WM
activation
space-color readout
converted image

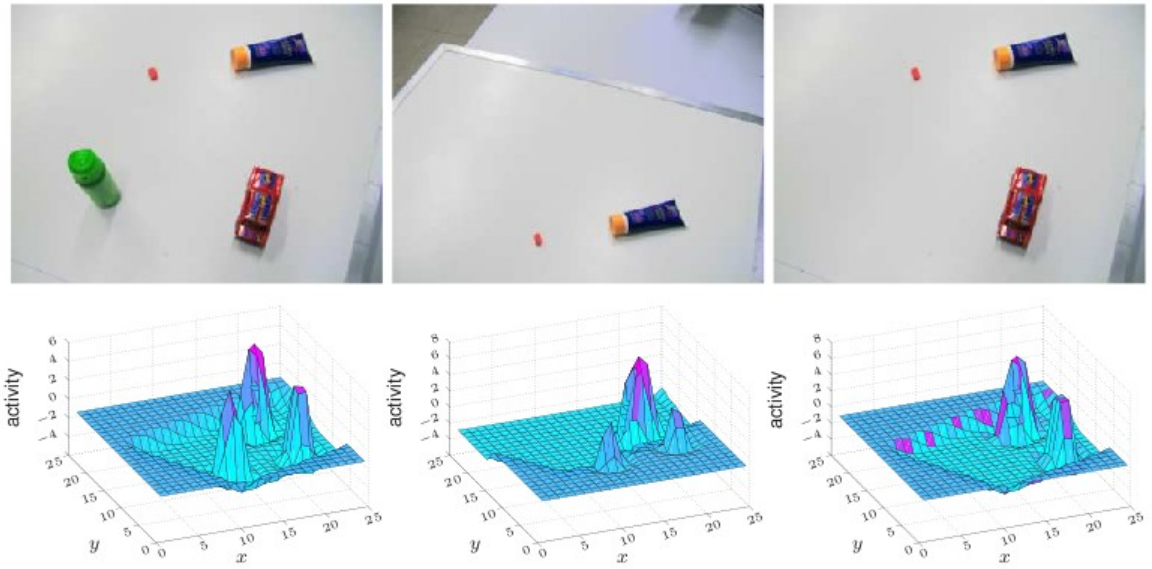
Video

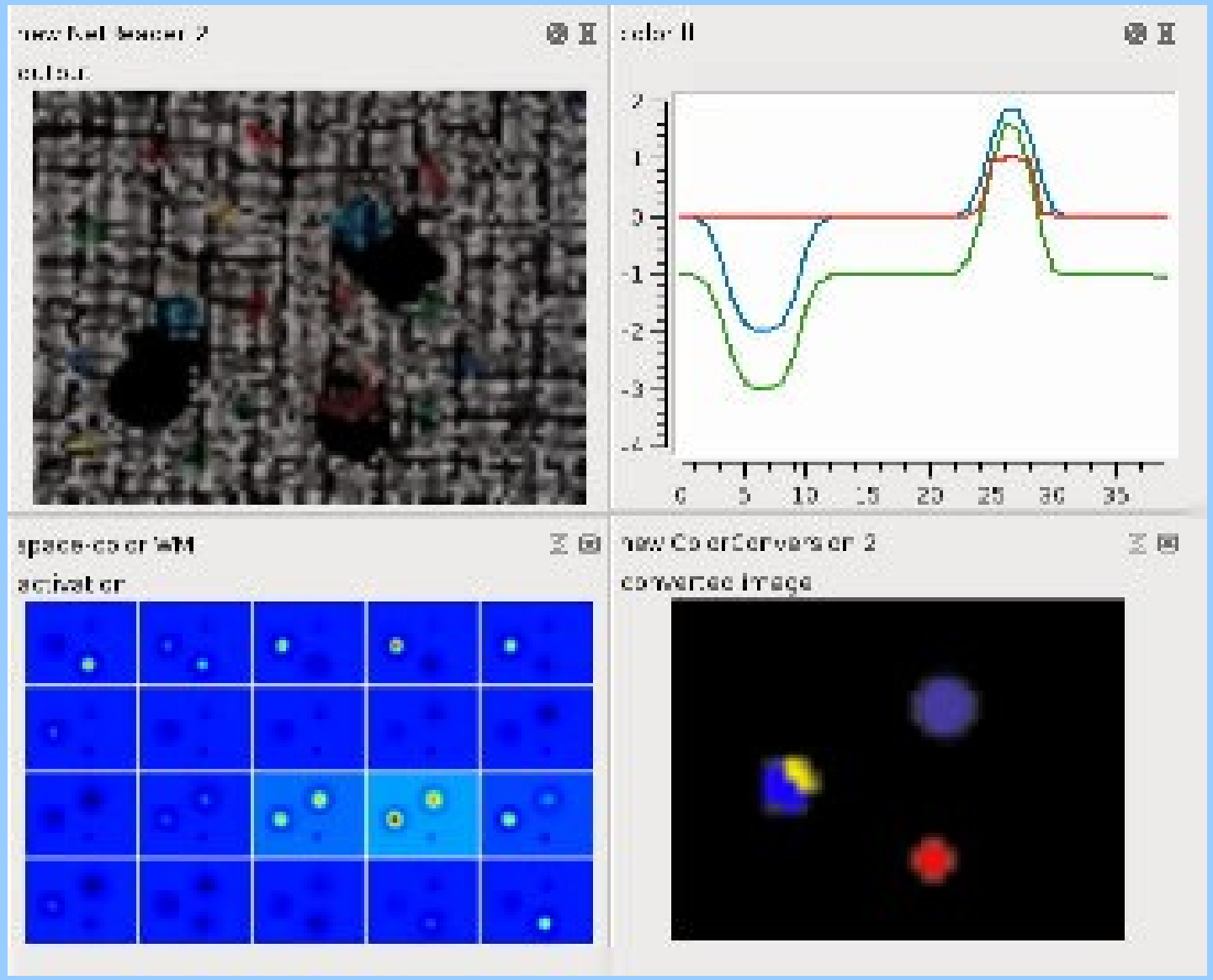
Maintenance Behavior



tracking

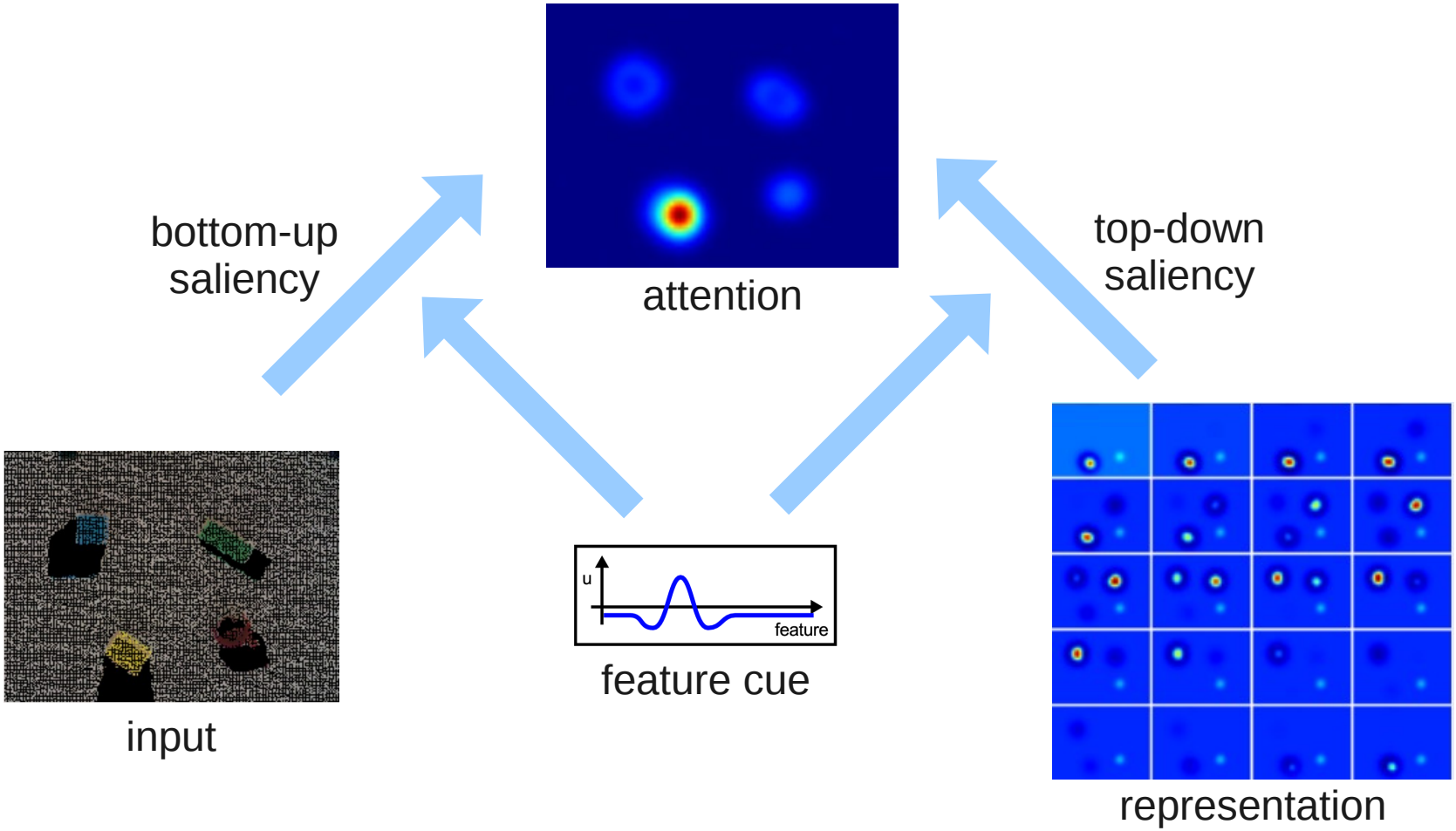
working memory
vs.
updating



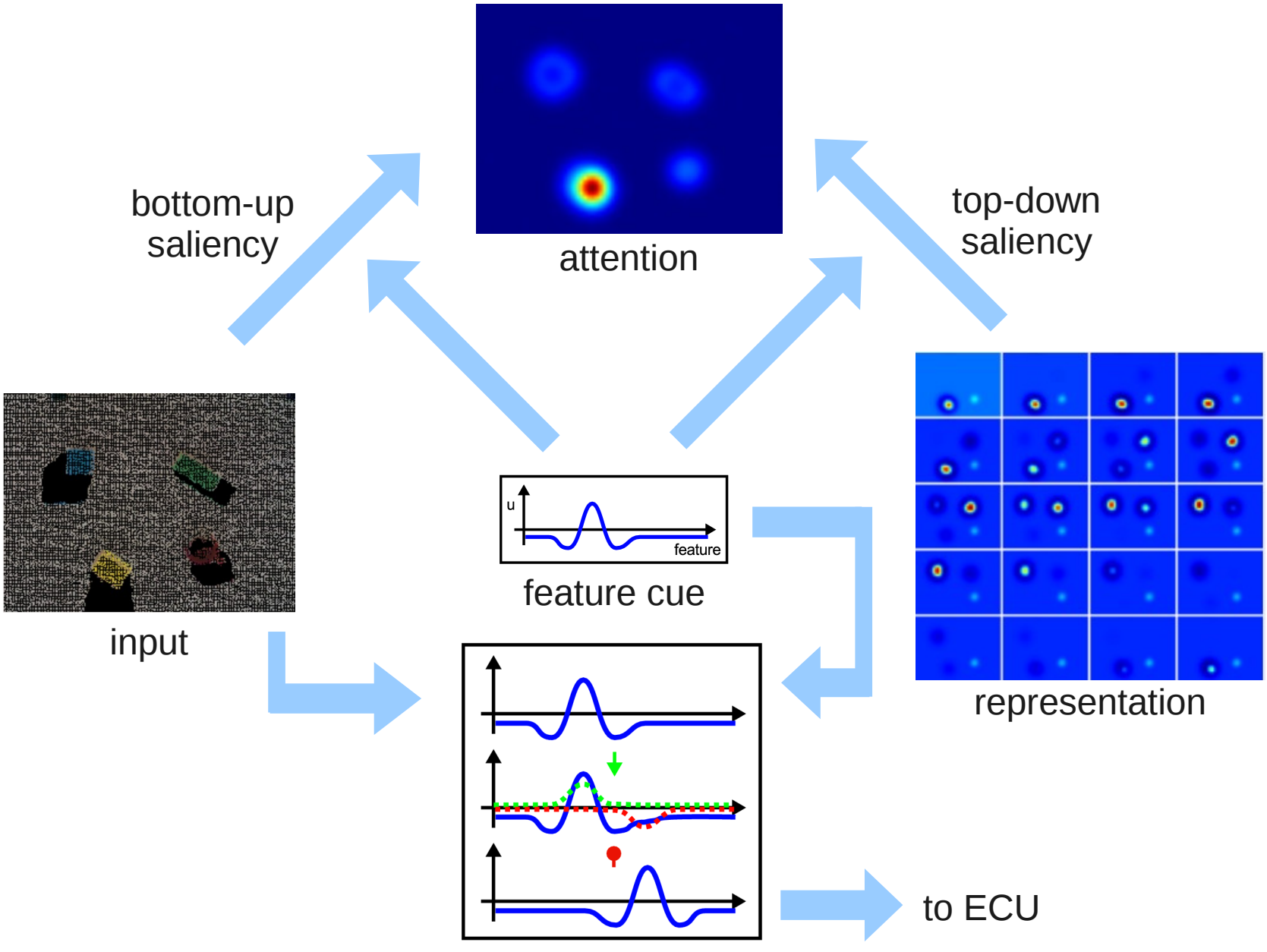


Video

Query Behavior



Autonomy of Query

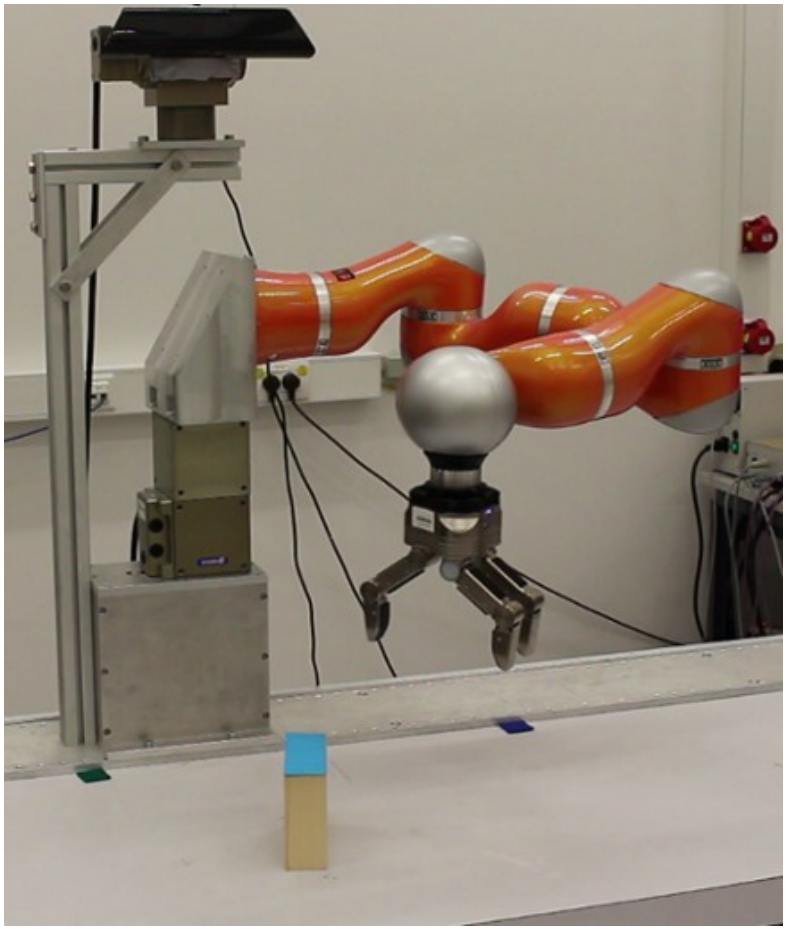


Reaching and Grasping

Challenges

focus attention

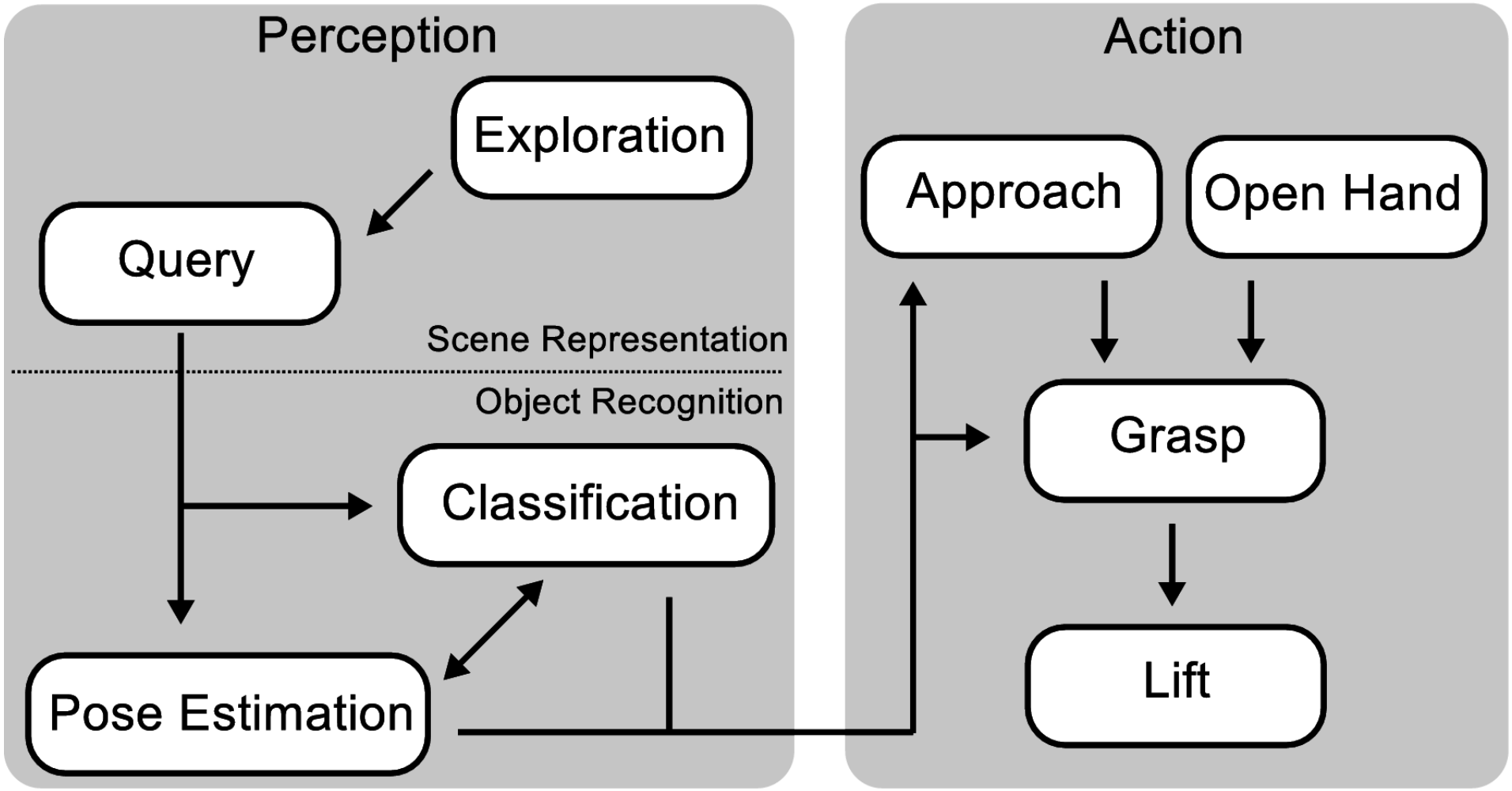
estimate shape



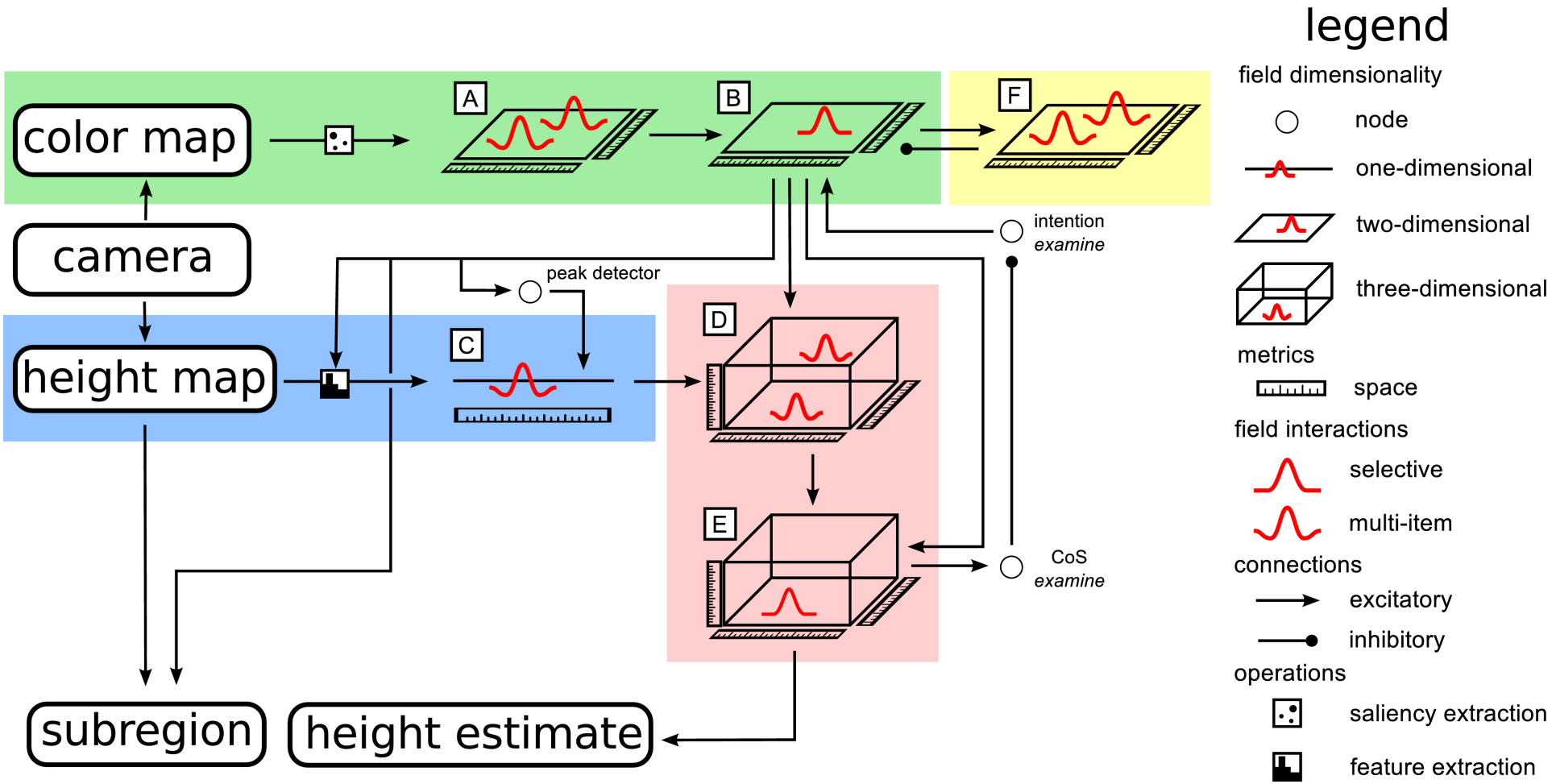
multiple behaviors

arm movement

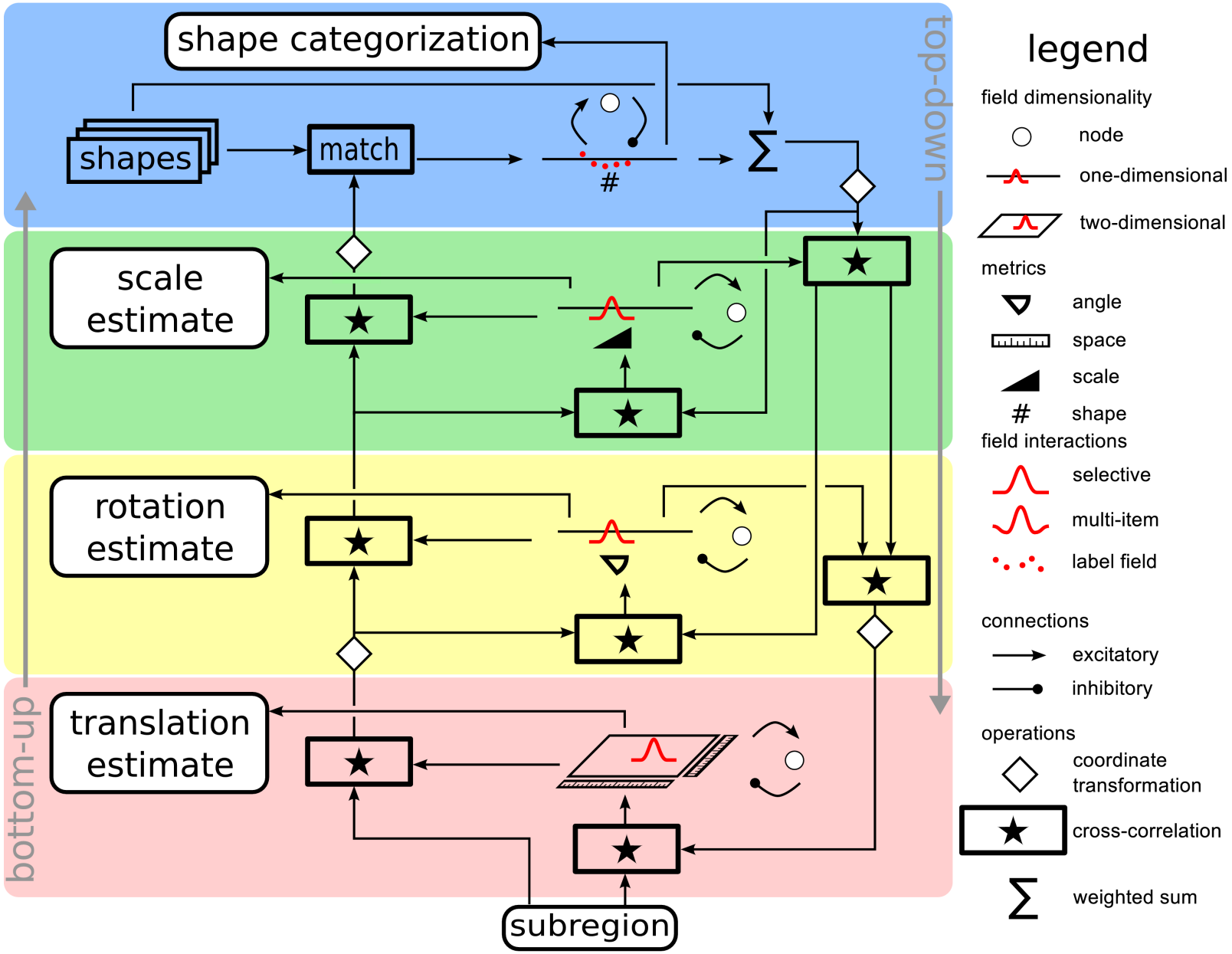
Behaviors Involved in Grasping



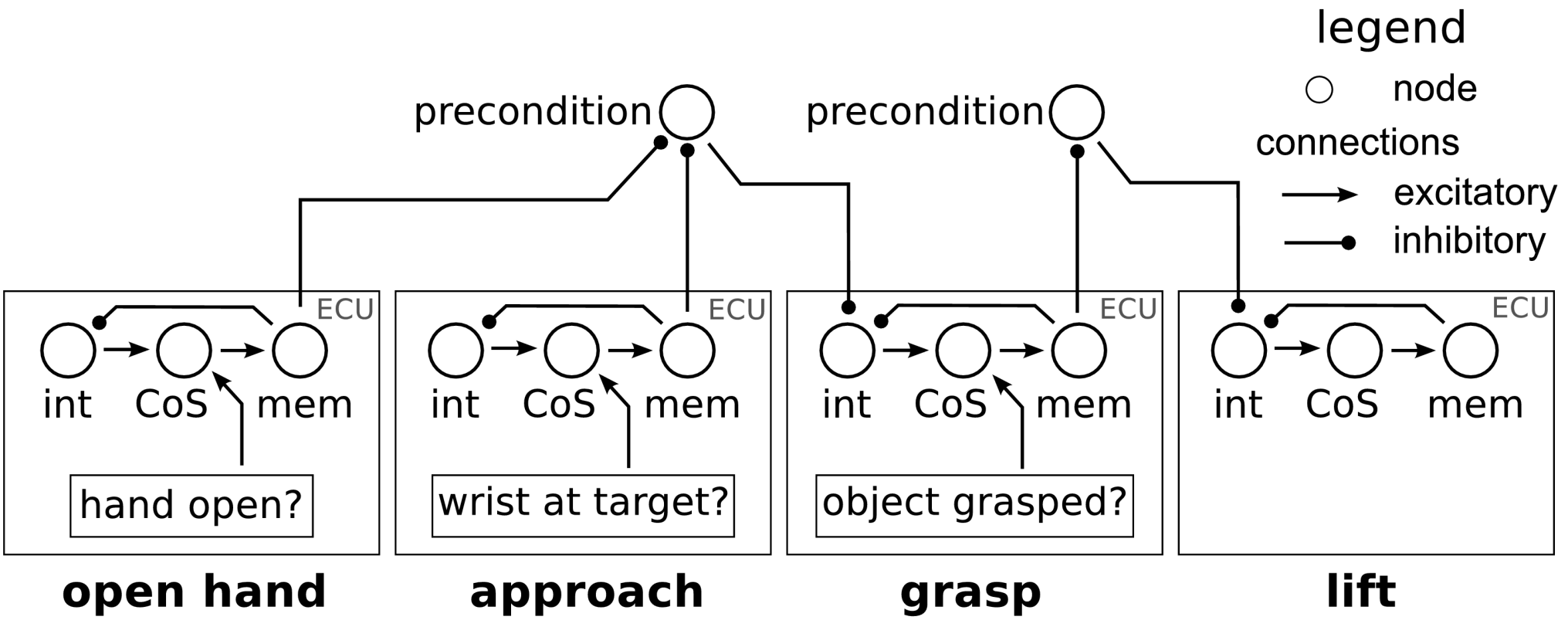
From Camera to a Height Map



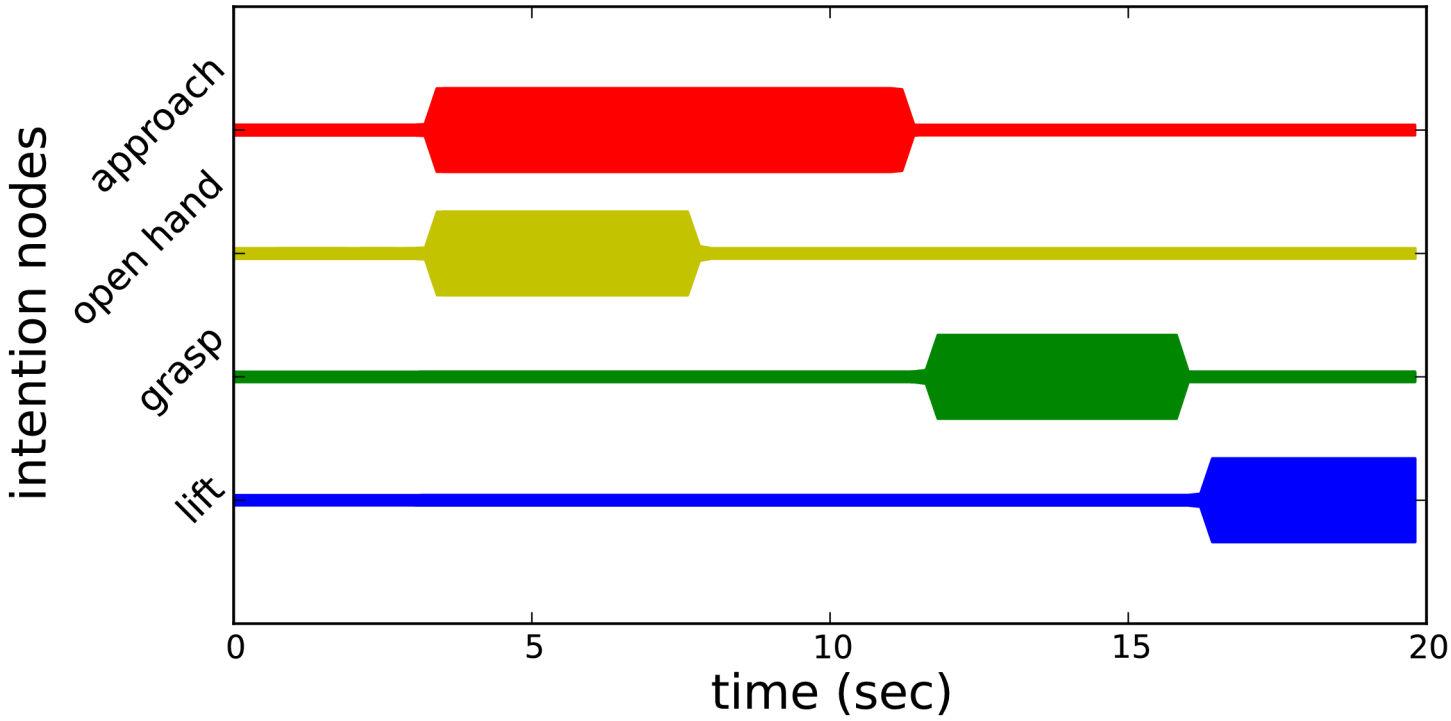
From Height Map to Grasp Parameters

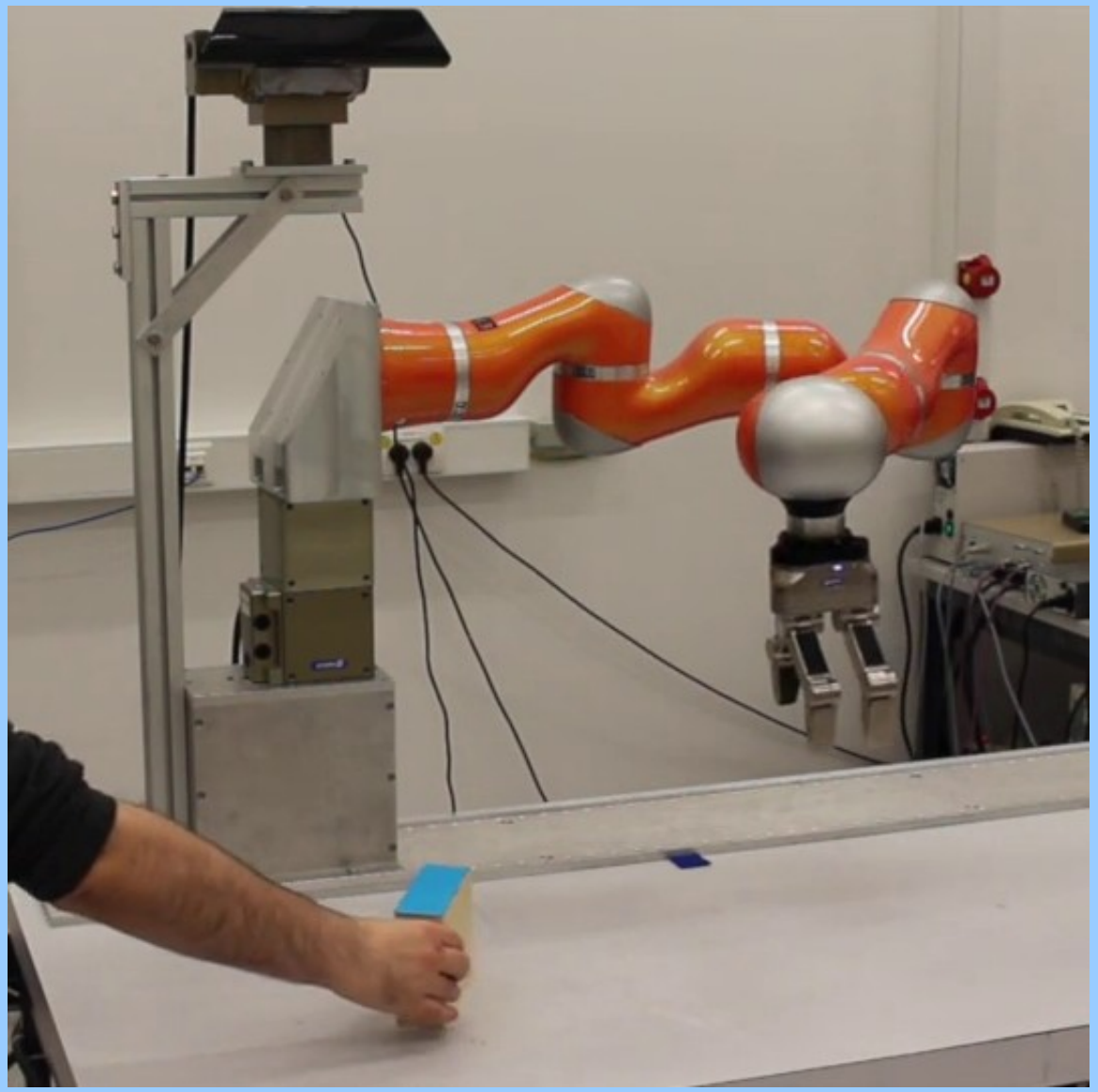


Grasp Execution



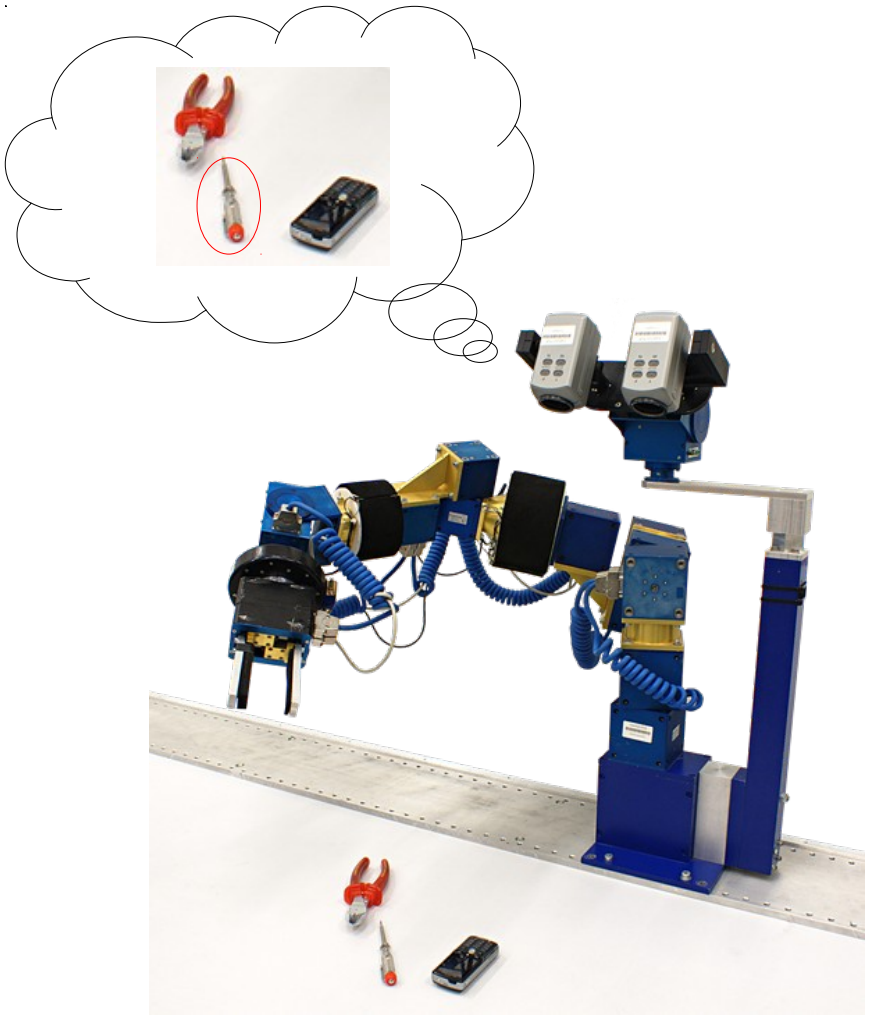
Grasp Execution





Video

Take-home Message



- exploration, maintenance and query are the core behaviors of scene representation
- change detection is a driving force for autonomy
- integration with other DFT architectures yields complex behaviors such as grasping
- integration is facilitated by DFT framework

Thanks for your attention!