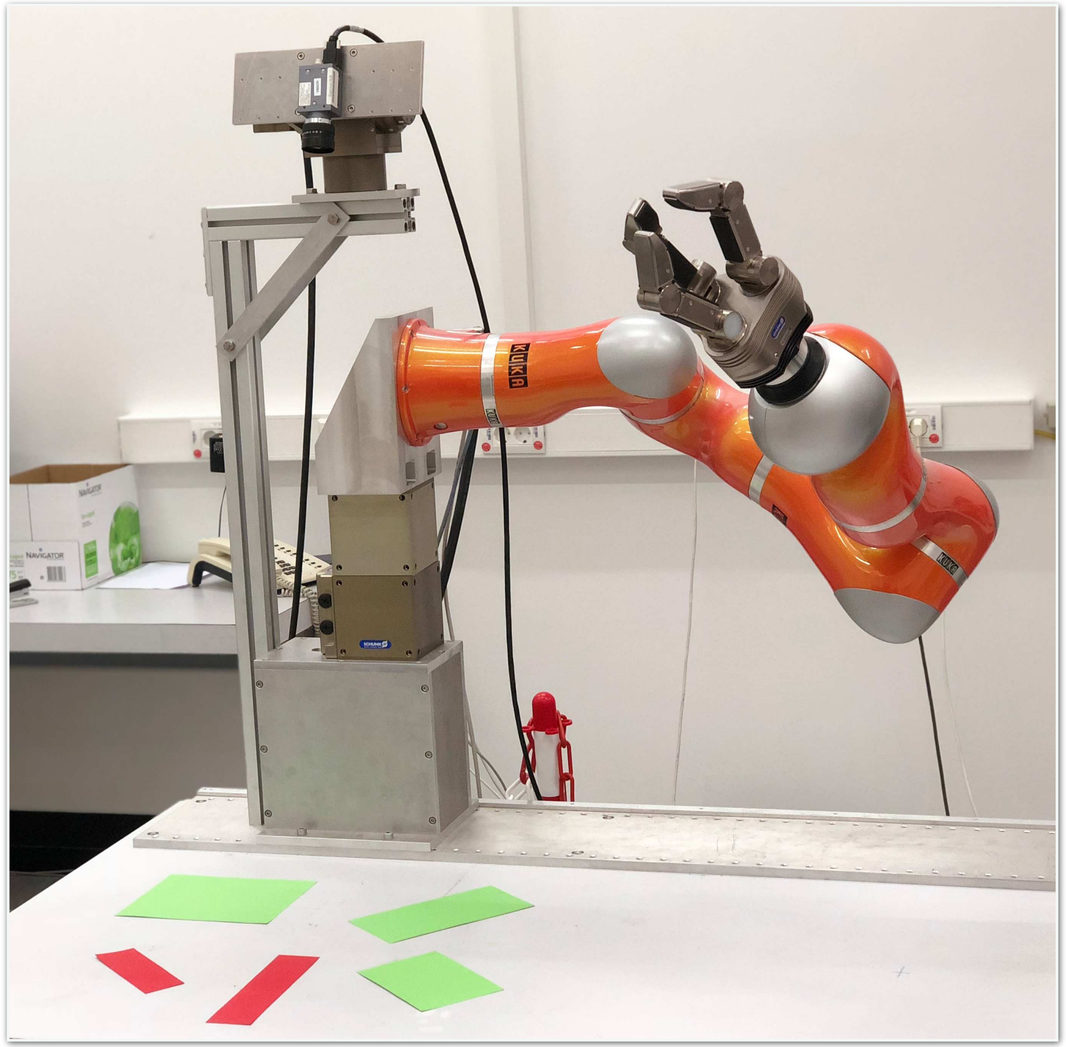
Scene Representation

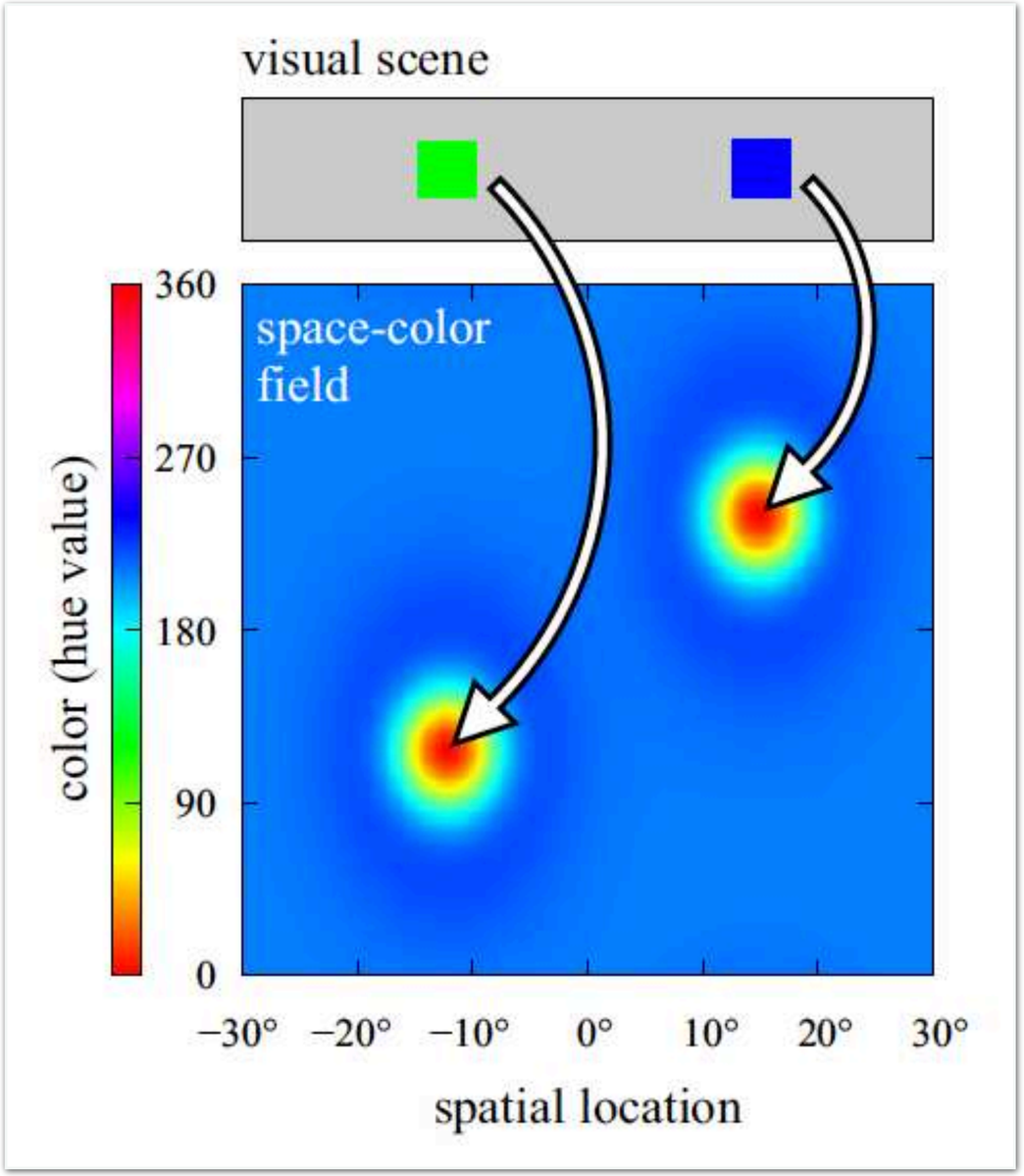
Raul Grieben

Scene Representation



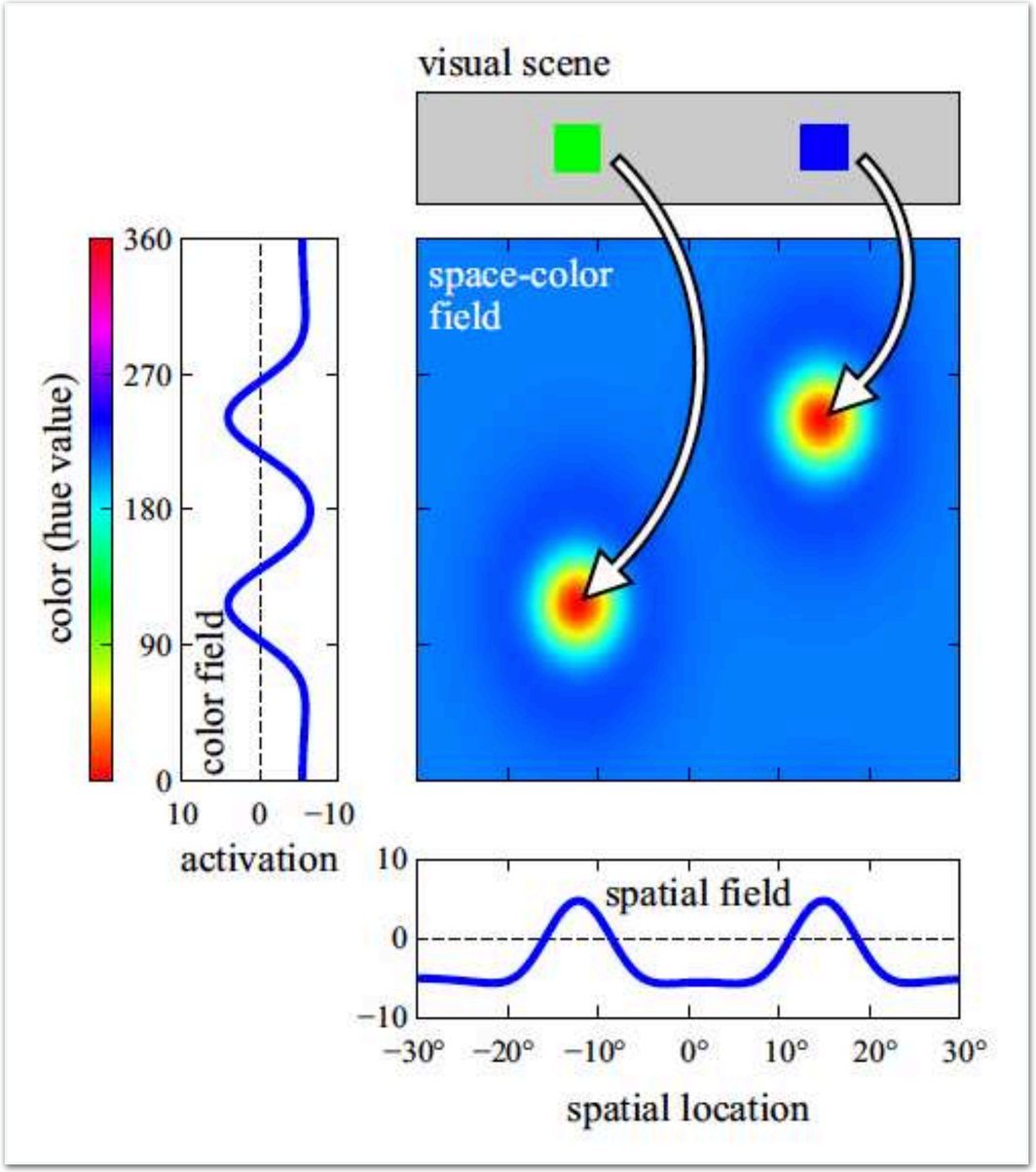
- internal representation of environment
- foundation for every higher cognitive operation and action
- stable despite eye and body movements
- limited capacity, link to long-term memory

Feature Binding



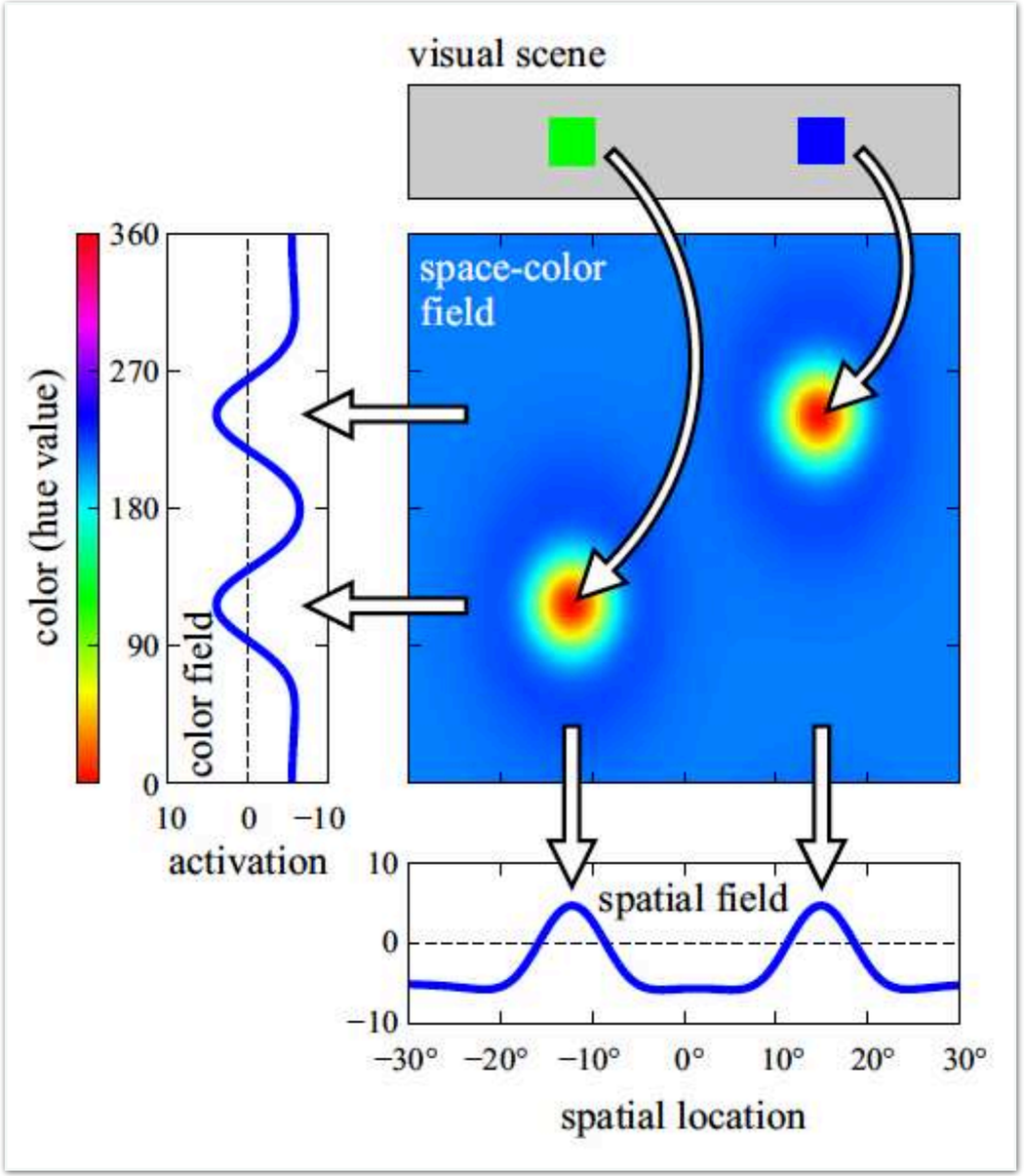
- 1D spatial location (for illustration)
- 1D color dimension (hue)
- visual input: 2D
- => 2D peaks

2D input



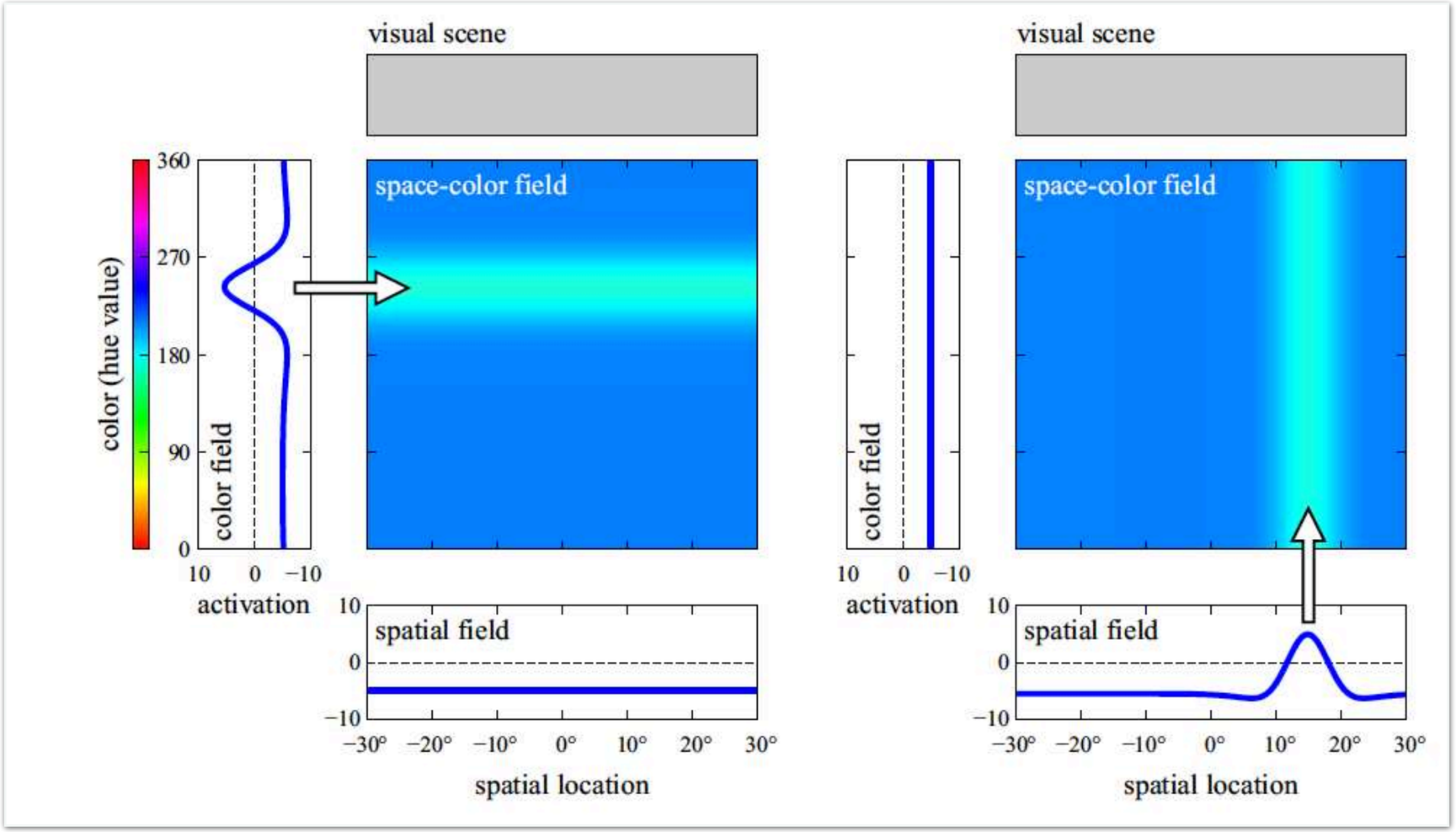
- creates 2D peaks that form combined (bound) representations of objects

extracting features

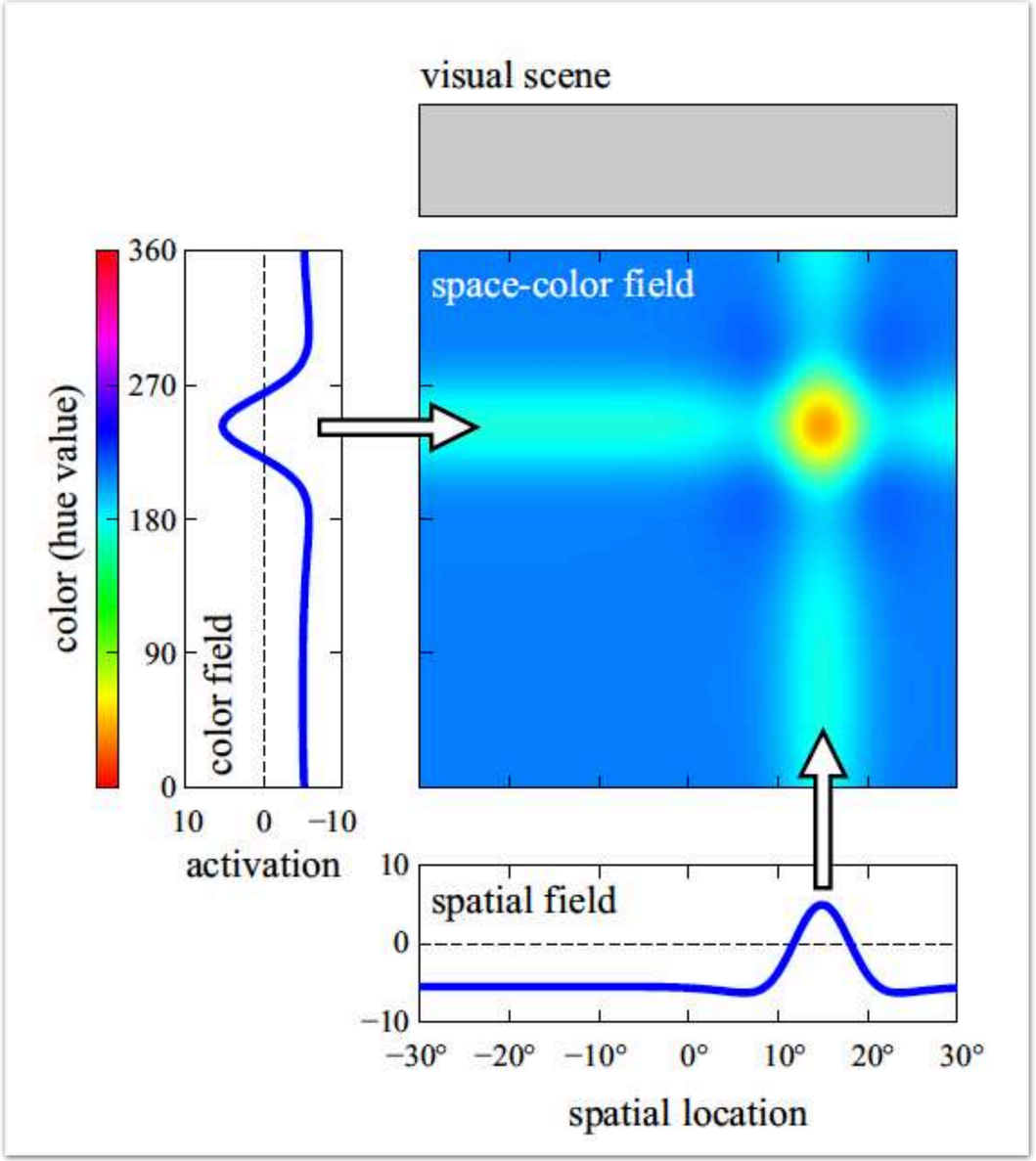


- read-out from 2D to 1D by projection
  - by summing along the other dimension (marginalization)
  - or by taking the (soft)max

assembling bound representations

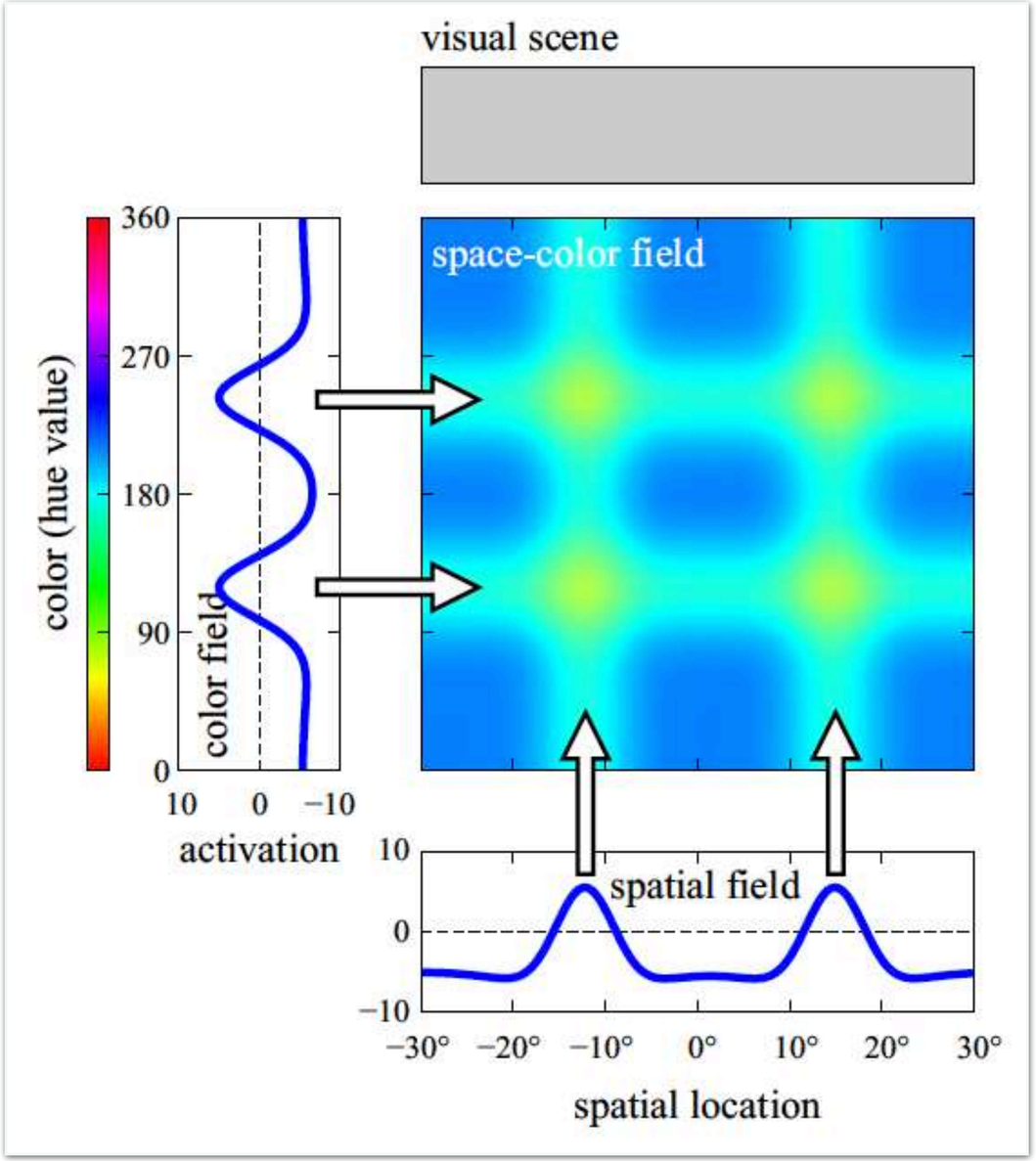


assembling bound representations



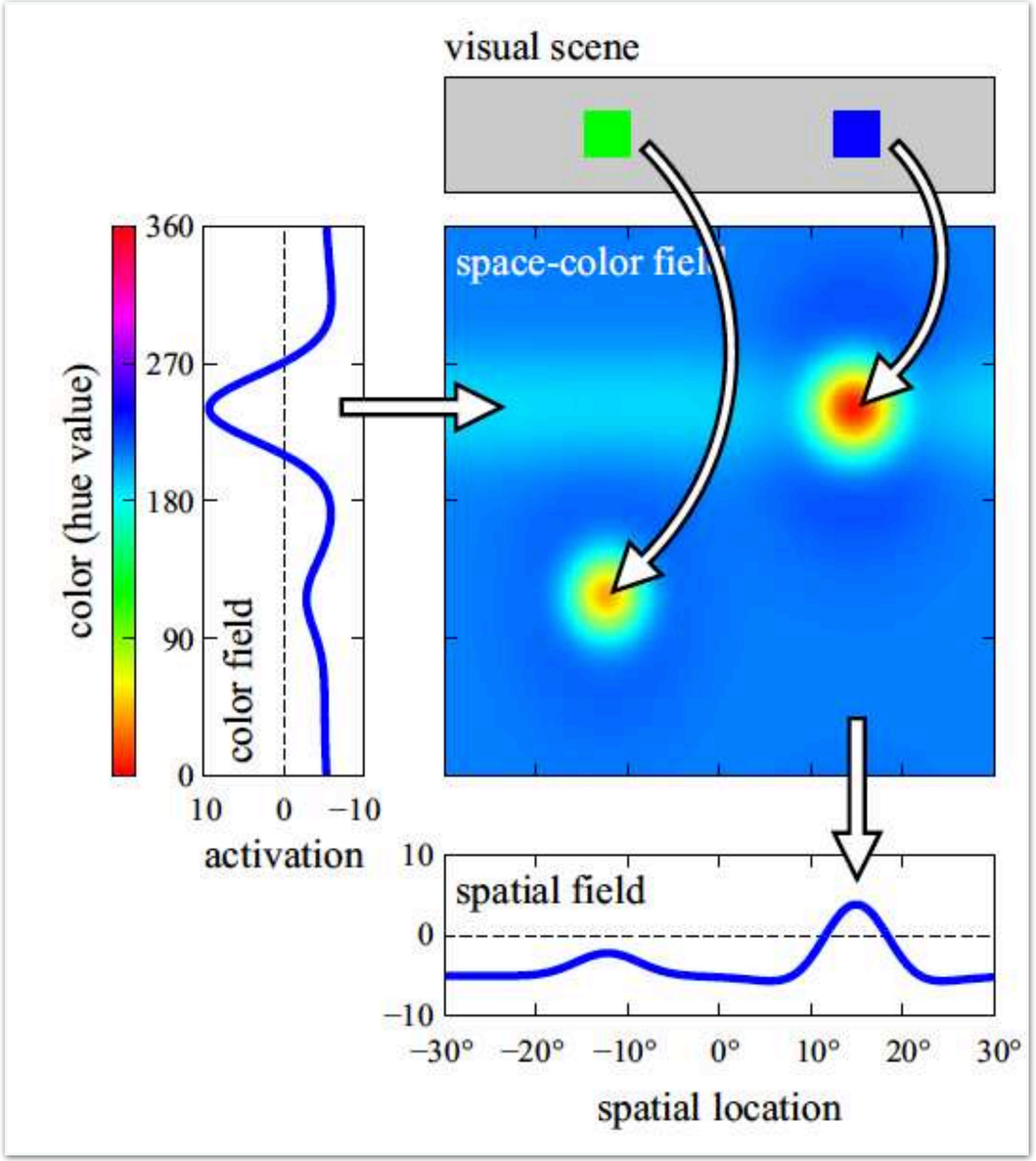
- peaks form at the intersections of ridges and form bound representations of the two dimensions

assembling bound representations



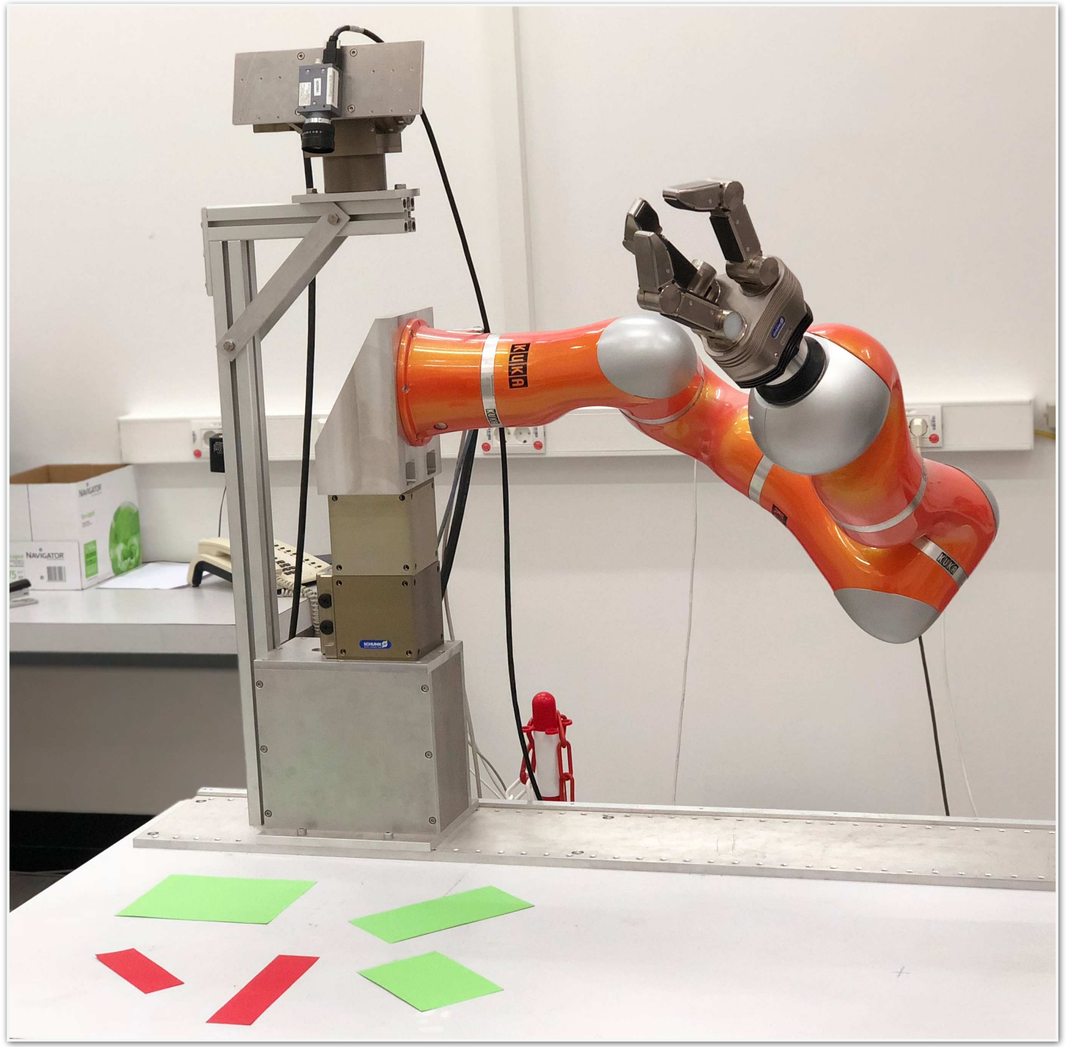
- binding problem: multiple ridges lead to a correspondence problem
- => assemble one object at a time
  - sequentiality bottleneck

visual search



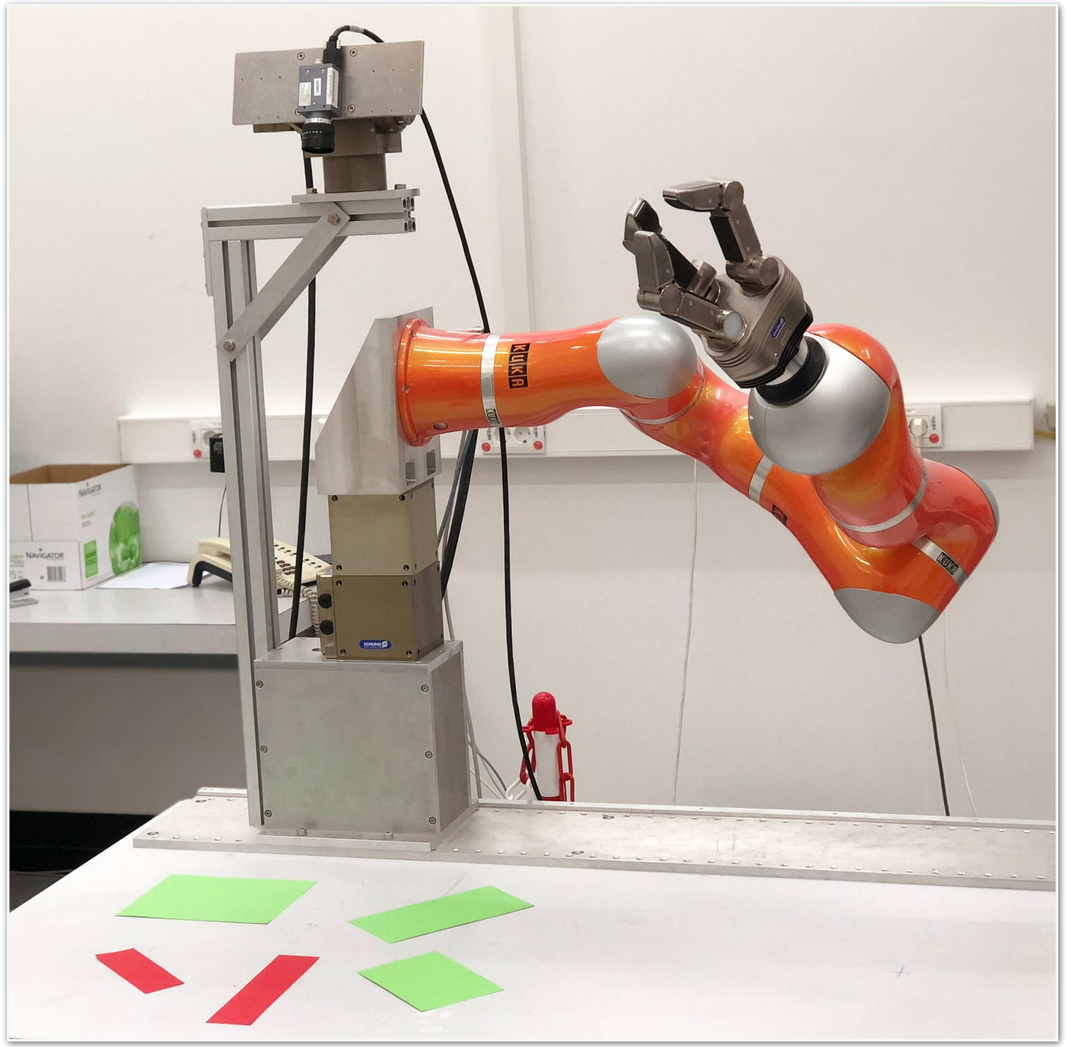
- combine 1D (ridge) input with 2D input
- so that only those 2D locations can form peaks that overlap with ridge (boost driven detection)
- activate objects consistent with 1D feature value

Robotic Scenario



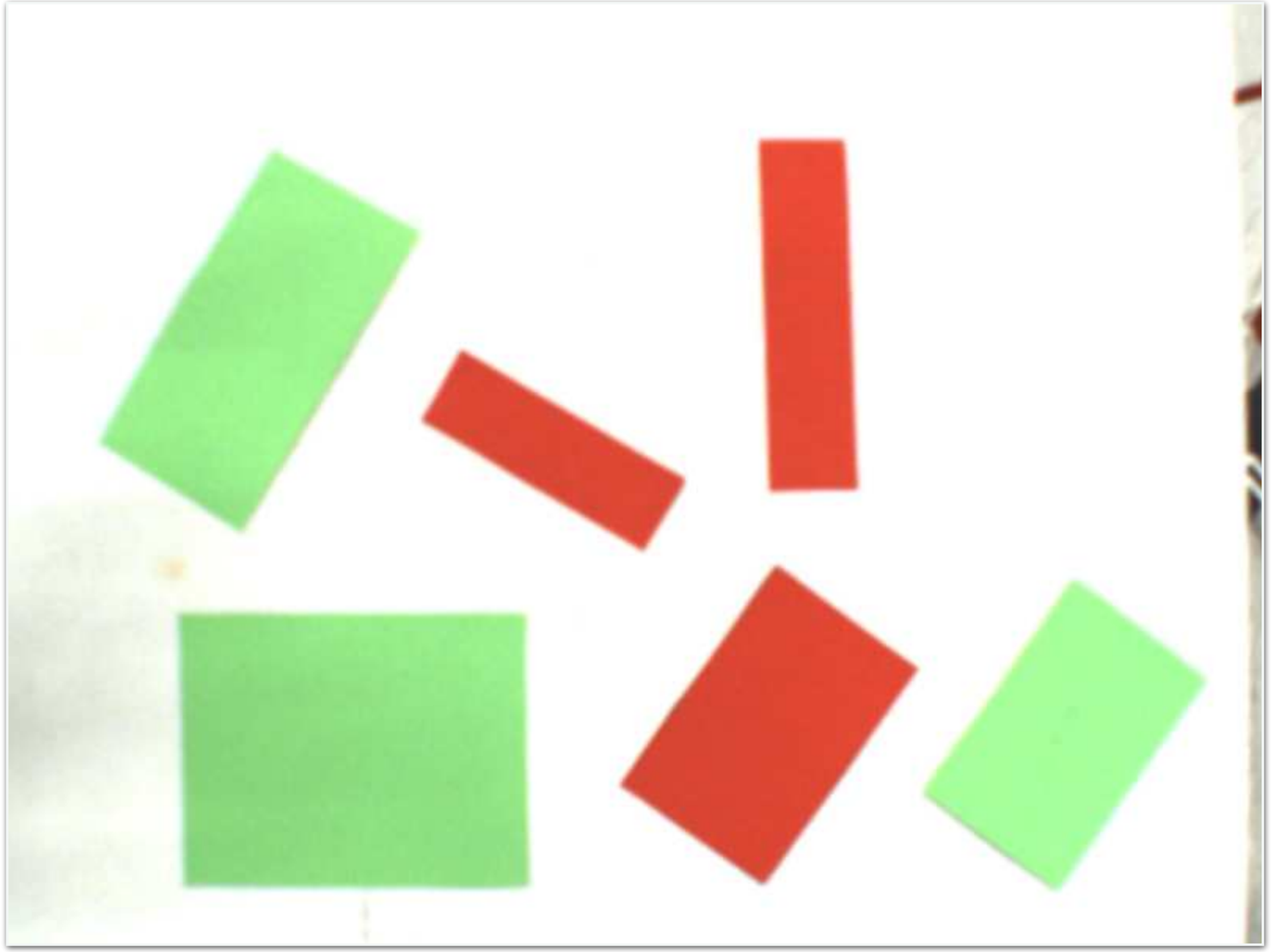
- apply to table-top scenario and human interaction
- use the internal representation for behavior generation
- interact with humans

Robotic Scenario: Behaviors

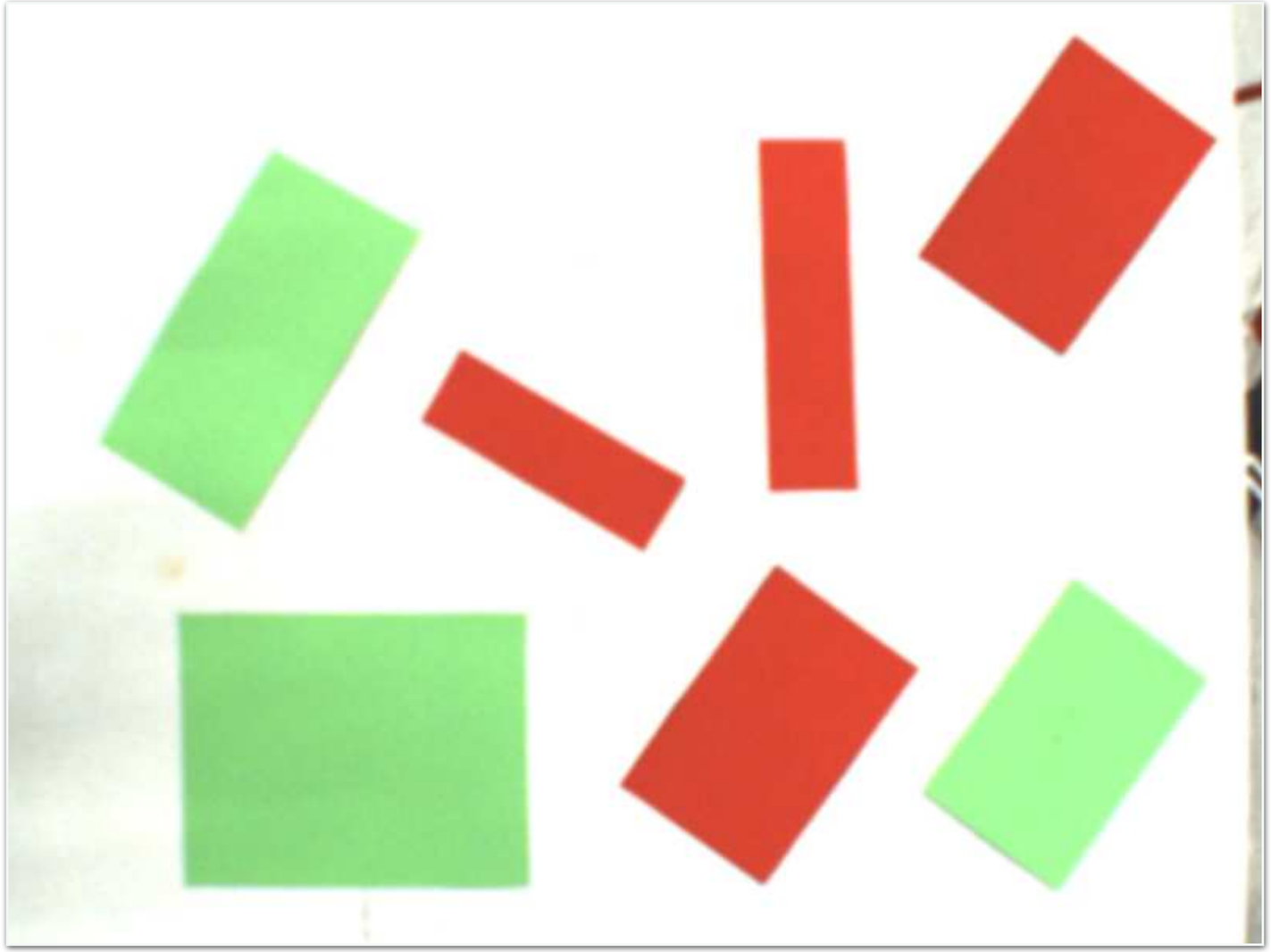


- explore the environment and store objects and their features internally
- maintain the internal representation
- query the representation to create autonomous action

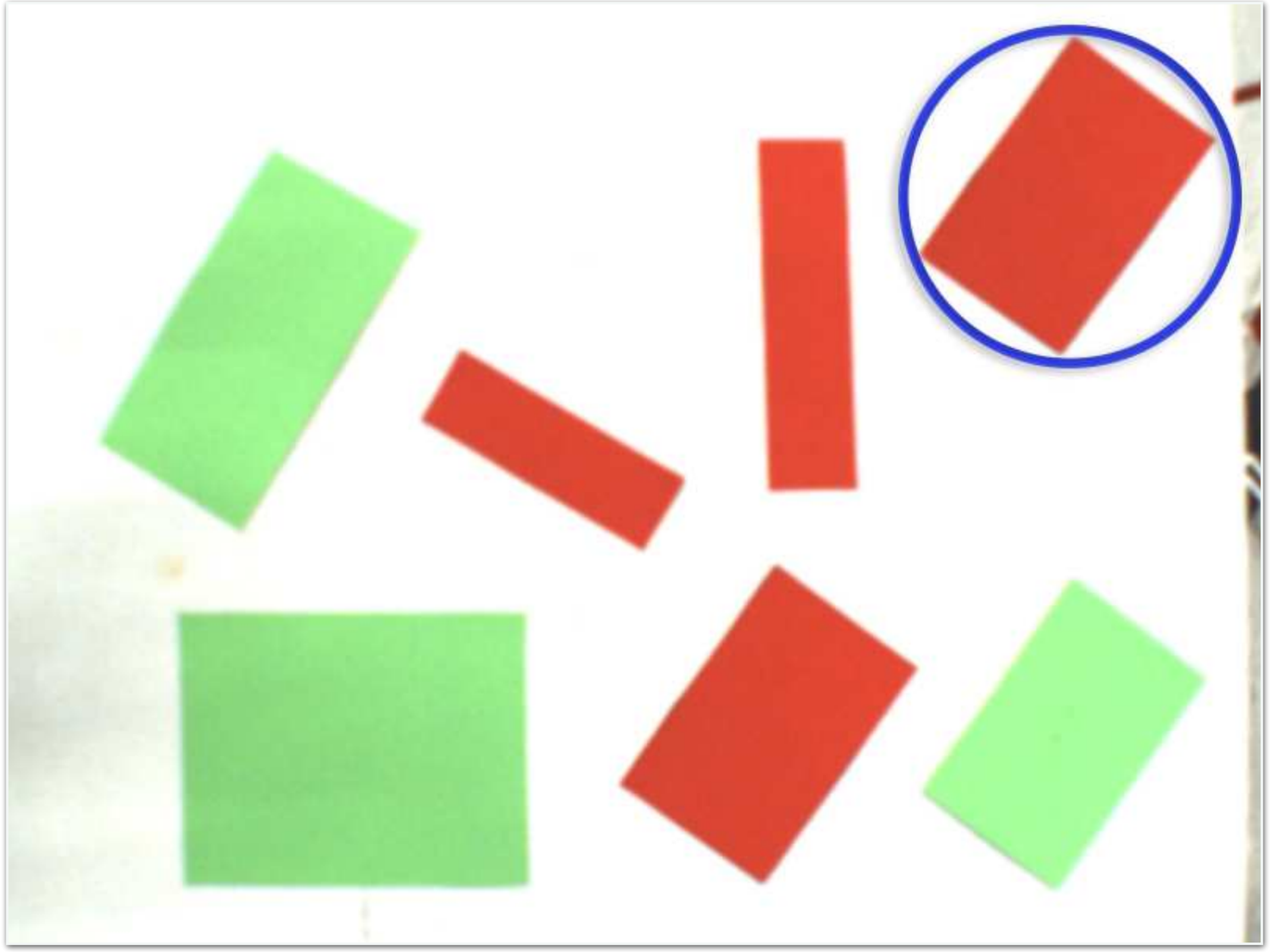
Scenario: Conjunctive Visual Search



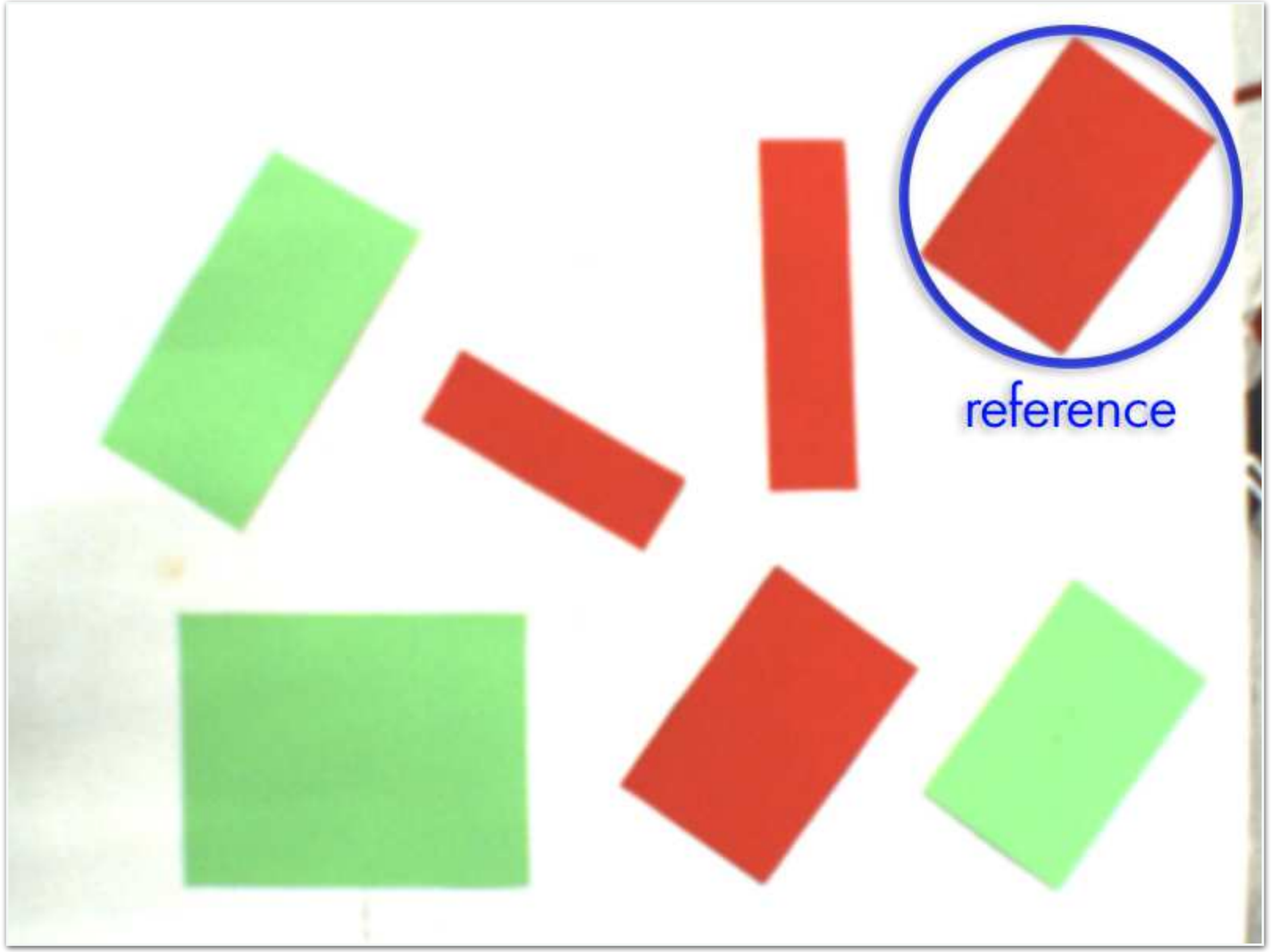
a new object is added to the scene



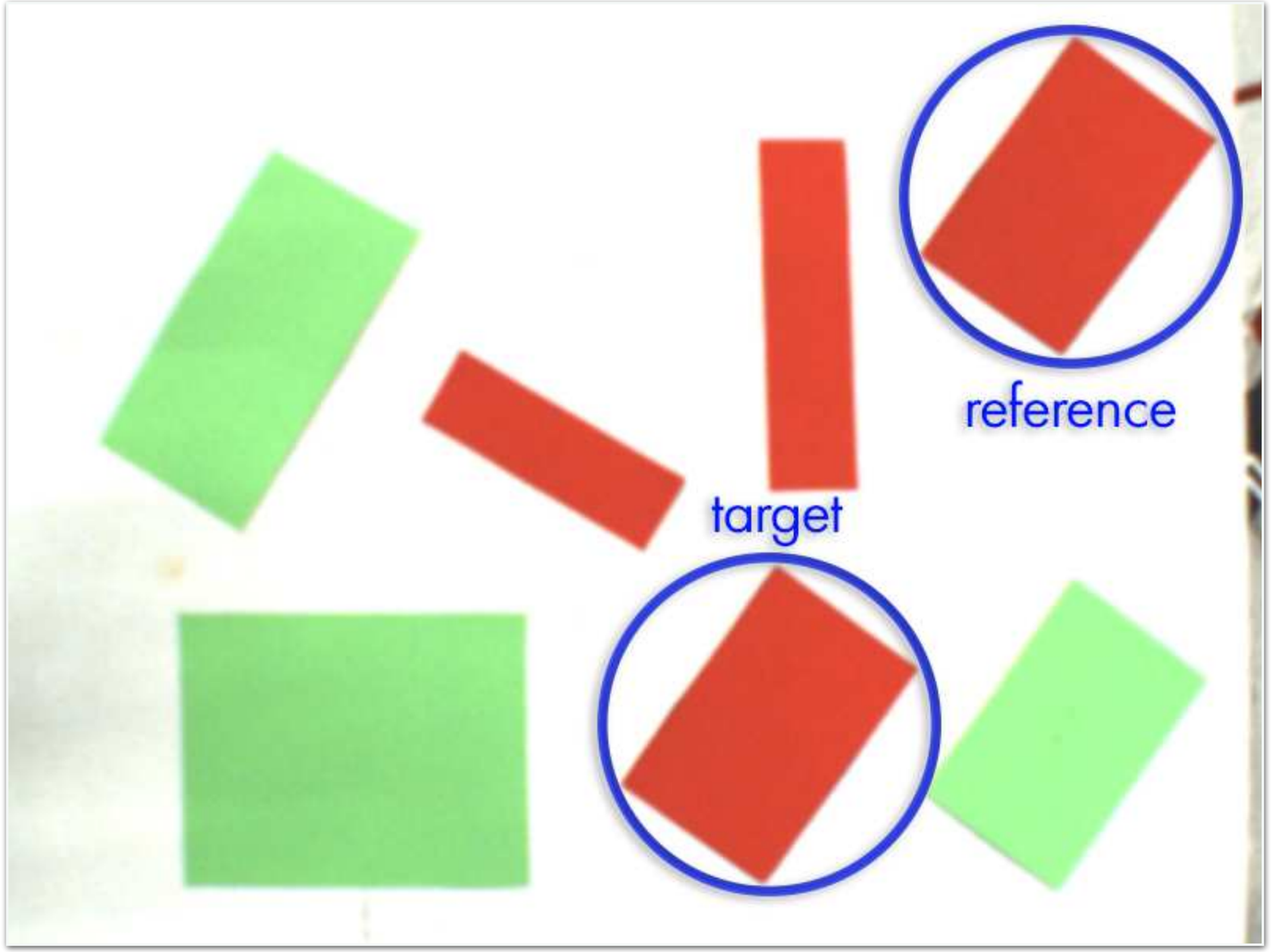
... and detected



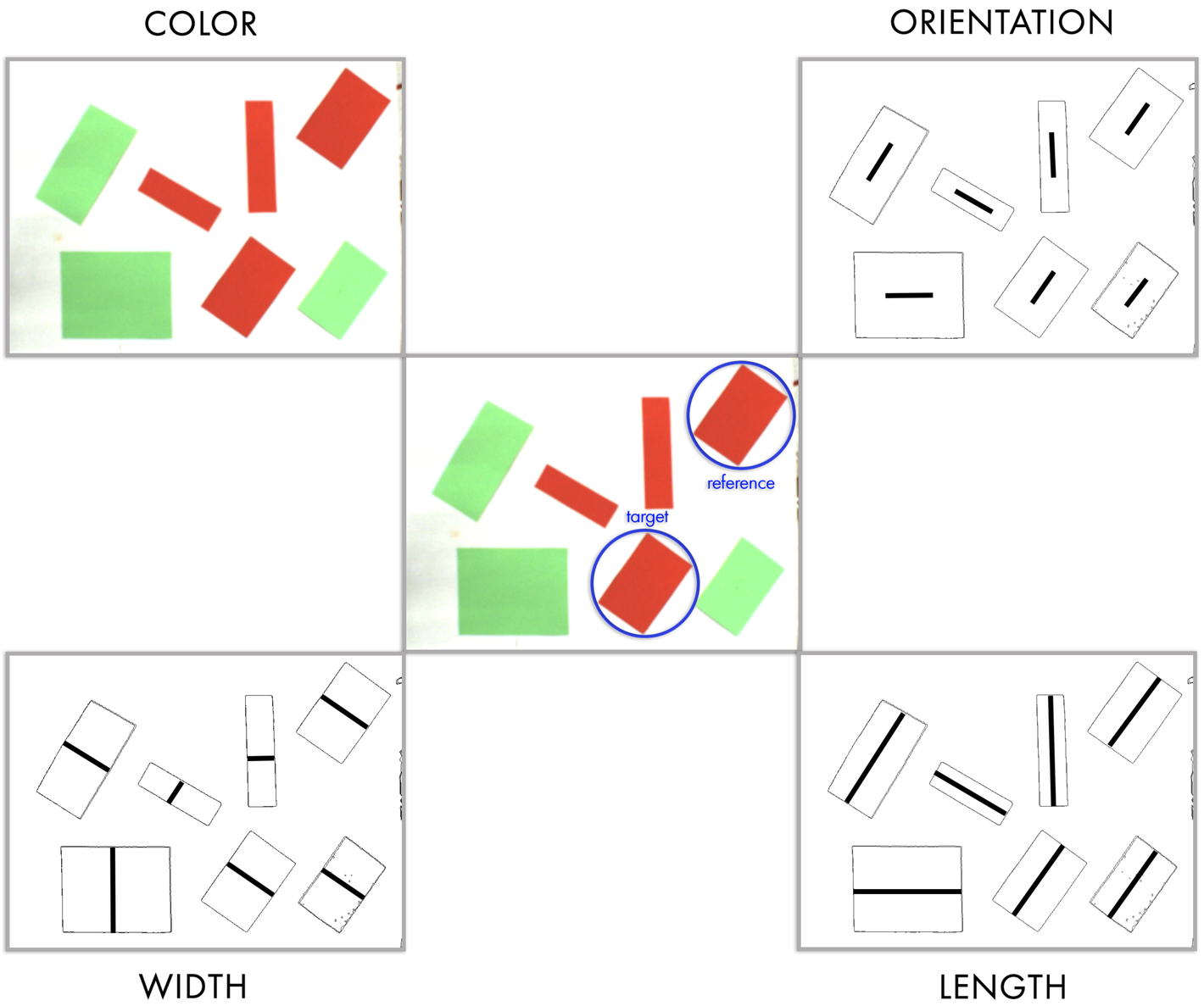
it will serve as the reference object



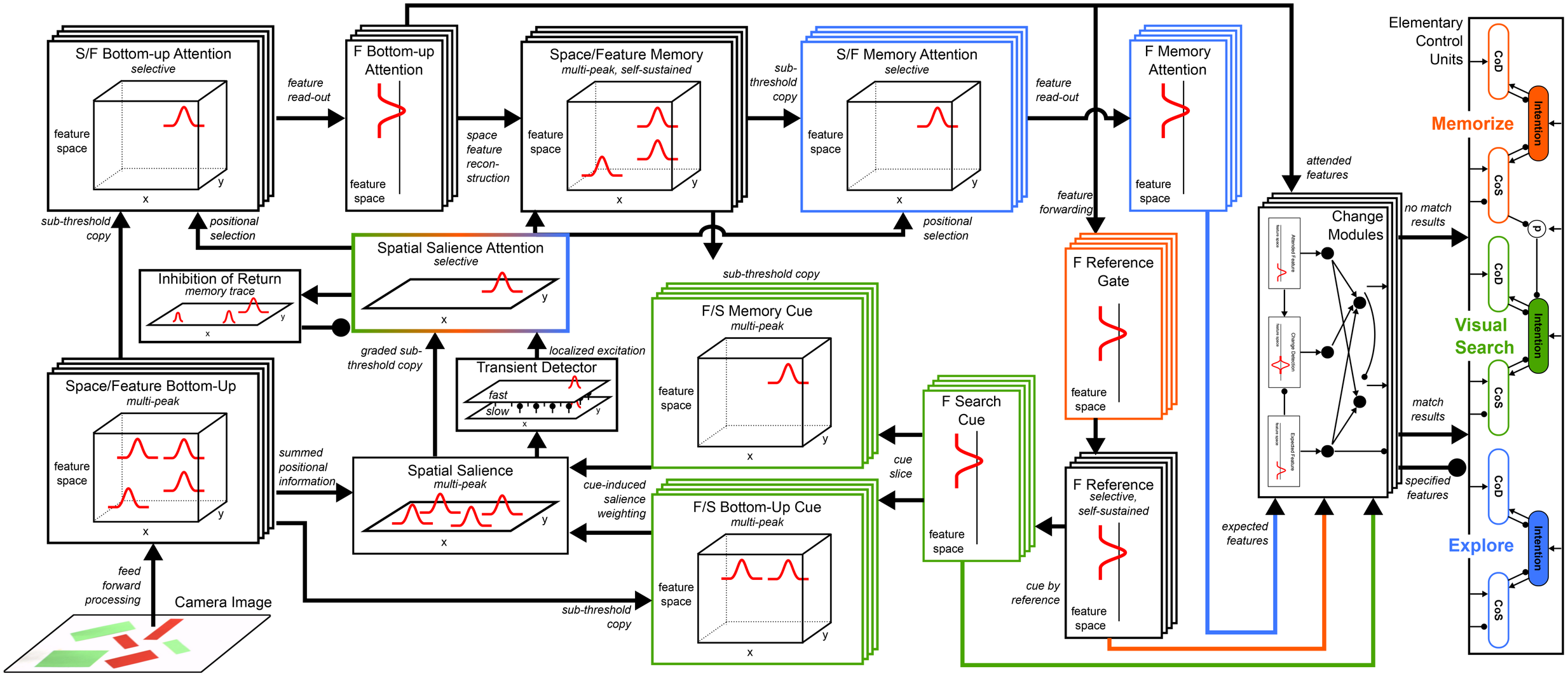
and the target object is brought into the attentional foreground



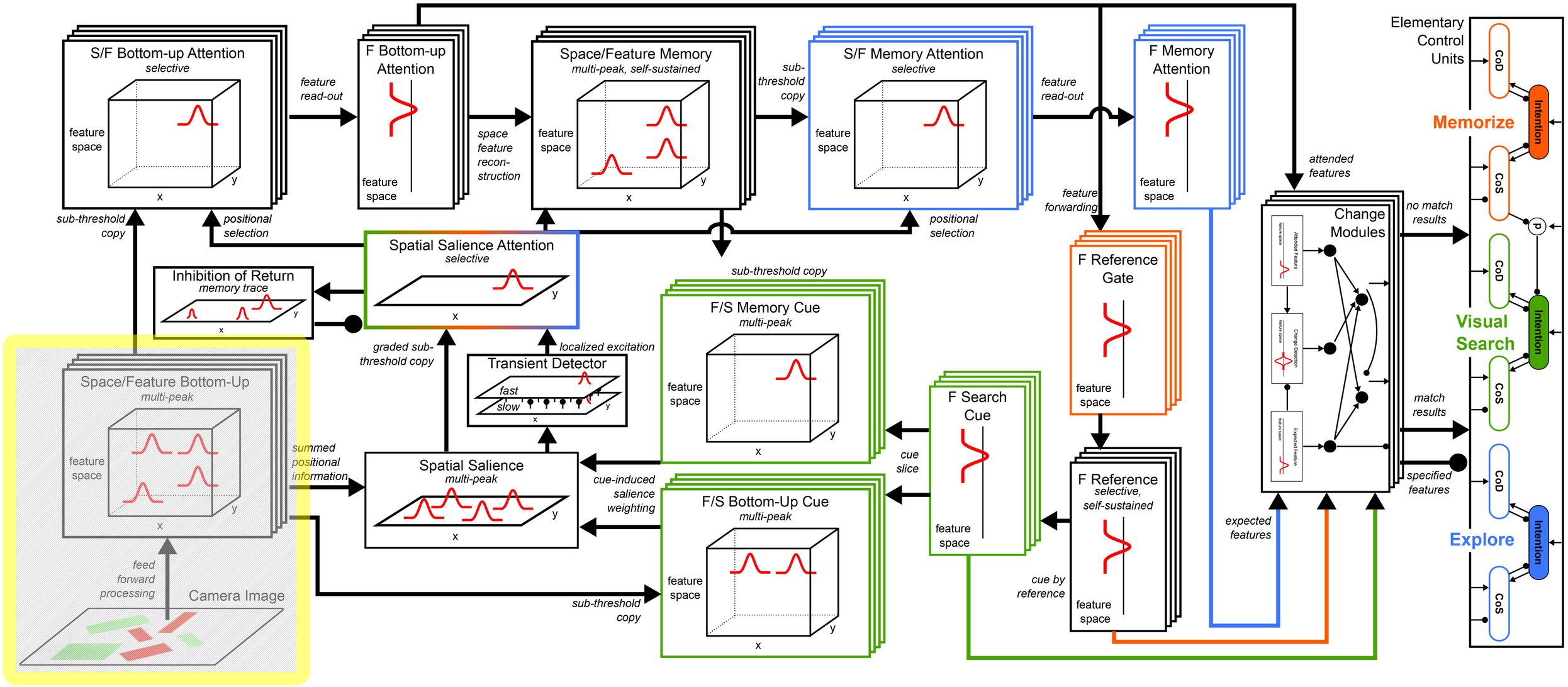
the features are extracted and stored in WM



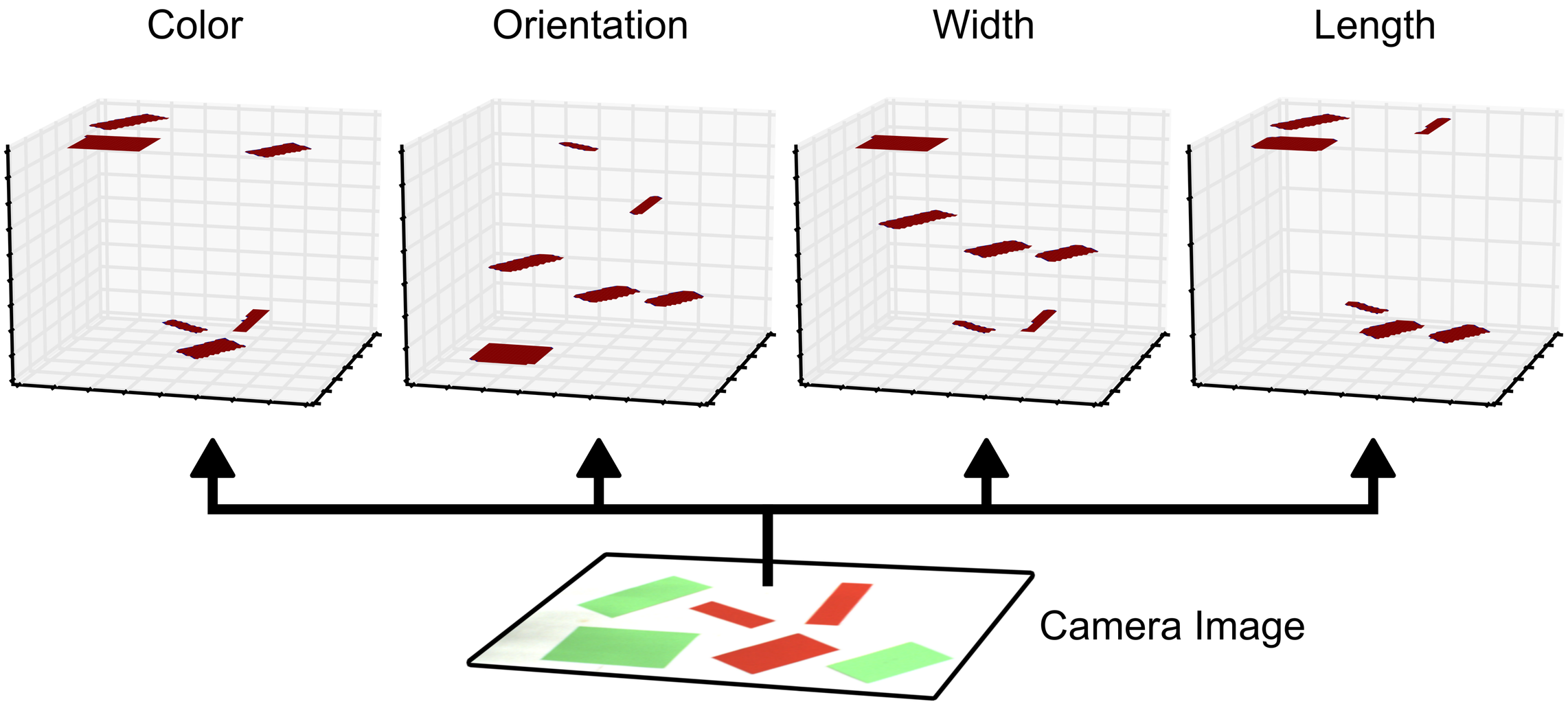
the neural dynamic architecture



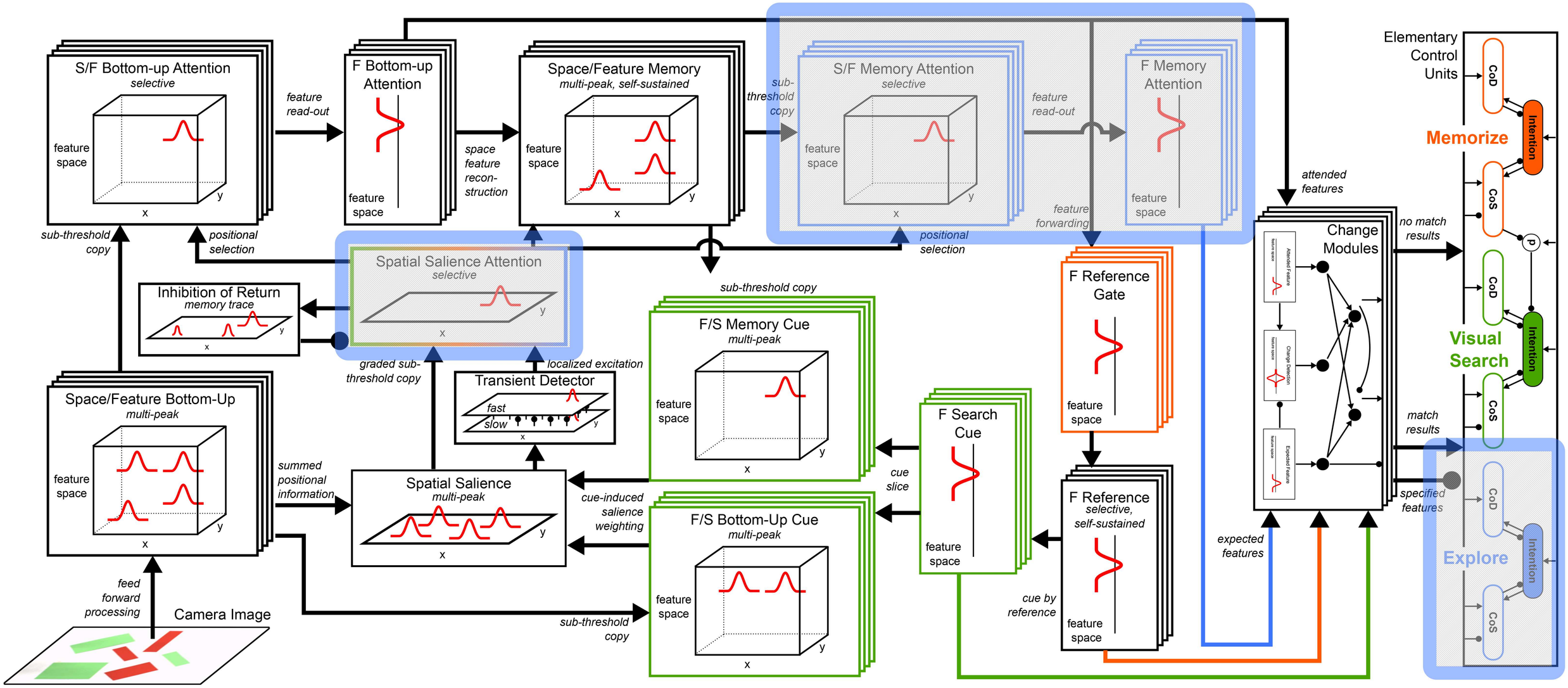
live camera input



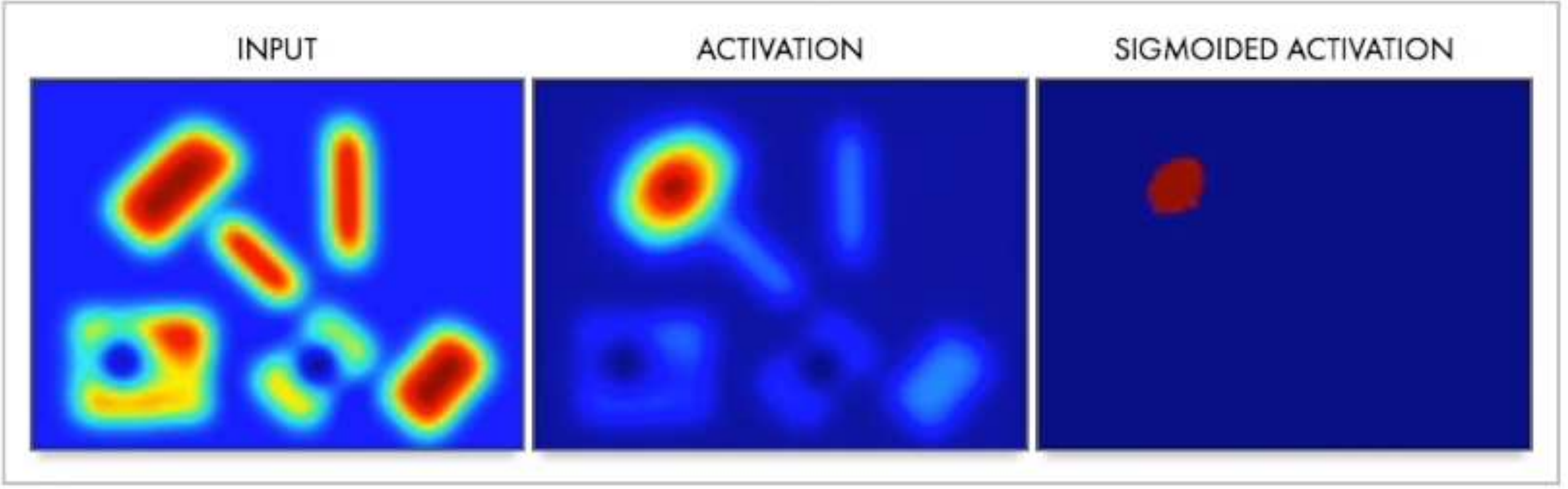
feature extraction



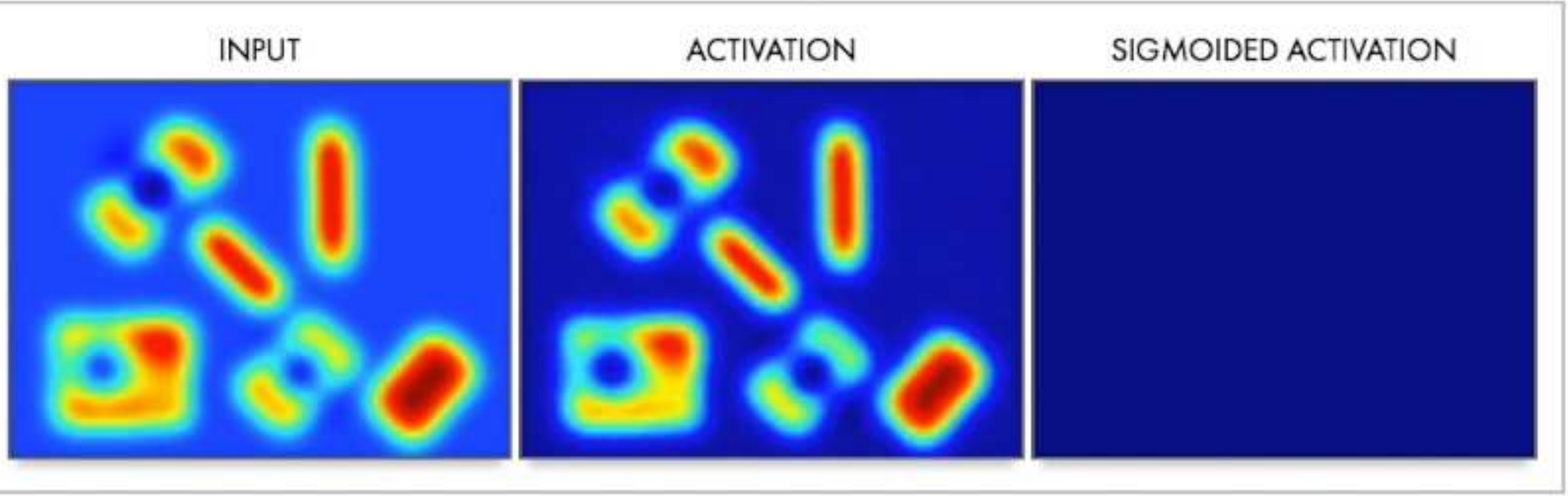
autonomous sequential exploration of the scene, building the scene memory



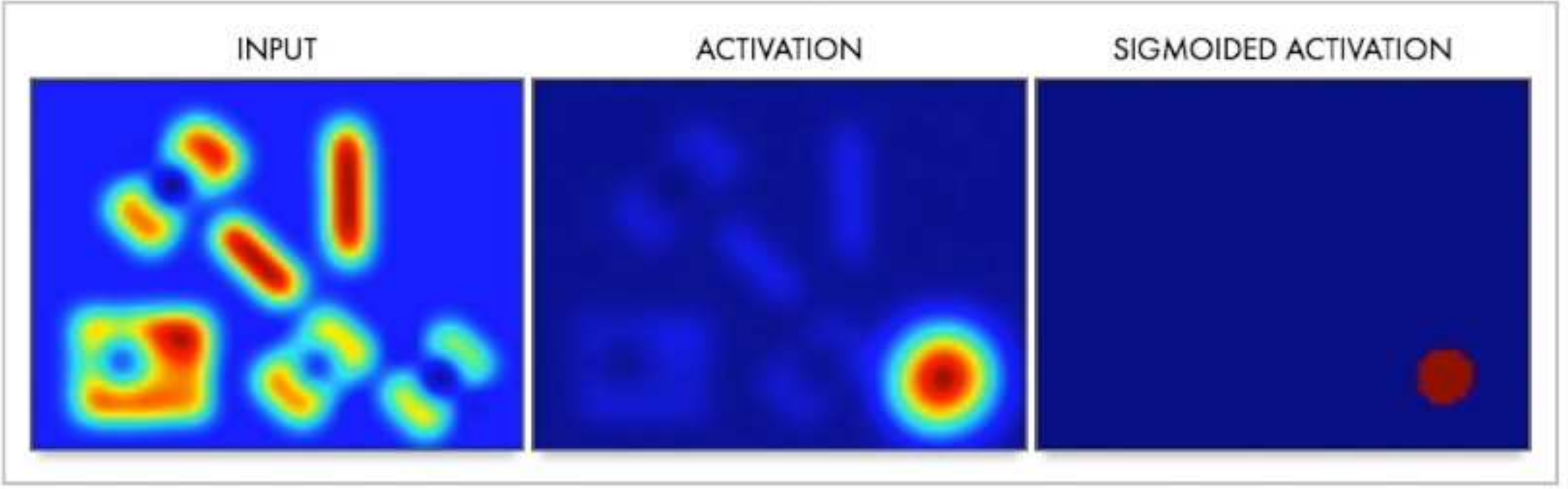
ATTENTION



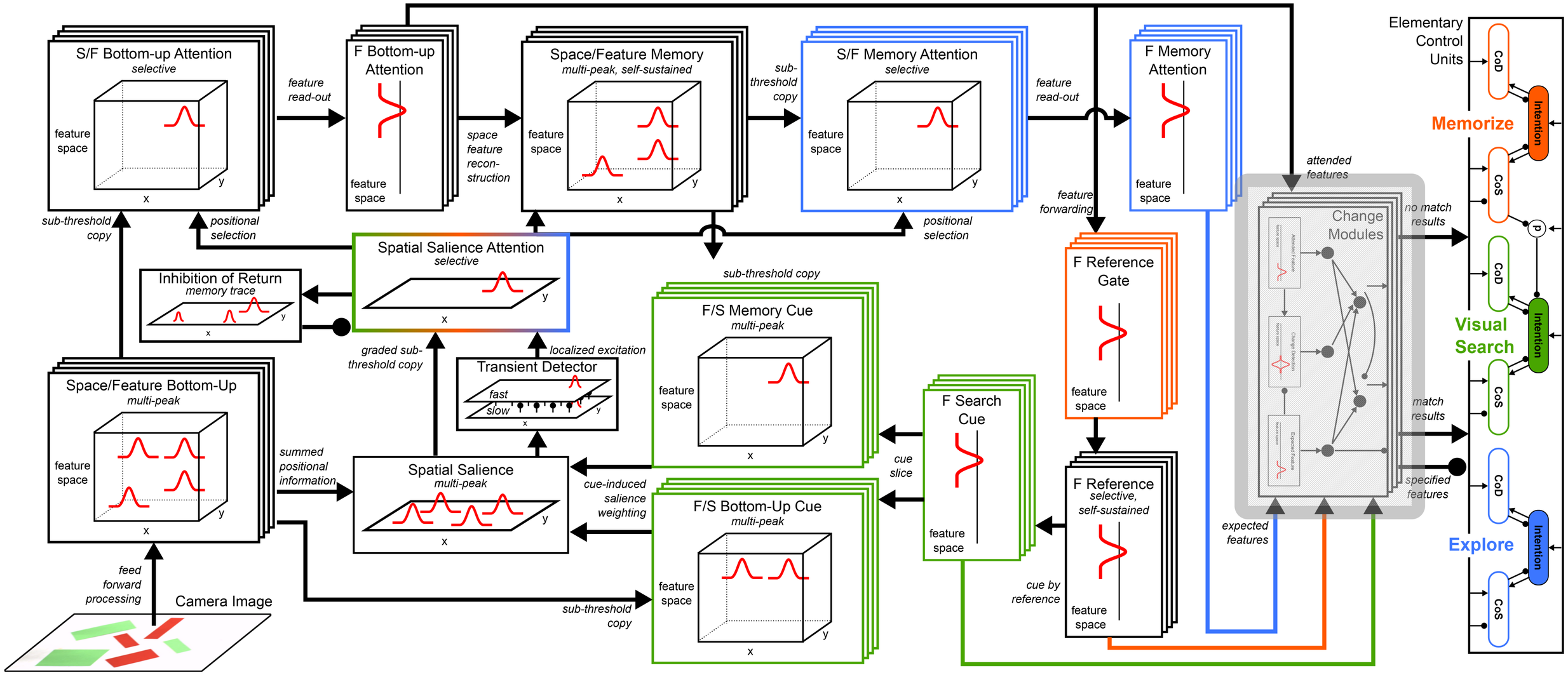
ATTENTION



ATTENTION

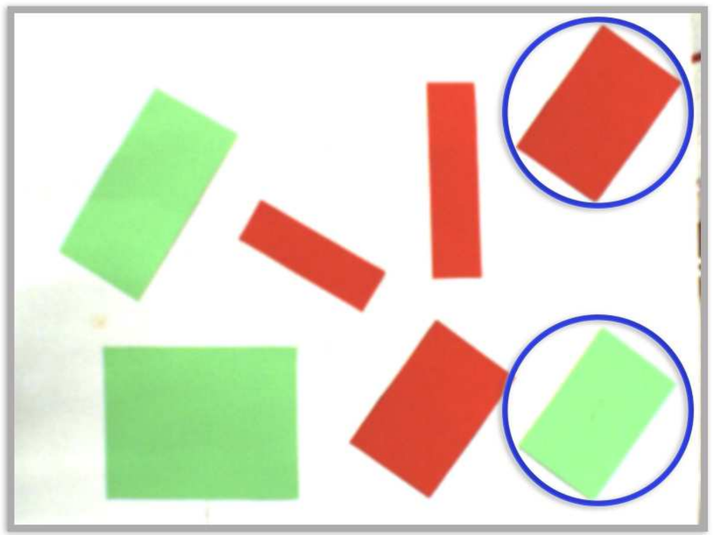


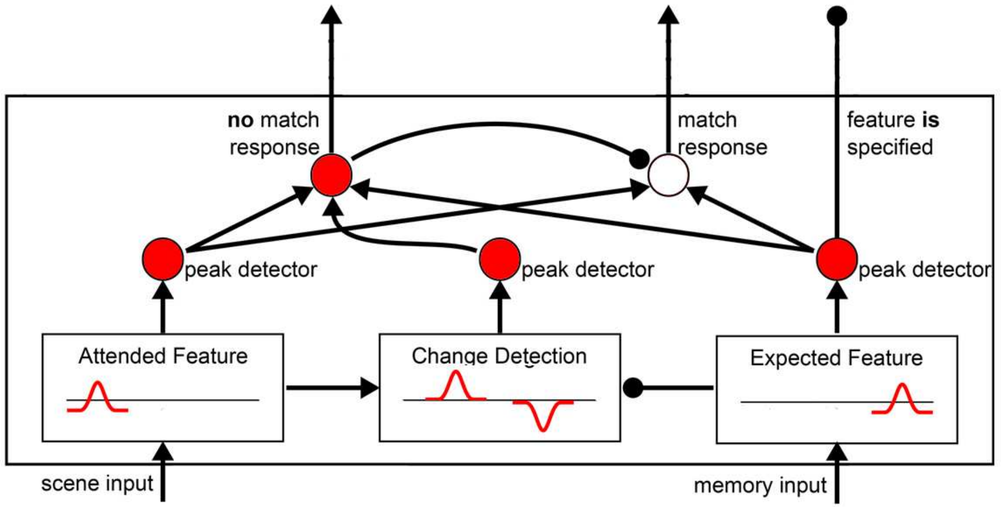
change detection, shared core control process

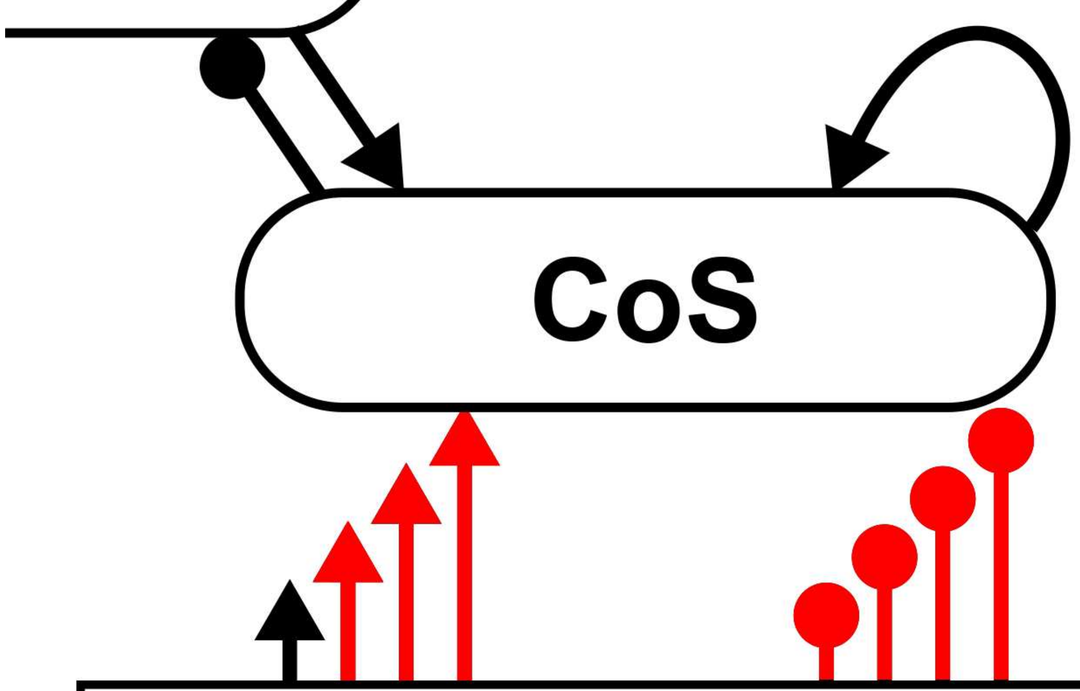


CoD

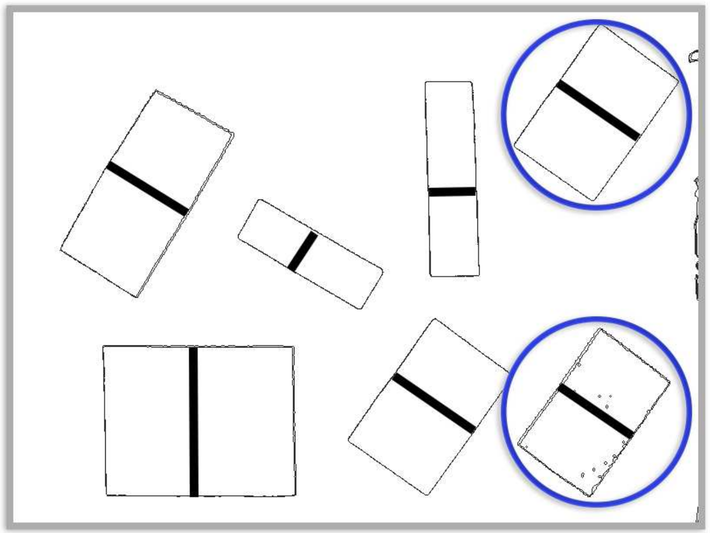
COLOR

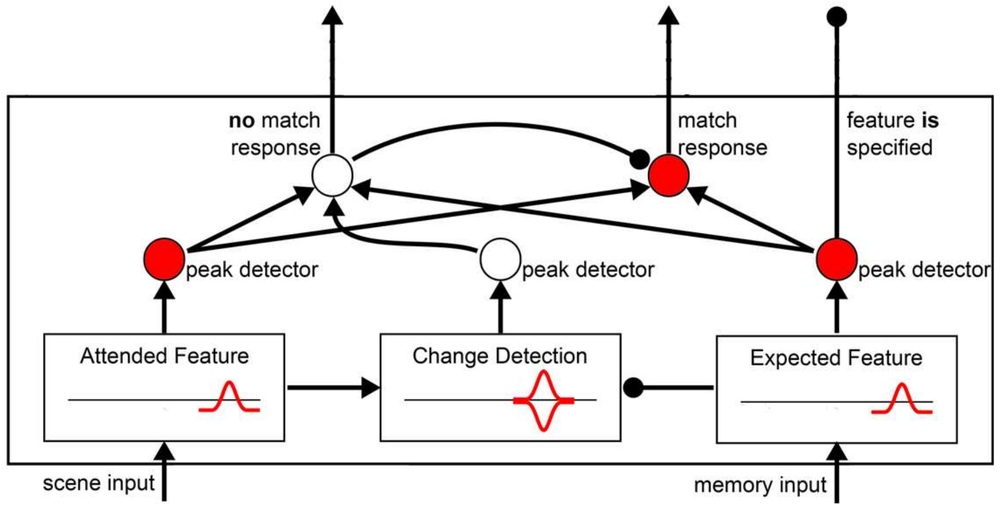


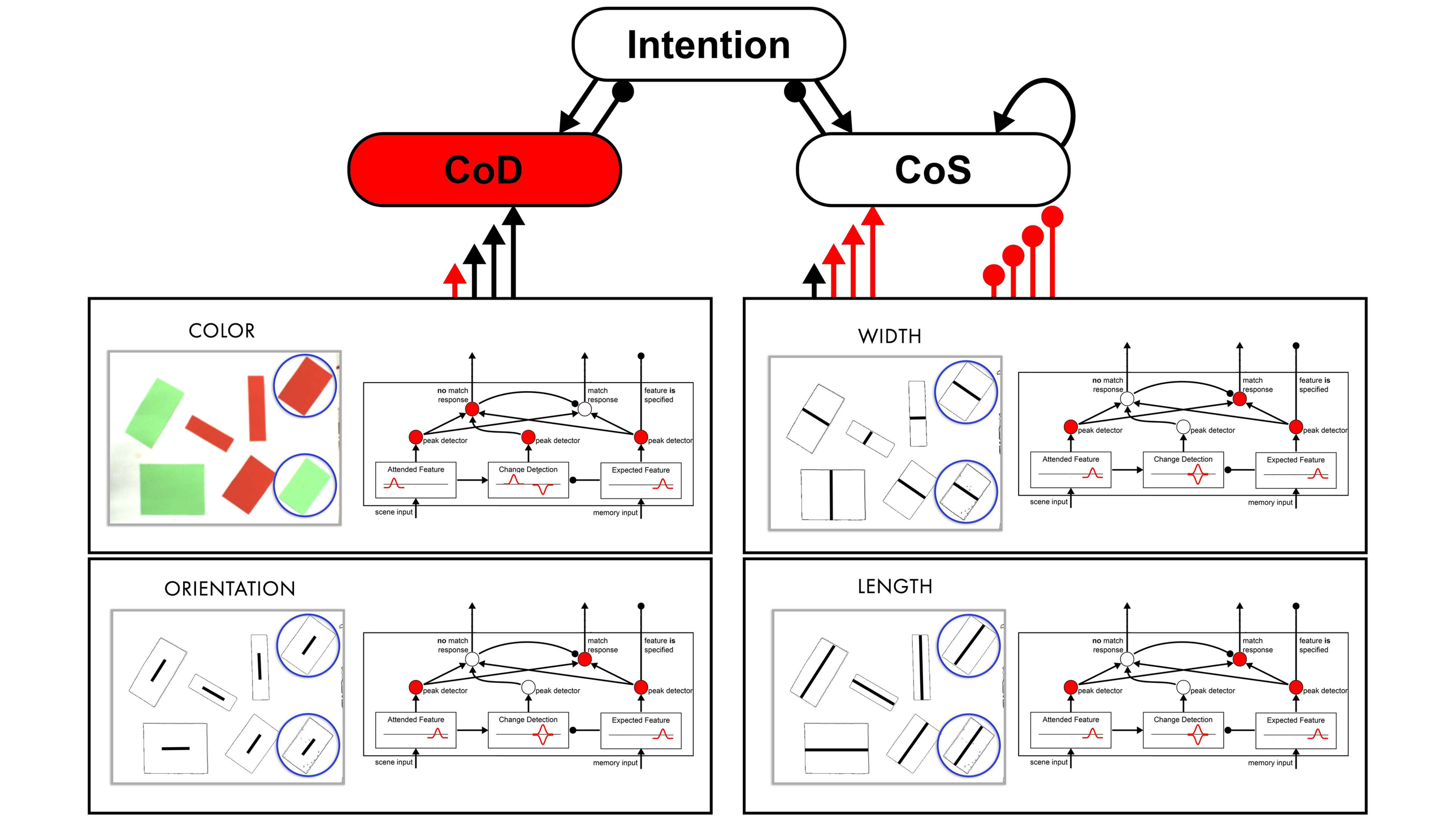

no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input


CoS

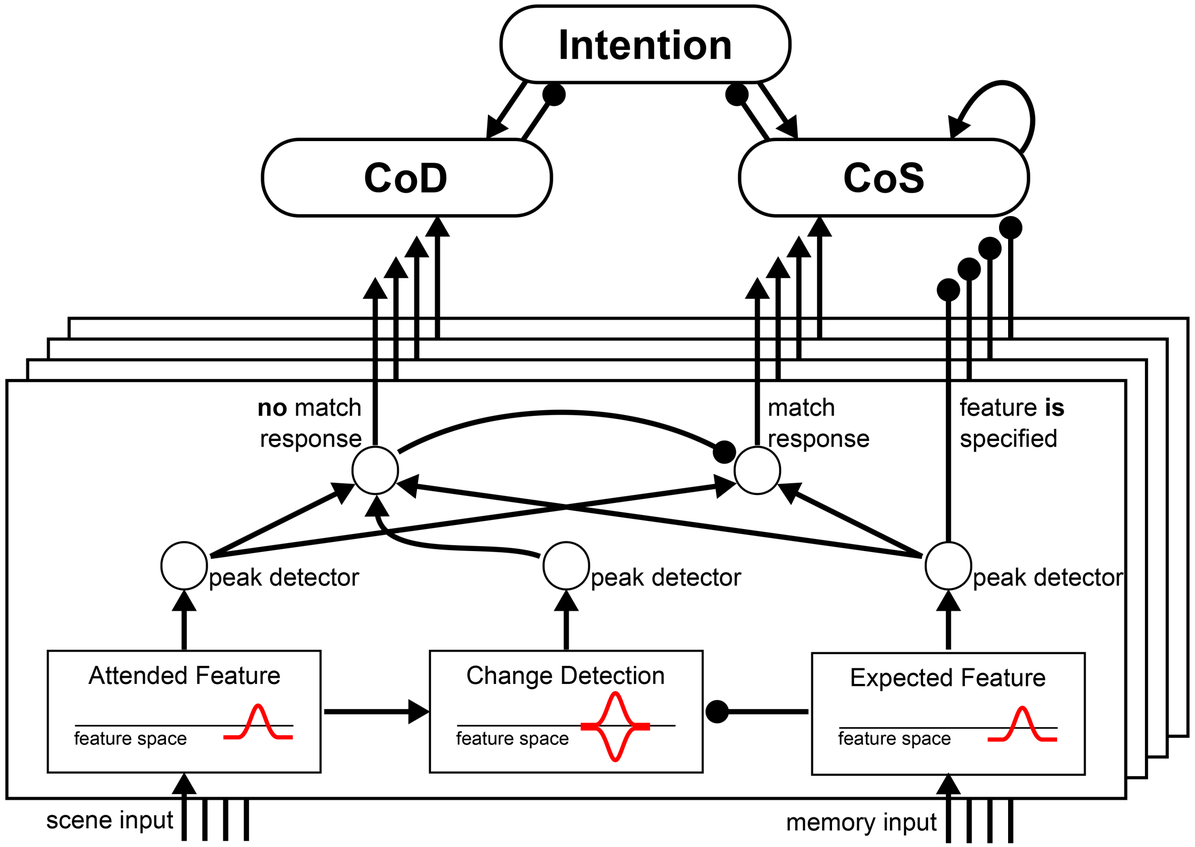
WIDTH




no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input


Intention
CoD
CoS
COLOR
no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input
WIDTH
no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input
ORIENTATION
no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input
LENGTH
no match response
match response
feature is specified
peak detector
peak detector
peak detector
Attended Feature
Change Detection
Expected Feature
scene input
memory input

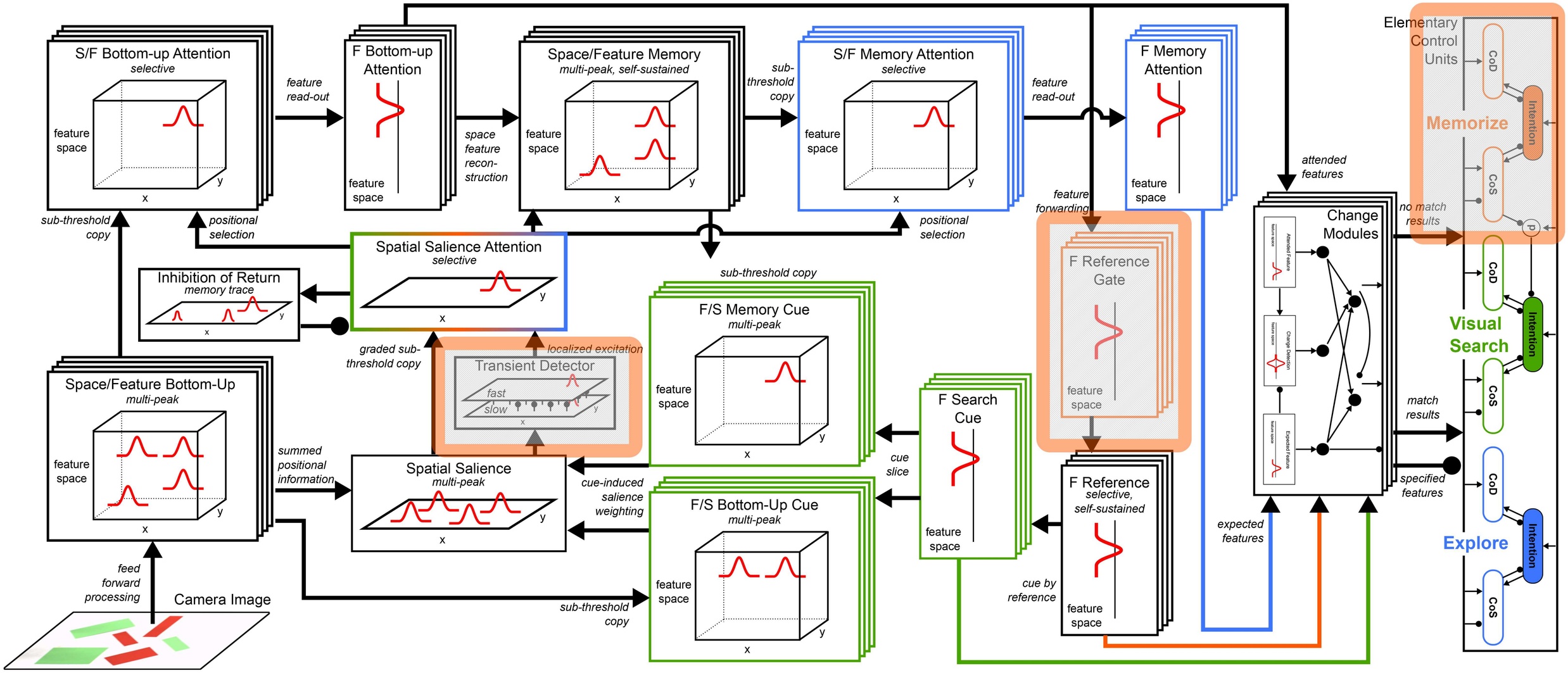
change detection



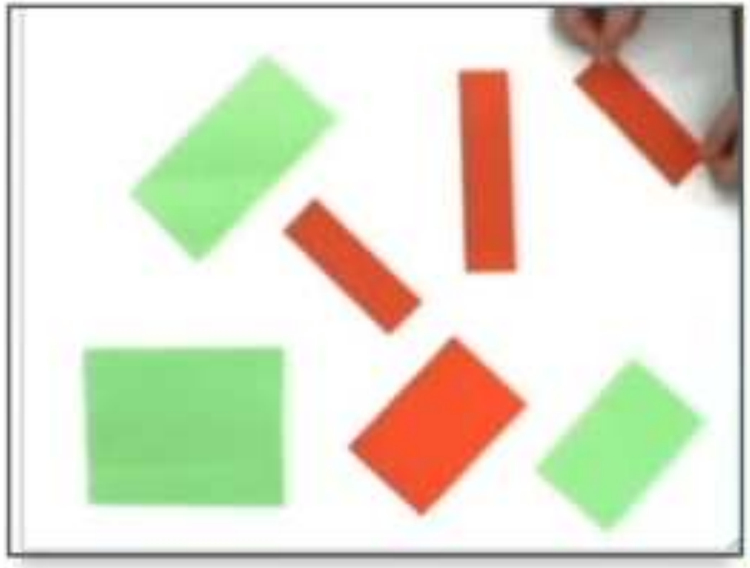
- comparison
- binding
- sequentiality
- process organization

Johnson, Spencer, Luck, & Schöner (Psychological Science, 2009)
Richter, Sandamirskaya, & Schöner (IROS, 2012)

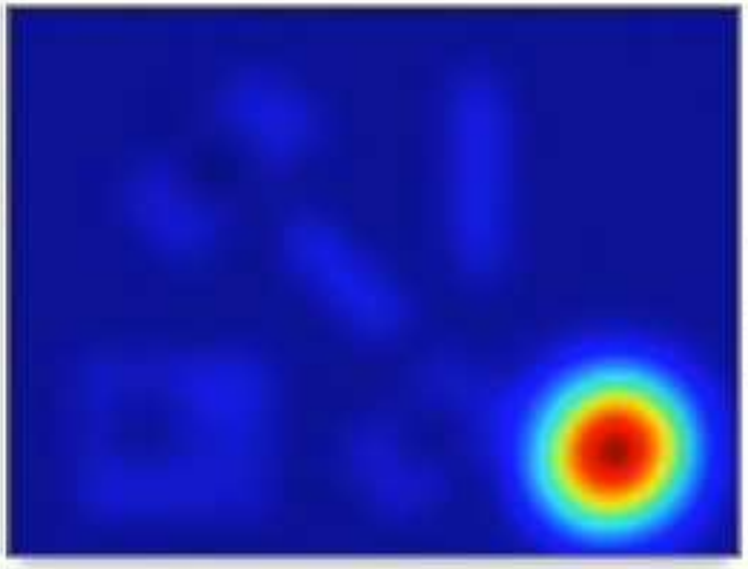
memorize, ...


S/F Bottom-up Attention
selective
feature space
x
y
feature read-out
F Bottom-up Attention
feature space
space feature recon-struction
Space/Feature Memory
multi-peak, self-sustained
sub-threshold copy
S/F Memory Attention
selective
feature read-out
F Memory Attention
feature space
attended features
Elementary Control Units
CoD
Intention
Memorize
CoS
sub-threshold copy
positional selection
Spatial Salience Attention
selective
feature forwarding
F Reference Gate
Change Modules
no match results
Inhibition of Return
memory trace
graded sub-threshold copy
localized excitation
Transient Detector
fast
slow
sub-threshold copy
F/S Memory Cue
multi-peak
Visual Search
Space/Feature Bottom-Up
multi-peak
summed positional information
Spatial Salience
multi-peak
cue-induced salience weighting
F/S Bottom-Up Cue
multi-peak
cue slice
F Search Cue
match results
specified features
F Reference
selective, self-sustained
cue by reference
expected features
Explore
feed forward processing
Camera Image
sub-threshold copy

CAMERA INPUT



ACTIVATION

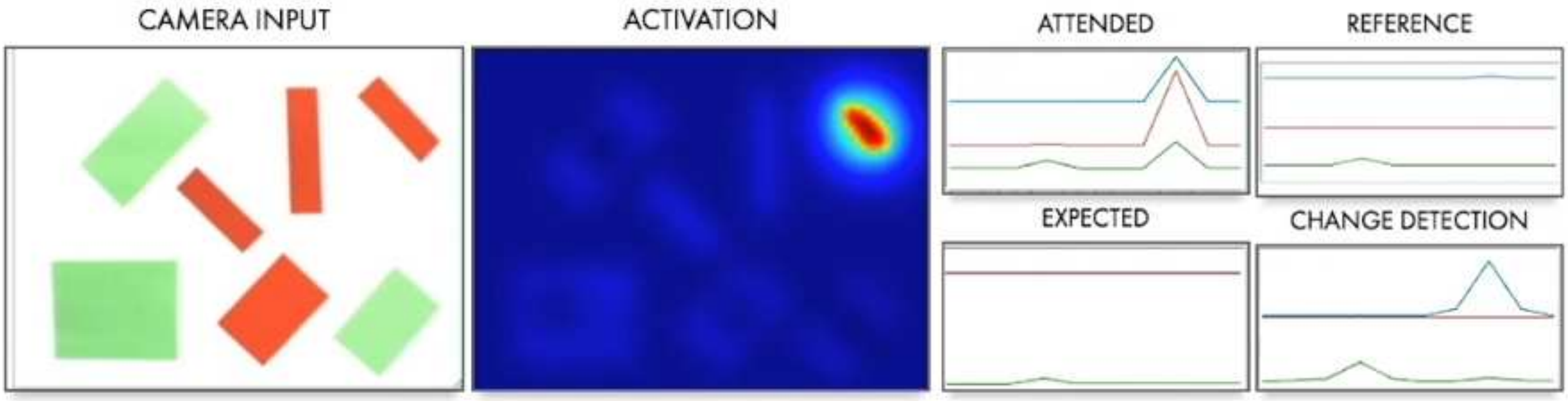


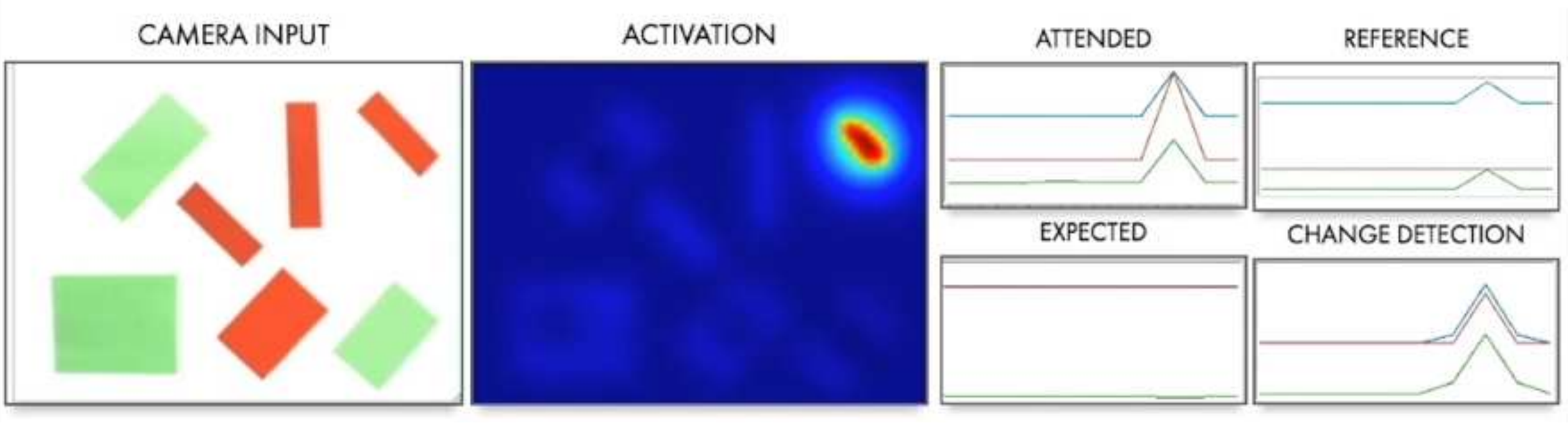
ATTENDED

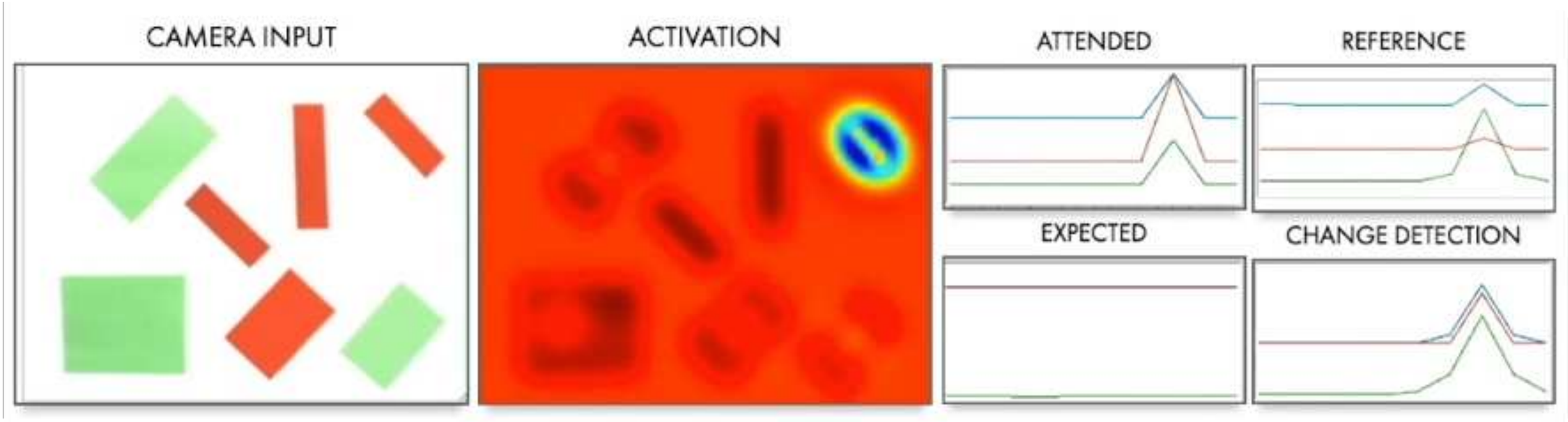
REFERENCE

EXPECTED

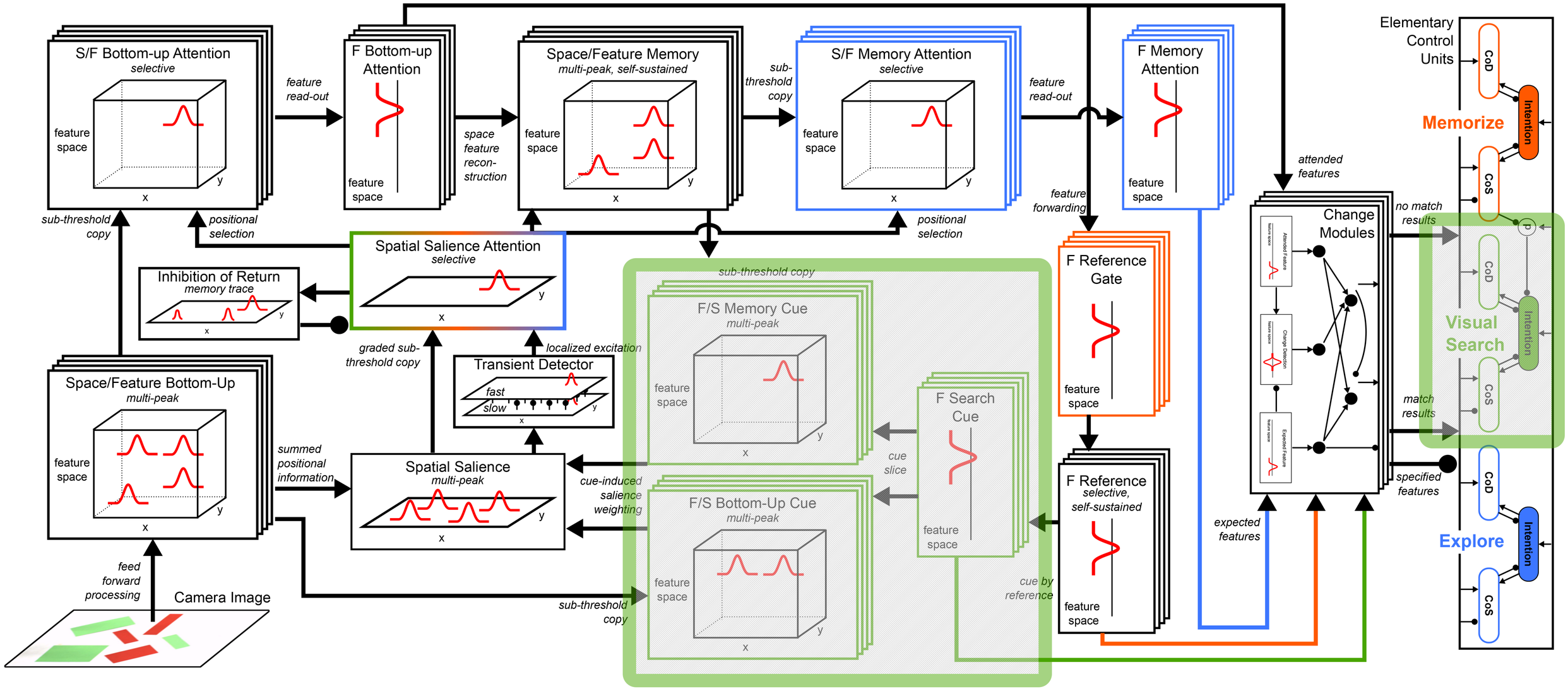
CHANGE DETECTION

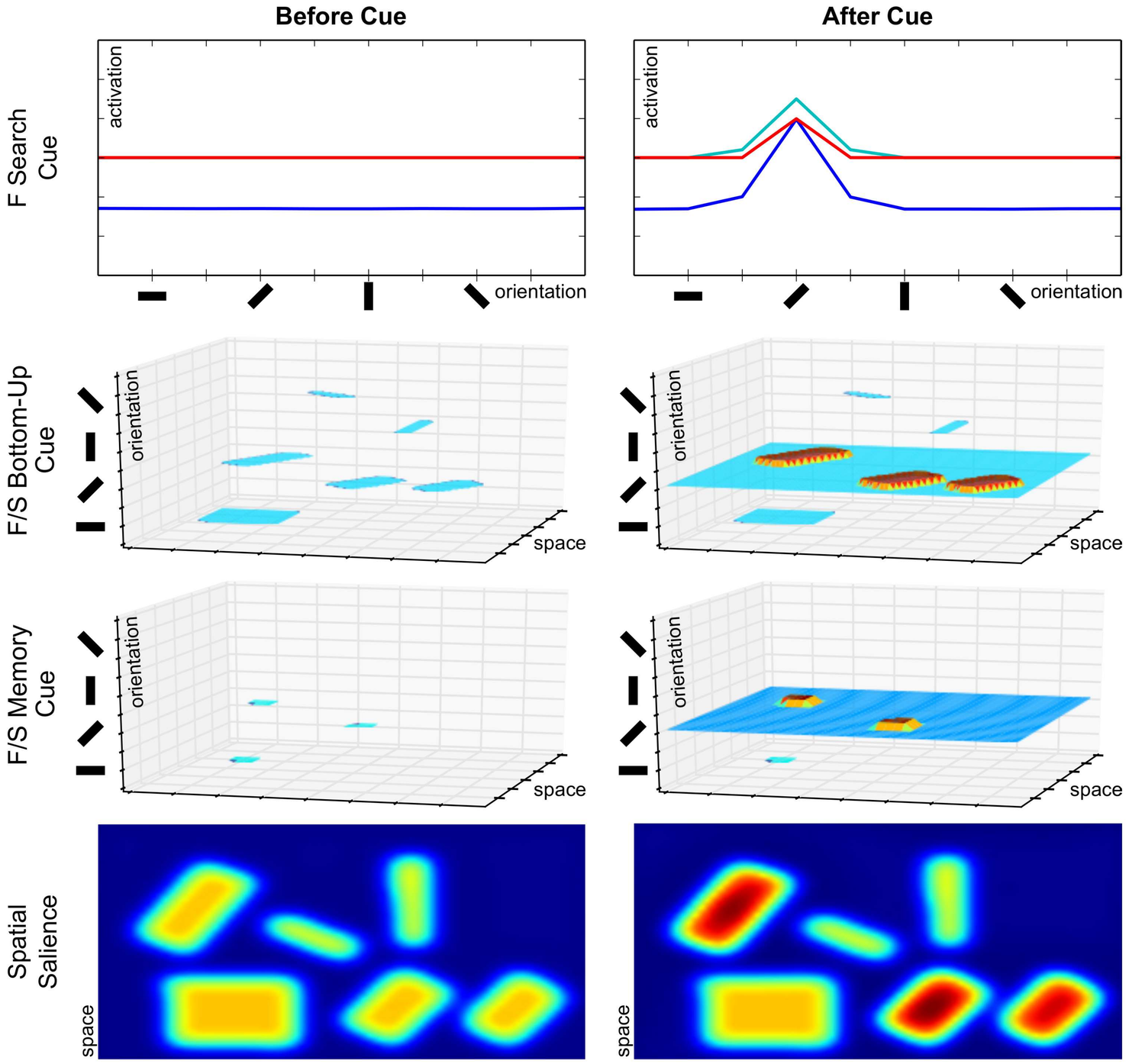

CAMERA INPUT
ACTIVATION
ATTENDED
REFERENCE
EXPECTED
CHANGE DETECTION


CAMERA INPUT
ACTIVATION
ATTENDED
REFERENCE
EXPECTED
CHANGE DETECTION


CAMERA INPUT
ACTIVATION
ATTENDED
REFERENCE
EXPECTED
CHANGE DETECTION

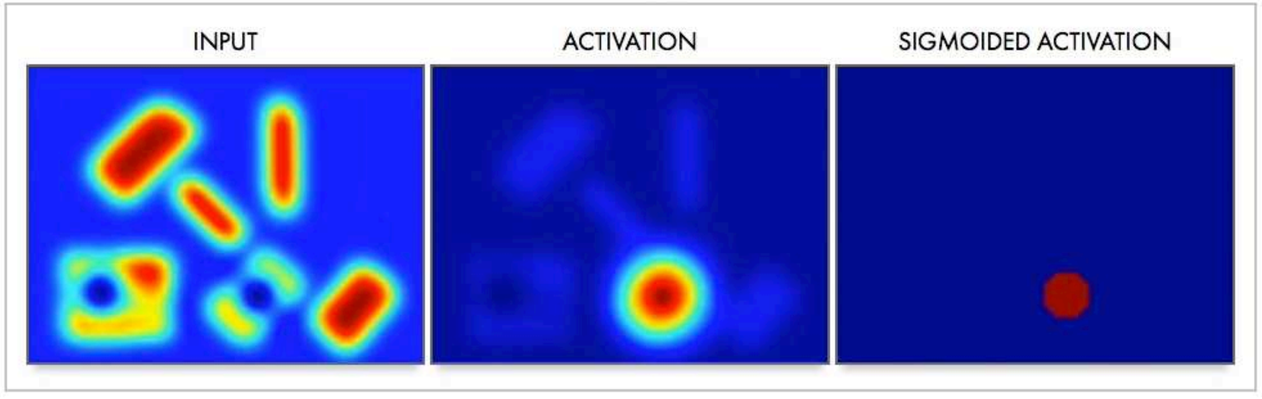
visual search



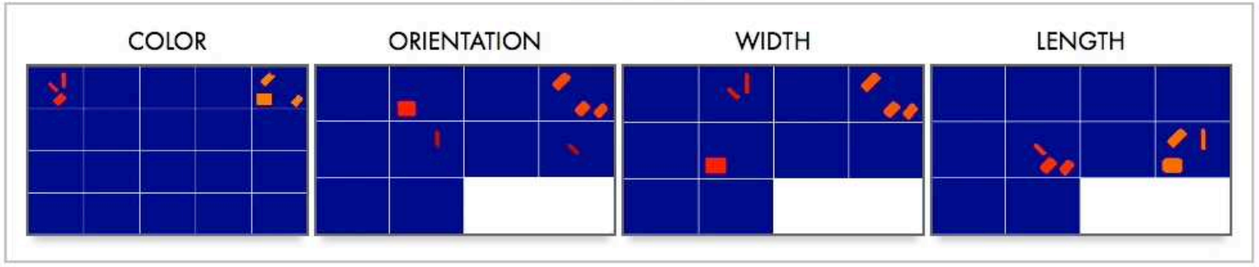
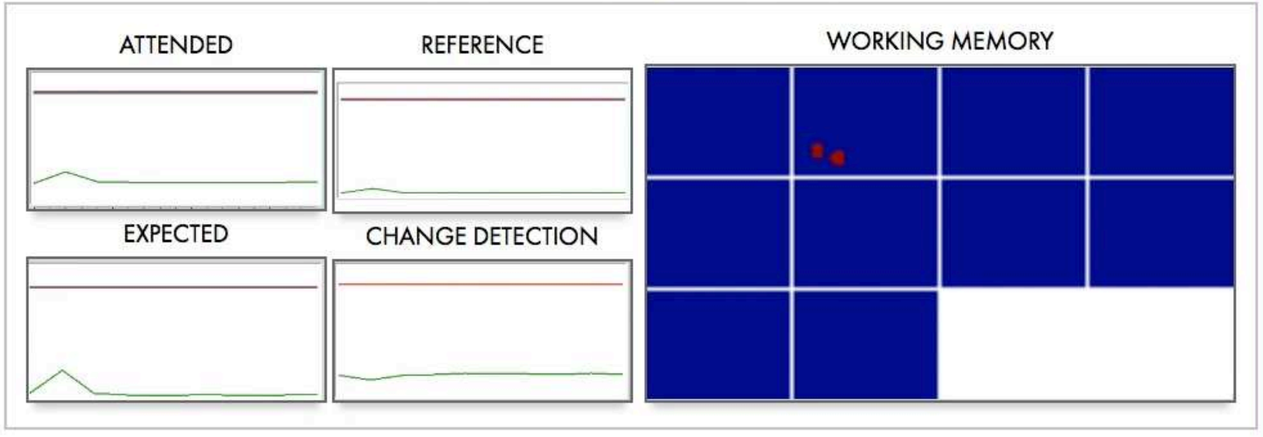

Before Cue
After Cue
F Search Cue
activation
orientation
F/S Bottom-Up Cue
F/S Memory Cue
space
Spatial Salience

SALIENCY MAP

ATTENTION



FEATURE MAPS



FEATURE PROCESSING (ORIENTATION)