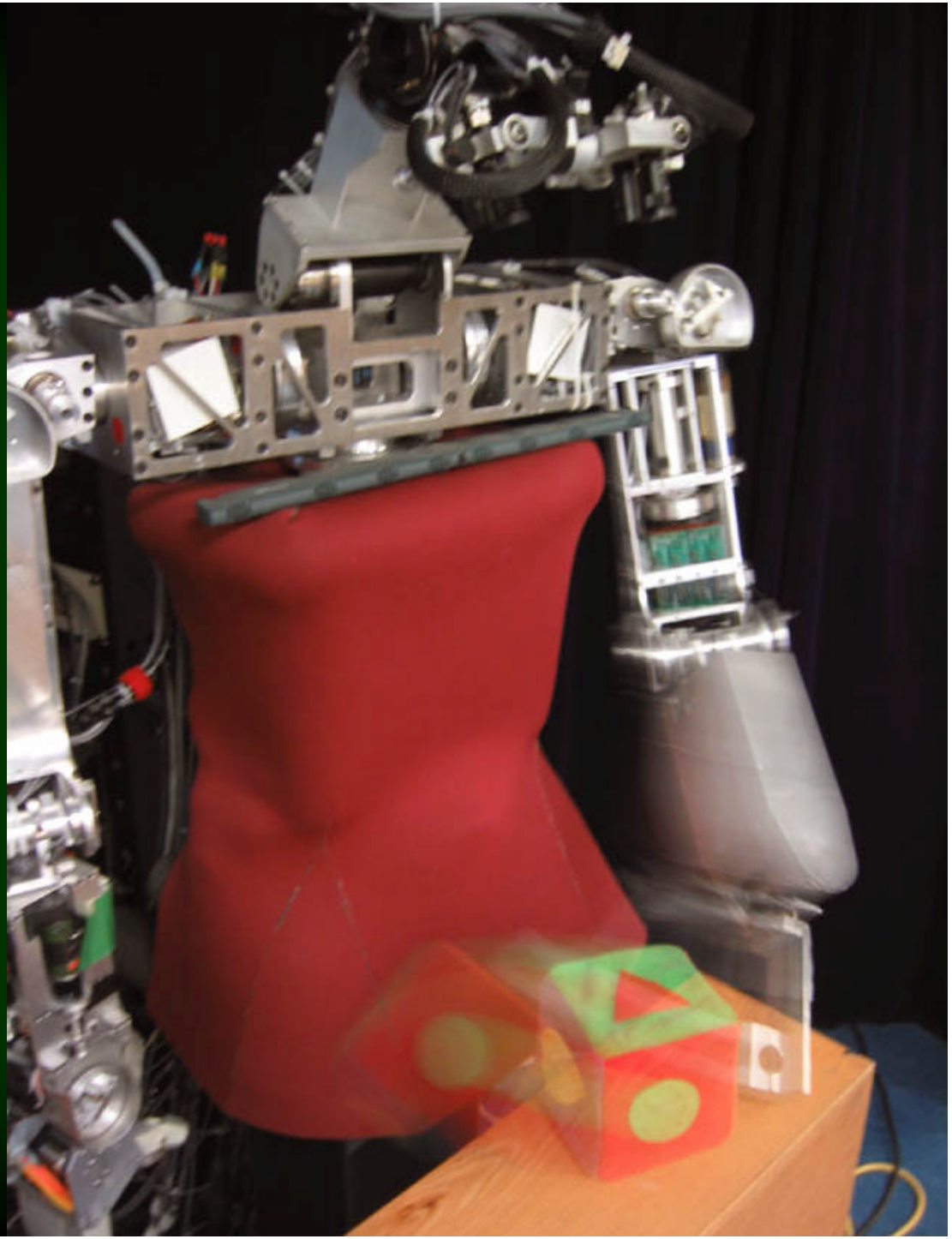
Object Lesson:

Discovering and Learning to Recognize Objects

– Paul Fitzpatrick –

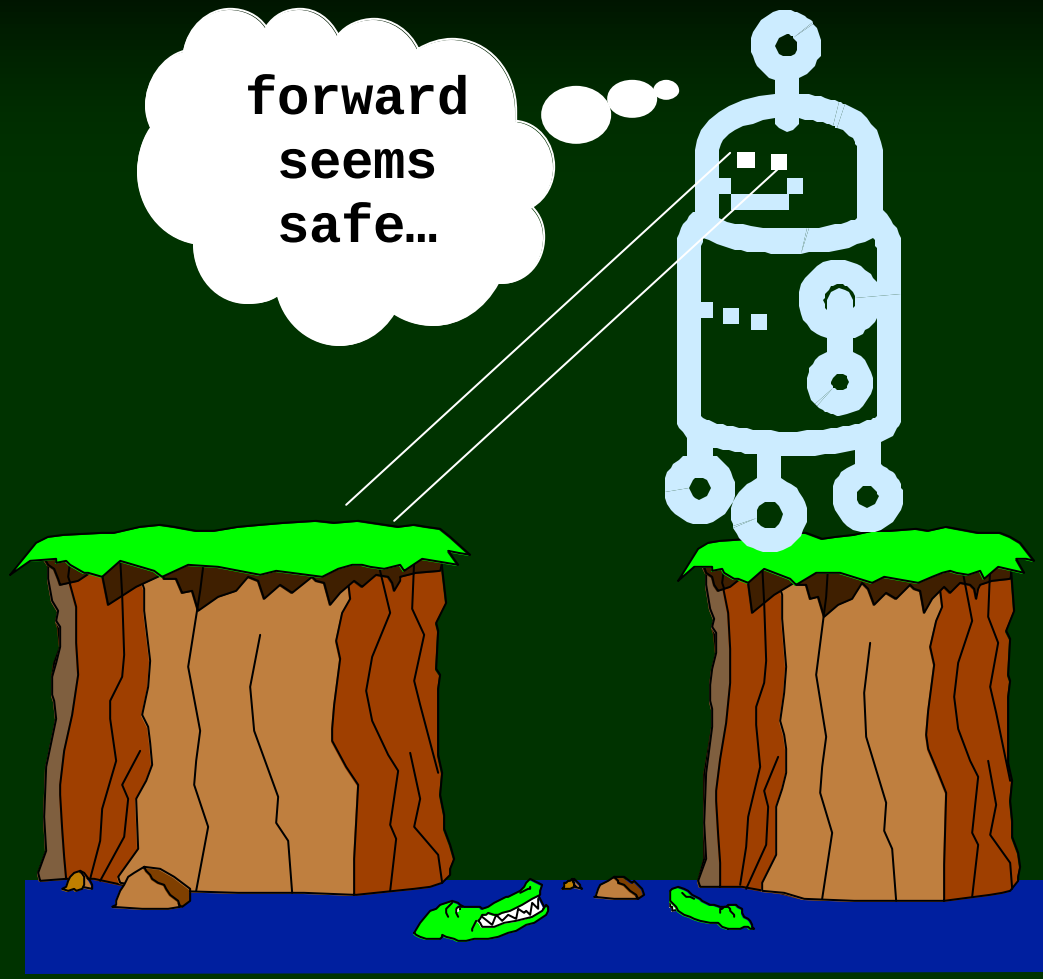
MIT CSAIL

USA



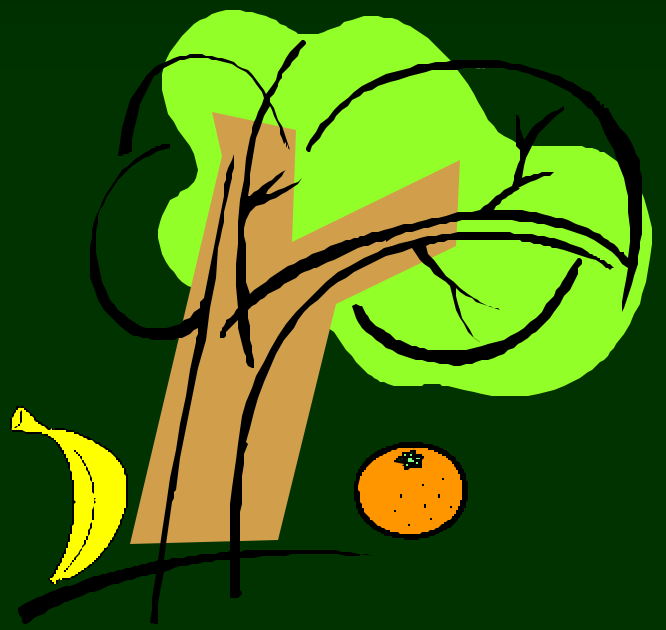
robots and learning

- Robots have access to physics, and physics is a good teacher
- Physics won't let you believe the wrong thing for long
- Robot perception should ideally integrate experimentation, or at least learn from (non-fatal) mistakes



a challenge: object perception

- Object perception is a key enabling technology
- Many components:
  - Object detection
  - Object segmentation
  - Object recognition
- Typical systems require human-prepared training data; can we use autonomous experimentation?



a challenge: object perception

- Object perception is a key enabling technology
- Many components:
  - Object detection
  - Object segmentation
  - Object recognition
- Typical systems require human-prepared training data; can we use autonomous experimentation?

Fruit detection

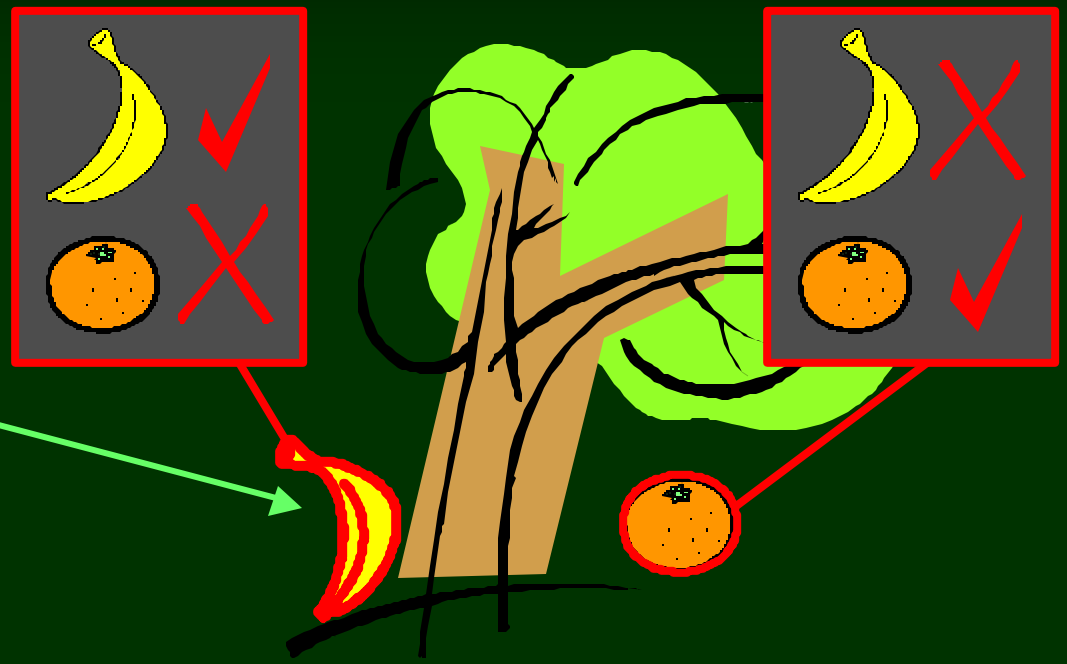
a challenge: object perception

- Object perception is a key enabling technology
- Many components:
  - Object detection
  - Object segmentation
  - Object recognition
- Typical systems require human-prepared training data; can we use autonomous experimentation?

Fruit segmentation

a challenge: object perception

- Object perception is a key enabling technology
- Many components:
  - Object detection
  - Object segmentation
  - Object recognition
- Typical systems require human-prepared training data – can't adapt to new situations autonomously



Fruit recognition

talk overview

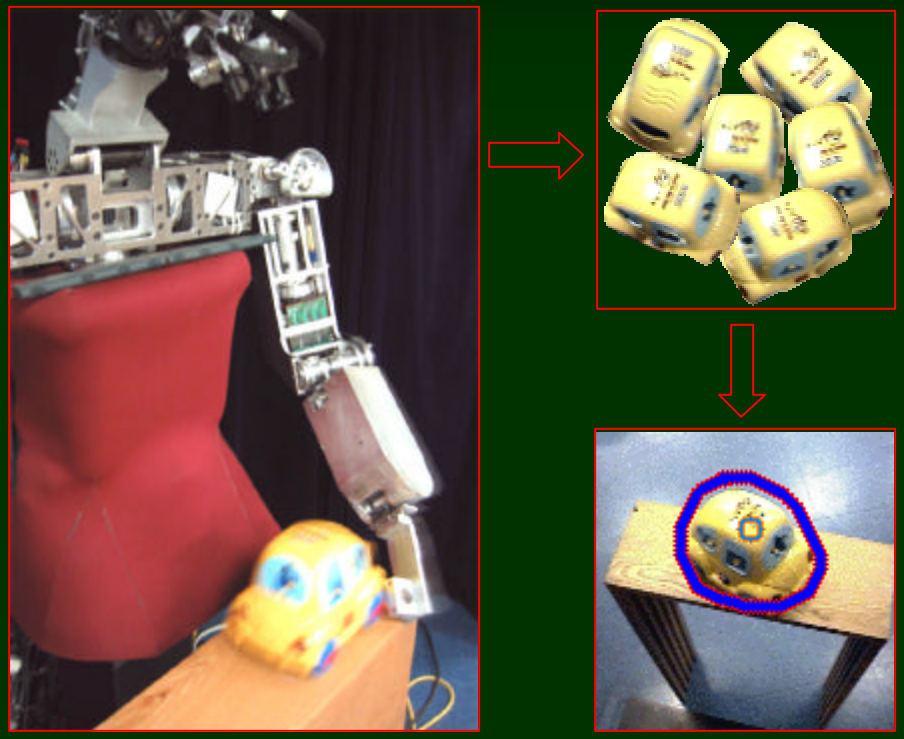
- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

talk overview

- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

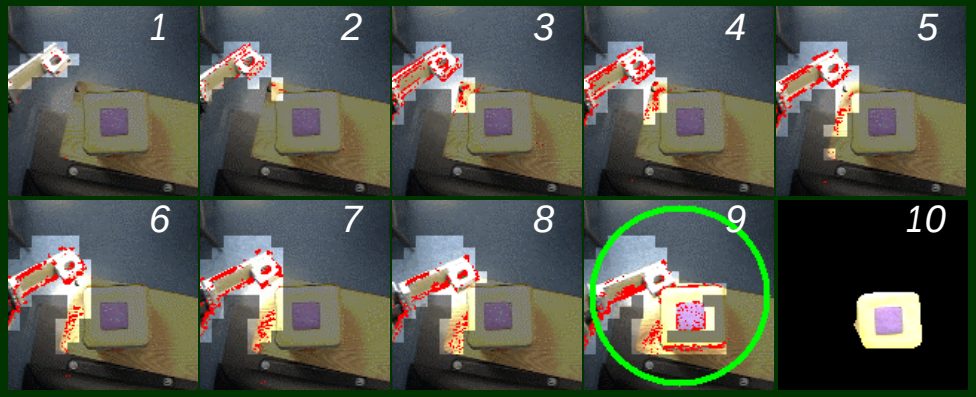
active segmentation

- Object boundaries are not always easy to detect visually
- Solution: *Cog* sweeps arm through ambiguous area
- Any resulting object motion helps segmentation
- Robot can learn to recognize and segment object without further contact



active segmentation

- Detect contact between arm and object using fast, coarse processing on optic flow signal
- Do detailed comparison of motion immediately before and after collision
- Use minimum-cut algorithm to generate best segmentation

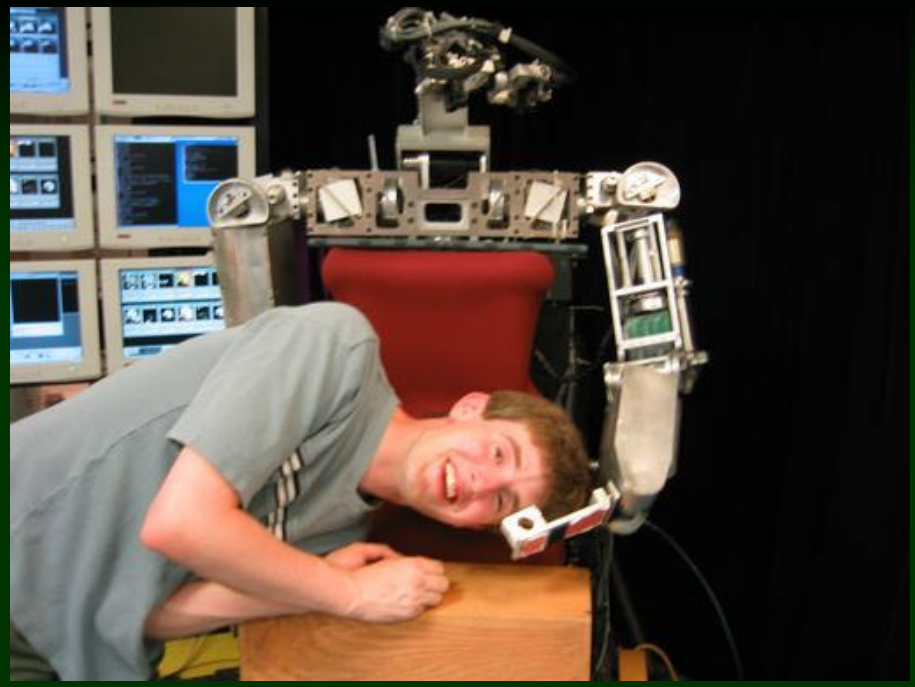


active segmentation

Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

active segmentation

- Not always practical!
- No good for objects the robot can view but not touch
- No good for very big or very small objects
- But fine for objects the robot is expected to manipulate

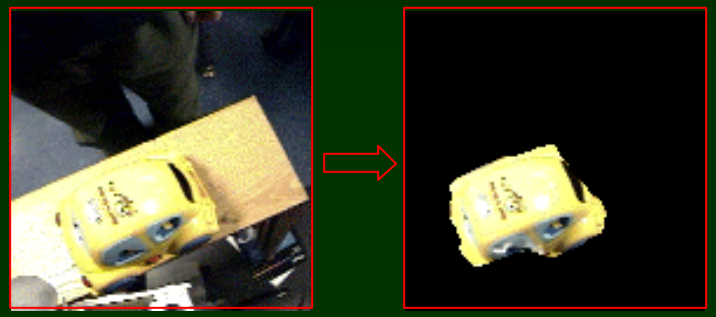


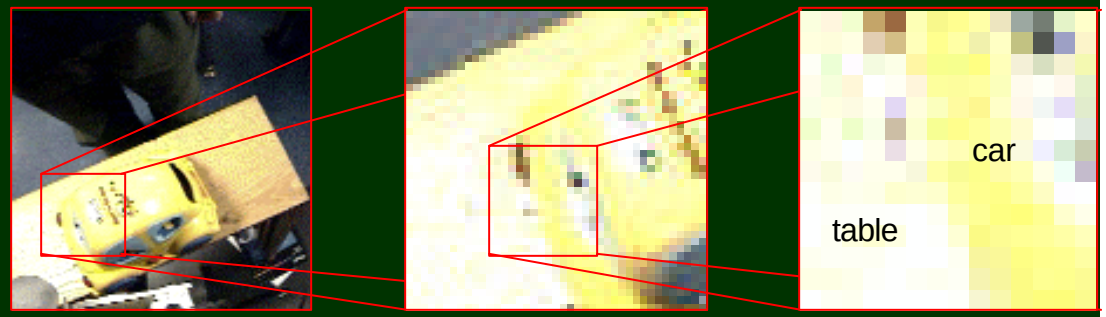
Head segmentation the hard way!

Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

listening to physics

- Active segmentation is useful even if robot normally depends on other segmentation cues (color, stereo)
- If passive segmentation is incorrect and robot fails to grasp object, active segmentation can use even clumsy collision to get truth
- Seems silly *not* to use this feedback from physics and keep making the same mistake





talk overview

- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

talk overview

- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

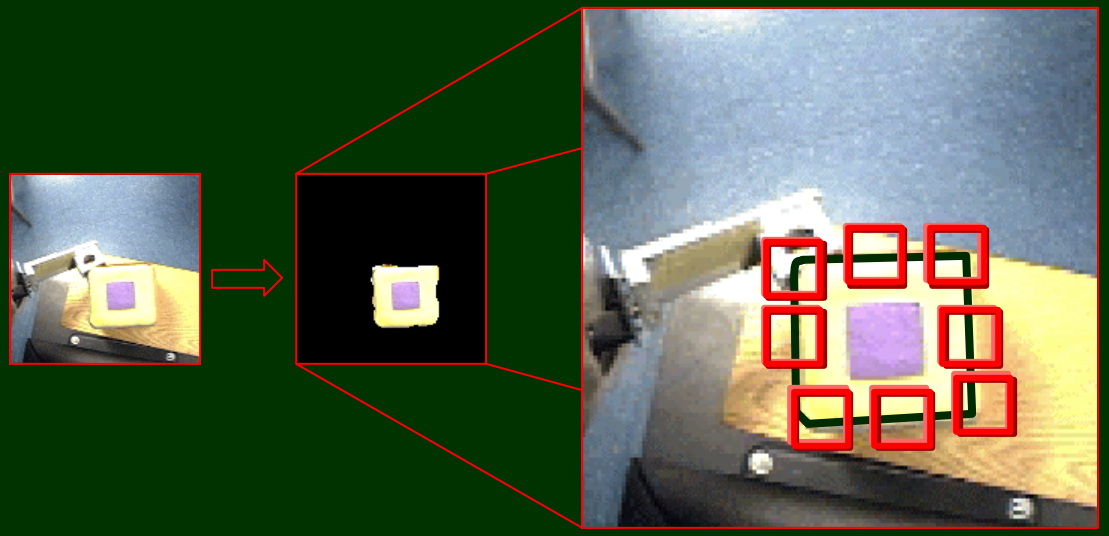
opportunistic learning

- To begin with, Cog has three categories of perceptual abilities:
  - Judgements it can currently make (e.g. about color, motion, time)
  - Judgements it can sometimes make (e.g. boundary of object, identity of object)
  - Judgements it cannot currently make (e.g. counting objects)

- With opportunistic learning, the robot takes judgements in the sometimes category and works to promote them to the can category by finding reliable correlated features that are more frequently available
  - Example: analysis of boundaries detected through motion yields purely visual features that are predictive of edge orientation
  - Example: assuming a non-hostile environment (some continuity in time and space) segmented views of objects can be grouped and purely visual features inferred that are characteristic of distinct objects

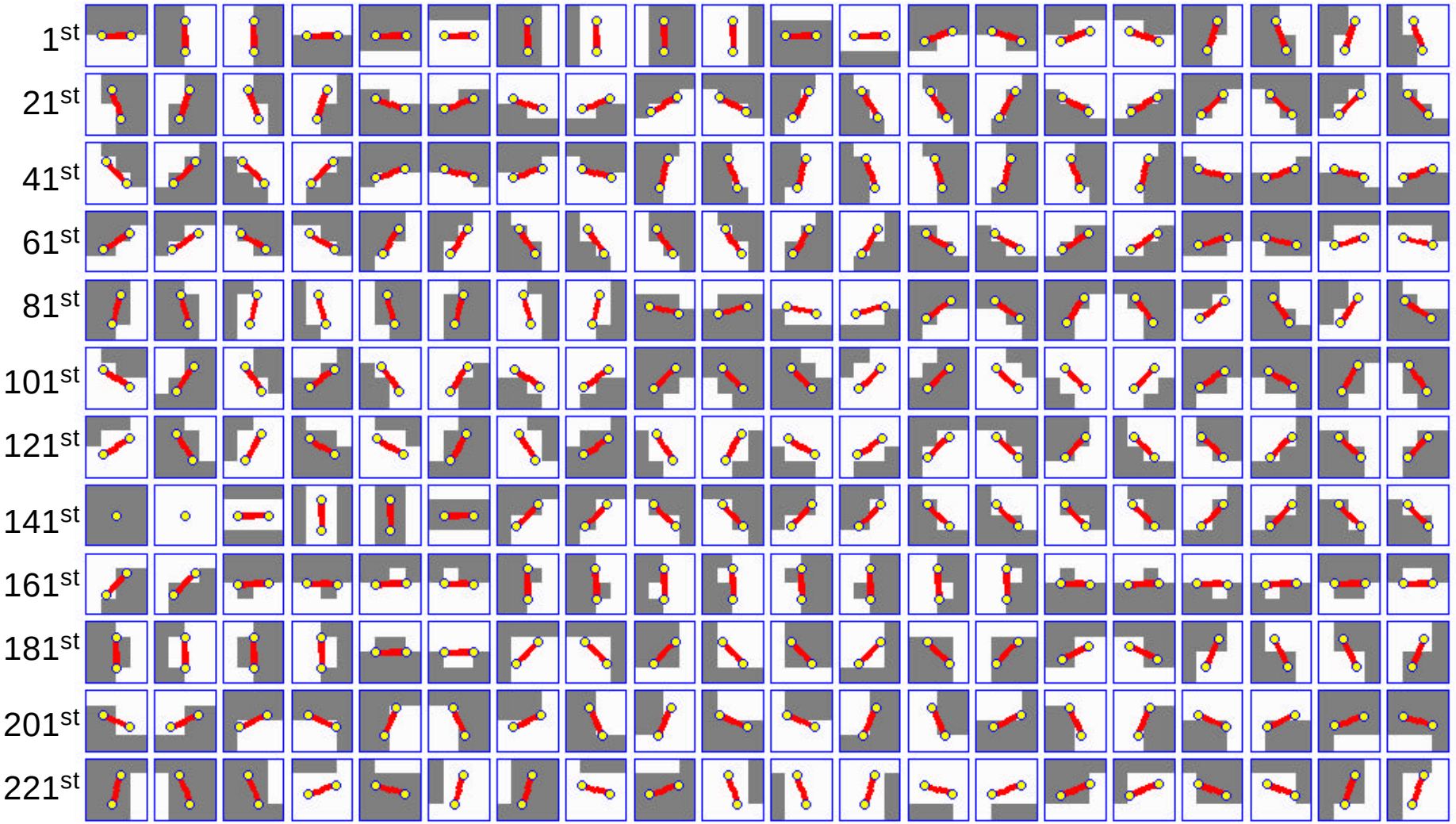
training a model of edge appearance

- Robot initially only perceives oriented edges through active segmentation procedure
- Robot collects samples of edge appearance along boundary, and builds a look-up table from appearance to orientation angle
- Now can perceive orientation directly
- This is often built in, but it doesn't have to be

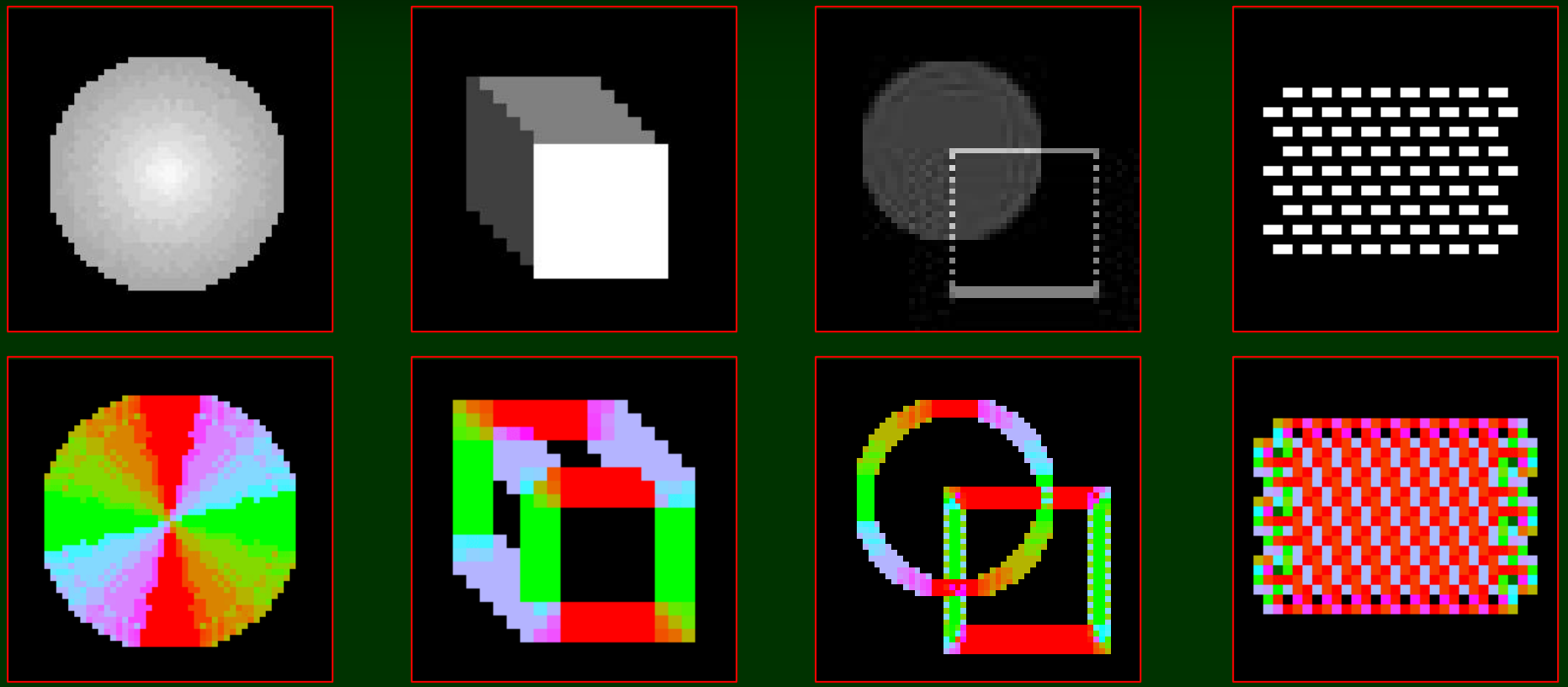


Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

most frequent samples

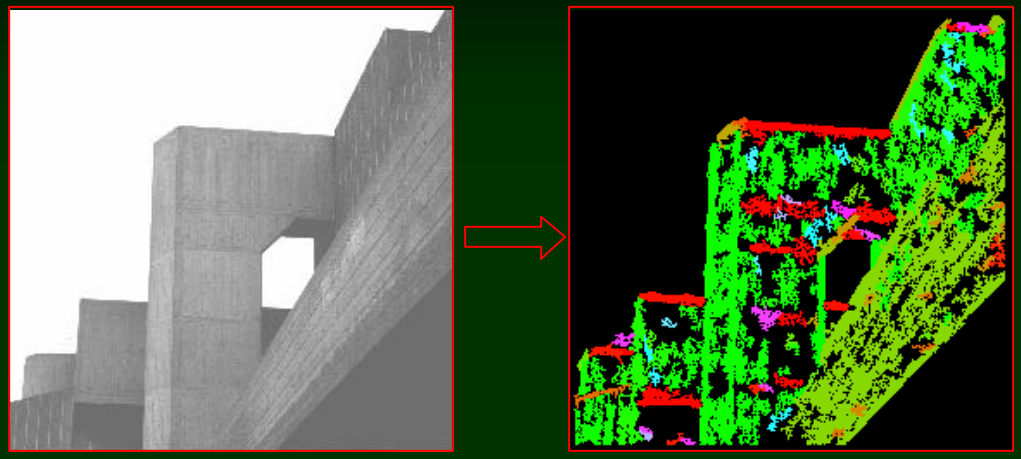


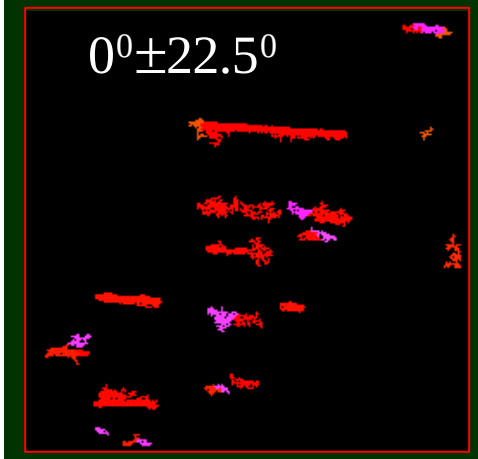
some tests

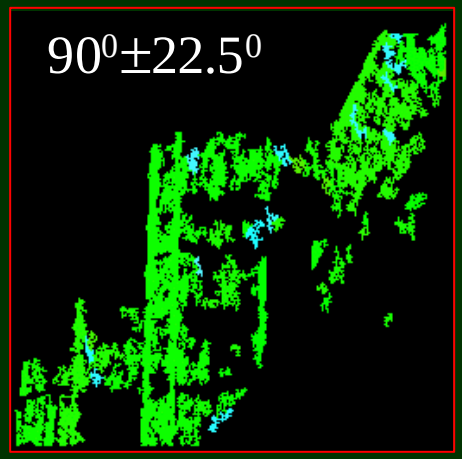


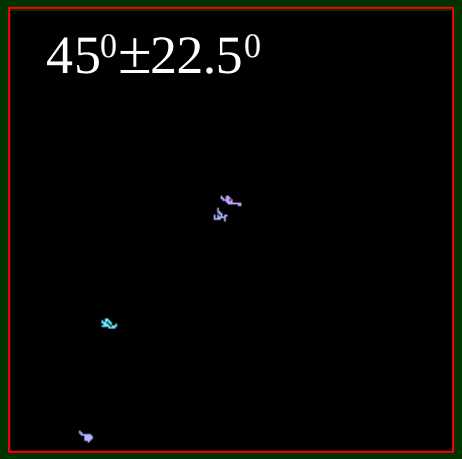
Red = horizontal Green = vertical

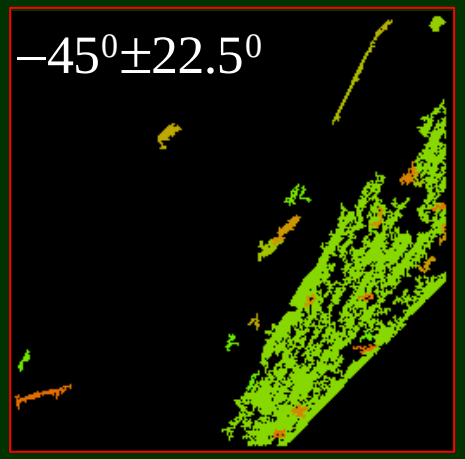
natural images









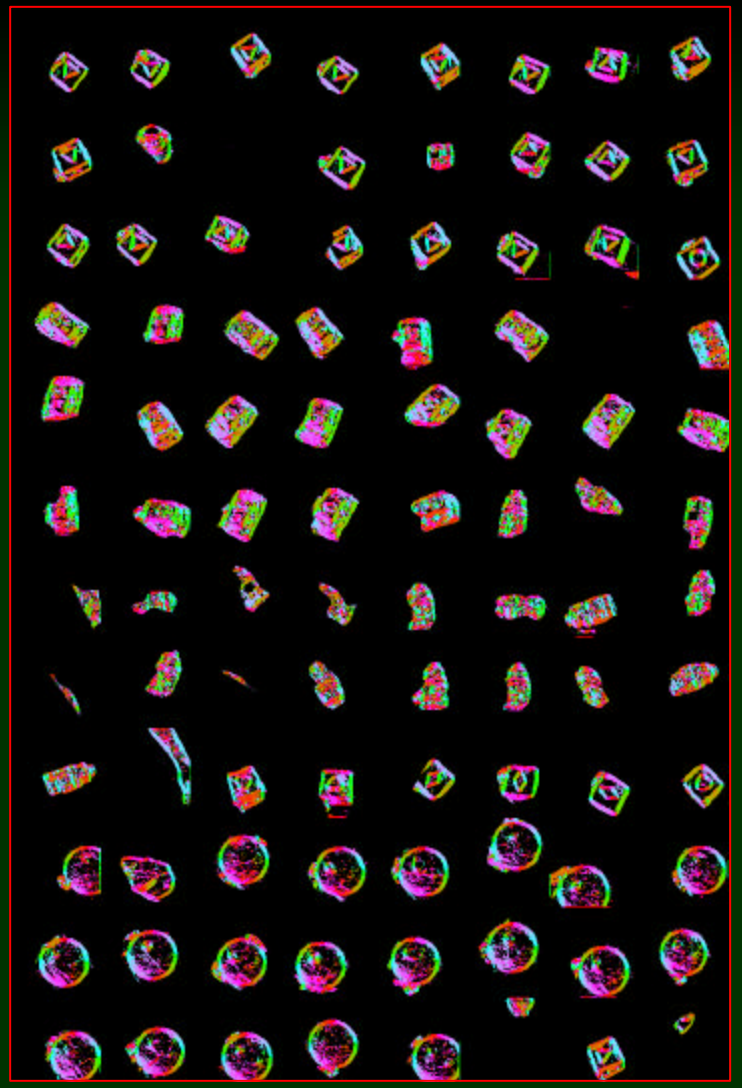


Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

talk overview

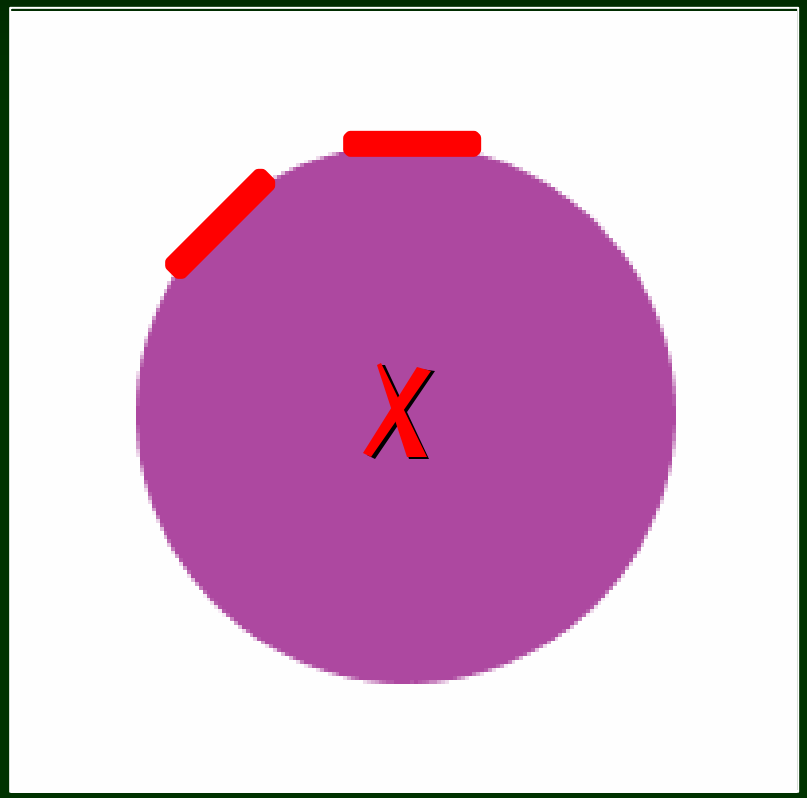
- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

on to object detection…



Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

on to object detection…

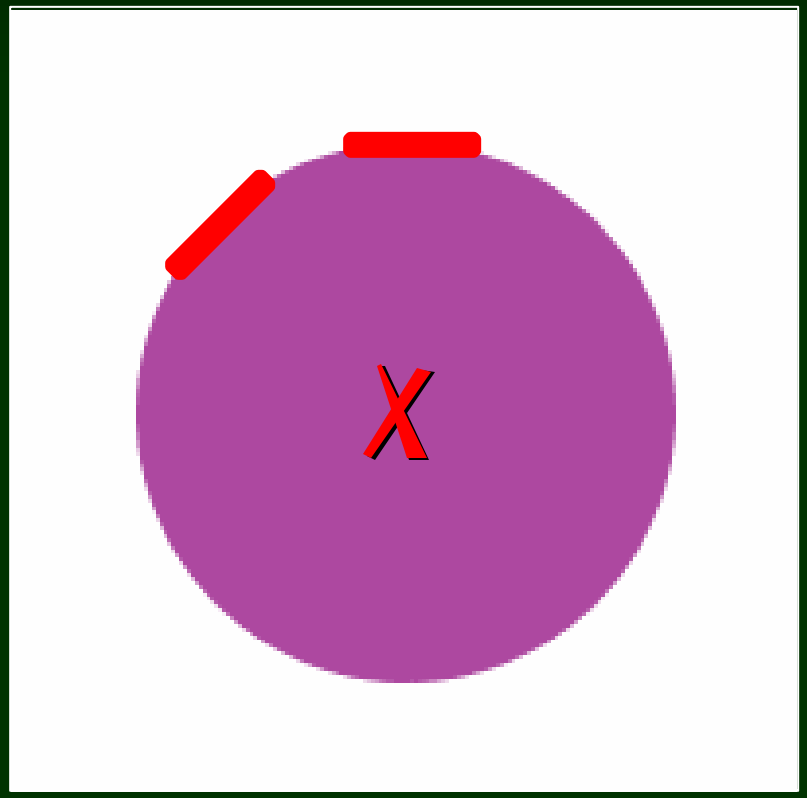


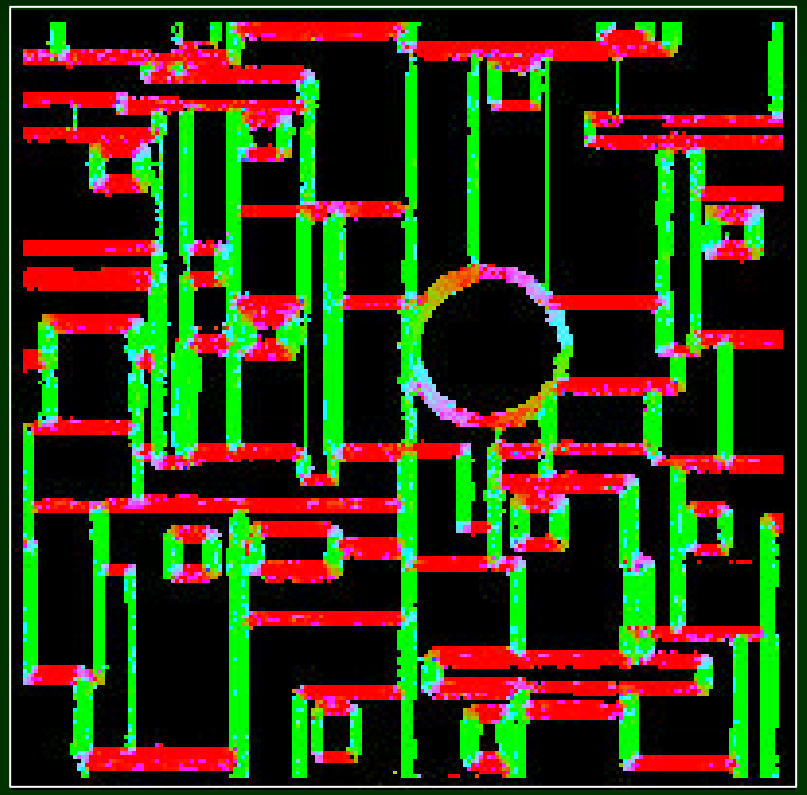
look for this…

…in this

Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

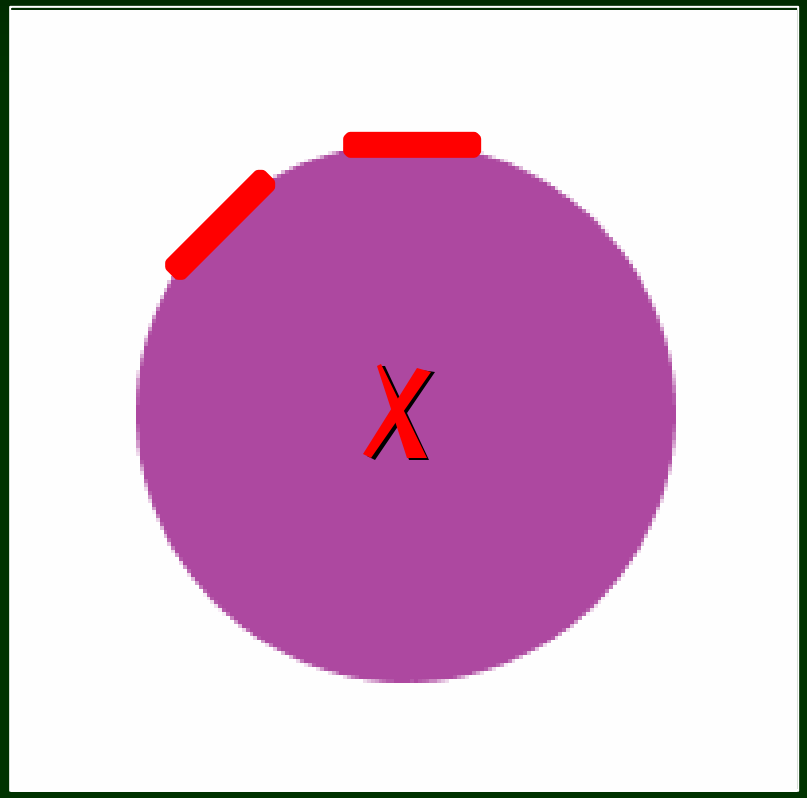
on to object detection…

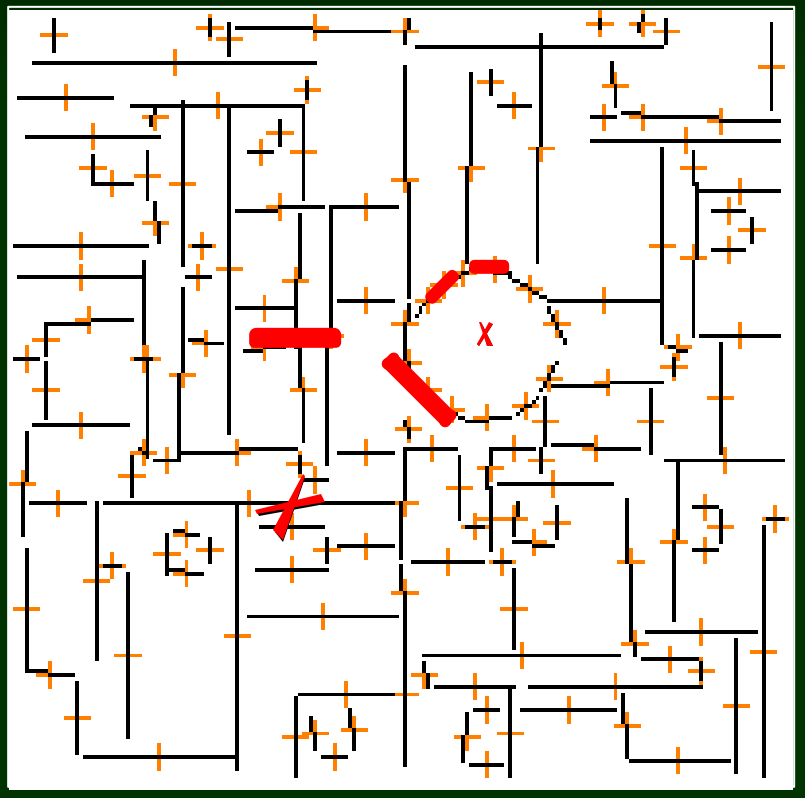




Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

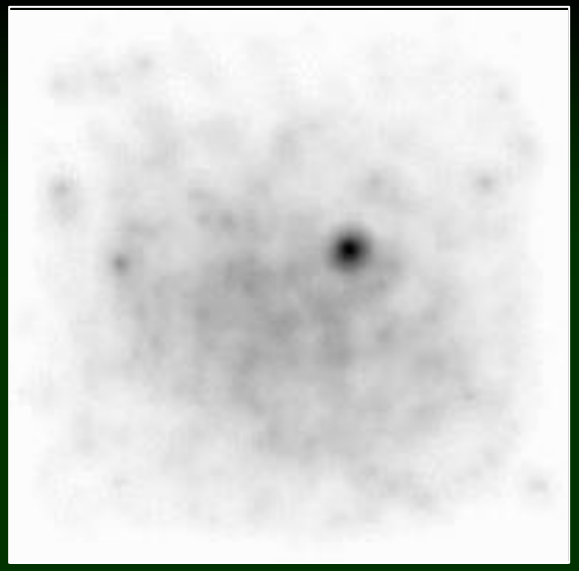
on to object detection…



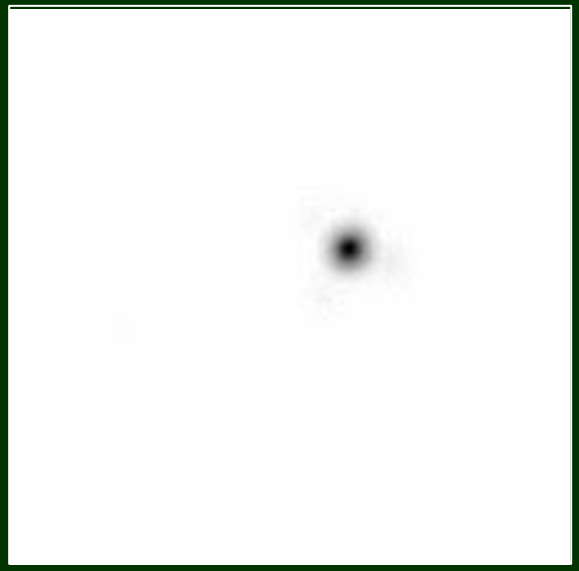


on to object detection…

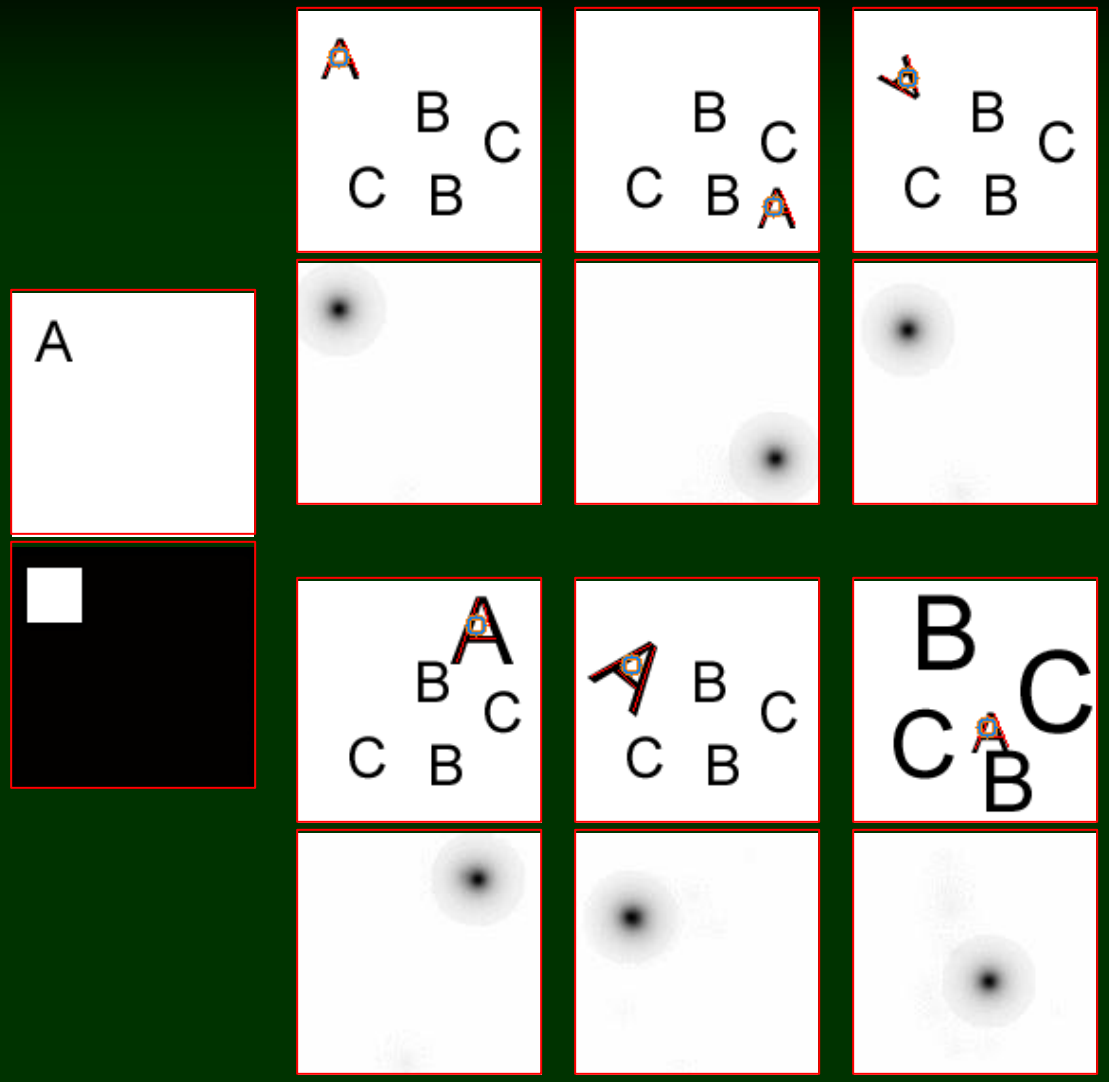
geometry alone



geometry + color

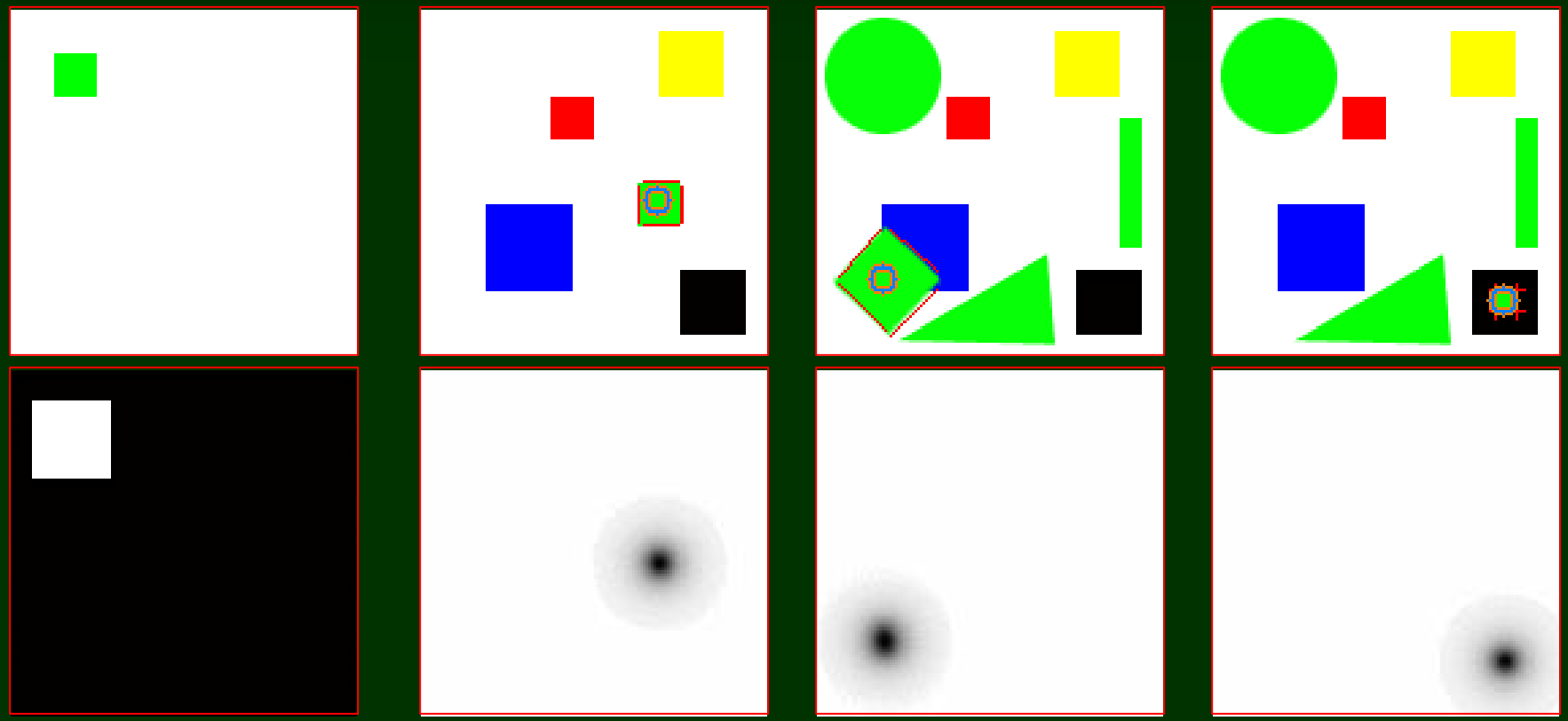


other examples

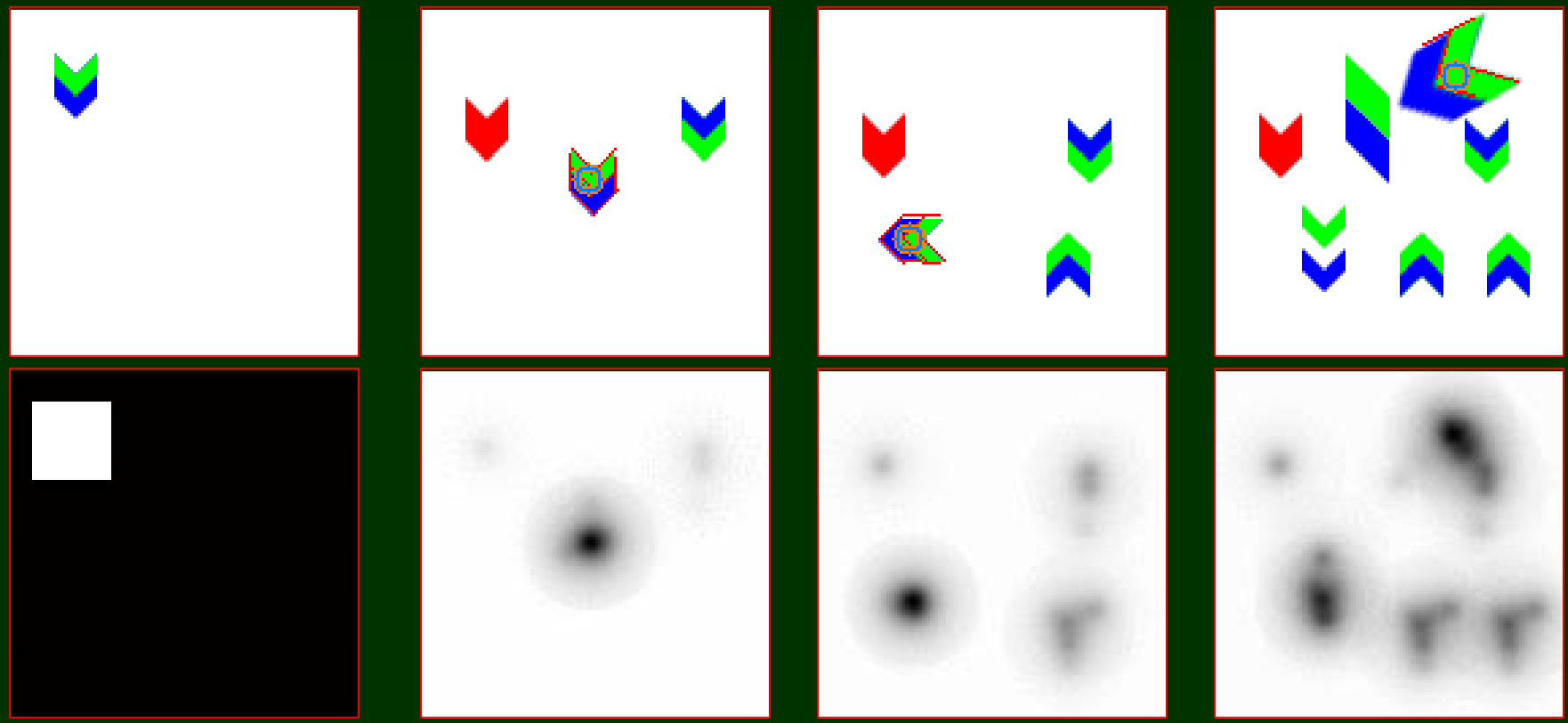


Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

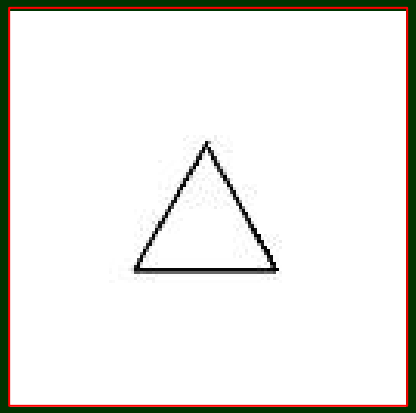
other examples



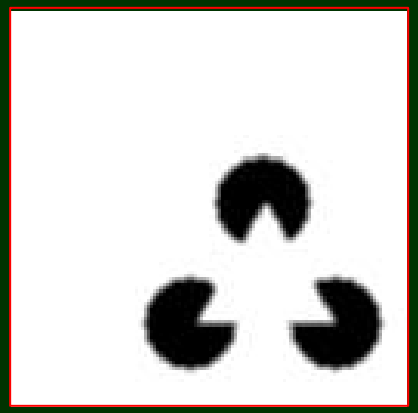
other examples



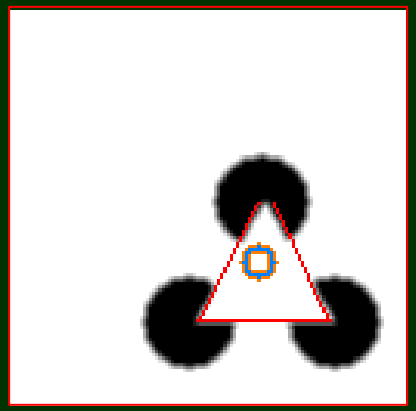
just for fun

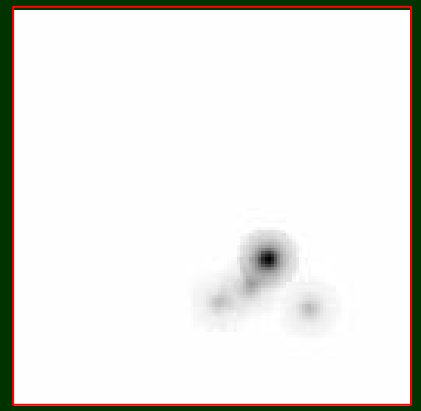


look for this…



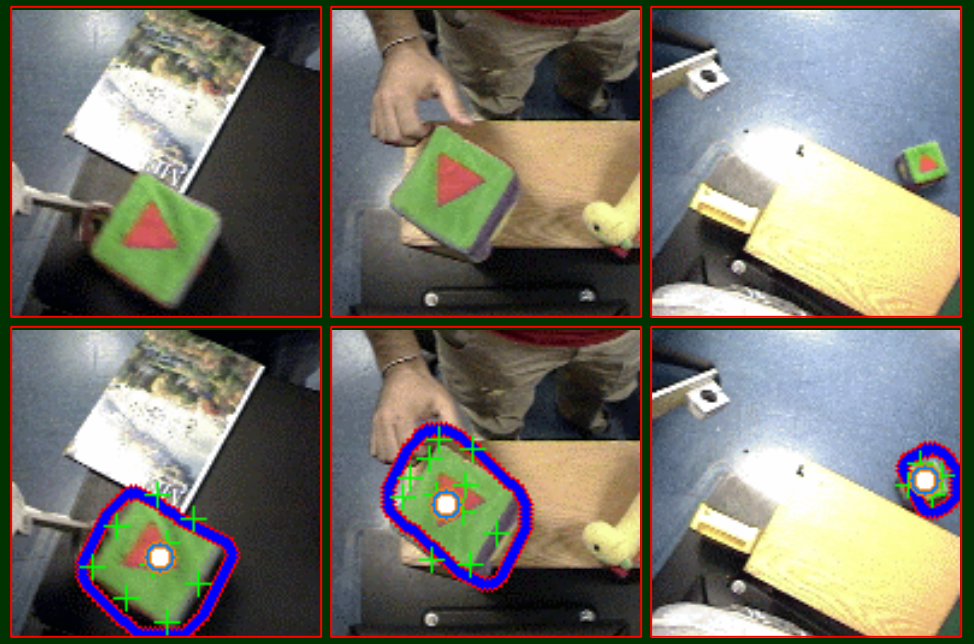
…in this





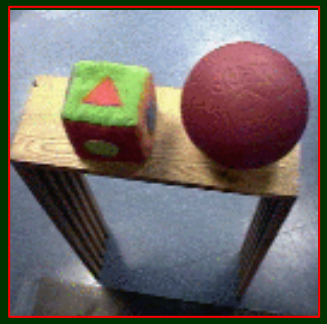
result

real object in real images

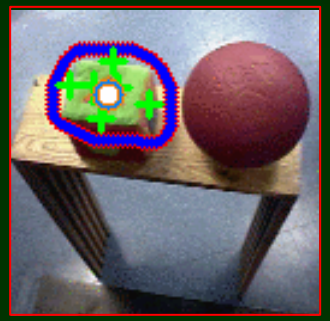


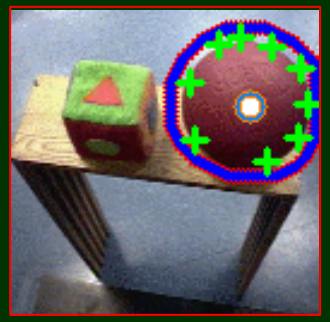
yellow on yellow

multiple objects

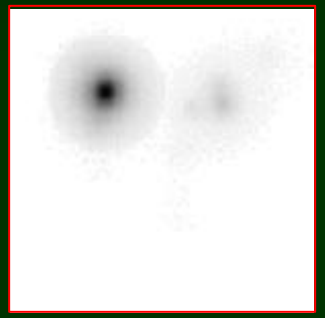


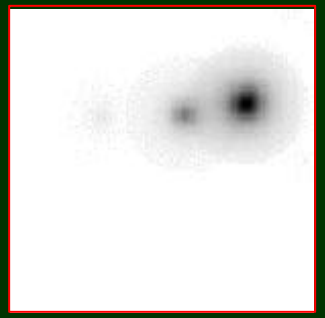
camera image





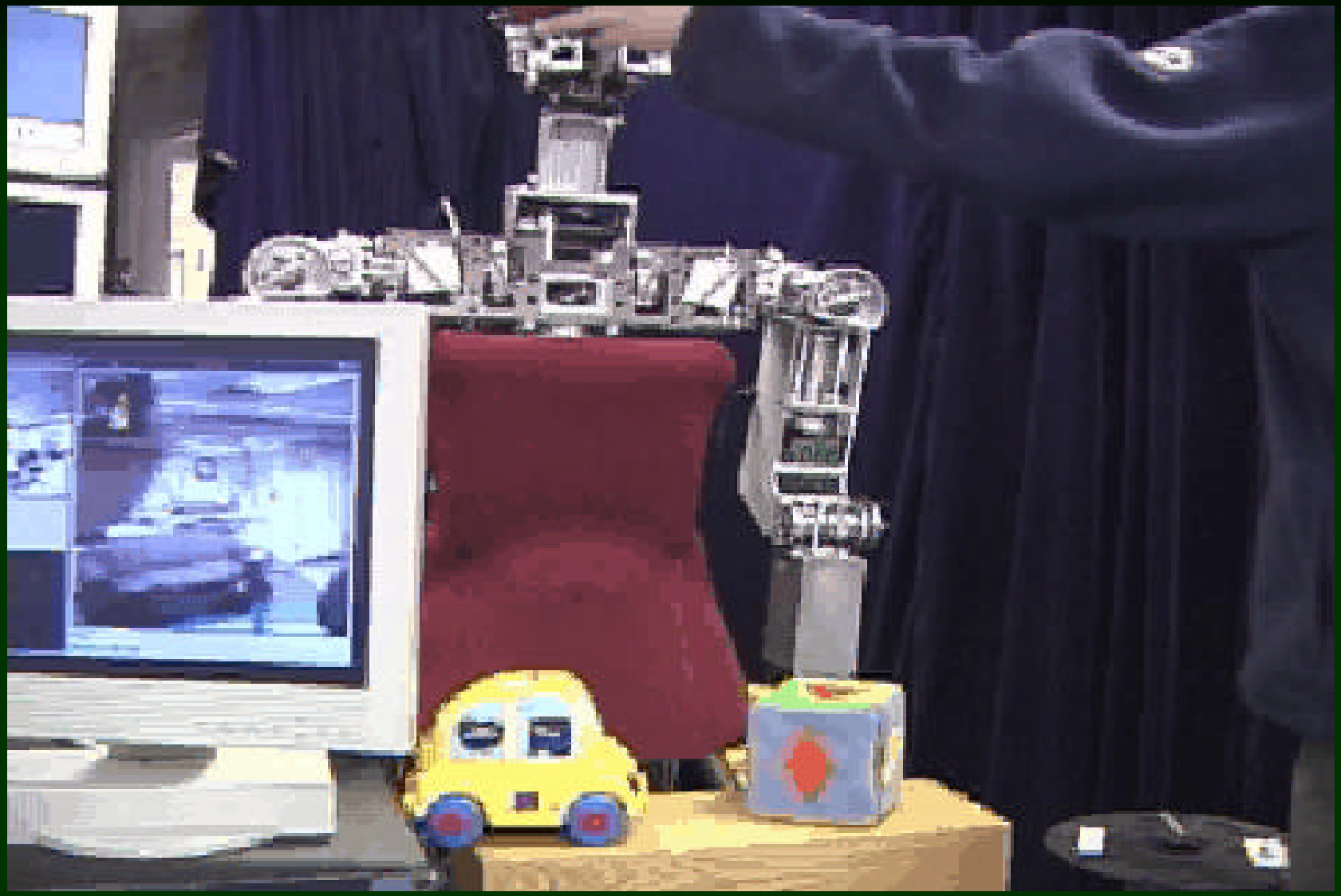
implicated edges
found and grouped





response for
each object

attention

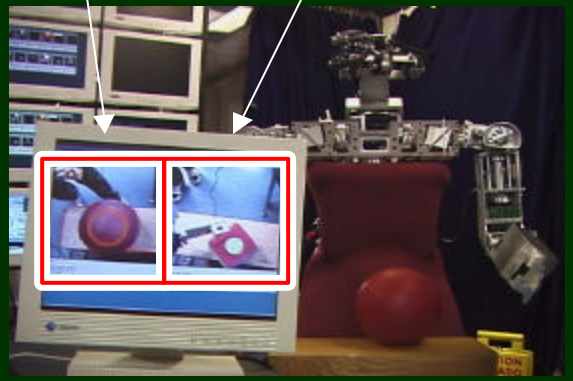


Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

first time seeing a ball

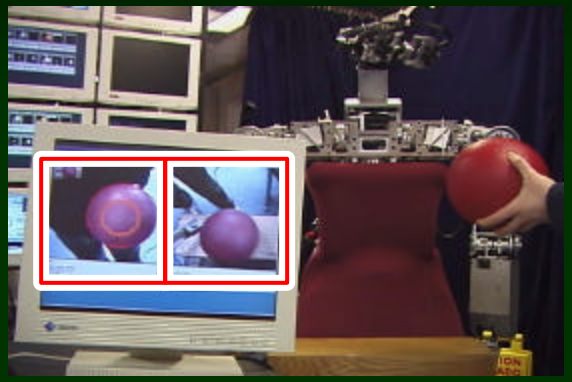
robot's current view

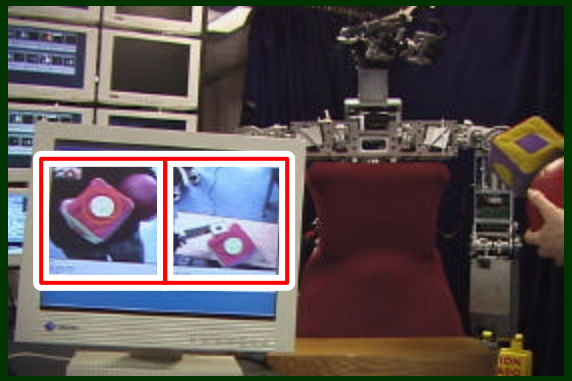
recognized object (as seen during poking)



sees ball, "thinks" it is cube

pokes, segments ball





correctly differentiates ball and cube

open object recognition

Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

talk overview

- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

talk overview

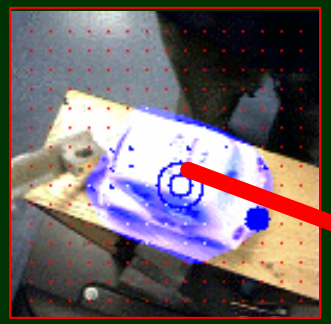
- Perceiving through experiment
  - Example: active segmentation
- Learning new perceptual abilities opportunistically
  - Example: detecting edge orientation
  - Example: object detection, segmentation, recognition
- An architecture for opportunistic learning
  - Example: learning about, and through, search activity

physically-grounded perception

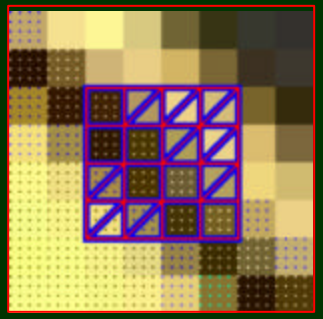
object segmentation

active segmentation

affordance exploitation
[with Giorgio Metta]
(rolling)



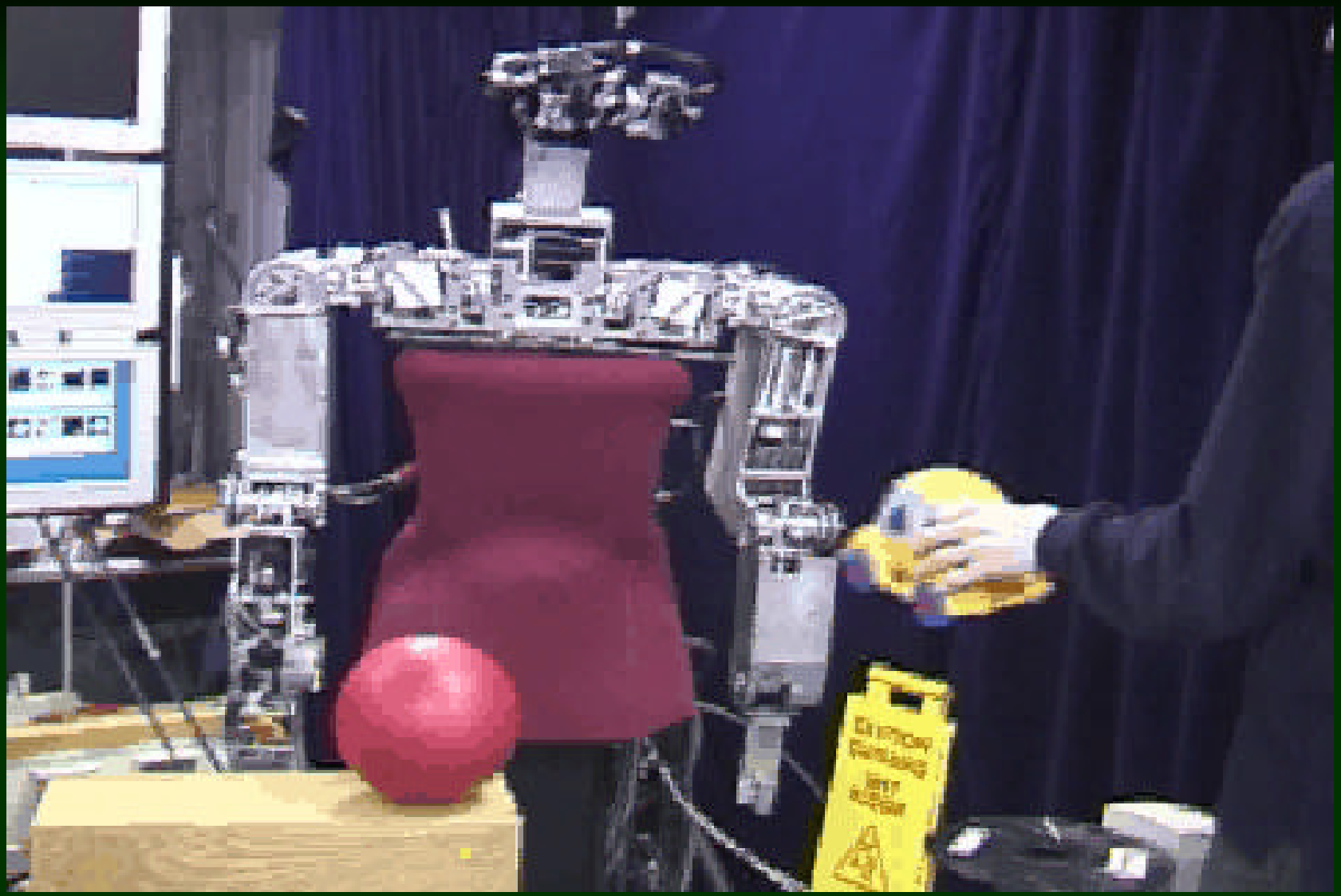
edge catalog



object detection
(recognition, localization,
contact-free segmentation)

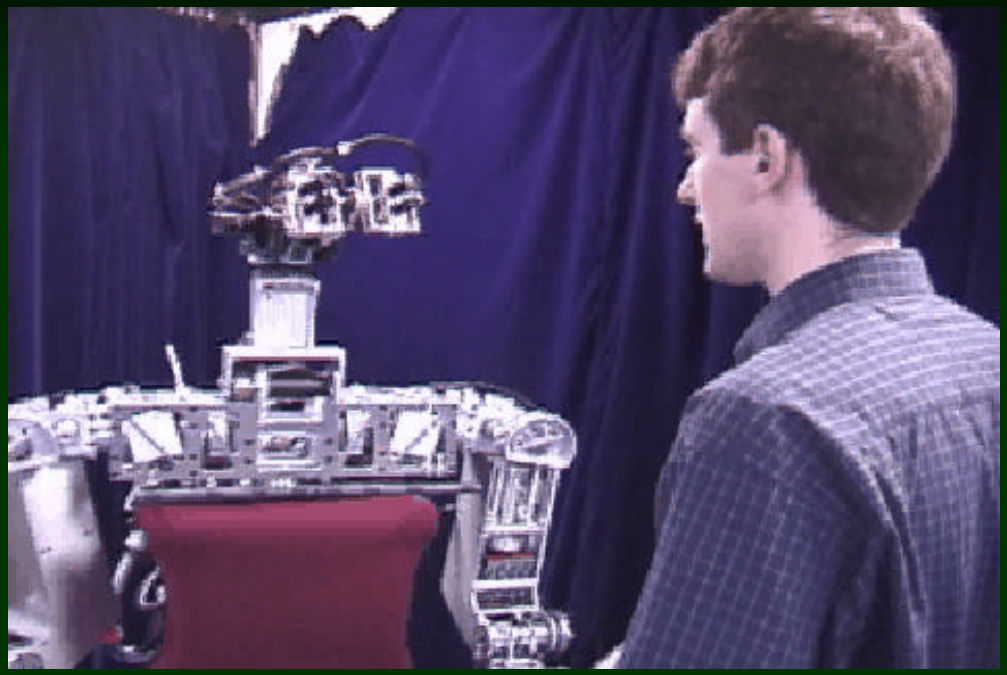
manipulator detection
(robot, human)

socially-grounded perception



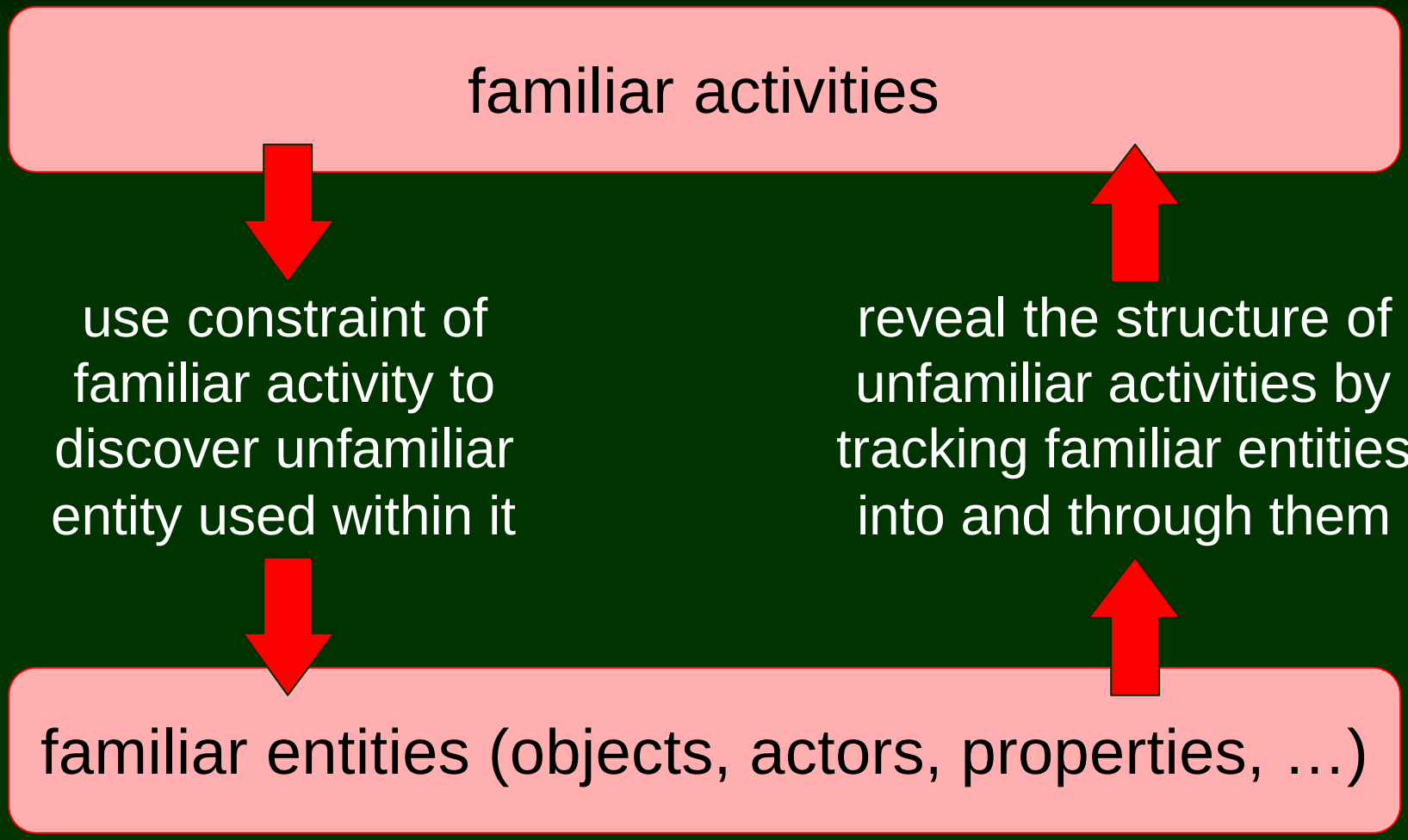
Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

socially-grounded perception



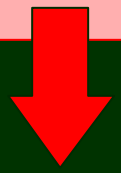
Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

– opportunistic architecture: a virtuous circle –



a virtuous circle

poking, chatting



discover car, ball, and cube through poking; discover their names through chatting

car, ball, cube, and their names

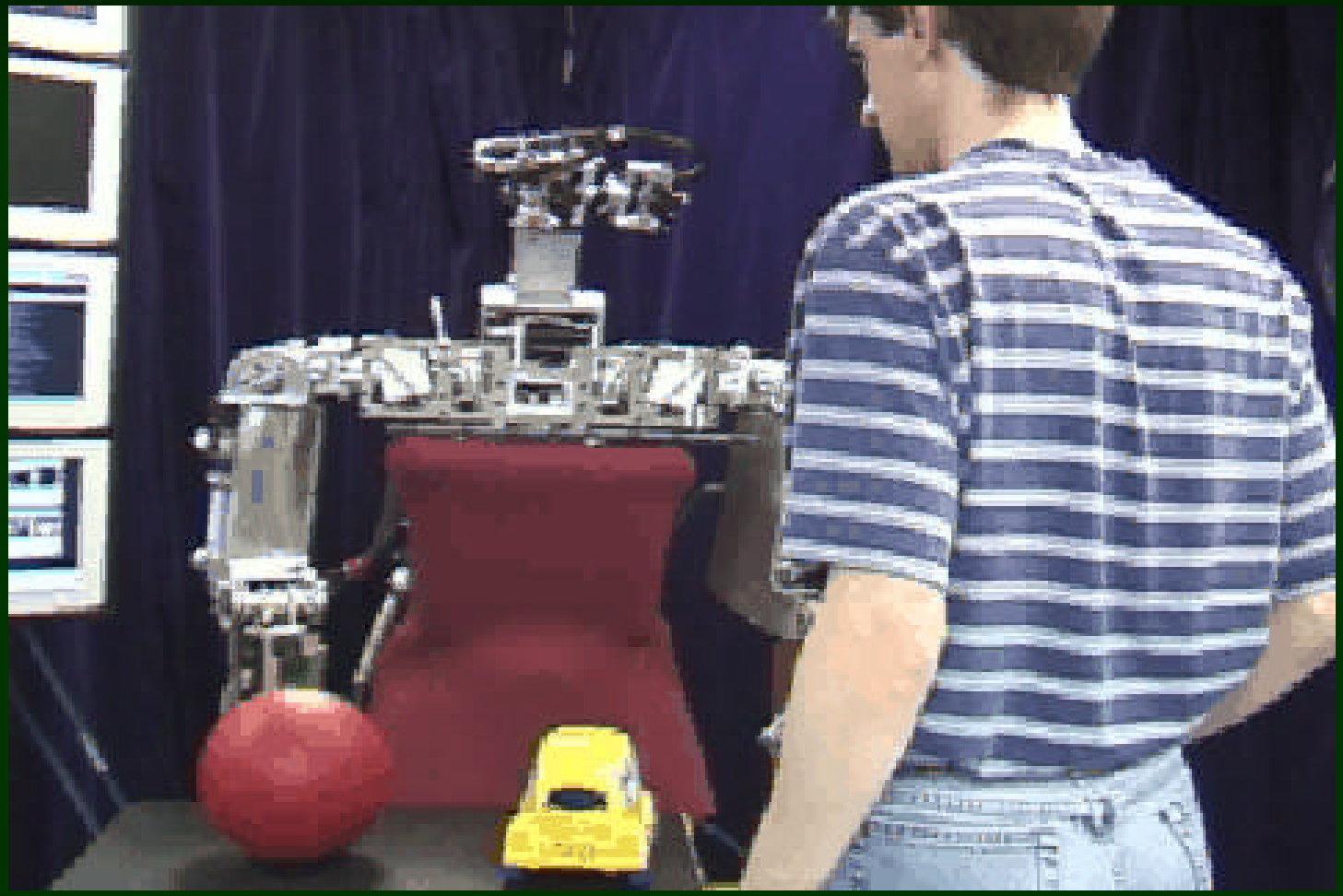
a virtuous circle

poking, chatting, **search**

follow named objects into search activity, and observe the structure of search

car, ball, cube, and their names

learning about search



Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

a virtuous circle

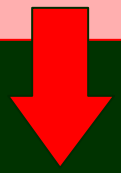
poking, chatting, **search**

follow named objects into search activity, and observe the structure of search

car, ball, cube, and their names

a virtuous circle

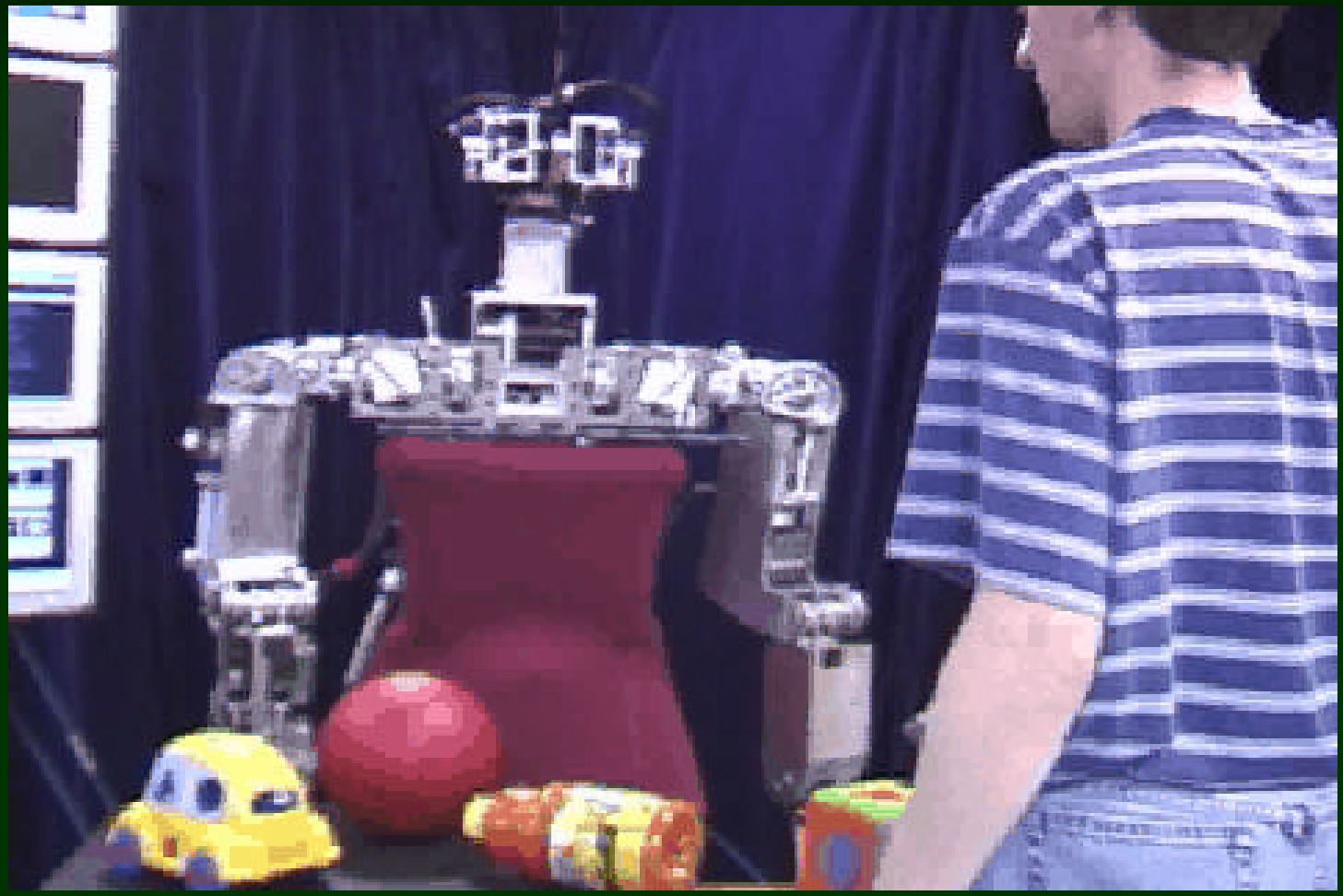
poking, chatting, searching



discover novel object
through poking, learn
its name (e.g. ‘toma’)
indirectly during search

car, ball, cube, **toma**, and their names

finding the toma



Paul Fitzpatrick, MIT CSAIL, Humanoids 2003

conclusion: why do this?

- The quest for truly flexible robots
- Humanoid form is general-purpose, mechanically flexible
- Robots that really live and work amongst us will need to be as general-purpose and adaptive perceptually as they are mechanically