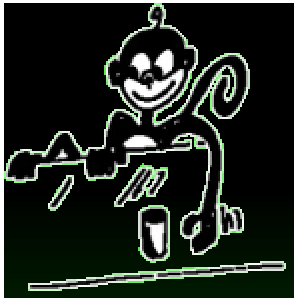
Better Vision through Manipulation

Giorgio Metta · Paul Fitzpatrick

Humanoid Robotics Group

MIT AI Lab



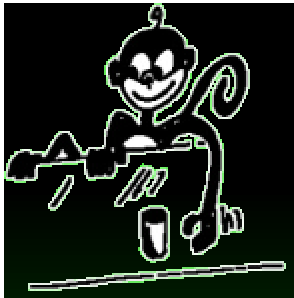
Vision & Manipulation

In robotics, vision is often used to guide manipulation

But manipulation can also guide vision

Important for…

- Correction recovering when perception is misleading
- Experimentation progressing when perception is ambiguous
- Development bootstrapping when perception is dumb



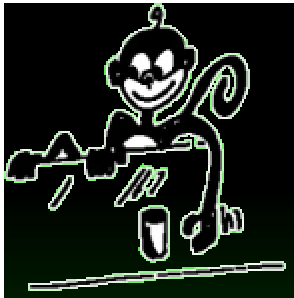
Linking Vision & Manipulation

A link from robotics

- Active vision: Good motor strategies can simplify perceptual problems

A link from neuroscience

- Mirror neurons: Relating perceived actions of others with own action may simplify learning tasks



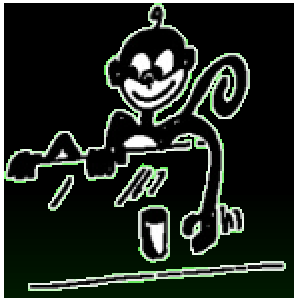
Linking Vision & Manipulation

A link from robotics

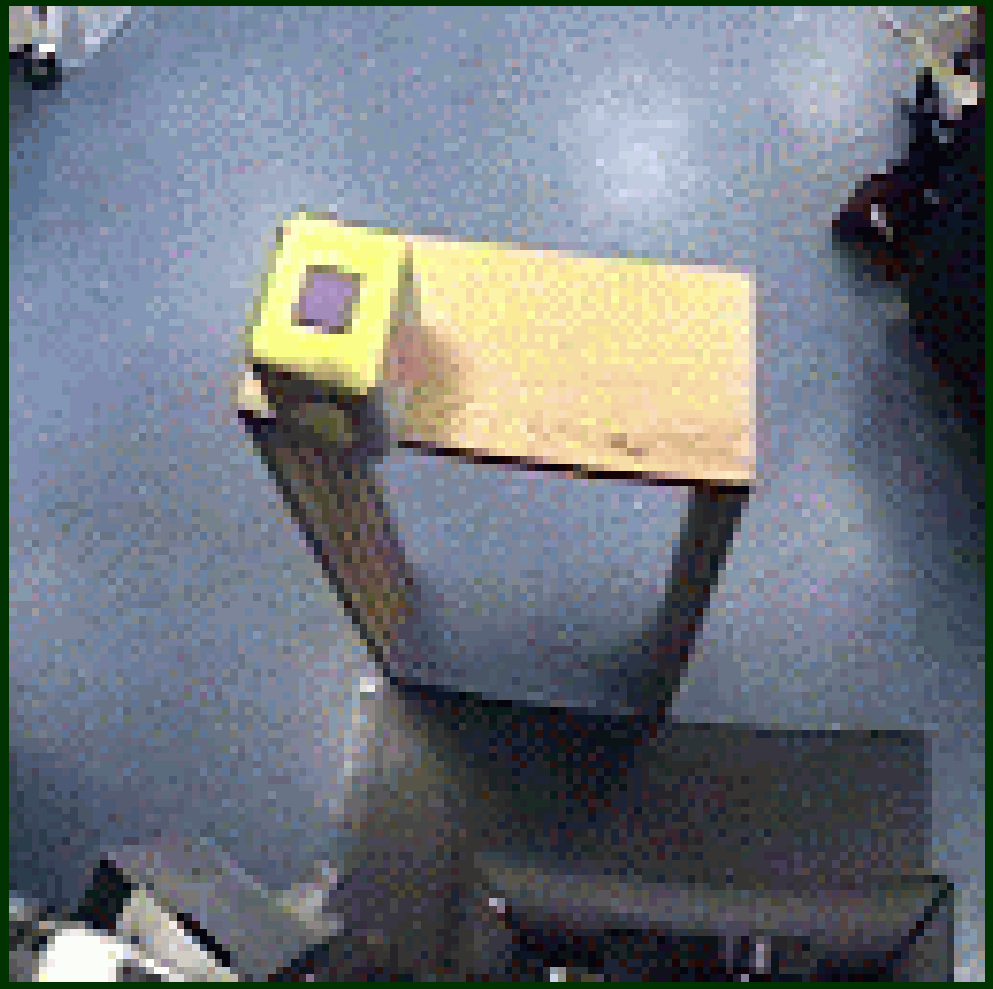
- Active vision: Good motor strategies can simplify perceptual problems

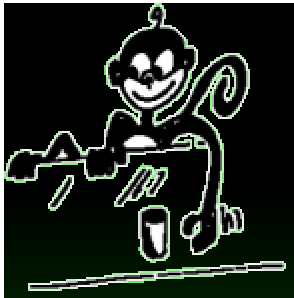
A link from neuroscience

- Mirror neurons: Relating perceived actions of others with own action may simplify learning tasks

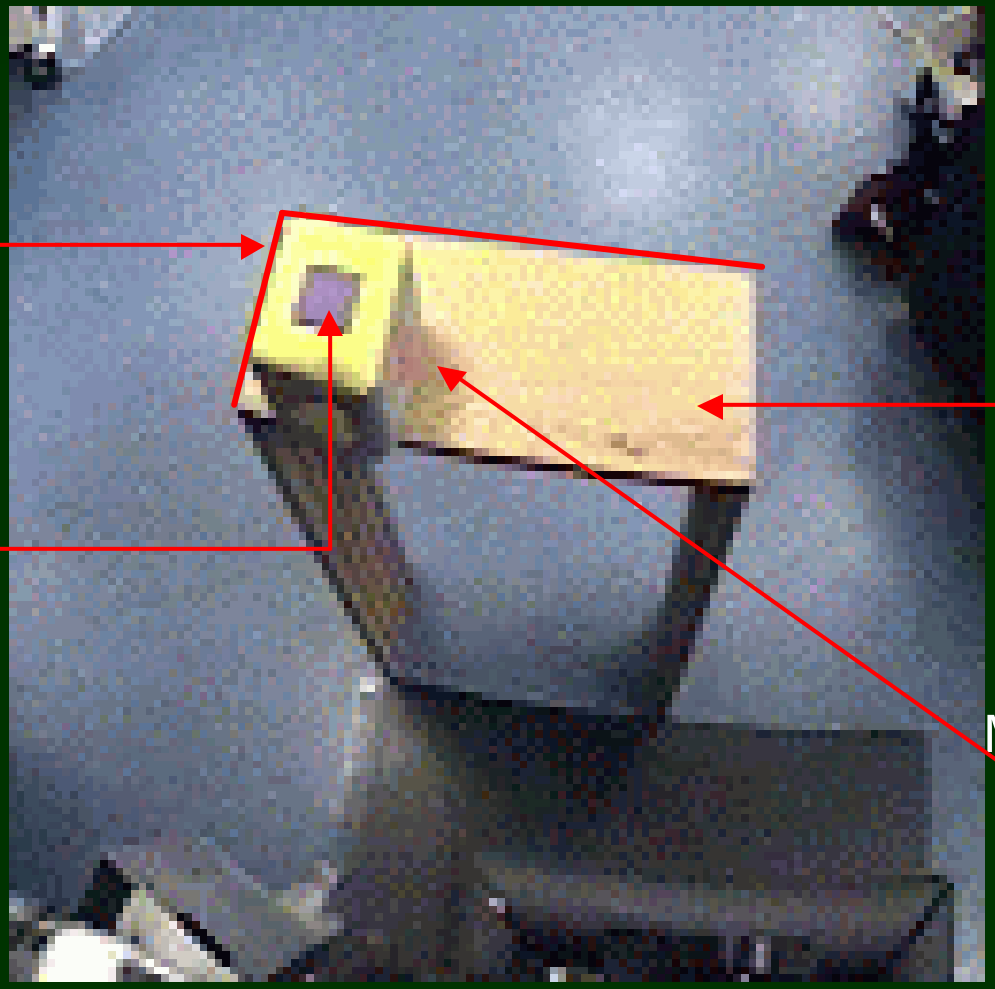


A Simple Scene?





A Simple Scene?

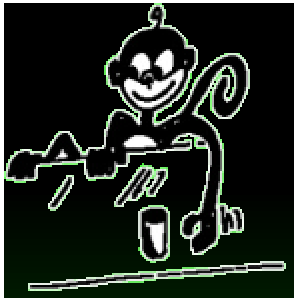


Edges of table and cube overlap

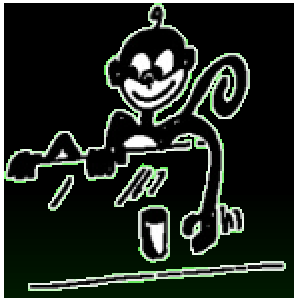
Cube has misleading surface pattern

Color of cube and table are poorly separated

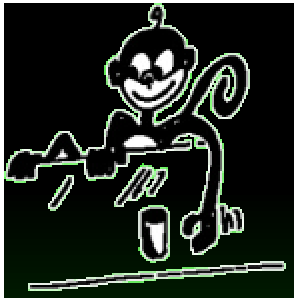
Maybe some cruel grad-student glued the cube to the table



Active Segmentation

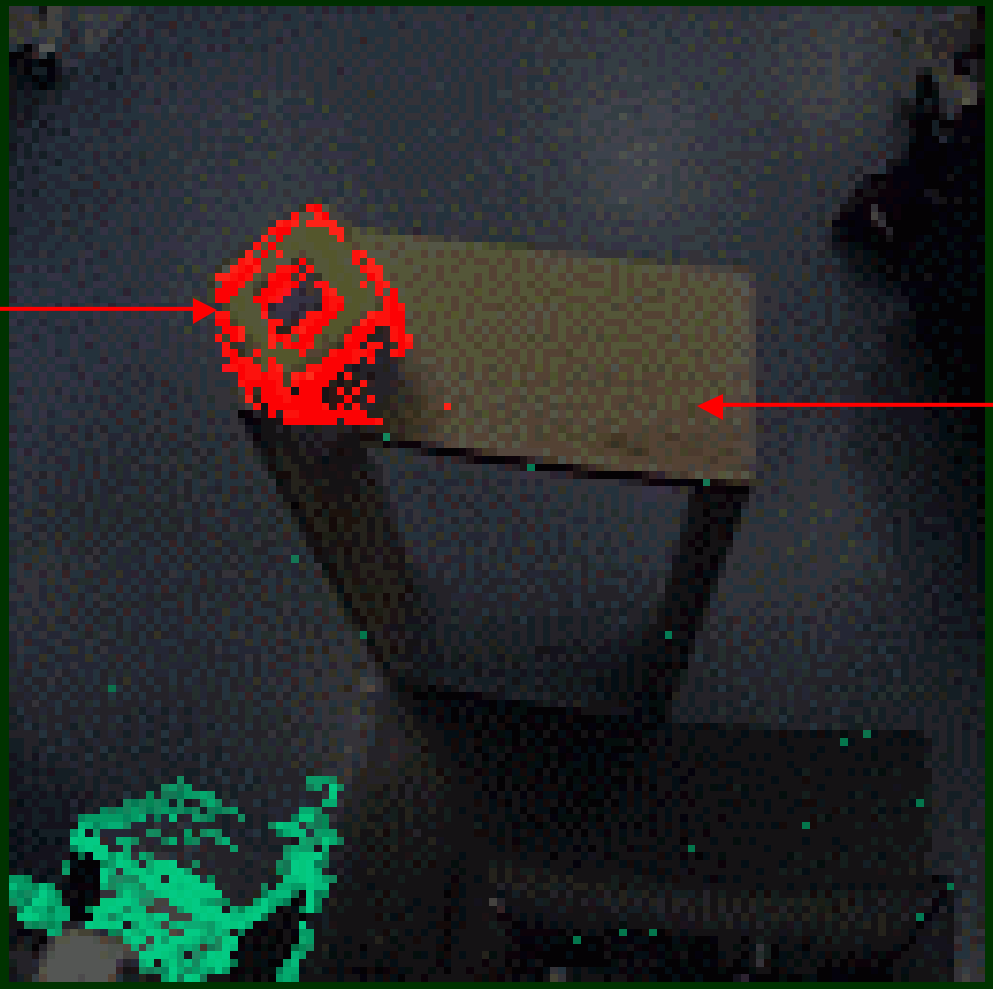


Active Segmentation



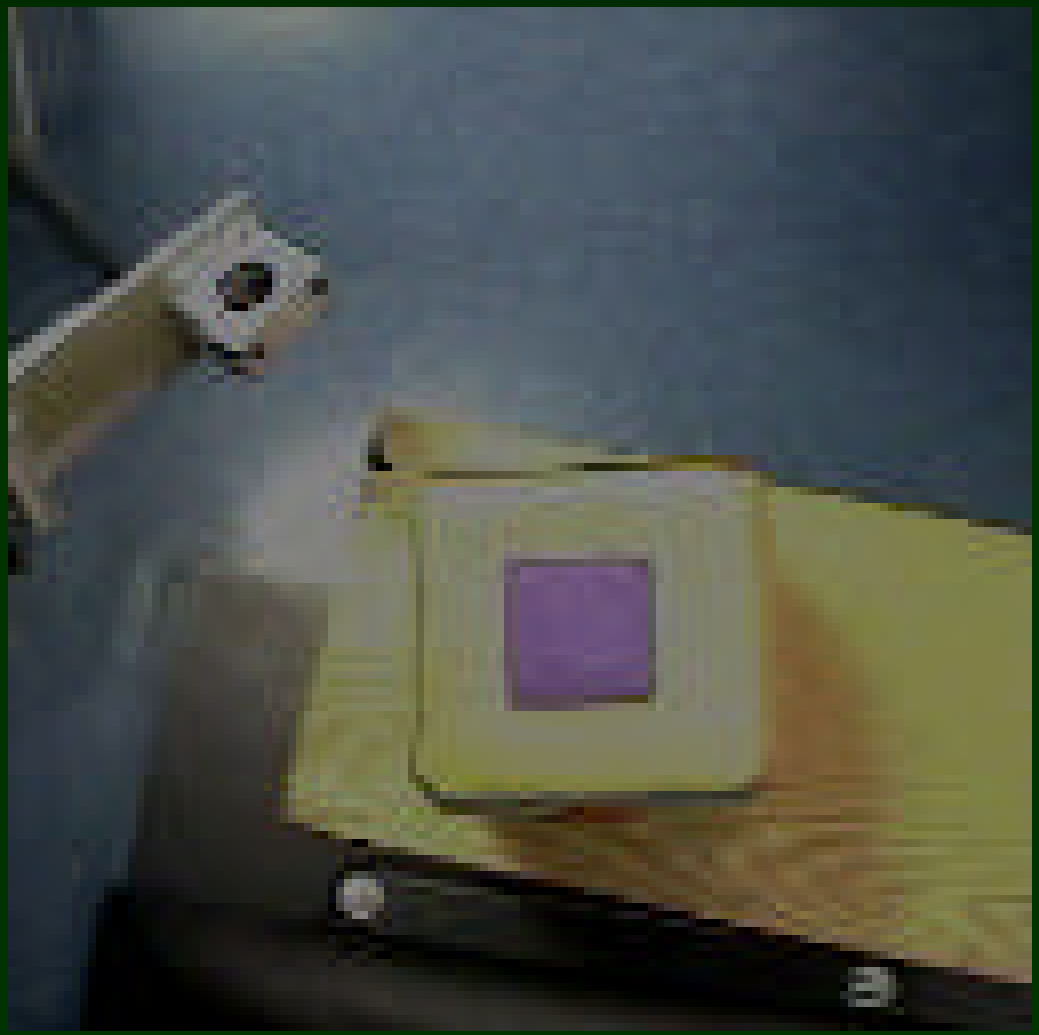
Result

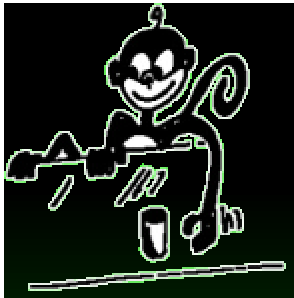
No confusion between cube and own texture



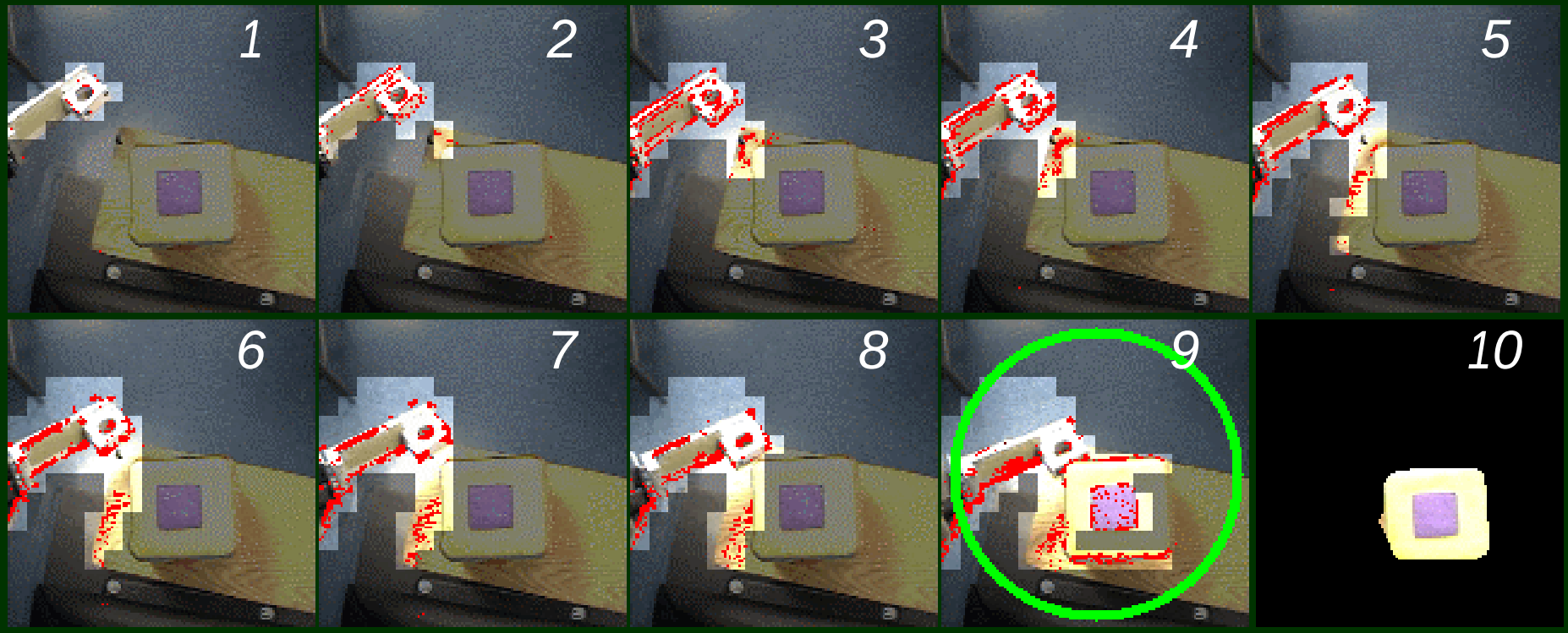
No confusion between cube and table

Point of Contact



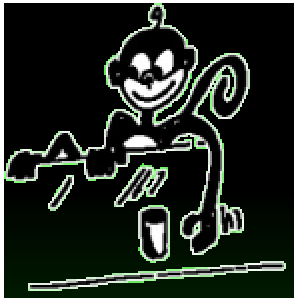


Point of Contact



Motion spreads continuously
(arm or its shadow)

Motion spreads suddenly, faster than the arm itself → contact



Segmentation

Side tap

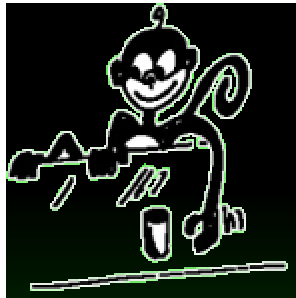
Back slap

Prior to impact

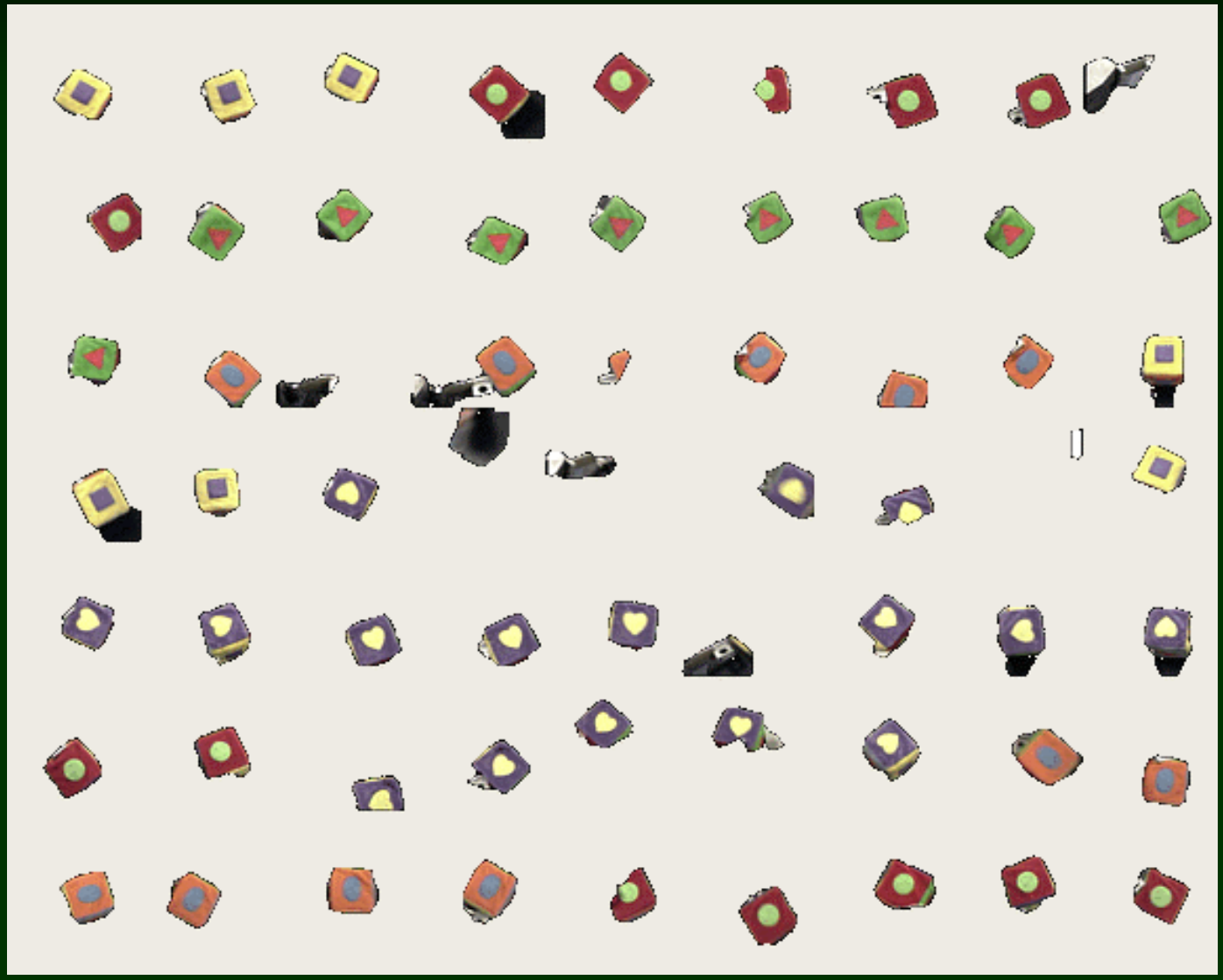
Impact event

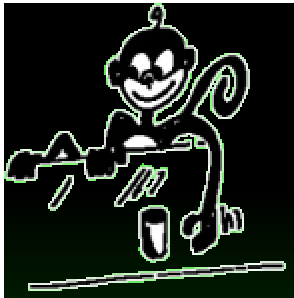
Motion caused
(red = novel,
Purple/blue = discounted)

Segmentation
(green/yellow)

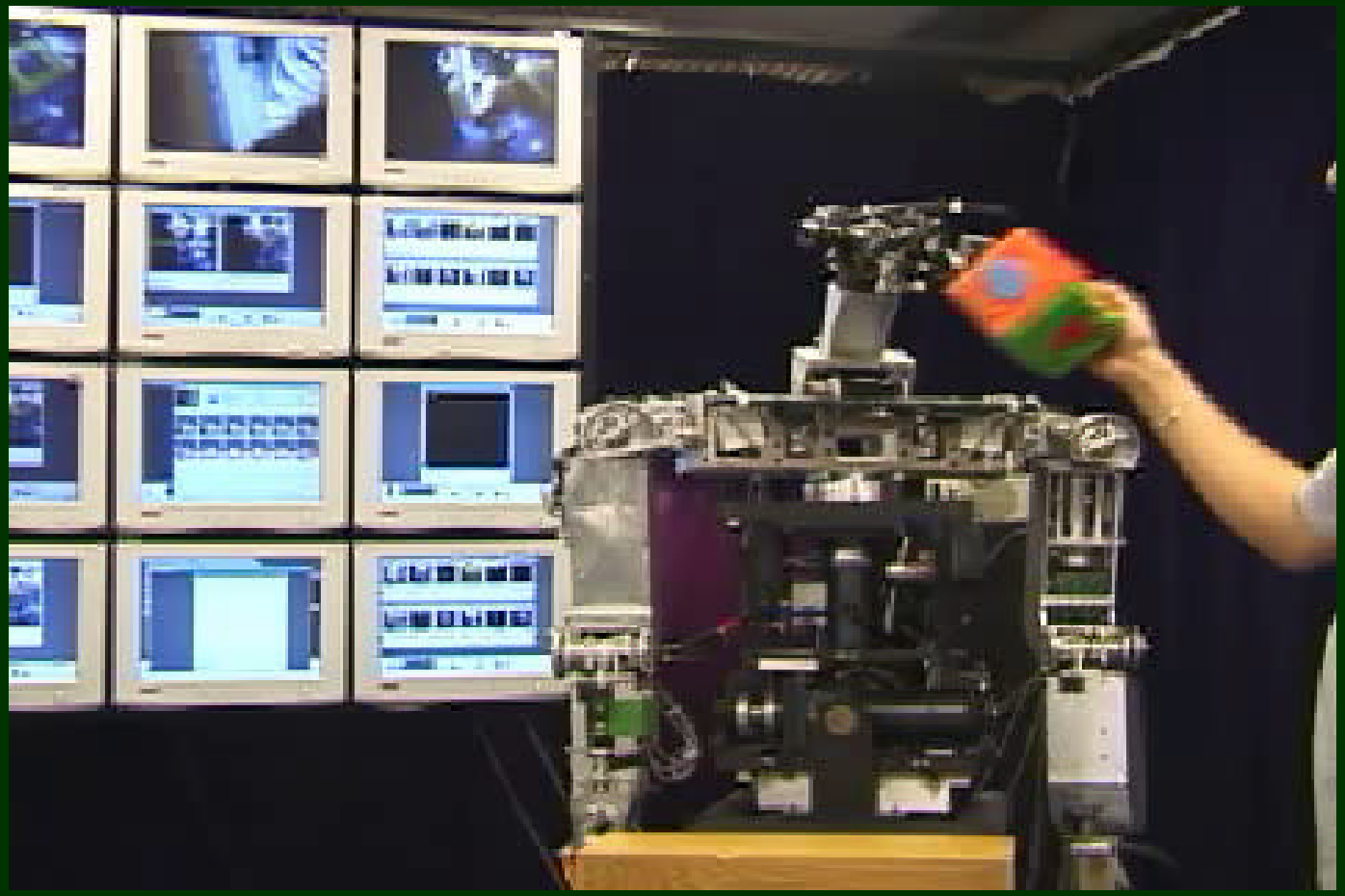


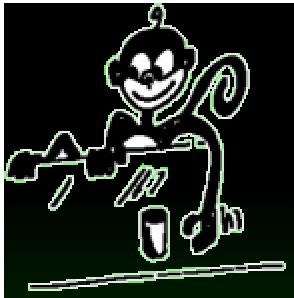
Typical results





A Complete Example





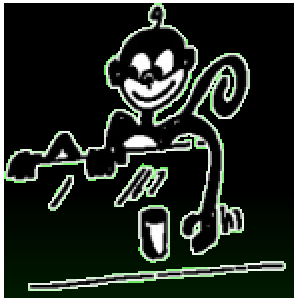
Linking Vision & Manipulation

A link from robotics

- Active vision: Good motor strategies can simplify perceptual problems

A link from neuroscience

- Mirror neurons: Relating perceived actions of others with own action may simplify learning tasks



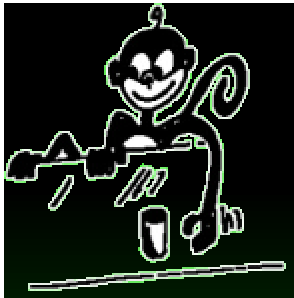
Linking Vision & Manipulation

A link from robotics

- Active vision: Good motor strategies can simplify perceptual problems

A link from neuroscience

- Mirror neurons: Relating perceived actions of others with own action may simplify learning tasks



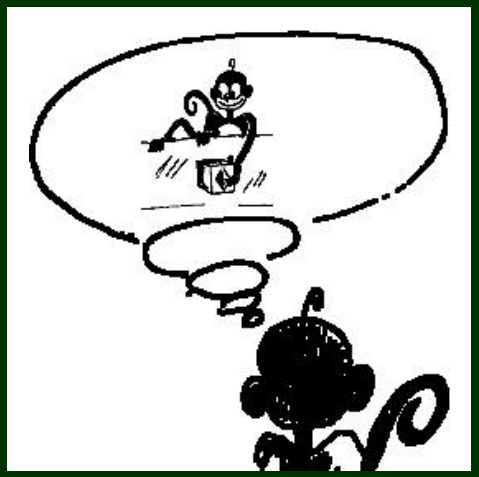
Viewing Manipulation

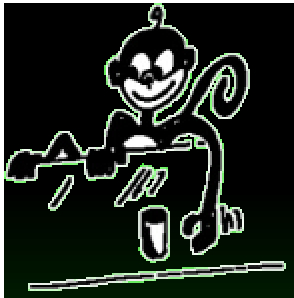
"Canonical neurons"

Active when manipulable objects are presented visually

"Mirror neurons"

Active when another individual is seen performing manipulative gestures



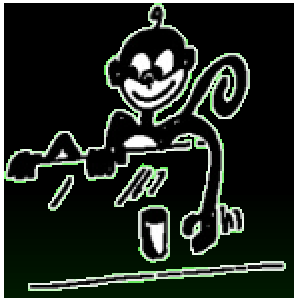


Simplest Form of Manipulation

What is the simplest possible manipulative gesture?

- Contact with object is necessary; can't do much without it
- Contact with object is sufficient for certain classes of affordances to come into play (e.g. rolling)
- So can use various styles of poking/prodding/tapping/swiping as basic manipulative gestures

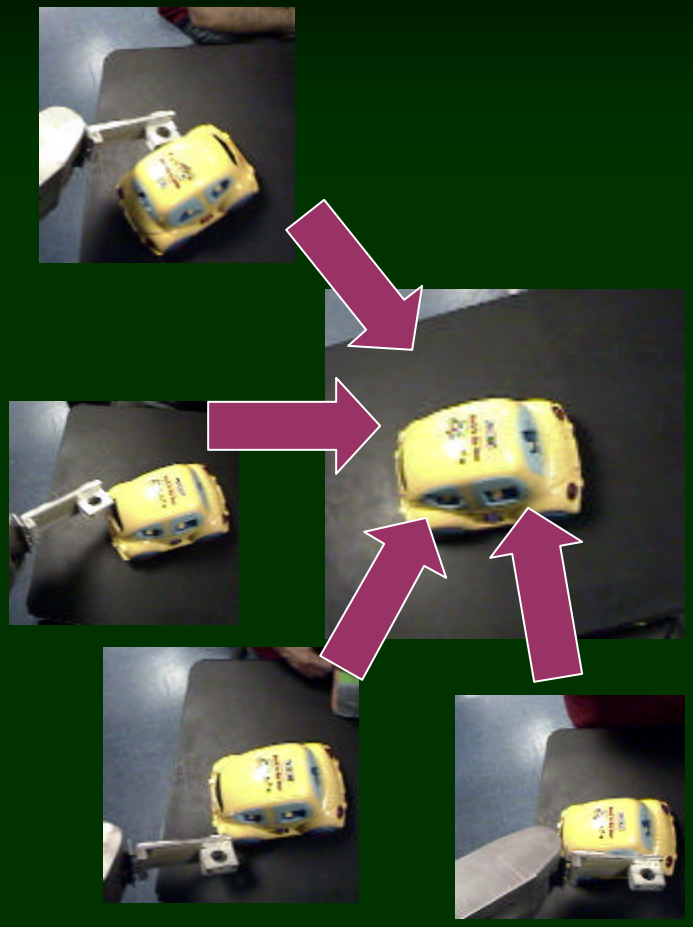
(if willing to omit the *manus* from manipulation…)



Gesture “Vocabulary”

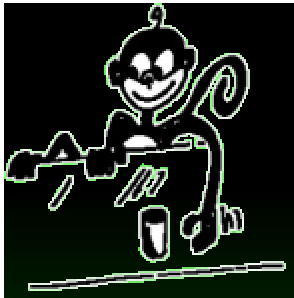
pull in

side tap

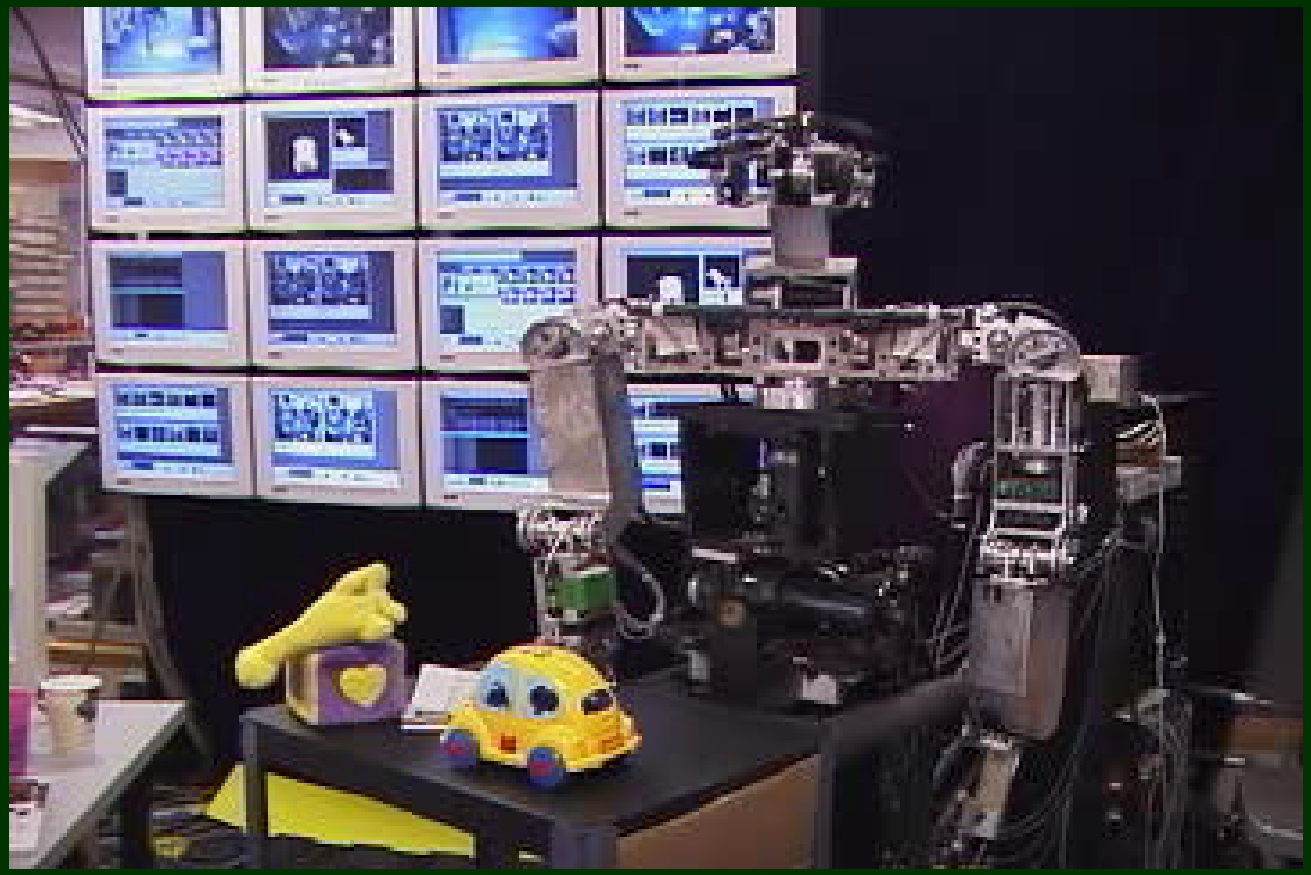


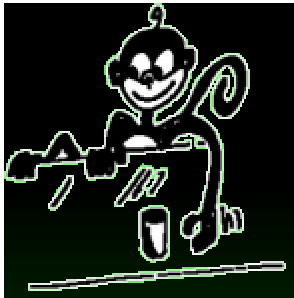
push away

back slap

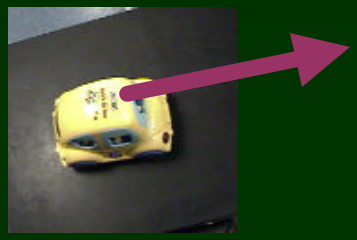


Exploring an Affordance: Rolling

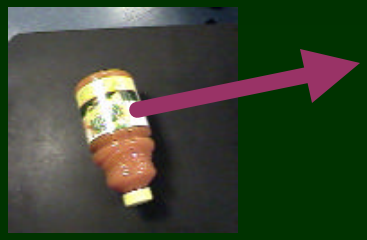




Exploring an Affordance: Rolling

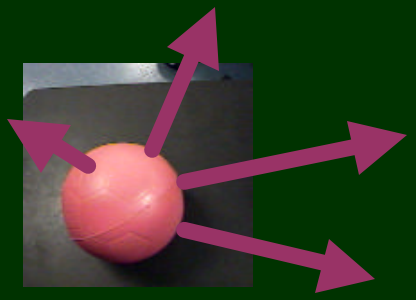


A toy car: it rolls in the direction of its principal axis

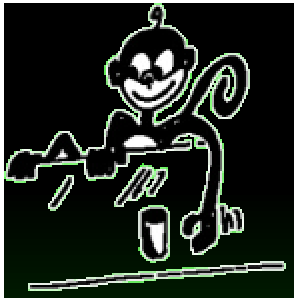


A bottle: it rolls orthogonal to the direction of its principal axis

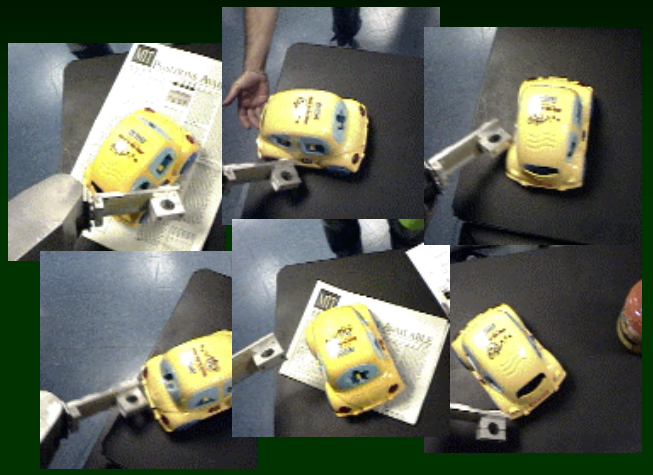
A toy cube: it doesn't roll, it doesn't have a principal axis

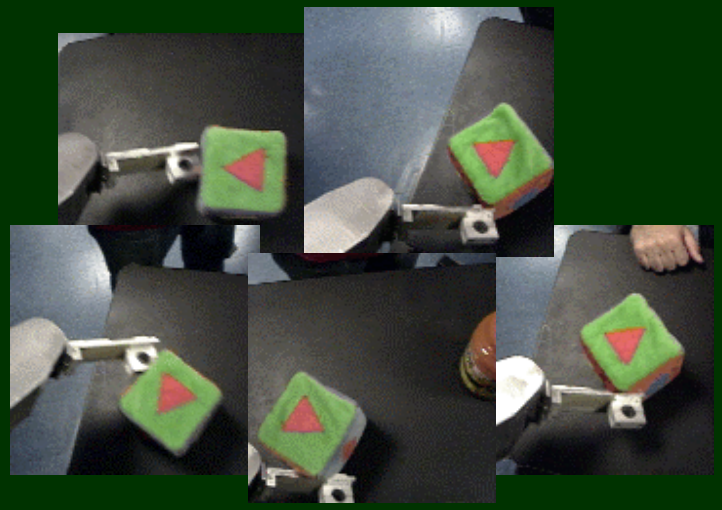


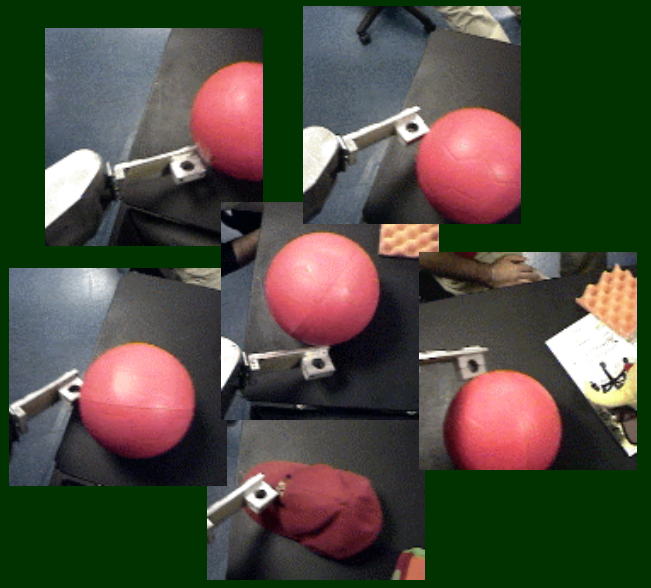
A ball: it rolls, it doesn't have a principal axis

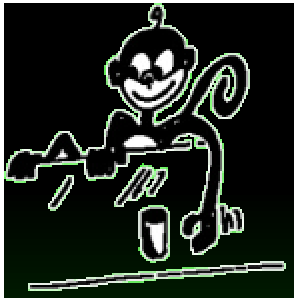


Forming Object Clusters





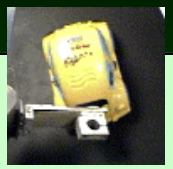




Preferred Direction of Motion

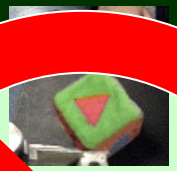
Bottle, “pointiness”=0.13

Rolls at right angles to principal axis

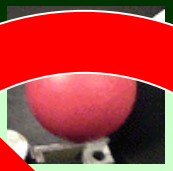


Car, “pointiness”=0.07

Rolls along principal axis



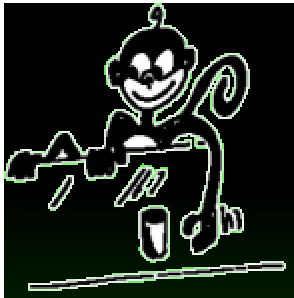
Cube, “pointiness”=0.0



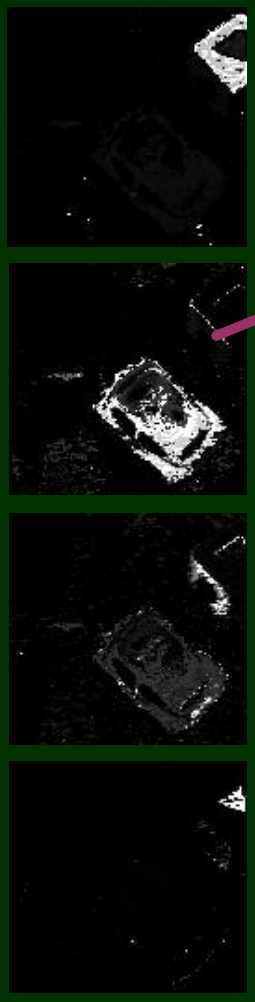
Ball, “pointiness”=0.02

estimated probability of occurrence

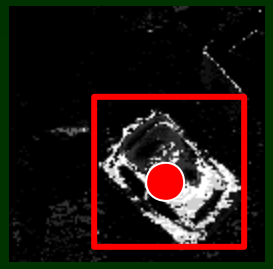
difference between angle of motion and principal axis of object [degrees]



Closing the Loop

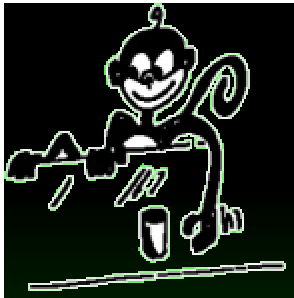


identify and localize object

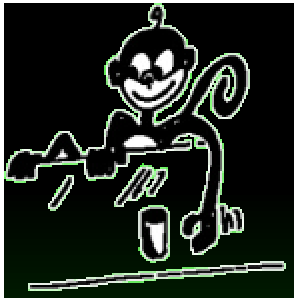


search rotation

Previously-poked prototypes



Closing The Loop: Very Preliminary!



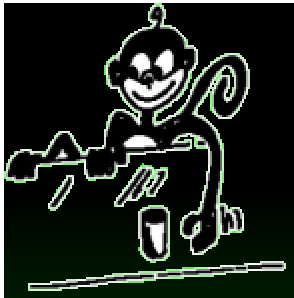
Conclusions

Poking works!

Will always be an important perceptual fall-back

Simple, yet already enough to let robot explore world of objects and motion

Stepping stone to greater things?



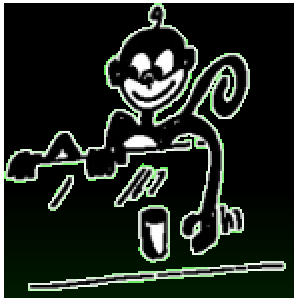
Acknowledgements

This work was funded by

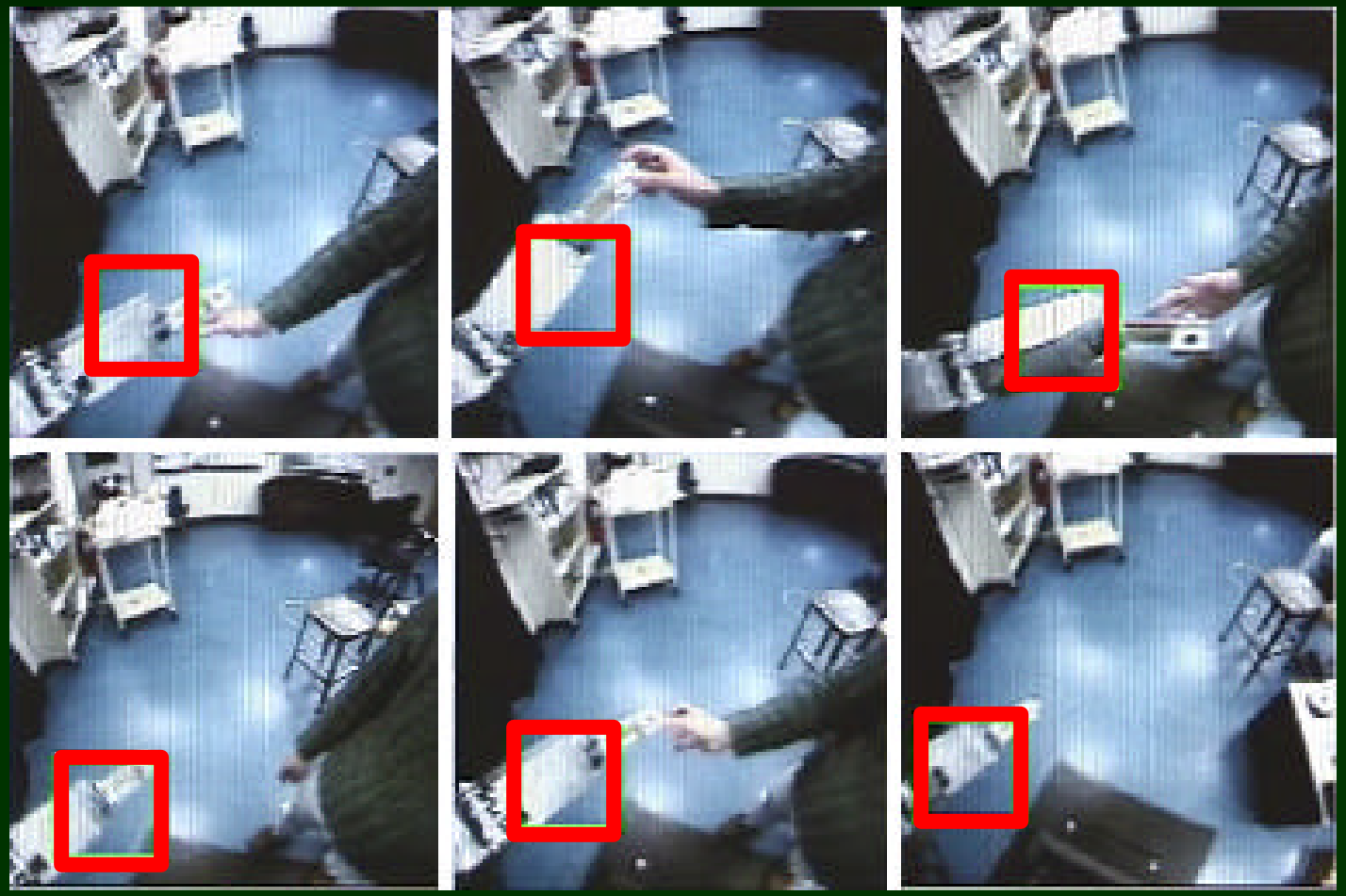
DARPA
as part of the
"Natural Tasking of Robots Based on Human Interaction Cues"
project under contract number DABT 63-00-C-10102

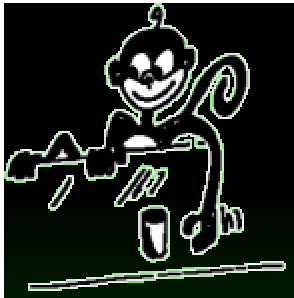
and by

NTT
as part of the
NTT/MIT Collaboration Agreement

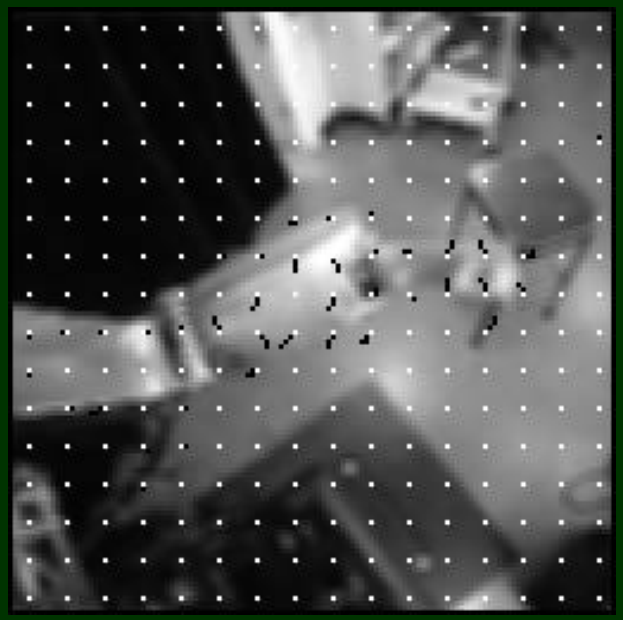


Training Visual Predictor

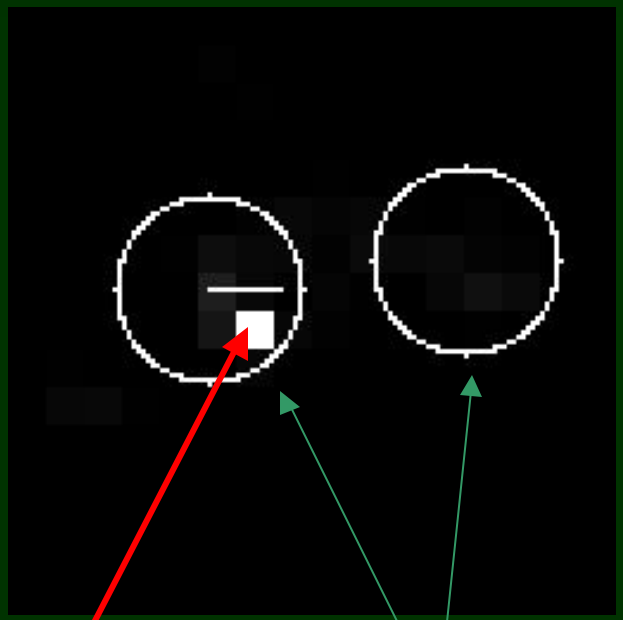




Locating Arm without Appearance Model

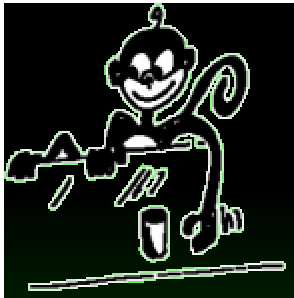


Optical flow

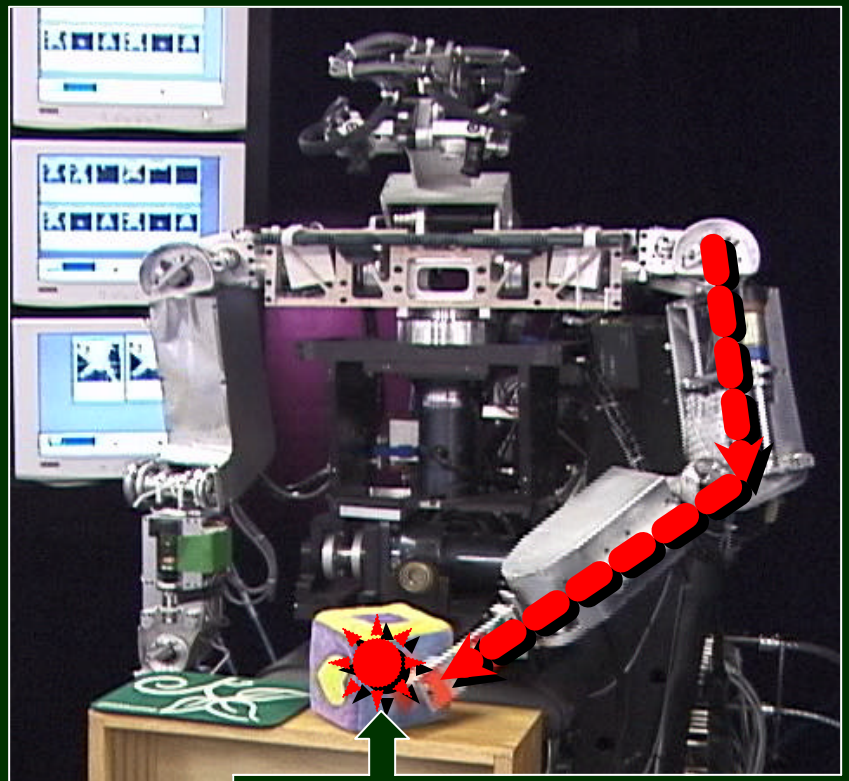


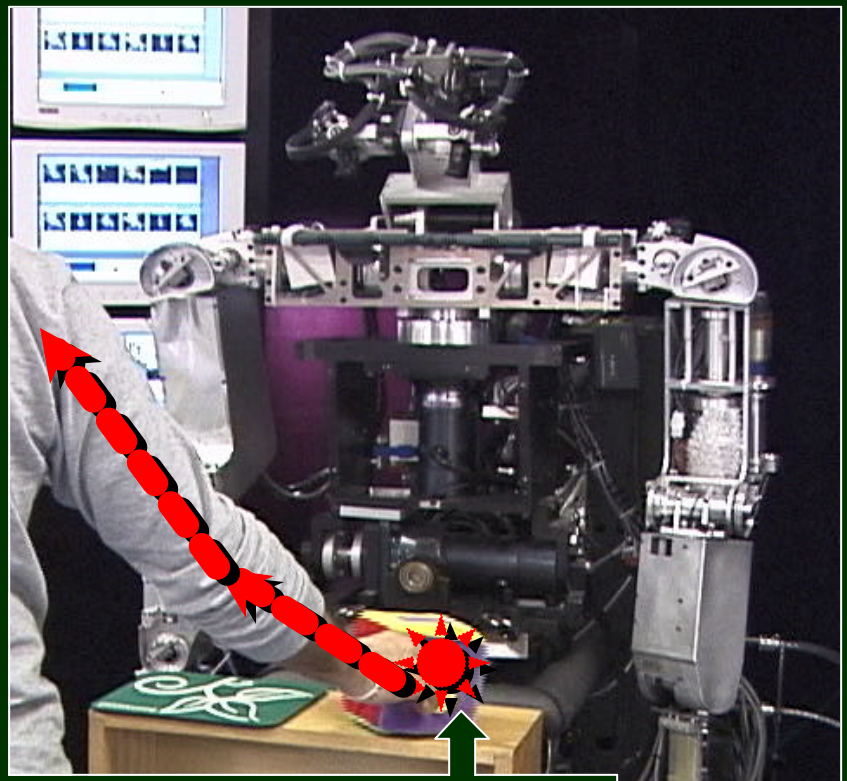
Maximum

Segmented regions



Tracing Cause and Effect





Object, goal connects robot and human action