

Better Vision through Poking

A close-up of a metallic robotic hand with a white sensor tip, pointing towards the right side of the slide.

Paul Fitzpatrick

Humanoid Robotics Group

MIT AI Lab

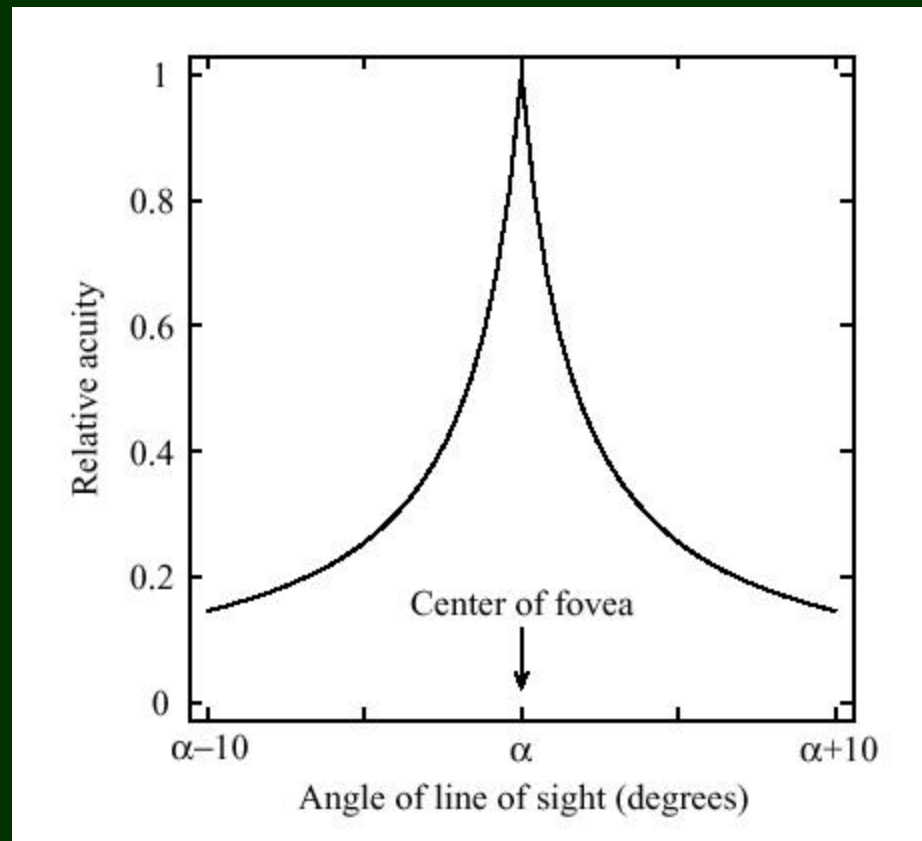
Strategies for Sensing

- Sensing capability is finite

Many choices in how to allocate it

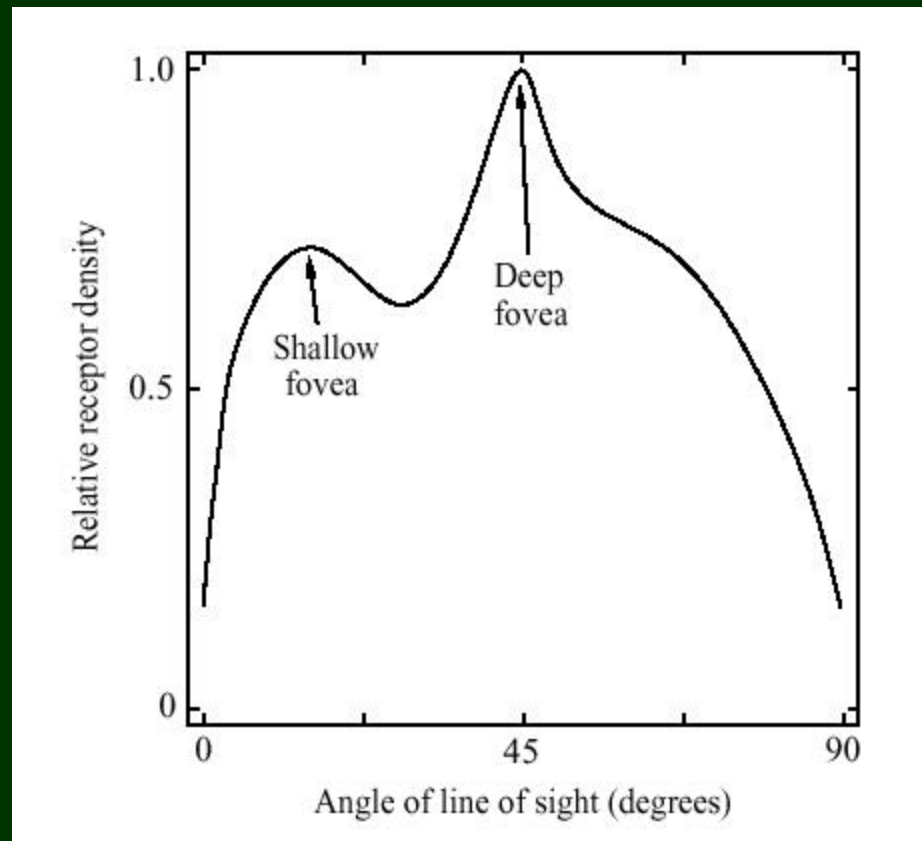
Smart choices simplify processing

Example 1: Human Foveation



(Adapted from V. Tucker '00)

Example 1: Falcon Foveation

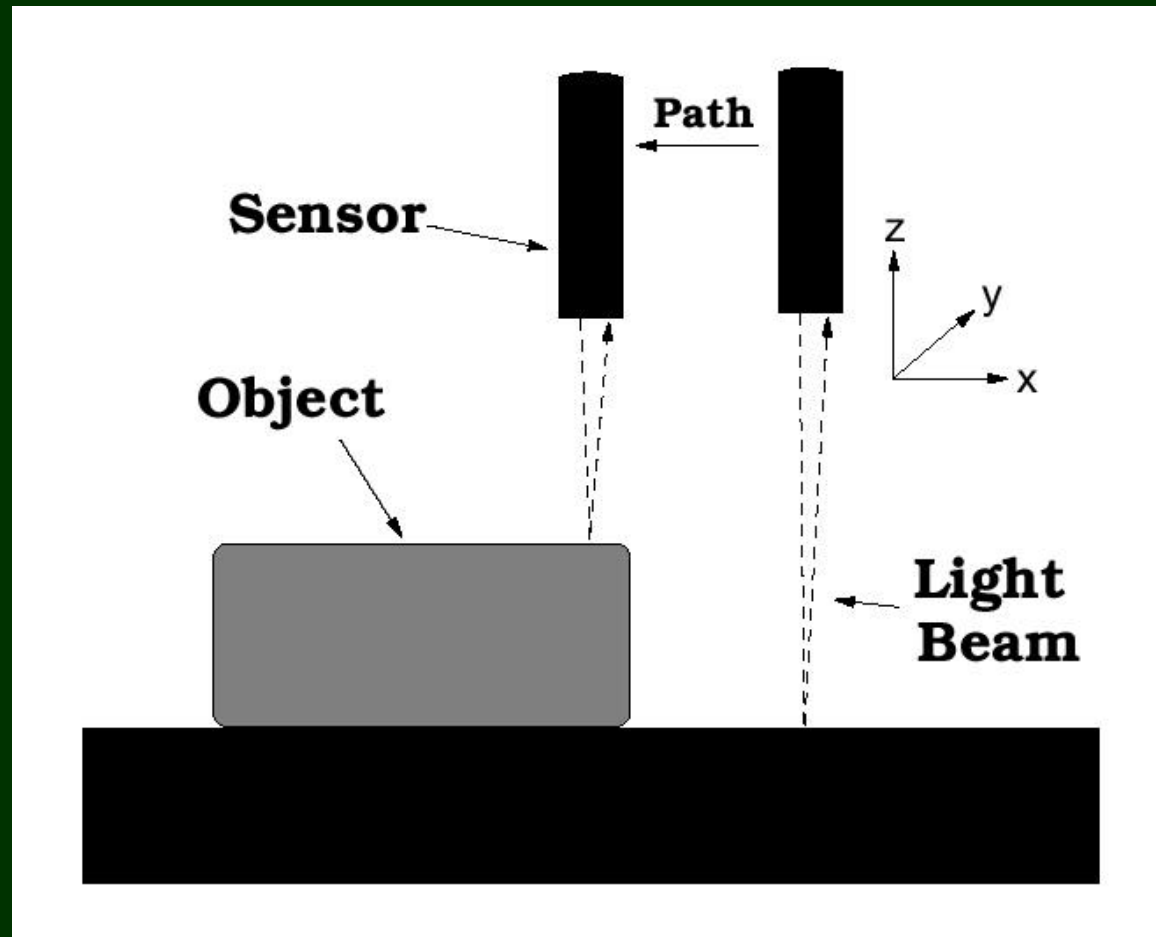


(Adapted from V. Tucker '00)

Example 1: Robot Foveation

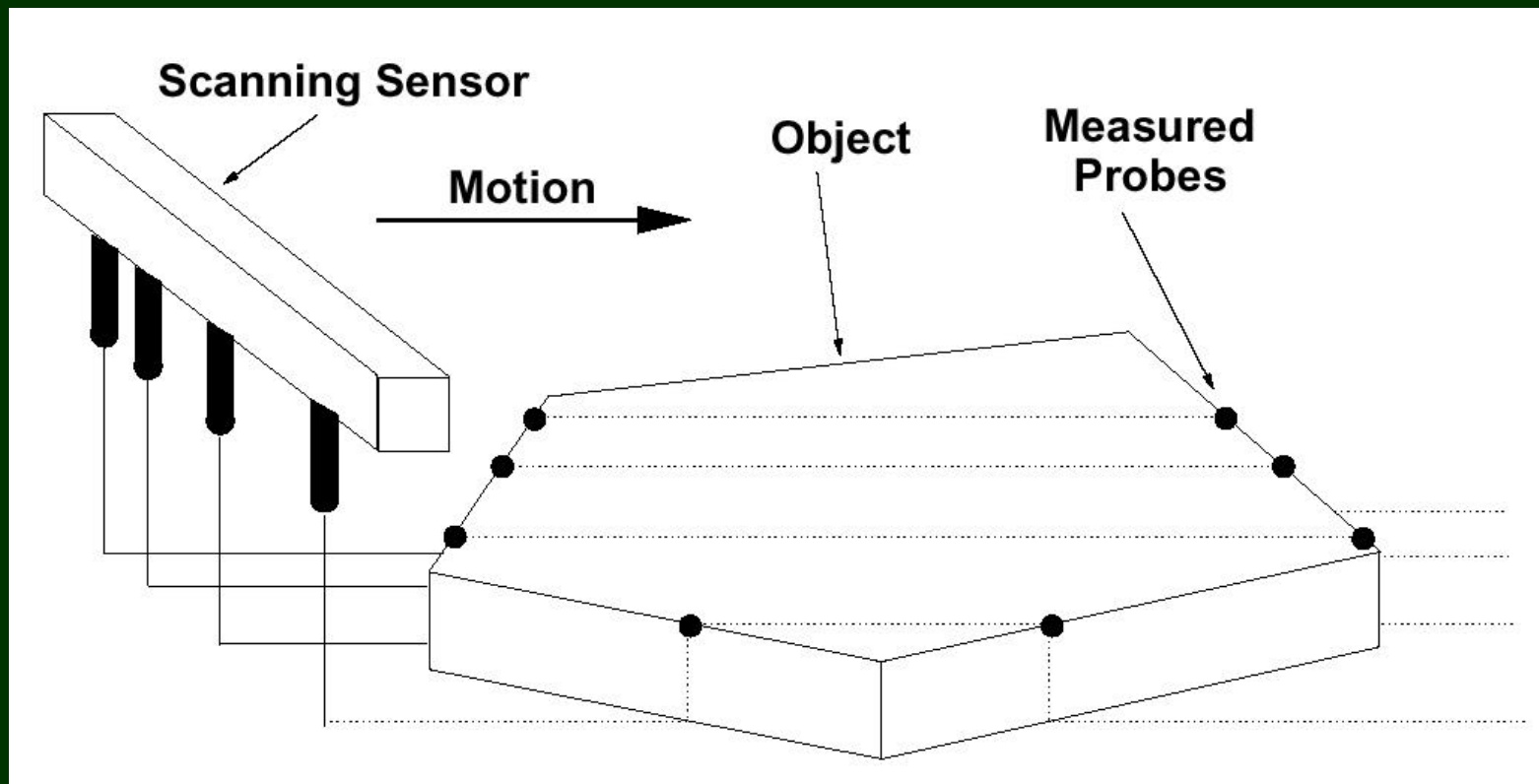


Example 2: Shape from Probing



(Figure from E. Paulos '99)

Example 2: Shape from Probing



(Figure from E. Paulos '99)

Image Segmentation

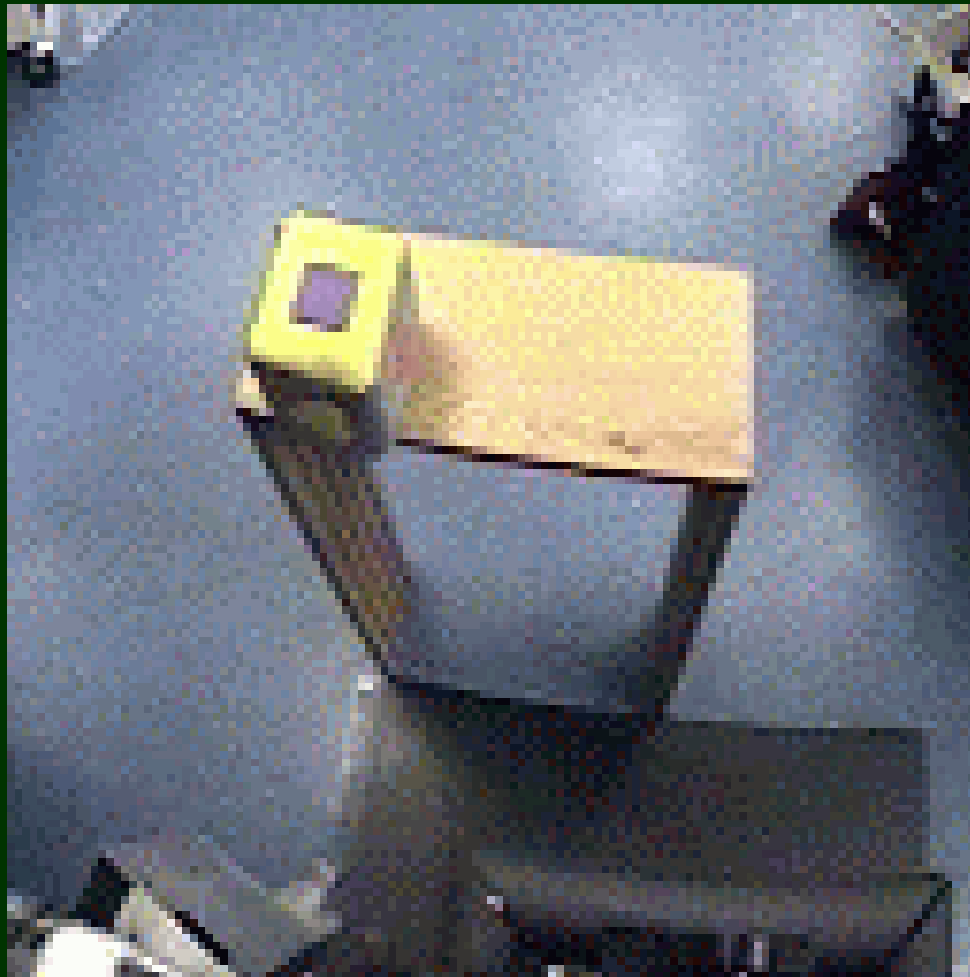
- Image segmentation is subtle, ambiguous
Physical poking is direct, to the point



Active Segmentation

- Unsure of an object's boundaries?
 - Poke it gently
 - Thump it savagely
 - Try to put your arm/hand/flipper beside it
 - Try to put your arm/hand/flipper behind it
 - Move your head for a different view
 - Get help

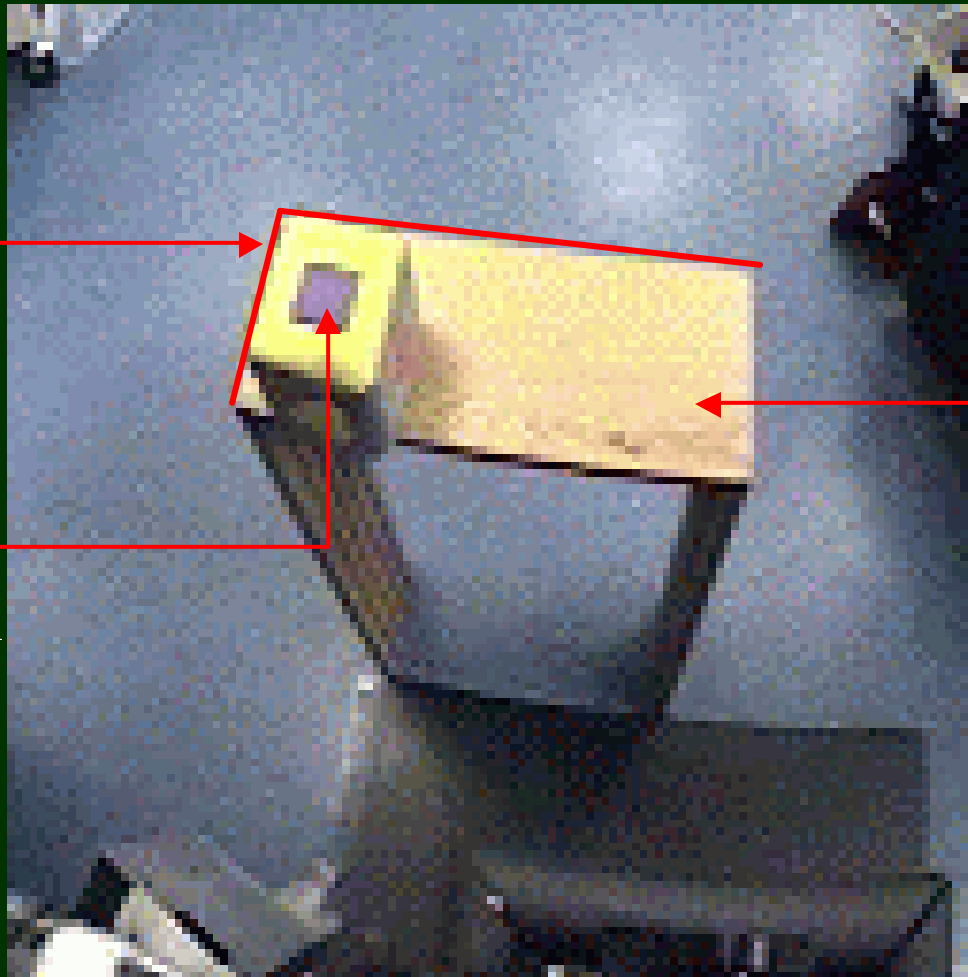
A Simple Scene?



A Simple Scene?

Edges of table
and cube
overlap

Cube has
misleading
surface pattern



Color of cube
and table are
poorly separated

Active Segmentation

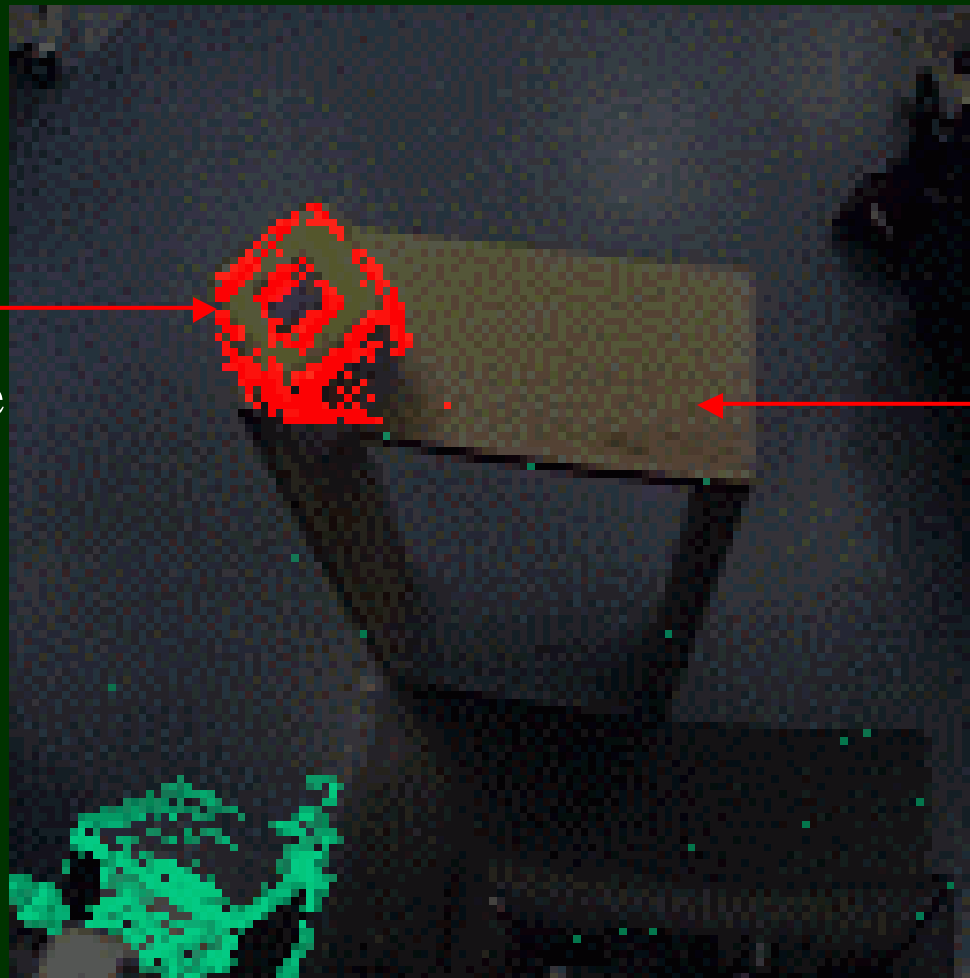


Active Segmentation



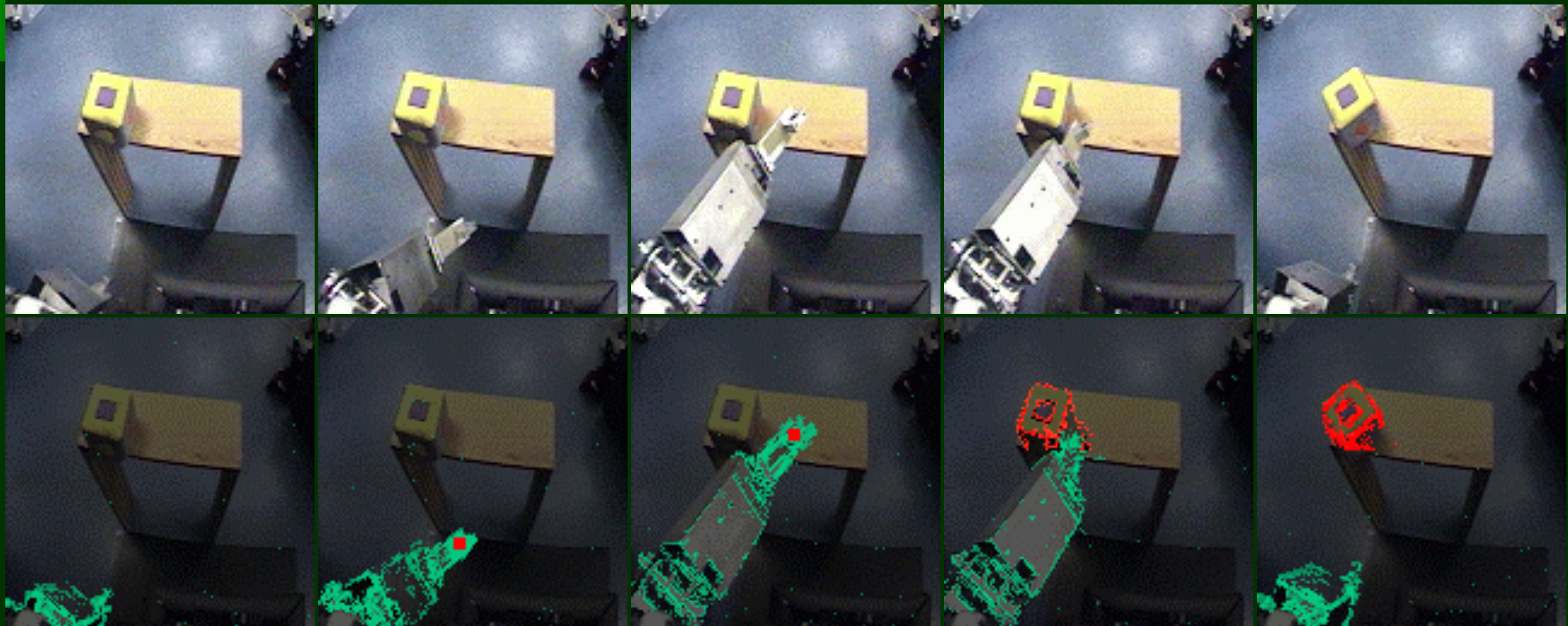
Result

No confusion
between cube
and own texture



No confusion
between cube
and table

Anatomy of a Poke



Begin

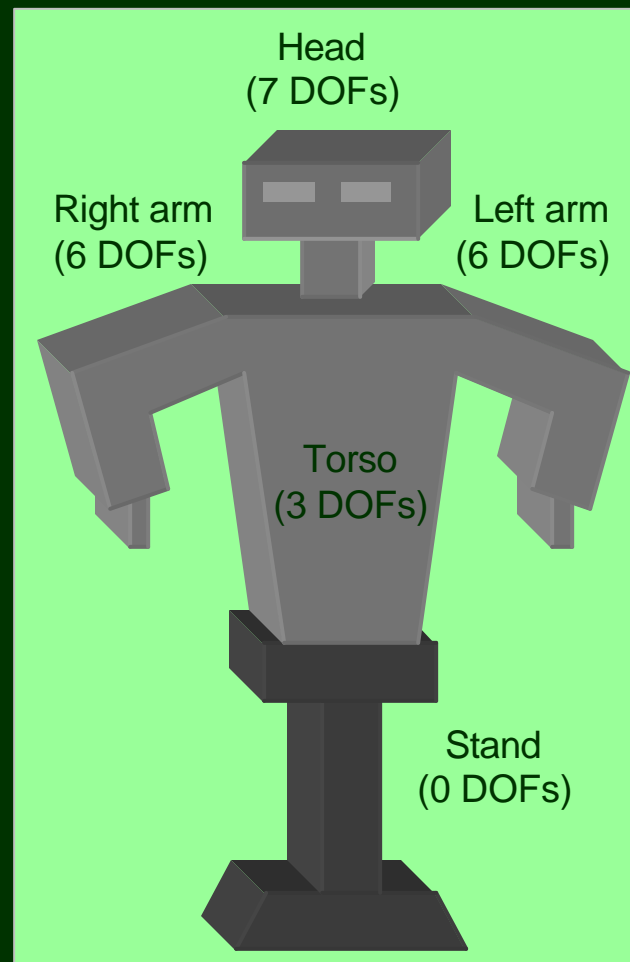
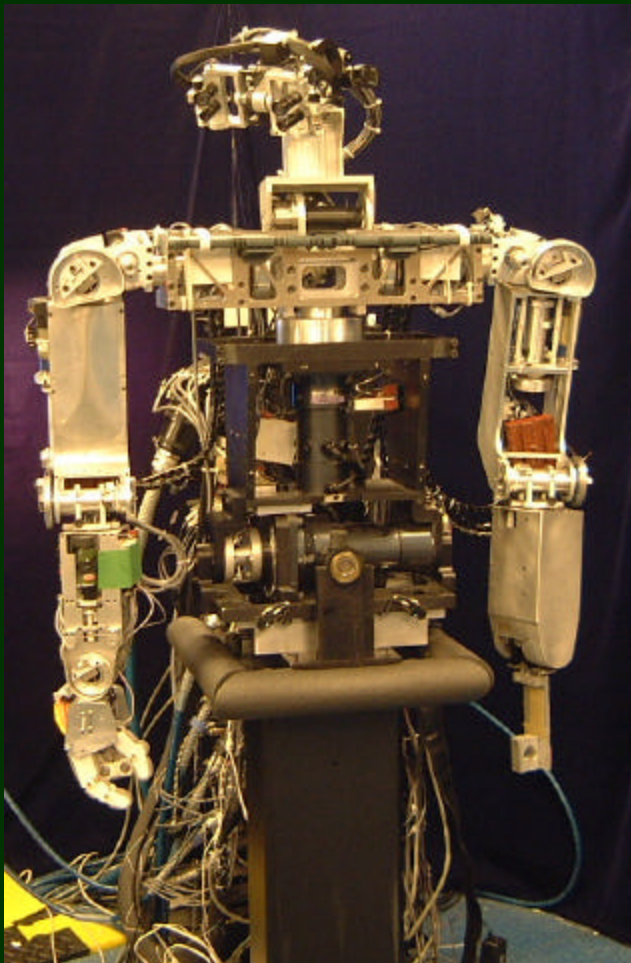
Find hand

Sweep

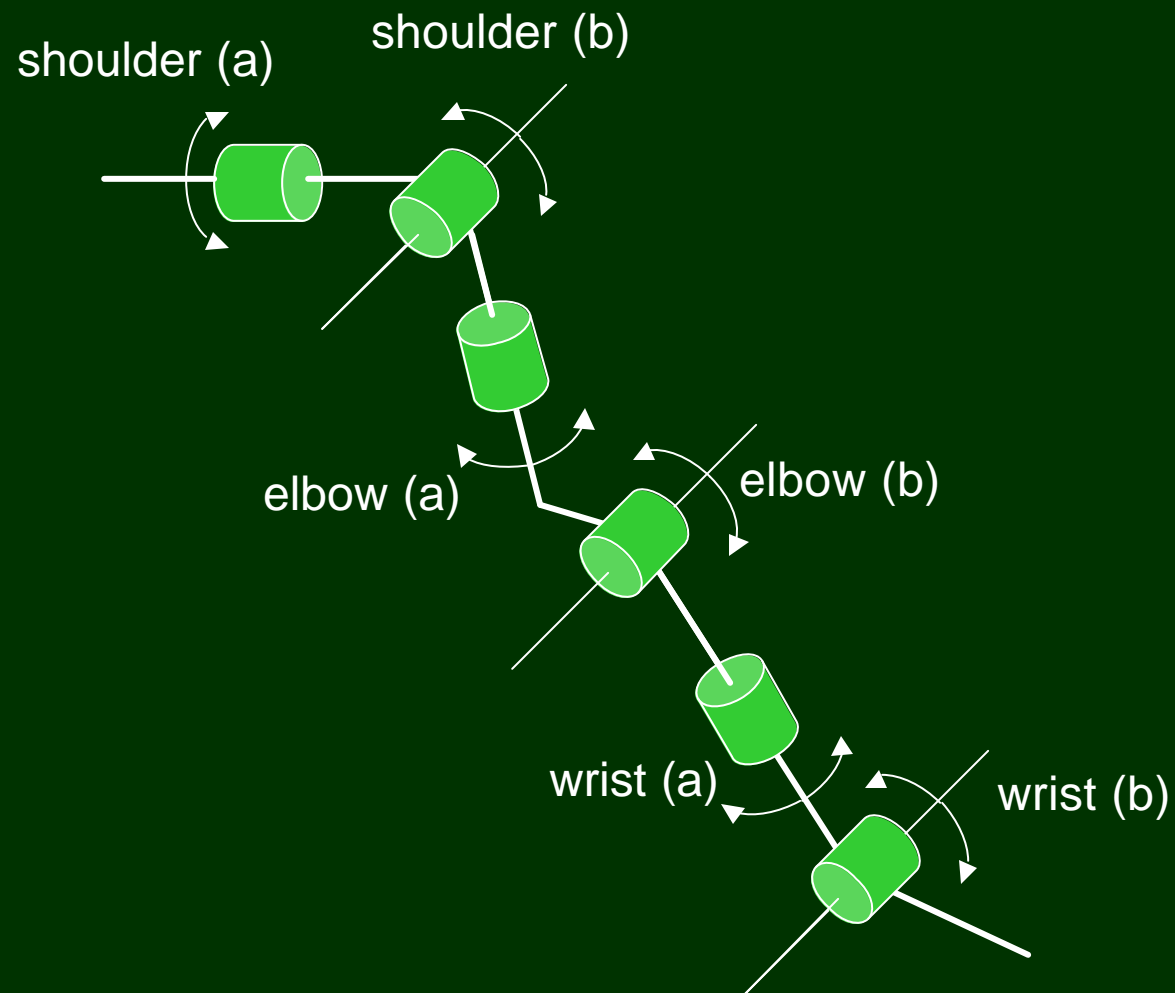
Contact!

Withdraw

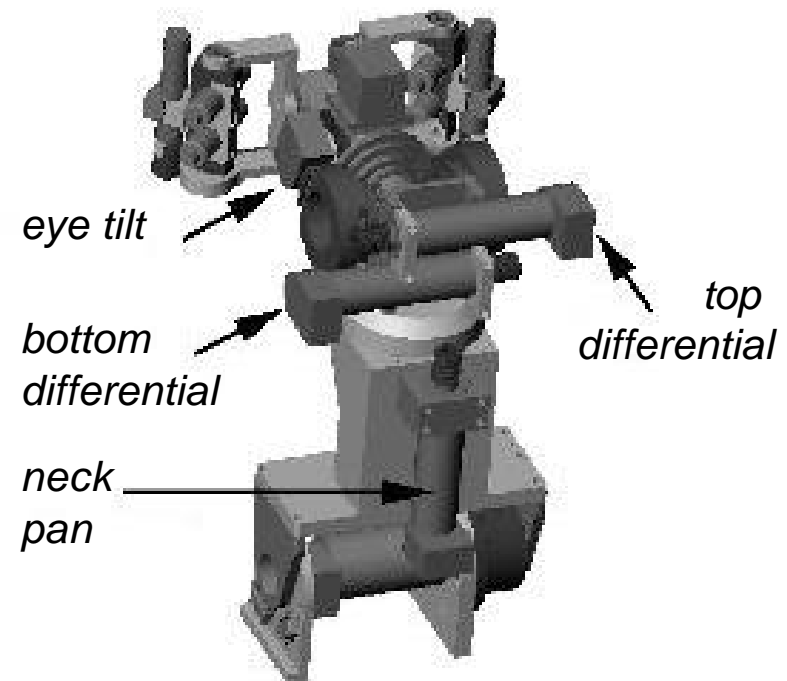
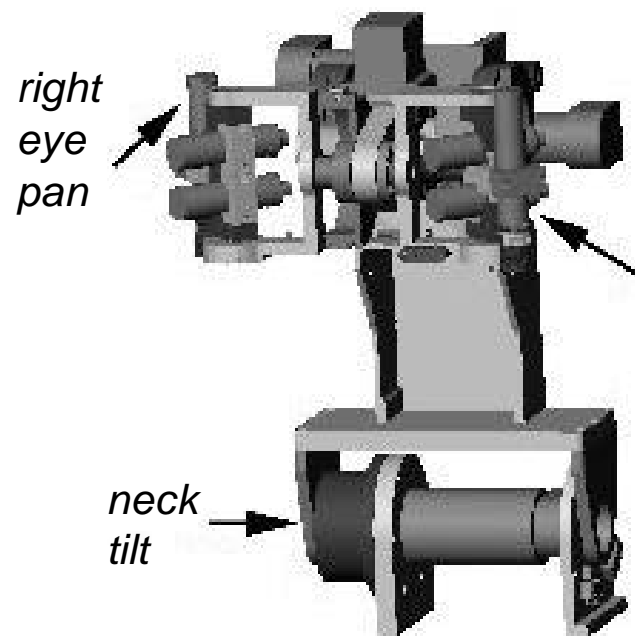
The Robot



The Arm



The Head

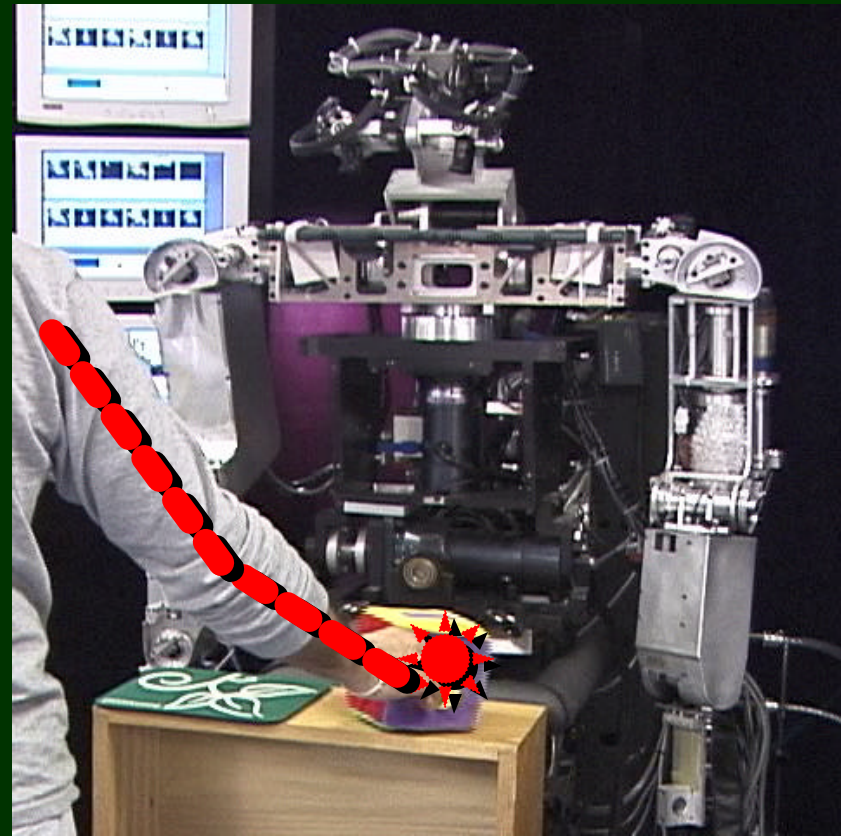
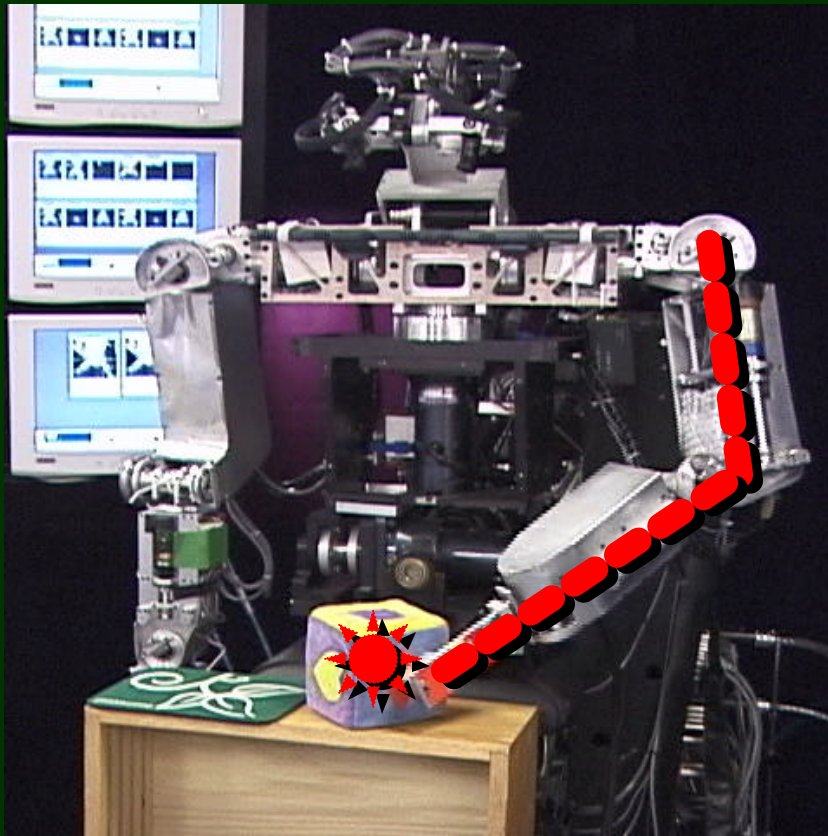


Tracing Cause and Effect

- Goal: to relate robot and human action without prior knowledge of visual appearance

- Determine appearance of own arm in motion
- Follow the causal chain outwards to determine the appearance of actions on objects
- Then follow the chain “up” a human’s arm when they move an object after the robot

Tracing Cause and Effect



Locating Arm without Appearance Model

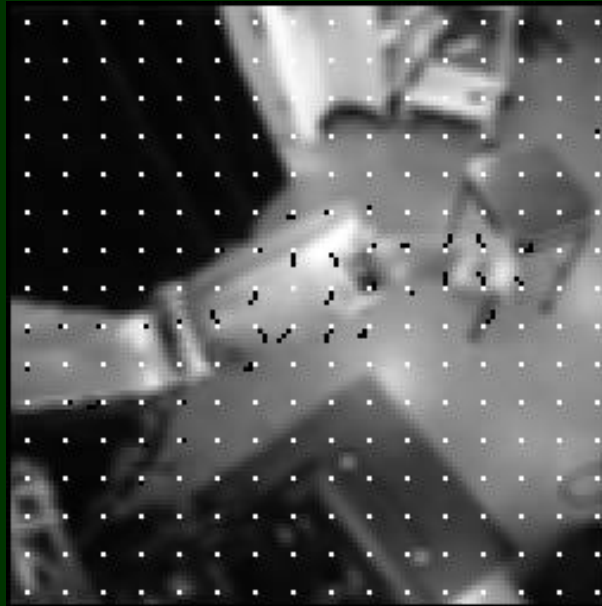
- Shake it

Correlate commanded motion with optic flow

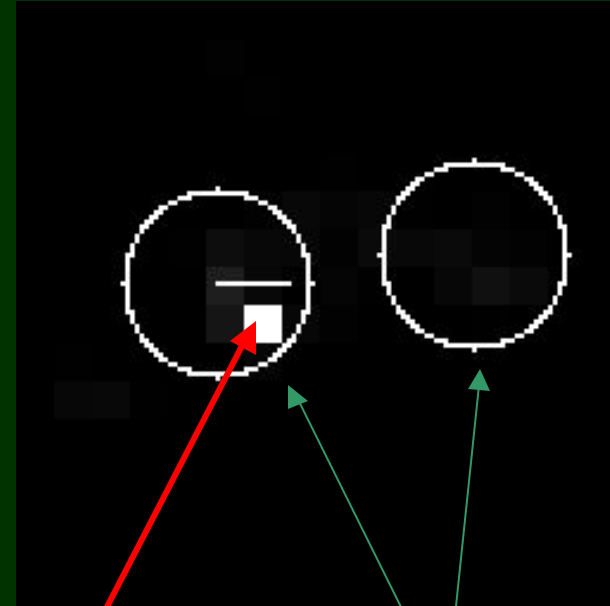
Ignore uncorrelated motion

(Giorgio Metta)

Locating Arm without Appearance Model



Optical flow

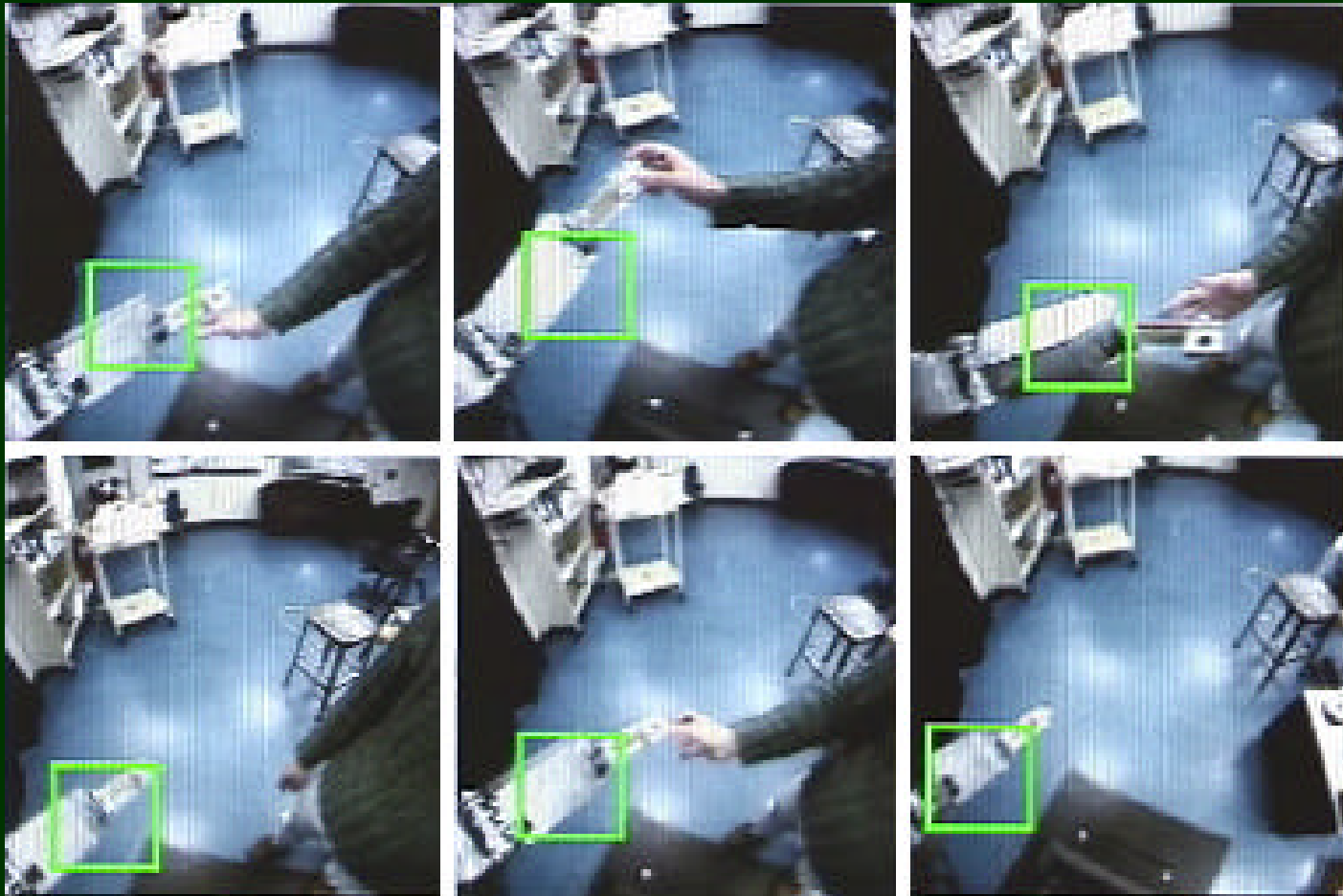


Maximum

Segmented regions

(Giorgio Metta)

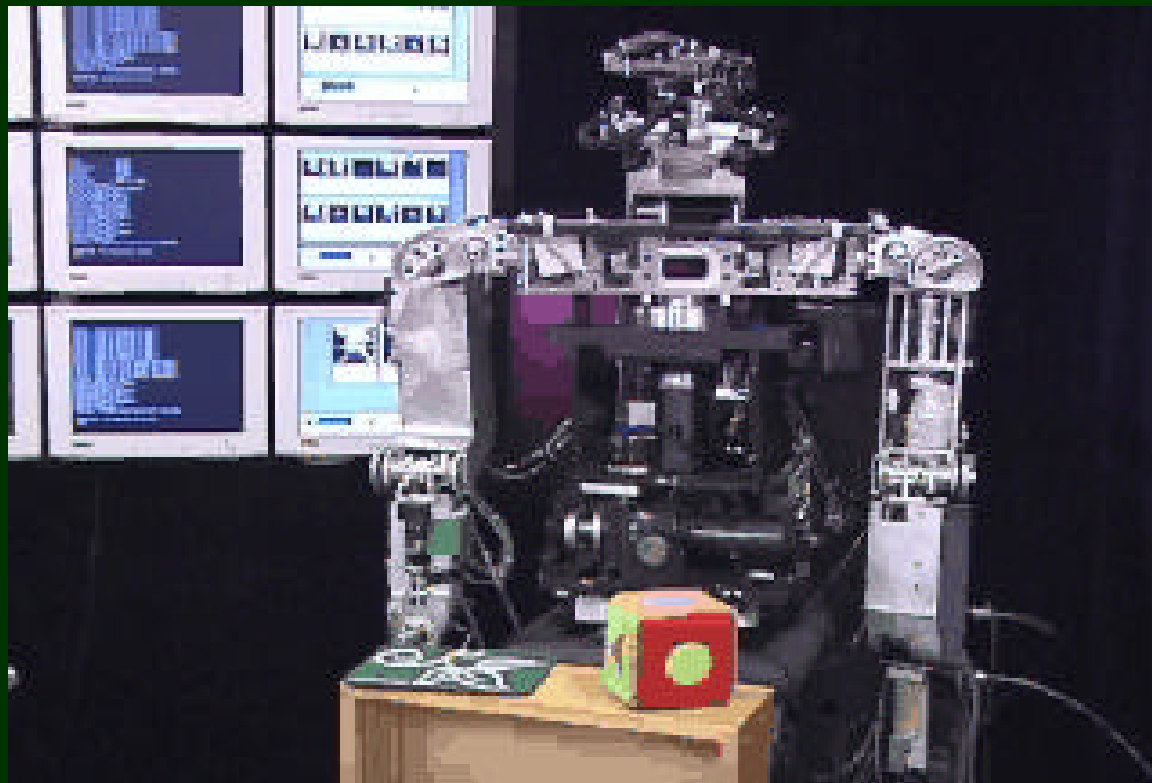
Training Visual Predictor



(Giorgio Metta)

Then...

Start poking things!



Things To Do



- Segment completely visually ambiguous scenes

Characterize non-rigid objects –
should lead to pragmatic, realistic
object model