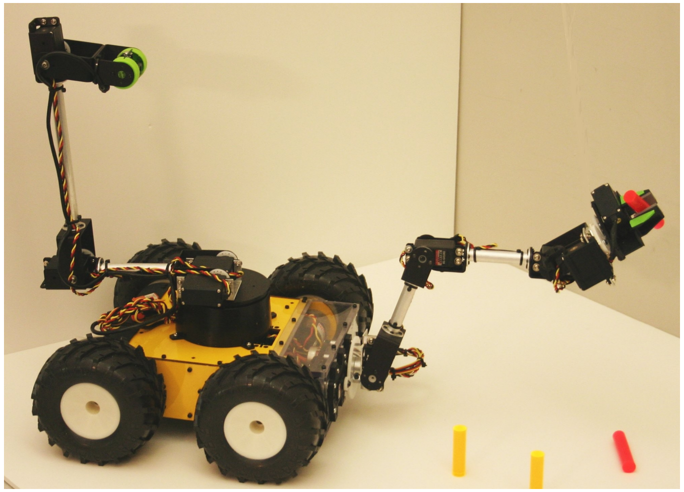
Robot Vision in Tekkotsu

David S. Touretzky
Computer Science Department
Carnegie Mellon

March, 2008

Tekkotsu is an application development framework for mobile robots. Principal architect: Ethan Tira-Thompson at CMU.



Funded in part by National Science Foundation award DUE-0717705.

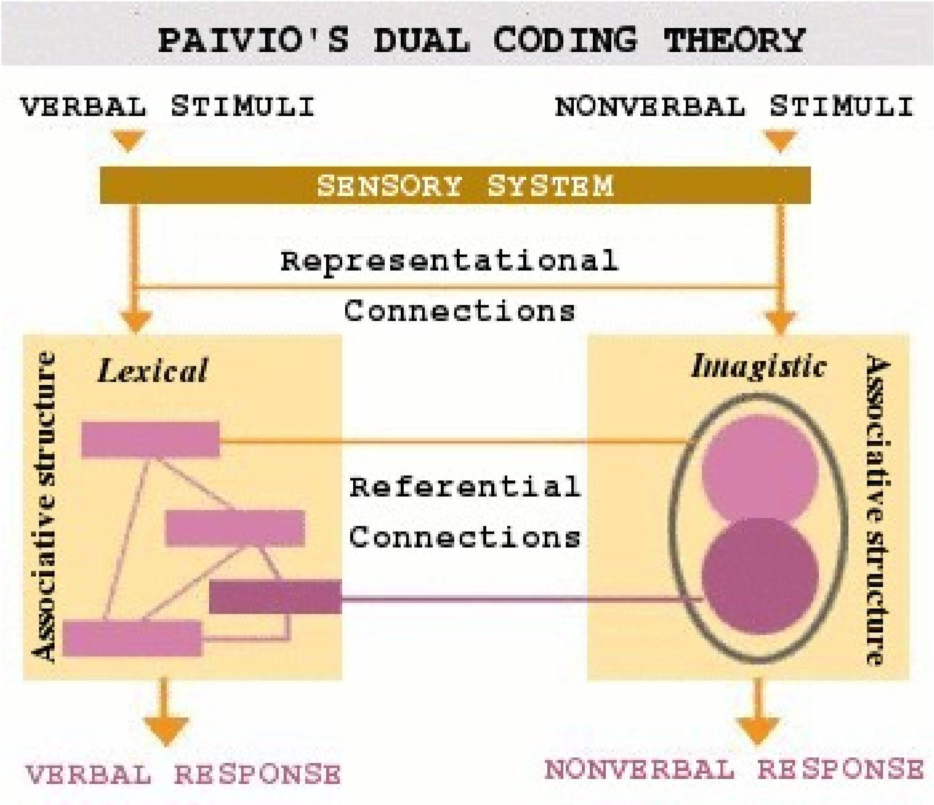
Dual-Coding Representations

- Paivio's “dual-coding theory”:

  People use both iconic (picture) and lexical (symbolic) mental representations.

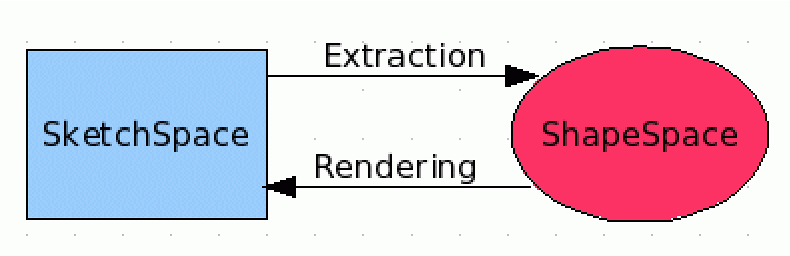
  They can convert between them when necessary...

  But at a cost of increased processing time and error rate.



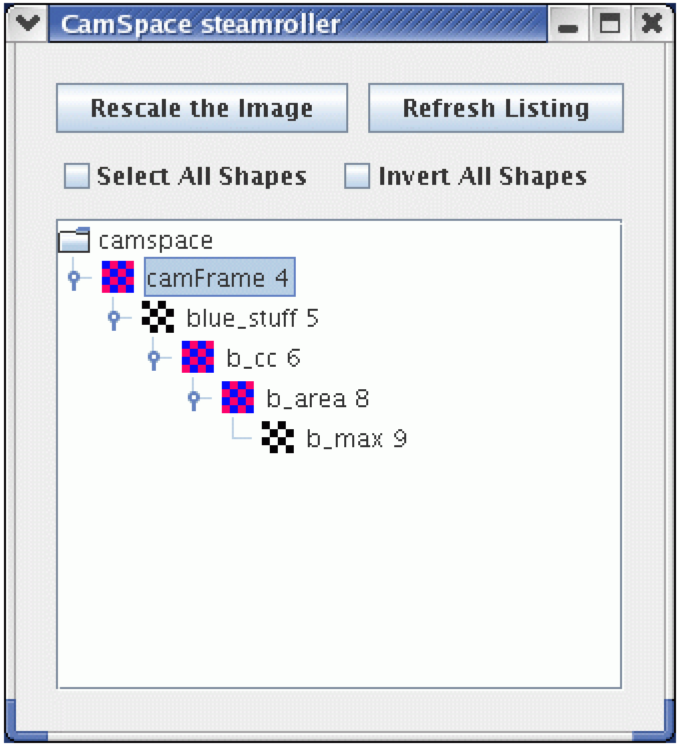
Dual-Coding in Tekkotsu

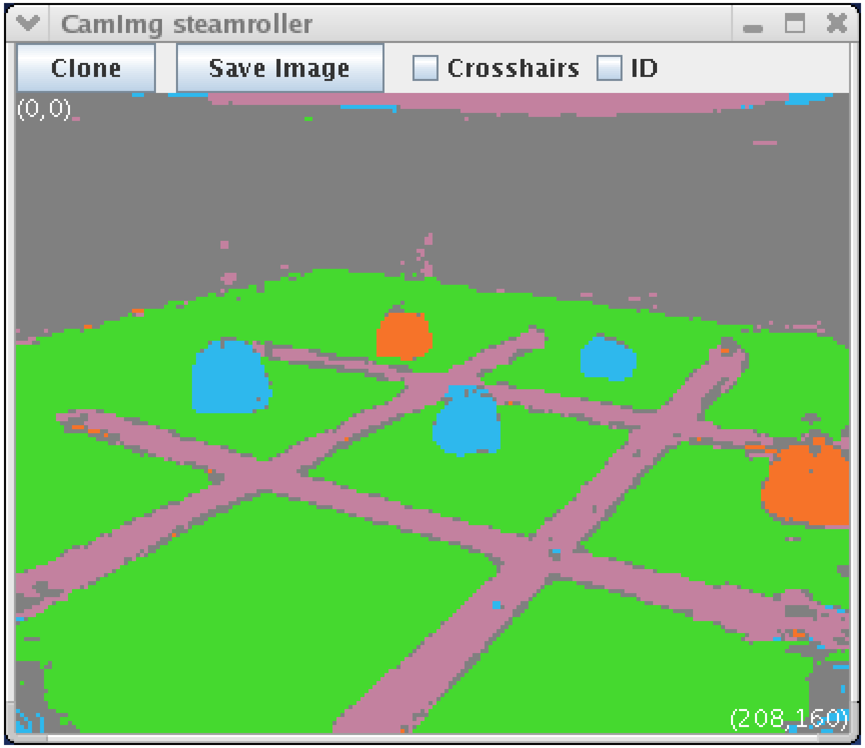
- Tekkotsu implements a version of Paivio's dual-coding idea:



- “Sketches” are iconic representations: pixel arrays.
- “Shapes” are lexical representations: algebraic descriptions of geometric objects.
- “Visual routines” (Ullman 1984) operate on both sketches and shapes.

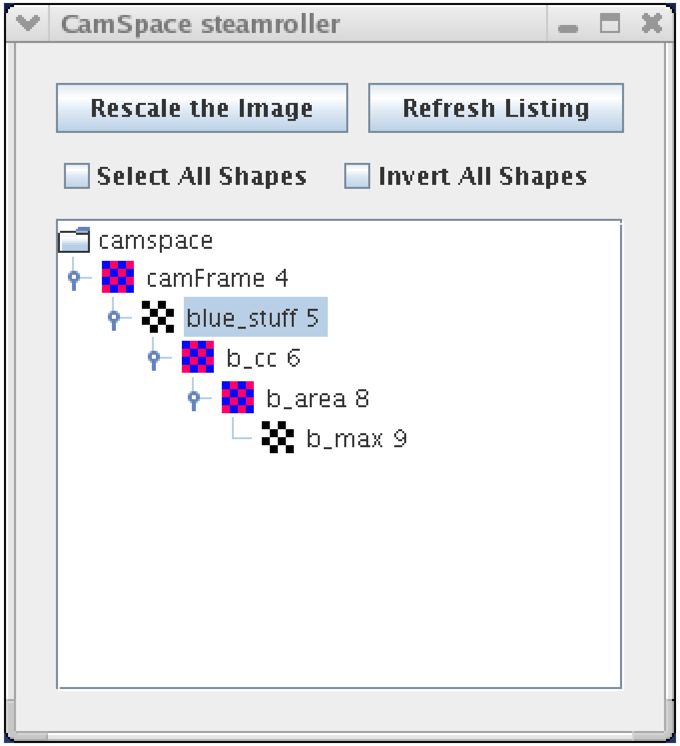
Problem: “Find the orange blob closest to the largest blue blob.”

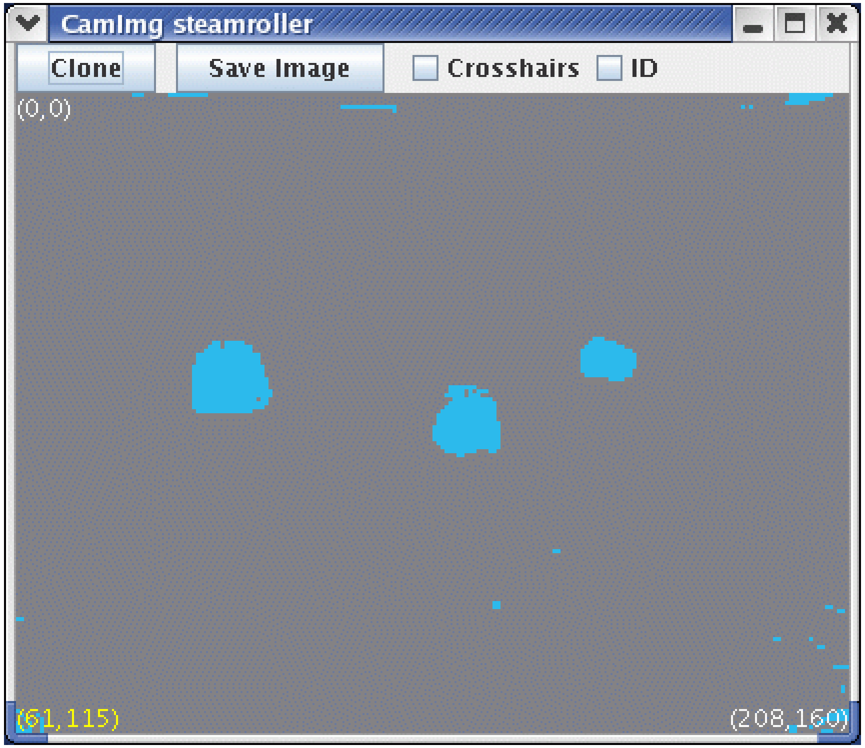




Color image segmentation done automatically by CMVision.

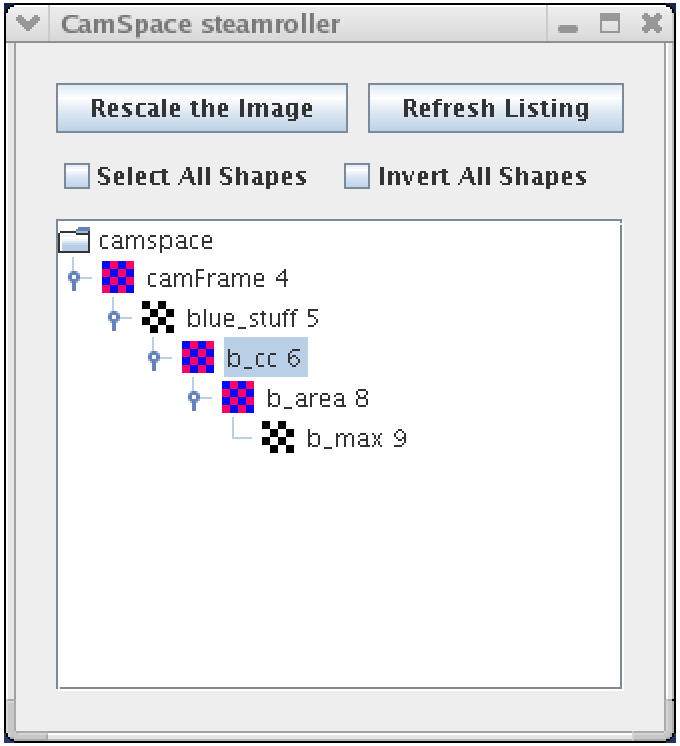
visops::colormask

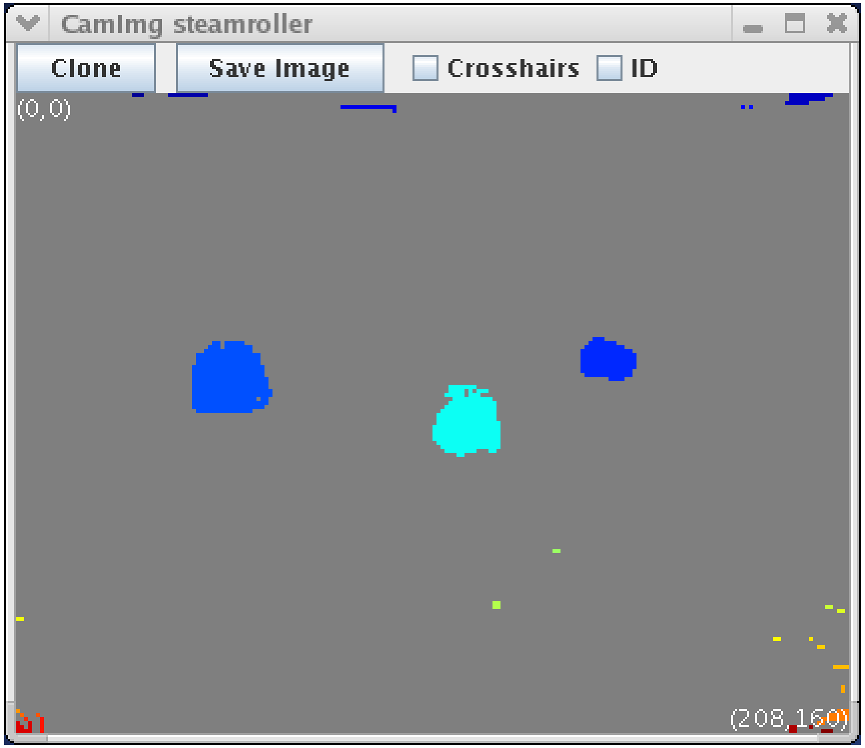




Step 1: find the blue pixels.

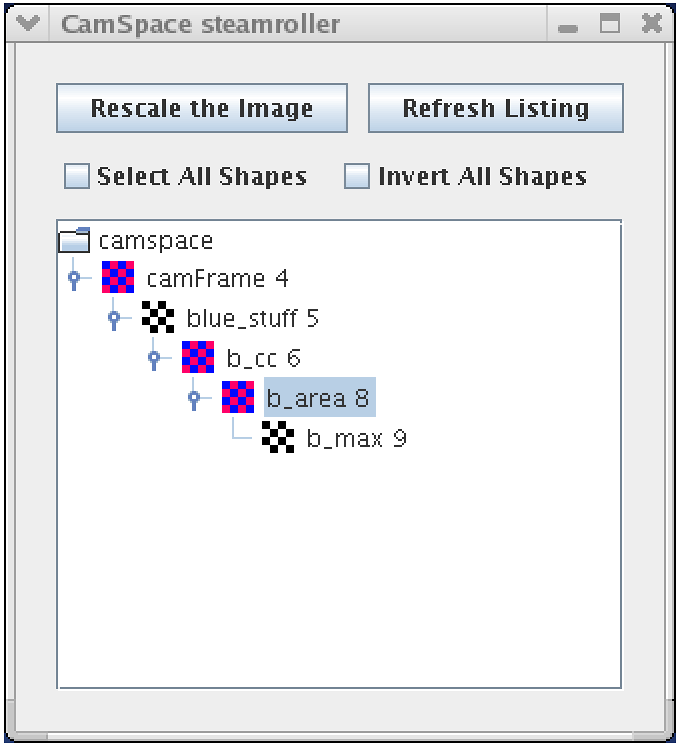
visops::labelcc

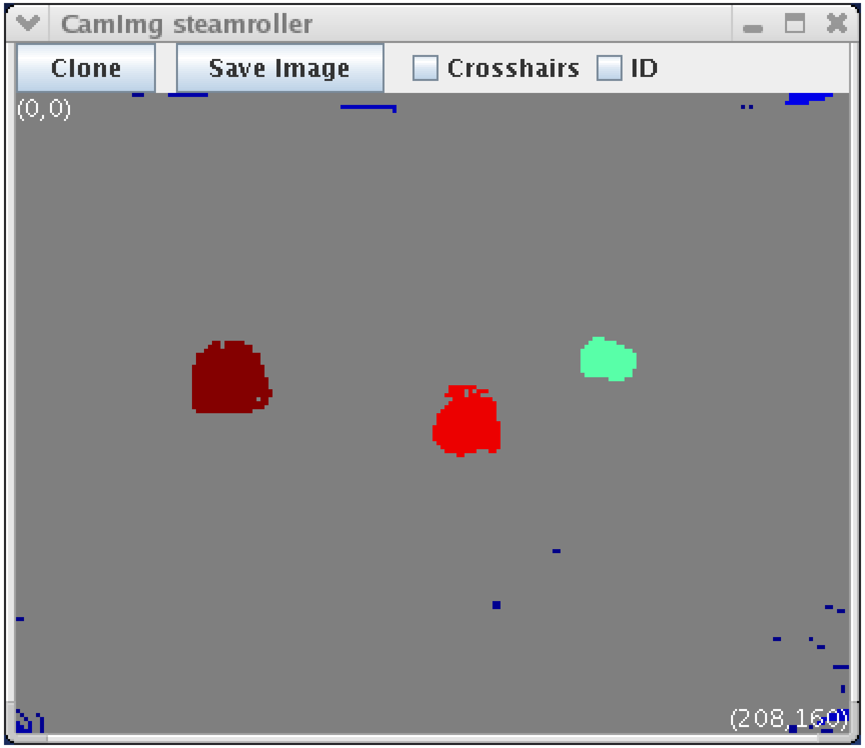




Step 2: label connected components.
Start from 1 in upper left; max label in lower right.

visops::areacc

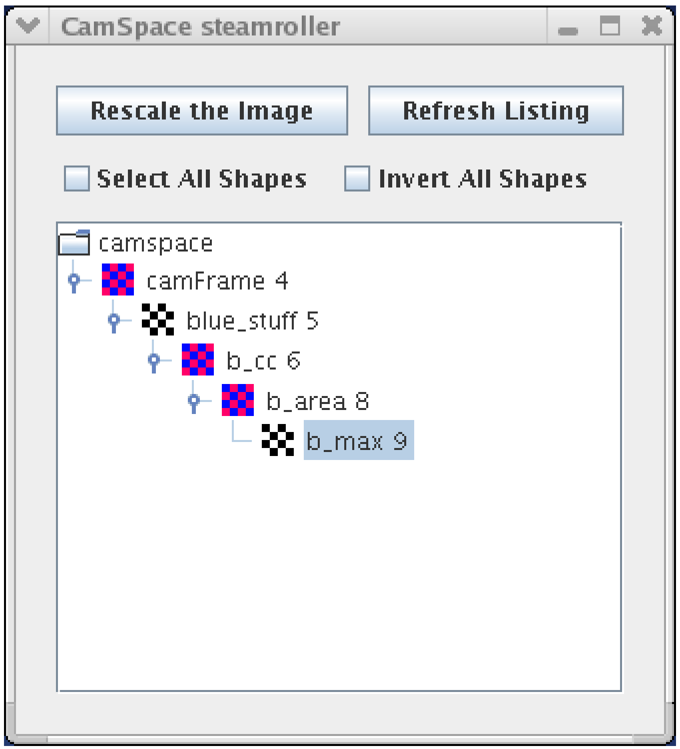


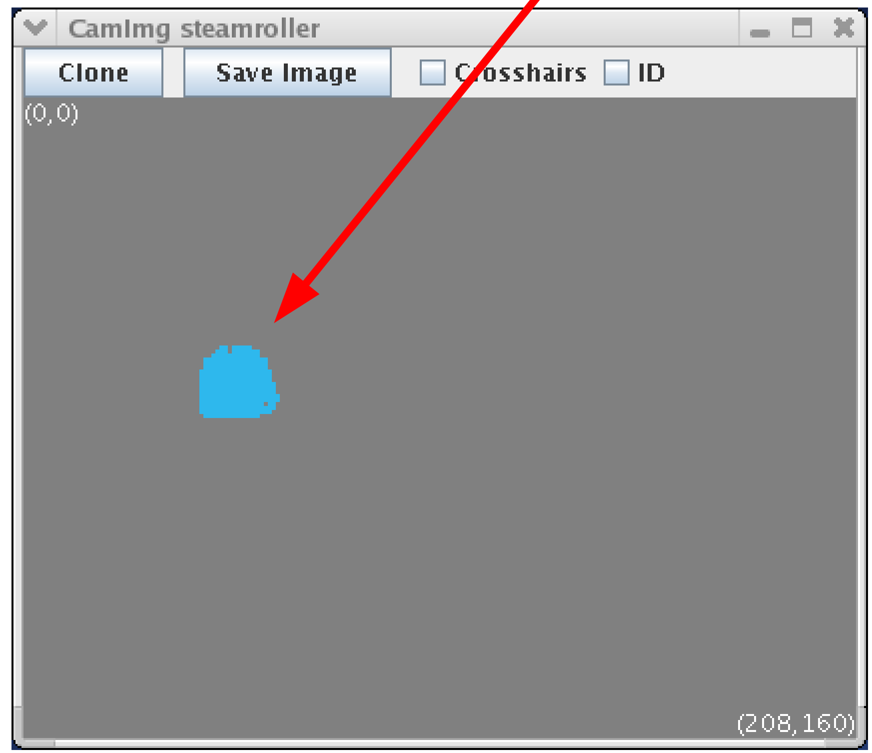


Step 3: relabel each connected component by its area.

b_area == b_area->max()

Largest blue blob

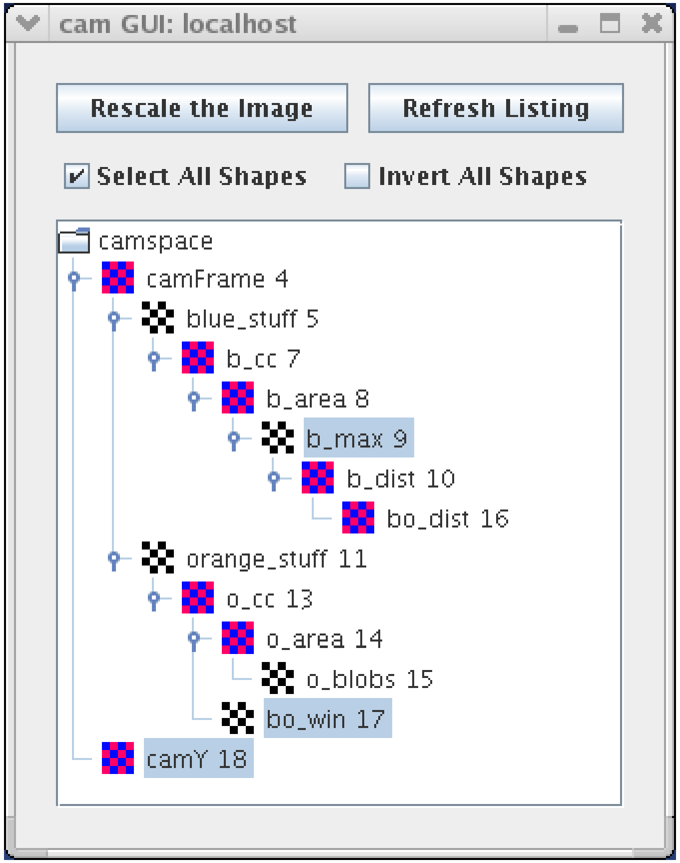


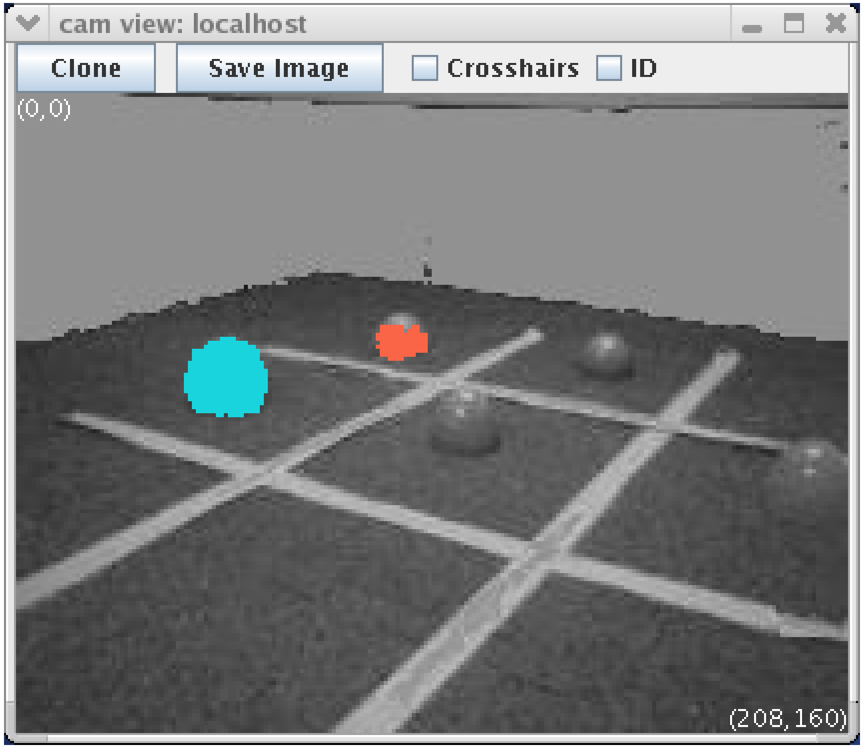


Step 4: mark all pixels bearing the largest area label.

Orange Blob Closest to the Largest Blue Blob

```
NEW_SKETCH(bo_win, bool, o_cc == min_label);
```



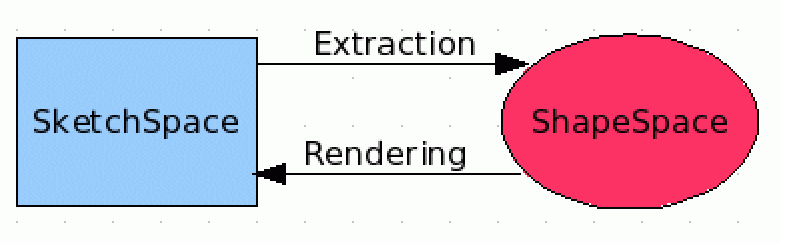


Steps 5-9 complete the computation.

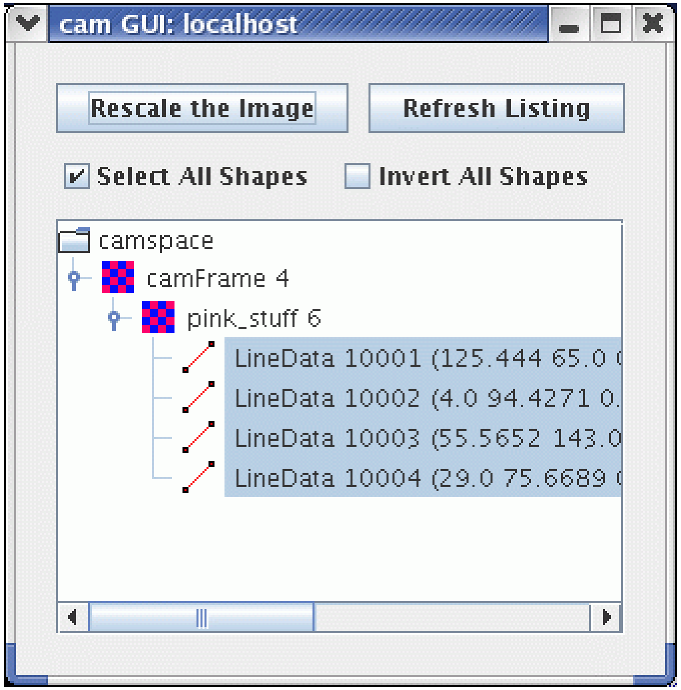
Shapes

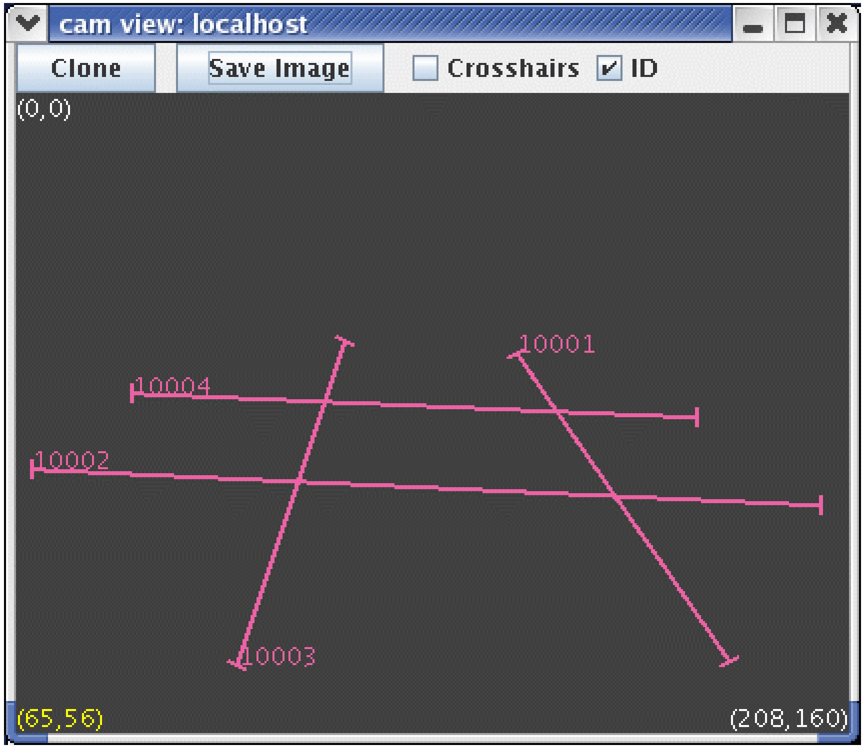
- Shapes are symbolic descriptions of geometric objects:
  - Point, Line, Ellipse
  - Polygon, Blob
  - Agent
  - Sphere, Brick, Pyramid

- Extraction and rendering algorithms map between sketch and shape representations.



Extracted Line Shapes



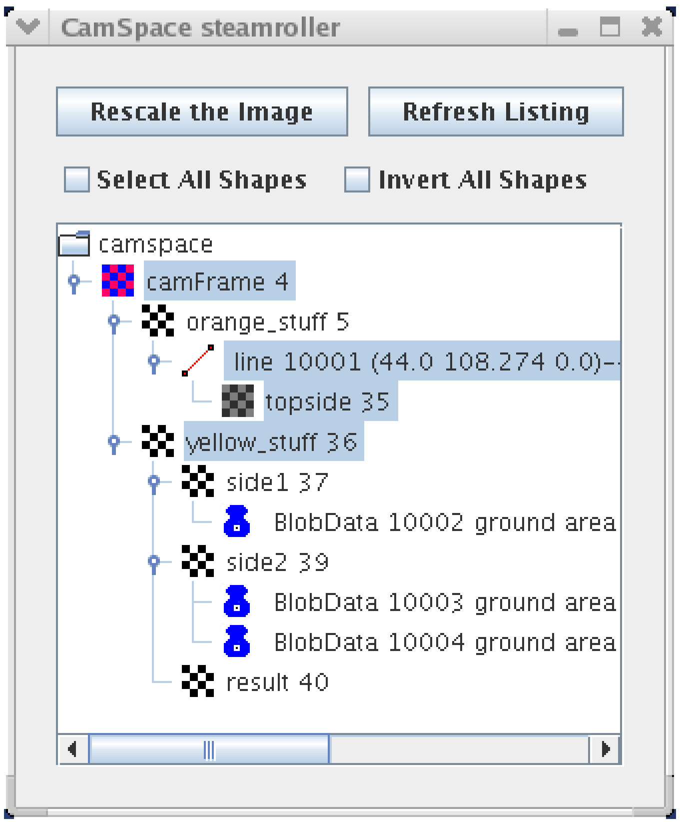


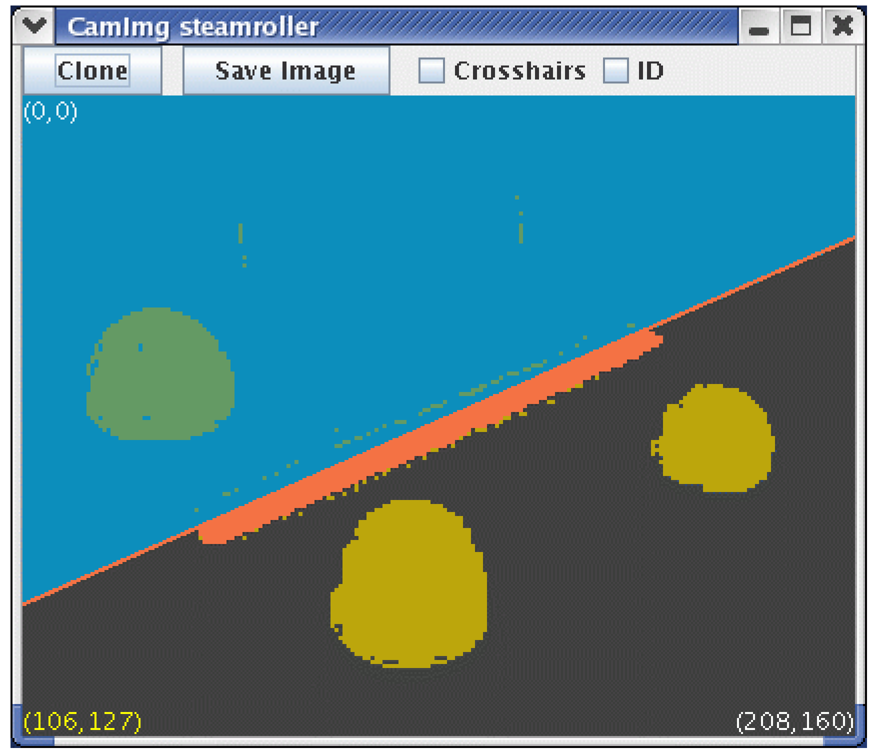
Mixing Sketches and Shapes

- The strength of the dual-coding approach comes from mixing sketch and shape operations.
- Example: which side of the orange line has more yellow blobs?

- If all we have is a line <u>segment</u>, people can still interpret it as an infinite “barrier”.
- How do we make the robot do this?

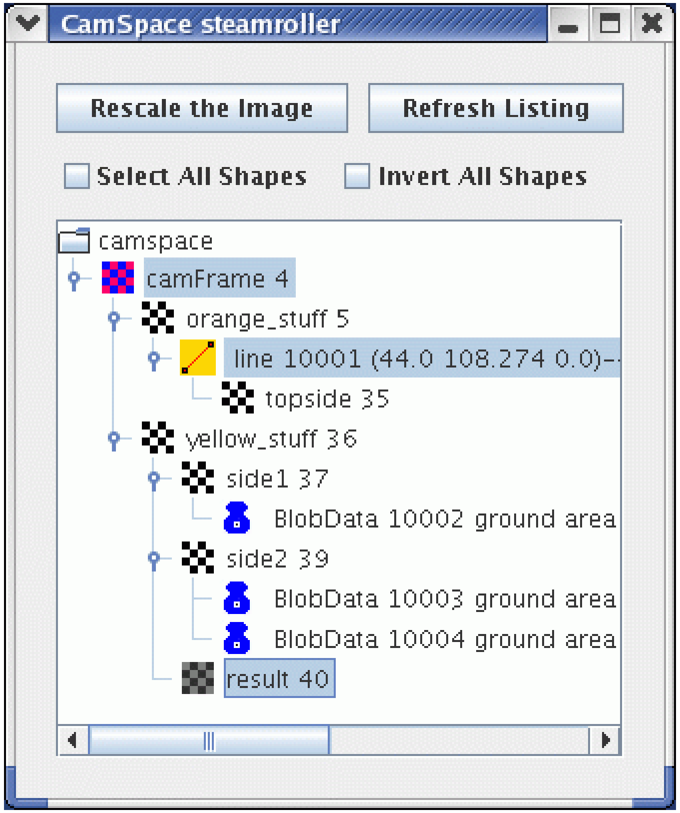
Lines As Barriers

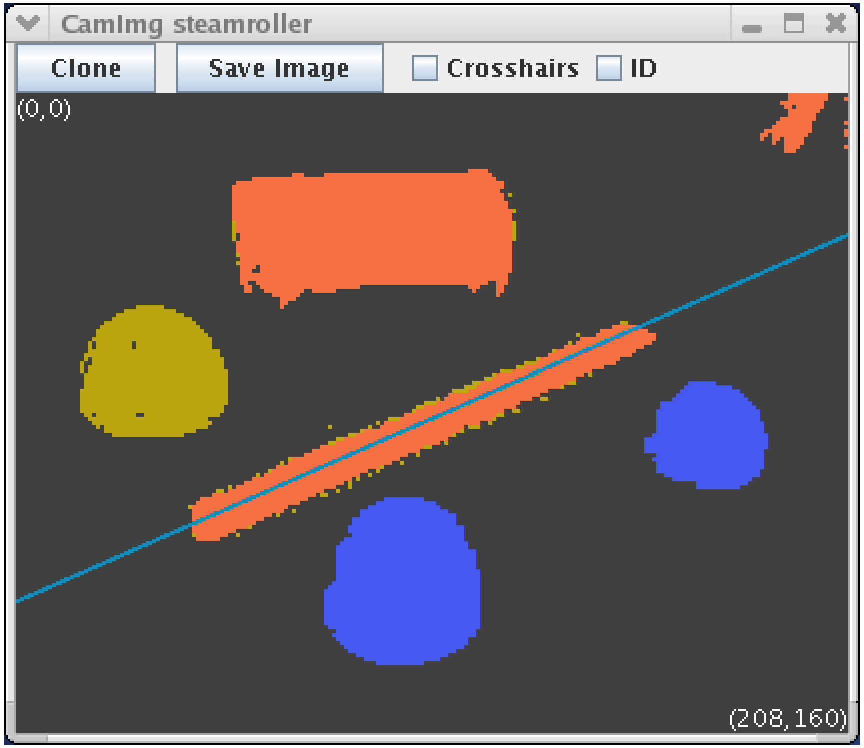




Subtle point: bool overrides uchar in the SketchGUI, so selecting yellow_stuff allows the top yellow blob to display even though the inverted (orange) *topside* is covering its appearance in *camFrame*. (Competing bools are averaged.)

Lines As Barriers

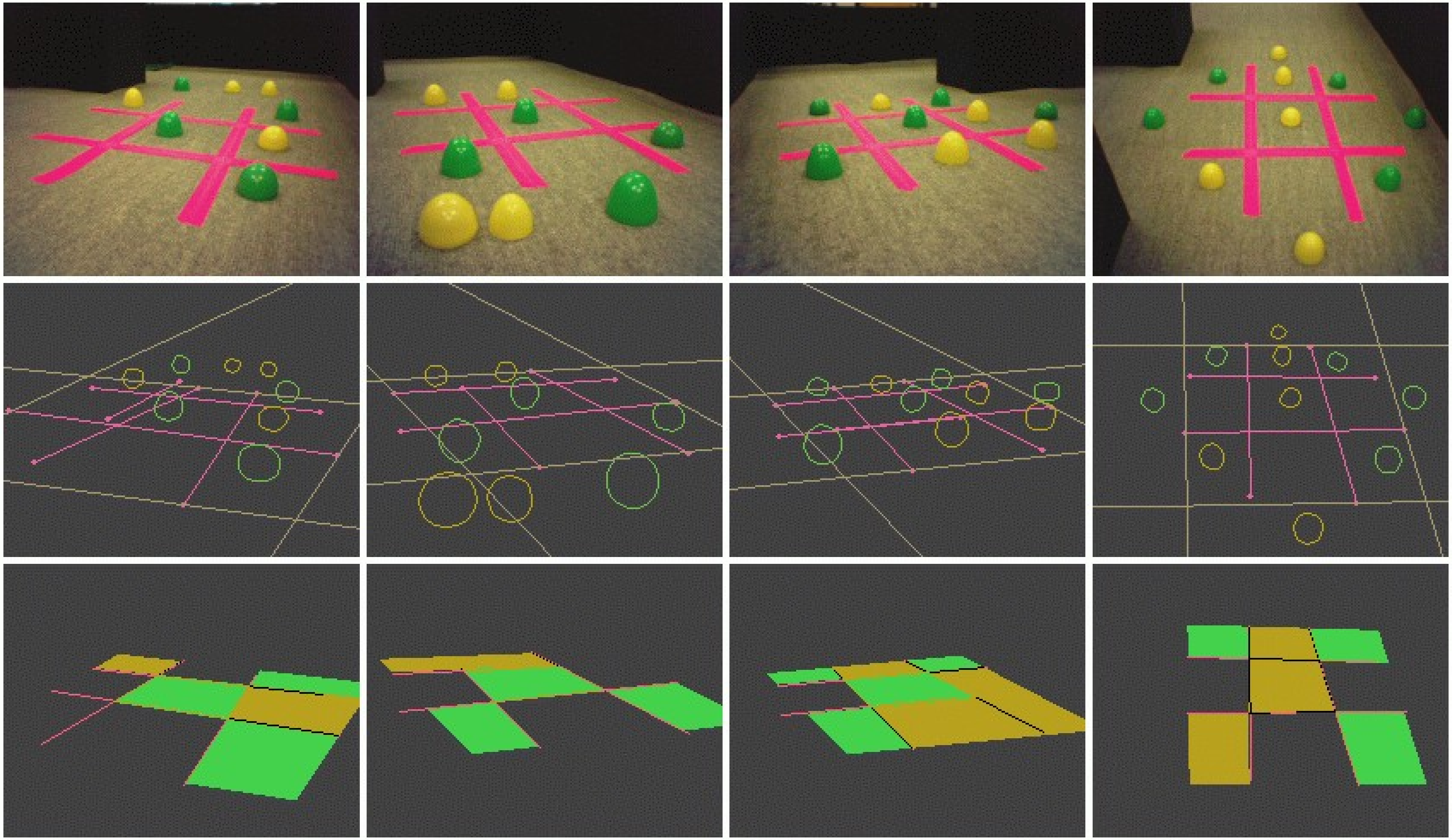




Parsing Tic-Tac-Toe Boards

Must infer board boundaries to exclude off-board game pieces.

Line extractor is programmed to handle occlusions.

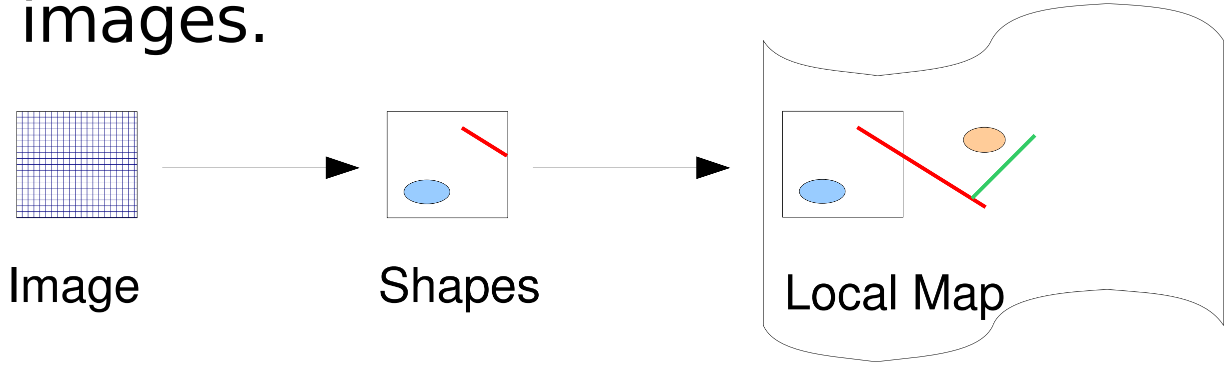


Seeing A Bigger Picture

- How can we assemble an accurate view of the robot's surroundings from a series of narrow camera frames?

- First, convert each image to symbolic form: shapes.

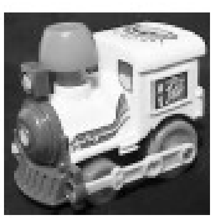
- Then, match the shapes in one image against the shapes in previous images.



- Construct a “local map” by matching up a series of camera images. This requires first translating from camera coordinates to groundplane coordinates to correct for perspective effects.

SIFT (Lowe, 2004)

- Scale-Invariant Feature Transform
- Can recognize objects independent of scale, translation, rotation, or occlusion.
- Can segment cluttered scenes.
- Slow training, but fast recognition.



How Does SIFT Work?

- Generate large numbers of features that densely cover each training object at various scales and orientations.

- A 500 x 500 pixel image may generate 2000 stable features.
- Store these features in a library.
- For recognition, find clusters of features present in the image that agree on the object position, orientation, and scale.

Keypoint Descriptors

A SIFT feature (or “keypoint”) is a complex structure containing a grid of histograms of local image gradients.

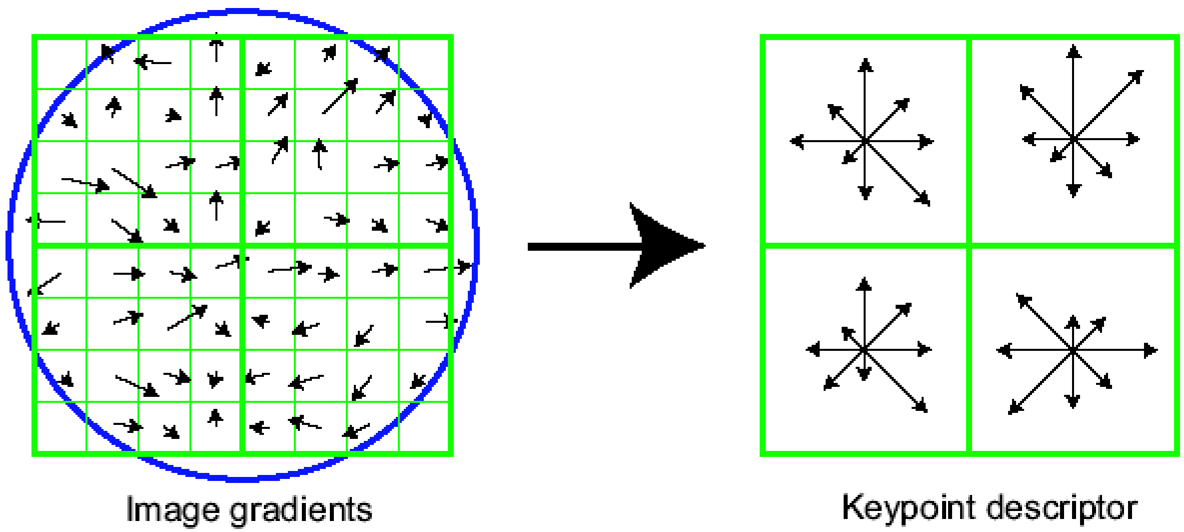


Figure 7: A keypoint descriptor is created by first computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location, as shown on the left. These are weighted by a Gaussian window, indicated by the overlaid circle. These samples are then accumulated into orientation histograms summarizing the contents over 4x4 subregions, as shown on the right, with the length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region. This figure shows a 2x2 descriptor array computed from an 8x8 set of samples, whereas the experiments in this paper use 4x4 descriptors computed from a 16x16 sample array.

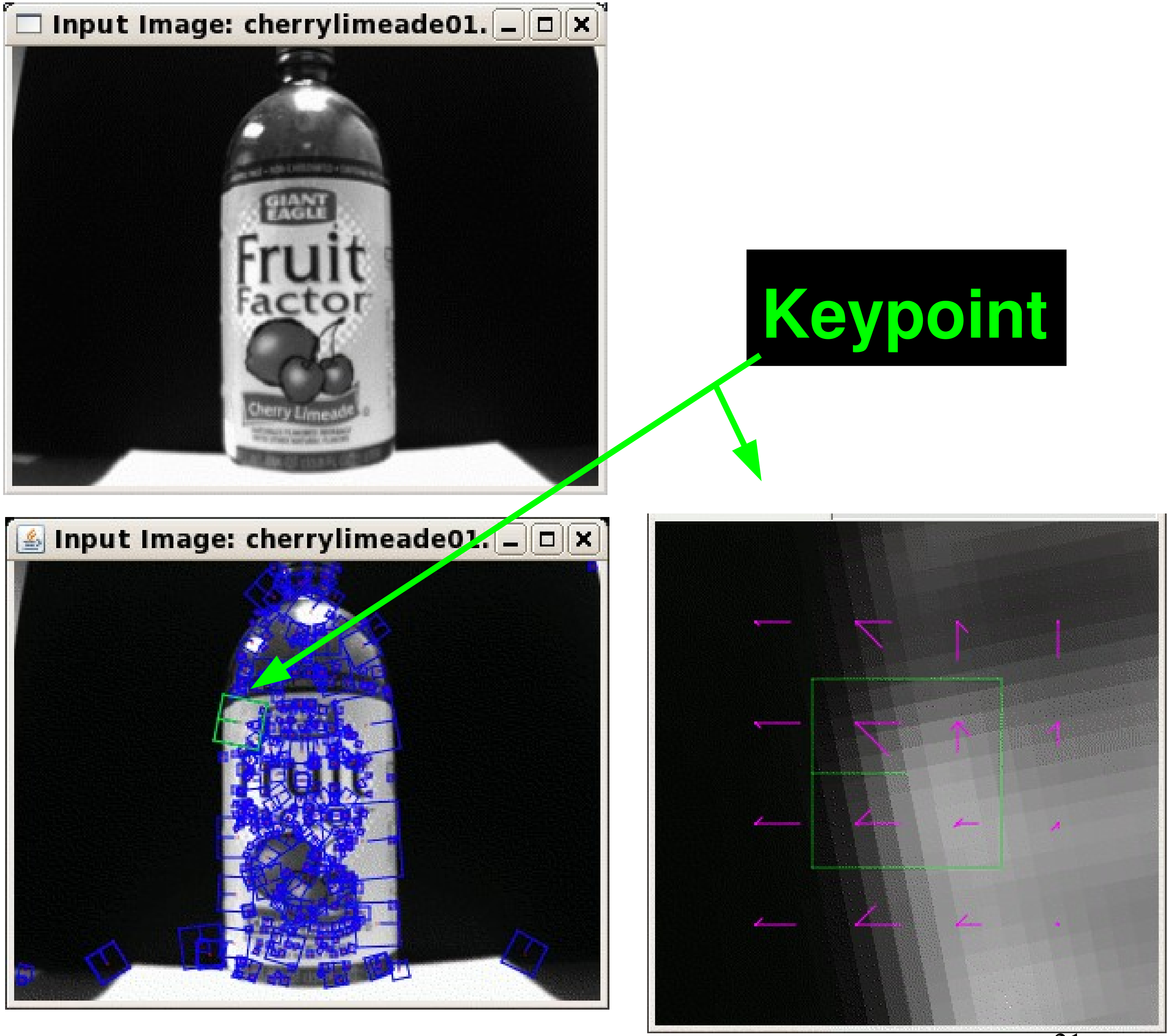
SIFT in Tekkotsu

- Implemented by Xinghao Pan at CMU.
- Built on the SIFT++ feature extractor by Andrea Vedaldi at UCLA.
- Matcher coded by Xinghao Pan based on algorithm by David Lowe.
- Java tool allows you to interactively assemble a library of object models, test new images against the models, and export models for use in Tekkotsu applications.

Pan's SIFT Demo

Image 1:

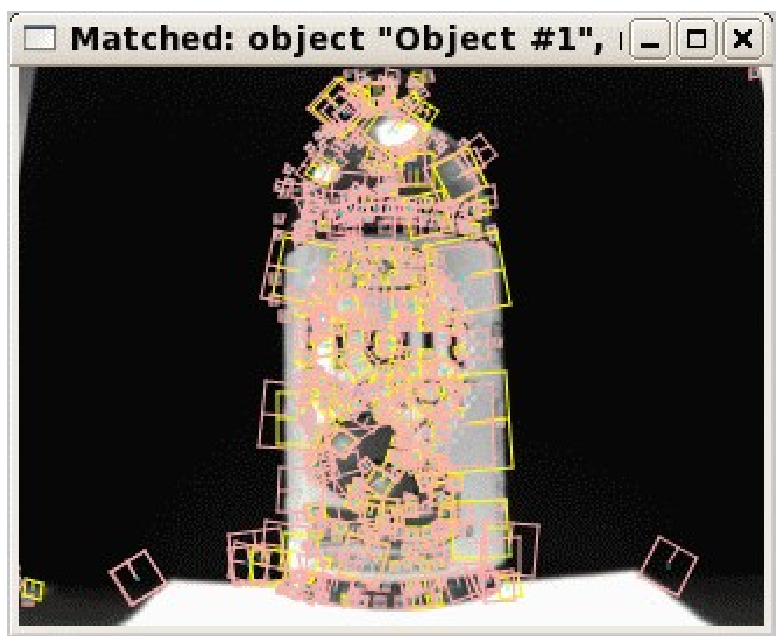
Model:



Pan's SIFT Demo

Image 2:

Model:

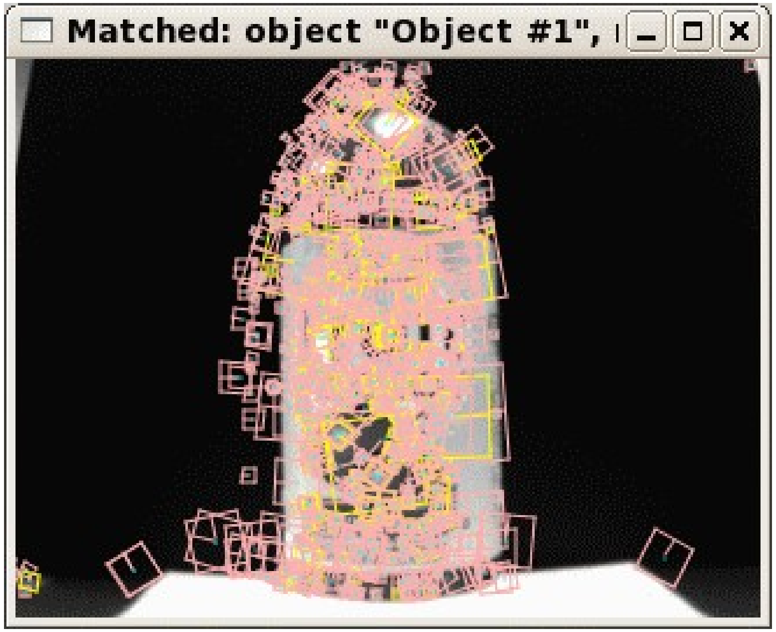


Yellow keypoints are matched in the image.

Pan's SIFT Demo

Image 3:

Model:



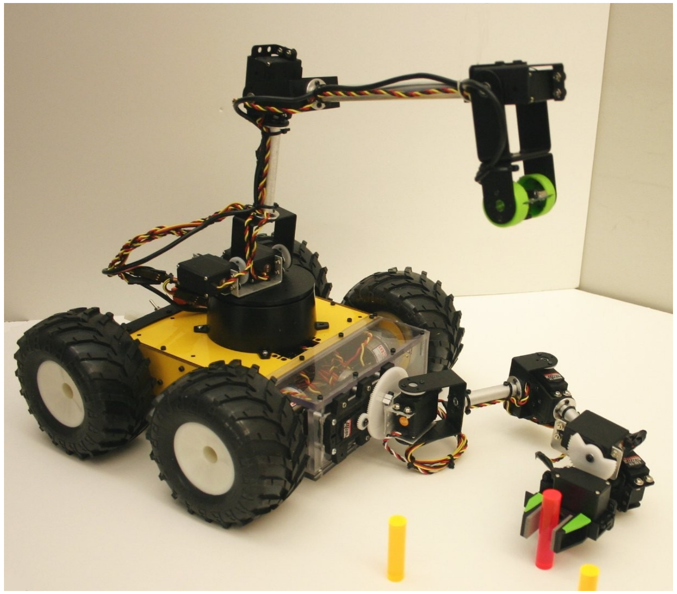
Model grows with each new training image.

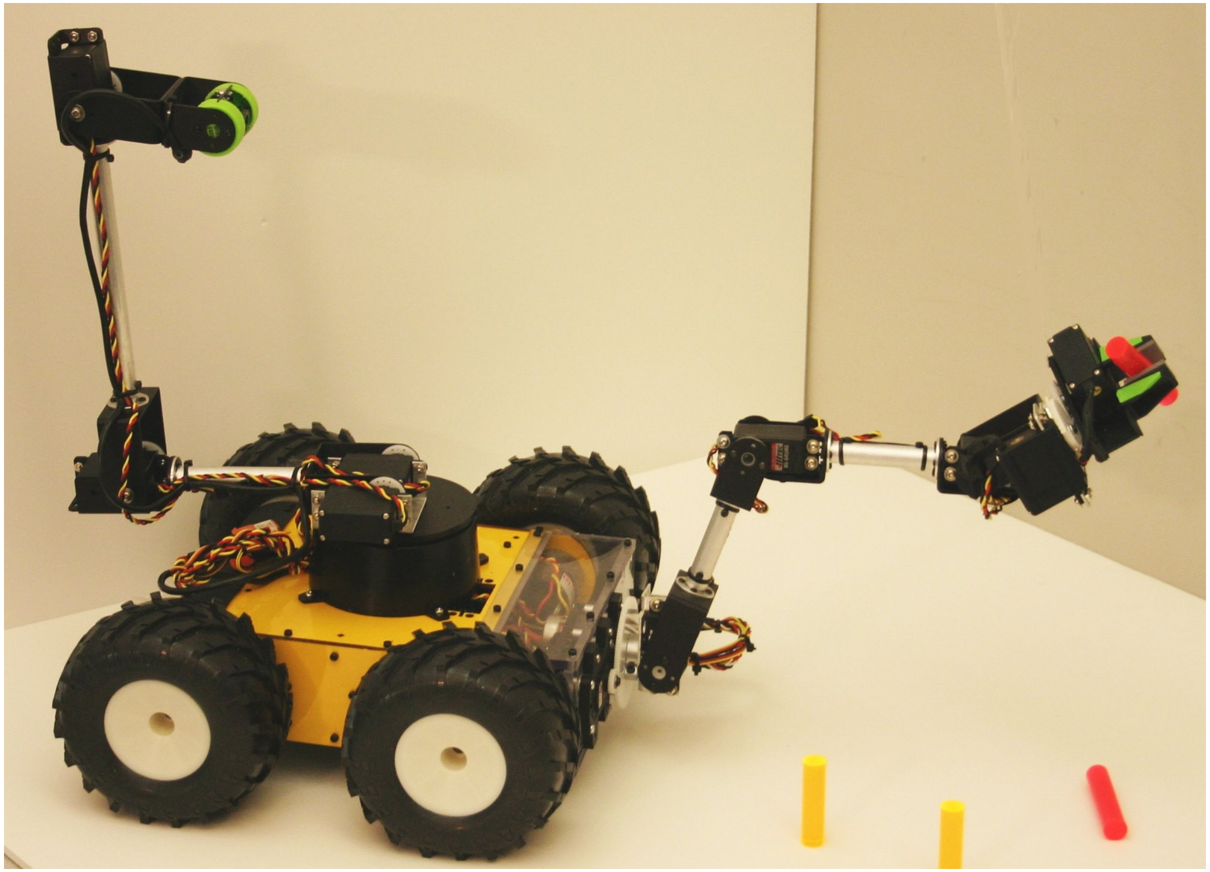
Beyond the AIBO: Qwerk

- Qwerkbot+ developed by Illah Nourbakhsh at CMU.
- Uses Qwerk controller board from Charmed Labs: $350.
- Robot recipes on the web:
  http://www.terk.ri.cmu.edu

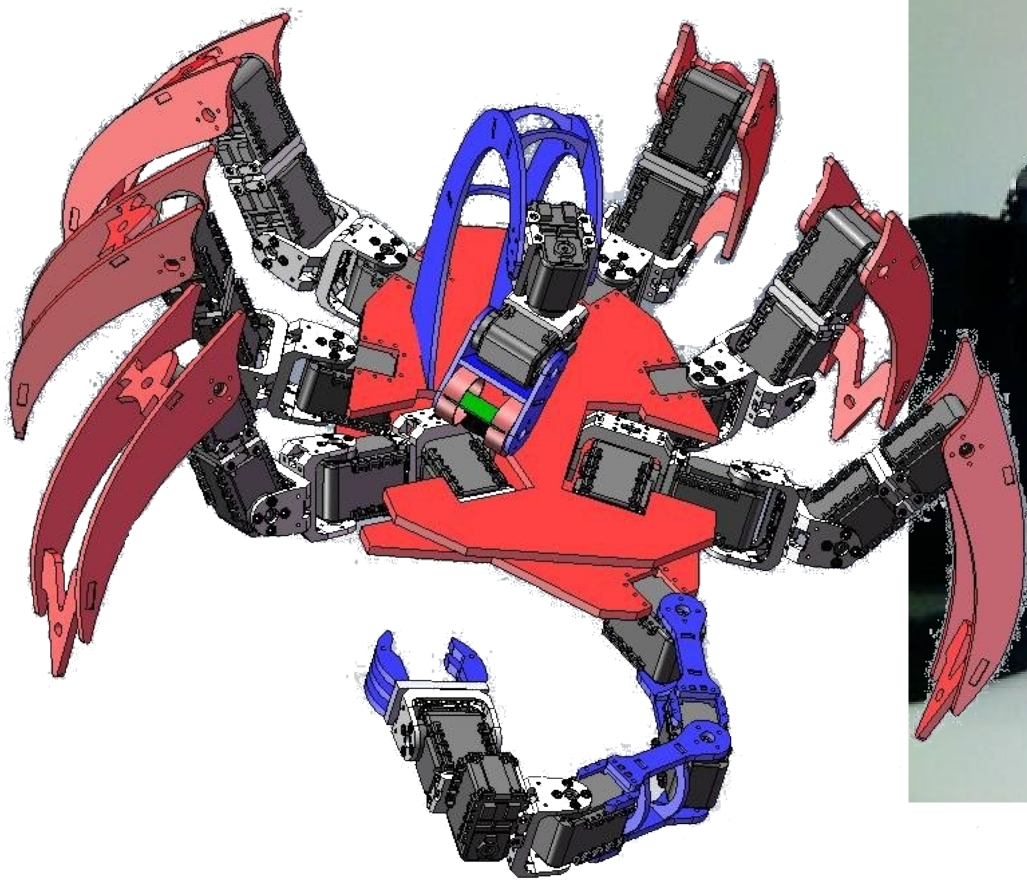
"Regis" Debuts at AAAI-07

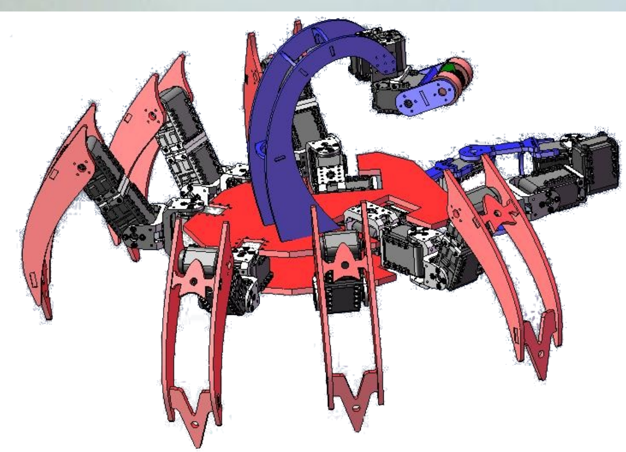
- Modified Lynx Motion 4WD3 base, SES arms
- "Goose neck" webcam
- Crab arm w/gripper
- 600 MHz Gumstix processor





In Development: Kathie Lee





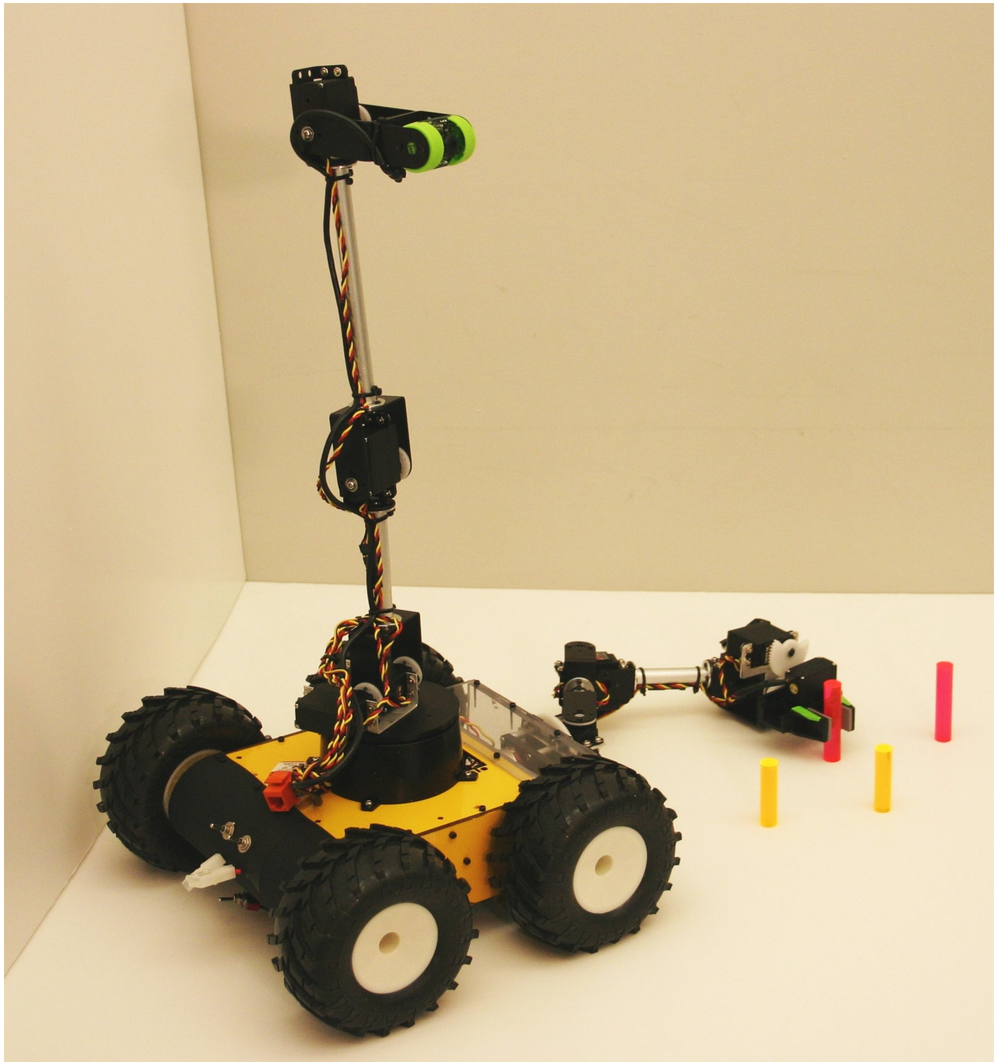
Robotis Dynamixel servos.

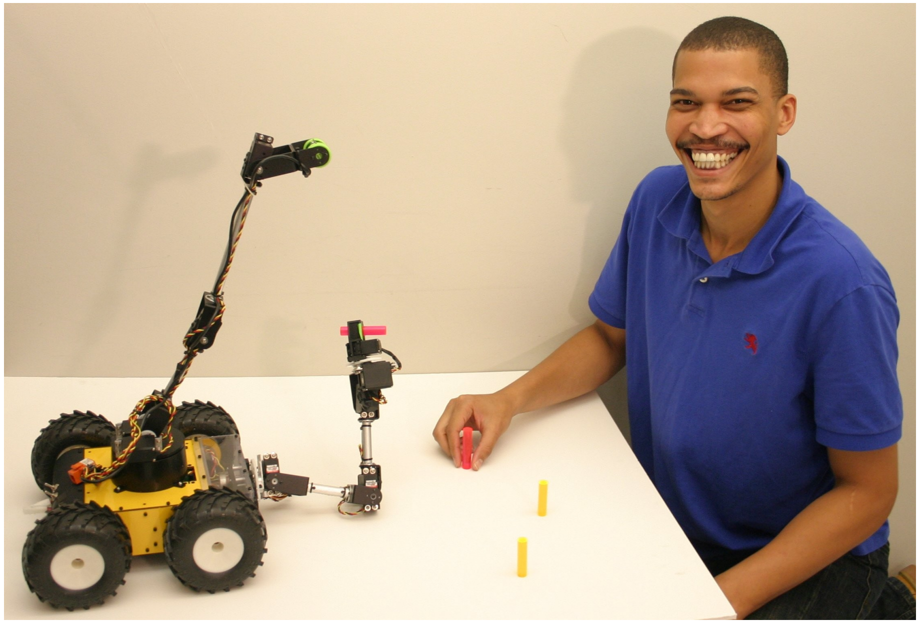
Pico-ITX proc: 1 GHz + 1 GB.

ARTSI: Advancing Robotics Technology for Societal Impact

- Alliance of 9 HBCUs and 7 R1 universities
- NSF funded: $2 million over 3 years
  - Student research projects
  - Faculty development
  - Outreach to middle and high schools
  - Viral marketing: YouTube videos
- Partners include Seagate, iRobot, Microsoft, and Apple

“Here's Looking At You”

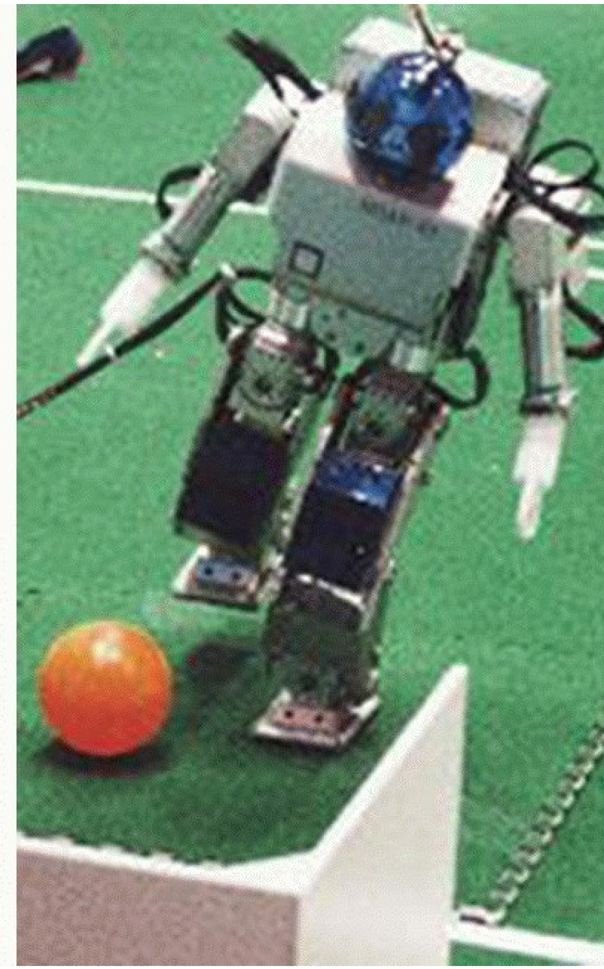




The Future of Cognitive Robotics

Honda's Asimo

Sony's QRIO (defunct)



Fujitsu's HOAP-2