Aveiro's Team in SRVC'2008 – software league

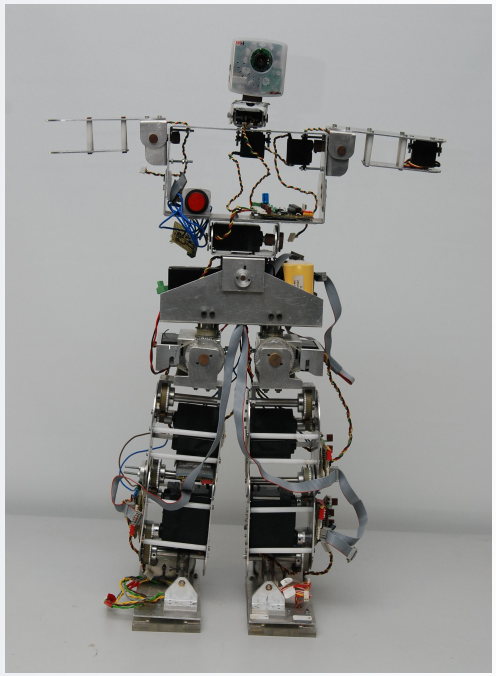
Luís Seabra Lopes, Frederico Valente,
Rui Pereira, Luís Ribeiro

Transverse Activity on Intelligent Robotics
IEETA/DETI - Universidade de Aveiro, Portugal

Transverse activity of on intelligent robotics

- The mission of the Transverse Activity on Intelligent Robotics (ATRI) is to study integrated architectures for intelligent robotic systems
  - The group builds on synergies between the main research areas of our institute
  - Created in 1999, acknowledging the strategic importance of developing this field of research in the institute
  - ATRI is involved in EURON – European Robotics Research Network and its interest groups

Interest in human-robot interaction



Interest in semantic vision for human-robot interaction

- Contribute to the development of robots that are easy to command and instruct by naïve users
- Spoken language dialogs with robots
- Integration with basic navigation, including guided navigation
- Knowledge acquisition and question answering
- Symbol grounding and language acquisition
- Visual object learning and recognition

Scenario in human-robot interaction

- Our robot, Carl – "Hi, how may I help you?"
- A user – "Hi Carl, please bring me the Spam in the kitchen table"
  - [but there are several objects in the table …]
- Carl – "Hum … "
  - [Carl searches the web to find out what Spam is]
- Carl – "Ok, I'll pick it up!"

Interest in semantic vision for robotic soccer

- According to the latest rules for the Middle-Size League of RoboCup, an arbitrary soccer ball can be used
  - Not necessarily orange!

Aveiro's team in SRVC

- Luís Seabra Lopes, faculty, team leader
- Undergraduate students
  - Frederico Valente
  - Rui Pereira
  - Luís Ribeiro
- Other faculty involved in supervision
  - Armando Pinho
  - Augusto Silva

Overview of entry

- Implemented in C/C++ and scripting languages
- A set of independent programs that store their results in the file system
  - Good for robustness of the overall system
- The learning phase (i.e. web search and model building) takes around 30 minutes
- The object detection and classification phase takes about 5 minutes!

General versus specific categories

- General categories tend to be more heterogeneous and more dependent on shape analysis
- Specific entities tend to be more homogeneous and provide enough local features for detection
  - particularly when they are branded and/or have elaborate graphic design

Global versus local descriptors

- General and specific categories are handled differently, according to the following heuristic
- Instances of specific categories are represented with SIFT features
  - The name of the category contains capital letters and/or quotes
- Instances of general categories are represented by log-polar descriptors

Web search and model building

- Search is done via Google only using English language only
- Unsupervised subset selection
- Category models are simply collections of instance descriptors

Unsupervised subset selection

- K-means clustering of fecthed images repeated several times
- Fetched images ranked by the number of times they occur in the largest cluster
- Number of instances included in the model depends on an evaluation of category heterogeneity
  - More instances are stored for more heterogeneous categories

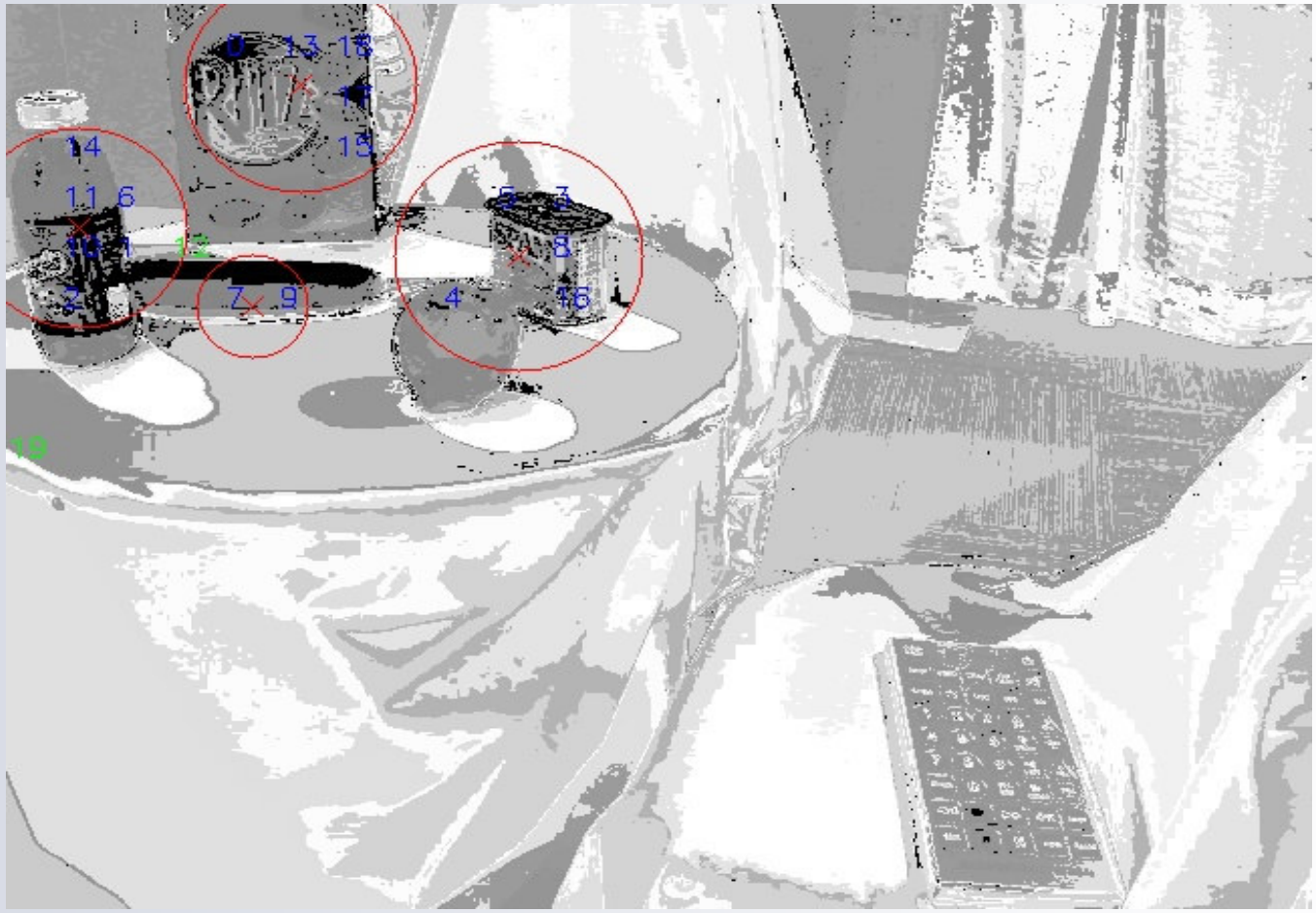
Performance phase

- For each image in the workset ..
  - Color-based saliency detection
  - Edge detection and object segmentation
  - Nearest-neighbor (or maximum similarity) evaluation of category membership
  - A voting scheme was also implemented, but not used in the competition

Color-based saliency detection

- Color histogram of the target image
- Saliency image is obained by rating each pixel of the original image proportionally to the relative frequency of the respective color
- Sparse analysis of color saliencies of the image, providing a set of key points
- Unsupervised clustering of keypoints

Color-based saliency detection



Results in SRVC'2008

- Trial run:
  - 2 instances of general categories: "blue hat", "coffee cup"
  - 4 instances of specific categories
- Competition run:
  - 33 minutes + 5 minutes
  - 1 instance of a general category: "fax machine"
  - 4 instances of specific categories: "Doritos", "Spam", "Ritz", "I-Robot"