RSS 2010

Towards Closing the Loop:
Active Learning for Robotics

Active Scene Analysis

N.Bergström, M.Björkman, J.Bohg,
M.Johnson-Roberson, G.Kootstra D.Kragic
CAS/CVAP, KTH Stockholm, Sweden

ROYAL INSTITUTE OF TECHNOLOGY

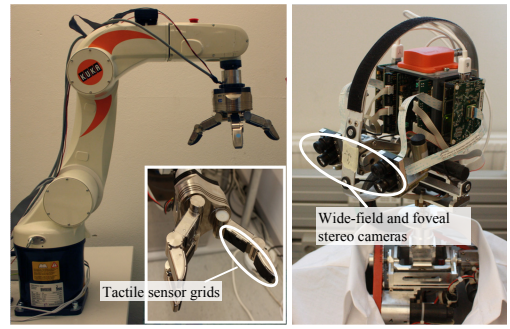
Introduction

In this work we want to perceive and represent unknown scenes with unknown objects.

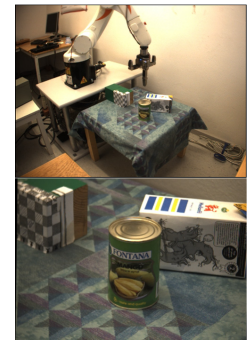
- "Unknown" means that only bottom-up mechanisms can be used for analyzing the scene
- The robot's active capabilities are used to gather more information about the scene.
  - The robot's gaze is controlled to verge on object hypotheses
  - The robot's manipulator with tactile sensing capabilities is used to explore parts of the scene not seen from the robot's head
- Through integration of different sensor modalities, a more complete model of the scene can be achieved.

The result is a spatial representation of the scene containing objects segmented from the background

Setup



The robot consists of an Armar III, 7 DOF active head and a 7 DOF Schunk Dexterous Hand with Weiss tactile sensors, mounted on a 6 DOF KUKA arm.

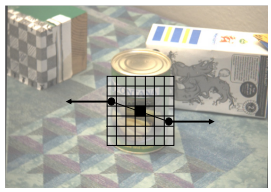


A scene as seen from the robots left wide-field (upper) and foveal (lower) cameras.

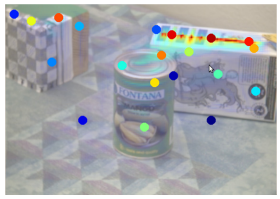
Methods

1 – Object Hypotheses

- Generates object hypotheses in the foveal view.
- Uses one of the Gestalt principles, *symmetry,* as a queue.
- High probability that a hypothesis belongs to an actual object, and that they are found close to the center of the object.



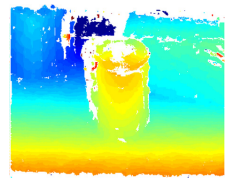
The symmetry kernel is run over several scales and the results are superposed. Gradients of opposing grid elements are compared.



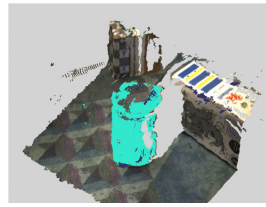
Final result. Object hypotheses correspond to colored circles. Hypothesis are sorted from strongest (red) to weakest (blue).

2 – Segmenting Objects

- Given an object hypothesis, fixate on that point and segments the object belonging to that hypothesis from the background.
- Uses color, contrast and disparity information.
- Depth information helps to deal with heterogeneously colored objects.
- Assumes a table top scenario and segments three classes: foreground, background and table.



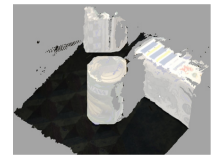
Disparity map of the scene. Red is closer, blue more distant.



Reconstructed view with one of the objects segmented.

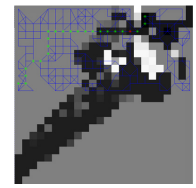
3 – Exploring the Scene

- By visiting several object hypotheses belonging to different objects, a segmented 3D scene is reconstructed.
- A 2D occupancy grid represents the scene from the top – missing parts due to occlusion.
- A Gaussian process is used to predict unknown scene parts, as well as guide tactile exploration towards areas with high uncertainty.

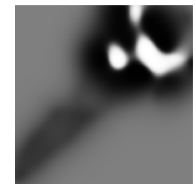


State of the scene. White is occupied space, black is free space.

Occupancy grid. Shows top view of the scene.



Exploration plan. Blue lines show possible paths. Green stars show chosen path. Red star shows robot position.



Predicted occupancy grid. The brighter intensity, the more probable that the area is occupied.

References

[1] G. Kootstra, N. Bergström and D. Kragic, "Using symmetry to select fixation points for segmentation", ICPR, 2010
[2] M. Björkman and D. Kragic, "Active 3d scene segmentation and detection of unknown objects", ICRA, 2010
[3] J. Bohg, M. Johnson-Robertson and D. Kragic, "Strategies for Multi-Modal Scene Exploration", IROS, 2010 (To appear)

This research has been supported by the EC through

IST-FP7-IP-215821-GRASP and Swedish Foundation for Strategic Research.