2D/3D Object Categorization for Task Based Grasping

Marianna Madry, Dan Song and Danica Kragic

We present an object categorization system integrated with a grasp planning and reasoning system (see Figure 1). The main motivation for the work is to equip robots with the ability of transferring grasping knowledge between objects that belong to the same category. The categories are defined based on their geometric properties and functionality, relating to the idea of affordances.

In the heart of the system, there is the *Object Categorization System* (OCS) using camera images as input. The system employs both 2D (RGB image) and 3D (point cloud representing the visible part of an object) information about an object. In designing the 2D/3D Object Categorization System a two-fold approach was adapted. First, we build the classic single cue OCS for each of the descriptors capturing the following object properties: (a) color (opponentSIFT), (b) 2D shape (HoG) and (c) 3D shape (FPFH), and then these systems are integrated to provide the final decision based on all cues. We present and test several 2D/3D integration strategies. The system is evaluated on real data collected using an active stereo head, capable of vergence and foveation. The data is generated in natural scenes, for a number of household object categories. The results showed that the proposed system achieved high object recognition rate (up to 91%), significantly better than the classic single cue OCS in the same task. The system is built upon an active scene segmentation module, able of generating object hypotheses thus segmenting them from the background in real-time [1]. We integrated the OCS with a task-constraint model for robot grasping [2], [3]. The results showed that the object categorization is very useful for reasoning and planning of goal-directed grasps in natural scenes with multiple objects.

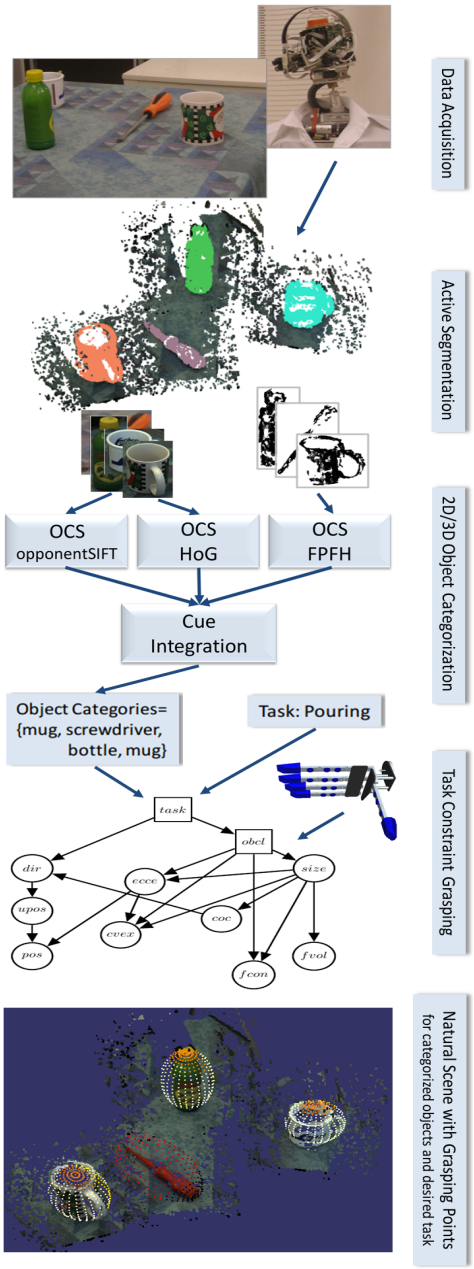

Fig. 1. System outline. First row: Data acquisition using ARMAR III robot head and view of a typical experimental scene. Second row: Segmented objects in the same scene. Third row: Integrated 2D and 3D Object Categorization Systems (OCSs). Fourth row: Generation of grasping points by Bayesian network. Fifth row: Experimental scene with grasping points for categorized objects and desired task (the lighter is the point the higher is the probability). For the accurate color information we kindly direct to the electronic version of the abstract.

REFERENCES

[1] M. Bjorkman and D. Kragic, "Active 3D scene segmentation and detection of unknown objects," in *ICRA*, 2010.

[2] D. Song, K. Huebner, V. Kyrki, and D. Kragic, "Learning Task Constraints for Robot Grasping using Graphical Models," in *IROS*, 2010.

[3] D. Song, C.-H. Ek, K. Huebner, and D. Kragic, "Multivariate discretization for bayesian network structure learning in robot grasping," in *ICRA*, 2011, to appear.

The authors are with KTH – Royal Institute of Technology, Stockholm, Sweden, as members of the Computer Vision & Active Perception Lab., Centre for Autonomous Systems, e-mail addresses: madry,dsong,danik@csc.kth.se. This work was supported by EU through the project GRASP, IST-FP7-IP-215821 and Swedish Foundation for Strategic Research.