Range Image Segmentation for Object-based Attention in the Context of Service Robotics

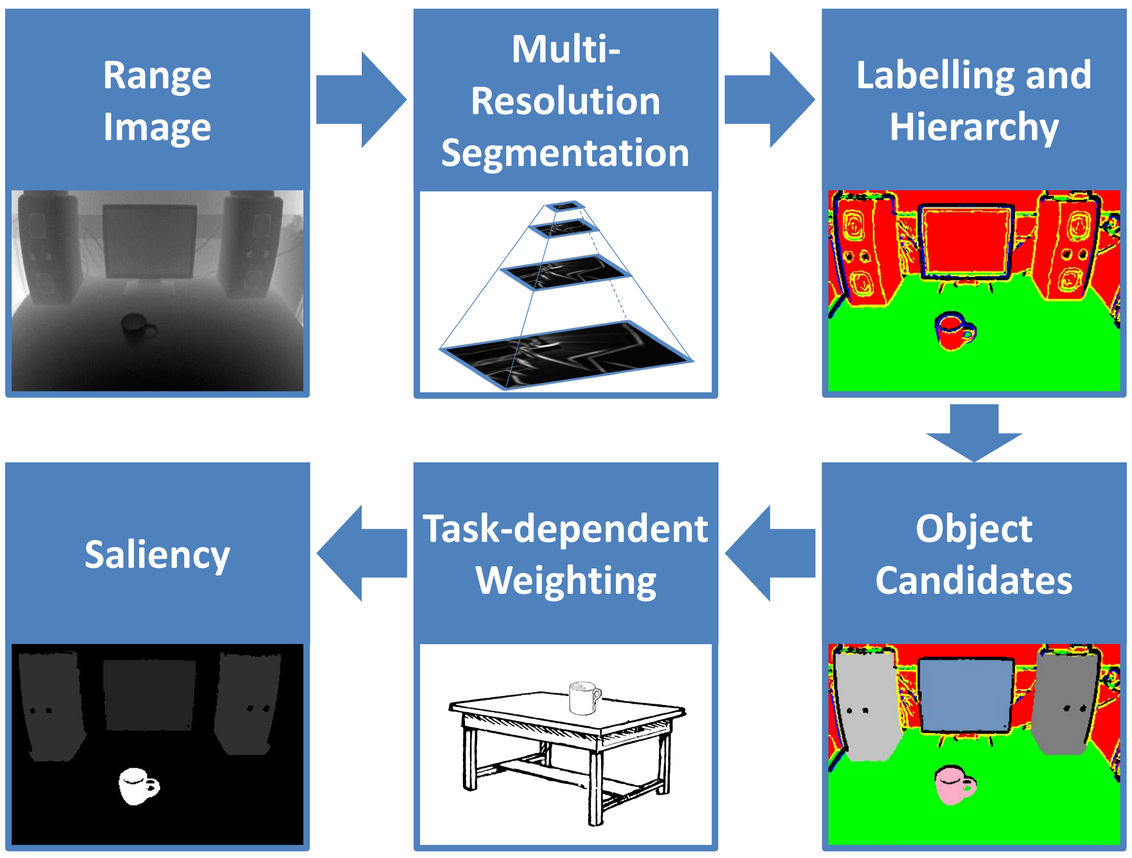
Peter Einramhof, Robert Schwarz, Markus Vincze
Vienna University of Technology, Vienna, Austria

Motivation

Service robots must be able to safely navigate in cluttered and dynamic home environments. Furthermore, they need attention mechanisms to efficiently search and detect task-relevant objects. Since the physical world is three-dimensional, the robot's perception system requires input data encoding structural information about the observed scene – typically range images. Bottom-up visual attention by Itti et al. [1] has been extended to accept range images as input, but despite their different meaning the same operators are applied as to 2D (color) images. Further extensions aim at deploying attention at object level by first computing a saliency map and then estimating a proto-object region at the most salient location [2]. We propose a more "robot-compatible" approach.

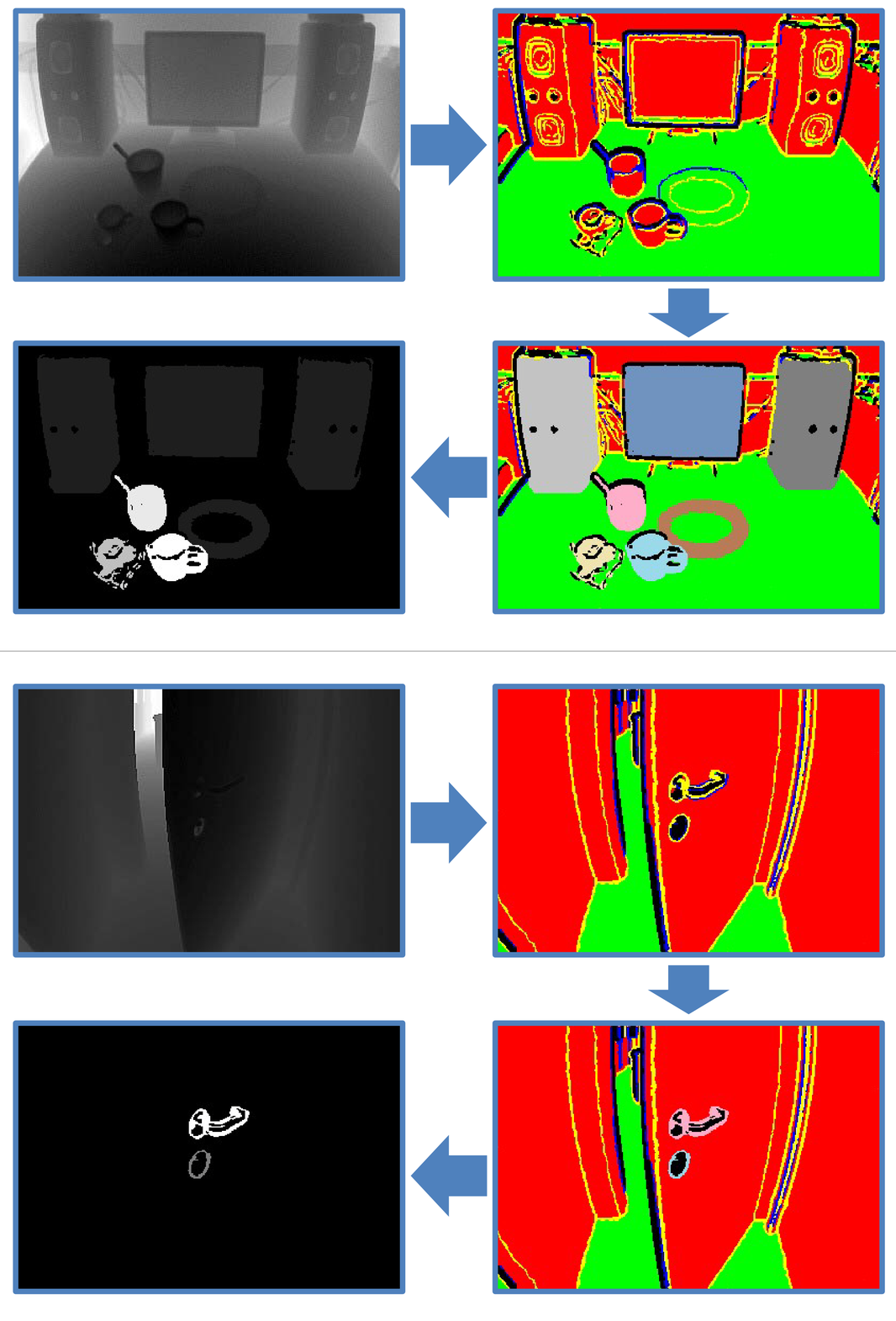
Approach

Range images are segmented into features, smooth and planar surface patches, step edges, and convex and concave roof edges [3]. Surface patches represent the boundaries of the environment, support planes and faces of objects, while edges indicate contours of and transitions between objects or parts thereof. Segmentation is done at multiple resolutions – coarse to fine – for selectable level of detail. The segmentation results serve as immediate input for navigation and are also used to establish a background-foreground hierarchy of object candidates. The latter possess attributes such as a 3D bounding box, pose within the scene and distance to the robot. Depending on how well their attributes fit those of the target object, the candidates receive a "saliency value" and are sorted in descending order. Finally, the object candidates with the highest saliency values are sequentially attended.



Results

Two examples for robotic tasks are shown in the images below: grasping a cup on a table (top) and opening a door (bottom). Each example shows the range image, the labelled image as segmentation result, object candidates and saliency of these candidates based on the relevance for the task. The table scene is representative for objects on a horizontal support plane, and the door scene for an object on a vertical support plane.



References

[1] Itti, L. and Koch, C.: Computational modelling of visual attention. Nat. Rev. Neurosci, 2(3):194-203, March 2001

[2] Walther, D. and Koch, C.: Modeling attention to salient proto-objects. Neural Networks 19(9): 1395-1407 (2006)

[3] Einramhof, P., Schwarz, R. and Vincze, M.: Fast Range Image Segmentation for a Domestic Service Robot. International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD), Brno, Czech Republic; Oct 5-7, 2011

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°248623, TACO (Three-dimensional Adaptive Camera with Object Detection and Foveation)