Knowledge-Enabled Robotic Agents for Shelf Replenishment in Cluttered Retail Environments

(Extended Abstract)

Jan Winkler, Ferenc Bálint-Benczédi, Thiemo Wiedemeyer, Michael Beetz
Institute for Artificial Intelligence
University of Bremen, Germany

Narunas Vaskevicius, Christian A. Mueller, Tobias Fromm, Andreas Birk
Robotics, CS & EE
Jacobs University Bremen, Germany

ABSTRACT

Autonomous robots in unstructured and dynamically changing retail environments have to master complex perception, knowledge processing, and manipulation tasks. To enable them to act competently, we propose a framework based on three core components: (∘) a knowledge-enabled perception system, capable of combining diverse information sources to cope with occlusions and stacked objects with a variety of textures and shapes, (∘) knowledge processing methods produce strategies for tidying up supermarket racks, and (∘) the necessary manipulation skills in confined spaces to arrange objects in semi-accessible rack shelves. We demonstrate our framework in an simulated environment as well as on a real shopping rack using a PR2 robot. Typical supermarket products are detected and rearranged in the retail rack, tidying up what was found to be misplaced items.

General Terms

Algorithms, Experimentation

Keywords

robot perception; knowledge-based reasoning;

1. INTRODUCTION

Robotic agents that are embodied and can manipulate the world in order to accomplish human-scale manipulation tasks in open and realistic environments are still a largely unresolved challenge.

Lately robotic solutions for retail environments are becoming more and more popular. Common tasks include refilling product shelves, and putting misplaced products back to where they belong. Retail scenarios are semi-structured, but unpredictable in details, products might be missing, requirements of where to place which item change, or products and shelves are partially obstructed. The products are placed such that they can be easily seen, the front side is typically visually distinctive but they also constitute challenges as identical items are placed directly next to each other, complicating object segmentation as well as manipulating objects without causing side effects on the neighboring items. The tasks that the robotic agents are to accomplish include loading an empty rack, restocking sold items, cleaning up unordered shelves, rearranging product configurations, etc.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, S. Marsella, C. Jonker (eds.), May 9–13, 2016, Singapore.*

Figure 1: Retail environment showing the used robot next to the perceived objects of an example scene

In this paper, we outline our systematic solution for robotic agents that perform a limited kind of shelf reordering. The robotic agent takes a qualitative spatial description of how the items in a shelf should be re-ordered, such as the cereal should be placed to the right of the coffee. The robotic agent then tessellates the target region of the items into variable size grid cells that are allocated for the individual product groups. The items are to be placed in the respective grid cell, next to each others and facing the front.

The solution presented herein is unique on the first hand in the way that objects can be recognized and grasped even if they are subject to heavy occlusion, or similar objects are stacked or aligned with each other (see Fig. 1). Multiple instances per object are not a limitation either. Second, our approach focuses on reasoning about objects' semantics which is crucial for everyday scenarios where the requirement is to achieve a certain final object configuration. Additionally, since we want to avoid to accidentally clear up the whole shelf, awareness and avoidance of clutter objects are a central requirement in such a scenario.

2. SYSTEM OVERVIEW

Addressing the problems stated in this paper would not be possible without a tight integration between specific modules that are needed by an autonomous robotic agent.

Three main components were used to build up the system, each of them introduced recently, mainly because they are open-source software and because they meet the required needs of being modular and easy to extend: (∘) CRAM [2] – a high-level robot planning

Set 1		Set 2	
Camera	Processed	Camera	Processed

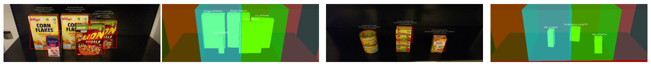

Figure 2: Example experiment arrangements of objects in a shopping rack. The left images show the source camera data. The right images depict the processed object instances recognized by the perception system.

and reasoning system (◦) KNOWROB [4] – a centralized knowledge processing and inference framework (◦) ROBOSHERLOCK [1] – a knowledge-enabled robotic perception framework.

Furthermore, ROBOSHERLOCK was extended with two essential perception components and tasked with the gist of the recognition process: a texture [5] and a shape-based [3] recognition system.

The entry point to the system is a generalized plan for rearranging the shelf, generated by CRAM and the target arrangement of the shelves in a specific aisle defined in our knowledge base. As a first step, ROBOSHERLOCK is queried for the detected items. At this point, using knowledge about the environment stored in the semantic map of KNOWROB, the raw data is filtered in order to contain only the regions of interest for the current task, that is, the respective shelf. Based on the filtered images, the instance and shape recognition then hypothesize about possible products and their locations, and transmit these back to CRAM through ROBOSHERLOCK.

Internally, the plan generation evaluates different strategies for executing the rearrangement task, which are ranked based on their intrinsic cost as calculated by an A*-based planner implemented in CRAM. The best ranked strategy is then chosen to be executed.

Once the best strategy has been chosen, the actual manipulation plan is generated and executed by the robot. In case of failures that occur during the manipulation (e.g. unable to grasp object, dropped object, unable to plan manipulation trajectory, etc.), control is handed back to CRAM where the contingency is resolved and an alternative plan is generated.

3. RESULTS

To evaluate the importance of harmonizing the different system components of our approach when dealing with complex shopping rack scenarios, we present a series of example cases. In these, we demonstrate the effects of slight variations in perception data on the output of the planning algorithm, and thus, in the manipulation.

Figure 2 depicts two examples of 20 shopping rack situations we created, separated into two qualitatively different series with varying pose, obstacles, unknown objects and stacking of similar and different objects. In some cases, objects were not or wrongly detected. The planner's goal was to group the available objects and evenly distribute them over two of the rack's shelves.

The cost shown in Table 1 represents the accumulated cost of all individual actions inside of a planned sequence as used by A*. Picking and placing are rather simple actions, while moving the robot's torso takes quite some time. As execution time is a quality criterion while tidying up shopping racks, it is more expensive. Moving the robot's base is rather fast, and handing over an object from one hand to the other takes longer than picking and placing.

As experiments suggest, the cost does not exceedingly increase when meeting difficulties, such as obstructions, or stacked objects.

Failures in perception are therefore not directly reflected in the planner's output. In one case, two stacked objects were identified as only one object, leading to a simpler, and wrong plan. In another

	Time [s]	# Pick	# Place	# Move Torso	# Move Base	Cost
Set 1	5.35	4	4	6	2	23.6
	(18.6 ± 37.8)	(4.4 ± 0.5)	(4.4 ± 0.5)	(5.6 ± 1.6)	(1.2 ± 1.0)	(23.0 ± 2.7)
Set 2	4.3	4.5	4.5	5.0	2.0	22.0
	(20.0 ± 32.5)	(4.5 ± 0.8)	(4.5 ± 0.8)	(5.3 ± 1.3)	(1.4 ± 1.0)	(22.8 ± 3.9)

Table 1: Summary of details from action sequences as generated based on the data from two qualitatively different sets, with 10 scenes each. "Time" is the time taken to plan the action sequence, "Cost" is the accumulated cost of all individual actions (top: median, bottom: mean ± standard deviation).

case, both lower objects on the stacks are not detected at all, also resulting in a wrong configuration. In these cases, we rely on the failure detection and recovery mechanisms which are implemented in CRAM. After every manipulation action, the performing robot re-perceived and validated the current scene, and re-plans its strategy if inconsistencies are detected. The video we created in the course of this work[1] shows the system being executed on a PR2 robot.

4. CONCLUSION

We have shown a novel application of knowledge-based manipulation in an everyday retail scenario which requires extended perception and reasoning capabilities.

In the current state of the art it is hard to measure the competence of full robotic systems other than evaluating individual components. In our experiments we show how the results of visual perception reflect in the generation of manipulation plans for robotic agents. Having a quantifiable connection between beliefs about the real world and the quality of the plans for the robotic agents allows for further investigation of other modalities for improving the robot behavior.

Acknowledgements

This work was partly supported by the EU FP7 project *RoboHow* (Grant Agreement Number 288533).

REFERENCES

[1] M. Beetz, F. Balint-Benczedi, N. Blodow, D. Nyga, T. Wiedemeyer, and Z.-C. Marton. RoboSherlock: Unstructured Information Processing for Robot Perception. In *International Conference on Robotics and Automation*, 2015.

[2] M. Beetz, L. Moesenlechner, and M. Tenorth. CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. In *International Conference on Intelligent Robots and Systems*, 2010.

[3] C. Mueller, K. Pathak, and A. Birk. Object shape categorization in RGBD images using hierarchical graph constellation models based on unsupervisedly learned shape parts described by a set of shape specificity levels. In *International Conference on Intelligent Robots and Systems*, 2014.

[4] M. Tenorth and M. Beetz. KnowRob: A knowledge processing infrastructure for cognition-enabled robots. *International Journal of Robotics Research*, 32(5):566–590, 2013.

[5] N. Vaskevicius, K. Pathak, A. Ichim, and A. Birk. The Jacobs robotics approach to object recognition and localization in the context of the ICRA'11 Solutions in Perception Challenge. In *International Conference on Robotics and Automation*, 2012.

[1] `https://youtu.be/xFwinZAHrnA`