WASP Project Course 2023

Object Picking and Constrained Placement by Visual Reasoning

Background

Pick-and-place is a common task that robot manipulators are applied for today. The objective of this project is to develop a robotics system that is capable of picking objects and placing them according to predefined constraints. An example is shown in Figure 1. Three colored (red, green, blue) boxes are placed. Each box is equipped with three differently shaped holes, specifically triangular, square, and circular. The objects (tetrahedrons, cubes and cylinders) must be placed into holes with certain color and shape constraints. For instance, the predefined constraints can be putting objects into the box with same color.

To pick-and-place with certain predefined constraints, we divide the task into two stages:

- The first stage is task planning, which generates a high-level plan using visual reasoning [1, 2, 3], containing two modules:
    1. Visual detection module to detect all objects and container attributes in the scene.
    2. High-level task planner to reason and generate a plan adhering to the given constraints. A plan comprises high-level actions wherein each action involves picking and placing objects. For example, one action might entail picking up a blue cube and placing it into a blue container.
- The second stage is task execution. Each action in the generated plan is performed through three distinct steps.
    1. Picking a target object from a given position.
    2. Estimating the in-hand 6D pose of the grasped object.
    3. Placing the grasped object in the target container.

There are several challenges in this project. In the first stage, for instance, the output from the planning phase depends on the visual detection module's ability to accurately estimate the object's shape and color. If the detection module is faulty, the planning output may become infeasible. Furthermore, the generated plan must adhere to the constraints.

In the second stage, there are several works [4, 5, 6, 7] focusing on the 6D object pose estimation. However, their object is placed on the table or grasped in the two-finger gripper. Grasping by the multi-finger hand may cause a more complex and larger region of occlusions.

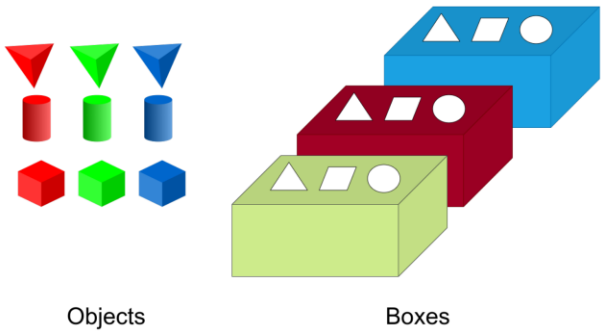

Figure 1: An example of experiment setup. The task is to pick objects (**left**) and place them into corresponding boxes (**right**) under predefined constraints.

Participants

Industrial partner: ABB

Industrial supervisor:

Sichao Liu, sichao.liu@se.abb.com

Academic supervisor:

Martin Magnusson, martin.magnusson@oru.se, Örebro University

Coordinating WARA representative:

jonas.larsson@se.abb.com, WARA-robotics

Suggested WASP PhD students:

Yufei Zhu, yufei.zhu@oru.se, Örebro University

Rishi Hazra, rishi.hazra@oru.se, Örebro University

Shih-Min Yang, shih-min.yang@oru.se, Örebro University

Challenges to investigate

- Estimating object attributes (shape, color) and reasoning to find a feasible plan
- Estimating small object pose with certain level occlusion

Resources

Örebro university: a Kinect Azure camera, a Franka Emika robot arm, a two-finger gripper and a multi-finger soft hand.
WARA- robotics

Deliverables

- A visual reasoning module as a high-level task planner
- A pose estimation module to estimate in-hand object pose
- A manipulation module for picking, changing the object pose, and placing
- A real robot demonstration

References

[1] The Neuro-Symbolic Concept Learner: interpreting Scenes, Words, and Sentences from Natural Supervision (ICLR 2019)
[2] CLEVRER: Collision Events for video representation and reasoning, ICLR 2020
[3] Text2Motion: From Natural Language Instructions to Feasible Plans, arXiv 2023
[4] Hu, Y., Fua, P., Wang, W., & Salzmann, M. (2020). Single-stage 6d object pose estimation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2930-2939).
[5] Hoang, D. C., Stork, J. A., & Stoyanov, T. (2022). Voting and Attention-Based Pose Relation Learning for Object Pose Estimation From 3D Point Clouds. IEEE Robotics and Automation Letters, 7(4), 8980-8987.
[6] Nguyen, V. N., Hu, Y., Xiao, Y., Salzmann, M., & Lepetit, V. (2022). Templates for 3D object pose estimation revisited: generalization to new objects and robustness to occlusions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6771-6780).

[7] Wen, B., Mitash, C., Soorian, S., Kimmel, A., Sintov, A., & Bekris, K. E. (2020, May). Robust, occlusion-aware pose estimation for objects grasped by adaptive hands. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 6210-6217). IEEE.

Keywords

Object Pose Estimation / Visual Reasoning / Computer Vision / Robotic Manipulation