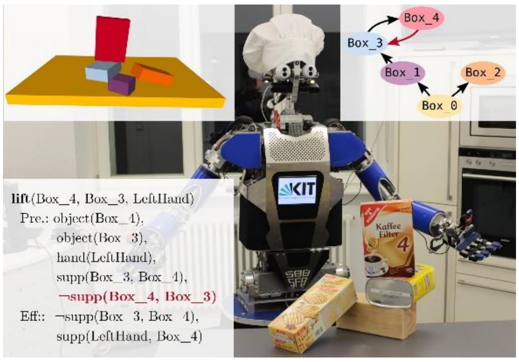
Subsymbolic Prediction of Bimanual Manipulation Action Effects

A robot needs to understand the possible effects of its actions on the environment in order to achieve manipulation goals and avoid undesired side effects. For example, pushing an object on a pile of other objects has the goal of moving the pushed object to a desired pose. However, due to interactions between the pushed object and the pile, other objects might move as well. In a recent work, we used a semantic scene representation consisting of physically plausible support relations between objects to identify possible unsafe actions, which might cause other objects to fall. This work utilizes symbolic preconditions and effects of the executed actions, to detect and handle these situations.



The humanoid robot ARMAR-III extracts objects and physically plausible support relations between them from its depth camera. Uncertain support relations can be used to predict undesired action effects.

In contrast to symbolic prediction, recent works consider the effects of robot actions on the perceptual level, e.g. a robot perceives its environment through a depth camera and executes a push action [1]. The goal is to predict the perceived depth image after the push action. Subsymbolic prediction has been done on multiple perceptual modalities including camera images [2], depth images [1] and optical flow [3]. These approaches use deep neural networks to learn the respective prediction models. However, they only consider single actions, like pushing in a specific direction.

We now want to research whether these subsymbolic prediction models can be used for more sophisticated robot actions. In particular, we are interested in bimanual manipulation actions, in which a humanoid robot uses one hand to execute a primary action and the second hand is used to prevent unwanted side effects.

References

[1] A. Byravan and D. Fox, "SE3-nets: Learning rigid body motion using deep neural networks," IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017, pp. 173-180.

[2] A. Eitel, N. Hauff, and W. Burgard, “Learning to singulate objects using a push proposal network,” In Proc. of the International Symposium on Robotics Research (ISRR), 2017.

[3] A. Dosovitskiy et al., "FlowNet: Learning Optical Flow with Convolutional Networks," IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 2758-2766.