

Masterarbeit

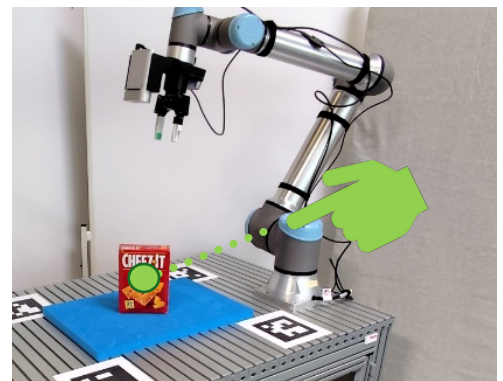
Point-to-Generalize: Visual Pointing for Object Selection and Bottleneck Specification in Single-Shot Imitation

Problem formulation

Single-shot imitation learning methods enable robots to learn new manipulation tasks from only one demonstration [1]. However, these approaches often assume a single relevant object in the scene and may require manually specified bottleneck poses or key intermediate states. In real environments with multiple objects, these assumptions limit practical usability. Humans naturally resolve such ambiguities by pointing to objects or locations of interest. This thesis investigates whether visual pointing can serve as an intuitive interaction modality to specify target objects and intermediate poses for single-shot imitation learning.

Task definition

The student will develop a system that integrates visual pointing with a single-shot imitation learning framework. Using a camera-based pointing interface or image-based clicking, the system will estimate the 3D location of the pointed object or target pose within the robot workspace. This information will be used to ground task execution, for example, by selecting the object to manipulate or defining bottleneck waypoints. The approach will be implemented and evaluated on a real robot platform across manipulation tasks such as pick-and-place and object relocation.



Requirements

- Solid knowledge base and experience in deep learning and robotics.
- Good coding skills in Python.
- Experience with ROS2 is beneficial

Literature

[1] K. Dreczkowski, P. Vitiello, V. Vosylius, and E. Johns, "Learning a thousand tasks in a day," Sci. Robot., vol. 10, no. 108, p. eadv7594, Nov. 2025, doi: 10.1126/scirobotics.adv7594.

Betreuer: M.Sc. Edgar Welte

Betr. Hochschullehrerin: Prof. Dr.-Ing. Rania Rayyes

Bearbeitungszeitraum: SS 2027

Bearbeiter: offen