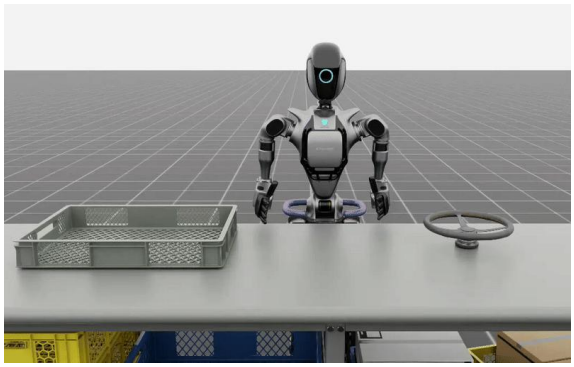
[MA/BA] Humanoid dual-arm manipulation with LLM code generation

Introduction



Recent progress in humanoid dual-arm manipulation has been largely driven by embodied learning approaches that rely on massive datasets and imitation learning at scale, such as vision-language-action (VLA) models. While effective, these methods are extremely data-hungry and demand substantial computational resources, which significantly limits their accessibility, scalability, and adaptability to new tasks and environments. In contrast, large language models (LLMs) inherently possess rich common-sense knowledge, strong reasoning ability, and powerful code-generation capabilities. Rather than learning control policies purely from data, LLMs offer the possibility of directly perceiving the surrounding world, reasoning about task objectives and generating executable control code for robotic systems. This paradigm opens a promising alternative path for humanoid manipulation that emphasizes generalization, interpretability, and reduced training cost.

Building upon recent advances in code generation and embodied reasoning, this thesis aims to explore a novel humanoid manipulation framework in which vision-language models or large language models directly control a humanoid robot via code generation. This approach is intended as a complementary alternative to current VLA-based methods rather than a replacement. Beyond manipulation itself, such a framework has significant potential for automatic data collection, explainable robotic decision-making, and safe manipulation monitoring, enabling more transparent and adaptable humanoid systems.

Thesis Objectives

1. Develop a humanoid manipulation environment using Isaac Sim and/or MuJoCo.
2. Design and implement basic control and perception APIs that enable LLM-based control of humanoid robots.
3. Implement a demonstration of dual-arm collaborative manipulation via code generation (e.g., object handover or coordinated grasping).
4. (Optional) Incorporate semantic scene understanding using existing foundation models or an automated pipeline, allowing the humanoid to operate across diverse environments.

Student Requirements

- Solid background in Python programming and Linux OS (e.g., Ubuntu).
- Prior experience with MuJoCo and/or Isaac Sim is a plus, but not mandatory.
- Good English communication and technical writing skills.
- Strong academic performance (grade < 2.3).

How to apply

If you are interested in this thesis topic, please feel free to contact me via email at **y.meng(at)tum.de.**
Kindly attach the following documents:

- Curriculum Vitae (CV)
- Academic transcripts
- A brief introduction describing your background, relevant skills, and motivation for this thesis topic

Self-motivated students with a strong interest in robotics, embodied AI, and LLM-based control are highly encouraged to apply. You will have the opportunity to work on cutting-edge research problems and actively contribute to ongoing projects in humanoid manipulation.

Thesis advisor: Yuan Meng (y.meng(at)tum.de), https://www.ce.cit.tum.de/air/people/yuan-meng/