Integration of Machine Learning and Robotic Vision in CoppeliaSim

Kong Zhiwei, Liu Guanhua, Nie Yiding
Huang Shuo, Fan Yuheng, Wang Junlin
Group 3
Southern University of Science and Technology

***Abstract*—This manuscript delineates a robotics simulation project using CoppeliaSim (V-REP) that encompasses the complete spectrum from data acquisition to result presentation. We present a novel methodology employing machine learning to obtain the optimal postures for robotic arms and grippers. This technique minimizes forces and torques while maximizing stability during the grasping process. The approach leverages environmental data of the target object, captured through vision sensors, to enhance the effectiveness of the robotic maneuvers.**

I. Introduction

This project is conducted within the CoppeliaSim (V-REP) environment, an adaptable framework designed for the simulation of intricate robotics scenarios. Our goal is to develop a model capable of utilizing processed images as input to ascertain the appropriate postures for a robotic arm (UR5) and gripper (RG2). This model aims to pinpoint two corresponding grasping points that maximize efficiency.

II. Problem Statement

The core objective of this project is to enhance the efficiency and stability of robotic arms during the grasping process in a simulated environment. We use CoppeliaSim to simulate the conditions and utilize a machine learning model trained on image data to predict optimal grasping points. The expected result is an algorithm that accurately determines arm and gripper positions to maximize efficiency, tested and validated within this simulation environment.

III. Literature Review

The application of machine learning techniques in robotic vision is well-documented. Our approach is inspired by the methodologies used in [2] and [3], which involve processing visual data to improve robotic interactions. Further inspiration is drawn from [1], highlighting the computation efficiency in robotic movements.

IV. Technical Approach

A. Simulation Setup

Our project leverages CoppeliaSim to create a highly controlled simulation environment tailored for the precise evaluation of robotic interactions. This setup includes vision sensors, UR5 robotic arms, RG2 grippers, and a specified target object, all orchestrated to mimic realistic interactions necessary for grasping tasks.

B. Data Acquisition and Processing

Vision sensors are deployed to capture environmental data around the target objects. This data is transmitted to external processing units via port protocols and a remote API. The initial processing steps include image denoising and regularization, which are crucial for preparing the data for subsequent modeling. These procedures are based on methods discussed in previous studies by Kumra and Zeng [2, 3].

C. Model Development and Output

Our methodology integrates a machine learning model designed to optimize a grasping stability function:

$$S = \int\int\int \exp(-F_x)\cdot\exp(-F_y)\cdot\exp(-M)\cdot dF_x\cdot dF_y\cdot dM$$

This function evaluates the efficiency of potential grasping points based on the forces and moments exerted at these points. The model starts with two initially random points and refines them through a gradient descent method to maximize the stability, inspired by techniques outlined in [1].

Once optimal grasping points are identified, CoppeliaSim's tools are used to compute the inverse kinematics for the UR5 and RG2, allowing the system to determine the appropriate postures for interaction with the target object. This reciprocal process ensures that the output from the machine learning model directly informs the physical adjustments required by the robotic arms and grippers.

D. Performance Evaluation

The model's performance will be assessed both quantitatively and qualitatively. Quantitative evaluations will involve observing and analyzing the precision and reliability of the robotic arm's grasping actions. Qualitatively, we will evaluate the stability and efficiency of the grasp during operation. Results will be documented in tables and potentially supported by video evidence to demonstrate the effectiveness of our approach.

V. Intermediate/Preliminary Results

To date, our team has successfully configured the initial simulation environment in CoppeliaSim, which is essential for conducting our experiments. This environment is fully functional, allowing seamless integration with external simulation software through a remote API. In terms of algorithmic development, we have achieved the capability to determine optimal gripping points for objects in two-dimensional scenarios. This foundational work is progressing methodically and adheres to our project timeline.

Preliminary tests indicate that the algorithm is capable of identifying grasp points with reasonable accuracy. However, our evaluations suggest that while the performance of the model shows promise, there are areas needing improvement. Specifically, we aim to enhance the precision and reliability of the gripping point calculations. These adjustments are crucial as we prepare to extend our model to handle more complex, three-dimensional tasks. Continued research and refinement are planned to address these challenges, with the goal of achieving higher fidelity in simulation results and more robust object handling in diverse scenarios.

VI. Conclusion

Integrating machine learning with robotic vision systems has shown potential in improving the efficiency and reliability of robotic arms in simulated environments. Ongoing work will focus on refining these models to enhance their precision and applicability.

Acknowledgments

The authors would like to thank their peers and professors at the Southern University of Science and Technology for their support and guidance.

References

[1] Computation of multi-fingered grasping force with linear combination. https://max.book118.com/html/2019/0115/6101222011002002.shtm, 2019. Accessed: 2024-04-14.

[2] S. Kumra and C. Kanan. Robotic grasp detection using deep convolutional neural networks. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 769–776, Vancouver, BC, Canada, 2017. doi: 10.1109/IROS.2017.8202237.

[3] A. Zeng, K.T. Yu, S. Song, D. Suo, E. Walker, A. Rodriguez, and J. Xiao. Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1386–1383. IEEE, May 2017.