Deep Reinforcement Learning for Robotic Grasping from Octrees

Learning Manipulation from Compact 3D Observations

Andrej Orsula
Robotics

Master's Thesis

AALBORG UNIVERSITY
STUDENT REPORT

Department of Electronic Systems
Aalborg University
https://es.aau.dk

Title:
Deep Reinforcement Learning for Robotic Grasping from Octrees

Theme:
Master's Thesis

Programme:
Robotics

Author:
Andrej Orsula[1]

Supervisor:
Simon Bøgh

Number of Pages:
12

Submission Date:
April 12, 2021

Abstract:

TODO: Abstract

[1]aorsul16@student.aau.dk

Contents

Resumé

TODO: Resumé in Danish

Preface

This Master's Thesis is written by Andrej Orsula as the final work of M.Sc. programme in Robotics at Aalborg University during the academic year 2020/21.

Acknowledgements

Special thanks goes to Simon Bøgh for his supervision, guidance and numerous discussions throughout the whole process that helped to shape this project. Moreover, I must express a very profound gratitude to my mum, dad, sister and brother for their love and everlasting support.

Additional Resources

The primary source code developed during this project is available on the following *GitHub* repository.

- https://github.com/andrejorsula/drl_grasping

All readers interested in reproducing the results from this work are welcome to use pre-built *Docker* images that can be found inside the *Docker Hub* repository below.

- https://hub.docker.com/r/andrejorsula/drl_grasping

This manuscript and additional resources such as raw data acquired during the testing can be accessed using the following *GitHub* repository.

- https://github.com/andrejorsula/master_thesis

Glossary

2D	Two-dimensional
3D	Three-dimensional
MDP	Markov Decision Process
RL	Reinforcement Learning
DRL	Deep Reinforcement Learning
DDPG	Deep Deterministic Policy Gradient
TD3	Twin Delayed Deep Deterministic
SAC	Soft Actor Critic
TQC	Truncated Quantile Critics
CNN	Convolutional Neural Network
TD	Temporal Difference

1 Introduction

2 Related Work

2.1 Robotic Grasping

2.1.1 Empirical Approaches

2.1.2 Learning-Based Approaches

2.2 Learning from 3D

2.2.1 3D Data Representations

Mesh

Point Cloud

Voxel Grid

Octree

3 Background

3.1 Model-Free Reinforcement Learning

3.1.1 Markov Property

3.1.2 Markov Chain

(Sutton and Barto, 2018)

3.1.3 Markov Decision Process

3.1.4 Q-Learning

3.1.5 Value-Based Reinforcement Learning

3.1.6 Policy-Based Reinforcement Learning

3.2 Actor Critic

3.2.1 Deep Deterministic Policy Gradient (DDPG)

3.2.2 Twin Delayed Deep Deterministic (TD3)

3.2.3 Soft Actor Critic (SAC)

3.2.4 Truncated Quantile Critics (TQC)

3.3 Function Approximation

3.3.1 Neural Networks

4 Problem Formulation

4.1 Observation Space

4.1.1 Observation Stacking

4.2 Action Space

4.3 Reward Function

5 Methods

5.1 Curriculum Learning

5.2 Demonstration Bootstrapping

5.3 Domain Randomization

6 Implementation

6.1 Simulation Environment

6.1.1 Selection

MuJoCo

PyBullet

Gazebo Classic

Ignition Gazebo

6.1.2 Simulating with Ignition Gazebo

Controller

Middleware - ROS 2

Motion Planning - MoveIt 2

6.2 OpenAI Gym Environment

6.2.1 Gym-Ignition

6.3 Stable Baselines3

6.4 Network Architecture

6.4.1 PyTorch

6.4.2 Feature Extractor

6.4.3 Actor Critic Networks

6.5 Hyperparameter Optimisation with Optuna

7 Experimental Evaluation

7.1 Experimental Setup

7.2 Results

7.3 Ablation Studies

8 Discussion

9 Conclusion

10 Future Work

Bibliography

Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA.

Appendices

A Low-Level Controller

B Dataset

C Hyperparameters

D Full Results