

Deep Reinforcement Learning for Robotic Grasping from Octrees

Learning Manipulation from Compact 3D Observations

Andrej Orsula
Robotics

Master's Thesis





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

AALBORG UNIVERSITY

STUDENT REPORT

Title:

Deep Reinforcement Learning for
Robotic Grasping from Octrees

Abstract:

TODO: Abstract

Theme:

Master's Thesis

Programme:

Robotics

Author:

Andrej Orsula¹

Supervisor:

Simon Bøgh

Number of Pages:

12

Submission Date:

April 10, 2021

¹aorsul16@student.aau.dk

Contents

Resumé	v
Preface	vi
1 Introduction	1
2 Related Work	2
2.1 Robotic Grasping	2
2.1.1 Empirical Approaches	2
2.1.2 Learning-Based Approaches	2
2.2 Learning from 3D	2
2.2.1 3D Data Representations	2
3 Background	3
3.1 Model-Free Reinforcement Learning	3
3.1.1 Markov Chain	3
3.1.2 Markov Decision Process	3
3.1.3 Q-Learning	3
3.1.4 Value-Based Reinforcement Learning	3
3.1.5 Policy-Based Reinforcement Learning	3
3.2 Actor Critic	3
3.2.1 Deep Deterministic Policy Gradient (DDPG)	3
3.2.2 Twin Delayed Deep Deterministic (TD3)	3
3.2.3 Soft Actor Critic (SAC)	3
3.2.4 Truncated Quantile Critics (TQC)	3
3.3 Function Approximation	3
3.3.1 Neural Networks	3
4 Problem Formulation	4
4.1 Observation Space	4
4.1.1 Observation Stacking	4
4.2 Action Space	4
4.3 Reward Function	4
5 Methods	5
5.1 Curriculum Learning	5
5.2 Demonstration Bootstrapping	5
5.3 Domain Randomization	5
6 Implementation	6

6.1	Simulation Environment	6
6.1.1	Selection	6
6.1.2	Simulating with Ignition Gazebo	6
6.2	OpenAI Gym Environment	6
6.2.1	Gym-Ignition	6
6.3	Stable Baselines3	6
6.4	Network Architecture	6
6.4.1	PyTorch	6
6.4.2	Feature Extractor	6
6.4.3	Actor Critic Networks	6
6.5	Hyperparameter Optimisation with Optuna	6
7	Experimental Evaluation	7
7.1	Experimental Setup	7
7.2	Results	7
7.3	Ablation Studies	7
8	Discussion	8
9	Conclusion	9
10	Future Work	10
	Bibliography	11
	Appendices	12
A	Low-Level Controller	12
B	Dataset	12
C	Hyperparameters	12
D	Full Results	12

Resumé

TODO: Resumé in Danish

Preface

TODO: Preface

Test citation (TODO: remove test citation): ([Name, 2000](#))

Glossary

- 2D** Two-dimensional
- 3D** Three-dimensional

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 Chain

3.1.2 Markov Decision Process

3.1.3 Q-Learning

3.1.4 Value-Based Reinforcement Learning

3.1.5 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

Name, R. (2000). Fascinating title. *Journal of Testing* 420(69), 69–420.

Appendices

A Low-Level Controller

B Dataset

C Hyperparameters

D Full Results