Deep Reinforcement Learning for Robotic Grasping from Octrees

Learning Manipulation from Compact 3D Observations

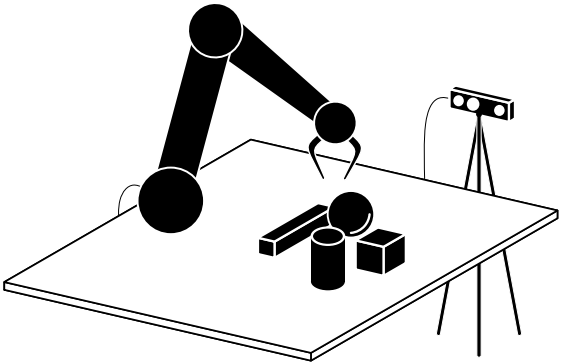
June 25, 2021

Andrej Orsula

Aalborg University
Denmark

AALBORG UNIVERSITY

Vision-Based Robotic Grasping of Diverse Objects



Vision-Based Robotic Grasping of Diverse Objects

Approach

Approaches

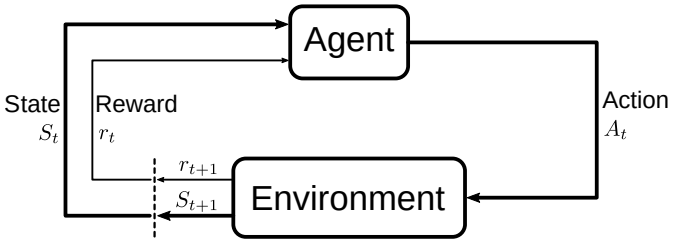
- Analytical
- Empirical
  - Supervised Learning
  - Imitation Learning
  - Reinforcement Learning

Vision-Based Robotic Grasping of Diverse Objects

Approach

Approaches

- Analytical
- Empirical
  - Supervised Learning
  - Imitation Learning
  - **Reinforcement Learning**



Task Definition

Agent

- High-level controller
  - Gripper pose
  - Gripper action

Environment

- Objects
- Robot
  - Low-level controllers
- Physics and visuals

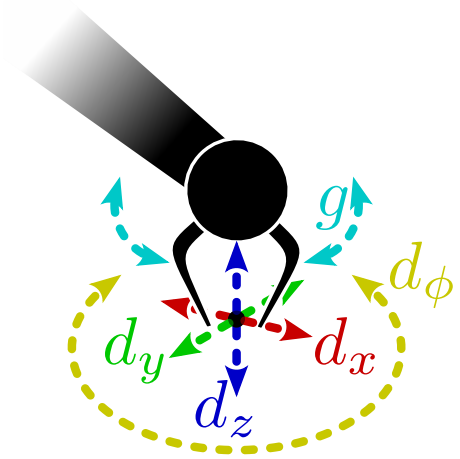
Episodic Task

- Success
  - Lifting an object
- Failure
  - Pushing all objects away
- Max 100 time steps
  - ~40 s (simulation)

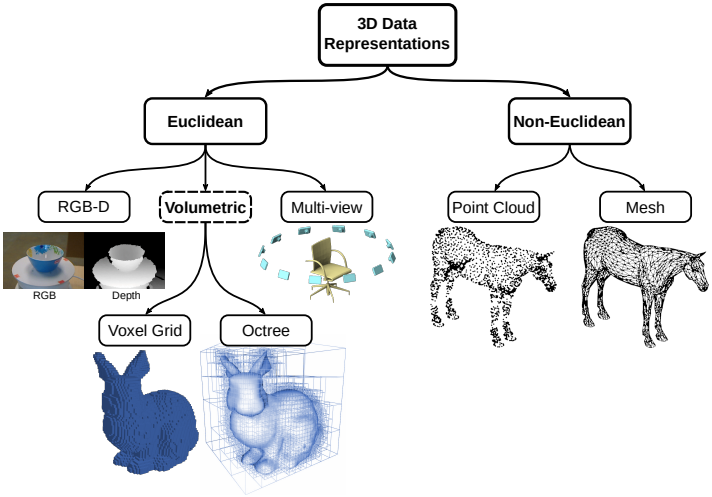
Action Space

Actions in Cartesian space

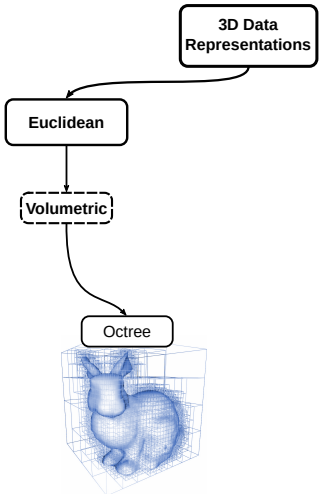
- Translational displacement
  - d_x
  - d_y
  - d_z
- Gripper rotation
  - d_ϕ
- Gripper actions (open/close)
  - g



Observation Space

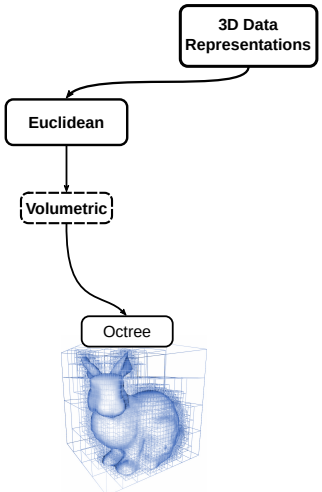


Observation Space



Wang et al. 2017. O-CNN: Octree-based Convolutional Neural Networks for 3D Shape Analysis. ACM Trans. Graph. (SIGGRAPH) 36, 4, Article 72.

Observation Space



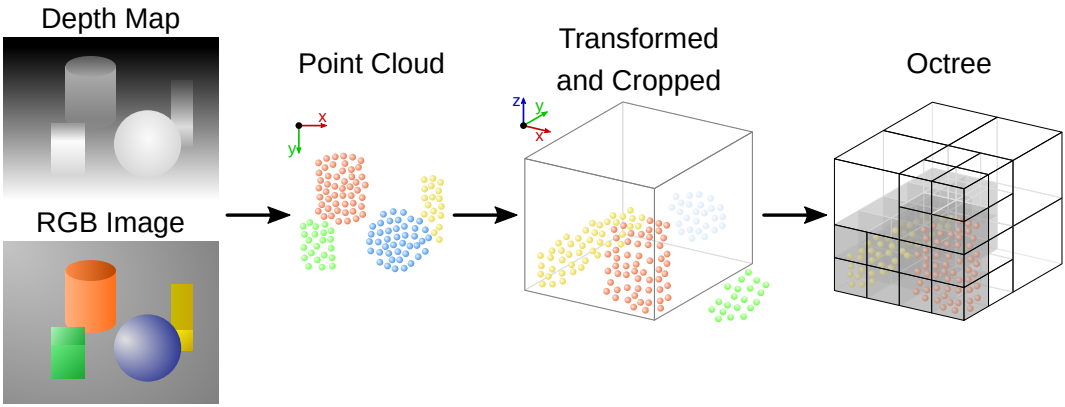
Proprioceptive Observations

- Gripper position
- Gripper rotation
- Gripper state

Wang et al. 2017. O-CNN: Octree-based Convolutional Neural Networks for 3D Shape Analysis. ACM Trans. Graph. (SIGGRAPH) 36, 4, Article 72.

Observation Space

Construction of Octree



Observation Space

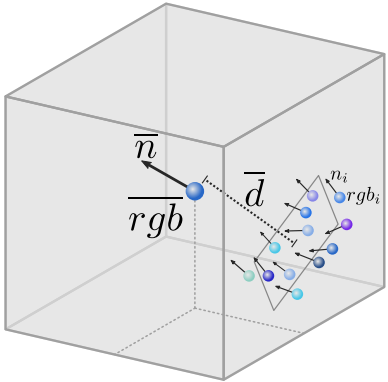
Features and Stacks

Features

- Spatial
  - Average normal vector $\overline{n}$
  - Average distance to points $\overline{d}$
- Colour
  - Average intensity of RGB channels $\overline{rgb}$

Observation Stacking

- Three consecutive stacks



Reward Function

Composite Reward

- Reach
  - $+1$ (7^0)
- Touch
  - $+7$ (7^1)
- Grasp
  - $+49$ (7^2)
- Lift
  - $+343$ (7^3)

Recurring Reward

- Collision with ground/table
  - -1
- Incentive to act quickly
  - -0.005

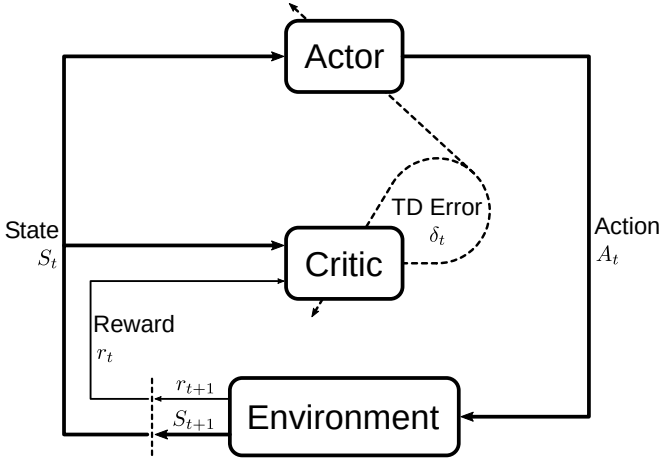
Reinforcement Learning
Algorithms

Actor-Critic Algorithms

- TD3
- SAC
- TQC

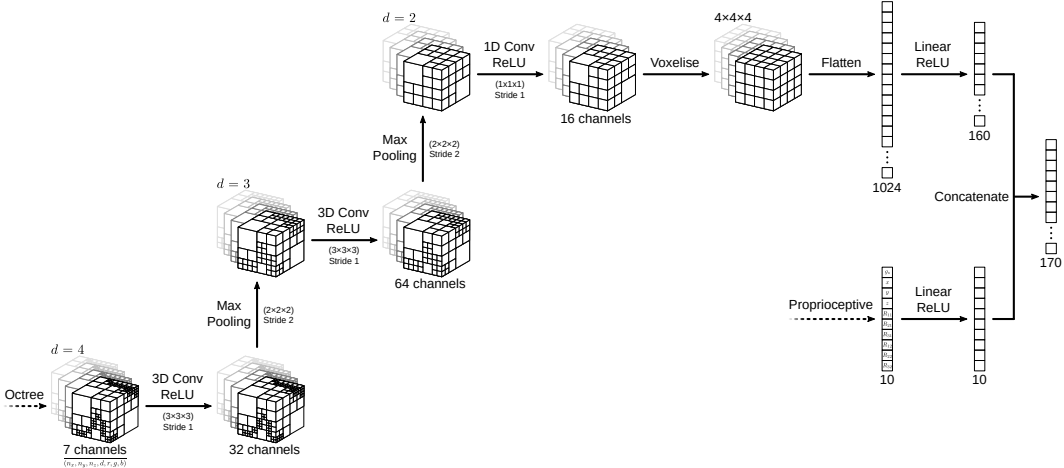
Implementation

- Stable Baselines3



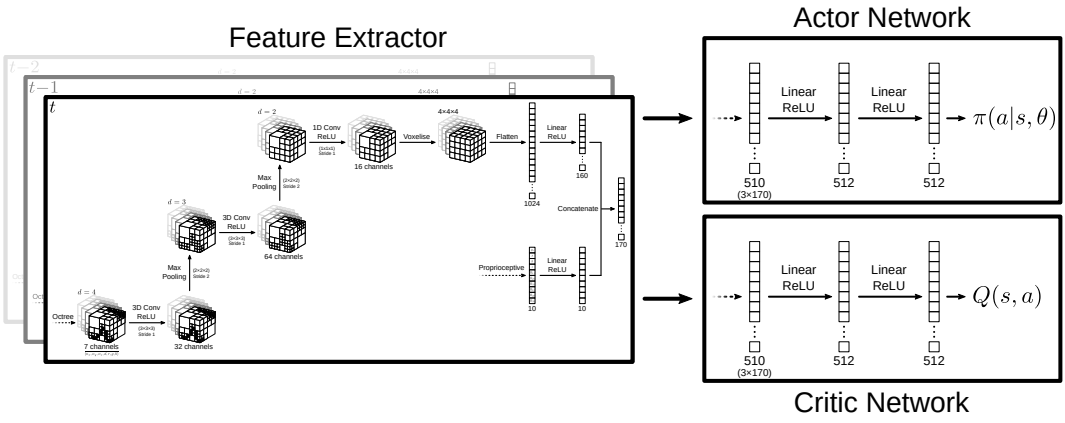
Deep Reinforcement Learning

Octree-Based Feature Extractor



Deep Reinforcement Learning

Full Actor-Critic Network Architecture



Thank you for your time

AALBORG UNIVERSITY