ICRA 2019

Robust Object Grasping in Clutter via Singulation

Marios Kiatos[1,2], Sotiris Malassiotis[1]

[1]Information Technologies Institute(ITI), Center of Research and Technology, Thessaloniki, Greece

[2]Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece

Introduction

- Grasping in cluttered environments is challenging due to collision free grasp affordances.
- We propose a pushing strategy that **singulates** a target object from its surrounding clutter by means of lateral pushing movements.
- We employ **reinforcement learning** in order to obtain optimal push policies.

Problem Formulation

We formulate the problem as a **Markov Decision Process**.

States

- The state is approximated by a feature vector that describes the topography of the given scene.
- Depth features are extracted from the scene heightmap in the neighborhood of the target object(Fig. 1).
- To avoid pushing the target near to bin walls, we add the position of the target to the state.

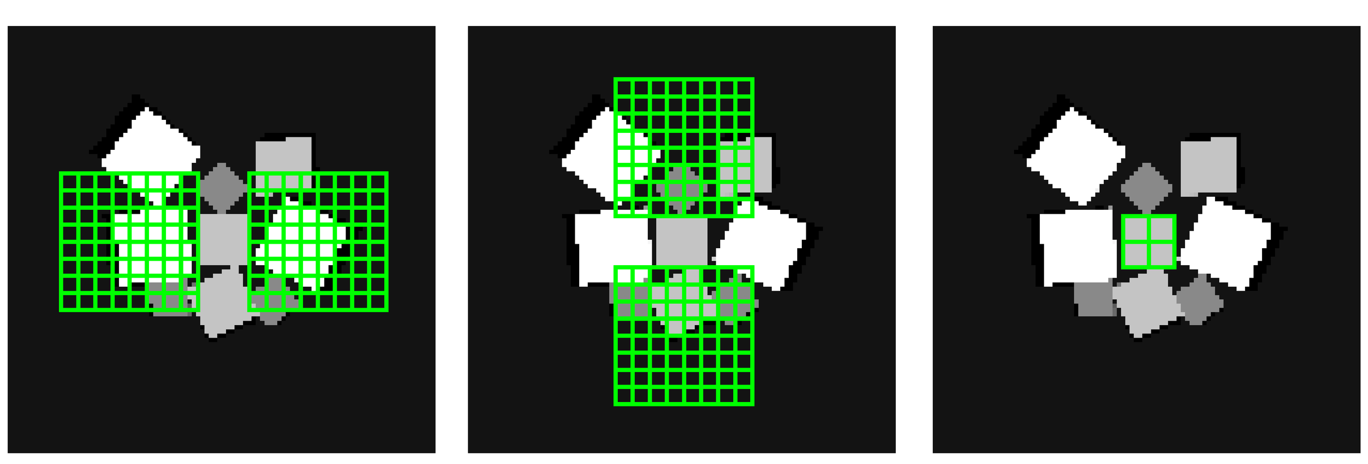
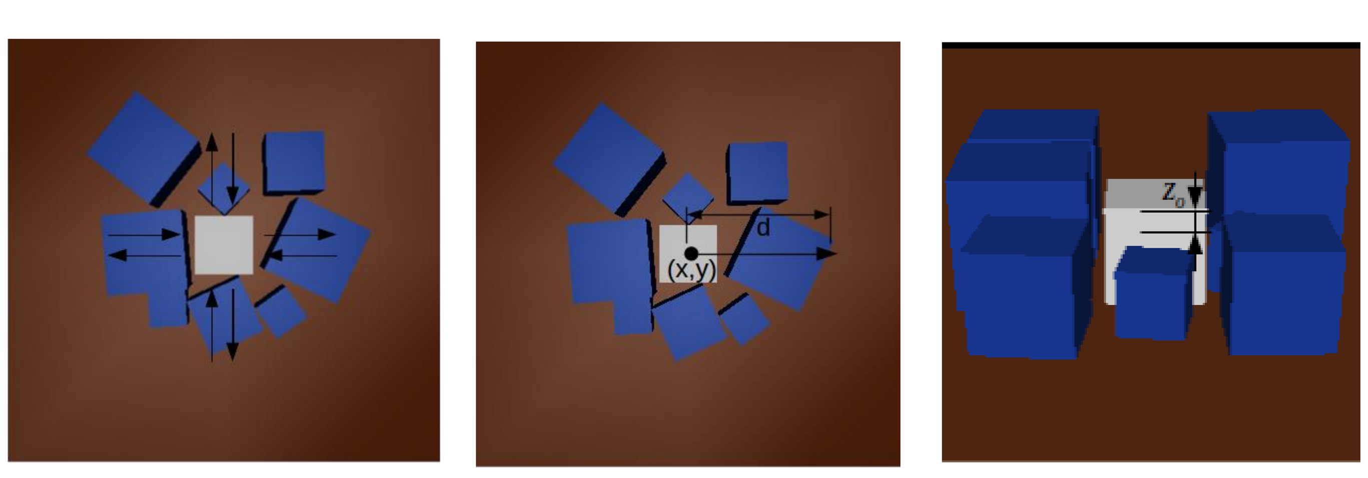


Figure 1: The regions of heightmap where the depth features are computed.

Actions

- The action space consists of pushes in different directions and heights around the target object.
- Each push can be parametrized by:
  - the initial push point,
  - the push distance and
  - the direction of push.



(a) (b) (c)

Figure 2: (a)Actions w.r.t. the target, (b), (c) the initial push point and the push distance

Rewards

- A sparse reward function is defined.
- We penalize the total number of pushes.
- The target is free if it is separated from the closest object by a minimum distance of 3cm.

Transition matrix

- Model-free problem.
- The robot learns the transition dynamics from trial and error.

Learning

- Objective: maximize the expected reward.
- Episodic task.
- The RL agent is trained with the **Deep Q-Learning** algorithm.
  - A shallow deep network is used as function approximator.
  - A replay buffer and a second target network are used to improve the network's convergence.
  - An ϵ-*greedy* exploration strategy is adopted.

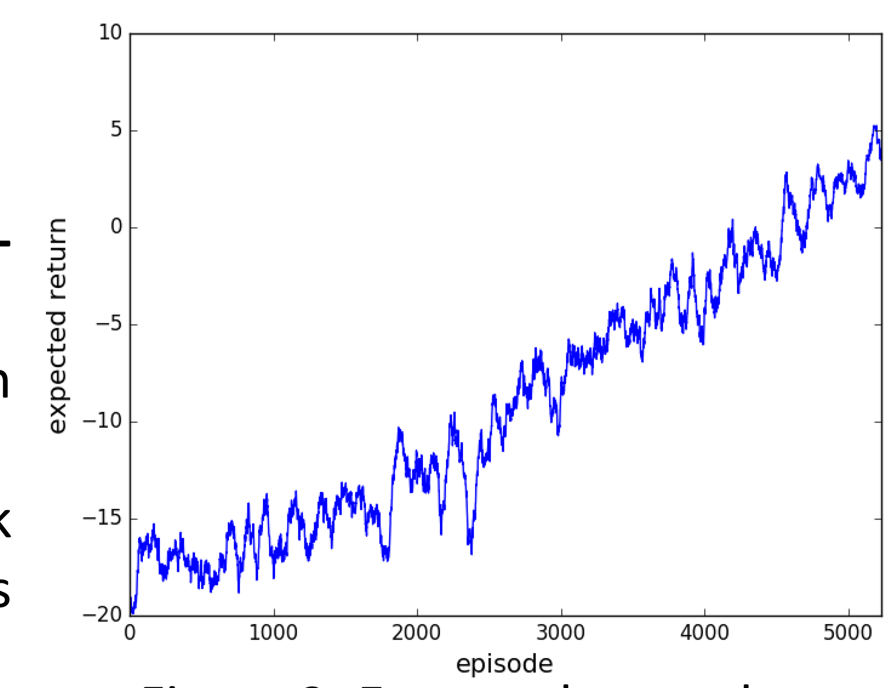


Figure 3: Expected reward per episode.

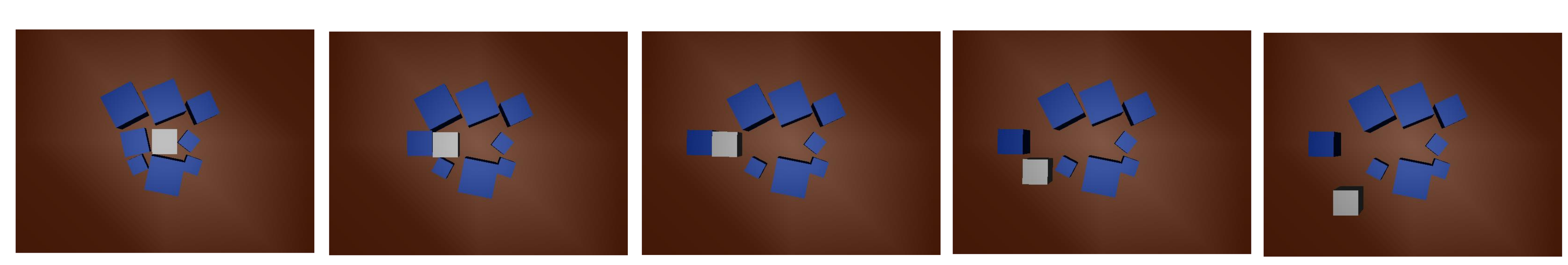
Experiments

Simulation

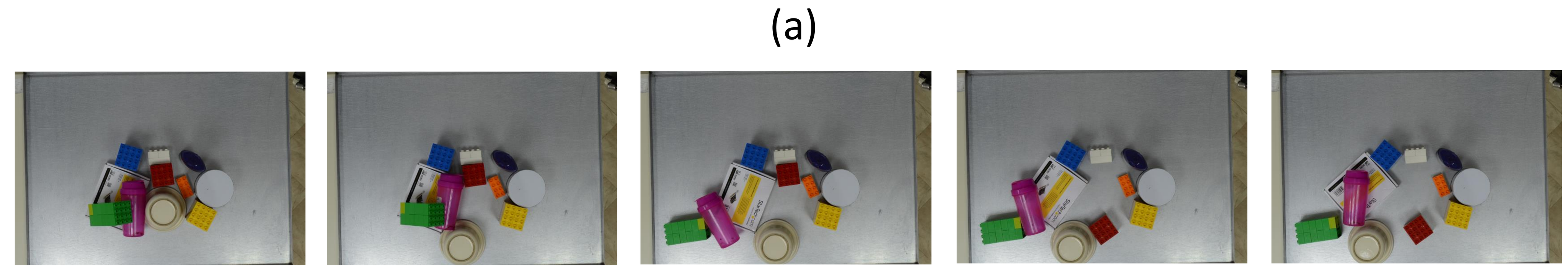
- The RL agent is trained and evaluated in MuJoCo physics simulator.
- We approximated the objects with cubes of random dimensions
- The RL agent is trained for 5000 episodes.

Robotic

- We transfer the learned policy in a real world scenario.
- The object set is composed of 10 different objects.



(a)



(b)

Figure 4: Pushing sequence that leads to the successful singulation of the (a) white cube in simulation and (b) red cube in a real world scenario.

Environment	Singulation success	Average pushes
Simulation	97%	3.35
Robotic	75%	4

Table 1: Experiment results. The difference in singulation success is due to different physics in simulation and real world.

Acknowledgements: The research leading to these results has received funding from the European Community's Framework Programme Horizon 2020-under grant agreement No 820767 – CoLLaboratE.