

QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation

arXiv:1806.10293, Kalashnikov et al, 2018.

Summarized by Hyecheol (Jerry) Jang

Department of Computer Sciences
University of Wisconsin–Madison

RL Paper Study, Jun. 29. 2020





1 Motivation

2 Goal

Motivation: Why Robotics + Reinforcement Learning

- Usually, Robots are good at **repetitive tasks** (e.g. Assembly Line)

Motivation: Why Robotics + Reinforcement Learning

- Usually, Robots are good at **repetitive tasks** (e.g. Assembly Line)
- Want to make Robots that **identifies surroundings** and **behave accordingly**, but it is difficult

Motivation: Why Robotics + Reinforcement Learning

- Usually, Robots are good at **repetitive tasks** (e.g. Assembly Line)
- Want to make Robots that **identifies surroundings** and **behave accordingly**, but it is difficult
 - **Deep Learning**
Provide ability to handling real-world scenarios
 - **Reinforcement Learning**
Provide ability to make decision in long-term, using previous experiences in complex and robust scenarios

Motivation: Why Robotics + Reinforcement Learning

- Usually, Robots are good at **repetitive tasks** (e.g. Assembly Line)
- Want to make Robots that **identifies surroundings** and **behave accordingly**, but it is difficult
 - **Deep Learning**
Provide ability to handling real-world scenarios
 - **Reinforcement Learning**
Provide ability to make decision in long-term, using previous experiences in complex and robust scenarios
- Combining two techniques
 - Able to learn policy continuously from their experience
 - No need for manual engineering, use data they collects



- Variance in **visual and physical property of objects**



- Variance in **visual and physical property of objects**
 - Hardness of object (Soft or Hard)
 - Surface Characteristics (Slippery, Sticky, ...)
 - Color Variation
 - Shape Variation
 - ...



- Variance in **visual and physical property of objects**
 - Hardness of object (Soft or Hard)
 - Surface Characteristics (Slippery, Sticky, ...)
 - Color Variation
 - Shape Variation
 - ...
- **Noise** of sensors

- Variance in **visual and physical property of objects**
 - Hardness of object (Soft or Hard)
 - Surface Characteristics (Slippery, Sticky, ...)
 - Color Variation
 - Shape Variation
 - ...
- **Noise** of sensors
 - ⇒ Still hard to handle though we have sufficiently large training set
 - ⇒ Collecting those training set is expensive (real experiments)

- Variance in **visual and physical property of objects**
 - Hardness of object (Soft or Hard)
 - Surface Characteristics (Slippery, Sticky, ...)
 - Color Variation
 - Shape Variation
 - ...
- **Noise** of sensors
 - ⇒ Still hard to handle though we have sufficiently large training set
 - ⇒ Collecting those training set is expensive (real experiments)
 - ⇒ Lots of researchers focused on reusing previous experiences



- Focused on learning narrow, individual tasks
 - hitting a ball
 - opening door
 - throwing objects
 - ...



- Focused on learning narrow, individual tasks
 - hitting a ball
 - opening door
 - throwing objects
 - ...
- ⇒ Use **Grasping** to achieve *generalization*

- Focused on learning narrow, individual tasks
 - hitting a ball
 - opening door
 - throwing objects
 - ...
- ⇒ Use **Grasping** to achieve *generalization*
- Approached the grasping task as predicting a *grasp pose*
 - 1 Observe the scene (*Normally, using a depth camera*)
 - 2 Choose best location to grasp
 - 3 Reach the location (open-loop setting)

- Focused on learning narrow, individual tasks

- hitting a ball
- opening door
- throwing objects
- ...

⇒ Use **Grasping** to achieve *generalization*

- Approached the grasping task as predicting a *grasp pose*

- 1 Observe the scene (*Normally, using a depth camera*)
 - 2 Choose best location to grasp
 - 3 Reach the location (open-loop setting)
- Different with how humans and animals behave
 - Grasp is a **dynamical process** that sense and control at each stage

- Focused on learning narrow, individual tasks

- hitting a ball
- opening door
- throwing objects
- ...

⇒ Use **Grasping** to achieve *generalization*

- Approached the grasping task as predicting a *grasp pose*

- 1 Observe the scene (*Normally, using a depth camera*)
- 2 Choose best location to grasp
- 3 Reach the location (open-loop setting)
 - Different with how humans and animals behave
 - Grasp is a **dynamical process** that sense and control at each stage

⇒ **Where this researches start!!**



1 Motivation

2 Goal

Use Reinforcement Learning with Deep Neural Network
to **perform pre-grasp manipulation**,
response to dynamic disturbances,
and **learn grasping in a generic framework**
that makes minimal assumptions about the task



- **Closed-loop condition** (With feedback, *Morrison, et al.*)
 - For the other papers work on closed-loop grasping, they deals with servoing problems.
 - This paper focuses on making generalized RL algorithm
 - In practice, it makes Kalashnikov et al.'s method (this method) to autonomously acquire complicated grasping strategy



- **Closed-loop condition** (With feedback, *Morrison, et al.*)
 - For the other papers work on closed-loop grasping, they deals with servoing problems.
 - This paper focuses on making generalized RL algorithm
 - In practice, it makes Kalashnikov et al.'s method (this method) to autonomously acquire complicated grasping strategy
- **Self-supervised** learning task
 - Compare to prevoius work(by Zeng et al.), Kalashnikov et al. utilize more general action space
 - Actions consist of end-effector **Cartesian motion** and **gripper opening/closing**

- **Closed-loop condition** (With feedback, *Morrison, et al.*)
 - For the other papers work on closed-loop grasping, they deal with servoing problems.
 - This paper focuses on making a generalized RL algorithm
 - In practice, it makes Kalashnikov et al.'s method (this method) to autonomously acquire a complicated grasping strategy
- **Self-supervised** learning task
 - Compare to previous work (by Zeng et al.), Kalashnikov et al. utilize more general action space
 - Actions consist of end-effector **Cartesian motion** and **gripper opening/closing**
- Observation comes from **a single RGB camera** over the shoulder
 - Many current grasping systems utilize depth sensing
 - Using wrist-mounted cameras