Parallelized deep reinforcement learning for robotic manipulation

Master thesis description

Isac Arnekvist
isacar@kth.se
Supervisor: Johannes A. Stork

January 16, 2017

1 Background

Recently a number of papers were published were robots learn manipulation tasks through shared experiences [1–4]. It would be interesting to see whether these methods can be used and altered to generalize for other contexts such as changing the type of manipulators and performing different tasks. The use cases are for robotic manipulation with camera as feedback where exact relative positions of objects, manipulators, and sensors need not be fixed, also where resources exists to use several robots for speeding up the learning process. The experiments needed for this thesis would be held at the Robotics, Perception, and Learning lab at KTH.

This study would mainly require choice, implementation and scientific evaluation of parallelized machine learning algorithms, coding for robotic manipulation, and ground truth generation. The machine learning algorithms would include both supervised learning for computer vision and reinforcement learning for performing tasks. This makes the project not only suitable for a computer science degree project, but also for the master program in machine learning.

1.1 Research question

How can deep reinforcement learning be used for performing manipulation where multiple robots are used in parallel for training?

2 Method

A manipulation task would be set up with multiple cameras, computers, and robotic arms. Building on the ideas mainly from the articles [1–4] I would try out this with robots that are different in the way their motion is controlled (torque vs. cartesian). The algoritms for controlling the arm would be end-to-end, i.e. that the entire pipeline from camera data to control signals to the arm would be learned.

3 Limitations

A first part of the work would to pre-train a part of the neural network for pose estimation (and training data generation). If this turns out to need a lot of time, the method for accomplishing the manipulation task might be chosen to better match the remaining time.

4 Students background

I have a bachelor's degree in computer science and finishing the machine learning master program. I have taken both first cycle and second cycle statistics courses, several machine learning courses including computer vision and several kinds of supervised- and

unsupervised learning. I have experience with implementing convolutional and regular neural networks from courses and from working. I have also recently finished a course in robotics and autonomous systems. I have no courses remaining to be finished except those that are planned until the start of the thesis project.

The courses that are ongoing and required for graduating are:

- Introduction to the Philosophy of Science and Research Methodology (DA2205)
- Program Integrating Course in Machine Learning (DD2301)
- Robotics and Autonomous Systems (DD2425)
- Statistical Methods in Applied Computer Science (DD2447)

References

[1] Yevgen Chebotar, Mrinal Kalakrishnan, Ali Yahya, Adrian Li, Stefan Schaal, and Sergey Levine. Path integral guided policy search. *arXiv preprint arXiv:1610.00529*, 2016.

[2] Chelsea Finn and Sergey Levine. Deep visual foresight for planning robot motion. *arXiv preprint arXiv:1610.00696*, 2016.

[3] Shixiang Gu, Ethan Holly, Timothy Lillicrap, and Sergey Levine. Deep reinforcement learning for robotic manipulation. *arXiv preprint arXiv:1610.00633*, 2016.

[4] Ali Yahya, Adrian Li, Mrinal Kalakrishnan, Yevgen Chebotar, and Sergey Levine. Collective robot reinforcement learning with distributed asynchronous guided policy search. *arXiv preprint arXiv:1610.00673*, 2016.