

Planning Report

Reinforcement Learning for an Object Search Task*

Oskar Lundin
<osklu414@student.liu.se>
070-756 80 58

February 7, 2022

1 Description

Object search is a perceptual task in which an environment is searched for a particular target object. In this project, an instance of this task will be considered where a fixed pan-tilt-zoom camera is to locate one or more target objects in its surroundings.

As the task involves learning a correlation between observations of the environment and actions that control the camera, (deep) reinforcement learning (RL) [9] will be used as the learning paradigm.

There are several aspects to the problem. There might be distractors in the environment that are not targets and have to be ignored. These should be visually distinct from the target objects. The target objects might also be partially hidden behind distractors. Furthermore, as the entire environment will not be visible at once, the agent should employ an effective scanning pattern. For this it has to learn where the probability of finding target objects is high. If there are multiple targets in the scene, the agent will need to indicate that it thinks a target has been found and then move on to the next object. Environments might be vary slightly in appearance and target objects are placed at random locations.

The following research questions will be answered in this thesis project:

1. What behaviours are desirable during object search, and how can they be tested?
2. How can object search be solved under an RL framework? Getting an agent to exhibit desirable searching behavior is not expected to be trivial. For example, certain aspects of the task might put requirements on the algorithm used.

*Preliminary title.

3. How do different methods compare in terms of search time and localization accuracy? Things that can be compared include RL algorithm, observation space, action space, reward signal, etc.

These questions can likely be split into parts, but exactly how depends on the evolution of the project.

2 Approach

The plan is to first do a thorough study of available literature and plan out the approach. This includes finding and experimenting with available RL algorithms, planning the task setup and determining what tools and frameworks to use, etc.

If the task is immediately tackled in its entirety, there is some risk that issues will be difficult to isolate. For example, there might be some error in the implementation, the problem setup might be too difficult for an agent to learn to solve, some hyperparameter might be crucial, etc. For this reason, the task will be approached in an incremental manner. First, a grid world-like environment will be implemented where aspects of the object search problem will be added incrementally. The hope is that this will lead to an increased practical understanding of what is required to solve the problem.

Once the problem has been solved and understood in the grid world environment, there are a couple of options. If the results are promising, the task could be tried in more visually-rich environment. FOI has an in-house simulator for outdoor 3D environments. The goal is to use this for the object search task, but it is quite resource demanding and a training run can take multiple days.

Another option is to ignore high-dimensional inputs entirely, and focus on the behavior of the agent. The final vision is to have an agent that is capable of surveilling an area and keep track of multiple targets. For this, some kind of memory mechanism is required. What happens if the targets start moving?

3 Literature

The object search is sometimes also referred to as visual search and is considered to be part of active vision. Some recent developments of this are surveyed in [2].

Minut and Mahadevan (2001) make an early attempt to imitate gaze patterns in humans with RL and a pan-tilt-zoom camera [8]. In a more recent work by Wu et al. (2020), an agent is trained to both move and look around in a scene to find a target described by a word [10] in a simulated in-door environment. Schmid et al. (2019) train an RL agent to approach target objects in and report when they have been found using real-world images collected by a robot. Druon et al. (2020) transform visual information into an intermediate representation in the form of a grid with high-level semantic information, and use this for object search [3]. This intermediate representation is almost analogous to a grid world,

and the method is one possible way to move from a grid world to a more visually rich task environment.

Ghesu et al. use RL to locate landmarks in medical images and achieve state of the art results [5, 6]. Although medical images such as CT scans exhibit little variance and might not be directly comparable to more

Many aspects of the object search problem can be found in other research areas. The problem of learning effective scan patterns is closely related to coverage path planning [4]. In computer vision, object detection is the task of classifying and determining the bounding box of a visible object. Some works train an agent to deform input images with a series of linear transformations to select an accurate bounding box for an object [1, 7]. Translating and scaling a bounding box is similar to panning/tilting and zooming with a camera.

4 Time Plan

The time plan is presented in Table 1. This is slightly shifted, so that week 1 starts at 2022-01-31.

References

- [1] Juan C. Caicedo and Svetlana Lazebnik. Active object localization with deep reinforcement learning. *arXiv:1511.06015 [cs]*, Nov 2015. arXiv: 1511.06015.
- [2] Shengyong Chen, Youfu Li, and Ngai Ming Kwok. Active vision in robotic systems: A survey of recent developments. *The International Journal of Robotics Research*, 30(11):1343–1377, Sep 2011.
- [3] Raphael Druon, Yusuke Yoshiyasu, Asako Kanezaki, and Alassane Watt. Visual object search by learning spatial context. *IEEE Robotics and Automation Letters*, 5(2):1279–1286, Apr 2020.
- [4] Enric Galceran and Marc Carreras. A survey on coverage path planning for robotics. *Robotics and Autonomous Systems*, 61(12):1258–1276, Dec 2013.
- [5] Florin C. Ghesu, Bogdan Georgescu, Tommaso Mansi, Dominik Neumann, Joachim Hornegger, and Dorin Comaniciu. *An Artificial Agent for Anatomical Landmark Detection in Medical Images*, volume 9902 of *Lecture Notes in Computer Science*, page 229–237. Springer International Publishing, 2016.
- [6] Florin-Cristian Ghesu, Bogdan Georgescu, Yefeng Zheng, Sasa Grbic, Andreas Maier, Joachim Hornegger, and Dorin Comaniciu. Multi-scale deep reinforcement learning for real-time 3d-landmark detection in ct scans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):176–189, Jan 2019.

Table 1: Planned activities and milestones of the work with the resolution of a week.

| Week | Activity |
|-------|---|
| 1 | Write planning report. Discuss ideas. |
| 2 | Get set up in office. Research and evaluate ideas, familiarize with hardware and software. |
| 3 | Implement grid world environment and RL algorithms. Run simple experiments to verify that algorithms work. |
| 4 | Change distribution of environment |
| 5 | |
| 6 | Milestone: Agent trained. Should now have some idea of what is doable in the time available. Additional investigations determined. |
| 7 | |
| 8 | |
| 9 | |
| 10 | Spend some time on report. Make sure everything done so far is written down clearly. |
| 11 | Milestone: Mid-thesis review. Grid world phase complete. Introduction, Theory and Method chapters written. |
| 12-15 | Buffer time. Unlikely that everything went smoothly, work on things that need more time. Otherwise, perform additional investigations. |
| 16 | Milestone: Results collected. Results chapter written. |
| 17 | Evaluate and interpret results. Collect thoughts. Make sure that research questions have been addressed. |
| 18 | Write discussion and conclusion. |
| 19 | Fix typographical errors in report, verify that requirements are satisfied. Submit draft report. |
| 20 | Prepare presentation. Polish report according to feedback. |
| 21 | Milestone: Final submission. Submit report. Final presentation scheduling and preparation. |

- [7] Zequn Jie, Xiaodan Liang, Jiashi Feng, Xiaojie Jin, Wen Feng Lu, and Shuicheng Yan. Tree-structured reinforcement learning for sequential object localization. *arXiv:1703.02710 [cs]*, Mar 2017. arXiv: 1703.02710.
- [8] Silviu Minut and Sridhar Mahadevan. A reinforcement learning model of selective visual attention. In *Proceedings of the fifth international conference on Autonomous agents - AGENTS '01*, page 457–464. ACM Press, 2001.
- [9] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. Adaptive computation and machine learning series. The MIT Press, second edition edition, 2018.
- [10] Jie Wu, Tianshui Chen, Lishan Huang, Hefeng Wu, Guanbin Li, Ling Tian, and Liang Lin. Active object search. In *Proceedings of the 28th ACM International Conference on Multimedia*, page 973–981. ACM, Oct 2020.