Planning Report
Reinforcement Learning for a Visual Search Task*

Oskar Lundin

\<osklu414@student.liu.se\>

070-756 80 58

March 2, 2022

1 Description

Visual 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.

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

1. How can this instance of the visual search problem be solved with RL? The focus here will be on observation space, action space and reward signal.

2. How do different RL algorithms compare for the problem? Focus on classes of algorithm, feature extraction architectures, etc.

3. How does an RL method compare to some baseline method? Here the method will be compared to one or more non-RL agents (rule-based, human controlled, random actions, etc.).

---

*Preliminary title.

2 Approach

The thesis project will be split into three phases. In the first phase, the problem will be formulated and properly restricted. A simple environment that tests the desired agent characteristics will be implemented and used for rapid experimentation and an initial comparison of methodologies. These tests can also help identify what kinds of results are reasonable. The plan is to do the main experiments in a simulator developed by FOI that simulates realistic outdoor environments. However, initial tests with the realistic environment show that training an agent for a sufficient number of time steps can take multiple days. For this reason, the simpler environment will be used to test ideas and understand the problem. If it is not too cumbersome, the realistic environment will then be used for the remaining phases. A backup solution is to flesh out the simpler environment and use that for the remainder of the thesis. At the end of this phase, the problem to be solved should be completely specified, some initial methods should have been found, and the simulator should be decided. The time allocated for this is approximately 5 weeks.

In the second phase, the methodology will be developed (first and second research question). This will be the bulk of the work, approximately 10 weeks. Various algorithms from relevant literature will be implemented and compared. Some of this literature is covered in the next section. After this phase, the method should be decided and an agent should be trained (possibly some variations). Finally, in the third phase, the best RL agent will be evaluated in depth and compared to some baseline methods (second research question). The current idea is to implement a rule-based agent using domain knowledge. The allocated time for this phase is approximately 5 weeks.

3 Literature

The considered instance of the object search problem is sometimes also referred to as object 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. Week 1 starts at 2022-01-31, and week 20 ends at 2022-06-19.

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.

[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.

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

Week	Activity
1	Discuss project and expectations. Write planning report. Collect relevant literature.
2	Get set up in office. Research and evaluate ideas, familiarize with hardware and software. Start working on simple environment.
3	Experiment in simple environment. Test various aspects of problem. Train agents that solve simplified instances of problem. Write part of introduction chapter of thesis report.
4	Continue experiments. Do some tests with realistic simulator. Write part of theory chapter of thesis report.
5	Consider what results and task complexity seems reasonable. Discuss with supervisors. Decide environment and finish task implementation.
6	**Milestone: Phase 1.** Task well-defined. Environment for task decided and ready for use. Update report accordingly. Start looking for suitable opponent.
7	Go back to collected literature. Check which methods and approaches are promising for the now fully specified task. Look for more literature.
8	Implement most promising methods. Experiment with them. Relate back to tests in simpler environment.
9	Same as last week. Hopefully some method shows promise. Set up a good way to track experiments and configurations.
10	Spend some time on report. Make sure everything done so far is written down clearly. Schedule meeting with examiner.
11	**Milestone: Mid-thesis review.** Grid world phase complete. Introduction, Theory and Method chapters written.
12	Spend some time on figuring out how the results should be presented. Focus on replication, illustrating performance.
13	Hyperparameter tuning can be important for some algorithms. Spend some time on this for most promising algorithms.
14	**Milestone: Phase 2.** One or multiple RL methodologies determined. Results for these are collected, and can be replicated.
15	Describe method in report. Plan baseline approaches.
16	Implement baselines and collect results for them.
17	Finish results chapter of report. See over report so far. Make sure research questions have been addressed. Think about discussion and conclusions.
18	Draft discussion and conclusion chapters.
19	Fix typographical errors in report, verify that requirements are met. Submit draft report.
20	Prepare presentation. Polish report according to feedback. Schedule final presentation.
21	**Milestone: Final submission**. Submit report. Finish presentation.

[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.