Reinforcement Learning for Sensorless Astronomical Imaging

Juri Morisse, Ioannis Koutalios, Rico Landman⋆, Thomas Moerland⋆

Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands
⋆ supervisor

1 Introduction

Adaptive optics (AO) is a technique used to correct for disturbances in optical systems, by adjusting a set of deformable mirrors (DM). Controlling a telescope's DM is a high-dimensional decision problem that follows a constant loop of adjusting mirrors and checking the adjustment's effect on the acquired image. Due to the sequential nature of this process and a clearly defined goal measure (reducing the noise), prior work increasingly focused on applying reinforcement learning (RL) to AO control [1,2,3,4,5,6].

In this work, we demonstrate that RL can also be used for sensor-free AO control in the context of astronomical imaging. We argue that this approach is advantageous, as it reduces the overall system complexity by removing the sensor component and with it any potential noise or bias such a sensor can introduce.

2 Methodology

The number of mirrors in an AO can be high, thus creating a high-dimensional space. Atmospheric distortions can be described using a set of Zernike polynomials [7]. We can therefore use the same polynomials to shape the set of DM and correct aberrations. This way we are able to naturally capture the action space dimensionality (to the number of used polynomials).

We use Soft Actor-Critic (SAC), an off-policy RL algorithm suited for continuous action spaces. For our application, states are images, actions are adjustments of the DMs, and the reward captures image sharpness or centering. Our implementation and the full paper can be accessed via GitHub*.

The environment we use has the focal image of the telescope (48x48 pixels) as the observation space, and the reward function is a measure of image sharpness. The action space is continuous with the dimensionality depending on the number of Zernike modes used. To simplify the problem, atmosphere distortions are filtered so that they can always be described using the same number of Zernike modes that define the action dimensionality. This makes the environment harder to solve as we increase the number of Zernike modes used. Higher degrees of the polynomials produce smaller distortions, which means that after a certain point,

*https://github.com/johnkou97/AdaptiveOptics

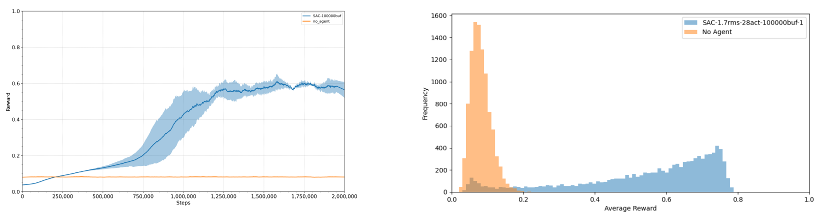

Fig. 1. Performance of the agent in the environment using 27 Zernike modes. The "no-agent" baseline is the performance of the environment without any agent controlling the deformable mirror and is represented by the orange line. Left panel: Learning Curve during training. Right panel: Evaluation Performance of the agent. The x-axis is the average reward over each episode and the y-axis is the frequency of each reward.

further increasing the number of modes will not make any significant difference to the environment and the filtering will no longer play a crucial role.

3 Results

We tested our agent in different environment versions, using a different number of Zernike modes. Each time we increase the number of modes, the environment gets harder to solve. In fig. 1 we show the learning curve and evaluation for our agent in the "hardest" version of the environment we were able to train at successfully. For the 27 Zernike modes environment, we see that our agent is able to learn and reaches a plateau around 0.6 in the learning curve. The evaluation shows that higher performance can be achieved, there are however episodes with worse performance. The agent was also tested in the environment with a smaller number of modes and was always successful in learning a policy with an average return above 0.6. Above 27 Zernike modes, we were unable to see any, probably due to limitations in computational resources (e.g., buffer size and training time).

4 Discussion & Conclusion

This work serves as a proof-of-concept for the suitability of RL for controlling sensorless AO systems. Using an environment with controlled action dimensionality and noise, we iteratively increased complexity and achieved success for up to 27 Zernike modes. While this does not match the full complexity of real-world AO systems, it proves that RL can, in principle, effectively control sensorless AO.

We see our work as a promising first step towards RL-based AO control. Future work could build on this and, for example, explore a curriculum learning approach where an agent starts learning a simple task and gradually evolves towards solving the full task. Additionally, we see value in comparing our method to that of a sensor-based approach or testing how it performs on a real-world telescope.

References

1. Durech, E., Newberry, W., Franke, J., Sarunic, M.V.: Wavefront sensor-less adaptive optics using deep reinforcement learning. Biomed. Opt. Express **12**(9), 5423–5438 (Sep 2021). `https://doi.org/10.1364/BOE.427970`, `https://opg.optica.org/boe/abstract.cfm?URI=boe-12-9-5423`
2. Landman, R., Haffert, S.Y., Radhakrishnan, V.M., Keller, C.U.: Self-optimizing adaptive optics control with reinforcement learning for high-contrast imaging. Journal of Astronomical Telescopes, Instruments, and Systems **7**(3), 039002–039002 (2021)
3. Nousiainen, J., Rajani, C., Kasper, M., Helin, T.: Adaptive optics control using model-based reinforcement learning. Optics Express **29**(10), 15327–15344 (2021)
4. Parvizi, P., Zou, R., Bellinger, C., Cheriton, R., Spinello, D.: Reinforcement learning-based wavefront sensorless adaptive optics approaches for satellite-to-ground laser communication (2023)
5. Parvizi, P., Zou, R., Bellinger, C., Cheriton, R., Spinello, D.: Reinforcement learning environment for wavefront sensorless adaptive optics in single-mode fiber coupled optical satellite communications downlinks. Photonics **10**(12) (2023). `https://doi.org/10.3390/photonics10121371`, `https://www.mdpi.com/2304-6732/10/12/1371`
6. Pou, B., Ferreira, F., Quinones, E., Gratadour, D., Martin, M.: Adaptive optics control with multi-agent model-free reinforcement learning. Opt. Express **30**(2), 2991–3015 (Jan 2022). `https://doi.org/10.1364/OE.444099`, `https://opg.optica.org/oe/abstract.cfm?URI=oe-30-2-2991`
7. Roddier, F.: Adaptive Optics in Astronomy. Cambridge University Press (2004)