Demonstrating reinforcement-learned gaits with two small quadrupeds

Leon Eshuijs[1,2,3]*, Gijs de Jong[3]*, and Arnoud Visser[3]

[1] Intelligence Systems Group, Universiteit Utrecht
[2] Computational Linguistics and Text Mining Lab, Vrije Universiteit
[3] Intelligent Robotics Lab, Universiteit van Amsterdam

Abstract. This demonstration shows how well the learned walking behaviors, based on two reinforcement techniques, could be applied on real robots. Two quadrupeds, the SpotMicroAI and the mini-Pupper 2 robot, are used to demonstrate the sim-to-real gap. The learned behaviors are not faster than the baseline gait, but include fall prevention, which make them potentially more stable.

Keywords: reinforcement learning · gait generation · robotics

1 Introduction

Legged robotics have the ability to walk through rough terrain, due to their small selection of contact points [2]. However, locomotion for legged robots is a challenging control problem as it requires precise control of actuators combined with good coordination between all four legs. With the recent advancements in machine learning, learning-based methods have shown promising results in both simulation and real-world robots [6,4,3].

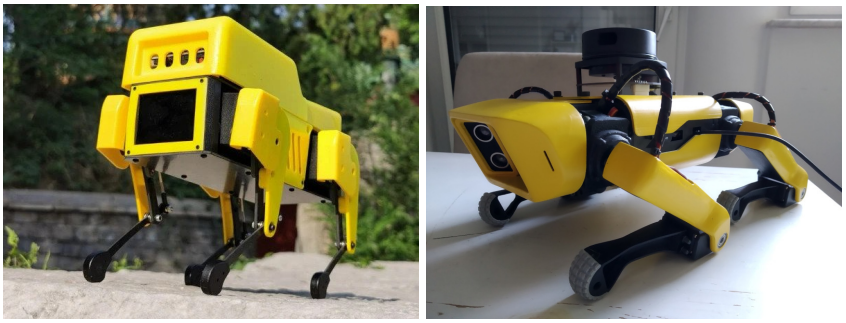
In the study presented at the BNAIC conference [1] two small robots were used to learn to walk with two reinforcement learning algorithms: Augmented Random Search and Soft Actor-Critic. The resulting gaits were tested both in simulation (on dynamic randomized terrains) and on a real robot (the MicroSpotAI). At BNAIC we like to demonstrate the learned gaits for both robots (the MicroSpotAI and mini-Pupper 2, see Fig. 1) live to the participants.

While these platforms do not come equipped with state-of-the-art actuators and sensors, Rahme *et al* demonstrated that state-of-the-art actuators and sensors are not required when applying learning-based methods to optimize robot's walking behaviors [7]. In this demonstration we show that this generalizes to both the MicroSpotAI and mini-Pupper 2.

2 System requirements

The goal of the demonstration is to let the robots walk a track of one meter. To reduce the impact of the legs touching the floor, the track will be softened of a

* Equal contribution



(a) mini-Pupper (b) SpotMicroAI

Fig. 1: Two low-cost open-source robot platforms. At the left the mini-Pupper 2 robot developed by Mangdang Robotics based on the Stanford Pupper [5]. At the right is the SpotMicroAI robot developed by Deok-yeon and further improved by [7].

thin flexible mat with a height of 0.5 cm. So mainly an open space is needed, together with a power supply.

3 Working of the system

Three short representative clips of each gait are available[4] are available for the MicroSpotAI. Unfortunately, the KickStarter program from MangDang robotics failed to finalize the design of the mini-Pupper 2 Pro in the timeline advertised at the beginning of the campaign, so no equivalent clip is available for the mini-Pupper 2. Luckily, we received a mini-Pupper 2 well before the BNAIC conference, which will allow us to demonstrate the walking behavior on both robots.

4 Conclusion

The experiments with the real MicroSpotAI quadruped robot showed that the learned gaits exhibited falling prevention behaviors. The resulting walking behavior were not faster than the baseline gait, but potentially more stable. Future research should demonstrate if an increase in the traversal speed is possible by adjusting the reward the agent receives so that is less concerned with minor deviations in the body orientation.

[4] https://youtu.be/Ji8rXjD60mU

References

1. Eshuijs, L., de Jong, G., Visser, A.: Learning to walk with a soft actor-critic approach. In: Proceedings of the 35th Benelux Conference on Artificial Intelligence (BNAIC) (Nov 2023)
2. Fukuhara, A., Gunji, M., Masuda, Y.: Comparative anatomy of quadruped robots and animals: a review. Advanced Robotics **36**(13), 612–630 (Jun 2022)
3. Haarnoja, T., Ha, S., Zhou, A., Tan, J., Tucker, G., Levine, S.: Learning to walk via deep reinforcement learning. In: Robotics: Science and Systems XV (June 2019)
4. Iscen, A., Caluwaerts, K., Tan, J., Zhang, T., Coumans, E., Sindhwani, V., Vanhoucke, V.: Policies modulating trajectory generators. In: Proceedings of the 2nd Conference on Robot Learning. pp. 916–926 (Oct 2018)
5. Kau, N.: Stanford pupper: A low-cost agile quadruped robot for benchmarking and education. preprint arXiv 2110.00736 (Feb 2022)
6. Kohl, N., Stone, P.: Policy gradient reinforcement learning for fast quadrupedal locomotion. In: IEEE International Conference on Robotics and Automation (ICRA '04). vol. 3, pp. 2619–2624 (Apr 2004)
7. Rahme, M., Abraham, I., Elwin, M.L., Murphey, T.D.: Dynamics and domain randomized gait modulation with bezier curves for sim-to-real legged locomotion. preprint arXiv 2010.12070 (Oct 2020)