

[11.1.2021]

[Learning Agile Robotic Locomotion Skills by Imitating Animals]

Summary

This paper proposed a new method for computing the trajectory of the robot facing unknown tough environments. The algorithm is implemented online and the moving speed of the robot can be as high as 4 m/s.

Some key points of this paper:

- extract key points from animal (or other outside animations) and use that as a target to learn
- use adaption strategy to learn the policy, see thoughts one for details.

Major Analysis and Comparison

- 1) The paper tries to imitate the motion from the animals → what's the advantage?
 - One advantage is that it is able to ease the pressure of hand craft gaits
 - But extra work has to be done to extract key points from the frame, and also based on this method, it seems that the targeted animal should have similar size as the simulation/ experiment animal → does this paper handle the problem?
 - I think not.
- 2)

Thoughts

- 1) What is the adaptive strategy in this paper? → why it is called sample efficient domain adaption?
 - I don't understand quite well, I suspect that if randomization is not necessary, this adaption might not needed.
 - The parameters along with the state decides the policy $\pi(a|s,z)$ where z is the representation of the parameter after an encoder. The paper says it is possible that the representation z can over fit to the parameter z (which now seems sensible to me, because those parameters are only a few) → that's why this paper says we need to

encourage the robustness to dynamics uncertainty (in the sense that the maximization of the expected reward should go along with that the KL divergence between the parameter distribution μ and a prior M should not be too big (should not look too alike)

- 2) Dynamic randomization works in this paper, but another paper (from the same author) shows that randomization is neither necessary nor sufficient.