Apollo3D: Training Goalie by Using Deep Reinforcement Learning

Xiangrui Jiang[1], Kuihan Chen[1], Chu Li[2], *Zhiwei Liang[1] and Tianjian Jiang[1]

College of Automation and College of Aritifical Intelligence,Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210023, China
apollo3d@foxmail.com
College of Portland,Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210023, China

1 Training Goalie by Deep Reinforcement Learning

In this paper, we propose a method for training goalkeepers using deep reinforcement learning. Compared to the previous goalkeepers who were judged by the physics engine, the goalkeepers trained with DRL can make more complex and varied saves depending on the position of the ball, their own condition and the environment. In the test, we can see that the neural network-based goalie saves improved by about 30 percent compared to traditional saves on mid-range kick saves.

1.1 Observation State

Like the other behaviors, the Goalie's observations mainly include the number of steps, the robot's z-axis position and velocity, IMU data, gyroscope data, acceleration data, and the robot's whole-body joint velocity and position. Also, we add the relative position of ball to the observation. At the same time, we tested adding the predicted ball point information for the pair of balls together with the observation, but the test effect was not significantly improved compared to the observation without it.

1.2 Training Method

We train neural networks for controlling robots by Proximal Policy Optimization(PPO) based on a deeply improved FCP code base [1]. The training environment is shown in Fig.1 and is divided into a kicking intelligence as well as a trained goalkeeper. The kicker completes shots of different lengths and angles in different areas, and the goalkeeper tries to stop the kicker from winning the goal. More detailed process could see the video(https://www.youtube.com/watch?v=rIcRcwHgsVQ). The reward in the test is defined as $Reward = (t_{sucess} - t_{fail})/t_{total}$.

The part of training process is shown in Fig.2. Although the rewards have not yet converged during training, they have shown better results in testing

Fig. 1. Training Environment.

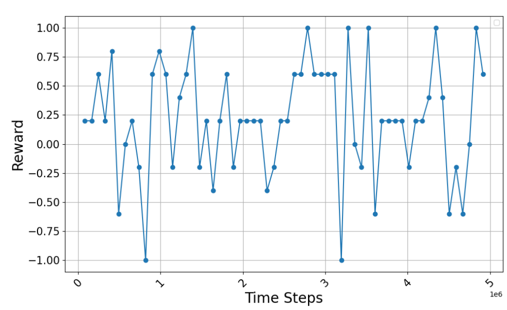

Fig. 2. Training Process.

References

1. M. Abreu, L. P. Reis, and N. Lau, "Designing a skilled soccer team for robocup: Exploring skill-set-primitives through reinforcement learning," *arXiv preprint arXiv:2312.14360*, 2023.