

Design of a reinforcement learning framework for terrain-adaptive locomotion of a tracked mobile robot in nuclear power plants

Seungjun Lee^{a,b}, Wonseo Lee^{a*}, Dongseok Ryu^a, and Jeakweon Han^b

^a Extreme Robotics Team, Korea Atomic Energy Research Institute (KAERI), Daejeon, Republic of Korea, 34057

^b Department of Robotics, Hanyang University, Ansan, Republic of Korea

*Corresponding author: wonseo@kaeri.re.kr

***Keywords :** Tracked mobile robot, Reinforcement learning, Terrain-adaptive locomotion, Nuclear power plants

1. Introduction

Today, mobile robots are widely used for disaster response, facility inspection, and maintenance because they can reduce risks to workers and perform tasks in environments that are difficult for humans to access. In particular, access to nuclear power plants may be limited during emergency situations due to the risk of radiation exposure and complex internal structures. In addition, stairs, steps, and irregular terrain inside nuclear power plants can reduce the locomotion stability of mobile robots. Response tasks in nuclear power plants require not only locomotion but also equipment transport and manipulation. Therefore, a mobile platform that can carry heavy work equipment and traverse complex terrain is required. Tracked mobile robots are suitable for such tasks because their large ground contact area and continuous tracks provide good mobility even when carrying heavy equipment. However, heavy equipment such as a manipulator can raise the center of gravity of the robot, reducing posture stability and increasing the risk of overturning when traversing slopes and steps. Therefore, a control method is required to adjust the robot posture using flippers according to the surrounding terrain and robot state to improve locomotion stability. To address these issues, terrain-aware reinforcement learning (RL) methods that use terrain information for robot control have been actively studied.

Pan *et al.* used NuBot (a rescue tracked robot, China) and converted surrounding terrain point clouds into local terrain height information. The terrain information was combined with robot states, such as flipper angles and chassis pitch angle, to construct the RL state for autonomous flipper control [1]. Gangapurwala *et al.* developed a terrain-aware locomotion framework for the ANYmal quadruped robot, in which the surrounding terrain elevation map was converted into terrain features and combined with robot states as inputs to the RL policy to determine suitable footholds for locomotion over rough terrain [2]. However, previous studies either used limited terrain information and symmetric flipper control or mainly focused on quadruped robots, which have different mechanical and control structures from KAERI-developed tracked mobile platform, TALON [3]. Therefore, independent flipper control based on terrain and robot states is

required for stable rough-terrain locomotion of TALON with a high center of gravity.

In this study, we propose a terrain-aware reinforcement learning framework to improve the locomotion stability of TALON with an elevated center of gravity caused by heavy equipment. The policy independently controls the four flippers based on surrounding terrain information and robot states. It is trained to limit body roll and pitch while maintaining locomotion toward the target direction. Through this framework, we aim to achieve stable terrain-adaptive locomotion of TALON under a high center-of-gravity condition.

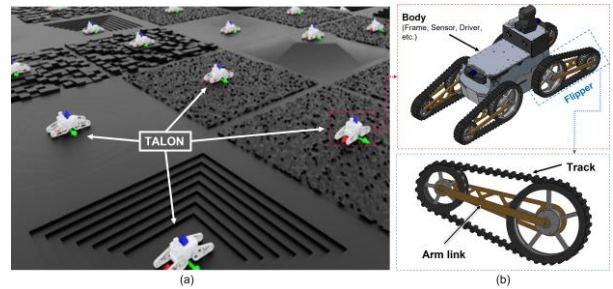


Fig. 1. Simulation environment and mechanical configuration of TALON: (a) rough-terrain training environment in Isaac Lab and (b) body and Flipper configuration of TALON.

2. Reinforcement learning framework for terrain-adaptive locomotion of a tracked mobile robot

2.1 Environment settings

TALON is a tracked mobile robot in which each track arm is equipped with a track drive motor and a flipper motor. The robot can traverse various terrains using the obstacle-crossing capability of the tracks and the posture adjustment of the flippers. Fig. 1 shows the rough-terrain environment constructed for reinforcement learning and the overall configuration of TALON.

The simulation model of TALON was validated in Isaac Sim, and the reinforcement learning environment was constructed based on Isaac Lab. The policy was trained using the PPO algorithm provided by RSL-RL. Multiple TALON environments were simulated in parallel during training, while a single-robot

environment was used to evaluate the behavior of the trained policy.

To provide terrain geometry information to the policy, the simulation environment was configured to acquire terrain information around the robot. The acquired terrain information is processed in the robot coordinate frame and used together with IMU and actuator states to construct the policy observation. In the current simulation environment, Isaac Lab RayCaster is used to acquire terrain information around TALON. The simulation uses generic rough terrain rather than reproducing a specific damaged nuclear facility.

2.2 Framework design

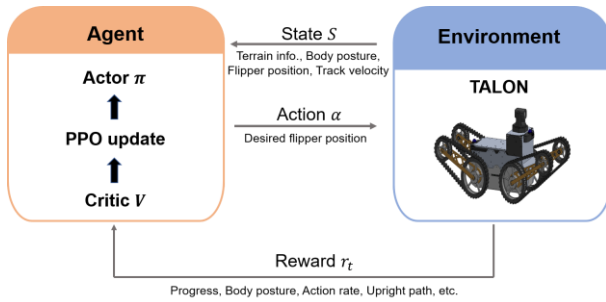


Fig. 2. Overall structure of the proposed RL framework.

Fig. 2 shows the overall reinforcement learning framework for terrain-adaptive locomotion of TALON. Based on the interaction between TALON and the rough terrain, the policy receives observations consisting of surrounding terrain information and robot posture and actuator states, and determines actions for the four flippers. After the actions are applied to the robot, rewards are calculated based on the resulting locomotion behavior. By repeatedly interacting with the environment, the policy learns a flipper control strategy suitable for the terrain and robot state.

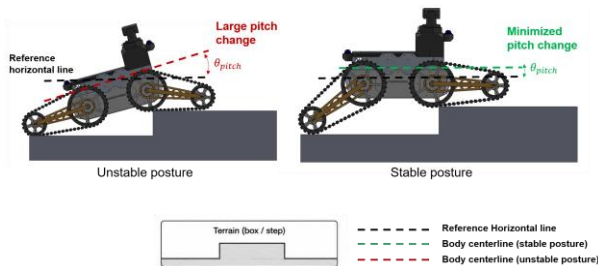


Fig. 3. Conceptual example of body pitch stabilization using adaptive flipper control during step traversal.

Fig. 3 shows a conceptual example of posture stabilization using flipper control during step traversal. Without appropriate flipper adjustment, a large change in body pitch may occur when the robot passes over a step, while uneven terrain can also induce body roll. In this study, the policy independently determines the actions of the four flippers based on terrain information and robot states to limit excessive roll and pitch changes and maintain a stable posture.

A PPO-based symmetric actor-critic architecture is used for policy training. The actor and critic use the same observations. Considering future deployment on the real TALON platform, the policy inputs are mainly composed of terrain information and robot states that can be obtained from the IMU and encoders. The forward velocity of the tracks is maintained at a constant value, while the policy independently outputs control actions for the four flippers, FL, FR, RL, and RR.

The reward function is mainly designed to promote forward progress toward the target direction and maintain robot posture stability. In particular, considering the high center of gravity caused by heavy equipment on TALON, the reward is designed to limit excessive body roll and pitch, deviation from the desired direction, and abrupt changes in actions. The required flipper posture for a specific terrain is not predefined. Instead, the policy learns an appropriate flipper control strategy for each terrain condition based on rewards from the resulting locomotion behavior.

3. Conclusions

In this study, a terrain-aware reinforcement learning framework was designed for stable locomotion of the tracked mobile robot TALON over rough terrain in nuclear power plants. A reinforcement learning environment consisting of TALON and various rough terrains was constructed using Isaac Sim and Isaac Lab. Surrounding terrain information, IMU measurements, and actuator states were used as policy observations. In addition, the framework was designed to independently control the four flippers using a PPO-based policy and to learn an appropriate flipper control strategy by considering forward progress and robot posture stability.

In future work, the reward design will be further optimized for stable flipper control over various rough terrains. Adaptive flipper control according to terrain changes and its effect on robot posture stability will first be evaluated in simulation. The trained policy will then be deployed on the real TALON platform, and adaptive flipper control based on actual terrain sensor data will be evaluated. Through these experiments, the effect of the proposed framework on improving the locomotion stability of TALON under a high center-of-gravity condition will be validated.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00144385).

REFERENCES

- [1] H. Pan, X. Chen, J. Ren, B. Chen, K. Huang, H. Zhang, and H. Lu, Deep Reinforcement Learning for Flipper Control of Tracked Robots in Urban Rescuing Environments, Remote Sensing, Vol. 15, 4616, 2023.

[2] S. Gangapurwala, M. Geisert, R. Orsolino, M. Fallon, and I. Havoutis, RLOC: Terrain-Aware Legged Locomotion Using Reinforcement Learning and Optimal Control, *IEEE Transactions on Robotics*, Vol. 38, No. 5, pp. 2908-2927, 2022.

[3] W. Lee, J. Park, D. Hyun, and D. Ryu, Development of a tracked mobile platform with adaptive locomotion capability for operations in nuclear power plants, *Korean Nuclear Society Spring Meeting*, May 7-8, 2026, Jeju, Korea.