Learning Climbing Controllers for Physics-Based Characters

Kyungwon Kang, Taehong Gu and Taesoo Kwon

Hanyang University

INTRODUCTION

Capturing both motion and the environment simultaneously is a challenging task. This is especially true for recording climbing movements, where it is crucial to capture both the motion and the wall structure.

Recent studies have explored motion synthesis using techniques like character control in simulated environments, with a focus on reinforcement learning [1, 2].

We propose a two-stage reinforcement learning-based method that utilizes AMP [2] to synthesize climbing motions when given climbing wall structures in a physically simulated environment.

SIMULATION ENVIRONMENT

Climbing Wall – The climbing wall has a slope of approximately 80°, with 32 hemispherical holds, each with a 10cm radius, arranged in a grid pattern.

Grip using Constraints – A limb's grip signal is generated by policy and its end-effector is within distance d of the center of a hold. A ball joint is temporarily activated to attach the end-effector to its current position.

State and Action Space – The control policy is defined as $\pi(\boldsymbol{a}|\boldsymbol{s}, \boldsymbol{o})$.

- State $\boldsymbol{s}$: root rotation, velocity, angular velocity, joint angles, joint velocities, relative position of the end-effectors, and the x-axis root-relative position from the starting point.
- Action $\boldsymbol{a}$: target angles of each joint(28 DoF) and the 4-dimensional signals to determine grip or release.

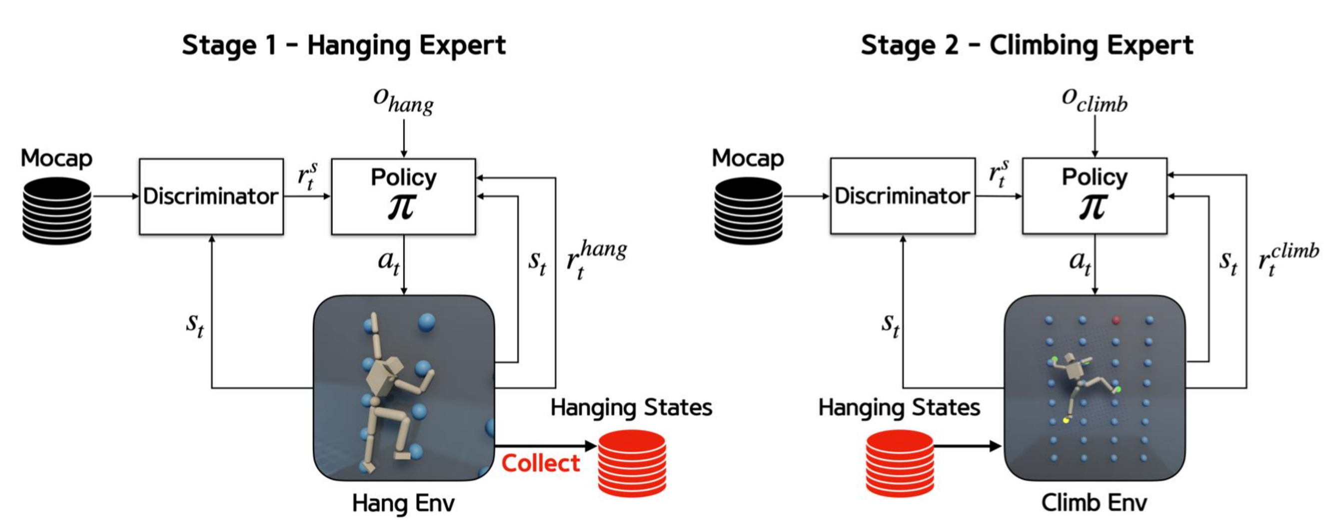

Figure 1 : Overview of our method

Hang Policy

In stage 1, the hanging policy learns to grip as many holds as possible when the character is placed at a random position around the climbing wall in a pose randomly extracted from the motion capture dataset.

Once the hanging policy training is completed, the character successfully performs a 4-grip simultaneously at the end of each episode, and the pose and grip status are collected. Using these data as the initial state dataset of the climbing policy allows the character to experience more diverse poses and holds.

- Task observation $\boldsymbol{o}_{hang}$: depth of the climbing wall, the signals of each end-effector indicating grip or release state and whether each end-effector is actually gripping or releasing.
- The reward function :

$$r_{hang} = r_{num\_grip} + r_{close} + r_{chest} * r_{up}$$

Climbing Policy

In stage 2, the climbing policy learns to maneuver the character's root towards a specified target location on the climbing wall. In this stage, the hanging state dataset obtained from the previous stage is used for initializing the episodes.

- Task observation $\boldsymbol{o}_{climb}$: $\boldsymbol{o}_{hang}$, the duration of gripping of releasing states for end-effectors and target position on yz-plane expressed relative to the root.
- The reward function :

$$r_{climb} = r_{pos} + r_{mov} + r_{speed} + r_{grip} + r_{chest} * r_{up} + r_{force}$$

Contacts

- Kyungwon Kang – kang361973@gmail.com
- Taehong Gu – gestoru@gmail.com
- Taesoo Kwon – taesoobear@gmail.com

RESULTS

Dataset – Some clips from Mixamo [3] and CIMI4D [4] were used for training. The size of hanging state dataset is 50.

Training – PPO was used as RL Algorithm. The grip distance d is experimentally chosen to be 15 cm and 20 cm for the hanging policy and climbing policy, respectively.

Evaluations – Figures 2 and 3 show the qualitative evaluations of the trained models for the hanging policy and climbing policy, respectively. Figure 4 and Table 1 show the ablation study results that verify the effectiveness of our method in the climbing task. In Figure 4, our method shows more natural motions with more even use of limbs than the others. Table 1 shows the success rates for the three methods over 1000 episodes.

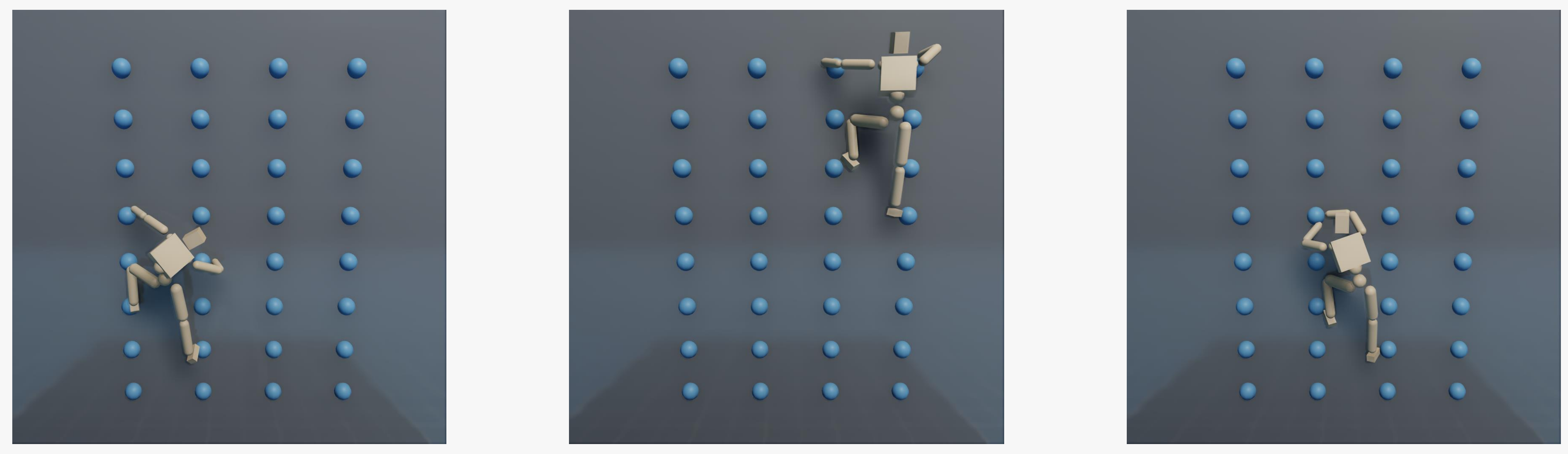

Figure 2 : Examples of learning outcomes of the hanging policy.

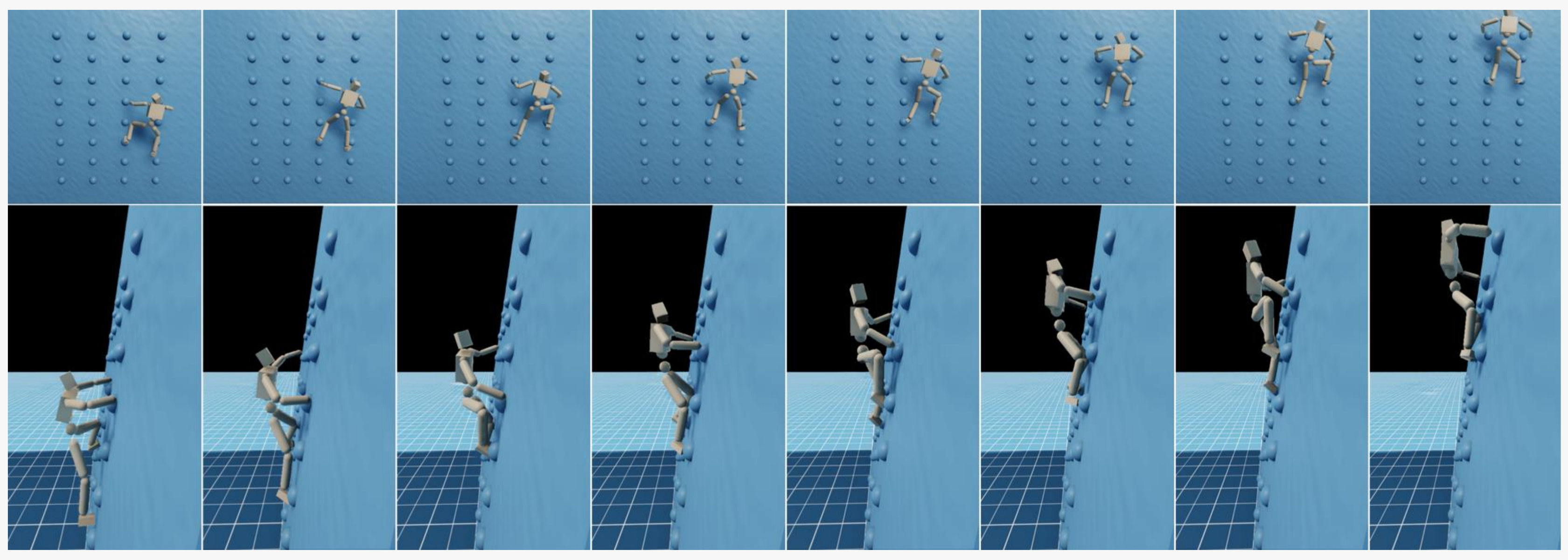

Figure 3 : Motion sequences obtained from the learned climbing expert model.
Top: frontal views of the resulting motions. Bottom: side views.

(a) Without Hanging State Collection

(b) Without r_{grip}

(c) Ours

Figure 4 : Ablation Results. The lights attached to the body indicate whether the body is actually anchored. Green : anchored, Yellow : released.

	Success Rate
Without HSC	0.89
Without r_{grip}	0.90
Ours	**0.99**

Table 1 : Comparison of success rate

Future Work

- Validation in more diverse environments.
- Simplification of the reward function.
- Interaction with holds using real hands and feet.

Demo Video

REFERENCES

[1] LNADERI K., BABADI A., ROOHI S., HÄMÄLÄINEN P.: A reinforcement learning approach to synthesizing climbing movements. In 2019 IEEE Conference on Games (CoG) (2019), pp. 1–7.
[2] PENG X. B., MA Z., ABBEEL P., LEVINE S., KANAZAWA A.: Amp: Adversarial motion priors for stylized physics-based character control. ACM Trans. Graph. 40, 4 (July 2021).
[3] ADOBE: mixamo. https://www.mixamo.com (2020).
[4] YAN M., WANG X., DAI Y., SHEN S., WEN C., XU L., MA Y., WANG C.: Cimi4d: A large multimodal climbing motion dataset un- der human-scene interactions. In Proceedings of the IEEE/CVF Confer ence on Computer Vision and Pattern Recognition (CVPR) (June 2023), pp. 12977–12988.