ACM SIGGRAPH / Eurographics Symposium on Computer Animation (2024)
M. Skouras and H. Wang (Editors) *Poster*

Learning Climbing Controllers for Physics-Based Characters

Kyungwon Kang[1], Taehong Gu[1] and Taesoo Kwon[1]

[1] Hanyang University, South Korea

Abstract

We propose a physics-based climbing controller that consists of two learning stages. Firstly, a hanging policy is trained to grasp holds in a natural posture. Once the policy is obtained, it is used to extract the positions of the holds, postures, and grip states, thus forming a dataset of favorable hanging poses. Subsequently, a climbing policy is trained to execute actual climbing maneuvers using this hanging state dataset. The climbing policy allows the character to move to the target location using limbs more evenly. Experiments have shown that the proposed method can effectively explore the space of good postures for climbing.

CCS Concepts

• ***Computing methodologies*** → *Physical simulation; Motion capture; Reinforcement learning;*

1. 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. Once a wall is created, modifying it becomes difficult, making it challenging to capture a variety of movements. Consequently, generating motion for different scenes requires post-processing work by animators.

Recent studies have explored motion synthesis using techniques like character control in simulated environments, with a focus on reinforcement learning [NBRH19, PMA*21]. These methods generate natural, diverse movements and are useful in interactive applications such as game and movie prototyping. We propose a reinforcement learning-based method capable of synthesizing realistic climbing motions when provided with climbing wall structures in a physically simulated environment.

Climbing involves using all limbs to ascend a wall or a steep slope, with movements restricted to predetermined holds, unlike general locomotion. The character needs to grasp handholds or step on footholds to support the body without falling, which presents challenges in routing. To address this, we propose a two-stage curriculum learning approach, as shown in Figure 2 The first stage involves obtaining the hanging policy, which learns to effectively grasp holds on the wall. After the completion of training for the hanging policy, a random initial pose can transition into a stable pose where multiple holds are grasped simultaneously in a natural stance after each episode. Using the trained policy, a dataset is compiled by gathering these poses along with the corresponding grasp statuses of the holds. The second stage is climbing policy training. In this stage, a character controller is trained to allow the character to climb to the designated position. The dataset obtained through the hanging policy is used to determine the initial pose and grip status, enabling more challenging tasks. Both stages use AMP architecture [PMA*21] for natural motion. As a result of training, the character uses its limbs at an adequate frequency to move to the designated position of the given climbing wall structure.

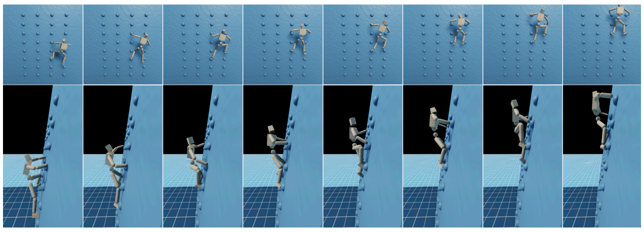

Figure 1: *Motion sequences obtained from the learned climbing expert model. Top: frontal views of the resulting motions. Bottom: side views.*

2. Learning Climbing Control Policy

Simulation Environment. In simulation, the climbing wall has a slope of approximately 80°, with 32 hemispherical holds, each with a 10cm radius, arranged in a grid pattern. The character has 13 joints and 28 degrees of freedom, stands 1.75m tall, weighs 47kg, and is controlled by a PD controller for each joint. The character's gripping and stepping onto holds are modeled as temporary constraints in the simulation.

When a limb's grip signal is generated 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. Note that this ball joint can generate contact forces in all directions, un-

DOI: 10.2312/sca.20241165

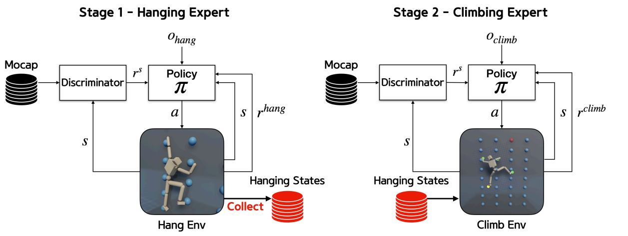

Figure 2: *The system comprises two stages of training. During the first stage, it learns to hang in natural poses and collects the hanging states using the acquired policy. In the second stage, it learns a climbing policy using the collected hanging state dataset.*

like a realistic contact model. When the grip signal is turned off, the corresponding joint is deactivated. In all experiments, wrists and toes serve as the end-effectors for the upper and lower body, respectively.

State and Action Space. The control policy $\pi(\mathbf{a}|\mathbf{s},\mathbf{o})$, where $\mathbf{a}$, $\mathbf{s}$, and $\mathbf{o}$ represent the action, state, and task-specific observation, respectively, governs the character's movements. The character state $\mathbf{s}$ includes the character'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 to determine the character's location. The action $\mathbf{a}$ comprises the target angles of each joint and the 4-dimensional signals to determine grip or release.

Hanging Policy. 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. The observation for the hanging policy $\mathbf{o}_{hang}$ comprises depth of the climbing wall frame - a 1.9m x 2.5m rectangular grid centered at pelvis, with an accuracy of 10cm, the signals of each end-effector indicating grip or release state and whether each end-effector is actually gripping or releasing. The reward function is defined as follows:

$$r_{hang} = r_{num\_grip} + r_{close} + r_{chest} * r_{up}. \quad (1)$$

r_{num_grip} encourages the character to grip as many holds as possible at the same time. r_{close} encourages the model to stay close to the wall while gripping holds. r_{chest} guides the facing direction of the upper body to align with the slope of the wall. r_{up} prevents the upper body from lying down.

Climbing Policy. 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. The observation for the climbing policy $\mathbf{o}_{climb}$ comprises $\mathbf{o}_{hang}$, the duration of gripping of releasing states for end-effectors and target position on the yz plane expressed relative to the pelvis. The reward function is defined as follows:

$$r_{climb} = r_{pos} + r_{mov} + r_{speed} + r_{grip} + r_{chest} * r_{up} + r_{force}. \quad (2)$$

r_{pos} minimizes the distance between the target position. r_{mov} aligns the character's movement direction. r_{speed} guides the root to reach the target speed. r_{grip} encourages the character to grip as many holds as possible simultaneously while also exploring new holds. The last term r_{force} penalizes using only arms, by using a higher penalty when the sum of arm torques are greater.

	Success Rate
Without HSC	0.89
Without grip reward	0.90
Ours	**0.99**

Table 1: *Comparison of success rate of our method with methods without Hanging State Collection(HSC) and without grip reward.*

3. Results and Conclusions

The grip distance d is experimentally chosen to be 15 cm and 20 cm for the hanging policy and climbing policy, respectively. Some clips from Mixamo [Ado20] and CIMI4D [YWD*23] were used for training the AMP discriminators in the hanging and climbing policies. Proximal Policy Optimization(PPO) was used as the reinforcement learning algorithm. Figure 1 represents the results of the trained climbing model. The character can move to desired position on the wall. Table 1 shows the success rates for the three methods. To demonstrate the effectiveness of the proposed method in climbing tasks, the success rate was measured over 1000 episodes, with success defined as the root being within 10 cm of the target position. The proposed method achieved the highest success rate of 0.99.

We proposed a two-stage climbing controller training methodology for physically-based character control. The hanging policy produces a dataset of favorable hanging poses, which is subsequently used to train the climbing policy. In the future, we aim to reduce the complexity of the reward equations by integrating a wide range of motion capture clips.

References

[Ado20] ADOBE: mixamo. *https://www.mixamo.com* (2020). 2

[NBRH19] NADERI 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. doi:10.1109/CIG.2019.8848127. 1

[PMA*21] 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). 1

[YWD*23] YAN M., WANG X., DAI Y., SHEN S., WEN C., XU L., MA Y., WANG C.: Cimi4d: A large multimodal climbing motion dataset under human-scene interactions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2023), pp. 12977–12988. 2