

LadderMan: Learning Humanoid Perceptive Ladder Climbing

Siheng Zhao^{1,2}, Yuanhang Zhang¹, Ziqi Lu¹, Pieter Abbeel^{1,3}, Rocky Duan¹,
Koushil Sreenath^{1,3}, Yue Wang², C. Karen Liu^{1,4†}, Guanya Shi^{1,5†}

¹Amazon FAR, ²USC, ³UC Berkeley, ⁴Stanford University, ⁵CMU, [†]Amazon FAR Co-Lead



Figure 1: **LadderMan** enables a Unitree G1 humanoid to robustly climb diverse ladders using a single perceptive policy via zero-shot sim-to-real transfer. Building on the learned climbing capability, LadderMan further supports stable whole-body on-ladder manipulation via teleoperation, including adjusting paintings, replacing light bulbs, and box handover.

Abstract: Humanoid robots hold great promise for operating in human-centered environments, yet ladder climbing remains one of the most challenging tasks due to sparse footholds and handholds, complex whole-body coordination, and sensitivity to perception and control errors. We present **LadderMan**, a unified system that enables humanoid robots to robustly climb diverse ladders and perform manipulation under such constrained conditions. Our climbing policy is built on a scalable two-stage learning pipeline, where we use hybrid motion tracking to learn multiple climbing experts from a single reference motion, and distill these experts into a unified depth-based visuomotor climbing policy via hybrid imitation and reinforcement learning. To enable real-world deployment, we leverage vision foundation models to bridge the sim-to-real gap in depth perception. Building on the learned climbing policy, we further train a separate manipulation policy using a dual-agent formulation, allowing stable on-ladder manipulation via teleoperation. Experiments demonstrate that LadderMan achieves robust ladder climbing across a wide range of geometries, successfully transfers to real-world hardware in a zero-shot manner, and supports various manipulation tasks under challenging ladder constraints. Video results are available on the [project website](#).

Keywords: Robot Ladder Climbing, Humanoid Perceptive Loco-Manipulation

1 Introduction

Enabling humanoid robots to robustly climb and operate on ladders is a key milestone toward deploying them in industrial environments such as construction sites, warehouses, and maintenance tasks. However, unlike walking on flat ground or stairs, ladder climbing requires precise and sustained coordination between upper and lower body contacts under sparse, discontinuous, and narrow footholds and handholds [1]. Small perception errors or control inaccuracies can easily lead to catastrophic failure. This challenge is further amplified in real-world environments, where ladders vary significantly in inclination, rung spacing, and structural design.

Despite recent advances in humanoid locomotion over uneven terrain [2, 3, 4], most existing approaches remain dominated by lower-body behaviors and do not fully exploit the whole-body capabilities of humanoids. Prior work on ladder climbing [5, 6, 7, 8, 9, 10, 11] largely relies on carefully engineered model-based control without incorporating perception into the control loop, and requires accurate environment modeling, task-specific tuning, and specialized hardware design. These approaches typically exhibit slow climbing speeds, limited robustness, and poor generalization beyond specific ladder configurations. Moreover, they do not extend to on-ladder manipulation [12, 13].

In this work, we ask: *can we learn a visuomotor system that enables humanoid robots to robustly climb diverse ladders and perform on-ladder manipulation while maintaining stable balance, without additional hardware modifications?*

To this end, we propose **LadderMan**, a unified system for perceptive humanoid ladder climbing and on-ladder manipulation. (1) Our climbing policy is built on a scalable two-stage learning pipeline. In the first stage, we introduce hybrid motion tracking to learn multiple expert policies from a single reference motion without requiring large motion datasets. To achieve this, we use relaxed upper-body tracking and incorporate ladder-centric contact tracking and climbing rewards. In the second stage, we distill these experts into a single depth-based visuomotor policy via hybrid imitation and reinforcement learning (RL), resulting in a unified policy that generalizes robustly across diverse ladder geometries. (2) To enable real-world deployment on narrow and sparse ladder structures under background clutter, we leverage a vision foundation model [14] (VFM) to bridge the gap between simulated and real-world depth, instead of relying on extensive depth randomization and manual tuning. (3) Building on the ladder climbing capability, we further investigate whole-body manipulation under ladder constraints. We adopt a dual-agent learning framework that decouples lower-body stabilization from upper-body manipulation [15], enabling the humanoid to perform manipulation tasks such as adjusting paintings and tightening light bulbs, while maintaining balance on ladder, which is difficult to achieve with off-the-shelf whole-body teleoperation policies [16, 17].

As shown in Fig. 1, our experiments demonstrate that LadderMan enables robust zero-shot sim-to-real ladder climbing across diverse ladder geometries, including variations in inclination, rung spacing, and structural design. The learned climbing policy achieves strong generalization in simulation, competitive climbing speeds compared to humans, and successful deployment on real-world humanoid hardware without additional hardware modifications. Furthermore, LadderMan supports stable on-ladder manipulation through a dual-agent teleoperation framework, outperforming off-the-shelf whole-body teleoperation policies under ladder constraints. All training and inference code, together with the deployable models, will be open-sourced.

2 Related Works

Robot Ladder Climbing. Robot ladder climbing has long been considered a challenging problem in robotics, and was included as a key task in DARPA Robotics Challenge [18]. A large body of prior work [5, 6, 7, 8, 9, 10, 11] has explored enabling humanoid robots to climb ladders using model-based control. These approaches typically assume access to accurate environment models and rely on task-specific tuning to ensure stable contacts. As a result, they are often limited to specific ladder configurations and do not incorporate perception as part of the control loop, making generalization to unseen ladders challenging. More recently, RL has been explored for ladder climbing. Vogel

et al. [19] demonstrates quadruped ladder climbing using an end-to-end policy, highlighting the potential of learning-based approaches for ladder traversal. However, compared with quadrupeds, humanoid ladder climbing introduces substantially greater challenges due to heterogeneous limbs, unstable support polygons, and the need for coordinated whole-body contacts between hands and feet. In addition, their method relies on privileged inputs such as motion capture and known ladder geometry, without incorporating perception directly into the control policy, and further requires task-specific hardware modifications. In contrast, our work learns a perceptive humanoid ladder climbing policy that directly operates on depth observations and generalizes across diverse ladder geometries. Building on this capability, we further enable manipulation under ladder constraints.

Perceptive Locomotion for Legged Robots. Perceptive locomotion enables legged robots to traverse complex terrains using exteroceptive sensing such as LiDAR and depth cameras. Early approaches typically rely on LiDAR-based mapping [4, 20, 21, 22, 23, 24, 25], where raw sensor data are converted into terrain representations such as elevation maps to guide foothold selection and control. While effective for moderately structured environments, these methods depend heavily on accurate calibration, state estimation, and online mapping, which can introduce additional latency and sim-to-real gap. Furthermore, elevation maps struggle to capture thin and sparse structures such as ladder rungs with sufficient precision. To overcome these limitations, more recent work has explored end-to-end visuomotor policies that directly use depth images [2, 3, 26, 27, 28, 29] for locomotion control. These approaches have demonstrated impressive performance on challenging terrains including stairs [26], stepping stones [4], and parkour-style environments [27]. However, most existing methods rely heavily on reward shaping, which tends to produce lower-body-dominated behaviors when applied to humanoids [2, 3, 4, 21, 25]. As a result, they do not fully exploit the whole-body coordination capabilities required for humanoid systems. Concurrent work on perceptive whole-body humanoid locomotion [30, 31] demonstrates promising results on tasks such as obstacle climbing and vaulting which mainly focus on agility and dynamic interactions. In contrast, ladder climbing requires fine-grained and highly constrained multi-contact coordination on thin and sparse structures, placing significantly higher demands on perception and control accuracy.

3 Method

Our goal is to learn a visuomotor system that enables humanoid robots to robustly climb diverse ladders and further perform manipulation under such constrained settings. This problem is particularly challenging due to four coupled factors: (1) ladder climbing requires precise whole-body coordination with tightly coupled hand-foot contacts, (2) ladder geometries vary significantly in inclination and rung spacing, making generalization difficult, (3) perception and control errors during sim-to-real on thin and sparse structures such as ladder rungs can easily accumulate and lead to catastrophic failures, and (4) manipulation under ladder constraints requires maintaining stable balance while executing upper-body interaction tasks. To address these challenges, we design LadderMan as a system with four key components: learning multiple expert climbing policies from a single reference motion via hybrid motion tracking (Sec. 3.1); distilling expert policies into a unified visuomotor climbing policy through hybrid imitation and RL (Sec. 3.2); bridging the sim-to-real perception gap using a vision foundation model (Sec. 3.3); enabling stable on-ladder manipulation through dual-agent learning (Sec. 3.4). An overview of LadderMan is shown in Fig. 2.

3.1 Learning Multiple Expert Policies from Single Reference Motion

Directly learning a single climbing policy that generalizes across diverse ladders is challenging. Therefore, in the first stage, we instead learn multiple state-based expert policies $\pi_{\phi, z}^{\text{expert}}$ corresponding to different ladder configurations (ϕ, z) , where ϕ denotes ladder inclination and z rung spacing. These expert policies are later distilled into a unified visuomotor policy π^{visual} . However, existing approaches are insufficient for humanoid ladder climbing: RL with reward shaping often leads to lower-body-dominated behaviors [2], while standard motion tracking methods [32] require large motion datasets to cover diverse ladder geometries, which are difficult to obtain in practice.

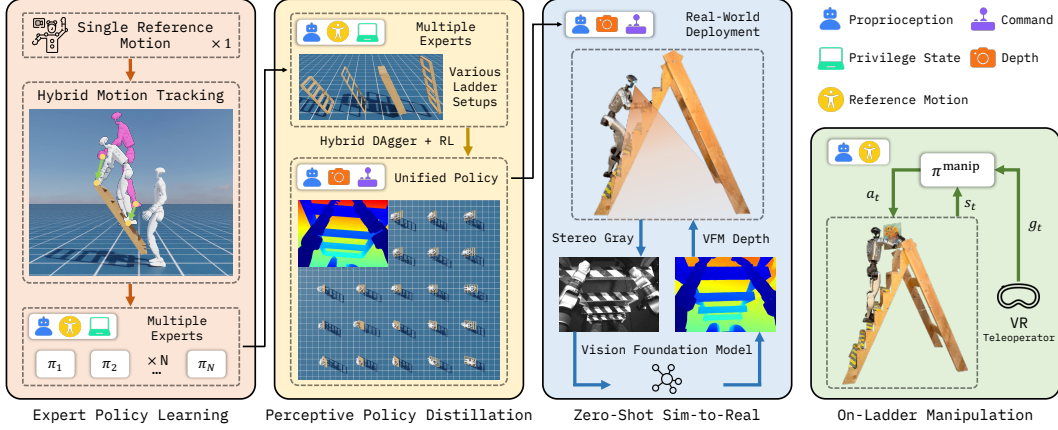


Figure 2: **Overview of LadderMan.** Starting from a single reference motion, we first learn multiple expert climbing policies for different ladder geometries using hybrid motion tracking. These experts are then distilled into a unified visuomotor policy through hybrid DAgger and RL. Building on the climbing policy, we further train a dual-agent manipulation policy for stable on-ladder teleoperation. To bridge the sim-to-real perception gap, we use VFM for robust real-world deployment.

To address these challenges, we introduce hybrid motion tracking, which combines relaxed upper-body tracking with task-driven adaptation to learn multiple expert climbing policies from a single reference. Details such as observation space and full reward terms are provided in Appendix A.1.

Reward. Standard motion tracking [33] reward is $r_X^{\text{tracking}} = \omega \exp(-\sigma \|\hat{X}[K] - X[K]\|_2)$, where K denotes all tracked keypoints, X denotes quantities such as joint positions and velocities, and \hat{X} denotes the reference quantities. In hybrid motion tracking, we partition the body into upper-body keypoints K_u and lower-body keypoints K_l , and use the following asymmetric tracking reward:

$$r_X^{\text{tracking}} = \frac{\omega}{2} \exp\left(-\frac{\sigma}{2} \|\hat{X}[K_u] - X[K_u]\|_2\right) + \omega \exp\left(-\sigma \|\hat{X}[K_l] - X[K_l]\|_2\right). \quad (1)$$

This asymmetric formulation encourages faithful reproduction of stable lower-body climbing behaviors while allowing the upper body to adapt contact locations according to ladder geometry. To further guide adaptation, we introduce ladder-centric contact and climbing rewards:

$$r^{\text{contact}} = \omega \hat{c} \exp\left(-\sigma \|P - \hat{P}\|_2\right) + (1 - \hat{c}), \quad r^{\text{task}} = \|R - \mathcal{G}\|_2, \quad (2)$$

where \hat{c} denotes the reference contact indicator, P the current end-effector position, \hat{P} the target contact position, R the current root pose, and \mathcal{G} the target root pose. The reference contact indicator is manually annotated once for the reference motion. The target contact position is automatically generated from the ladder geometry by assigning predefined contact locations on the target ladder rung for each hand and foot. This hybrid formulation enables successful expert policy learning across diverse ladders from a single reference motion.

Observation. We extend the standard motion tracking observations (proprioception, reference motion, and privileged robot states) with relative poses between the humanoid pelvis and ladder.

Termination and Domain Randomization. For termination, standard motion tracking terminates when either the local or global tracking error exceeds a predefined threshold. We further introduce a ladder-specific contact termination condition, where episodes terminate if required ladder contacts are lost for more than 30 consecutive frames. For domain randomization, in addition to physical parameter and initialization state randomization, we apply local perturbations to ladder configurations around each target configuration (ϕ, z) , together with rung geometry and friction properties, to improve robustness during expert policy learning.

3.2 Unified Climbing Policy via Hybrid Imitation and Reinforcement Learning

While individual expert policies $\pi_{\phi, z}^{\text{expert}}$ can handle specific ladder configurations (ϕ, z) , practical deployment requires a unified policy that generalizes across diverse ladders and remains robust

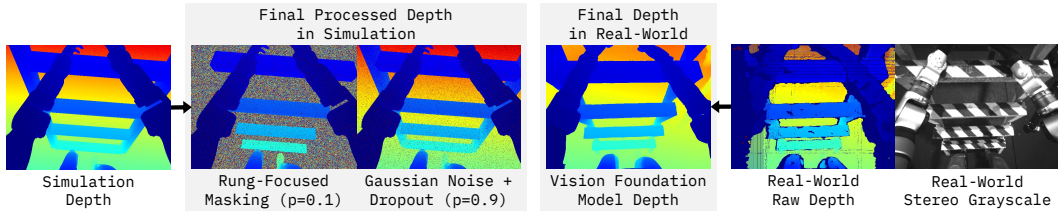


Figure 3: We bridge the visual gap between simulation and real world by applying rung-focused masking, vision foundation model [14], and minimalist noise augmentation.

under compounding errors. A natural approach is to distill multiple experts into a single visuomotor policy π^{visual} using DAgger-style imitation learning [34]. However, imitation alone is insufficient for this highly dynamic and contact-rich task: expert policies do not fully cover the configuration space, and small imitation errors can accumulate and lead to failure. To address this issue, we combine DAgger with RL during policy distillation. The unified climbing policy is optimized with:

$$L(\pi^{\text{visual}}) = L_{\text{RL}}(\pi^{\text{visual}}) + \lambda D_{\text{KL}}(\pi^{\text{visual}} \parallel \pi^{\text{expert}}), \quad (3)$$

where L_{RL} is the PPO [35] objective using the same rewards as the expert policies, and the KL term encourages imitation of expert behaviors. The weight λ is gradually annealed during training, allowing the policy to transition from imitation-driven initialization to task-driven optimization. Details such as observation space and model architecture are provided in Appendix A.2.

Observation and Depth Rendering. The unified policy π^{visual} takes proprioception, depth observations, and a binary climbing direction command as input, and outputs whole-body actions. During training, depth images are rendered using NVIDIA Warp [36] for high-throughput simulation. The binary command specifies whether the humanoid should climb upward or downward.

3.3 Bridging Sim-to-Real Perception Gap with Vision Foundation Model

A major challenge in deploying perceptive humanoid policies is the sim-to-real gap in depth observations [37, 38]. For ladder climbing, this challenge is particularly severe due to the thin and sparse nature of ladder rungs, where depth inaccuracies can easily lead to contact failures. In practice, we observe two major sources of discrepancy between simulated and real-world depth observations: (1) *unstructured noise* caused by sensor artifacts such as missing pixels and measurement noise, and (2) *structured perception gap* arising from ladder-specific geometric variations that are difficult to model accurately and exhaustively in simulation.

Unstructured Noise. Prior works [2, 30, 39] typically address depth sim-to-real gap through extensive randomization, which often requires careful tuning of many hyperparameters. Instead, we use Fast-FoundationStereo [14], a pretrained vision foundation model (VFM) for stereo depth estimation, to obtain real-world depth observations. As shown in Fig. 3, compared with raw depth sensing, VFM produces cleaner and more geometrically consistent depth observations with real-time inference speed, which is critical for humanoid ladder climbing. This substantially reduces the reliance on manually engineered depth randomization and improves robustness against sensor artifacts. During training, we still apply lightweight depth randomization consisting of Gaussian noise $\mathcal{N}(0, 0.005)$ and pixel dropout with probability 0.05. The resulting depth values are clipped to the range [0.1 m, 2 m].

Structured Perception Gap. To improve robustness against ladder-specific geometric variations such as handrails and supports, we introduce rung-focused masking (RFM), which masks non-rung regions during training with probability $p = 0.1$. As illustrated in Fig. 3, this encourages the policy to focus on task-relevant climbing structures rather than ladder-specific background geometry.

3.4 On-Ladder Manipulation via Dual-Agent Learning

Building on the learned climbing policy, we further enable manipulation under ladder constraints. This setting is challenging because the humanoid must maintain stable balance at the top of the

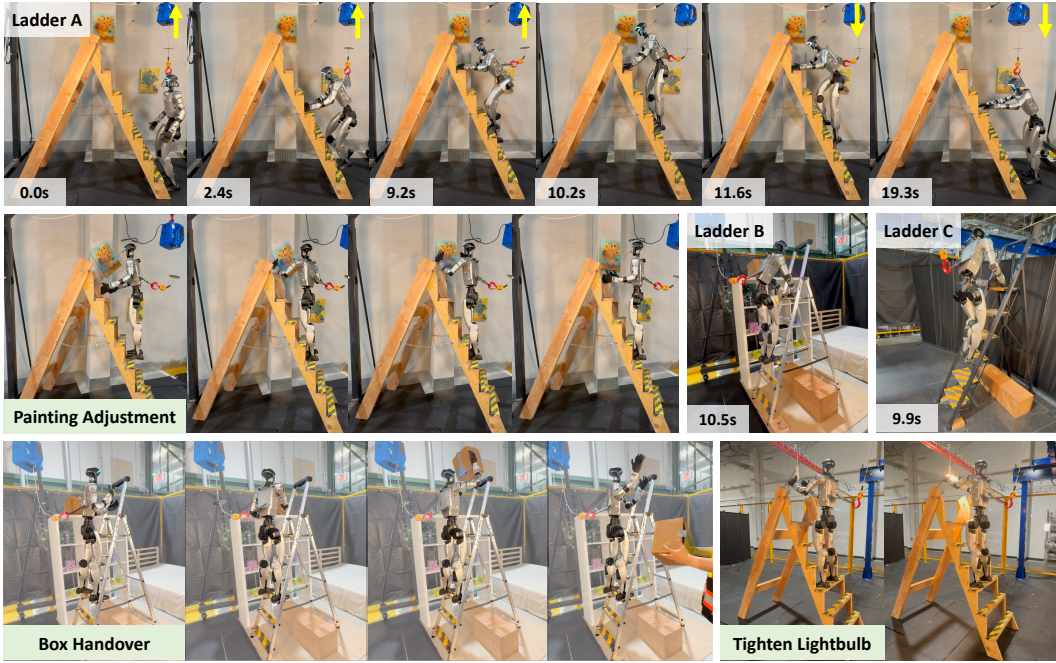


Figure 4: Real-world results of **LadderMan**. The humanoid performs zero-shot sim-to-real ladder climbing across diverse real-world ladders (A, B, and C) and executes various on-ladder manipulation tasks while maintaining stable balance. Additional video results are available on our [website](#).

ladder while executing manipulation tasks. Directly applying an off-the-shelf teleoperation policy often leads to unstable whole-body behaviors during policy transition and manipulation.

To address this issue, we train a separate manipulation policy $\pi^{\text{manip}}(s_t, g_t) = [\pi^u(s_t, g_t); \pi^l(s_t)]$, using a dual-agent formulation [15]. During training, the two agents use separate actor and critic networks. The lower-body agent π^l is responsible for maintaining stable balance, while the upper-body agent π^u tracks teleoperated manipulation targets. This decomposition decouples stabilization from manipulation, reducing interference between the two objectives and improving manipulation stability under ladder constraints. Refer to Appendix A.4 for more details.

Observation, Action, and Reward. Both agents receive whole-body proprioception s_t as input. π^u additionally receives teleoperated upper-body motion targets g_t , including arm, wrist, and waist poses. During training, upper-body target joint poses are randomly sampled from the AMASS [40] dataset, while during deployment they are generated through IK from teleoperation. The final action is formed by concatenating the outputs of the two agents: $a_t = [a_t^l; a_t^u]$. π^u is trained using the upper-body tracking reward $r_X^{\text{tracking},u}$ defined in Eq. 1. π^l is trained using the contact reward r^{contact} and stabilization reward r^{task} from Eq. 2, which encourage stable ladder contacts and maintain the pelvis near the target on-ladder pose. To increase manipulation workspace during training, upper-body collisions with the ladder are disabled in simulation.

4 Experimental Results

We evaluate LadderMan through extensive simulation and real-world experiments to answer:

- Q1:** Can LadderMan robustly perform climbing and manipulation across diverse ladders?
- Q2:** Does the learned dual-agent policy improve manipulation stability under ladder constraints?
- Q3:** Can hybrid motion tracking learn expert policies across diverse ladders from single reference?
- Q4:** Which system components are critical for robust sim-to-real transfer?

We use IsaacSim [41] for policy training and evaluation. All policies are deployed on a 1.3 m 29-DoF Unitree G1 humanoid without hardware modifications, using an onboard Intel RealSense D435i camera for perception and an onboard NVIDIA Jetson Orin for policy inference. We use only

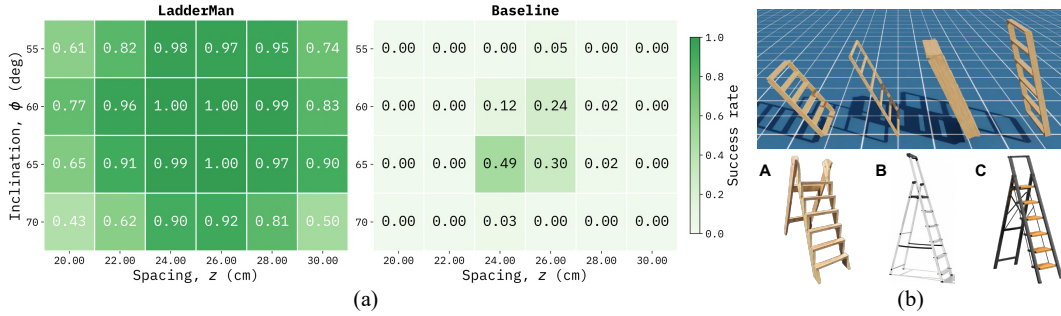


Figure 5: (a) Success rates of LadderMan and the blind motion tracking baseline across ladder configurations with varying inclination angles and rung spacings. (b) Representative ladder geometries used in simulation (*top*) and real-world evaluation (*bottom*).

a single reference motion for expert policy learning, collected using an OptiTrack motion capture system on a ladder with inclination angle $\theta = 65.5^\circ$ and rung spacing $s = 24.8$ cm. Additional experimental details are provided in Appendix B.

4.1 Real-World Evaluation

Experiment Setup. To answer Q1, we evaluate our climbing policy on three real-world ladders, including one custom-built ladder and two commercially available ladders with diverse geometries and materials, as shown in Fig. 5(b). We parameterize each ladder configuration as a tuple (ϕ, z, l, w) , where ϕ denotes inclination angle, z rung spacing, and l, w the rung dimensions. The evaluated ladders correspond to: A = $(65.5^\circ, 24.8\text{cm}, 60\text{cm}, 8\text{cm})$, B = $(69.2^\circ, 23.1\text{cm}, 28\text{cm}, 8\text{cm})$, C = $(64.8^\circ, 25.5\text{cm}, 30\text{cm}, 20\text{cm})$. The reference motion is collected on Ladder A.

Ladder Climbing Results. As shown in Fig. 4 and Tab. 2, LadderMan successfully climbs both upward and downward on Ladder A within 20 seconds in a zero-shot sim-to-real setting. We further demonstrate successful transfer across additional real-world ladders (Ladders B and C) with different geometries and materials. Despite significant variations across ladders, the learned policy maintains stable whole-body coordination and reliable multi-contact transitions throughout the process.

On-Ladder Manipulation Results. Building on the learned climbing policy, we further evaluate on-ladder manipulation in the real world to answer Q1. We use a PICO 4 Ultra VR headset for teleoperation, which provides whole-body motion streaming through a head-mounted display, handheld controllers, and two ankle-mounted trackers without requiring external motion capture systems. As shown in Fig. 4, the humanoid performs diverse on-ladder manipulation tasks, including adjusting painting, replacing light bulb, and box handover, while maintaining stable balance.



Figure 6: Failure case of the off-the-shelf teleoperation policy TWIST2 under ladder constraints.

while our dual-agent policy maintains stable balance throughout manipulation due to decoupling lower-body stabilization from upper-body manipulation.

To answer Q2, we compare the learned dual-agent policy against a state-of-the-art whole-body teleoperation policy TWIST2 [16]. After reaching the top of the ladder using the climbing policy, control is switched to the teleoperation policy. As shown in Fig. 6, TWIST2 frequently loses balance, while our dual-agent policy maintains stable balance throughout manipulation due to decoupling lower-body stabilization from upper-body manipulation.

Human Comparison. To evaluate practical performance, we compare the climbing speed of LadderMan against human measured for one male and one female subject across three different ladders. Humans achieve average climbing speed of ~ 3.2 s/rung, while LadderMan achieves a comparable speed of ~ 3.4 s/rung. Side-by-side comparison is provided in Appendix B.2.

Method	(55°, 20cm)	(65.5°, 24.8cm)*	(70°, 30cm)
DeepMimic [32]	✗	✓	✗
Ours	✓	✓	✓
w/o contact tracking	✗	✓	✗
w/o climb reward	✓	✓	✗

Table 1: Ablation study on expert policy learning methods across different ladder geometries. * denotes the ladder configuration where the reference motion was collected.

Method	Ladder A	Ladder B	Ladder C
Ours	9/10	6/10	7/10
w/o RL	2/10	0/10	0/10
w/o VFM	3/10	0/10	2/10
w/o RFM	0/10	0/10	0/10

Table 2: Real-world ablation of key components for robust sim-to-real transfer across 3 different ladders.

4.2 Simulation Evaluation

Experiment Setup. To further answer **Q1** quantitatively, we evaluate the learned climbing policy across ladders with varying rung spacing z and inclination angle ϕ . Specifically, we evaluate the policy on a discrete grid of ladder configurations with inclination angles $\phi \in \{55^\circ, 60^\circ, 65^\circ, 70^\circ\}$ and rung spacings $z \in \{20, 22, 24, 26, 28, 30\}$ cm, covering a broad range of real-world ladder geometries. For each ladder configuration, we further randomize rung size and shape (cuboid or cylindrical). Representative ladder geometries used in simulation are shown in Fig. 5(b). We report success rates over 100 evaluation episodes for each configuration.

Baseline. We compare the climbing policy of LadderMan against a blind motion tracking baseline trained without perceptual observations following the BeyondMimic [33] training setup.

As shown in Fig. 5(a), LadderMan demonstrates strong generalization across a wide range of ladder configurations, achieving over 95% success rate for ladders with spacing between 24cm and 28cm and inclination angles between 55° and 65°. Even under more extreme configurations near the boundary of the evaluation space, the policy still maintains an average success rate of $\sim 57\%$. In contrast, the blind motion tracking baseline achieves a peak success rate of only 49% near the reference configuration and rapidly fails as ladder geometry deviates from the training setup, yielding near-zero success rates for most unseen configurations.

4.3 Ablation Studies

Hybrid Motion Tracking Enables Expert Policy Learning Across Diverse Ladder Geometries.

To answer **Q3**, we ablate hybrid motion tracking during expert policy learning. We compare against: (1) standard DeepMimic-style motion tracking, (2) hybrid motion tracking without contact tracking, and (3) hybrid motion tracking without task rewards. We use only a single reference motion captured on a ladder with $(\phi, z) = (65.5^\circ, 24.8\text{cm})$. For each method, we independently train expert policies on three representative ladder configurations: $(55^\circ, 20\text{cm})$, $(65.5^\circ, 24.8\text{cm})$, $(70^\circ, 30\text{cm})$ corresponding to ladders with smaller, identical, and larger geometry relative to the reference setup. After training for 8,000 iterations, we evaluate whether each method can successfully learn a stable climbing policy for the target ladder configuration. The evaluation metric is binary, indicating whether successful climbing behavior emerges during training. As shown in Tab. 1, standard motion tracking only succeeds when training on the reference ladder geometry and fails to learn stable climbing behaviors on substantially different configurations. In contrast, hybrid motion tracking successfully learns climbing behaviors across all evaluated ladder geometries from a single reference motion. Removing either contact tracking or task rewards degrades learning performance, highlighting the importance of task-driven adaptation for robust whole-body climbing skill acquisition.

Key Components for Robust Sim-to-Real Transfer. To answer **Q4**, we ablate several key system components critical for real-world deployment: (1) hybrid DAgger and RL, (2) vision foundation model (VFM)-based depth perception, and (3) rung-focused masking (RFM). We evaluate all variants in the real world over 10 trials with randomized initial poses. As shown in Tab. 2, incorporating RL substantially improves robustness and recovery behaviors, resulting in significantly higher real-world success rates than pure DAgger. In addition, both VFM-based perception and rung-focused masking contribute significantly to sim-to-real perception robustness and generalization. In particular, masking task-irrelevant ladder structures improves perception consistency and reduces failures caused by unseen ladder rail geometries that are difficult to model accurately in simulation.

5 Conclusion and Limitation

In this work, we present LadderMan, a unified perceptive system for humanoid ladder climbing and on-ladder manipulation. We propose a scalable two-stage learning pipeline for learning a unified visuomotor climbing policy. To enable zero-shot real-world deployment, we leverage vision foundation models to bridge the sim-to-real gap in depth perception. Our results demonstrate robust climbing performance across diverse ladder geometries and stable on-ladder manipulation through a dual-agent teleoperation framework.

Despite these promising results, several limitations remain and open directions for future research. (1) Our climbing policy is primarily evaluated on inclined ladders up to 75° and does not yet extend to vertical ladders, which require larger contact forces, tighter stability margins, and potentially different climbing strategies. (2) Our manipulation capability is currently limited by the absence of dexterous end-effectors. Extending the system with dexterous hands is an important direction toward more practical real-world applications.

References

- [1] W. Hammer and U. Schmalz. Human behaviour when climbing ladders with varying inclinations. *Safety Science*, 15(1):21–38, 1992.
- [2] Y. Zhang, Y. Seo, J. Chen, Y. Yuan, K. Sreenath, P. Abbeel, C. Sferrazza, K. Liu, R. Duan, and G. Shi. Rpl: Learning robust humanoid perceptive locomotion on challenging terrains. *arXiv preprint arXiv:2602.03002*, 2026.
- [3] Z. Zhuang, S. Yao, and H. Zhao. Humanoid parkour learning. *arXiv preprint arXiv:2406.10759*, 2024.
- [4] H. Wang, Z. Wang, J. Ren, Q. Ben, T. Huang, W. Zhang, and J. Pang. Beamdojo: Learning agile humanoid locomotion on sparse footholds. In *Robotics: Science and Systems (RSS)*, 2025.
- [5] H. Yoneda, K. Sekiyama, Y. Hasegawa, and T. Fukuda. Vertical ladder climbing motion with posture control for multi-locomotion robot. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3579–3584, 2008.
- [6] J. Luo, Y. Zhang, K. Hauser, H. A. Park, M. Paldhe, C. S. G. Lee, M. Grey, M. Stilman, J. H. Oh, J. Lee, I. Kim, and P. Oh. Robust ladder-climbing with a humanoid robot with application to the darpa robotics challenge. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2792–2798, 2014.
- [7] J. Vaillant, A. Kheddar, H. Audren, F. Keith, S. Brossette, A. Escande, K. Kaneko, M. Morisawa, P. Gergondet, E. Yoshida, S. Kajita, and F. Kanehiro. Multi-contact vertical ladder climbing with an hrp-2 humanoid. *Autonomous Robots*, 2016.
- [8] T. Yoshiike, M. Kuroda, R. Ujino, H. Kaneko, H. Higuchi, S. Iwasaki, Y. Kanemoto, M. Asatani, and T. Koshiishi. Development of experimental legged robot for inspection and disaster response in plants. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4869–4876, 2017.
- [9] X. Sun, K. Hashimoto, T. Teramachi, T. Matsuzawa, S. Kimura, N. Sakai, S. Hayashi, Y. Yoshida, and A. Takanishi. Planning and control of stable ladder climbing motion for the four-limbed robot “warec-1”. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6547–6554, 2017.
- [10] A. A. Saputra, Y. Toda, N. Takesue, and N. Kubota. A novel capabilities of quadruped robot moving through vertical ladder without handrail support. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1448–1453, 2019.
- [11] X. Sun, K. Hashimoto, S. Hayashi, M. Okawara, T. Mastuzawa, and A. Takanishi. Stable vertical ladder climbing with rung recognition for a four-limbed robot. *Journal of Bionic Engineering*, 2021.
- [12] H. Weng, Y. Li, N. Sobanbabu, Z. Wang, Z. Luo, T. He, D. Ramanan, and G. Shi. Hdmi: Learning interactive humanoid whole-body control from human videos. *arXiv preprint arXiv:2509.16757*, 2025.
- [13] S. Zhao, Y. Ze, Y. Wang, C. K. Liu, P. Abbeel, G. Shi, and R. Duan. Resmimic: From general motion tracking to humanoid whole-body loco-manipulation via residual learning. *arXiv preprint arXiv:2510.05070*, 2025.
- [14] B. Wen, S. Dewan, and S. Birchfield. Fast-FoundationStereo: Real-time zero-shot stereo matching. *CVPR*, 2026.

- [15] Y. Zhang, Y. Yuan, P. Gurunath, I. Gupta, S. Omidshafiei, A.-a. Agha-mohammadi, M. Vazquez-Chanlatte, L. Pedersen, T. He, and G. Shi. Falcon: Learning force-adaptive humanoid loco-manipulation. *arXiv preprint arXiv:2505.06776*, 2025.
- [16] Y. Ze, S. Zhao, W. Wang, A. Kanazawa, R. Duan, P. Abbeel, G. Shi, J. Wu, and C. K. Liu. Twist2: Scalable, portable, and holistic humanoid data collection system. *arXiv preprint arXiv:2511.02832*, 2025.
- [17] Z. Luo, Y. Yuan, T. Wang, C. Li, S. Chen, F. Castañeda, Z.-A. Cao, J. Li, D. Minor, Q. Ben, X. Da, R. Ding, C. Hogg, L. Song, E. Lim, E. Jeong, T. He, H. Xue, W. Xiao, Z. Wang, S. Yuen, J. Kautz, Y. Chang, U. Iqbal, L. Fan, and Y. Zhu. Sonic: Supersizing motion tracking for natural humanoid whole-body control. *arXiv preprint arXiv:2511.07820*, 2025.
- [18] Defense Advanced Research Projects Agency (DARPA). Darpa robotics challenge (drc). <https://www.darpa.mil/research/programs/darpa-robotics-challenge>, 2026.
- [19] D. Vogel, R. Baines, J. Church, J. Lotzer, K. Werner, and M. Hutter. Robust ladder climbing with a quadrupedal robot. *arXiv preprint arXiv:2409.17731*, 2025.
- [20] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, 7(62):eabk2822, 2022.
- [21] J. Long, J. Ren, M. Shi, Z. Wang, T. Huang, P. Luo, and J. Pang. Learning humanoid locomotion with perceptive internal model. *arXiv preprint arXiv:2411.14386*, 2024.
- [22] D. Hoeller, N. Rudin, D. Sako, and M. Hutter. Anymal parkour: Learning agile navigation for quadrupedal robots. *Science Robotics*, 9(88):eadi7566, 2024.
- [23] J. He, C. Zhang, F. Jenelten, R. Grandia, M. Bächer, and M. Hutter. Attention-based map encoding for learning generalized legged locomotion. *Science Robotics*, 10(105):eadv3604, 2025.
- [24] P. Fankhauser, M. Bloesch, and M. Hutter. Probabilistic terrain mapping for mobile robots with uncertain localization. *IEEE Robotics and Automation Letters*, 3(4):3019–3026, 2018.
- [25] Q. Ben, B. Xu, K. Li, F. Jia, W. Zhang, J. Wang, J. Wang, D. Lin, and J. Pang. Gallant: Voxel grid-based humanoid locomotion and local-navigation across 3d constrained terrains. *arXiv preprint arXiv:2511.14625*, 2025.
- [26] A. Agarwal, A. Kumar, J. Malik, and D. Pathak. Legged locomotion in challenging terrains using egocentric vision. In *Proceedings of The 6th Conference on Robot Learning*, volume 205 of *Proceedings of Machine Learning Research*, pages 403–415. PMLR, 14–18 Dec 2023.
- [27] X. Cheng, K. Shi, A. Agarwal, and D. Pathak. Extreme parkour with legged robots. *arXiv preprint arXiv:2309.14341*, 2023.
- [28] Z. Zhuang, Z. Fu, J. Wang, C. Atkeson, S. Schwertfeger, C. Finn, and H. Zhao. Robot parkour learning. *arXiv preprint arXiv:2309.05665*, 2023.
- [29] R. Yang, M. Zhang, N. Hansen, H. Xu, and X. Wang. Learning vision-guided quadrupedal locomotion end-to-end with cross-modal transformers. *arXiv preprint arXiv:2107.03996*, 2021.
- [30] Z. Wu, X. Huang, L. Yang, Y. Zhang, X. Chen, P. Abbeel, R. Duan, A. Kanazawa, C. Sferrazza, G. Shi, and C. K. Liu. Perceptive humanoid parkour: Chaining dynamic human skills via motion matching. *arXiv preprint arXiv:2602.15827*, 2026.
- [31] Z. Zhang, K. Wen, M. Xu, J. He, C. Li, T. Miki, C. Schwarke, C. Zhang, X. B. Peng, and M. Hutter. Learning whole-body humanoid locomotion via motion generation and motion tracking. *arXiv preprint arXiv:2604.17335*, 2026.

- [32] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Trans. Graph.*, 37(4):143:1–143:14, 2018.
- [33] Q. Liao, T. E. Truong, X. Huang, Y. Gao, G. Tevet, K. Sreenath, and C. K. Liu. Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv preprint arXiv:2508.08241*, 2025.
- [34] S. Ross, G. Gordon, and D. Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pages 627–635. PMLR, 2011.
- [35] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- [36] M. Macklin. Warp: A high-performance python framework for gpu simulation and graphics, Mar. 2022. URL <https://github.com/NVIDIA/warp>. NVIDIA GPU Technology Conference (GTC).
- [37] S. Zhu, Z. Zhuang, M. Zhao, K.-Y. Lee, and H. Zhao. Hiking in the wild: A scalable perceptive parkour framework for humanoids. *arXiv preprint arXiv:2601.07718*, 2026.
- [38] Z. Zhuang, S. Zhu, M. Zhao, and H. Zhao. Deep whole-body parkour. *arXiv preprint arXiv:2601.07701*, 2026.
- [39] N. Rudin, J. He, J. Aurand, and M. Hutter. Parkour in the wild: Learning a general and extensible agile locomotion policy using multi-expert distillation and rl fine-tuning. *arXiv preprint arXiv:2505.11164*, 2025.
- [40] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black. AMASS: Archive of motion capture as surface shapes. In *International Conference on Computer Vision*, pages 5442–5451, 2019.
- [41] NVIDIA. Isaac Sim. URL <https://github.com/isaac-sim/IsaacSim>.
- [42] L. Yang, X. Huang, Z. Wu, A. Kanazawa, P. Abbeel, C. Sferrazza, C. K. Liu, R. Duan, and G. Shi. Omniretarget: Interaction-preserving data generation for humanoid whole-body locomanipulation and scene interaction. *arXiv preprint arXiv:2509.26633*, 2025.

A Method

A.1 Learning Multiple Expert Policies from a Single Reference Motion

In this stage, our goal is to learn state-based expert policies $\pi_{\phi,z}^{\text{expert}}$ corresponding to different ladder configurations (ϕ, z) , where ϕ denotes the ladder inclination and z the rung spacing. We formulate the problem as a goal-conditioned reinforcement learning problem within a Markov Decision Process (MDP)

$$\mathcal{M} = \langle \mathcal{O}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle, \quad (4)$$

where \mathcal{O} is the observation space, \mathcal{A} the action space, \mathcal{T} the transition dynamics, \mathcal{R} the reward function, and γ the discount factor. At time step t , the policy π takes observation $o_t \in \mathcal{O}$ as input and outputs action a_t , which specifies target joint positions executed through a PD controller. The reward is defined as

$$r_t = \mathcal{R}(o_t, a_t), \quad (5)$$

and the training objective is to maximize the expected cumulative discounted reward:

$$\mathbb{E} \left[\sum_{t=1}^T \gamma^{t-1} r_t \right]. \quad (6)$$

Reference Motion Collection. Hybrid motion tracking requires only a single reference motion. We collect this motion using an OptiTrack motion capture system on a ladder with inclination angle $\phi = 65.5^\circ$ and rung spacing $z = 24.8$ cm. The reference motion consists of a complete climbing sequence and is used for learning all expert policies across different ladder configurations. In addition to the motion trajectory, we manually annotate contact events for the four end-effectors (both hands and feet) once on the reference motion. These annotations are used to construct the reference contact indicator \hat{c} and the ladder-centric contact targets used in the contact tracking reward. Despite using only a single reference motion, our hybrid motion tracking formulation enables successful expert policy learning across a wide range of ladder geometries.

Observation Space and Reward. The expert policy $\pi_{\phi,z}^{\text{expert}}$ takes (i) reference motion, (ii) proprioception, and (iii) privileged state information as observations, as shown in Table 3. The reward consists of three categories: motion tracking, contact tracking, and regularization rewards, as summarized in Table 4. Motion tracking rewards follow the exponential form

$$r = \omega \exp\left(-\frac{E}{\tau}\right), \quad (7)$$

where ω and τ denote the reward weight and temperature, respectively. Specifically, the contact tracking reward is defined as

$$r^{\text{contact}} = \omega \hat{c} \exp\left(-\frac{\|P - \hat{P}\|_2}{\tau}\right) + (1 - \hat{c}), \quad (8)$$

where \hat{c} denotes the reference contact indicator, P the current end-effector position, and \hat{P} the target contact position. The reference contact indicator is manually annotated once for the reference motion. The target contact position is automatically generated from the ladder geometry by assigning predefined contact locations on the target ladder rung for each hand and foot.

Observation Group	Description	Equation	Observation Group	Description	Equation
Reference Motion	reference body local pose	\hat{p}_t	Privilege State	ladder relative pose	$p_t^{\text{root}} - p_t^{\text{ladder}}$
	reference joint position	\hat{q}_t		body local pose	p_t
	reference motion phase	t/T		body velocity	\dot{p}_t
Proprioception	root angle velocity	w_t		root velocity	v_t
	root orientation	θ_t		body pose error	$p_t - \hat{p}_t$
	joint position	$q_{t-8:t}$		body velocity error	$\dot{p}_t - \dot{\hat{p}}_t$
	previous actions	$a_{t-3:t-1}$	joint position error	$q_t - \hat{q}_t$	

Table 3: Observation space of the expert climbing policy.

Reward		Weight	Temperature	Reward	Weight
<i>Motion Tracking</i>				<i>Regularization Penalties</i>	
Body Local Pose	Upper Body	1.0	1.0	Action Rate L2	0.1
	Lower Body	2.0	0.5	Joint Position Limits L1	10.0
Body Velocity	Upper Body	0.5	5.0	Joint Velocity Penalty	$5e^{-4}$
	Lower Body	1.0	2.5	Joint Torque Limits L1	0.01
Joint Tracking	Upper Body	0.5	0.5	Feet Impact Force L2	1.0
	Lower Body	1.0	0.25	Feet Slip	0.5
Root Pose	-	1.0	0.5	Feet Air Time	5.0
<i>Contact Tracking</i>					
Hands Contact Position		0.3	0.3		
Feet Contact Position		0.3	0.3		

Table 4: Reward terms used for expert policy learning.

Domain Randomization	Sampling Distribution	Domain Randomization	Sampling Distribution
<i>Physical parameters</i>		<i>Reference initialization</i>	
Link mass	$\text{mass} \sim \mathcal{U}[0.9, 1.1] \times \text{mass}$	Reference root position [m]	$\Delta xyz \sim \mathcal{U}[-0.01, 0.01]$
Torso COM offset [m]	$\Delta xyz \sim \mathcal{U}[-0.02, 0.02]$	Reference root rotation [rad]	$\Delta rpy \sim \mathcal{U}[-0.05, 0.05]$
Joint position offset [rad]	$\Delta q \sim \mathcal{U}[-0.01, 0.01]$	Reference joint position [rad]	$\Delta q \sim \mathcal{U}[-0.05, 0.05]$
Static friction coefficient	$\mu_{\text{static}} \sim \mathcal{U}[0.3, 7.2]$	<i>Ladder Configuration</i>	
Dynamic friction coefficient	$\mu_{\text{dynamic}} \sim \mathcal{U}[0.3, 3.6]$	Ladder inclination [°]	$\Delta \phi \sim \mathcal{U}[-1.0, 1.0]$
Restitution coefficient	$e_{\text{rest}} \sim \mathcal{U}[0.0, 0.2]$	Ladder rung spacing [m]	$\Delta z \sim \mathcal{U}[-0.01, 0.01]$

Table 5: Domain randomization parameters used during expert policy training.

Domain Randomization and Termination. As shown in Tab. 5, in addition to physical parameters and reference-state randomization, we further randomize ladder configurations by perturbing both ladder inclination ϕ and rung spacing z . Training episodes terminate under three conditions: (i) root global pose error exceeds 0.5 m or 1.2 rad for 25 consecutive simulation steps; (ii) body local pose error exceeds 0.5 m or 1.2 rad for 25 consecutive simulation steps; (iii) required ladder contacts are lost for more than 30 consecutive simulation steps.

Training Details. We use PPO [35] to train the expert policies. Training is performed in IsaacSim [41] with 4096 parallel environments on a single NVIDIA L40S GPU for 8,000 epochs. The actor and critic networks use MLP architectures of [512, 256, 256] and [512, 256, 128], respectively.

A.2 Unified Climbing Policy via Hybrid Imitation and Reinforcement Learning

In this stage, our goal is to learn a unified visuomotor climbing policy π^{visual} that generalizes across diverse ladder geometries. The policy takes climbing commands, proprioception (as defined in Tab. 3), and depth observations as input.

For depth rendering, we use NVIDIA Warp [36] instead of the built-in IsaacSim [41] renderer to achieve higher rendering throughput during training [2]. The depth image resolution is 60×80 . To improve robustness, we apply minimal depth randomization consisting of Gaussian noise $\mathcal{N}(0, 0.005)$ and pixel dropout with probability 0.05. The resulting depth values are clipped to the range [0.1 m, 2 m]. Since the onboard camera of Unitree G1 is not rigidly mounted due to hardware constraints, we additionally randomize camera extrinsics with translational perturbations in $[-0.01 \text{ m}, 0.01 \text{ m}]$ and rotational perturbations in $[-1.5^\circ, 1.5^\circ]$.

The actor and critic networks use MLP architectures of [2048, 1024, 512, 256, 128] and [512, 256, 128], respectively. The depth encoder consists of a three-layer convolutional network with channel dimensions [8, 8, 8] and kernel size 5.

A.3 Bridging the Sim-to-Real Gap

The unified visuomotor climbing policy faces two primary sources of sim-to-real discrepancy: perception and dynamics. To mitigate the perception gap, we apply depth randomization during training

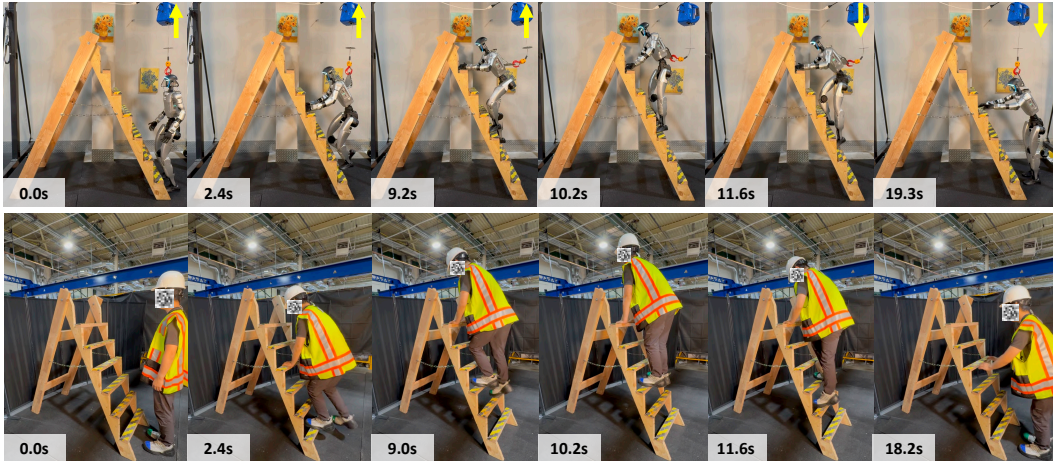


Figure 7: Side-by-side comparison between a human and LadderMan performing the same ladder climbing task. LadderMan achieves human-comparable climbing speed.

and further leverage FastFoundationStereo [14] to obtain cleaner and more geometrically consistent depth observations in the real world. On the dynamics side, we identify collision geometry as a critical yet often overlooked factor for successful sim-to-real transfer. This effect is particularly pronounced for ladder climbing, which involves frequent whole-body contacts and requires precise interaction between the robot and ladder.

Collision Geometry. We investigate several publicly available Unitree G1 models used in recent humanoid projects and observe substantial differences in their collision geometry. In particular, as shown in Fig. 8, the collision models used by TWIST2 [16] and HDMI [12] are heavily simplified, making accurate ladder contact simulation difficult. We therefore adopt the collision geometry from OmniRetarget [42] as our starting point due to its substantially higher geometric fidelity. However, we further observe that several body parts, particularly the shins, hands, and feet, are still approximated using simple collision primitives. While such approximations are often sufficient for locomotion on flat terrain, they introduce noticeable sim-to-real discrepancies for ladder climbing, where small contact errors can accumulate and lead to failure. These observations highlight the importance of high-fidelity collision geometry for contact-rich humanoid tasks and motivate the collision model used in LadderMan.

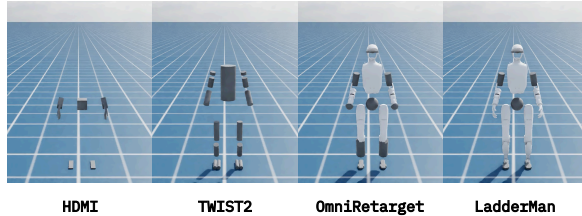


Figure 8: Comparison of Unitree G1 collision geometries used in different humanoid projects.

A.4 On-Ladder Manipulation via Dual-Agent Learning

In this stage, our goal is to train a manipulation policy $\pi^{\text{manip}}(s_t, g_t) = [\pi^u(s_t, g_t); \pi^l(s_t)]$, using a dual-agent formulation that maintains stable balance on the ladder while tracking teleoperated manipulation targets. The two agents use separate actor and critic networks. The actor and critic networks use MLP architectures of [512, 256, 256] and [512, 256, 128], respectively.

B Experimental Results

B.1 Experiment Setup

Climbing and Manipulation Policy. Both policies are deployed on the onboard NVIDIA Jetson Orin computer with a control frequency of 50 Hz. To generate the teleoperation targets g_t during

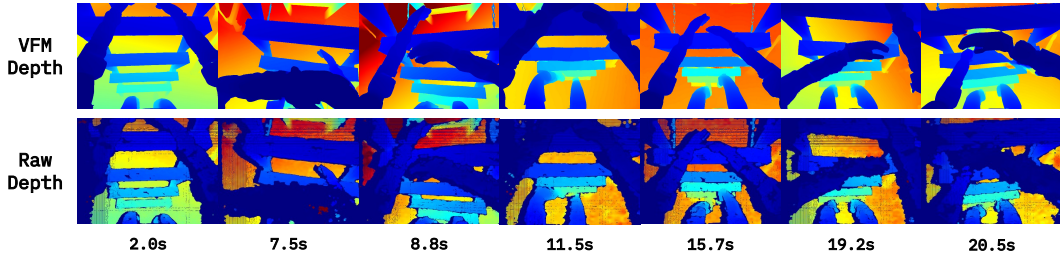


Figure 9: Comparison between raw depth observations and depth predictions generated by Fast FoundationStereo. The foundation model significantly reduces sensor artifacts and improves geometric consistency, facilitating sim-to-real transfer.

deployment, we use a PICO 4 Ultra VR headset. The system provides whole-body motion streaming through a head-mounted display, handheld controllers, and two ankle-mounted trackers, without requiring external motion capture systems.

Vision Foundation Model. To obtain cleaner depth observations in the real world, we use Fast FoundationStereo [14] to process stereo grayscale images captured by the onboard Intel RealSense D435i camera. During data acquisition, the infrared emitter and automatic exposure control are disabled. Real-time inference is performed on an NVIDIA RTX 4090 GPU during deployment.

B.2 Additional Results

Hybrid Motion Tracking. We provide a side-by-side comparison between the collected reference motion and the learned policy execution on a ladder configuration with $(\theta, s) = (55^\circ, 20 \text{ cm})$, as shown in Fig. 10. Despite being trained from a single reference motion collected on a different ladder configuration, the learned policy successfully adapts the contact locations of both hands and feet while preserving the overall climbing pattern. This result qualitatively demonstrates the ability of hybrid motion tracking to generalize across ladder geometries without requiring additional reference motions.

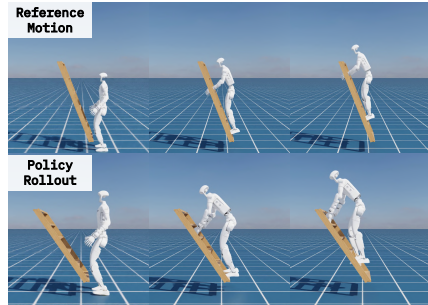


Figure 10: Comparison between the reference motion and the learned climbing behavior on a ladder different from that used during motion capture.

Human Comparison. As shown in Fig. 7, we provide a side-by-side comparison between a human and the humanoid robot performing the same ladder climbing task. For a complete traversal consisting of climbing three rungs up and three rungs down, the robot requires

approximately 19.3 s, while the human completes the task in approximately 18.2 s. Although the robot adopts a more conservative climbing strategy for safety and robustness, its climbing speed is already comparable to that of a human operator. This result suggests that LadderMan achieves practical climbing efficiency while maintaining reliable whole-body coordination.

Depth Visualization. As shown in Fig. 9, we provide side-by-side visualizations of the raw depth observations and the corresponding depth predictions generated by Fast FoundationStereo [14]. Compared with raw depth observations, the predicted depth maps exhibit significantly reduced sensor artifacts and improved geometric consistency, particularly around ladder rungs and contact regions. This qualitative result supports our observation that vision foundation models reduce the need for extensive manually engineered depth randomization during sim-to-real transfer.