ResTrack: Residual Joint Refinement for Improving Whole-Body Tracking in Humanoid Loco-Manipulation

Yuming Huang*†, Zhidan Liu‡†, Yixiong Lin§†, Di Guo*¶ and Xiaojian Ma†¶
*School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, China
†Beijing Institute for General Artificial Intelligence (BIGAI)
‡Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China
§School of Computer Science and Technology, Harbin Institute of Technology, China
¶Corresponding Author: Di Guo (guodi.gd@gmail.com) and Xiaojian Ma (maxiaojian@bigai.ai)

***Abstract*—Large-scale pretrained humanoid tracking policies provide strong priors for locomotion and whole-body coordination, but loco-manipulation motions can still produce residual errors that limit tracking fidelity. Although directly fine-tuning the full controller may reduce these errors, it can also weaken the robustness acquired during large-scale pretraining. We present ResTrack, a residual joint refinement framework that augments a frozen pretrained humanoid controller with a lightweight trainable residual policy in joint-action space. Conditioned on the current tracking state together with the nominal base-policy output, the residual policy learns to compensate for the errors left by the pretrained controller while preserving its stability prior. We evaluate ResTrack on a daily mobile-manipulation benchmark and show that it consistently reduces whole-body tracking errors relative to the pretrained controller and a direct fine-tuning baseline. These results indicate that residual refinement provides an effective and low-overhead way to improve whole-body tracking performance on top of pretrained humanoid controllers.**

I. Introduction

Recent advances in humanoid control have enabled large-scale policies capable of robust whole-body motion tracking across diverse locomotion and manipulation behaviors. By leveraging large motion datasets and reinforcement learning, these controllers [2, 6] can achieve stable coordination of the entire body and exhibit strong generalization across motion sequences. Such pretrained policies have become an increasingly important foundation for humanoid loco-manipulation systems [14].

Even with these advances, there remains substantial room to improve whole-body tracking fidelity. On loco-manipulation motions, pretrained controllers may still accumulate noticeable anchor drift or local pose mismatch, which lowers replay quality and can degrade downstream loco-manipulation behavior [1]. A straightforward remedy is to further fine-tune the pretrained controller on the target data. However, directly updating all policy parameters may weaken the locomotion and coordination priors learned during large-scale pretraining, making optimization less stable and less predictable, as shown in Fig. 1.

Residual control provides a practical alternative [5, 4, 14]. Rather than replacing the pretrained controller, a residual policy learns corrective actions that refine the behavior of a frozen base policy. This design preserves the robustness of the pretrained controller while allocating learning capacity specifically to the remaining tracking errors. In this way, the residual module acts as a lightweight refinement layer whose role is to improve tracking fidelity rather than relearn the entire control problem.

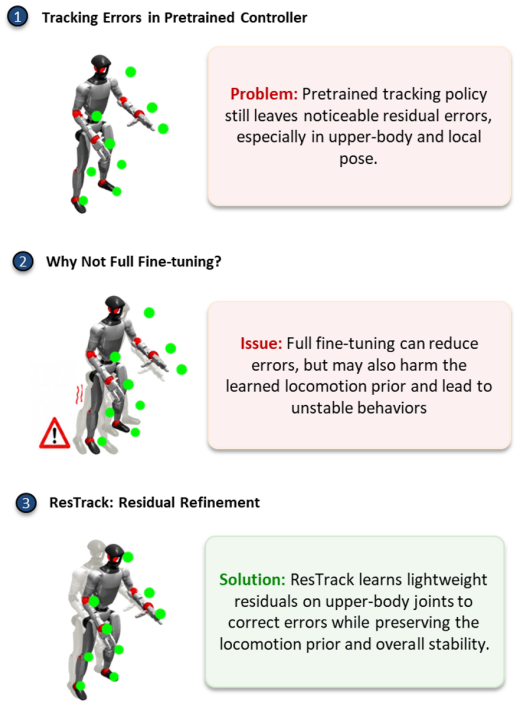

Fig. 1. Motivation of ResTrack. Pretrained humanoid tracking policies can still exhibit residual tracking errors on loco-manipulation motions. Direct fine-tuning may reduce these errors but can also weaken the locomotion prior learned during pretraining. ResTrack addresses this issue by applying local residual refinement primarily to upper-body joints while preserving the frozen base policy.

In this paper, we introduce **ResTrack**, a residual joint refinement framework for humanoid motion tracking. ResTrack augments a frozen pretrained controller [6] with a trainable residual module in joint-action space. The residual policy reasons about the current tracking state together with the

nominal output of the base controller, and predicts corrective actions that improve the executed motion while respecting the stability prior of the frozen policy. The resulting system is simple to integrate with existing pretrained humanoid trackers and requires no redesign of the underlying base controller.

We evaluate the proposed approach on an in-house daily mobile-manipulation benchmark comprising table-approach and object-grasping behaviors. On this benchmark, ResTrack consistently outperforms both the pretrained controller and a direct fine-tuning baseline. These results indicate that residual refinement is an effective way to further improve whole-body tracking quality on top of pretrained humanoid controllers.

The main contributions of this work are:

- A residual joint refinement framework that improves whole-body motion tracking on top of a frozen pretrained humanoid controller.
- A base-aware residual observation design that focuses learning on the tracking errors that remain after the nominal pretrained action is produced.
- Experimental results on a daily mobile-manipulation benchmark demonstrating consistent improvements in whole-body tracking quality.

II. Related Work

A. Humanoid Loco-Manipulation

Humanoid loco-manipulation aims to achieve coordinated locomotion and manipulation behaviors while maintaining whole-body balance and motion naturalness. Recent reinforcement learning and motion imitation approaches have demonstrated remarkable progress in enabling humanoids to perform complex tasks involving simultaneous locomotion and object interaction.

A common paradigm is to learn a robust motion prior from large-scale human motion datasets and subsequently adapt the policy to downstream manipulation tasks. DeepMimic [7] established a representative formulation in which a physics-based policy is trained with reinforcement learning to track reference motion clips from motion capture, using imitation rewards defined over pose, velocity, and end-effector consistency, optionally together with task rewards. Subsequent works extend this paradigm in several directions: AMP [8] replaces manually designed imitation terms with adversarially learned motion priors, while more recent general motion-tracking systems [2, 12, 13] scale training to larger and more diverse whole-body motion corpora and improve the robustness and coverage of a unified controller. SONIC [6] further scales humanoid motion training to more than 100 million frames and demonstrates that a single policy can acquire robust and transferable motion skills across diverse behaviors. Building upon these motion priors, recent loco-manipulation frameworks have incorporated task-specific objectives to achieve accurate object interaction while preserving stable locomotion.

Despite these advances, maintaining high whole-body tracking quality on difference motions remains challenging because locomotion, balance, and manipulation constraints interact strongly. ResMimic [14] trains a general motion tracking policy [2] and adapts it to loco-manipulation tasks through task-specific refinement. Our work is complementary to these efforts: rather than redesigning the controller hierarchy, we study how to further improve whole-body tracking fidelity of an already strong pretrained controller through a residual refinement layer.

B. Residual Control for Motion Tracking

Residual learning has emerged as an effective strategy for adapting pretrained controllers by learning corrective actions on top of them, thereby preserving their underlying control priors. Early residual reinforcement learning methods combine a conventional controller with a learned residual policy, allowing the residual component to compensate for modeling errors and task-specific deficiencies [5, 9].

This idea has recently been adopted in humanoid control and motion tracking. HiWET [1] uses a hierarchical residual framework to optimize world-frame end-effector trajectories while maintaining locomotion stability through a low-level controller. RobotDancing [10] predicts residual joint targets to compensate for dynamics mismatch during long-horizon humanoid motion imitation. ASAP [3] and MOSAIC [11] perform rapid residual adaptation on top of a general motion tracking policy to improve sim-to-real transfer. SteadyTray [4] and ResMimic [14] further demonstrate that task-specific residual policies can effectively refine a pretrained motion prior for object manipulation tasks. Our work follows this direction, but targets a different regime: improving the whole-body tracking fidelity of an already strong frozen pretrained controller through a lightweight joint-space residual module. The experiment reported in this paper supports the effectiveness of this formulation on a daily mobile-manipulation benchmark.

III. Method

A. Overview

The objective of this work is to improve whole-body tracking performance while preserving the robustness of a pretrained humanoid tracking controller. To achieve this objective, we introduce **ResTrack**, a residual joint refinement framework that augments a frozen pretrained base policy with a trainable residual policy. Figure 2 illustrates the overall architecture. During execution, the base controller produces a nominal full-body action, while the residual policy predicts a corrective action designed to reduce the remaining tracking error. Because the base controller remains frozen, the residual policy does not need to relearn stable locomotion or global coordination from scratch. Instead, it can focus on refining the executed motion under the strong prior supplied by the pretrained controller.

B. Joint Residual Control

At each time step, the frozen base policy produces a nominal action, and the residual policy predicts a corrective action to refine the executed motion. The base and residual policies

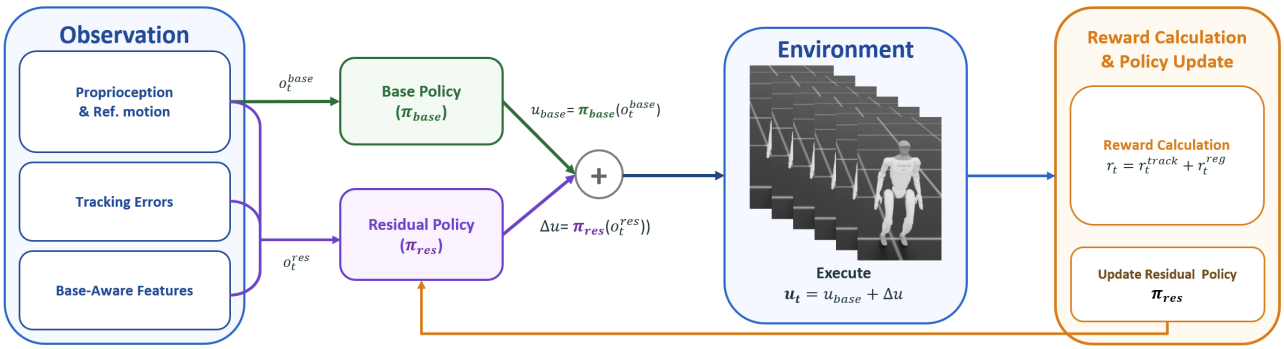

Fig. 2. Overview of the ResTrack framework. A frozen pretrained base controller produces nominal whole-body actions, while a lightweight residual policy predicts joint-space corrections to improve tracking fidelity.

receive observations o_t^{base} and o_t^{res}, respectively, where o_t^{res} is tailored to residual correction.

The nominal action generated by the frozen base policy is defined as

$$u_t^{base} = \pi_{base}(o_t^{base}), \tag{1}$$

Based on the residual-specific observation, the residual policy predicts an additive correction

$$\Delta u_t = \pi_\theta(o_t^{res}), \tag{2}$$

where θ denotes the trainable parameters of the residual policy. The final executed action is obtained by combining the nominal action with the residual correction

$$u_t = u_t^{base} + \Delta u_t, \tag{3}$$

To confine corrective behavior to task-relevant joints, the residual output is applied only to a selected subset of action dimensions. In our implementation, residual corrections are applied to upper-body joints, comprising the waist and arm joints (shoulders, elbows, and wrists), while the lower body remains governed by the pretrained controller. This design preserves the locomotion prior encoded in the base policy while enabling targeted refinement of the executed motion.

C. Residual Observation Model

Effective residual learning requires access to information that describes both the current tracking state and the behavior of the base controller. The base and residual observations are derived from the same underlying simulator state, but they are not identical in content. In particular, the residual observation augments the core control state with tracking-error and base-aware features:

$$o_t^{res} = \left[o_t^{prop}; b_t^{base}; e_t^{track}\right]. \tag{4}$$

The first component, o_t^{prop}, contains proprioceptive information such as gravity direction, base angular velocity, joint positions, joint velocities, and previous actions. For clarity, we first define the base-aware term b_t^{base}, which exposes the behavior of the pretrained controller to the residual policy:

$$b_t^{base} = \left[u_t^{base}; q_t^* - q_t^{base}\right], \tag{5}$$

where q_t^* is the reference joint configuration and q_t^{base} denotes the nominal joint-space target implied by the base controller. We then define the tracking-error term e_t^{track}, which summarizes the current discrepancy between the nominal motion and the reference motion over selected tracking entities:

$$e_t^{track} = \left[\{\delta p_t^{(k)}\}_{k=1}^K; \{\delta v_t^{(k)}\}_{k=1}^K\right], \tag{6}$$

where

$$\delta p_t^{(k)} = p_t^{*(k)} - p_t^{(k)}, \qquad \delta v_t^{(k)} = v_t^{*(k)} - v_t^{(k)}, \tag{7}$$

with $p_t^{(k)}$ and $v_t^{(k)}$ denoting the position and velocity of the k-th tracking point, and $p_t^{*(k)}$, $v_t^{*(k)}$ denoting their reference values.

This observation design allows the residual policy to reason explicitly about the correction that remains necessary after the base controller has produced its action. Rather than learning the entire tracking problem, the residual module learns only the refinement required to reduce the remaining error.

D. Training Objective

The residual policy is trained with a tracking objective together with regularization terms. The overall per-step reward is written as

$$r_t = r_t^{track} + r_t^{reg}, \tag{8}$$

where r_t^{track} represents motion-tracking rewards and r_t^{reg} contains regularization terms. The tracking portion aggregates anchor-position, anchor-orientation, body-pose, body-velocity, and local-alignment rewards, while the regularization portion penalizes unstable or physically implausible behavior.

This formulation makes the role of the residual module explicit: the pretrained controller supplies a stable nominal action prior, and the residual policy is trained only to produce corrections that improve whole-body tracking quality under that fixed prior. The reward weights used in training are summarized in Table I.

TABLE I
REWARD TERMS AND WEIGHTS USED FOR RESIDUAL-POLICY TRAINING.

Term	Weight
Tracking rewards	
Anchor position	0.5
Anchor orientation	0.5
Relative body position	1.0
Relative body orientation	1.0
Body linear-velocity	1.0
Body angular-velocity	1.0
Local five-point alignment	2.0
Local wrist position	1.0
Local wrist orientation	1.0
Regularization and penalties	
Action-rate	-0.1
Joint-limit	-10.0
Undesired-contact	-0.1
Anti-shake angular-velocity	-0.005
Ankle-acceleration	-2.5e-7

TABLE II
RESULTS ON THE DAILY MOBILE-MANIPULATION BENCHMARK. LOWER IS BETTER FOR ALL METRICS.

Metric	SONIC	Fine-tune	ResTrack
$E_{\text{Anchor}} \downarrow$	0.517	0.672	**0.500**
$E_{\text{MPJPE-L}} \downarrow$	31.3	33.514	**27.2**

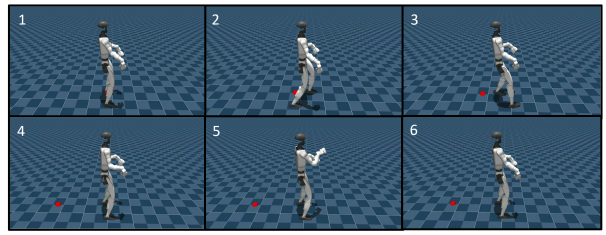

Fig. 3. Qualitative MuJoCo snapshots from the daily mobile-manipulation benchmark. From left to right, the humanoid approaches the table, aligns its upper body with the tabletop region, reaches forward, executes the grasping motion, completes the grasp posture, and transitions into the final phase of the sequence.

IV. EXPERIMENTS

A. Experimental Setup

a) Baselines: We compare three settings: (1) the pretrained SONIC [6] tracking policy, (2) direct fine-tuning from the same pretrained checkpoint, and (3) the proposed residual policy. For our method, the SONIC controller is frozen while a lightweight residual MLP with hidden dimensions $(512, 512, 256)$ predicts 29-dimensional action corrections that are added to the base actions with $\lambda = 1.0$. Under the `upper_body` setting, residual actions are masked so that only the 17 upper-body joints (waist and both arms) receive corrections.

b) Training Configuration: The residual policy is trained using PPO with the tracking objective described in Section III and uses the same training dataset as SONIC. Training is conducted on a single NVIDIA RTX 5090 GPU and typically requires approximately two to three days of optimization. Direct fine-tuning uses the same pretrained checkpoint, reward function, and PPO hyperparameters as the proposed method.

c) Evaluation Dataset: We evaluate the proposed framework on a held-out daily mobile-manipulation dataset that does not appear during training. The daily mobile-manipulation benchmark contains 30 motion-capture sequences in which a subject approaches a table, grasps a banana, and walks away. This dataset emphasizes task-oriented loco-manipulation motion and provides a focused testbed for evaluating tracking quality on daily manipulation behaviors.

d) Metrics: We report the following metrics: Anchor Position Error (m), denoted by E_{Anchor}, computed as the Euclidean distance between the robot anchor position and the reference anchor position in the world frame, thereby quantifying global tracking drift; Local Mean Per-Joint Position Error (MPJPE-L, mm), denoted by $E_{\text{MPJPE-L}}$, computed as the average Euclidean joint-position error over the tracked body links in the root-relative frame, thereby quantifying the fidelity of the reproduced whole-body pose in the robot's local frame. These metrics characterize complementary aspects of tracking quality: the former captures global anchor accuracy, while the latter measures the fidelity of the reproduced whole-body pose.

B. Main Results

Table II presents the results on the daily mobile-manipulation benchmark. ResTrack achieves the lowest error on both metrics. Relative to SONIC, it reduces E_{Anchor} from 0.517 to 0.500 and lowers $E_{\text{MPJPE-L}}$ from 31.3 mm to 27.2 mm. The gains are larger when compared with direct fine-tuning, which yields 0.672 and 33.514 mm on the two metrics, respectively. These results indicate that residual refinement improves both global anchor tracking and local whole-body pose fidelity on task-oriented mobile-manipulation motions, while direct fine-tuning does not preserve the pretrained tracking prior as effectively on this benchmark.

Figure 3 provides a qualitative view of the target behavior on the daily mobile-manipulation benchmark. The six MuJoCo snapshots illustrate a representative sequence in which the humanoid approaches the table, aligns the upper body with the tabletop region, and executes a grasping motion. This sequence highlights the coordinated loco-manipulation behavior considered in this benchmark, where stable locomotion and accurate upper-body motion must be achieved simultaneously.

V. CONCLUSION

We presented ResTrack, a residual joint refinement framework for enhancing whole-body tracking quality in humanoid loco-manipulation. Built on top of a frozen pretrained controller, ResTrack learns joint-space residual corrections that refine the executed motion while preserving the robustness of the original motion prior. Experiments on a daily mobile-manipulation benchmark show consistent gains over both the pretrained controller and a direct fine-tuning baseline in Anchor Position Error and MPJPE-L. Overall, the results suggest that residual refinement is a simple and effective way to further improve pretrained humanoid tracking policies.

References

[1] Zhanxiang Cao, Liyun Yan, Yang Zhang, Sirui Chen, Jianming Ma, Tianyue Zhan, Shengcheng Fu, Yufei Jia, Cewu Lu, and Yue Gao. Hiwet: Hierarchical world-frame end-effector tracking for long-horizon humanoid loco-manipulation. *arXiv preprint arXiv:2602.06341*, 2026.

[2] Zixuan Chen, Mazeyu Ji, Xuxin Cheng, Xuanbin Peng, Xue Bin Peng, and Xiaolong Wang. Gmt: General motion tracking for humanoid whole-body control. *arXiv preprint arXiv:2506.14770*, 2025.

[3] Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbab, Chaoyi Pan, et al. Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills. *arXiv preprint arXiv:2502.01143*, 2025.

[4] Anlun Huang, Zhenyu Wu, Soofiyan Atar, Yuheng Zhi, and Michael Yip. Steadytray: Learning object balancing tasks in humanoid tray transport via residual reinforcement learning. *arXiv preprint arXiv:2603.10306*, 2026.

[5] Tobias Johannink, Shikhar Bahl, Ashvin Nair, Jianlan Luo, Avinash Kumar, Matthias Loskyll, Juan Aparicio Ojea, Eugen Solowjow, and Sergey Levine. Residual reinforcement learning for robot control. In *2019 international conference on robotics and automation (ICRA)*, pages 6023–6029. IEEE, 2019.

[6] Zhengyi Luo, Ye Yuan, Tingwu Wang, Chenran Li, Sirui Chen, Fernando Castaneda, Zi-Ang Cao, Jiefeng Li, David Minor, Qingwei Ben, et al. Sonic: Supersizing motion tracking for natural humanoid whole-body control. *arXiv preprint arXiv:2511.07820*, 2025.

[7] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel Van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions On Graphics (TOG)*, 37(4):1–14, 2018.

[8] Xue Bin Peng, Ze Ma, Pieter Abbeel, Sergey Levine, and Angjoo Kanazawa. Amp: Adversarial motion priors for stylized physics-based character control. *ACM Transactions on Graphics (ToG)*, 40(4):1–20, 2021.

[9] Tom Silver, Kelsey Allen, Josh Tenenbaum, and Leslie Kaelbling. Residual policy learning. *arXiv preprint arXiv:1812.06298*, 2018.

[10] Zhenguo Sun, Yibo Peng, Yuan Meng, Xukun Li, Bo-Sheng Huang, Zhenshan Bing, Xinlong Wang, and Alois Knoll. Robotdancing: Residual-action reinforcement learning enables robust long-horizon humanoid motion tracking. *arXiv preprint arXiv:2509.20717*, 2025.

[11] Zhenguo Sun, Bo-Sheng Huang, Yibo Peng, Xukun Li, Jingyu Ma, Yu Sun, Zhe Li, Haojun Jiang, Biao Gao, Zhenshan Bing, et al. Mosaic: Bridging the sim-to-real gap in generalist humanoid motion tracking and teleoperation with rapid residual adaptation. *arXiv preprint arXiv:2602.08594*, 2026.

[12] Yanjie Ze, Zixuan Chen, Joao Pedro Araujo, Zi ang Cao, Xue Bin Peng, Jiajun Wu, and Karen Liu. TWIST: Teleoperated whole-body imitation system. In *9th Annual Conference on Robot Learning*, 2025.

[13] Yanjie Ze, Siheng Zhao, Weizhuo Wang, Angjoo Kanazawa, Rocky Duan, Pieter Abbeel, Guanya Shi, Jiajun Wu, and C Karen Liu. Twist2: Scalable, portable, and holistic humanoid data collection system. *arXiv preprint arXiv:2511.02832*, 2025.

[14] Siheng Zhao, Yanjie Ze, Yue Wang, C Karen Liu, Pieter Abbeel, Guanya Shi, and Rocky Duan. Resmimic: From general motion tracking to humanoid whole-body loco-manipulation via residual learning. *arXiv preprint arXiv:2510.05070*, 2025.