Robot Skill Mimicking and Future: A Survey

Warner Wu [1]

Abstract

This survey reviews recent progress in robot skill mimicking and learning-based locomotion and manipulation systems. This survey serves as a preliminary overview for new research directions in robot learning methodologies.

1. Relative Works

1.1. Blind locomotion

Blind locomotion is enabled by simulator-based software such as Legged Gym. Blind locomotion has gone through two stages: pure reinforcement learning with reward constraints and demonstration mimicking. DeepMimic and AMP are two major advances in the heuristic locomotion with human prior.

Tricks like Domain Randomizations and privilege info teacher-student policy learning are applied to increase robustness respectively for unseen environments and partially-observed environments.

1.2. Blind WBC

Blind WBC follows the same pattern as blind locomotion. The difficulty here was the mechanism to transfer human motion to robotics motion, which is name **Motion Retargeting**. After retargeting, an RL-based trajectory tracker is trained to overfit and provide robustness for the motion that have been generated. Some works like BeyondMimic(Liao et al., 2025) and GMR (Araujo et al., 2025)

Problem remains as how to make robots follows robust movement generating at any time with zero-shot deployment, and also how to make close-loop action with feedbacks from vision and tactile or, in a sense, contact.

OmniRetarget preserve the contact and provide the dataset for the humanoid-object interaction (Yang et al., 2025) But still lack of real-time interaction feedback from the scene and object.

1.3. Vision locomotion

Vision locomotion works using hardware like LiDAR and depth camera, with tools like Elevation Map built from LiDAR Data. PIM are typical work of these.

1.4. Vision WBC/manipulation

Isaac Gym does not provide realistic visual sensing simulation, meaning that visual system dynamics are unavailable in the simulator.

Some recent works therefore utilize behavioral cloning to enable the learning process, such as distilling a learned blind locomotion policy into another network (Zhengmao He, 2024). But the policy still faces unsuccessful demostration and the morphology. This pipeline shares the common things that train an RL policy robust enougn and then find another policy that scale the visual input. One big problem towards visual training is the huge amount of training data needs(He et al., 2025), which is probably a right way to solve the problem, according to the experience learned from Large Language Model.

Some other works avoid the visual training problem by using waypoint in training, and use deploy-time perception to navigate the waypoints. works like PhysHSI (Wang et al., 2025) is example of these.The second half of BeyondMimic (Liao et al., 2025) is blind but using external MoCap room as a powerful state estimator. Other downstream applications like Hitter(Su et al., 2025)

For earlier navigation problem, the author also avoid this problem by multi-stage training without visual simulation. Because only partial information of visual input is important. Same dogma was imply in the project

1.5. Tele-Operation

Researchers, especially those in the industry, due to the pressure of delivering product

1.6. VLA and BFM

Now, people are seeing the advances in LLM, so one thing to transfer the idea into robotics is the idea of **Vision-**

[1] Department of XXX, University of YYY, Location, Country. Correspondence to: Warner Wu <email@domain.com>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

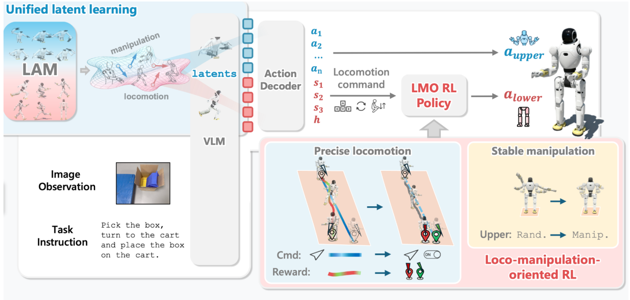

Figure 1. pipeline of wholebodyVLA, where VLM is decoding the tokens for later upper&lower policy for manipulation and locomotion. While locomotion is trained on RL, the manipulation decoder training is supervised learning such as an MLP

Language-Action model. The idea is simple: take vision and language instruction as input, and action as output.

$$a = \pi(o, l) \tag{1}$$

The model architecture is a Transformer, which concate o and l, self-attend with each other, and output to the action head. The action head is normally an MLP. But in terms of humanoid robots, this is an RL policy, So VLA acts like a task-level policy.

2. Industry and Big Tech Labs

Industry like SundayRobotics have already shipped robots that really act autonomously in 0.1 speed, using ACT derived from (Zhao et al., 2023)

3. Datasets

4. Questions about safety and agility

I see people drawing curves in previous paper pipeline, which is man-made trajectory. in the demos before. Is it possible for **generative model** to work on manipulation or even loco-manipulation task?

References

Araujo, J. P., Ze, Y., Xu, P., Wu, J., and Liu, C. K. Retargeting matters: General motion retargeting for humanoid motion tracking, 2025. URL https://arxiv.org/abs/2510.02252.

He, T., Wang, Z., Xue, H., Ben, Q., Luo, Z., Xiao, W., Yuan, Y., Da, X., Castañeda, F., Sastry, S., Liu, C., Shi, G., Fan, L., and Zhu, Y. Viral: Visual sim-to-real at scale for humanoid loco-manipulation, 2025. URL https://arxiv.org/abs/2511.15200.

Liao, Q., Truong, T. E., Huang, X., Gao, Y., Tevet, G., Sreenath, K., and Liu, C. K. Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv preprint arXiv:2508.08241*, 2025.

Su, Z., Zhang, B., Rahmanian, N., Gao, Y., Liao, Q., Regan, C., Sreenath, K., and Sastry, S. S. Hitter: A humanoid table tennis robot via hierarchical planning and learning, 2025. URL https://arxiv.org/abs/2508.21043.

Wang, H., Zhang, W., Yu, R., Huang, T., Ren, J., Jia, F., Wang, Z., Niu, X., Chen, X., Chen, J., Chen, Q., Wang, J., and Pang, J. Physhsi: Towards a real-world generalizable and natural humanoid-scene interaction system, 2025. URL https://arxiv.org/abs/2510.11072.

Yang, L., Huang, X., Wu, Z., Kanazawa, A., Abbeel, P., Sferrazza, C., Liu, C. K., Duan, R., and Shi, G. Omniretarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction, 2025. URL https://arxiv.org/abs/2509.26633.

Zhao, T. Z., Kumar, V., Levine, S., and Finn, C. Learning fine-grained bimanual manipulation with low-cost hardware. *arXiv preprint arXiv:2304.13705*, 2023.

Zhengmao He, Kun Lei, Y. Z. K. S. Z. L. H. X. Learning visual quadrupedal loco-manipulation from demonstrations. 2024.