Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.37515

Bipedal Navigation and Locomotion Control Algorithms for Humanoid Robots

Xueyan Fu

School of Mechanical Engineering, University of Science and Technology Beijing, Beijing, China
u202340757@xs.ustb.edu.cn

Abstract. With their anthropomorphic design, humanoid robots are naturally adapted to human environments; however, high degrees of freedom, severe nonlinear dynamics, and small stability margins make autonomous locomotion a major bottleneck for practical applications. This paper surveys the key algorithms for humanoid robot locomotion, navigation, and motion control. In legged navigation, this survey traces the development from simplified-model gait generation to search-based footstep planning, and finally to online dynamic replanning and, ultimately, a hierarchical "Footstep Planning to Model Predictive Control to Whole-Body Control" framework. In motion control, this paper reviews two primary approaches: model-based optimization and reinforcement learning. The former has gone from simple models to real-time model predictive control and whole-body multi-contact optimization, and the latter has developed from proprioceptive blind walking to terrain adaptation through vision and synthesis of anthropomorphic motion. Next, a comparative analysis of these two paradigms on three aspects: stability, real-time performance and environmental adaptability, reveals that the "model-based reference with learning-based residual" hybrid scheme has emerged as a prominent trend. Finally, open challenges are mentioned in a systematic fashion, such as the interpretability and safety of learned policies, long-horizon robust locomotion in unstructured environments and whole-body integration of locomotion and manipulation.

Keywords: Humanoid robot, walking navigation, footstep planning, motion control

1. Introduction

Humanoid robots are autonomous systems that are inspired by the human morphology and locomotion. Their bipedal design is well adapted to the human living and working environment, allowing them to step over obstacles, go upstairs and go through uneven terrains. But the bipedal systems inherently have large degrees of freedom, high dynamic nonlinearities and small stability margins. To address the dual core challenges of "where and how to navigate" and "how to maintain dynamic stability," humanoid robots require the concurrent development of legged navigation and motion control systems. In particular, planning and execution of walking paths towards target destinations in response to environmental stimuli are addressed as legged navigation, while the stable execution of each planned step is addressed as motion control. Moreover, they involve close coupling between the modules of visual perception, generation of states, gait generation and whole-

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.37515

body control, with errors in one module being dynamically propagated to others and resulting in loss of stability or catastrophic falls. Thus, autonomous bipedal locomotion, navigation and motion control are fundamental challenges in humanoid robotics, and are essential prerequisites for real-world applications, such as domestic service, disaster response and industrial inspection. In this context, research has been conducted on multiple parallel tracks, ranging from the early simplified models, to the recent learning-based controllers.

Initial works on humanoid robots focused on obtaining stable gaits using a simplified model, for example, the center of mass dynamics were linearized using the 3D Linear Inverted Pendulum Model to generate bipedal walking gaits quickly [1]. On this basis, footstep planning approaches developed a formulation of the footstep planning problem as a search and optimization problem for discrete footsteps, and proved the validity of searching based footstep generation on robotic platforms like ASIMO [2]. As a result of the developments of Model Predictive Control, researchers started solving and updating whole-body motion trajectories with optimization over a receding horizon, allowing the robot to deal with external disturbances and to adapt to dynamic environments [3].

Recently, reinforcement learning has made it possible to learn robust locomotion policies directly without modeling dynamics and achieved significant improvements in adaptability to a variety of environments, including different outdoor terrains. Humanoids trained with policies using domain randomization can achieve zero-shot transfer to real-world robots, not only to robust locomotion and dynamic disturbance rejection in challenging outdoor terrain, but to allow online adaptation to contextual information, laying a solid foundation for future real-world applications of humanoids in industrial and home environments [4]. But these advances are spread out in different layers, namely: planning, model-based control and learning.

Hence, this paper aims to give a thorough review focusing on two fundamental aspects: navigation planning and motion control, which includes both classical search- and optimization-based methods as well as data-driven learning paradigms. First, this survey outlines the development of gait generation and footstep planning to autonomous navigation in dynamic environments technically, and systematically explains the hierarchical theoretical logic behind mobile locomotion control. This paper then makes a systematic comparison of the two approaches, model-based optimization and reinforcement learning in terms of stability, real-time performance, and environmental adaptability, pinpointing their respective strengths, weaknesses, and how they can complement each other to help guide their applicability in real systems. Finally, this paper summarizes the current state of the art and projects future research directions, aiming to offer a valuable reference for subsequent studies in this field.

2. Bipedal locomotion and navigation planning for humanoid robots

2.1. Gait generation and footstep planning

The autonomous mobility of humanoid robots primarily confronts two fundamental challenges: "how to walk stably" and "where to step"; the former corresponds to gait generation, while the latter pertains to footstep planning.

This subsection first addresses gait generation, i.e., how a humanoid robot can walk stably. The key to solving the question "how to walk stably" is dynamic balance control: a closed-loop control process in which the system detects the states and changes in the environment and dynamically modifies joint actuation torques and contact forces to reject perturbations and keep the center of mass and contact points within safe stability limits. However, for balance control, the first step is to

DOI: 10.54254/2755-2721/2026.37515

introduce criteria which are able to predict the risks of instability accurately. The Zero Moment Point criterion provided the basic theory for dynamic bipedal walking [5]. The ZMP is the point at which the horizontal components of the ground reaction moment vanish; if this point lies within the support polygon, dynamic balance can be maintained. This criterion was originally introduced with the "semi-inverse method" for gait synthesis and is still one of the most important criteria for evaluating and constraining bipedal balance. While the Zero Moment Point criterion provides balance constraints, it is difficult to directly resolve the high-degree-of-freedom whole-body trajectories from it. To overcome this limitation, the 3D Linear Inverted Pendulum Model was proposed. By using this method, the center of mass of the robot is confined to a plane of fixed height thus simplifying the complicated nonlinear whole-body dynamics to an analytically tractable linear one. This greatly reduces online computational overhead and thus makes it the most popular reduced order model for the pattern generation of bipedal walking. However, more sophisticated balance quantities than simple planar models are needed to realize real-time balance control over the complex 3D terrains. The concept of Divergent Component of Motion was extended in three dimensions [6] to describe the direction and magnitude of external and resultant forces, which can be used for real-time trajectory planning and tracking for improved adaptive balancing on undulating surfaces.

Once stable locomotion is achieved, one should be able to detect stable footstep placements via footstep planning. One discrete footstep planning method that was developed on the Honda ASIMO platform was based on graph search [2]. The approach is to set a finite horizon and search for a sequence of step lengths and turning angles that can be performed discretely and that will avoid obstacles using A*. Departing from the "obstacle-avoidance" concept of the work described above, later works further explored the potential of crossing obstacles only available for humanoids [7]. They cast it as a problem of "humanoid terrain traversal" to let the robot step onto obstacles or step over the obstacles instead of detouring around the obstacles. To overcome the high computational complexity of obstacle-crossing planning, they added algorithmic optimizations and verified the feasibility of dynamic obstacle crossing on the HRP-2 platform, thus expanding the range of applications of footstep planning beyond planar collision avoidance to dynamic negotiation of non-flat environments.

Summarizing, gait generation research has developed following the technical path from "stability criteria to tractable models, and further to 3D real-time control", where the Zero Moment Point criterion provided a unified way to assess dynamic balance, the 3D Linear Inverted Pendulum Model reduced the high-dimensional and nonlinear whole-body dynamics into an analytically solvable linear model, and the Divergent Component of Motion framework extended the balance control from planar surfaces to 3D space, allowing humanoid robots to walk steadily on increasingly complex terrains. At the same time, footstep planning can be used to link "stable walking" with "autonomous navigation" and give the robot the ability to either evade or overcome obstacles on its path to reach a desired goal location. However, these approaches are largely based on simplified models of center of mass dynamics and are not well equipped to simultaneously consider full-body joint torques and contact constraints and are not sufficiently robust against changing dynamics in the environment. These limitations are what precisely drive the next steps towards online planning and whole-body motion control.

2.2. Online motion planning in dynamic environments

The above footstep planning assumes that the environment is unchanged during one planning cycle. But if, for example, obstacles are moved or the terrain changes, the predetermined static foothold

DOI: 10.54254/2755-2721/2026.37515

sequence is no longer valid. This means that robots have to compute online in real-time and have to replan quickly while they are moving. To achieve this, two approaches have been investigated, namely: making the algorithms for replanning more efficient; and combining replanning and motion control into hierarchical frameworks. Along the first track, the incremental heuristic search algorithm D* Lite was introduced into footstep planning [8]. Since a replanning query often exhibits only localized discrepancies compared with the prior search graph, the D* Lite algorithm efficiently reuses most of the previous search tree, thereby achieving replanning at a computational cost significantly lower than re-solving from scratch. Furthermore, they realized continuous foothold modeling alongside efficient footstep collision checking, validating the online adaptation of footstep paths across both simulated and physical Nao humanoid robots. This work was a milestone paradigm change in footstep planning: from "one-shot offline computation" to "on-the-fly online correction", and thus set algorithmic groundwork for real-time planning in dynamic environments.

Distinct from incremental graph reuse, an alternative paradigm prioritizes obtaining an initial feasible solution within a strict time budget, followed by progressive optimization. Both Anytime Repairing A* and Randomized A* were introduced into footstep planning [9]. It returns a feasible path quickly, and then iterates it towards optimality, if time allows, with a provable suboptimality bound with respect to the optimal path. These anytime search methods maintain the goal-directedness of the search, unlike purely randomized sampling-based search methods like Rapidly-exploring Random Trees, which expand many nodes in areas with a large number of local minima. The authors systematically evaluated the proposed framework for various scenarios and heuristic formulations, and confirmed its ability to provide paths that have a theoretical suboptimality bound in extremely short planning horizons. This is important for optimising bipedal locomotion: the robot does not have to wait for the whole optimal sequence of footholds to be computed before it takes its first step on the currently possible path. At the same time, the planner polishes up the following footholds as the motion is executed, balancing both for a quick start and for good path quality.

After footstep planning, center of mass trajectories and contact forces need to be resolved and mapped to joint torques. So, a hierarchical framework was proposed that consists of "Footstep Planning—Model Predictive Control—Whole-Body Control." In particular, footstep placements are optimized online at the top layer, model predictive control is used to generate center-of-mass trajectories and contact forces at the intermediate layer, and finally the motions are tracked and executed at the bottom layer using whole-body control. This framework was then used in dynamic obstacle avoidance in which the locations of the footholds were modulated online based on the position of the moving obstacles and the robot"s dynamics [10]. They verified on the simulated DRACO3 platform the real-time collision-avoidance effectiveness of this hierarchical scheme.

3. Motion control algorithms for humanoid robots

In the previous section, gait generation and footstep planning were studied to guarantee the preservation of balance and guidance of foot placement, but in the current section, the main attention is paid to the core motion control layer. This survey systematically breaks down two of the dominant control paradigms, model-based optimization and reinforcement learning, and compares them along important aspects such as dynamic stability, real-time computational feasibility and environmental adaptability.

DOI: 10.54254/2755-2721/2026.37515

3.1. Model- and optimization-based motion control

Model-based approaches are based, at their outset, on simplified models. Reduced-order models, such as the 3D Linear Inverted Pendulum Model, approximate full-body dynamics using center-of-mass dynamics, offering high computational efficiency and structural simplicity. However, they only retain the center of mass state, losing the whole-body joint limits and contact constraints and thus making them incapable of dealing with complex motions. In addition, they are conventionally modelled in an open-loop paradigm, in which an offline computed trajectory is followed in an open-loop manner with little robustness to real-time perturbations. To balance between the real-time performance and dynamic fidelity, researchers formulated the control problem in a receding horizon iterative optimization framework, which resulted in model predictive control. A typical early example of this has been an extensive model predictive controller that was executed in real-time on the HRP-2 humanoid robot [3]. They evaluated the feasibility of full body model predictive control on real humanoid robots by introducing custom-made engineering treatments to reduce computation times and modeling errors.

As tasks go beyond stable walking to multi-contact manipulation and very dynamic maneuvers, full-body nonlinear optimal control over the complete dynamics becomes necessary. Crocoddyl, an open-source framework for multi-contact optimal control, was proposed [11]. It makes use of sparse analytical derivatives, structural sharing and differential geometry to represent the states of floating-base systems, and thus to solve multi-contact optimal control problems in an efficient manner. They also suggested the Feasibility-Driven Differential Dynamic Programming algorithm, which removes the constraint of infeasibility in the backward pass and keeps gaps open in the early stages of exploration, which offers better convergence in a global sense as compared to the classical Differential Dynamic Programming algorithm. Previously, even multi-contact optimal control problems were considered "computationally intractable", and Crocoddyl took multi-contact optimal control from "computationally intractable" to "real-time solvable", offering a versatile solver for whole-body optimal control.

On top of such a generic solver framework, the nonlinear model predictive control architecture BiConMP was added where the whole-body problem is projected back onto the centroidal one. BiConMP essentially decreases the dimension of the optimization problem by only considering the relationship between the robot center of mass and the contact forces [12]. The resulting formulation has a built-in biconvex structure, by fixing the contact forces the problem becomes convex with respect to the center-of-mass trajectory and vice versa, allowing for alternating optimization. BiConMP efficiently decomposes the original problem into convex subproblems that can be solved iteratively and scales up full-body trajectory generation to online receding-horizon optimization thanks to the use of efficient optimization algorithms like Alternating Direction Method of Multipliers and proximal gradient methods. This method was validated on the physical Solo12 quadruped for periodic gaits, online gait transitions, and push recovery under external perturbations. Aperiodic whole-body dynamic motions were demonstrated in simulation on the Talos humanoid and the ANYmal quadruped. BiConMP validates the use of real-time nonlinear whole-body model predictive control for legged/humanoid platforms and takes the Crocoddyl-style optimization one step towards "online, real-time".

In addition, these optimization techniques were put into practice in a complete system running on real humanoid hardware [13]. Combining convex optimization, mixed-integer programming, and sparse nonlinear optimization in footstep planning, whole-body motion control, and state estimation, they were able to execute long-range, precise locomotion on non-flat, complex ground on the full-

DOI: 10.54254/2755-2721/2026.37515

scale hydraulic humanoid robot Atlas developed by Boston Dynamics. This work took whole-body optimization from individual algorithms to a real-world deployed system, thus proving that such approaches can efficiently solve complex agile maneuvers and can be reliably executed on physical humanoid platforms.

From the above, it is clear that model-based methodologies have progressively evolved toward increasingly complete dynamic models and higher behavioral complexity. Reduced order models have the greatest computational tractability, but lack dynamic fidelity, while model predictive control is able to balance real-time feasibility and accuracy. On the other hand, full-body nonlinear optimization can obtain the highest dynamic fidelity, naturally handling multi-contact, while being extremely hard to satisfy any online real-time requirements unless having a very efficient numerical solver.

3.2. Reinforcement learning-based adaptive locomotion control

In contrast to model-dependent approaches, reinforcement learning has been used to learn the locomotion policies directly from trial-and-error interactions in simulation without requiring explicit dynamic modeling. A parametric reinforcement learning framework was proposed in which one policy can be trained to learn various gaits with different parameters [14] and the sim-to-real gap is filled with domain randomization, a formal randomisation of the parameters like mass, friction etc. during training. This controller was implemented on the real Cassie bipedal robot and showed the ability to control online the walking velocity, the body height and the yaw angle of the robot, as well as a push recovery ability to withstand external perturbations. The parameterized reinforcement learning controllers outperform traditional controllers and previous residual learning methods in terms of robustness. In this work, the authors show that reinforcement learning can actually learn a unified gait policy that can be used to execute any number of tunable gaits, without requiring to code specific gait controllers for each gait.

To this, proprioceptive policies were scaled-up from the standard lightweight networks to the large-scale Causal Transformer networks [4]. This enables that history of observations and actions can be encoded as a context token, which can be used to adapt the environment without explicit, online weight updates, but rather making use of implicit representations of the environment. The learned model was trained with large-scale domain randomization in the simulation and directly transferred zero-shot to the physical humanoid platform and was shown to be more robust on a variety of outdoor terrains. Later, modeling of sequences was pre-trained on flat ground trajectories and followed by reinforcement learning fine-tuning on rugged terrains, which allowed the general framework to blindly navigate a 4-mile mountain trail and steep hiking slopes in San Francisco [15]. The work is an important first step towards scaling up purely proprioceptive "blind" locomotion to complex and highly dynamic and unstructured physical environments.

When this is coupled with pure proprioception, vision exteroception can give "look-ahead" terrain geometry, which can shift the footfall from passive, reactive balance to proactive, feed-forward gait modulation. However, direct learning of control policies from visual inputs is known to have problems of sample inefficiency, convergence and the enormous visual domain gap between the simulated and real-world environments makes it hard to transfer the policy directly from simulation to real environments. To overcome such problem, a Vision-Assisted Teacher-Student framework was proposed to avoid the bottleneck of visual training [16]. In particular, a privileged teacher policy is first trained with ground-truth data to learn terrain features, and foothold selection is dynamically routed to different sub-networks, depending on the terrain. Then a student policy, which relies primarily on proprioception, is learnt to imitate the teacher, using an alignment

DOI: 10.54254/2755-2721/2026.37515

loss. Instead of pointwise matching, it relies on consistency between the student and the teacher's behaviors, and allows the robot to autonomously modify its gait in response to geometries of the terrains. Further, the teacher policy is derived from Mixture-of-Experts: distinct sub-networks are assigned to distinct terrain categories and different sub-networks are dynamically switch to the teacher, depending on features in the environment. In real-world experiments, the P1 bipedal robot by Limx Dynamics was shown to be able to walk in a stable and reliable manner across challenging unstructured terrains including staircases with single step-downs, slopes and compliant grassy surfaces, with the aid of this algorithm. This work not only systematically verifies the effectiveness of the exteroceptive perception for complex terrain adaptability of the bipeds, but also will be an important and pioneering paradigm shift in bipedal locomotion from proprioceptive blind walking to integrated visual-proprioceptive perception.

In addition to exteroceptive perception and strong stability, a parallel line of research uses human motion capture data as a reference prior, in order to counteract the unnatural stiffness of traditional gaits, and instead produce very anthropomorphic locomotion. HiLo was proposed, which breaks down whole-body control into an open-loop feed-forward component and a reinforcement learning residual component [17]. In particular, the open-loop controller determines the basic motion patterns, and the reinforcement learning policy only learns the corrective residuals. The framework is able to synthesize natural, agile and perturbation-resilient human-like gaits by adding this decomposition to simple domain randomization including stochastic force injections and actuation latency. An imitation learning approach based on Adversarial Motion Priors was also developed where an imitation learning discriminator is used to enforce outputs of the policy to be human-like motion distributions [18]. By leveraging human motion capture data, the authors achieve natural locomotion with near-human gait characteristics on Adam, a custom full-scale humanoid robot, marking the first demonstration of motion-capture-driven learning at this scale. Together these investigations collectively imply that the research goals in the field of bipedal locomotion control have shifted from the basic "environment-feedback-driven stable traversal" to more of a "human-prior-guided anthropomorphic synthesis".

When considered together, the aforementioned model-based paradigm and learning-based paradigm have different operation regions and are complementary. In terms of stability, the model-based approaches are very well interpreted and allow for constraint satisfaction, which is very suitable for safety-critical situations. On the other hand, reinforcement learning is empirically more robust, and stability guarantees a priori is challenging. As far as real-time tractability, convex optimization with simplified models is fast, while whole-body optimization requires special, high efficiency solvers; on the other hand, reinforcement-learning requires expensive offline computation, but fast online computation. However, model-based methods can only go as far as the models while reinforcement learning has been shown to be more robust in dealing with a variety and diversity of terrains through significant domain randomization. This makes model-based methods more useful in structured operational scenarios whereas reinforcement learning is useful in unstructured scenarios and changing environments. A hybrid paradigm to integrate both approaches is gradually becoming a significant research direction.

4. Conclusion

This review is focused on two fundamental questions for humanoid robots, namely "where to go" (humanoid navigation) and "how to walk stably" (whole-body motion control), and systematically summarizes the recent research advances in these two major aspects. For the case of bipedal navigation, the research starts with simple reduced order models like Zero Moment Point and the 3D

DOI: 10.54254/2755-2721/2026.37515

Linear Inverted Pendulum Model, which are used to achieve an analytical gait generation. These formulations then use search-based footstep planning to find discrete footholds and combine them with incremental search and anytime planning to allow for replanning in a real-time context when operating in dynamic environments, eventually leading to a hierarchical formulation of "Footstep Planning to Model Predictive Control to Whole-Body Control". The development of model-based optimization in motion control has been paralleled by the development of reinforcement learning first as proprioceptively blind locomotion and more recently as vision-based motion control and anthropomorphic motion synthesis. The two paradigms have complementary viewpoints and specific strengths with respect to dynamic stability, real-time computational tractability and environmental adaptability.

In the future, the state of the art of humanoid locomotion will move beyond the ability of simply maintaining balance to multi-modal convergence and more complex behavioral goals. With multi-modal perception and the use of human motion priors, the goal of locomotion control has shifted from dynamic stability to natural and anthropomorphic synthesis and smooth human-robot interaction. However, there are still important open challenges such as formal interpretability and safety verification of reinforcement learning policies, long-horizon robust locomotion control in unstructured environments and unified coordination of whole-body locomotion with multi-contact physical interaction and manipulation. Progress in solving these hurdles will be crucial to the potential of moving humanoid robots from the laboratory to real-world operational applications such as the home, industrial factories and disaster response.

References

[1] Kajita, S., Kanehiro, F., Kaneko, K., Yokoi, K. and Hirukawa, H. (2001) The 3D Linear Inverted Pendulum Mode: A Simple Modeling for a Biped Walking Pattern Generation. In 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, 239-246.

[2] Chestnutt, J., Lau, M., Cheung, G., Kuffner, J., Hodgins, J. and Kanade, T. (2005) Footstep Planning for the Honda ASIMO Humanoid. In 2005 IEEE International Conference on Robotics and Automation, 629-634.

[3] Koenemann, J., Del Prete, A., Tassa, Y., Todorov, E., Stasse, O., Bennewitz, M. and Mansard, N. (2015) Whole-Body Model-Predictive Control Applied to the HRP-2 Humanoid. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 3346-3351.

[4] Radosavovic, I., Xiao, T., Zhang, B., Darrell, T., Malik, J. and Sreenath, K. (2023) Real-World Humanoid Locomotion with Reinforcement Learning.

[5] Vukobratović, M. and Borovac, B. (2004) Zero-Moment Point: Thirty-Five Years of Its Life. International Journal of Humanoid Robotics, 1, 157-173.

[6] Englsberger, J., Ott, C. and Albu-Schäffer, A. (2015) Three-Dimensional Bipedal Walking Control Based on Divergent Component of Motion. IEEE Transactions on Robotics, 31, 355-368.

[7] Ayaz, Y., Owa, T., Tsujita, T., Konno, A., Munawar, K. and Uchiyama, M. (2009) Footstep Planning for Humanoid Robots among Obstacles of Various Types. In IEEE-RAS International Conference on Humanoid Robots, 361-366.

[8] Garimort, J., Hornung, A. and Bennewitz, M. (2011) Humanoid Navigation with Dynamic Footstep Plans. In IEEE International Conference on Robotics and Automation, 3982-3987.

[9] Hornung, A., Dornbush, A., Likhachev, M. and Bennewitz, M. (2012) Anytime Search-Based Footstep Planning with Suboptimality Bounds. In IEEE-RAS International Conference on Humanoid Robots, 674-679.

[10] Rossini, L., Hoffman, E.M., Bang, S.H., Sentis, L. and Tsagarakis, N.G. (2023) A Real-Time Approach for Humanoid Robot Walking Including Dynamic Obstacles Avoidance. In IEEE-RAS International Conference on Humanoid Robots, 1-8.

[11] Mastalli, C., Budhiraja, R., Merkt, W., Saurel, G., Hammoud, B., Naveau, M., Carpentier, J., Righetti, L., Vijayakumar, S. and Mansard, N. (2020) Crocoddyl: An Efficient and Versatile Framework for Multi-Contact Optimal Control. In IEEE International Conference on Robotics and Automation, 2536-2542.

[12] Meduri, A., Shah, P., Viereck, J., Khadiv, M., Havoutis, I. and Righetti, L. (2023) BiConMP: A Nonlinear Model Predictive Control Framework for Whole Body Motion Planning. IEEE Transactions on Robotics, 39, 905-922.

DOI: 10.54254/2755-2721/2026.37515

[13] Kuindersma, S., Deits, R., Fallon, M., Valenzuela, A., Dai, H., Permenter, F., Koolen, T., Marion, P. and Tedrake, R. (2016) Optimization-Based Locomotion Planning, Estimation, and Control Design for the Atlas Humanoid Robot. Autonomous Robots, 40, 429-455.
[14] Li, Z., Cheng, X., Peng, X.B., Abbeel, P., Levine, S., Berseth, G. and Sreenath, K. (2021) Reinforcement Learning for Robust Parameterized Locomotion Control of Bipedal Robots. In IEEE International Conference on Robotics and Automation, 2811-2817.
[15] Radosavovic, I., Kamat, S., Darrell, T. and Malik, J. (2024) Learning Humanoid Locomotion over Challenging Terrain. arXiv Preprint arXiv: 2410.03654.
[16] Chen, F., Wan, R., Liu, P., Zheng, N., Wang, B. and Zhou, B. (2025) VMTS: Vision-Assisted Teacher-Student Reinforcement Learning for Multi-Terrain Locomotion in Bipedal Robots. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 4759-4766.
[17] Zhang, Q., Weng, C., Li, G., He, F. and Cai, Y. (2025) HiLo: Learning Whole-Body Human-Like Locomotion with Motion Tracking Controller. arXiv Preprint arXiv: 2502.03122.
[18] Zhang, Q., Cui, P., Yan, D., Sun, J., Duan, Y., Han, G., Zhao, W., Zhang, W., Guo, Y., Zhang, A. and Xu, R. (2024) Whole-Body Humanoid Robot Locomotion with Human Reference. arXiv Preprint arXiv: 2402.18294.