Augmented Action-space Whole-Body Teleoperation of Mobile Manipulation Robots

Sophie Lueth **Georgia Chalvatzaki**
Department of Computer Sciences
Technical University of Darmstadt Germany
{sophie.lueth,georgia.chalvatzaki}@tu-darmstadt.de

Abstract: We present a novel approach for whole-body teleoperation of mobile manipulation robots, focusing on complex household tasks that require coordination of arms-manipulation and robot-body posture, while enabling multiple contact points along the robot's kinematic chain. Current state-of-the-art methods often overlook the synergy between different parts of the robot's body, particularly in humanoid robots, as they typically rely on end-effector actions alone. In contrast, we propose a teleoperation framework that supports multi-embodiment and multi-contact teleoperation, allowing for efficient whole-body control that enhances task execution in real-world scenarios and brings robots closer to human-like behaviors. Our augmented action-space compliant teleoperation interface enables users to control the robot and collect high-quality demonstrations for future imitation learning. Preliminary experiments demonstrate promising results, with high success rates in task completion, suggesting the potential of our approach for robust and efficient teleoperation. Videos of our real-robot experiments can be found on https://sites.google.com/view/aawbt/home.

Keywords: Whole-Body Manipulation, Teleoperation, Whole-Body Control

1 Introduction

Humanoid robots have long been envisioned as household assistants, yet the complexity of tasks within a human household and the unstructured nature of these environments have limited their integration into our daily life. While robot learning promises to provide the intelligence necessary for humanoid robots to fully exploit their capabilities, one major challenge remains: enabling these robots to coordinate whole-body actions in a way that mirrors human behavior. The human-like kinematics of humanoid robots enable them to solve tasks similar to humans, which is specifically interesting for cases in which humans use their whole body to solve a task, e.g., when closing a drawer or the dishwasher with a knee because both hands are occupied (Figure 1). Recent efforts in robot teleoperation [1, 2, 3, 4, 5, 6, 7, 8], show that in many cases the hardware is not the missing component, but rather the intelligence for coordinating multiple embodiments based on perceptual feedback.

Figure 1: Human closing a dishwasher with their knee because their hands are full of objects they just grasped from the dishwasher.

This paper introduces a novel teleoperation framework that augments the robot's action space to enable whole-body manipulation with multiple contact points along the robot's kinematic chain. Our approach allows for simultaneous interaction with the envi-

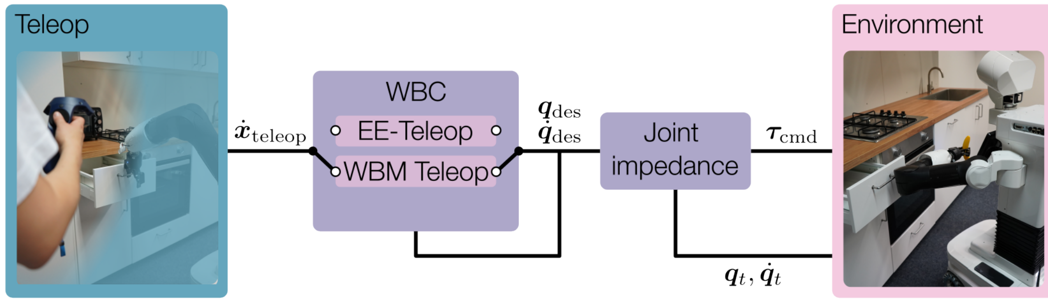

Figure 2: Teleoperation Framework Overview

ronment using parts of the robot's body beyond the end-effector, as exemplified by the scenario in Figure 1. This opens up new possibilities for more efficient and human-like task execution.

We are particularly interested in tasks where we allow manipulation to happen beyond the end-effector space. We characterize these tasks by exploiting an additional contact point with the environment along the kinematic chain, beyond the one possible by the robot's end-effector, which allows the execution of several sub-tasks at the same time, similar to the example of Figure 1. Notably, the idea of using multiple contact points has been identified previously in a taxonomy of whole-body motions of humans [9], yet the execution of the synergies of different embodiments beyond separated control [10] is limited. Augmenting the action space of humanoid or mobile manipulation robots with the ability to simultaneously interact with different parts of their body with the environment, can broaden the robot's capabilities and increase their efficiency in solving tasks that are intuitive to humans.

In this work, we extend the action space of mobile manipulation robots to whole-body motions and, in particular, we explore the interplay of base, end-effector, and elbow motion in human-teleoperated control through a handheld remote controller. We focus on long-horizon mobile manipulation tasks that require multiple contacts, inspired by the efficiency of humans in completing tasks as the one of Figure 1. To solve the aforementioned group of tasks, we propose a novel augmented action-space teleoperation framework for mobile manipulation robots that considers the control and synergy of expressive coordinate frames, such as the end-effector and the arm's elbow. Our approach allows the teleoperator to convey task-critical motion for the end-effector and any other previously defined task-relevant frame (for this work the elbow and base). Tracking the twist from the teleoperation handheld controller, we can extract either the desired motion for the end-effector or for elbow and base, depending on a chosen teleoperation mode by the teleoperator. From this desired velocity and chosen teleoperation mode, our compliant Whole-Body Controller computes the desired joint-space motion under consideration of additional task objectives and constraints. Crucially, our controller employs impedance control that is always active, unless high-precision end-effector operation is chosen, translating into a higher stiffness control of the robot. We evaluate the suitability and quality of our proposed teleoperation framework and whole-body controller in real-robot experiments in three different scenarios, in which the teleoperator has to control multiple embodiments and establish contacts with different parts of the robot's arm to successfully complete the tasks.

To summarize our contributions are twofold: **(i)** We identify and characterize a group of tasks that benefit from whole-body manipulation, where interacting with the environment through multiple contact points allows for simultaneous execution of several sub-tasks. **(ii)** We propose an augmented action-space teleoperation framework, supported by a compliant whole-body controller that balances multiple (sub)tasks, facilitating novel and more efficient task solutions in real-world scenarios.

2 Related work

Robotic teleoperation systems. Assessing how the human operator controls the robot's many DoFs, we identify three categories in decreasing order of direct influence that the operator's actions have on the robot's joint motion. In the first category, the operator directly controls the joint angles through kinesthetic teaching, coming in physical contact with the system [11] or teleoperates via a leader-follower system [4, 12, 13], that requires the human to be fully engaged, and, in case of mobile manipulation, the teleoperator has to follow the mobile robot; this might not be optimal for constrained spaces, or for teleoperating/collecting data in domains beyond households where humans cannot or should not be present. Crucially, such comparably easy-to-set-up systems requires robot-specific experts [14]. Secondly, we group a set of works [15, 6, 16, 17, 18, 19] that capture human motion via (implicit) keypoints in visual or depth input or from motion-tracking suits. These methods enable the extraction of high-dimensional motion data like kicking or back jumping but require extensive calibration of human body frames. Lastly, we point out virtual or augmented reality-based setups [1, 20, 5, 21, 7, 8, 22], allowing for decoupled motion signals for different frames like arm end-effector(s), base, torso or head by tracking a headset and hand controllers. While removing the need for fine-grained calibration, it limits the number of frames being teleoperated at the same time. At the same time, it introduces the need of some kind of whole-body controller to combine these motion signals into robot joint space.

Learning from teleoperated demonstrations. The current state-of-the-art of learning-based approaches that rely on teleoperated data are not yet representative of tasks that humans are efficient in doing in their households, perform anywhere close to humans for most household tasks [23]. The majority of tasks involve simple pick-and-place sequences. Most policies are formulated either in end-effector or joint space; however, both options come with some limitations: end-effector space does not contain any information about the kinematic chain on which the end-effector sits and, therefore, it cannot easily integrate knowledge about the manipulability and dynamics of the robotic mechanism. In joint space, however, we carry all that information. Nonetheless, teleoperating either for data collection or shared autonomy approaches [24] all robot joints in a robot-optimal manner, exploiting the robot kinematics, and avoiding unfavorable positions of low manipulability and singularities requires extensive expert knowledge and training of the teleoperators. An approach that is close to ours and combines a mapping from human capabilities to robotic execution is humanoid shadowing [25] that learns a mapping from human poses to robot joints; yet, the resulting policies can only handle locomotion and manipulation separately, thus, not fully exploiting the capabilites of robots. We also argue that the absence of embodied teleoperation might not enable the robot to solve a wide range of tasks, e.g., it might only refer to motion style re-targeting and learning without object interactions [19]. Crucially, none of the aforementioned methods consider teleoperation and control for multi-contact whole-body manipulation tasks.

3 Method

We consider a group of long-horizon mobile manipulation tasks that require several subtasks to be completed, some of which can be solved only by using the end-effector, while some require interaction with other parts of the body, while satisfying constraints about the end-effector. In this work, we focus on interacting with the elbow, for example, for closing a door when the hand is full. However, our control approach can be extended to arbitrary contact points along the kinematic chain on the robot, as long as sufficient DoFs are available. We call this type of control *augmented action-space whole-body control*.

We present a teleoperation framework to enable a teleoperator to solve such augmented action-space whole-body manipulation tasks using a handheld controller. The teleoperated twist $\dot{\boldsymbol{x}}_{\text{teleop}}$ from the Teleoperation interface (section 3.1) is given as input to our Whole-Body Controller (section 3.2), which computes the desired joint space motion $\boldsymbol{q}_{\text{des}}, \dot{\boldsymbol{q}}_{\text{des}}$. Depending on the task, we choose to use a current-based joint impedance controller with stiffness adjusted prior to task execution, trading off the impedance's compliance with the high stiffness's accuracy. An overview of the framework can be seen in figure 2.

3.1 Teleoperation Setup

To receive motion signals from the teleoperator, we use a hand controller, from which we track the twist $\dot{\boldsymbol{x}}_{\text{teleop}} = [\boldsymbol{v}_{\text{teleop}}, \boldsymbol{\omega}_{\text{teleop}}]^T \in \mathbb{R}^6$, with linear velocity $\boldsymbol{v}_{\text{teleop}} \in \mathbb{R}^3$ and angular velocity $\boldsymbol{\omega}_{\text{teleop}} \in \mathbb{R}^3$, and the gripper state $g \in [0, 1]$ indicating percentage of gripper opening, with 0 being fully closed and 1 fully open. Additionally, to differentiate between end-effector (EE) whole-body teleoperation (EE mode, $m = 0$) and elbow-based whole-body manipulation (WBM mode, $m = 1$), we track a mode $m \in \{0, 1\}$, resulting in an action space $\boldsymbol{a} \in \mathbb{R}^8$.

The teleoperation twist signal $\dot{\boldsymbol{x}}_{\text{teleop}}$ is understood as a first-person velocity from the robot's viewpoint. The mode m differentiates between the kinematic frames which $\dot{\boldsymbol{x}}_{\text{teleop}}$ influence, i.e., a change in the goal definition of desired twists and respective task-priorities with our WBC:

$$\dot{\boldsymbol{x}}_{\text{ee,wbc}} = \begin{cases} \dot{\boldsymbol{x}}_{\text{teleop}}, & \text{if} \quad m = 0 \\ \boldsymbol{0} & \text{if} \quad m = 1 \end{cases} \qquad \dot{\boldsymbol{x}}_{\text{elbow,wbc}} = \begin{cases} \boldsymbol{0}, & \text{if} \quad m = 0 \\ \dot{\boldsymbol{x}}_{\text{teleop}}, & \text{if} \quad m = 1 \end{cases}$$

3.2 Whole-Body Controller

Our Whole-Body Controller, i.e., the mapping of desired task-space motion $\dot{\boldsymbol{x}}_{\text{ee,wbc}} or \dot{\boldsymbol{x}}_{\text{elbow,wbc}}$ to joint space motion $\boldsymbol{q}_{\text{des}}, \dot{\boldsymbol{q}}_{\text{des}}$ under additional objectives/constrains is implemented as a Task Space Inverse Kinematics Formulation via Quadratic Programming [26]:

$$\min_{\ddot{\boldsymbol{q}}} \sum_i w_i J_i, \quad \text{s.t. } \boldsymbol{a}_j = \boldsymbol{0} \tag{1}$$

with weighted cost functions $w_i J_i$ and constraints $\boldsymbol{a}_j$ for $i, j \in \mathbb{N}$ We introduce an hierarchy of tasks by adding the task objectives either as a cost function with tunable weights or as a hard constraint function. Compared to redundancy resolution via null-space projection, the benefit of this approach is the fast computation time, allowing for frequencies up to 2.5 kHz.

As a result of prior experiments [27], we concluded the following design principles for designing the task objectives: (i) For the teleoperator to develop intuition about the robot's response to their motion signal, it is important that the teleoperated kinematic frame deviates as little as possible from the motion signal, e.g., (self-)collision avoidance. This means that the teleoperator is responsible for avoiding unwanted collisions in time for the teleoperated frame. (ii) For kinematic frames that are not directly teleoperated but whose position is computed via redundancy resolution according to the task hierarchy of the WBC, and can collide with the environment itself, it is important to integrate collision avoidance instead. This is absolutely necessary to allow the controller to compensate for the prevented motion with other degrees of freedom and to not get stuck in a deadlock. (iii) The degrees of freedom in which the whole-body manipulation kinematic frame is teleoperated have to be identified carefully so as to not over-constrain degrees of freedom.

Constraints and cost functions applied in EE mode. To prioritize directly teleoperating the end-effector motion, we set the following task objectives in decreasing order of priority. **(a)** *Joint Limits* in position, velocity, and acceleration are enforced as a hard constraint. **(b)** According to principle (ii), we *do not move the base* when not directly teleoperating it. For simplicity, we do not integrate base collision avoidance and, therefore, use a hard constraint to prevent base motion in this case. (c) The *total teleoperated twist* $\dot{\boldsymbol{x}}_{\text{ee,wbc}}$ is enforced as the cost function with the highest weight. According to principle (i), we refrain from constraining it by self-collision avoidance and trust the user to prevent unwanted collisions. (d) For redundancy resolution, we do *null-space damping* and favor a beneficial joint configuration. We deem it unnecessary to do self-collision avoidance in our current setup, as the kinematic design of our TIAGo robot prevents self-collision with the arm. However, this can be added for other types of robots when needed.

Constraints and cost functions applied in WBM mode. In this mode, we directly teleoperate the base motion and, with less priority, elbow motion. At the same time, we carry out an end-effector task with the same priority as the base motion. In decreasing order of priority, we list the task-objectives set: (a) *Joint Limits* in position, velocity, and acceleration are still enforced as a hard constraint. (b) Direct *teleoperation of the base* according to principle (i) in $\mathbb{SE}(2)$-corresponding velocity space is prioritized equally with the non-teleoperated *end-effector task*. Here, it is important to carefully think about the actively computed DoFs according to principle (iii). (c) To influence the *elbow pose*, we apply the same $\mathbb{SE}(2)$-corresponding velocity. With these task objectives in place, we refrain from adding any joint null-space damping.

4 Experiments

We conduct experiments in the real world with a dual-armed PAL Robotics TIAGo++ mobile manipulation robot with an omnidirectional base. The robot has 20 DoF in total (2 for the head, 1 for the torso, 3 for the holonomic base, and 7 for each arm). In this work, we only use 10 DoF from the left arm (7) and the base (3). For teleoperation, we use the HTC Vive Virtual Reality setup depicted in figure 3. We implement the Whole-Body Controller specified in section 3.2 with the Task Space Inverse Dynamics Framework [26] based on the Pinocchio library [28] and run it at 1 kHz, feeding the previously computed joint state $\dot{\boldsymbol{q}}_{\mathrm{des}}, \boldsymbol{q}_{\mathrm{des}}$ back as to not get stuck due to the trailing impedance controller. At 15 Hz, we forward the computed $\dot{\boldsymbol{q}}_{\mathrm{des}}, \boldsymbol{q}_{\mathrm{des}}$ to the joint impedance controller.

Figure 3: Button functionality of teleoperator's hand controller. The red dot marks the origin of the frame whose twist is tracked.

We evaluate our framework on three tasks taken from the household environment, for two of which we conduct a small user study with three participants, and an additional third task on which we demonstrate feasibility of our proposed approach. These experiments are to be understood as proof-of-concept.

Motivated by the task of carrying liquid in a cup, we choose the end-effector task during WBM mode to be no linear or angular twist in the x and y directions while still staying at the same pose relative to the robot's base.

4.1 Evaluation Setup

Demonstration: Lightswitch Activation This demonstration highlights the effectiveness of our approach in a practical scenario. The robot is tasked with activating a light switch on the wall using its elbow, while simultaneously and stably moving its end-effector, which is holding an object. This setup showcases the robot's ability to perform complex, multi-contact tasks that require precise coordination between different parts of its body. Video of this demonstration is available on `https://sites.google.com/view/aawbt/home`.

User Study Setup This evaluation aims to test both the usability of our teleoperation approach and the effectiveness of the whole-body controller (WBC). Three participants were asked to complete each task three times after an adequate introduction period. The following metrics were recorded:
a. *Success rate* Completion of the task without fatal collision with the environment or the robot's embodiment, showcasing the feasibility of our approach.
b. *Teleoperated distance* The total distance covered by the teleoperator to complete the task, highlighting the usability and efficiency of the system.
c. *Time until completion* The total time required to finish the task, including reorientations and pauses for decision-making, providing insights into the overall usability of the approach.
d. *Task objective error during WBC mode* The deviation from the desired task-specific motion (twist) due to the balancing of multiple objectives, showcasing the effectiveness and design of the WBC in maintaining task performance.

Task 1: Drawer Retrieval In this task, the robot must open a standard kitchen drawer located below the counter, retrieve a banana from inside, and then close the drawer using its elbow on the same arm that holds the banana. This task demonstrates the robot's ability to perform multiple subtasks with a single limb, utilizing whole-body coordination. A depiction of the optimal task completion as a result of successful teleoperation and control can be seen in Figure 4.

Figure 4: Desired execution of the Drawer Retrieval task

Task 2: Cupboard Retrieval In this task, the robot must close a cupboard door that extends from the floor to torso height, while keeping its end-effector stable. The task assumes that the robot has grasped a cup (potentially filled with liquid) from inside the cupboard. To complete the task, the robot must close the cupboard door using the elbow of the same arm that is holding the cup, ensuring stability and avoiding any spillage. This task demonstrates the robot's ability to perform delicate manipulation while simultaneously executing whole-body motions. A depiction of the optimal task completion as a result of successful teleoperation and control can be seen in Figure 5.

Figure 5: Desired execution of the Cupboard Retrieval task

4.2 Results

We report the successful execution of all evaluation tasks over all trials and participants. Videos from our experiments are available on our website.

For the teleoperated tasks, we report an average execution time of 3.76 minutes and a teleoperated distance of 2.05 meters for the drawer retrieval task. For the cupboard retrieval task, the average execution time was 5.25 minutes, with a teleoperated distance of 2.73 meters. Due to the limited number of participants, we refrain from reporting a standard deviation, although we observed high variability in these metrics. This variability, alongside the small sample size, highlights the low representativeness of the study. While these results demonstrate the feasibility of our approach, they also reveal areas for improvement in terms of intuitiveness, as no human would take this long to complete such tasks.

In Figure 6, we present the deviation in task objective velocity for the end-effector, elbow, and base during the whole-body manipulation mode. Aside from brief spikes in the end-effector plot—caused by re-orientation pauses, during which the system temporarily switches to end-effector mode—the task was successfully executed. We observed minimal deviation from the desired teleoperated velocity for both the base and elbow. However, in the final 10 seconds, the base motion exhibited significant deviation from the desired signal, suggesting a trade-off with the end-effector task within the whole-body controller.

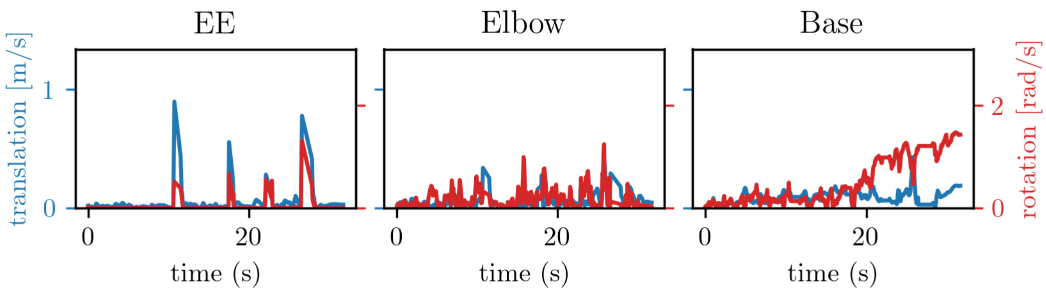

Figure 6: Deviation from desired task objective velocity for end-effector, elbow, and base, computed by the whole-body controller, only during WBM teleoperation mode from one trajectory.

Overall, these results provide proof of concept for our approach and the design of our whole-body controller. The insights gained from this study will guide future improvements, particularly in enhancing the intuitiveness and efficiency of our teleoperation system.

5 Conclusion

In this work, we tackled the challenge of creating human-like yet robot-favorable motions for mobile, multi-DOF manipulators, addressing complex household tasks that require whole-body coordination. Unlike most state-of-the-art methods, which often focus solely on teleoperating end-effector actions, we propose leveraging the entire robotic embodiment to enhance the interaction between the robot and its environment. Our teleoperation framework, supported by a compliant whole-body controller, provides a flexible platform for enabling these interactions, offering the potential for both learning from demonstrations and efficient task execution. We tested our approach on two household tasks while showcasing a complex light switch one, demonstrating the ability of our approach to handle multiple subtasks simultaneously, thus validating the framework's suitability for complex, long-horizon tasks. The ability to simultaneously control different parts of the robot's kinematic chain allowed for more fluid and natural task execution, resembling the way humans solve similar problems. Our preliminary results indicate that our method provides a significant step forward in enabling mobile manipulators to perform tasks combining efficiency and adaptability, bringing robots closer to achieving human-like behaviors in real-world environments.

Limitations and Future Work

One of the key *limitations* of our current approach is the lack of representativeness of the user study. Operators had limited time to familiarize themselves with the system. While a learning curve was observed, it impacted their ability to fully exploit the teleoperation framework within the study period. A more extensive user study with a more extended familiarization period is necessary to better understand the system's full potential and account for user interaction experience.

For *future work*, we plan to address several areas of improvement: **(i) Base motion in end-effector mode:** We wish to enable base motion during end-effector mode, which would allow the robot to reach areas farther away with improved manipulability. Currently, the teleoperator may experience situations where the end-effector doesn't move because it cannot reach the desired orientation. This can lead to confusion, as the teleoperator may not understand why the robot is not responding. Incorporating base collision avoidance will be essential to ensure safe and effective movement. **(ii) Increasing degrees of freedom for elbow control:** Another goal is to enable more degrees of freedom for the elbow, including teleoperating in x, y, and z dimensions and integrating torso movement. This will become feasible after implementing base collision avoidance, allowing more precise whole-body motion control during teleoperation. **(iii) Learning-based impedance adaptation:** We aim to explore automatic impedance adaptation to reduce strain on the robot's arm during interactions with articulated objects, such as doors and drawers. By learning to adapt impedance dynamically, the robot can improve both the precision and safety of its interactions while improving user experience during teleoperation.

Acknowledgments

This research received funding from the European Union's Horizon program under grant agreement no. 101135959, project ARISE, the German Research Foundation (DFG) Emmy Noether Programme (CH 2676/1-1) and the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and the Arts, hessian.AI.

References

[1] S. Dass, W. Ai, Y. Jiang, S. Singh, J. Hu, R. Zhang, P. Stone, B. Abbatematteo, and R. Martín-Martín. TeleMoMa: A Modular and Versatile Teleoperation System for Mobile Manipulation, Mar. 2024.

[2] J. Cheng, F. Abi-Farraj, F. Farshidian, and M. Hutter. Haptic Teleoperation of High-dimensional Robotic Systems Using a Feedback MPC Framework. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6197–6204, Oct. 2022. doi:10.1109/IROS47612.2022.9981290.

[3] K. Darvish, L. Penco, J. Ramos, R. Cisneros, J. Pratt, E. Yoshida, S. Ivaldi, and D. Pucci. Teleoperation of Humanoid Robots: A Survey. *IEEE Transactions on Robotics*, 39(3):1706–1727, June 2023. ISSN 1941-0468. doi:10.1109/TRO.2023.3236952.

[4] Z. Fu, T. Z. Zhao, and C. Finn. Mobile ALOHA: Learning Bimanual Mobile Manipulation with Low-Cost Whole-Body Teleoperation, Jan. 2024.

[5] B. Galarza, P. Ayala, S. Manzano, and M. Garcia. Virtual Reality Teleoperation System for Mobile Robot Manipulation. *Robotics*, 12:163, Nov. 2023. doi:10.3390/robotics12060163.

[6] T. He, Z. Luo, W. Xiao, C. Zhang, K. Kitani, C. Liu, and G. Shi. Learning Human-to-Humanoid Real-Time Whole-Body Teleoperation, Mar. 2024.

[7] L. Penco, K. Momose, S. McCrory, D. Anderson, N. Kitchel, D. Calvert, and R. J. Griffin. Mixed Reality Teleoperation Assistance for Direct Control of Humanoids. *IEEE Robotics and Automation Letters*, 9(2):1937–1944, Feb. 2024. ISSN 2377-3766. doi:10.1109/LRA.2024.3349809. URL `https://ieeexplore.ieee.org/document/10380694`.

[8] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang. Open-television: Teleoperation with immersive active visual feedback. *arXiv preprint arXiv:2407.01512*, 2024.

[9] J. Borràs and T. Asfour. A whole-body pose taxonomy for loco-manipulation tasks. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1578–1585, Sept. 2015. doi:10.1109/IROS.2015.7353578.

[10] H. Xiong, R. Mendonca, K. Shaw, and D. Pathak. Adaptive mobile manipulation for articulated objects in the open world. *arXiv preprint arXiv:2401.14403*, 2024.

[11] T. Yang, Y. Jing, H. Wu, J. Xu, K. Sima, G. Chen, Q. Sima, and T. Kong. MOMA-Force: Visual-Force Imitation for Real-World Mobile Manipulation, Aug. 2023. URL `http://arxiv.org/abs/2308.03624`.

[12] S. Yang, M. Liu, Y. Qin, R. Ding, J. Li, X. Cheng, R. Yang, S. Yi, and X. Wang. Ace: A cross-platform visual-exoskeletons system for low-cost dexterous teleoperation. *arXiv preprint arXiv:2408.11805*, 2024.

[13] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel. GELLO: A general, low-cost, and intuitive teleoperation framework for robot manipulators. *CoRR*, abs/2309.13037, 2023. doi:10.48550/ARXIV.2309.13037. URL `https://doi.org/10.48550/arXiv.2309.13037`.

[14] Z. Fu, X. Cheng, and D. Pathak. Deep Whole-Body Control: Learning a Unified Policy for Manipulation and Locomotion. In *Proceedings of The 6th Conference on Robot Learning*, pages 138–149. PMLR, Mar. 2023.

[15] M. Arduengo, A. Arduengo, A. Colomé, J. Lobo-Prat, and C. Torras. Human to Robot Whole-Body Motion Transfer. In *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, pages 299–305, July 2021. doi:10.1109/HUMANOIDS47582.2021.9555769. URL `http://arxiv.org/abs/1909.06278`.

[16] T. He, Z. Luo, X. He, W. Xiao, C. Zhang, W. Zhang, K. Kitani, C. Liu, and G. Shi. OmniH2O: Universal and Dexterous Human-to-Humanoid Whole-Body Teleoperation and Learning, June 2024.

[17] J. Humphreys, C. Peers, J. Li, Y. Wan, and C. Zhou. High Utility Teleoperation Framework for Legged Manipulators Through Leveraging Whole-Body Control. *Journal of Intelligent & Robotic Systems*, 108(3):57, July 2023. ISSN 1573-0409. doi:10.1007/s10846-023-01866-7.

[18] F. Wang, F. Chen, Y. Dong, Q. Yong, X. Yang, L. Zheng, X. Zhang, and H. Su. Whole-Body Teleoperation Control of Dual-Arm Robot Using Sensor Fusion. *Biomimetics*, 8(8):591, Dec. 2023. ISSN 2313-7673. doi:10.3390/biomimetics8080591.

[19] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang. Expressive whole-body control for humanoid robots. *arXiv preprint arXiv:2402.16796*, 2024.

[20] J. Nakanishi, S. Itadera, T. Aoyama, and Y. Hasegawa. Towards the development of an intuitive teleoperation system for human support robot using a VR device. *Advanced Robotics*, 34(19): 1239–1253, Oct. 2020. ISSN 0169-1864. doi:10.1080/01691864.2020.1813623.

[21] M. Seo, S. Han, K. Sim, S. H. Bang, C. Gonzalez, L. Sentis, and Y. Zhu. Deep Imitation Learning for Humanoid Loco-manipulation through Human Teleoperation, Nov. 2023. URL `http://arxiv.org/abs/2309.01952`.

[22] R. Ding, Y. Qin, J. Zhu, C. Jia, S. Yang, R. Yang, X. Qi, and X. Wang. Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning. *arXiv preprint arXiv:2407.03162*, 2024.

[23] O. X.-E. Collaboration, A. O'Neill, A. Rehman, A. Gupta, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta, A. Mandlekar, A. Jain, A. Tung, A. Bewley, A. Herzog, A. Irpan, A. Khazatsky, A. Rai, A. Gupta, A. Wang, A. Kolobov, A. Singh, A. Garg, A. Kembhavi, A. Xie, A. Brohan, A. Raffin, A. Sharma, A. Yavary, A. Jain, A. Balakrishna, A. Wahid, B. Burgess-Limerick, B. Kim, B. Schölkopf, B. Wulfe, B. Ichter, C. Lu, C. Xu, C. Le, C. Finn, C. Wang, C. Xu, C. Chi, C. Huang, C. Chan, C. Agia, C. Pan, C. Fu, C. Devin, D. Xu, D. Morton, D. Driess, D. Chen, D. Pathak, D. Shah, D. Büchler, D. Jayaraman, D. Kalashnikov, D. Sadigh, E. Johns, E. Foster, F. Liu, F. Ceola, F. Xia, F. Zhao, F. V. Frujeri, F. Stulp, G. Zhou, G. S. Sukhatme, G. Salhotra, G. Yan, G. Feng, G. Schiavi, G. Berseth, G. Kahn, G. Yang, G. Wang, H. Su, H.-S. Fang, H. Shi, H. Bao, H. B. Amor, H. I. Christensen, H. Furuta, H. Bharadhwaj, H. Walke, H. Fang, H. Ha, I. Mordatch, I. Radosavovic, I. Leal, J. Liang, J. Abou-Chakra, J. Kim, J. Drake, J. Peters, J. Schneider, J. Hsu, J. Vakil, J. Bohg, J. Bingham, J. Wu, J. Gao, J. Hu, J. Wu, J. Wu, J. Sun, J. Luo, J. Gu, J. Tan, J. Oh, J. Wu, J. Lu, J. Yang, J. Malik, J. Silvério, J. Hejna, J. Booher, J. Tompson, J. Yang, J. Salvador, J. J. Lim, J. Han, K. Wang, K. Rao, K. Pertsch, K. Hausman, K. Go, K. Gopalakrishnan, K. Goldberg, K. Byrne, K. Oslund, K. Kawaharazuka, K. Black, K. Lin, K. Zhang, K. Ehsani, K. Lekkala, K. Ellis, K. Rana, K. Srinivasan, K. Fang, K. P. Singh, K.-H. Zeng, K. Hatch, K. Hsu, L. Itti, L. Y. Chen, L. Pinto, L. Fei-Fei, L. Tan, L. J. Fan, L. Ott, L. Lee, L. Weihs, M. Chen, M. Lepert, M. Memmel, M. Tomizuka, M. Itkina, M. G. Castro, M. Spero, M. Du, M. Ahn, M. C. Yip, M. Zhang, M. Ding, M. Heo, M. K. Srirama, M. Sharma, M. J. Kim, N. Kanazawa, N. Hansen, N. Heess, N. J. Joshi, N. Suenderhauf, N. Liu, N. Di Palo, N. M. M. Shafiullah, O. Mees,

O. Kroemer, O. Bastani, P. R. Sanketi, P. T. Miller, P. Yin, P. Wohlhart, P. Xu, P. D. Fagan, P. Mitrano, P. Sermanet, P. Abbeel, P. Sundaresan, Q. Chen, Q. Vuong, R. Rafailov, R. Tian, R. Doshi, R. Mart'in-Mart'in, R. Baijal, R. Scalise, R. Hendrix, R. Lin, R. Qian, R. Zhang, R. Mendonca, R. Shah, R. Hoque, R. Julian, S. Bustamante, S. Kirmani, S. Levine, S. Lin, S. Moore, S. Bahl, S. Dass, S. Sonawani, S. Tulsiani, S. Song, S. Xu, S. Haldar, S. Karamcheti, S. Adebola, S. Guist, S. Nasiriany, S. Schaal, S. Welker, S. Tian, S. Ramamoorthy, S. Dasari, S. Belkhale, S. Park, S. Nair, S. Mirchandani, T. Osa, T. Gupta, T. Harada, T. Matsushima, T. Xiao, T. Kollar, T. Yu, T. Ding, T. Davchev, T. Z. Zhao, T. Armstrong, T. Darrell, T. Chung, V. Jain, V. Kumar, V. Vanhoucke, W. Zhan, W. Zhou, W. Burgard, X. Chen, X. Chen, X. Wang, X. Zhu, X. Geng, X. Liu, X. Liangwei, X. Li, Y. Pang, Y. Lu, Y. J. Ma, Y. Kim, Y. Chebotar, Y. Zhou, Y. Zhu, Y. Wu, Y. Xu, Y. Wang, Y. Bisk, Y. Dou, Y. Cho, Y. Lee, Y. Cui, Y. Cao, Y.-H. Wu, Y. Tang, Y. Zhu, Y. Zhang, Y. Jiang, Y. Li, Y. Li, Y. Iwasawa, Y. Matsuo, Z. Ma, Z. Xu, Z. J. Cui, Z. Zhang, Z. Fu, and Z. Lin. Open X-Embodiment: Robotic Learning Datasets and RT-X Models, June 2024. URL http://arxiv.org/abs/2310.08864.

[24] A. Gottardi, S. Tortora, E. Tosello, and E. Menegatti. Shared Control in Robot Teleoperation With Improved Potential Fields. *IEEE Transactions on Human-Machine Systems*, 52(3):410–422, June 2022. ISSN 2168-2305. doi:10.1109/THMS.2022.3155716. URL https://ieeexplore.ieee.org/abstract/document/9734752.

[25] Z. Fu, Q. Zhao, Q. Wu, G. Wetzstein, and C. Finn. Humanplus: Humanoid shadowing and imitation from humans. *arXiv preprint arXiv:2406.10454*, 2024.

[26] A. Del Prete, N. Mansard, O. Ramos, O. Stasse, and F. Nori. Implementing Torque Control with High-Ratio Gear Boxes and Without Joint-Torque Sensors. *International Journal of Humanoid Robotics*, 13, Mar. 2015. doi:10.1142/S0219843615500449.

[27] S. Lueth. Whole-body teleoperation of mobile manipulation robots. *Unpublished Master's Thesis, TU Darmstadt*, 2024.

[28] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiraux, O. Stasse, and N. Mansard. The Pinocchio C++ library : A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives. In *2019 IEEE/SICE International Symposium on System Integration (SII)*, pages 614–619, Paris, France, Jan. 2019. IEEE. ISBN 978-1-5386-3615-2. doi:10.1109/SII.2019.8700380. URL https://ieeexplore.ieee.org/document/8700380/.