Bimanual Interaction Technique for Controlling Humanoid Robot Walking and Working

Joon Hyub Lee*
DRB-KAIST SketchTheFuture Research Center
KAIST
Daejeon, Republic of Korea
joonhyub.lee@kaist.ac.kr

Siripon Sutthiwanna*
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
siripon.s@kaist.ac.kr

Sang-Hyun Lee
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
sang-hyun.lee@kaist.ac.kr

Jin Park
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
jin.park@kaist.ac.kr

Seok-Hyung Bae
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
seokhyung.bae@kaist.ac.kr

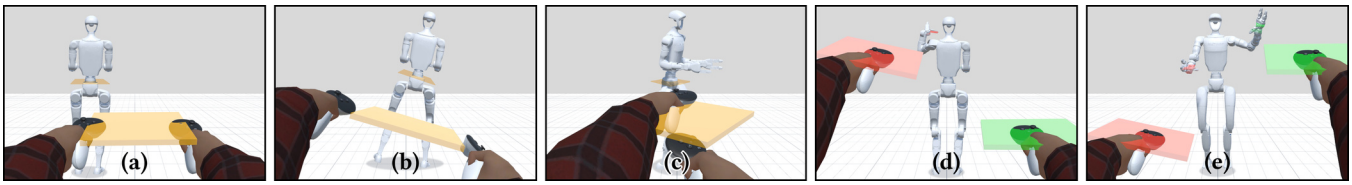


Figure 1: We propose a novel interaction technique for controlling a humanoid robot in real time using bimanual hand gestures. (a) The user activates the robot's *walking mode* by lowering hands and holding a virtual 3D plane corresponding to the plane at the robot's pelvis, (b) tilts the plane to translate the robot in the tilted direction, and (c) turns the plane to rotate the robot in the turned direction. (d) The user activates the robot's *working mode* by raising hands and holding two virtual 3D planes corresponding to those at the robot's hands, and translates and rotates each plane to move each robot hand by the same amount. (e) When the user and the robot face each other, the user's left hand controls the robot's right hand, and vice versa.

Abstract

Humanoid robots are difficult for novice users to control intuitively due to their high degrees of freedom. We propose a novel interaction technique that enables users to control the locomotion and manipulation of a humanoid robot by holding and moving a virtual 3D plane with both hands. We implemented our technique using a Unitree G1 humanoid robot model in a VR physics simulation and compared it against gamepad and exoskeleton baselines on tasks from an open-source humanoid robot benchmark in a user study with 24 participants. Our technique was 59% faster and improved benchmark results by 17% in a locomotion task, while it was 41% faster and improved benchmark results by 20% in a manipulation task. It also outperformed the baselines on most subjective workload measures.

CCS Concepts

• **Human-centered computing** → **Interaction techniques**.

*The authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST Adjunct '26, Detroit, MI, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2855-6/26/11
https://doi.org/10.1145/3830397.3841828

Keywords

Humanoid robot; locomotion; manipulation; hand gesture; VR

ACM Reference Format:
Joon Hyub Lee, Siripon Sutthiwanna, Sang-Hyun Lee, Jin Park, and Seok-Hyung Bae. 2026. Bimanual Interaction Technique for Controlling Humanoid Robot Walking and Working. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '26), November 02–05, 2026, Detroit, MI, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3830397.3841828

1 Introduction

Humanoid robots have recently become consumer products that can be purchased from online stores much like home appliances. Imagining scenes from science fiction films in which robots assist with everyday tasks, customers eagerly await their robot's arrival, which can be as soon as the next day. After unboxing the robot, they can operate it using the included gamepad, making it walk with thumbsticks or wave its arms by pressing dedicated buttons.

However, interaction between users and humanoid robots often ends there. Current humanoid robots are still limited in their ability to autonomously perform everyday tasks [2]. At the same time, there is no quick and easy way for users to directly control the robots' high degrees of freedom. As a result, expensive humanoid robots may end up standing idle like statues, collecting dust.

For humanoid robots to become practical, users should be able to communicate their intended actions with minimal physical and mental effort. In this study, we propose a novel interaction technique for doing so in real time (Figure 1). We implemented the technique in a VR physics simulation and quantitatively and qualitatively evaluated its usability against widely used baselines with an open-source humanoid robot task benchmark. We report promising preliminary results.

2 Interaction Technique

We propose a novel interaction technique that enables real-time humanoid robot control through bimanual gestures similar to holding and moving a virtual 3D plane. The user can move the robot to a desired location from a comfortable viewpoint, perform precise bimanual manipulation, and seamlessly switch between locomotion and manipulation without explicitly pressing buttons or invoking commands, simply by initiating the same gesture at different heights. Robot control using a 3D plane metaphor has previously been shown to be effective for collaborative robots [3], and this work is the first to apply it to humanoid robots.

Walking mode. The user activates the robot's walking mode by lowering hands to a comfortable resting height and holding a virtual 3D plane horizontally (Figure 1a). This plane acts as a remote handle for the robot's pelvis, which serves as a natural frame of reference for locomotion. The user commands the robot to move by tilting (Figure 1b) or turning (Figure 1c) the plane.

The direction of tilt or turn is interpreted in the global coordinate system shared by the user and the robot, so the user can intuitively express the desired movement direction without the mental cost of transforming it to the robot's frame of reference. Larger tilt or turn angles result in faster robot movements.

Working mode. The user activates the robot's working mode by raising hands to a height where manipulation typically occurs and holding a virtual 3D plane in each hand. These planes act as remote handles for the robot's hands. Moving each plane moves the corresponding robot hand by the same positional and orientational changes, again in the global coordinate system (Figure 1d).

By default, the user's left and right hands control the robot's left and right hands, respectively. However, when the user and the robot face each other, this mapping is reversed so that the user can work intuitively as if looking into a mirror (Figure 1e). Releasing the plane disconnects the user's hand from the robot's hand, allowing the user to clutch by repositioning the hand and then re-engaging the gesture, thereby reducing wrist strain. Squeezing the index trigger of the VR controller closes the robot's hand.

3 Implementation

To prototype and simulate the user experience before applying our technique to a physical robot, we implemented our technique using Unity, its built-in physics engine, a model of the Unitree G1 humanoid robot, and a Meta Quest 3 VR headset. We used BioAnimation [6] to generate natural walking motion, and BioIK [5] to generate entire arm poses from the robot hand positions and orientations. The system ran on a PC with an Intel Core i9-14900HX CPU, an NVIDIA GeForce RTX 4070 GPU, and 32 GB of RAM.

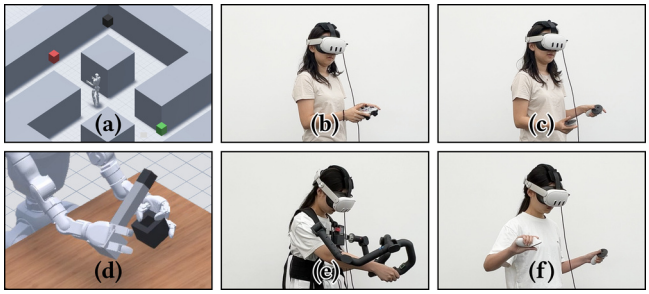


Figure 2: Participants performed (a) the Maze task of passing through waypoints without touching walls, using (b) a gamepad baseline and (c) our technique, and (d) the Insert task of assembling color-matched blocks onto both ends of a stick, using (e) an exoskeleton baseline and (f) our technique.

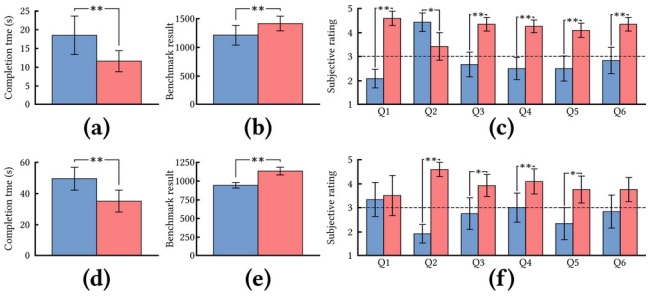


Figure 3: Our technique completed the Maze task (a) 59% faster and (b) with 17% higher benchmark results, and (c) scored better except on Q2. Our technique also completed the Insert task (d) 41% faster and (e) with 20% higher benchmark results, and (f) scored better except on Q1 and Q6 of the NASA-TLX 5-point Likert scale questions (Q1: low mental demand, Q2: low physical demand, Q3: low temporal demand, Q4: high performance, Q5: low effort, and Q6: low frustration), where higher is better. Blue: baseline, red: our technique, *: $p < 0.05$, **: $p < 0.01$, error bars: ±2 SE.

4 Evaluation

We conducted a user study based on *HumanoidBench* [4], an open-source humanoid robot task benchmark, to quantitatively and qualitatively compare the usability of our technique against baselines.

HumanoidBench was originally designed to evaluate AI task performance. However, because its task environments, assets, objectives, and scoring formulas are well-defined, it is also suitable for comparing human task performance across different interaction techniques. We compared our technique with a gamepad, which is widely used for robot locomotion [7], in the *Maze* task, and with an exoskeleton, which is widely used for bimanual robot manipulation [1], in the *Insert* task (Figure 2).

We recruited 24 participants in total (12 male and 12 female), with 12 participants performing each task. After a brief tutorial, participants performed ten trials using both our technique and the baseline in a counterbalanced order. We measured task completion times and benchmark results, and recorded participants' subjective workload using NASA-TLX (Figure 3).

5 Conclusion & Future Work

We proposed an interaction technique that attaches virtual 3D planes to the pelvis and two hands of a humanoid robot as remote handles for controlling locomotion and bimanual manipulation, and seamlessly switching between them. Our technique enables users to observe the robot from their desired viewpoints and quickly and easily control its movements in real time through hand motions in directions that are comfortable and intuitive to them.

In the user study, all participants were able to learn to use the proposed technique proficiently within a short time. Preliminary results show that our technique outperformed the gamepad baseline specialized for robot locomotion in the locomotion task, and also outperformed the exoskeleton baseline specialized for bimanual robot manipulation in the manipulation task. Our technique also achieved better results on most task workload measures.

Future work should further analyze these early findings in depth and evaluate the proposed technique in composite tasks that require frequent switching between locomotion and manipulation, using a full-body motion-capture baseline [8] for comparison. In addition, the proposed technique should be deployed on a physical humanoid robot and tested across a wide range of real-world use scenarios.

Acknowledgments

This research was supported by the DRB-KAIST SketchTheFuture Research Center, and the Technology Commercialization Support Program through the Korea Innovation Foundation funded by the Ministry of Science and ICT.

References

[1] Qingwei Ben, Feiyu Jia, Jia Zeng, Junting Dong, Dahua Lin, and Jiangmiao Pang. 2025. HOMIE: humanoid loco-manipulation with isomorphic exoskeleton cockpit. In *Proc. RSS '25*. Article 70, 21 pages.

[2] Zhaoyuan Gu, Junheng Li, Wenlan Shen, Wenhao Yu, Zhaoming Xie, Stephen McCrory, Xianyi Cheng, Abdulaziz Shamsah, Robert Griffin, C. Karen Liu, Abderrahmane Kheddar, Xue Bin Peng, Yuke Zhu, Guanya Shi, Quan Nguyen, Gordon Cheng, Huijun Gao, and Ye Zhao. 2026. Humanoid locomotion and manipulation: current progress and challenges in control, planning, and learning. *TMECH* 31, 2 (2026), 2300–2330.

[3] Joon Hyub Lee, Yongkwan Kim, Sang-Gyun An, and Seok-Hyung Bae. 2020. Robot telekinesis: application of a unimanual and bimanual object manipulation technique to robot control. In *Proc. ICRA '20*. 9866–9872.

[4] Carmelo Sferrazza, Dun-Ming Huang, Xingyu Lin, Youngwoon Lee, and Pieter Abbeel. 2024. HumanoidBench: simulated humanoid benchmark for whole-body locomotion and manipulation. In *Proc. RSS '24*. Article 61, 14 pages.

[5] Sebastian Starke, Norman Hendrich, Dennis Krupke, and Jianwei Zhang. 2017. Evolutionary multi-objective inverse kinematics on highly articulated and humanoid robots. In *Proc. IROS '17*. 6959–6966.

[6] Sebastian Starke, He Zhang, Taku Komura, and Jun Saito. 2019. Neural state machine for character-scene interactions. *TOG* 38, 6, Article 178 (2019), 14 pages.

[7] Unitree Robotics. 2025. *G1 Remote Control: User Manual*. Retrieved July 10, 2026 from https://marketing.unitree.com/article/en/G1/Remote_Control.html

[8] Yanjie Ze, Siheng Zhao, Weizhuo Wang, Angjoo Kanazawa, Rocky Duan, Pieter Abbeel, Guanya Shi, Jiajun Wu, and C. Karen Liu. 2026. TWIST2: scalable, portable, and holistic humanoid data collection system. In *Proc. ICRA '26*. To appear.