2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Developing a Framework for Natural Human Movement Mimicry of Low-Dynamic Motions in Mobile-Based Humanoids

Simon Gormuzov[1], Yushi Wang[2], Pinchu Yang[2], Tamon Miyake[2], Tetsuya Ogata[3,4], Shigeki Sugano[3]

Abstract— **In this work, we propose a framework that facilitates natural whole-body movements for a mobile-based humanoid. The framework takes bipedal human motion animation as input. After re-targeting the motion to the robot rig, the joint space data is applied to a physics-enabled robot model for balance confirmation and then deployed one-shot to the real robot. This method is beneficial for: 1) mapping expressive and low-dynamic whole-body motions, such as walking, to mobile-based robots; and 2) serving as a basis for training more complex control policies for more dynamic motions. Experiments were conducted by deploying human walking animations to the robot, assessing its ability to mirror the movements, and evaluating the subjective feelings of humans observing the robot performing the motions generated by both the proposed method and the traditional method. The results indicated that the proposed method is effective for mimicking human movements and consistently delivered a better overall impression in the natural appearance of the motion, the human-like factor, and friendliness.**

I. Introduction

The aging population has led to a labor shortage, particularly in hospitals and retirement homes. To address this issue, there is a growing need for the adoption of general-purpose humanoid robots. However, for these robots to be effectively integrated and widely adopted, advancements in intelligent human-robot interaction are essential. This includes enhancing the robots' ability to mimic human behavior and interactions, making them more relatable and efficient in caregiving roles.

General-purpose humanoids based on their primary functions can be broadly categorized into physically assistive robots and socially assistive robots [1][2][3]. Socially assistive robots are primarily used for emotional and cognitive care. Therefore, it is crucial to carefully design and implement natural expressions. Specifically, natural movements are characterized by smooth transitions, bio-mechanically accurate motion trajectories, realistic speed and acceleration profiles, and the absence of abrupt or discontinuous changes, aligning with the kinematic and dynamic patterns observed in human motion. On the other hand, physically assistive robots require higher power output, robust physical assistance capabilities, and coordination with human movements to provide effective support. When robots demonstrate human-like behaviors, people are more inclined to synchronize their movements, resulting in more effective and seamless human-robot interactions. Furthermore, robots that exhibit familiar, human-like behaviors are more readily accepted and perceived as approachable, enhancing comfort levels during such interactions [4]. It is necessary to develop methods for generating human-like movements to ensure effective use of such robots.

Recent work has exhibited the possibility of learning motion and skills from human data with a bipedal robot using a combination of real-time shadowing and imitation learning [5][6]. However, to ensure improved safety and eliminate the risk of erroneous or unstable motions resulting from model hallucinations, an empirical one-shot approach offers a robust framework for capturing and reproducing natural movements with high fidelity.

In terms of morphology, humanoid robots can also be broadly categorized into two types: bipedal robots and mobile-based robots. Bipedal robots, such as the renowned Atlas [7], are equipped with two legs and are particularly adept at tasks related to disaster relief, including navigating uneven outdoor terrain. In contrast, mobile-based robots often have a mechanically simplified lower body, which enhances the stability of the upper body, particularly the dual arms, enabling efficient object manipulation or physical assistance of humans, for example, the TWENDY-ONE robot [8]. Significant effort has been invested in the dynamic locomotion and motion generation of bipedal robots, and recent advancements have been made through the integration of deep learning techniques, especially by utilizing data from motion capture systems or computer graphics (CG) animations. Researchers have developed methods for bipedal robots to imitate human movements, which involve tracking the positions of key markers on the human or animation and solving joint angle trajectories. For bipedal humanoid robots with body structures similar to humans, this strategy inherently benefits from the ease of mapping joint positions. On the other hand, the motion of mobile-based humanoids which focus on manipulation using their two arms, are often taught by a dual-armed passive device, such as the mobile ALOHA [9]. However, such devices are insufficient for teaching whole-body motion. While Deep Reinforcement Learning (DRL) and other Machine Learning (ML) approaches can generate complex motions, they raise

*This work was supported by JST Moonshot RD, Grant No. JPMJMS2031.

[1]Simon Gormuzov is with Waseda University, Graduate School of Creative Science and Engineering, Tokyo, Japan.(Corresponding author: Simon Gormuzov, e-mail: s.gormuzov@fuji.waseda.jp)

[2]Yushi Wang, Pinchu Yang, and Tamon Miyake are with the Future Robotics Organization, Waseda University, Tokyo, Japan.

[3]Tetsuya Ogata and Shigeki Sugano are with the Faculty of Science and Engineering, Waseda University, Tokyo, Japan.

[4]Tetsuya Ogata is with The National Institute of Advanced Science and Technology, Tokyo, Japan.

significant safety concerns in the context of a mobile-based robot designed to interact with humans. These methods can produce unpredictable or inaccurate behaviors, which greatly increases risk of operating the robot around humans. An empirical one-shot approach that directly maps real or animated human motions to mobile-based humanoids enhances safety by ensuring the generated movements are predictable and physically feasible.

The potential to map human motions to mobile-based humanoids for data-driven whole-body natural motion generation has caught the author's attention .

In this study, we utilized a mobile-based humanoid robot, AIREC, as shown in Fig. 1, such robots are designed primarily to assist humans in homes or facilities such as hospitals and retirement homes [10]. The ability to exhibit natural movements as well as to provide effective physical assistance is vital for it.

Therefore, we aim to address the following issues:

- Solving the morphological discrepancy between human movements and mobile-based robot movements using proposed Mobile Anchor Position Tracking method (MAPT).
- Validating natural whole-body movements by creating a physics-enabled model of the robot before deploying the motion on the physical machine
- Developing a basis for safe and fluid low dynamic animation to full-body motion on mobile-based robots.

In this work, we present a framework that addresses structural differences between human and robotic movement using MAPT. The process begins with generating motions from a humanoid rig driven by CG animations. These motions are then tested and validated on a real-time physics-based robot model to ensure safety and realism. Once validated, the refined motions are deployed in real-time to the physical robot, enabling effective and natural movement. This framework has the potential to repeatably transfer human motion sequences into positional control for the base to control the robot's position while controlling the motions of the torso, head and arms of the robot. This is intended to contribute to more natural and expressive motion for mobile-based humanoids.

The remainder of this paper is organized as follows. Section II presents a review of related work. The proposed method is then introduced in Section III. Section IV describes the experiments and their results. Finally, future work is provided in Section V.

II. Related Work

Whole-body motion generation for bipedal robots is a complex and critical area of research. Optimization-based methods, such as those presented by [7] for the Atlas robot and [11] for natural-looking motion primitives, are commonly used but often struggle to adapt to new conditions after initial optimization. In contrast, using an explicit robot model allows for precise control by understanding kinematics and dynamics, enabling better planning and decision-making, as highlighted by [12][13][14]. However, approaches like

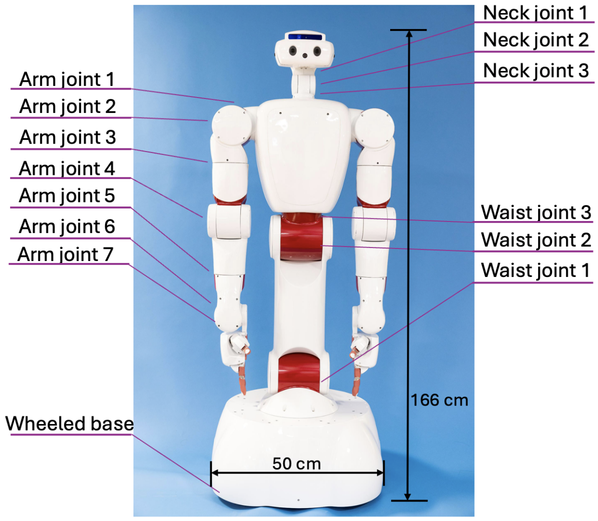

Fig. 1. Size and appearance of the AIREC robot

Model Predictive Control (MPC), which solve optimization problems at each step, can be computationally intensive. While balance for humanoid robots has been studied extensively [15] it remains a challenge in robot motion to this day. Research has paid little attention to the generation of expressive whole-body motions of mobile-based robots that go beyond this foundational issue. Additionally, the aforementioned methods are not well-suited for achieving expressive, low-dynamic motions, which are crucial for more natural and human-like robot movements.

On the other hand, robots with a fixed or mobile base, often used in controlled facility environments, do not face the same balance challenges as bipedal robots. As a result, achieving balance especially in low dynamic motions requires less computational effort. For these base robots, the focus shifts to generating more expressive and natural whole-body motions that mimic human movements, which is crucial for effective human-robot interaction.

In addition to model-based control, learning-based control methods are also rapidly advancing [16][17]. These learning-based methods can achieve robust performance in humanoid robots by training in highly diverse simulated environments, which enhances their adaptability, as demonstrated in [18][19][20].

For mobile-based robots, utilizing recent advancements in learning-based methods for human motion imitation could be beneficial for achieving natural and robust movements. However, directly applying the successful learning-based techniques developed for bipedal robots is impractical due to structural differences as well as unsafe in the context of operating around humans. Consequently, there is currently no effective foundation for applying human-like motions to base robots to achieve natural movement.

III. Methods

A. Utilized Tools

The primary platform for this research was the AIREC mobile-based humanoid robot (see, Fig. 1), designed to assist humans in environments such as homes, hospitals, and retirement homes. This robot was paired with a workstation PC to run software, generate, and transmit commands.

IK Rig facilitated accurate mapping of human movements to the virtual humanoid skeleton, calculating positions of connected body parts based on key joint movements [21]. Unity's physics engine ensured that the simulated motions adhered to real-world physics as well as the robot's physical properties, enhancing realism and applicability. ROS provided a reliable framework for translating motion data into actionable commands.

B. Human data mirroring

1) 3 kinds of Rigs: There are three kinds of rigs involved in the motion transfer flow for human movement mimicry in this work:

- Humanoid Rig Armature: A digital skeleton used in CG animation and 3D modeling to replicate human-like movements. It offers high flexibility with joints that can move freely in multiple directions, making it ideal for a wide range of animations. However, this flexibility is not suitable for real robot operations, where precise joint constraints are necessary.
- Human-Robot Rig Armature: An intermediary rig that adapts human-like movements to robot models. It captures human motion and transfers it to a robot's proportions while retaining the essence of the movement. Unlike real robots, its joints are not restricted to single-axis motions, providing flexibility for initial animations but lacking precision for practical robotic applications. It is used to bridge human motion to robotic form before finalizing with a more accurate robot rig.
- Robot Rig: A rig specifically designed to reflect the mechanical constraints and joint mechanics of a real robot. It accurately replicates the robot's joint locations, degrees of freedom, and movement limits to ensure that digital animations can be executed safely and realistically by a physical robot. The robot rig aligns closely with the robot's actual specifications, ensuring that all movements are feasible and operationally safe.

2) Mobile Anchor Position Tracking method (MAPT): We propose the Mobile Anchor Position Tracking (MAPT) method to align upper body motion with lower body movement. In this approach, we select the neck joint as the anchor point, serving as a stable, central reference near the upper torso's center of mass, providing a consistent basis for movement calculations (see Fig. 2). This aids in maintaining alignment and coherence between upper and lower body motions. By projecting the neck joint position onto the ground to determine the base position and calculating lower body joint angles using inverse kinematics, we dynamically adjust the robot's base and lower body to maintain balance and stability. Crucially, MAPT aligns upper body movements to facilitate lower body motion, addressing structural differences in mobile-based robots.

Initially, we imported human motion sequences from open-source repositories into Unity, mapping them onto humanoid rigs (see Fig. 2).

Alternatively, we developed a framework to transfer real-time human motion to the humanoid rig, but for most results in this paper, we used downloaded animations replicating real-world movements.

In our initial attempt to deploy whole-body bipedal human motion onto a mobile-based robot, we used a motion capture system to mirror human movements onto the AIREC robot rig in Unity. We employed the HTC VIVE system with four trackers—one each for the head, waist, and arms. The robot rig replicated the human's whole-body motion, as shown in Fig. 3. Despite the minimal trackers, the robot correctly mirrored human motion, particularly in the head and waist. This is evident from the sub-figures, where the robot's base moves correspondingly when the human walks sideways. Typically, robots mimicking human motions keep the lower body fixed to simplify balance control, limiting natural whole-body movements crucial for enhancing human-robot interaction. Our MAPT method shows significant potential in enabling fluid, expressive whole-body motions, thereby improving interaction capabilities.

Although the robot rig's motions include all necessary transformations to map human motion to the robot's kinematics, this is insufficient for commanding the actual robot due to its characteristic speed and inertia at each joint and base. A virtual physical environment allows simulation of realistic dynamics, including gravity, friction, and collisions.

C. Physics-enabled modeling

1) Base controller: In order to remedy the issue mentioned above, The physics model of the robot was constructed. For this, the robot rig configuration was used but instead of using a rig, the physics model is constructed out of a set of articulation bodies. Each articulation body representing a single link in the real robot has the appropriate joint ranges, link mass, joint friction, inertia and drive parameters. Theoretically, this ensures that the robot is properly represented in the Unity simulation while preventing unfeasible motions according to the physics limits. The angles from the robot rig are sent to the appropriate joints of the Physics model as target positions. The physics model then calculates the robot movement to reach the target positions of each joint resulting in the transfer of the animation movements to the physics-enabled robot model.

Using the proposed MAPT, the robot's base movement is controlled by tracking the anchor movement, and a custom base controller was developed to apply rotational and translational forces to the physics-enabled model's base for movement and rotation. This controller employs two separate PID controllers in parallel, one for translation and one for rotation. Each PID controller contains integrated smoothing parameters and velocity limits, allowing for further customization of the motion profile, as shown in Eq. 1.

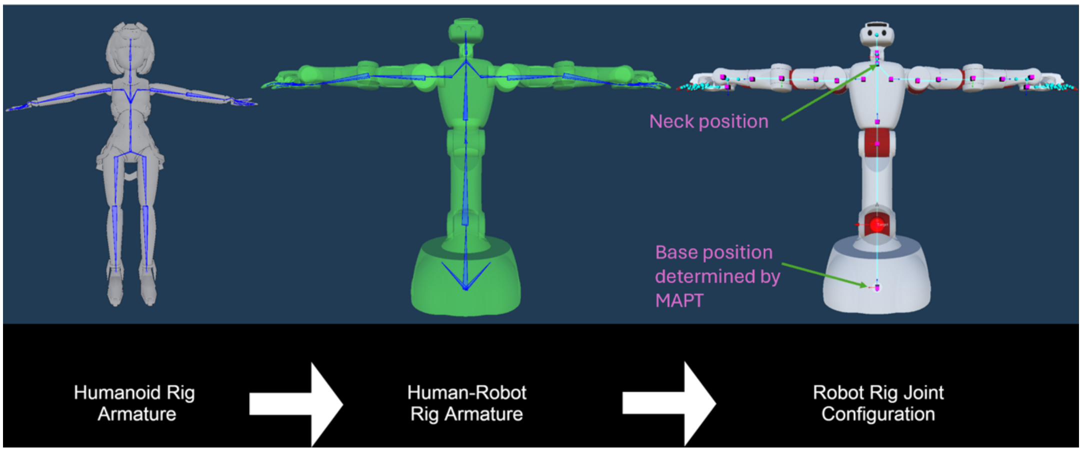

Fig. 2. The motion transfer flow is shown. Humanoid rig armature receives the original animation, providing high flexibility for natural human-like movements. Human-robot Rig Armature acts as an intermediary, adapting the animation to a robot's proportions while retaining human-like motion characteristics but without precise joint constraints. Robot rig transfers the adapted motion to a rig that mirrors the real robot's joint constraints and mechanics, ensuring animations are accurate, safe, and feasible for actual robotic execution. The upper body motion can be easily mapped due to the similar structure between the human rig armature and the robot rig. To address the morphological discrepancies in the lower body, the neck joint position is projected to the ground to determine the base position by MAPT, and the lower body movements, including waist joints 1 and 2, are calculated using inverse kinematics.

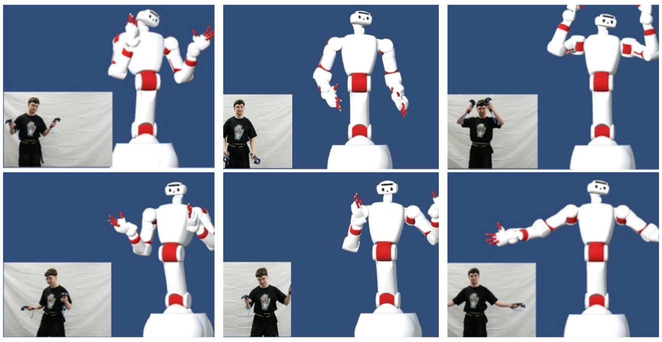

Fig. 3. AIREC robot rig mirroring whole-body motion in real-time from mocap data in Unity.

$$U(s) = \begin{cases} U_{\text{pos}}(s) = (1-\alpha_v)U_{\text{pos,prev}}(s) \\ +\alpha_v \left(K_p + \frac{K_i}{s} + K_d s\right) E_{\text{pos}}(s), \\ U_{\text{rot}}(s) = (1-\alpha_r)U_{\text{rot,prev}}(s) \\ +\alpha_r \left(K_{p,\text{rot}} + \frac{K_{i,\text{rot}}}{s} + K_{d,\text{rot}} s\right) E_{\text{rot}}(s) \end{cases}$$

with constraints:

$$U_{\text{pos}}(s) = \min(\max(U_{\text{pos}}(s), -V_{\max}), V_{\max})$$
$$U_{\text{rot}}(s) = \min(\max(U_{\text{rot}}(s), -\omega_{\max}), \omega_{\max}) \tag{1}$$

Where the control output for position is denoted by $U_{\text{pos}}(s)$. The position proportional gain is K_p, the position integral gain is K_i, and the position derivative gain is K_d. The error in position is represented by $E_{\text{pos}}(s)$. The velocity smoothing factor is α_v, and the maximum velocity is $V_{\max}$. The previous control output for position is $U_{\text{pos,prev}}(s)$.

For rotation, the control output is denoted by $U_{\text{rot}}(s)$. The rotation proportional gain is $K_{p,\text{rot}}$, the rotation integral gain is $K_{i,\text{rot}}$, and the rotation derivative gain is $K_{d,\text{rot}}$. The error in rotation is represented by $E_{\text{rot}}(s)$. The rotation smoothing factor is α_r, and the maximum rotation velocity is $\omega_{\max}$. The previous control output for rotation is $U_{\text{rot,prev}}(s)$.

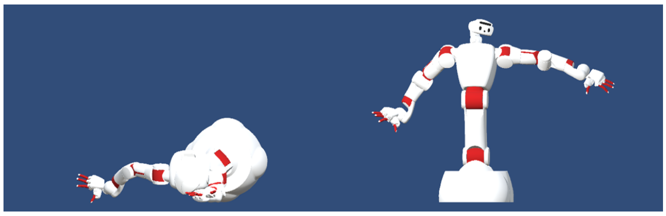

Fig. 4. The physics-enabled robot (left) falls due to real-world constraints, unlike the robot rig (right).

Besides, a contact condition was applied to this base controller to disable movement of the base if contact between the base and the floor was broken in case the robot tips over. The parameters for this controller were then hand-tuned which resulted in adequate performance of position tracking for the base movement of the physics-enabled model.

The combined body and base control allow the physics model of AIREC to follow the animation motion within the parameters of the real robot. This allows us to send the joint angles and base positions of the physics model as targets to the Robot Controller. Web-socket was used to achieve this, sending messages with all the joint positions.

2) Whole body motion validation: Various CG animation motions have been applied to both the robot rig and the physics-enabled rig. For low-dynamic motions, both tracked the original motion well. However, when high dynamic motion is applied, the physics-enabled robot falls down, as seen in Fig. 4. This occurs because the physics-enabled robot must contend with real-world physical constraints such as balance, inertia, and friction. These factors can cause it to lose stability and fall when subjected to rapid or complex movements, which are not accounted for in the input motion. In contrast, the robot rig on the right, which is not governed by real-world physics, can execute the movements without these stability issues. If human motions, particularly fast and dynamic ones, were transmitted directly to a robot, it would have difficulty keeping up due to the constraints of its actuators and control systems. In this work, we focus on lower dynamic motions so that this physics model can serve as a validation for the applicable motions suitable for direct real robot deployment.

D. One-shot deployment

After validating the application of the selected motion to the physics-enabled model, the joint angle data from the physics-enabled model in Unity is continuously transmitted to the real robot to control its movements. For the upper body, trajectory controller is used. For the robot's base, which only allows velocity control, the previously mentioned velocity controller is utilized to effectively manage the position-based commands received. This approach enables the robot to closely replicate the motions displayed in the Unity simulation with minimal delay, thereby facilitating real-time control of the AIREC robot and other similar mobile-based robots using human motion.

IV. Experiments and Results

A. Experiments

As discussed in Section I, it is essential for an assistive robot to exhibit natural, human-like motions to enhance its social acceptance, particularly in terms of fostering affection, intimacy, and cooperative willingness from users. We conducted the following two experiments to assess this methodology. The first experiment included a motion demonstration of the robot using one-shot, real-time robot control using a pre-recorded animation. In addition, we conducted a survey to score the motion generated in the first experiment, aiming to estimate the public's emotional response to such control methods for AIREC compared with a traditional robot motion control method.

The motion demonstration experiment was performed using a ready-to-download animation from Mixamo called: "Walk in Circle". This animation represents a good balance of animated upper body and character motion. The character appears relaxed, with his head looking around and arms swinging naturally. The upper body movements are characteristic of human motion as it was obtained from professional motion capture. Due to size constraints, the translation parameters of the movement were scaled to 0.5 of the original circle radius (from 1 meter to 0.5-meter radius) of the circle the robot walked in. Besides, the motion speed was scaled to 0.5 of the original animation speed to ensure the dynamic stability of the robot as we focus on low dynamic motion. The Odometry of the robot was reset each take of this experiment to prevent accumulating errors. Once the prerequisites were set as described above, the motion was performed as shown in Fig. 5.

For comparison, we controlled the robot to move in the same circle (radius of 0.5 meters) using a traditional approach focused solely on locomotion, without considering naturalness or expressiveness. This was achieved by simply sending position commands to the base to execute the motion. Afterwards, we conducted an online questionnaire on a 7-point scale (1 strongly disagree – 7 strongly agree) was asked by showing two types of videos; the motion of the proposed method and the motion of the traditional mobility. The questions were "How friendly the robot motion was," "How relieved the robot motion was," "How natural the robot motion was," "How trustworthy the robot motion was," and "How human-like the robot motion was". Subjective scores were obtained from 26 anonymous people with science and engineering backgrounds who are under 40 years old. Since the data were non-parametric, the Wilcoxon signed-rank sum test compared the scores to analyze whether the data differed significantly (the significance level was 0.05).

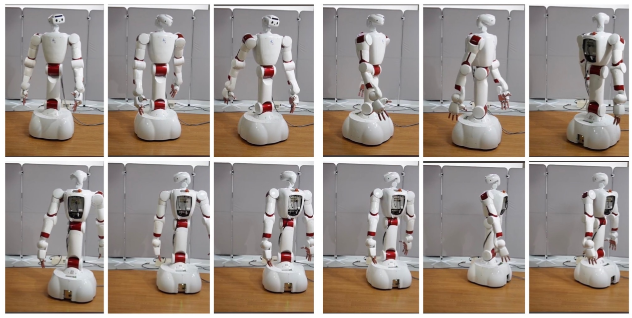

Fig. 5. Incremental snapshots of the robot performing the "walk in circle" animations using the proposed method. The snapshots show that the AIREC is expressing the positional change by using human body language such as arm swinging and torso and head movement

B. Results and discussion

The animation transfer experiment was able to follow the necessary motions of the target animation. As shown in Fig. 5, the motion produced by the robot was dynamic and resembled the "walk in circle" animation. In Fig. 6, we presented the left arm joint 1, torso joint 3 as well as head joint 3 as the representative data among all 23 joints since in the animation of human walking, the arm swing and waist twisting represent distinctive characteristics that markedly differ from those of traditional robotic movements. We can see that once the animation started, both the arm and waist joints closely followed the command references, albeit with some discrepancies.

First of all, there is minor but consistent undershooting of the command references by the real angles. This is explained by the control method used. Since we used target position control, there is a balance to be struck between undershooting of the animation target, and jitter of the final movement. The undershooting was not severe but persistent and would require a more sophisticated synchronization between the robot and the command program or a better control algorithm.

The second difference between the reference and command signal was the initial difference between the command and sensed angles. This is explained by the way the experiment was conducted. The initial state of the robot was slightly different from the state of the animation, resulting in the difference. Once the animation is initialized and sent to the robot after 20 seconds, the offset corrects itself, signaling a robust error feedback policy.

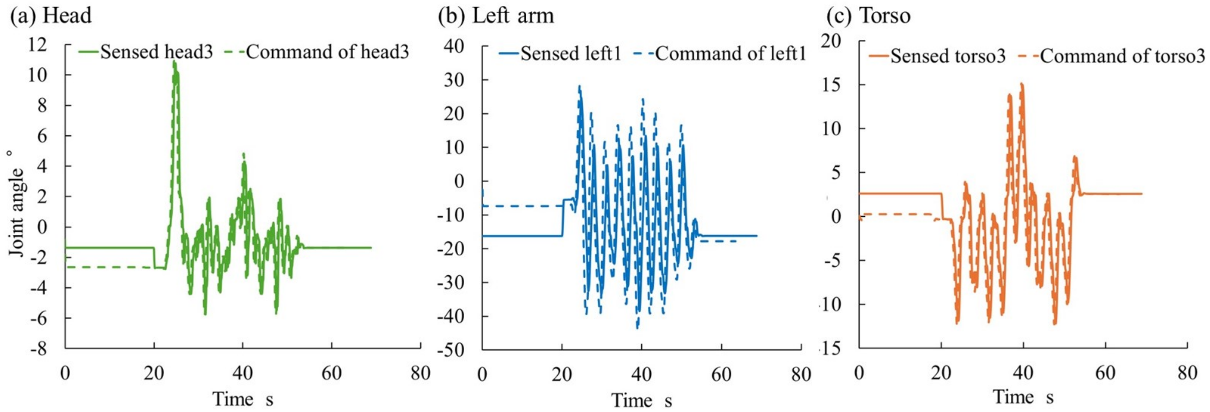

Fig. 6. A comparison between the command joint angles sent to AIREC and the recorded real joint angles of the left arm joint 1 and waist joint 3 throughout the motion experiment.

As shown in Fig. 7, tracking the position and rotation of the robot base, the rotation parameter is largely accurate, while the position variables show significant overshoot. This is explained by the hand-tuning which was done to the controller of the base. Since the variables are not mathematically optimised, some error is expected. Nonetheless, the overall shape of the sensed signals follows the command reference, once again, pointing to a promising approach.

Furthermore, the entire framework along with the motion was performed in real time using a one-shot approach. Within the Unity engine, there is a negligible delay between the humanoid animation and the output goals. The very slight delay seen in the data of Fig. 6 and Fig. 5 are related to the communication protocol with the robot and are almost imperceptible. It is, however, still possible to reduce it if a more robust data transfer method is used.

In general, performing the natural whole-body motion as planned using real-time one-shot control was safe and successful. The mobile-based robot AIREC acted as expected and all the discrepancies are explainable and are not inherent to the proposed framework.

The results collected from the subjective survey comparing motion generated by proposed framework to the traditional upper-body stationary mobility methods were insightful to the perception of robot body language by humans. Fig. 8 indicates the result of the investigation in subjective impressions about the robot motion video from 26 anonymous people. People have a significantly more positive impression, that is, more friendly, natural, and human-like motions, toward the motion of the proposed method. On the other hand, the survey found that the robot motion did not appear significantly more trustworthy or relieved.

The increased perception of friendliness, naturalness and human likeness suggests that incorporating more dynamic upper-body movements can make mobile-based robots appear more approachable and relatable to humans. This could be particularly beneficial in applications where social interaction is key, such as in customer service, healthcare, or educational settings, where a friendly and human-like demeanor can enhance user experience and engagement.

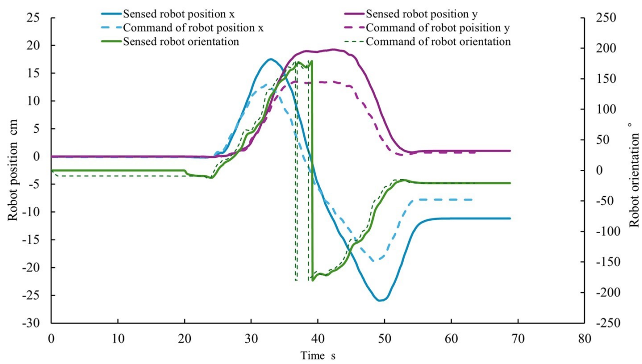

Fig. 7. A comparison between the commanded position and rotation of AIREC's base and the recorded actual position and rotation of the base throughout the motion experiment.

On the other hand, the discrepancy between the improved perception of the aforementioned traits and the stagnant perception of trustworthiness and relief might indicate that while motion can enhance superficial impressions, deeper trust-related factors may rely on other elements such as consistent performance, reliability, and clear communication of intent.

In general, these results suggest that the proposed framework for motion generation could offer an improvement over traditional methods that focus solely on functionality, such as mobility. In our experiments, the proposed approach enhances the naturalness of the robot's movements, leading to better human perception and interaction.

V. Conclusions and Future Work

In this work, we proposed a framework utilizes MAPT to address structural differences, employing a physics-enabled robot model to validate motions derived from a humanoid rig to which CG animations are applied. Upon confirming the safety and feasibility of these motions, they are subsequently deployed in real-time to the physical robot. This framework ensures the authenticity of whole-body movements and lays the foundation for a seamless transition from humanoid to full-body motion in robot-based systems.

Human walking animations were used to test the robot's capabilities, specifically its ability to replicate the natural

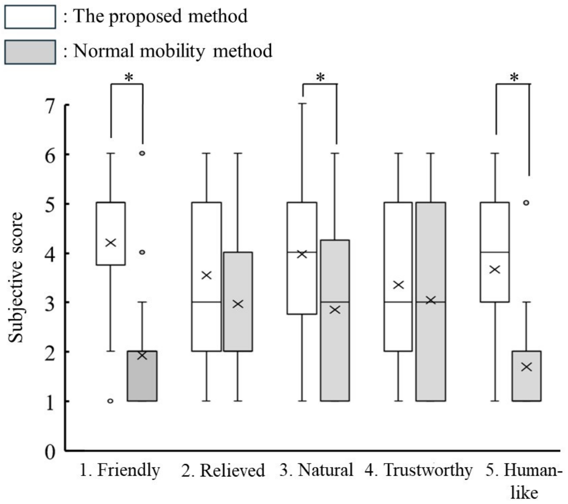

Fig. 8. Subjective score. 7-score means "strongly agree" and 1-score means "strongly disagree". * indicates a significant difference.

movements and how humans perceived these actions when comparing the performance of both the proposed and traditional methods. The findings demonstrated that the proposed method successfully mimics natural human motions, consistently achieving superior impressions in the natural appearance of the motion, the human-like factor, and friendliness. In the future, this research could lay the groundwork for developing a more robust framework from human motion to AIREC, utilizing reinforcement learning to facilitate high-dynamic motions. There are many improvements which could be made to the current method. Using velocity control for the upper body to avoid undershooting and achieve smoother movements would be a good next step. In addition, training the PID values for the base using machine learning approaches, instead of relying on manual tuning, to improve performance can be another necessary future upgrade. Next to the possible technical improvements, it is important to conduct more experiments using a variety of animations to clearly show the limits of how dynamic the movement has to be for this approach to fail. This will hopefully reveal what needs to be done to make this framework more robust.

Next to this, it is also an interesting perspective to keep the human reaction to the movements in mind. especially how human-like the generated movements are. If this questioning can be done on a larger scale, it could lead to insight about how to use animated robots to achieve more natural human-robot interaction.

References

[1] N. Maalouf *et al.*, "Robotics in nursing: a scoping review," *Journal of Nursing Scholarship*, vol. 50, no. 6, pp. 590–600, 2018.

[2] D. Feil-Seifer and M. J. Mataric, "Defining socially assistive robotics," in *9th IEEE International Conference on Rehabilitation Robotics*, 2005, pp. 465–468.

[3] M. J. Matarić and B. Scassellati, "Socially assistive robotics," in *Springer Handbook of Robotics*. Springer, 2016, pp. 1973–1994.

[4] E. C. Baek and E. Stolterman, "From robotphobia to robotophilia: The influence of social presence on robot acceptance," *International Journal of Social Robotics*, 2023.

[5] Z. Fu, Q. Zhao, Q. Wu, G. Wetzstein, and C. Finn, "Humanplus: Humanoid shadowing and imitation from humans," *arXiv preprint arXiv:2406.10454*, 2024, project website: this https URL. [Online]. Available: https://doi.org/10.48550/arXiv.2406.10454

[6] 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, submitted on 26 Feb 2024 (v1), last revised 6 Mar 2024 (this version, v2). [Online]. Available: https://doi.org/10.48550/arXiv.2402.16796

[7] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous Robots*, vol. 40, pp. 429–455, Mar 2016.

[8] H. Iwata and S. Sugano, "Design of human symbiotic robot twendy-one," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 580–586.

[9] Z. Fu, Q. Zhao, and C. Finn, "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," *arXiv preprint arXiv:2401.02117*, Jan 2024.

[10] T. Miyake, Y. Wang, G. Yan, and S. Sugano, "Skeleton recognition-based motion generation and user emotion evaluation with in-home rehabilitation assistive humanoid robot," in *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, 2022, pp. 616–621.

[11] K. Harada, K. Hauser, T. Bretl, and J.-C. Latombe, "Natural motion generation for humanoid robots," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 833–839.

[12] S. Katayama, M. Murooka, and Y. Tazaki, "Model predictive control of legged and humanoid robots: models and algorithms," *Advanced Robotics*, vol. 37, no. 5, pp. 298–315, 2023.

[13] N. Scianca, D. De Simone, L. Lanari, and G. Oriolo, "Mpc for humanoid gait generation: Stability and feasibility," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1171–1188, Aug 2020.

[14] T. Erez, K. Lowrey, Y. Tassa, V. Kumar, S. Kolev, and E. Todorov, "An integrated system for real-time model predictive control of humanoid robots," in *2013 IEEE-RAS International Conference on Humanoid Robots*. University of Washington, Feb 2015, pp. 292–299.

[15] K. Yamane and J. Hodgins, "Simultaneous tracking and balancing of humanoid robots for imitating human motion capture data," 11 2009, pp. 2510 – 2517.

[16] Z. Li, X. Cheng, X. B. Peng, P. Abbeel, S. Levine, G. Berseth, and K. Sreenath, "Reinforcement learning for robust parameterized locomotion control of bipedal robots," *arXiv preprint arXiv:2103.14295*, 2021, submitted on 26 Mar 2021. To appear in 2021 International Conference on Robotics and Automation (ICRA 2021). [Online]. Available: https://arxiv.org/abs/2103.14295

[17] Q. Nguyen, A. Agrawal, X. Da, W. Martin, H. Geyer, J. Grizzle, and K. Sreenath, "Dynamic walking on randomly-varying discrete terrain with one-step preview," in *Proceedings of Robotics: Science and Systems (RSS '17)*, July 2017.

[18] J. Siekmann, Y. Godse, A. Fern, and J. Hurst, "Sim-to-real learning of all common bipedal gaits via periodic reward composition," *arXiv preprint arXiv:2011.01387*, 2021, accepted for presentation at ICRA 2021. The first two authors contributed equally to this work. [Online]. Available: https://arxiv.org/abs/2011.01387

[19] Z. Xie, P. Clary, J. Dao, P. Morais, J. Hurst, and M. van de Panne, "Learning locomotion skills for cassie: Iterative design and sim-to-real," in *Proceedings of the Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, L. P. Kaelbling, D. Kragic, and K. Sugiura, Eds., vol. 100. PMLR, 30 Oct–01 Nov 2020, pp. 317–329. [Online]. Available: https://proceedings.mlr.press/v100/xie20a.html

[20] J. Siekmann, S. Valluri, J. Dao, L. Bermillo, H. Duan, A. Fern, and J. Hurst, "Learning memory-based control for human-scale bipedal locomotion," *arXiv preprint arXiv:2006.02402*, 2020, 8 pages, 5 figures, submitted to Robotics: Science and Systems 2020. [Online]. Available: https://arxiv.org/abs/2006.02402

[21] P. C. Yang, S. Funabashi, M. Al-Sada, and T. Ogata, "Generating humanoid robot motions based on a procedural animation ik rig method," in *2022 IEEE/SICE International Symposium on System Integration (SII)*, Narvik, Norway, 2022, pp. 491–498.