Human Factors and Systems Interaction, Vol. 192, 2025, 19–28
https://doi.org/10.54941/ahfe1006624

Marionette-Inspired Interface: Bridging Traditional Puppetry and Modern Avatar Control

Kazumi Inada[1] and Sangtae Kim[2]

[1]Doctoral Programs in Informatics, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Ibaraki, 305-0821, Japan
[2]University of Tsukuba, Ibaraki, 305-0821, Japan

ABSTRACT

With the advancement of virtual spaces, avatars and robots have become crucial for expressing human presence. However, current avatar control methods rely on expensive and complex technologies such as motion capture and image recognition, making implementation challenging. To address these issues, this study proposes an intuitive avatar control interface inspired by marionette manipulation techniques. Marionettes allow for expressive movements through simple operations. In this study, we applied these characteristics to a digital environment and developed a system enabling users to control avatars with one-handed operation. Experiments demonstrated that even novice users could effectively operate the system. Additionally, we explored methods to retain expressive richness while mitigating the learning difficulties associated with conventional marionette techniques. The findings suggest that a marionette-inspired interface is an effective method for precise and expressive avatar manipulation, bridging traditional and modern technologies and expanding interaction possibilities.

Keywords: Marionette control, Avatar manipulation, Virtual reality, Human-computer interaction, Motion capture alternatives, Digital puppetry

INTRODUCTION

Background

With the recent expansion of virtual spaces, avatars and robots have become increasingly important for representing human presence. Current avatar control methods primarily rely on motion capture and image recognition technologies, which suffer from high costs, complex setups, and environmental dependence, making them unsuitable for everyday use, especially by general users.

Moreover, while conventional avatar control excels at replicating simple actions, it lacks expressive flexibility and real-time adaptability. To address these limitations, this study proposes an interface inspired by traditional marionette manipulation techniques.

Research Objective

This study aims to overcome the limitations of traditional avatar control technologies—high costs, environmental constraints, and lack of operational flexibility—by developing a more intuitive and expressive interface.

Traditional marionette manipulation techniques are ingeniously designed to produce complex expressions through simple hand movements. By applying these characteristics to a digital environment, we propose a novel interface for avatar and robot manipulation.

Characteristics of Marionettes

Marionettes consist of a puppet and a controller, employing a unique mechanism that abstracts and simplifies human movements while maintaining a high level of expressiveness.

While marionettes exist in various cultural forms worldwide, most are designed to produce rich movement with minimal manipulation and can be intuitively controlled using hand movements alone (**Figure 1**).

Leveraging these characteristics, avatar and robot control can be made more intuitive and expressive. This research aims to apply marionette motion dynamics to a digital environment, developing an interface that enables sophisticated motion expression with one-handed operation.

Contributions of This Research

Unlike conventional motion capture and camera-based avatar control, this study uniquely utilizes marionette motion dynamics to enable expressive control through simple operations.

While previous methods have focused on precise motion reproduction, this study aims to facilitate intuitive and improvisational control.

Figure 1: A display of traditional marionettes from various cultures, from the collection of *Library of Puppetry*, exhibited at Suita City Museum, Japan. The collection showcases diverse string puppet designs and control mechanisms, highlighting the rich history and cultural variations in marionette performance.

The contributions of this research are as follows:

1. Proposal of a novel avatar control method leveraging marionette movement characteristics
2. Evaluation of usability and effectiveness compared to conventional methods
3. Exploration of potential applications for professional use.

RELATED WORK

Existing Avatar Control Technologies

Various techniques have been used for avatar control, including:

- Motion Capture (MoCap): MoCap is widely used in animation and gaming due to its high precision. However, its high cost and large space requirements make it impractical for everyday users.
- Camera-Based Gesture Recognition: This technique enables avatar control without special devices, but it is susceptible to lighting conditions and environmental factors, limiting detailed movement accuracy.
- Gamepad/Joystick Control: While commonly used in gaming and virtual environments, these controls rely on pre-programmed animations, making them unsuitable for expressive, improvisational performances.

Each of these conventional methods has advantages but struggles to balance expressive flexibility and intuitive usability. To overcome these challenges, this study explores an avatar control method inspired by traditional marionette techniques.

Digital Interfaces Utilizing Puppetry Techniques

Puppetry has evolved over centuries, providing a foundation for creating rich expressions with limited control mechanisms. Recent studies have explored applying these techniques to digital environments:

- **Marionette Motion Dynamics and Control:**
  - Chen et al. (2004) analyzed marionette structures and control mechanisms, demonstrating their applicability to robotics.
  - Yamane et al. (2003) introduced a method for controlling motor-driven marionettes using motion capture data, highlighting the potential for digital adaptation.
- **Application of Puppetry Techniques:**
  - Sakashita et al. (2017) developed a remote-controlled puppetry system, enhancing intuitive operability.
  - Yen et al. (2024) applied Taiwanese Budaixi puppet gestures to digital interfaces, enabling avatar control through hand movements.
- **Integration into VR and Robotics:**
  - Mazalek & Nitsche (2007) proposed Tangible User Interfaces (TUI) for manipulating virtual characters.

  - Hashimoto & Hirao (2024) introduced a marionette-inspired fingertip control system, combining real-time avatar manipulation with haptic feedback.

These studies suggest that various puppetry techniques, including marionettes, Guignol, and Budaixi, offer promising insights for avatar control interfaces. This study contributes to this field by exploring marionette-specific manipulative dynamics for expressive digital avatar control.

Positioning of This Research

Unlike conventional motion capture and camera-based techniques, this study aims to achieve an intuitive and expressive avatar interface leveraging marionette operation principles.

- **Comparison with Existing Methods:**
  - More affordable and expressive than MoCap or camera-based recognition
  - More intuitive than gamepads, with finer control than gesture recognition
  - Establishes a novel avatar control framework by applying marionette techniques to digital environments
- **Key Contributions:**
  - Introduction of a marionette-based avatar control method
  - Demonstration of potential applications in professional fields
  - Fusion of traditional puppetry techniques with digital technology for enhanced expression

This study does not merely extend existing techniques but proposes a fundamentally new approach to avatar control, potentially complementing conventional methods.

METHODOLOGY

Overview of the Proposed Method

By integrating marionette manipulation techniques with digital technology, we propose an interface that allows users to control avatars and robots using a marionette-like controller. This system enables one-handed operation, reducing the need for full-body motion tracking.

Unlike conventional motion capture, this method emphasizes skilled manipulation for expressive control rather than direct motion replication. It is particularly suitable for animation production and remote performance control.

Additionally, while marionette manipulation typically requires expertise, this study explores features that support beginners, potentially broadening its user base.

Advantages Over Conventional Methods

The proposed method offers the following advantages:

1. **Intuitive and Flexible Operation:**
    - Enables natural motion adjustments via a single hand
    - Allows real-time improvisation and nuanced gestures
2. **Minimal Environmental Constraints:**
    - Requires no cameras or motion capture suits
    - Operates effectively in small spaces
3. **Precision and Expressiveness:**
    - Provides fast response times with immediate feedback
    - Enables subtle emotional expressions (e.g., head tilts, shoulder shrugs)

This study explores how these advantages position the proposed method as a viable alternative to existing avatar control systems.

Potential Applications

The proposed method has potential applications in various fields, including:

1. **Stage Theater and Virtual Live Performances:** By incorporating traditional puppetry techniques into digital performances, this method can enhance artistic expression and broaden audience engagement. Additionally, it may encourage renewed interest in traditional performing arts.
2. **Remote Control and Customer Service:** The system enables natural gesture-based interactions using simple hardware. This makes it a viable option for remote-operated services, such as virtual assistants or unmanned store management, where expressive avatar control is beneficial.
3. **Games and Interactive Art:** By leveraging marionette-inspired movement, this interface introduces a novel interaction method for gaming and digital art. It also has the potential to improve conventional game controllers by offering a more dynamic and nuanced way to manipulate characters and objects.

EXPERIMENT

Purpose of Experiment

Through the creation of a simple prototype of the proposed interface and user experiments, we addressed the following:

- To establish a concrete method for integrating a string marionette controller with digital technology.
- To clarify issues necessary to improve usability.
- To study effective applications of the system.

Experimental Methodology

An initial prototype was implemented based on the proposed interface concept. The prototype is a 3DCG-rendered avatar head that can be manipulated with the proposed device. The control device mimics the shape of vertical controllers used in the Czech Republic and Germany and is equipped with an acceleration sensor to detect body tilting operations and a rotary encoder to measure the rotation of the handle operated by a finger (**Figure 2** and **Figure 3**). The outputs of the sensors were connected to a PC via Arduino, ensuring a stable frame rate of 60 fps or higher.

Figure 2: Prototype of the marionette-inspired controller, consisting of a handle, rotary encoder, and acceleration sensor.

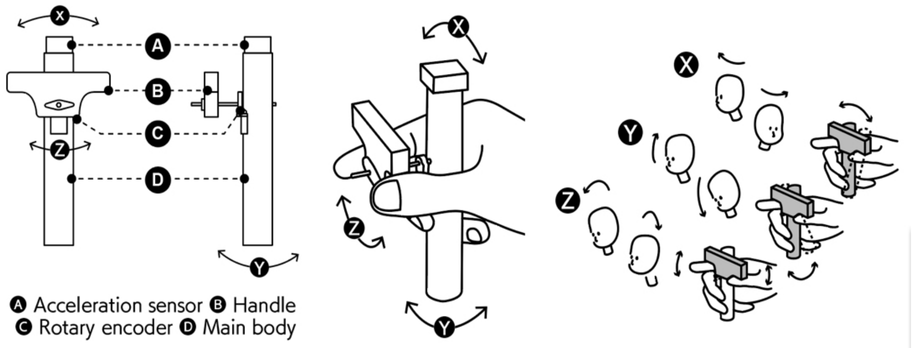

Figure 3: Diagram illustrating the structure and motion axes of the controller. The system captures rotational input from the handle and tilt input from the acceleration sensor to control the avatar's head orientation in 3D space.

User experiments were conducted using the prototype system (**Figure 4**). Participants were given the freedom to manipulate the avatar's head movements and respond to a one-minute scripted voice simulation of a call center interaction. As a reference for comparison, users also performed the same tasks using webcam-based motion capture and mouse-dragging control.

Evaluation

The experimenter observed the users' operations, the time required to learn and complete the tasks, and the process of skill acquisition.

After the experiment, a discussion was conducted, asking participants about what they found useful, what challenges they encountered, and the potential of the device. Users also rated the proposed method and comparison methods on a five-point scale across four categories: ease of use, intuitiveness, accuracy of movement, and overall satisfaction.

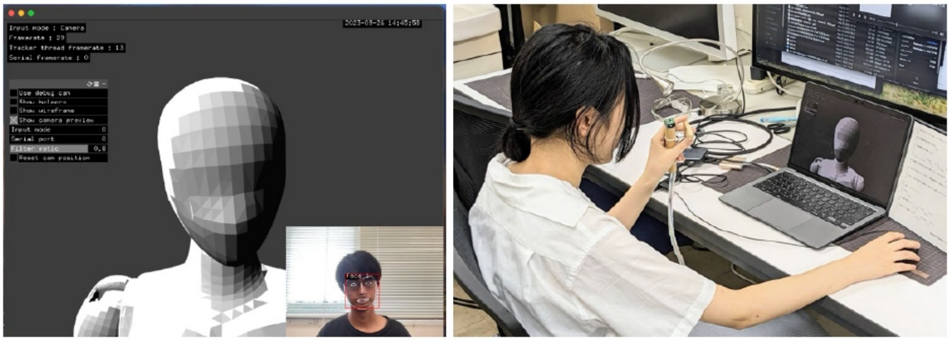

Figure 4: (Left) Interface displaying real-time avatar control. (Right) User operating the marionette-inspired interface prototype to manipulate an avatar using a physical controller.

RESULTS

Organization of Experimental Results

A total of six users in their 20s participated in the experiment. The time required to complete the tasks ranged from approximately 3 to 5 minutes for both the proposed and comparison methods. All participants were able to operate the system smoothly without significant difficulties.

In post-experiment interviews, participants noted that while camera-based control felt intuitive, it required large movements, making it unsuitable for prolonged use. Conversely, the proposed method was suggested to be advantageous in environments demanding prolonged operation and precise motion adjustments.

Analysis of Evaluation Data

In terms of average overall evaluation, the camera (4.2), controller (3.8), and mouse input (2.7) received the highest scores, in that order. Camera-based control was rated highest for intuitiveness, while the proposed method was evaluated favorably for operational stability and control accuracy.

Camera operation was rated highest for intuitiveness (AVG = 4.8, SD = 0.4). The particularly low standard deviation suggests a consistent perception among participants. However, while camera-based control was also rated highly for accuracy (AVG = 4.3, SD = 1.1), the larger standard deviation suggests variability in individual user experiences.

Mouse operation received the lowest intuitiveness rating (AVG = 2.8, SD = 1.6), indicating that while some participants found it manageable, others experienced significant discomfort.

The overall evaluation of the proposed method ranked second after the camera (AVG = 3.8, SD = 0.8). While its intuitiveness rating was lower than the camera's (AVG = 3.5, SD = 1.2), the relatively large standard deviation suggests variation in user experience based on proficiency levels.

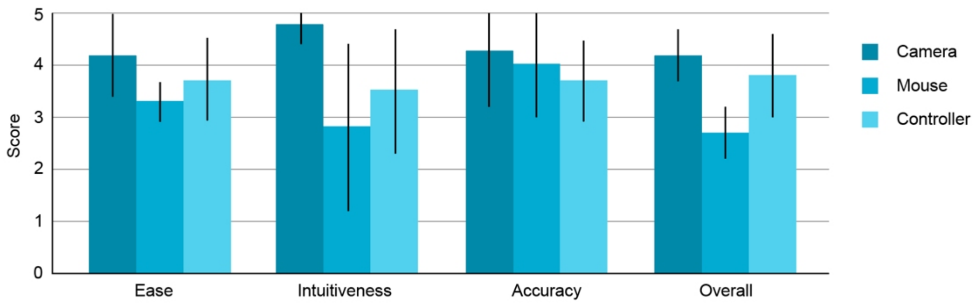

Figure 5: Mean scores and standard deviations of user evaluations for avatar control methods. The graph compares three methods—camera-based tracking, mouse input, and the proposed marionette-inspired controller—across four evaluation criteria: Ease of use, intuititiveness, accuracy, and overall satisfaction. Error bars indicate the standard deviation of user responses.

Issues and Future Prospects

Some users expressed concerns regarding the left/right movement correspondence (mirror image reversal) and the scale of movement amplification. These issues can likely be addressed by introducing an adaptive adjustment function tailored to individual users. Additionally, implementing guidance systems to assist beginners and enhancing real-time feedback will be crucial for improving usability.

Although the proposed method was rated lower than camera control in terms of intuitiveness, its stability in overall evaluation suggests that proficiency development can enhance user experience. Future improvements should focus on tutorials and feedback functions to aid skill acquisition.

Addressing these technical challenges is essential for future research. Enhancements in intuitiveness, learning support functions, and complementary technologies will be major steps toward practical implementation. These future research directions are discussed in the next chapter.

DISCUSSION

Contributions of the Study

The novelty of this research lies in its proposal of an avatar control method that leverages marionette manipulation characteristics to enable expressive, skill-based motion. While previous camera-based methods emphasize intuitive input, this approach provides precise control suitable for performance and animation.

This distinction suggests a complementary relationship with existing motion capture techniques, making the method promising for professional applications. Additionally, the proposed method offers advantages over traditional approaches by being less affected by environmental conditions and operable in limited spaces. As a result, it has potential applications in fields such as animation, stage performance, and remote services.

Unlike conventional interfaces that prioritize immediate ease of use, this study emphasizes depth of expression through proficiency. Webcams may be preferable for everyday video communication, whereas this method is more suited to applications requiring intricate control, such as animation fine-tuning and virtual performances. The study demonstrates that marionette-inspired control offers a distinctive alternative to camera-based methods, providing a practical solution for different use cases.

Future Issues

Experimental results indicate that user proficiency enhances operational accuracy. While the proposed method was rated lower in intuitiveness than camera control, it was favorably evaluated for stability and fine movement adjustments. This suggests its suitability for professional applications, and incorporating learning support features may make it more accessible to a broader audience.

Future research will address the following key challenges and improvements:

1. **Technical Enhancements:**
   - Improving action feedback (e.g., haptic feedback, latency optimization)
   - Implementing gesture-complementing technology (e.g., motion matching and ML-based motion synthesis)
   - Leveraging animation datasets to enhance expressiveness
2. **User Experience Improvements:**
   - Developing interactive tutorials and adaptive UI features
   - Integrating camera-based operation for hybrid control
   - Customizing movement settings based on user behavior

While the proposed method shows promise for professional applications requiring precision control, refining its intuitive usability and providing structured learning support will help expand its adoption.

CONCLUSION

Summary of Research

This study introduced a novel avatar control method based on marionette manipulation and experimentally evaluated its usability. Results confirmed that while camera-based control is more intuitive, the proposed method provides stable and precise control, making it suitable for specialized applications.

The research contribution lies in demonstrating an alternative avatar control technique that does not rely on motion capture or game controllers. The method's expressiveness is expected to improve with proficiency, making it a promising tool for professional applications.

Future Prospects

Future work should focus on enhancing tutorials and feedback mechanisms to improve intuitiveness. Additionally, integrating motion-matching technology can simplify operations and enhance expressiveness. Long-term development should position the proposed method as a professional tool, particularly in animation and virtual performance fields. The ultimate goal is to create a hybrid interface that balances intuitive operation with precise control.

ACKNOWLEDGMENT

This work was supported by JST SPRING, Grant Number JPMJSP2124.

REFERENCES

Chen, I. M., Tay, R., Xing, S., & Yeo, S. H. 2004. "Marionette: From Traditional Manipulation to Robotic Manipulation." In International Symposium on History of Machines and Mechanisms, edited by M. Ceccarelli. Dordrecht: Springer. doi: 10.1007/1-4020-2204-2_10.

Hashimoto, T., & Hirao, Y. 2024. "Selfrionette: A Fingertip Force-Input Controller for Continuous Full-Body Avatar Manipulation and Diverse Haptic Interactions." In Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24), Article 137, 1–14. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3654777.3676409.

Mazalek, A., & Nitsche, M. 2007. "Tangible Interfaces for Real-Time 3D Virtual Environments." In Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE '07), 155–162. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/1255047.1255080.

Sakashita, M., Minagawa, T., Koike, A., Suzuki, I., Kawahara, K., & Ochiai, Y. 2017. "You as a Puppet: Evaluation of Telepresence User Interface for Puppetry." In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), 217–228. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3126594.3126608.

Yamane, K., Hodgins, J. K., & Brown, H. B. 2003. "Controlling a Marionette with Human Motion Capture Data." 2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422), Taipei, Taiwan, 3834–3841 vol.3. doi: 10.1109/ROBOT.2003.1242185.

Yen, W. C., Liu, J. K., Peng, J. J., Luo, H. Y., Han, P. H., & Hsu, C. C. 2024. "Theater of Future Puppetry: An Immersive Puppeteering Experience Based on Hands Tracking and Gestures Recognition." In SIGGRAPH Asia 2024 XR (SA '24), Article 19, 1–2. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3681759.3688909.