

Puppeteer: Manipulating Human Avatar Actions with Intuitive Hand Gestures and Upper-Body Postures

Ching-Wen Hung
National Taiwan University
Taipei, Taiwan
amie12349@cmlab.csie.ntu.edu.tw

Ruei-Che Chang
University of Michigan
Ann Arbor, MI, USA
rueiche@umich.edu

Hong-Sheng Chen
National Taiwan University
Taipei, Taiwan
r09944076@ntu.edu.tw

Chung-Han Liang
National Taiwan University
Taipei, Taiwan
r09922a02@ntu.edu.tw

Liwei Chan
National Yang Ming Chiao
Tung University
Hsinchu, Taiwan
liweichan@cs.nycu.edu.tw

Bing-Yu Chen
National Taiwan University
Taipei, Taiwan
robin@ntu.edu.tw

ABSTRACT

We present Puppeteer, an input prototype system that allows players directly control their avatars through intuitive hand gestures and upper-body postures. We selected 17 avatar actions discovered in the pilot study and conducted a gesture elicitation study to invite 12 participants to design best representing hand gestures and upper-body postures for each action. Then we implemented a prototype system using the MediaPipe framework to detect keypoints and a self-trained model to recognize 17 hand gestures and 17 upper-body postures. Finally, three applications demonstrate the interactions enabled by Puppeteer.

CCS CONCEPTS

• Human-centered computing → Gestural input.

KEYWORDS

Body Posture, Hand Gesture, Camera system, User-Defined Gesture, Video Game, Input Techniques

ACM Reference Format:

Ching-Wen Hung, Ruei-Che Chang, Hong-Sheng Chen, Chung-Han Liang, Liwei Chan, and Bing-Yu Chen. 2022. Puppeteer: Manipulating Human Avatar Actions with Intuitive Hand Gestures and Upper-Body Postures. In *The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct)*, October 29–November 2, 2022, Bend, OR, USA. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3526114.3558689>

1 INTRODUCTION

Gamepad-controlled avatars are often used as the main avatar manipulation for gaming. However, the limited capabilities of gamepad-controlled manipulation lack intuitive control, which affects players' presence, enjoyment, and agency of avatar control in video

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UIST '22 Adjunct, October 29–November 2, 2022, Bend, OR, USA

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9321-8/22/10.

<https://doi.org/10.1145/3526114.3558689>

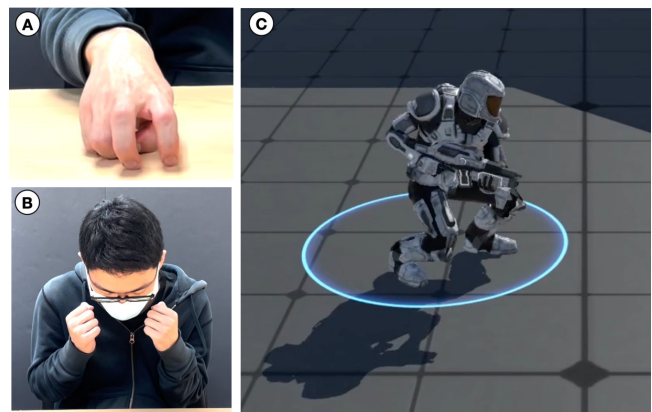


Figure 1: Puppeteer introduces the concept of combining hand gestures and upper-body postures to control avatars actions. Here is an example of the *crouch* action: (a) hand gesture (b) upper-body posture, and (c) avatar animation.

games [10]. Besides, the manipulation requires players to hold input devices during gaming, which restricts the freedom of hand movements to affect the game experience. Body-controlled avatars provide a more intuitive and free-hand manipulation that allows players to directly control their avatars in the virtual world through real-time body-to-body motion mapping, such as Kinect¹, Vicon², Optitrack³, etc. However, this type of manipulation is not appropriate to use in scenarios where players are in a narrow space or want to sit to play games because the manipulation needs more physical effort and interaction space [6, 7]. On the opposite, finger-controlled avatars provide dexterous and direct manipulation within a close-range space where players only use their hands to control avatar movement, including digital puppetry techniques [2, 4, 6, 8, 13] and iconic gestures [3, 9, 11, 12]. Although finger-controlled systems provide fewer sensory cues than body-controlled systems, they present less body fatigue and a more convenient method to explore the virtual environment [5]. The two intuitive manipulations have

¹<https://en.wikipedia.org/wiki/Kinect>

²<https://www.vicon.com/>

³<https://optitrack.com/>

their advantages and limitations, which motivates us to consider whether they have appropriate scenarios to represent. Could we combine the two techniques with their benefits as a new input technique?

Although previous works demonstrated either body-controlled or finger-controlled avatars [1, 2, 4, 11], none of them discussed tradeoffs and user preferences for body-controlled and hand-controlled manipulation that would significantly affect the perception of game experience. In response, to understand and compete the advantages of the two manipulations, we conducted a pilot study to know why, when, and how to use body postures or hand gestures for avatar manipulation in 20 games selected from three top-sell categories on STEAM⁴. Through the survey of the 20 games, we decided to focus on representing human avatars' motions, and we further interviewed participants about their preferences for hand-controlled and body-controlled avatars. According to the result of the pilot study, players use their bodies to represent avatar actions when the actions are easy to be mimicked by bodies. However, when players want to sit during gaming to reduce body fatigue, or the avatar actions are unrealistic and hard to be represented directly by bodies, they tend to use hands to control their avatars. We discovered that hand gestures provide an alternative to lower-body movement when players do not want to move exaggeratedly.

Based on the above results, we proposed Puppeteer, a novel input system that leverages hand gestures and upper-body postures as an intuitive manipulation to control avatar actions, which is shown in Figure 1. Puppeteer consists of a multi-camera system that can recognize the selected 17 upper-body postures and 17 hand gestures using our self-trained machine learning model, which achieves an average of 90% accuracy for upper-body postures and 91% for hand gestures detection. We performed the pilot study investigating users' preference between hand-based and body-based input. We then examined a gesture elicitation study to get gestures/postures users defined to manipulate avatar actions. Based on the defined gestures/postures, we collected data to create two datasets, implemented a prototype system for gesture recognition, and developed three game applications to demonstrate the Puppeteer system. Finally, we also discussed future applications of Puppeteer that goes beyond games and current limitations.

2 APPLICATION

Based on the studies' results, we created three game applications to demonstrate Puppeteer interaction, including three popular game genres – *Action-Adventure Game*, *RPG Game* and *Shooter Game*, as shown in Figure 2. During gaming, players can use their hands or upper body to control the avatar directly and switch their inputs based on their preference at any time. Such input manipulation provides more freedom for players to decide when and how to control their characters for gaming. Each application contains the actions appropriate to represent by hand gestures and upper-body postures. Players will frequently switch between gestures and postures during gaming and experience the combination of the two manipulations.

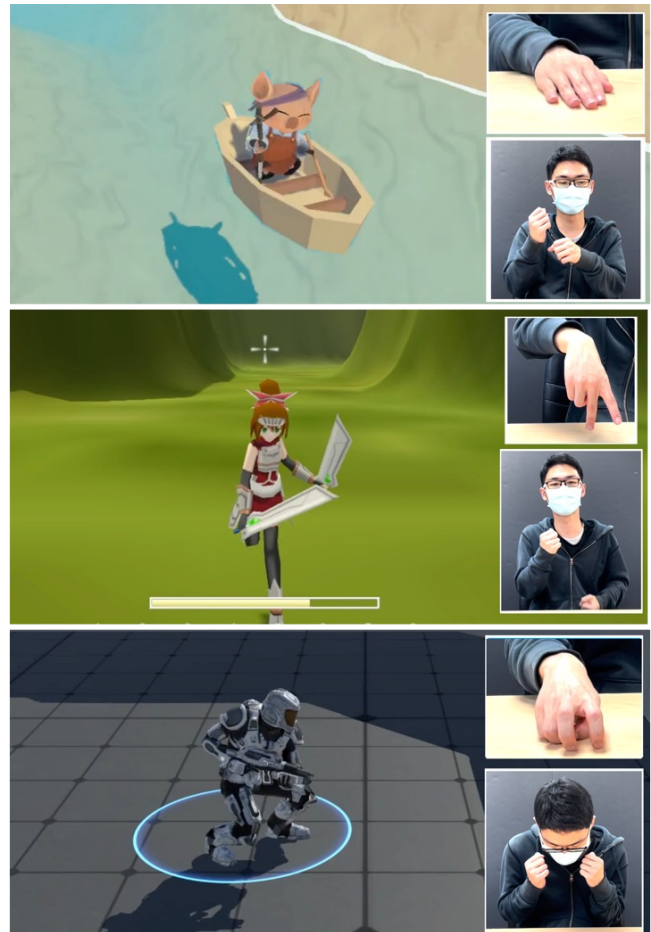


Figure 2: Three applications demonstrate usage of Puppeteer: Action-Adventure Game (upper), RPG Game (middle), and Shooter Game (bottom).

3 CONCLUSION

We present Puppeteer, a concept that combines hand gestures and upper-body postures to provide a new game input interaction by multiple cameras detection. We built a prototype using two cameras mounted on a screen and a tripod separately. Three demonstration applications enabled by Puppeteer allows participants to switch to input hand gestures and upper-body postures to manipulate their virtual avatars. We believe Puppeteer provides a new avatar manipulation for convenient and easy interaction of hands and upper bodies in video games.

ACKNOWLEDGMENTS

This research was supported in part by the Ministry of Science and Technology of Taiwan (MOST111-2221-E-002-145-MY3, 111-2218-E-002-028, 110-2634-F-002-051, 111-3111-E-002-002), and National Taiwan University.

⁴<https://store.steampowered.com/>

REFERENCES

- [1] Karan Ahuja, Eyal Ofek, Mar Gonzalez-Franco, Christian Holz, and Andrew D. Wilson. 2021. CoolMoves: User Motion Accentuation in Virtual Reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 52 (jun 2021), 23 pages. <https://doi.org/10.1145/3463499>
- [2] Yu-Ting Cheng, Timothy K Shih, and Chih-Yang Lin. 2017. Create a puppet play and interactive digital models with leap Motion. In *2017 10th International Conference on Ubi-media Computing and Workshops (Ubi-Media)*. IEEE, 1–6.
- [3] Christian Holz and Andrew Wilson. 2011. Data miming: inferring spatial object descriptions from human gesture. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 811–820.
- [4] An-Pin Huang, Fay Huang, and Jing-Siang Jhu. 2018. Unreal Interactive Puppet Game Development Using Leap Motion. In *Journal of Physics: Conference Series*, Vol. 1004. IOP Publishing, 012025.
- [5] Ji-Sun Kim, Denis Gračanin, Krešimir Matković, and Francis Quek. 2008. Finger walking in place (FWIP): A traveling technique in virtual environments. In *International Symposium on Smart Graphics*. Springer, 58–69.
- [6] Luis LEite and Veronica Orvalho. 2017. Mani-Pull-Action: Hand-Based Digital Puppetry. *Proc. ACM Hum.-Comput. Interact.* 1, EICS, Article 2 (jun 2017), 16 pages. <https://doi.org/10.1145/3095804>
- [7] Noah Lockwood and Karan Singh. 2012. Fingerwalking: motion editing with contact-based hand performance. In *Proceedings of the 11th ACM SIGGRAPH/Eurographics conference on Computer Animation*. 43–52.
- [8] Masaki Oshita, Yuta Senju, and Syun Morishige. 2013. Character Motion Control Interface with Hand Manipulation Inspired by Puppet Mechanism. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (Hong Kong, Hong Kong) (VRCAI '13)*. Association for Computing Machinery, New York, NY, USA, 131–138. <https://doi.org/10.1145/2534329.2534360>
- [9] Eunbi Seol and Gerard J Kim. 2019. Handytool: Object manipulation through metaphorical hand/fingers-to-tool mapping. In *International Conference on Human-Computer Interaction*. Springer, 432–439.
- [10] Paul Skalski, Ron Tamborini, Ashleigh Shelton, Michael Buncher, and Pete Lindmark. 2011. Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society* 13, 2 (2011), 224–242.
- [11] Christian Steins, Sean Gustafson, Christian Holz, and Patrick Baudisch. 2013. Imaginary Devices: Gesture-Based Interaction Mimicking Traditional Input Devices. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services (Munich, Germany) (MobileHCI '13)*. Association for Computing Machinery, New York, NY, USA, 123–126. <https://doi.org/10.1145/2493190.2493208>
- [12] Amato Tsuji, Keita Ushida, and Qiu Chen. 2018. Real Time Animation of 3D Models with Finger Plays and Hand Shadow. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. 441–444.
- [13] Bo-Xiang Wang, Yu-Wei Wang, Yen-Kai Chen, Chun-Miao Tseng, Min-Chien Hsu, Cheng An Hsieh, Hsin-Ying Lee, and Mike Y. Chen. 2020. *Miniature Haptics: Experiencing Haptic Feedback through Hand-Based and Embodied Avatars*. Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3313831.3376292>