Poppel: Manipulating a mid-air CG character using a puppet and human body

Miyu Fukuoka
The University of Electro-Commnications
Tokyo, Japan
fukuoka@media.lab.uec.ac.jp

Naoya Koizumi
The University of Electro-Commnications
Tokyo, Japan
koizumi.naoya@uec.ac.jp

***Abstract*—Mid-air imaging technology enables CG characters to pop out of the real world and share the same physical space with viewers. While methods exist for real-time CG character control, none allow full-body control of a mid-air CG character. Existing techniques that capture the operator's movements require significant space, and controlling via the puppet allows the operator to control the character in a smaller space; however, simultaneously controlling multiple body parts is challenging. We propose an integrated manipulation system that combines puppet motion capture and human body movement. It allows manipulation of the character's arm, neck, and leg movements associated with spatial movement. Arm and spatial movements are controlled via the puppet, whereas the neck movement is synchronized with that of the operator's neck. We confirmed that this system enables the operator to express complex and diverse movements combining neck, arm, and spatial movements, while directly controlling 3D spatial movement.**

***Index Terms*—Mid-air image, Puppet, Avatar controller, Motion capture, Interactive talk show type attraction, Multi-body part coordination**

I. Introduction

Puppets and CG characters are used in various environments, such as video content, communication [1] and educational tools [2], and theme park performances, and their emotions are expressed through body movements and facial expressions. Controlling these characters to display charming motions is crucial.

By bringing CG characters into real space, viewers can reach out and interact with the characters. Mid-air imaging technology allows CG characters to seamlessly connect to real space, resulting in the characters popping out, all while sharing the same physical space with the viewer.

This study aims to design and demonstrate a system for easily controlling the various body movements of a mid-air CG character, including the popping-out effect that is characteristic of mid-air imaging. The conventional method of capturing the operator's own body movement to control the body movements of a CG character in real time has limitations in terms of fatigue and space requirements. Therefore, we focused on capturing the motion of a puppet and controlling a CG character via the puppet's body movements. However, the operator can hold and control only a limited number of puppet body parts simultaneously.

Therefore, we propose an integrated manipulation system that combines puppet motion capture and human body movement. Specifically, simultaneous manipulation of the movements of both arms and neck and free spatial movements is performed by combining the manipulations of the puppet's palm and body positions with the operator's neck movement; we attempted to achieve a high degree of freedom of body movement. To assess the effectiveness of the proposed method, we analyzed the movements represented by the proposed method and evaluated the accuracy of the direct manipulation of 3D spatial movements.

II. System design

A. Design Requirements

To achieve the purpose of this study, which is a system for easily controlling the various body movements of a mid-air CG character, including the popping-out effect that is characteristic of mid-air imaging, we considered a design that satisfies the following three requirements.

1) The shape of the controller must be similar to that of the human body.
2) Must be capable of improvising movements.
3) Direct manipulation of 3D spatial movement must be possible.

Requirement 1 is necessary for allowing an amateur to control the character intuitively. According to Inami et al. [3], robots are computers with a physical body, and they proposed the Robotic User Interface (RUI) that uses enactive and iconic representations based on the body image that humans acquire during early growth processes. Based on this idea, for easy manipulation, we set Requirement 1. In addition, Requirement 2 is necessary for the mid-air CG character to flexibly express various movements in response to the communication between the viewers and the character. Requirement 3 allows the operator to move the mid-air CG character to the appropriate position according to the viewer's actions for controlling 3D spatial movement, which includes depth movement, the characteristic of mid-air imaging.

B. Manipulation of a mid-air CG character

The proposed system enables arm, neck, and leg movements associated with the spatial movement of a mid-air CG character. The arm and neck movements were selected as important movement elements with reference to the manipulation manual of the puppet show. The manipulation design was crafted to enable these movements to be executed simultaneously with spatial movement operations (Requirement 3). Requirement 2 requires the ability to express a variety of movements, and to allow a single person to perform an expressive character, each movement element is generated differently.

Arm movement is an important element within gestures known as non-verbal communication [4]. A significant number of the arm movements are associated with the palm positions, and controlling this position allows a high degree of freedom for expressing arm movement. Therefore, it is necessary to capture the palm position of the puppet.

Neck movement is important for expressing emotions [5]. The proposed system uses the operator's own neck movement to control the character's neck movement because using the puppet's neck makes it difficult to simultaneously control both the arms of the puppet while controlling its neck.

A mid-air image has the advantage of being able to pop out in the depth direction, allowing characters to approach in real space in addition to moving two-dimensionally (left, right, up, and down). Spatial movement enables walking and jumping movements that require leg movement associated with spatial movement. Based on the idea that movements of the upper body, such as arms and neck, are conscious movements, whereas those of the lower body are automatic, we determined that playing back the leg movement associated with spatial movement is sufficient.

III. Implementation

Fig. 1 illustrates the overall system structure[1]. The movements measured by the motion capture system are mapped to those of the CG character, and a mid-air CG character is displayed using an optical system. To realize communication between the mid-air CG character and the viewer, the operator controls the character while observing the viewer through the monitoring system.

A. Motion capture system and Integration of CG character's motion

The movements corresponding to the arm, spatial, and neck movements of a mid-air CG character are obtained using an optical motion capture system. The body movement data obtained by the motion capture system are mapped to the CG character movements.

The arm movement of the character is generated by inverse kinematics (IK) control based on the position and orientation data of the palms obtained through the motion capture system. To manipulate both arms, retroreflective markers were fixed to the puppet's palms, as illustrated in Fig. 2. The operator uses sticks, similar to using a rod puppet.

The neck movement of the CG character was generated by synchronizing the orientation data of the operator's head, who wears a catsuit with retroreflective markers fixed to their head, with the orientation parameter of the CG character's neck. To make neck movement more natural, the hip was rotated with the neck.

[1]Video: https://youtu.be/9sAyeChsxJM

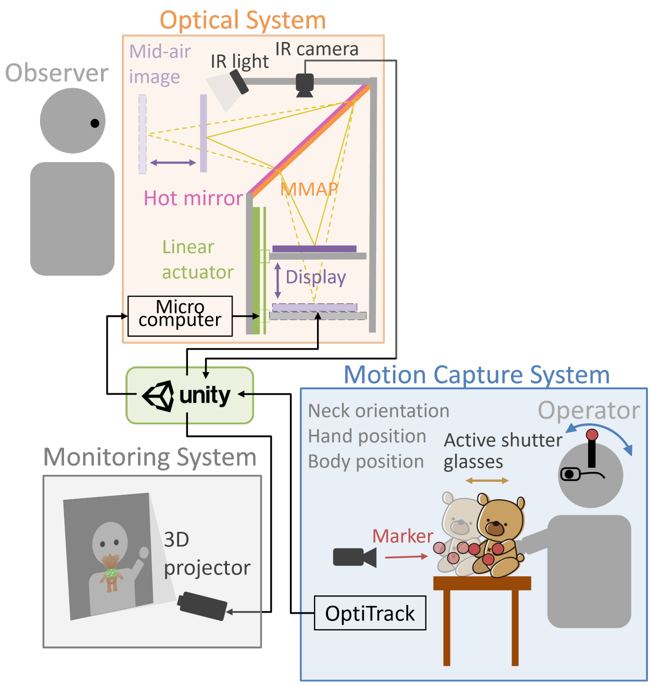

Fig. 1. Overall system structure. The movements measured by the motion capture system are mapped to those of the CG character, and a mid-air CG character is displayed using an optical system. The monitoring system allows the operator to observe the viewer and the mid-air CG character.

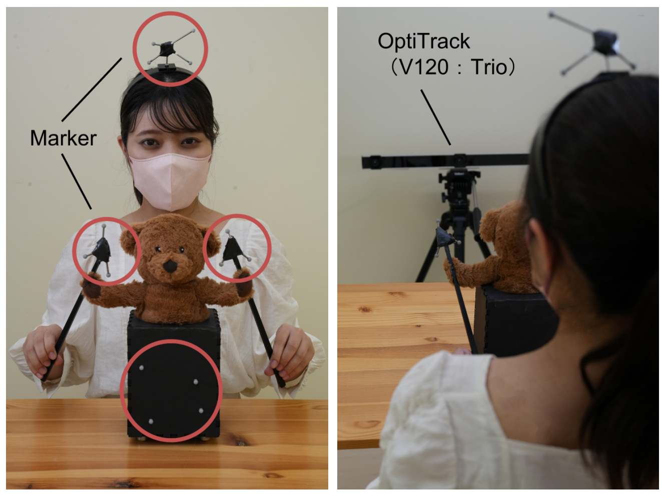

Fig. 2. Implementation of the motion capture system. An optical motion capture system, OptiTrack V120: Trio, was used to obtain the puppet's palm position, body position, and neck orientation in real-time.

The leg movement associated with spatial movement is realized by playing back a walking motion when the puppet is moved back and forth and left and right, and by playing back a jumping motion when the puppet is lifted upward. To acquire data regarding the puppet's body position, a box was attached to the lower body of the puppet and markers were attached to the front of the box.

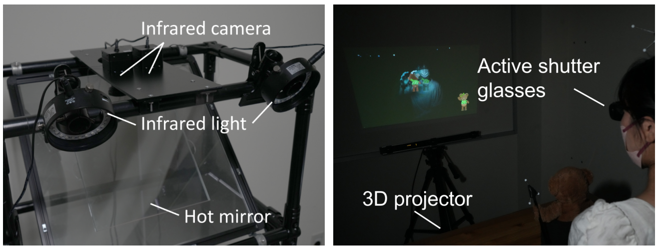

Fig. 3. Implementation of the monitoring system. Stereo cameras are placed at 64-mm intervals in the Floatagent method [7], and infrared lights are positioned to illuminate the viewer's face well (left). The operator wearing the active shutter glasses communicates with the viewer while watching the stereoscopic images displayed by the 3D projector (right).

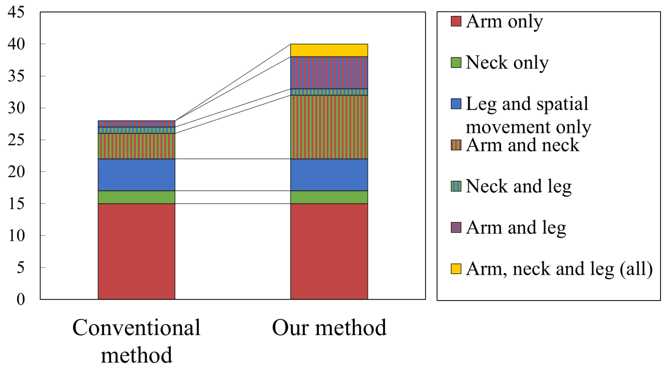

Fig. 4. Comparison of the breakdown of feasible body movements between the traditional puppet method and the proposed method.

B. Optical system

The optical system is shown in Fig. 1 wherein the optical design used for generating the mid-air CG character is illustrated. For the optical system used to form a mid-air image that can move in the depth direction, we adopt the MARIO method [6]. Our system comprises a display, MMAP, light shield, linear actuator, and microcomputer. The light emitted from the display is reflected by the MMAP, which is tilted at a 45-degree angle, and forms a mid-air image at a position that is plane-symmetrical to the MMAP. The microcomputer is used to communicate with the linear actuator. The display fixed to the linear actuator is moved vertically up and down, which shifts the position of the mid-air image in the depth direction. The mid-air CG character's vertical and horizontal movements are generated by moving the character's 2D position on the display.

C. Monitoring system

The monitoring system allows the operator to observe the viewer and enables communication between the mid-air CG character and the viewer. Fig. 3 shows the implementation of the monitoring system comprising an infrared camera, infrared light, hot mirror, projector, and active shutter glasses. For observing the viewer from the viewpoint behind the mid-air image, we adopted the Floagent method proposed by Ando et al. [7]. To confirm the physical movement of the mid-air CG character, a CG character is superimposed onto an infrared camera image showing the viewer. To determine the depth positions of the viewer and mid-air CG character, a stereoscopic image with binocular disparity is presented. As shown in Fig. 3, the infrared cameras for the left and right eyes are placed at 64-mm intervals, which is the average distance between the eyes. Additionally, two CG cameras were placed in a similar manner. A projector capable of projecting active shutter-type images is used to display stereoscopic images.

IV. Evaluation

A. Character's movement

In addition to confirming whether the system satisfied Requirement 2 : ***Must be capable of improvising movements***, we also investigated the system's limitations. Participants controlled the character while watching the viewer on a monitor, and designed the physical movements they wanted to express, regardless of whether they were feasible or not with the system. The participants included five students, with a mean age of 22.8 years, four of whom were male and one of whom was female.

We observed 40 types of movements that can be realized using this system (65.6 %), 4 that can be realized but are difficult to control (6.6 %), and 17 that cannot be realized (27.9 %). The Fig. 4 demonstrates that the proposed method is more versatile compared to the method where the puppet's body shape synchronizes with the CG character in a one-to-one correspondence (referred to as "conventional puppet-mediated manipulation method") [8], [9] in terms of feasible body movements. Fig. 4 shows that 45 % (18/40) of the feasible body movement expressions included body movements combined with either all parts or any of the arms, neck, and spatial movement. The proposed approach, which allows simultaneous manipulation of multiple parts, has demonstrated its enhanced versatility compared to conventional puppet-mediated manipulation methods.

B. Spatial movement

We experimented to confirm whether the proposed system enables direct manipulation of the spatial movement of a mid-air CG character, which is Requirement 3 of this system. The experiment aimed to evaluate whether the mid-air CG character could be accurately placed in the position intended by the operator. The participants moved the mid-air CG character while watching the monitor image to ensure that it was at the same depth as the landmark in real space. The markers are positioned at depth positions ranging from 5 cm to 30 cm, in 5 cm intervals, relative to the minimum projection distance of the mid-air imaging system as the origin. Additionally, markers are placed at 0 cm and 20 cm depth positions, each located 10 cm from the center in the left and right directions. Only one landmark was presented per trial, and the positions of the landmarks were presented in a random

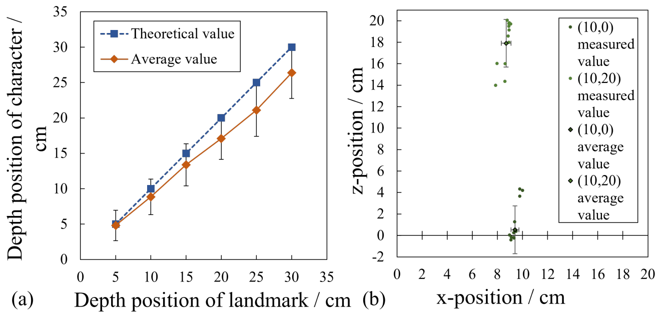

Fig. 5. (a) Target and character placement positions in the depth direction. (b) Placement position during spatial movement, including lateral movement.

order. Three trials were conducted for each placement, for a total of 24 trials. The participants included twelve students, with a mean age of 24 years, eleven of whom were male and one of whom was female.

Fig. 5 confirms that spatial movement, including that in the left and right directions, is possible with an average absolute error rate of approximately 10–15 %. Therefore, we confirmed that the system can be used to directly control 3D spatial movement within the range of interaction assumed in this system without any problems. We assume that the interaction scenarios between the viewer and the mid-air CG character include the character riding on the viewer's palm and following the movements of the viewer's hand. Considering the average length of a human palm, the error obtained in Section IV-B is less than 40 % of the palm length, indicating that it should not be a significant problem in the assumed interaction scenarios.

V. Discussion

This section discusses the validity of the manipulation of the body movement elements designed in Section II-B.

By controlling the arm, including the palm orientation, different body movements can be expressed using the same arm shape. For example, despite the similar arm shape used for expressing "cover eyes" and "place hands under the eyes as if crying," the two actions can be distinguished by adjusting the palm orientation. A total of 32 unique arm expressions, including various combined movements, were designed, demonstrating that controlling only the position and orientation of the palm is sufficient to realize a wide range of expressions.

The fact that all three axes of rotation (pitch, roll, and yaw) were used to express neck movement supports the assumption that the neck has three degrees of freedom.

Spatial movement is considered an important element because it was designed with the characteristics of mid-air imaging, the forward pop-out movement, as well as lateral movements, jumping, and levitating. In particular, the levitation movement is beyond human physical capabilities and is enabled by manipulation via a puppet. However, there were demands for incorporating leg movements such as "kick," "sit," "skip," "crouch," "stomp foot," and "cross legs." For this, the system must extend the leg movement associated with spatial movement, because "sit" and "crouch" can be interpreted as unconscious leg movements associated with the downward movement of the body.

VI. Conclusion

We proposed an integrated manipulation system that combines puppet motion capture and human body movement to easily control the various body movements, including the popping-out movement, which is characteristic of mid-air imaging. The arms and spatial movement of the mid-air CG character are controlled via the physical puppet, and the neck movement is synchronized with that of the operator's own neck. In addition, we investigated the physical movements of the mid-air CG character that can be represented using this system and evaluated the manipulation of spatial movement, including that in the depth direction. The investigation of the body movements of the mid-air CG character involved designing various body movements, including individual and complex body movements combining neck, arm, and spatial movements. Through evaluations of spatial movement manipulation, we confirmed that the mid-air CG character could be moved to the position intended by the operator with an average absolute error rate of approximately 10–15 % in the left/right and depth directions.

Acknowledgment

This research was supported by a JSPS Grant-in-Aid JP21K19821 and JP22H01079, and JST FOREST Program, Grant Number JPMJFR216L.

References

[1] C. Schaefer and L. Carey, *Family Play Therapy*, ser. Child therapy series. J. Aronson, 1994. [Online]. Available: https://books.google.co.jp/books?id=kElcV84iH-4C

[2] M. Jintapitak, "Use of animation characters to motivate students in a higher education class," in *2018 International Conference on Digital Arts, Media and Technology (ICDAMT)*, 2018, pp. 280–284.

[3] D. Sekiguchi, M. Inami, and S. Tachi, "RobotPHONE: Rui for interpersonal communication," in *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 277–278.

[4] N. Dael, M. Goudbeek, and K. R. Scherer, "Perceived gesture dynamics in nonverbal expression of emotion," *Perception*, vol. 42, no. 6, pp. 642–657, 2013, pMID: 24422246.

[5] A. Samanta and T. Guha, "On the role of head motion in affective expression," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 2886–2890.

[6] H. Kim, I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura, "MARIO: Mid-air augmented reality interaction with objects," *Entertainment Computing*, vol. 5, no. 4, pp. 233–241, 2014.

[7] S. Ando and N. Koizumi, "Floagent: Interaction with mid-air image via hidden sensors," in *SIGGRAPH Asia 2022 Emerging Technologies*, ser. SA '22. New York, NY, USA: Association for Computing Machinery, 2022.

[8] W. Yoshizaki, Y. Sugiura, A. C. Chiou, S. Hashimoto, M. Inami, T. Igarashi, Y. Akazawa, K. Kawachi, S. Kagami, and M. Mochimaru, "An actuated physical puppet as an input device for controlling a digital manikin," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 637–646.

[9] N. Shimizu, N. Koizumi, M. Sugimoto, H. Nii, D. Sekiguchi, and M. Inami, "A teddy-bear-based robotic user interface," *Comput. Entertain.*, vol. 4, no. 3, p. 8–es, jul 2006.