Evaluation of User's Psychological Sense in Tele-Immersion Robot Avatar

Yuki Kida[1(✉)], Shujie Li[1], Yuto Ueda[2], Naoto Takayanagi[2], Kei Matsuoka[3], and Tetsuro Ogi[1]

[1] Graduate School of System Design and Management, Keio University, Yokohama, Kanagawa, Japan
{ykida,lishujie}@keio.jp, ogi@sdm.keio.ac.jp

[2] Graduate School of Science and Technology, Keio University, Yokohama, Kanagawa, Japan
{ikimono-3,naoto20221106}@keio.jp

[3] Department of Multidisciplinary Design Science, Shonan Institute of Technology, Fujisawa, Kanagawa, Japan
matsuoka@md.shonan-it.ac.jp

Abstract. Robot avatars are attracting attention as a means of remote communication. In this study, a tele-immersion system was developed in which the user's own body transitions into the body of a robot avatar, enabling communication with a remote person as if they were in the same place. Its use of a VR headset, 360-degree camera, and motion tracking is in contrast to systems where a robot avatar is operated using a PC. A prototype of this system was developed and tested through evaluation experiments. As a result, users of the remote avatar felt as if they were in the same room with and collaborating with the other person, confirming the system's effectiveness as a tele-immersion system. In addition, users generally felt a sense of body ownership and a sense of agency towards the robot avatar, suggesting the generation of a sensation that their own body is transferred to the robot avatar's body.

1 Introduction

One approach to communicating with people in remote locations is to use robot avatars. Robot avatars are a technology for remote communication that utilize robots with physical entities as their own alter egos (avatars). For remote communication using avatars, methods using CG avatars in the metaverse have attracted attention and been actively researched in recent years [1–3]. However, robot avatars can interact with people and objects in real space because they have physical entities and are expected to be used in a variety of situations and applications.

For example, one currently commercialized robot avatar service is "newme" by avatarin Inc. [4]. This robot avatar is a combination of a display and a mobile robot and is capable of conversation and remote-controlled movement while projecting the user's face on the display. Users can visit and communicate in various remote locations via this robot avatar. Other research is aimed at supporting remote employment of disabled people by utilizing robot avatars to remotely transport goods and serve customers. With

L. Barolli (Ed.): NBiS 2023, LNDECT 183, pp. 325–333, 2023.
https://doi.org/10.1007/978-3-031-40978-3_35

this robot avatar, users operate the robot "OriHime-D" with a mouse or gaze input depending on their own disabilities, and actual remote work can be performed [5]. As described above, robot avatars are expected to be a means of changing the way people move and work in the future, and various studies and services are being conducted.

On the other hand, a common feature of current robot avatars is that the user often communicates by remotely operating the robot avatar on an information terminal such as a PC. Therefore, it is difficult to say that there has been sufficient discussion about the sense that one's own body is transferred to the body of the robot avatar in the process of communication.

Therefore, the purpose of this study is to realize a tele-immersion system that allows users to communicate with remote people as if they were in the same place, with the sensation that one's own body is transferred into the body of the robot avatar. In this study, a prototype was developed, and evaluation experiments were conducted. In the following sections, the concept and system design of a tele-immersion system using robotic avatars, as well as evaluation experiments using the developed prototype are discussed.

2 Tele-Immersion System

2.1 Concept and System Design

The concept of the system proposed in this research means that instead of remotely operating the robot avatar on a PC, the user can see the information captured by the robot avatar's eyes and express one's own will to move, as if one's own body were transferred to that robot avatar's body. Therefore, in this system, it is necessary to have the sensation that one's own body is transferred to the body of a robot avatar. In other words, it is necessary to generate a "sense of body ownership", in which one feels that the body of the robot avatar is one's own body, and a "sense of agency", in which one feels that one is moving the robot avatar's body [6, 7]. In addition, free viewing angles and speech were also required, and the following functions were necessary to realize this system.

- A function that allows the robot avatar's body to move freely as if it were the user's own body: motion function.
- A function that allows 360-degree free immersive viewing in the direction one wants to view: viewing function.
- A function that allows one to speak freely about what one wants to speak about: speech function.

The design of the components of this system to realize these functions will be described. First, for the motion function, we designed the system to be realized using Pepper [8] (Fig. 1(a)) and Azure Kinect (Fig. 1(b)). Pepper is a 121 cm-tall humanoid communication robot with actuators on its head, shoulders, elbows, wrists, hips, and knees, allowing it to perform a variety of human-like behaviors and gestures. There have been several studies on human interaction with Pepper [9–11], which led to this study choosing Pepper's use as an avatar. Azure Kinect can be used for body tracking and can acquire three-dimensional coordinate information of the user's joints. Based on the three-dimensional coordinate information of the user's joints, we designed to link the user's own motions with Pepper's motions.

Next, for the viewing function, we designed an immersive 360-degree free viewing angle for the user using a THETA V 360-degree camera (Fig. 1 (c)), and an Oculus Quest 2 head-mounted display (HMD) (Fig. 1 (d)).

For the speech function, since the specification of Pepper is to output speech from text data, and it cannot directly output the user's voice, we designed Pepper to speak freely by acquiring the user's voice with a microphone, converting it to text data, and then having Pepper speak.

(a) Pepper (b) Azure Kinect (c) THETA V (d) Oculus Quest 2

Fig. 1. System Components

2.2 Prototype

Figure 2 shows the system configuration of the prototype developed in this study.

In this system, when the user stands in front of Azure Kinect, the three-dimensional coordinate information of the user's joints is acquired by the body tracking function. The roll and pitch angles of both shoulders and the roll and yaw angles of both elbows are calculated from the acquired three-dimensional coordinate information of the user's joints and sent to the server via HTTP communication. Through the Pepper control software Choregraphe, the roll and pitch angles of the shoulders and the roll and yaw angles of the elbows are read from the server, and by controlling Pepper's shoulders and elbows, the user's arm motions and Pepper's arm motions are linked to each other. In this prototype, considering that Pepper does not have two legs, only Pepper's arms are linked to the user, without any lower body movement or mobility.

For voice, the user's voice is collected by the PC's microphone, converted to text data by the Web Speech API, and sent to the server via HTTP communication. The text data of the voice sent to the server is read through Choregraphe, and the voice is output in Pepper's voice.

For the viewing, the THETA V camera is attached to Pepper's chest, and real-time 360-degree video acquired from the THETA V is displayed on the Oculus Quest2 HMD using an application created with the game engine Unity. Specifically, a sphere object is created in Unity, and real-time 360-degree video acquired from the THETA V is pasted onto it. It is designed to allow the user to view a 360-degree image from inside the sphere, and to realize 360-degree free viewing. By building this Unity application on the Oculus Quest2, users can freely view the scene where the THETA V is located through the

HMD. The THETA V has the internal capability to communicate with devices in a local environment via wireless LAN and can display real-time 360-degree images directly to the Oculus Quest2. In this study, a prototype was developed using the wireless LAN communication feature of the THETA V, since it was decided to experimentally use different rooms in the same building as remote locations.

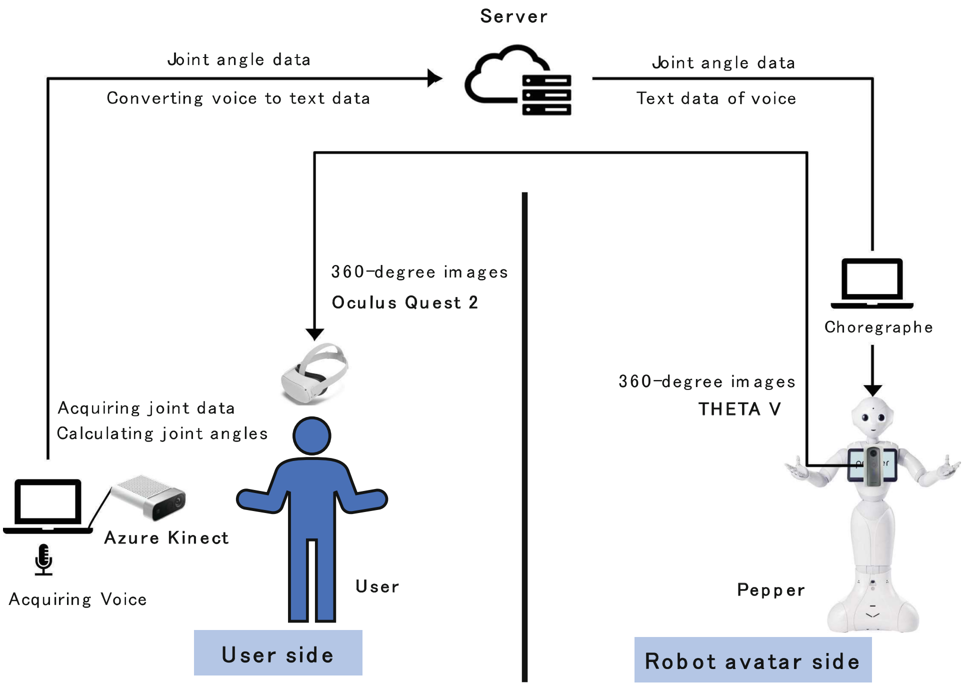
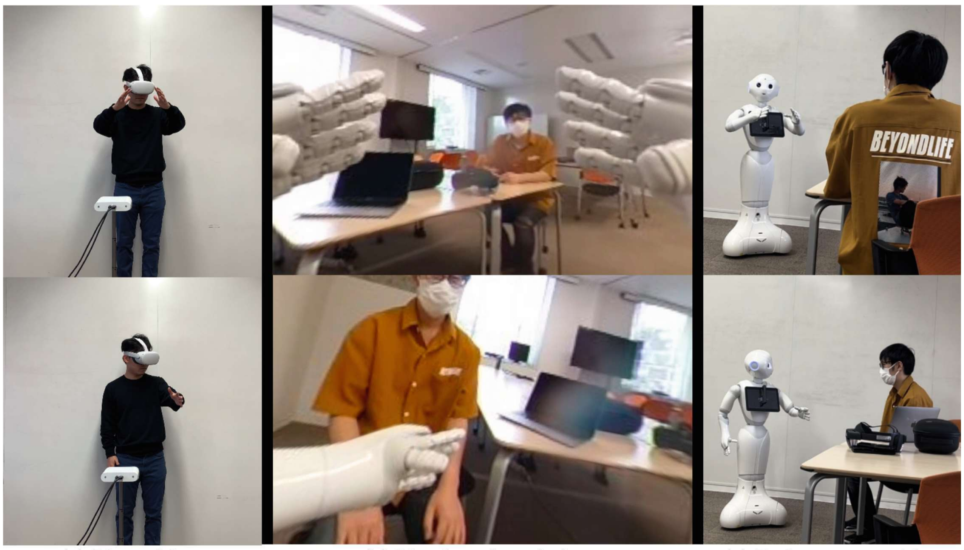


Fig. 2. System Configuration

Figure 3 shows the operation of the prototype developed. The user can move his or her body while viewing a 360-degree image on the HMD (Fig. 3(a)). Pepper's arm is visible from the user's viewing angle, so the user can communicate with the other person while watching Pepper's arm move in conjunction with the user's movements (Fig. 3 (b)). On the robot avatar side, the robot avatar can communicate with people on-site through gestures and speech (Fig. 3 (c)). In addition, the voice of the person on the robot avatar side is communicated with the user side using the call function of the ZOOM video conferencing tool on PCs.



(a) User side (b) User's point of view (c) Robot avatar side

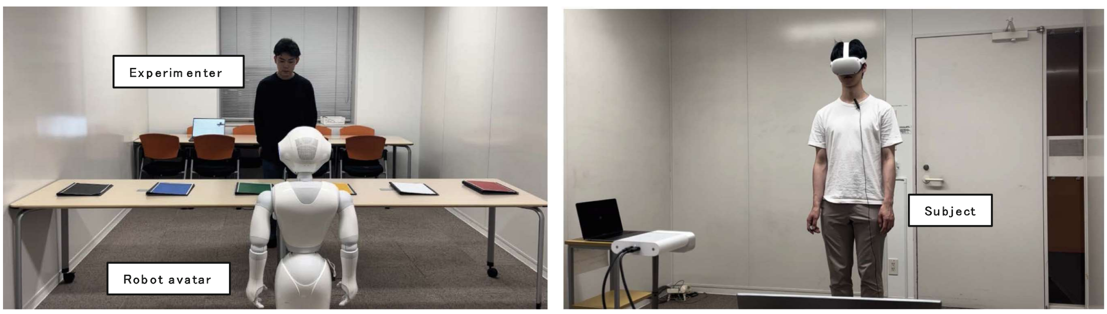
Fig. 3. Prototype of the developed tele-immersion system

3 Evaluation Experiment

As an evaluation experiment of this system, an experiment was conducted in which subjects performed collaborative tasks with the other person via a robot avatar. In this experiment, the effectiveness of the developed system and the sensation that one's own body is transferred to the body of the robot avatar were evaluated.

3.1 Experimental Environment

Figure 4 shows the experimental environment. The experimental environment consisted of two rooms, one with an experimenter and a robot avatar (Fig. 4(a)), and the other with a subject (Fig. 4(b)). As shown in Fig. 4(a), the experimenter and the robot avatar faced each other across a 70-cm-high table with six files of six different colors. The distance from the file to the robot avatar was 1 m, and the six files of six different colors were placed on the table at intervals of 50 cm.



(a) Room on the robot avatar side (b) Room on the subject side

Fig. 4. Experimental Environment

3.2 Experimental Method

Specific experimental methods are described below. First, the experimenter and the robot avatar begin the experiment facing each other across the table. The experimenter, by voice, instructs the subject's robot avatar to select a file of a particular color. The subject selects a file of the proper color and makes a gesture pointing in the direction of the correct file. For example, when the experimenter says, "Please select the red file," the subject, via the robot avatar, points with a gesture in the direction of the red file, and the robot avatar's voice says, "The red file is this one." This trial was performed three times, which constituted one set. The experiment was conducted in two sets, with participants completing a questionnaire after the first set and a questionnaire after the second set.

The experiment was conducted on a total of 10 subjects. The questionnaire items are shown in Table 1, with a 7-point Likert scale (1: "Strongly disagree", 2: "Disagree", 3: "Slightly disagree", 4: "Neither", 5: "Slightly agree", 6: "Agree", 7: "Strongly agree").

Table 1. Questionnaire items

Q1. Did you understand the instructions?
Q2. Did you feel as if yourself was in the room with the other person?
Q3. Did you feel as if the robot you were operating was in the same room with the other person, and not you?
Q4. Did you feel as if yourself collaborated with the other person?
Q5. Did you feel as if the robot you were operating collaborated with the other person, and not you?
Q6. Did you find yourself communicating easily with the other person?
Q7. Did you find the robot you operated easier to communicate with the other person, and not you?
Q8. Did you feel the robot's body like your own body?
Q9. Did you feel like you were moving the robot body yourself?
Q10. Did you feel like you were being a robot?
Q11. Did you feel like you had moved like a robot?

4 Results and Discussion

Figure 5 shows the results of the experiment. First, from the evaluation of Q1, regarding the understanding of instructions, it can be seen that the subject was able to communicate while understanding the experimenter's instructions through the robot avatar.

Next, a t-test was performed to compare the evaluation results of the two counties in terms of whether the subjects feel the sense of being in the same room (Q2, Q3), the sense of completing a task collaboratively (Q4, Q5), and the ease of communication (Q6, Q7), either for themselves or for the robot they are operating rather than themselves. The comparison of Q2 and Q3 by t-test showed a significant difference ($p = 0.020$), and the evaluation of the sense of oneself being in the same room with the other person in Q2 was higher than the sense of the robot I am operating being in the same room with the other person in Q3. The comparison of Q4 and Q5 by t-test showed a significant difference ($p = 0.006$), and the evaluation of the sense of oneself collaborating tasks with the other person in Q4 was higher than the sense of the robot I am operating collaborating tasks with the other person in Q5. The comparison of Q6 and Q7 by t-test showed a significant difference ($p = 0.003$), and the evaluation of the sense of oneself being easy to communicate with the other person in Q6 was higher than the sense of the robot I am operating being easy to communicate with the other person in Q7. From this, it can be said that this system enabled communication with a sense of oneself being in the same room as the other person and a sense of collaborating with the other person. In other words, the results suggest that this system is effective as a tele-immersion system.

Evaluations for the sense of body ownership in Q8 and the sense of agency in Q9 were over 4.5, indicating that the subjects generally felt the sense of body ownership and the sense of agency. In short, the robot avatar's body felt like one's own body, and one could feel that one was moving that body, and it is thought that there was the generation of the sensation that one's own body was transferred to the robot avatar's body.

In Q10 and Q11, the evaluations were also over 4.5, indicating that the subjects generally felt as if they were being a robot and that they were moving like a robot. The Proteus effect [12–15] is the effect of the avatar's appearance and physical characteristics on the user's psychological state and behavioral characteristics. It can be suggested that the Proteus effect was at work in this system.

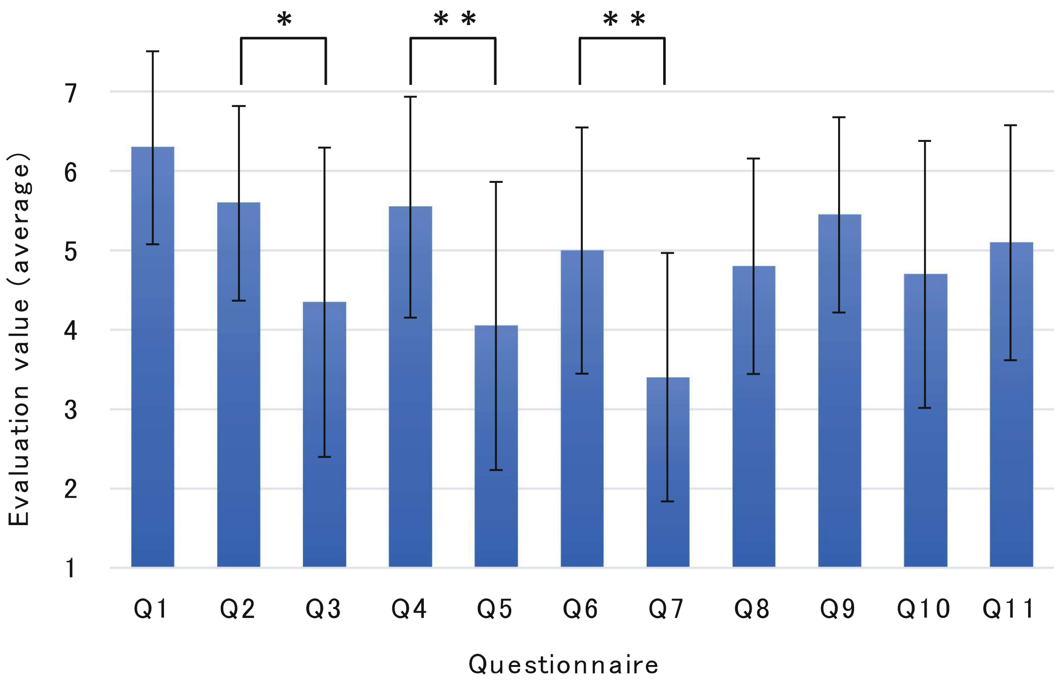


Fig. 5. Experimental Results

5 Conclusion

In this study, we aimed to develop a tele-immersion system that allows users to communicate with remote people as if they were in the same place, with the sensation that one's own body is transferred into the body of the robot avatar. In addition, a prototype was developed and evaluated it through evaluation experiments. As a result, users felt as if they were in the same room with the other person, and as if they were collaborating with the other person, and it was confirmed that this system was effective as a tele-immersion system. In addition, it was found that the user generally feels the sense of body ownership and the sense of agency toward the robot avatar in this system and that there is a generation of a sensation that the user's body is transferred to the robot avatar's body.

This system has the potential to be applied in various services and environments in the future. For example, it is expected to be used as a tele-immersion system in train stations, airports, tourist attractions, etc. In this case, the implementation of a mobility function for the robot avatar and its application to multilingual services are future prospects.

Acknowledgments. This research was partially supported by JKA promotion funds from KEIRIN RACE (2022M-272). This work was supported by JST SPRING, Grant Number JPMJSP2123.

References

1. Han, E., et al.: People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. J. Comput.-Mediat. Commun. **28**(2), 1–27 (2023)
2. Batnasan, G., Gochoo, M., Otgonbold, M.E., Alnajjar, F., Shih, T.K.: ArSL21L: arabic sign language letter dataset benchmarking and an educational avatar for metaverse applications. In: 2022 IEEE Global Engineering Education Conference (EDUCON), pp. 1814–1821 (2022)
3. Ferrer, C.D.R., Fujiwara, T.: Work-in-progress—avatar vs teams: co-presence and visual behavior in a videoconference cooperative task. In: 2022 8th International Conference of the Immersive Learning Research Network (iLRN), pp. 1–4 (2022)
4. Avatarin Inc. https://avatarin.com/
5. Takeuchi, K., Yamazaki, Y., Yoshifuji, K.: Avatar work: telework for disabled people unable to go outside by using avatar robots. In: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, pp. 53–60 (2020)
6. Kilteni, K., Groten, R., Slater, M.: The sense of embodiment in virtual reality. Presence: Teleoperators Virtual Env. **21**(4), 373–387 (2012)
7. Matsumiya, K.: Awareness of voluntary action, rather than body ownership, improves motor control. Sci. Rep. **11**(418), 1–14 (2021)
8. Pandey, A.K., Gelin, R.: A mass-produced sociable humanoid robot: pepper: the first machine of its kind. IEEE Robot. Autom. Mag. **25**(3), 40–48 (2018)
9. Carros, F., et al.: Care workers making use of robots: Results of a three-month study on human-robot interaction within a care home. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2022)
10. Mou, W., Ruocco, M., Zanatto, D., Cangelosi, A.: When would you trust a robot? a study on trust and theory of mind in human-robot interactions. In: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 956–962 (2020)

11. Hsieh, W.F., Sato-Shimokawara, E., Yamaguchi, T.: Investigation of robot expression style in human-robot interaction. J. Robot Mechatron. **32**(1), 224–235 (2020)
12. Yee, N., Bailenson, J.: The Proteus effect: the effect of transformed self-representation on behavior. Hum. Commun. Res. **33**(3), 271–290 (2007)
13. Fox, J., Bailenson, J.N., Tricase, L.: The embodiment of sexualized virtual selves: the Proteus effect and experiences of self-objectification via avatars. Comput. Hum. Behav. **29**(3), 930–938 (2013)
14. Praetorius, A.S., Görlich, D.: How avatars influence user behavior: a review on the proteus effect in virtual environments and video games. In: Proceedings of the 15th International Conference on the Foundations of Digital Games, pp. 1–9 (2020)
15. Stavropoulos, V., Pontes, H.M., Gomez, R., Schivinski, B., Griffiths, M.: Proteus effect profiles: how do they relate with disordered gaming behaviours? Psychiatr. Q. **91**, 615–628 (2020)