Effect of Embodiment Presentation by Humanoid Robot on Social Telepresence

Ikkaku Kawaguchi[*], Yuki Kodama[*], Hideaki Kuzuoka[*], Mai Otsuki[*], Yusuke Suzuki[†]

[*]University of Tsukuba, Tsukuba, Ibaraki, Japan,
s1430196@u.tsukuba.ac.jp, kodama.yuki49@gmail.com,
kuzuoka@iit.tsukuba.ac.jp, otsuki@emp.tsukuba.ac.jp

[†]OKI Electric Industry Co.Ltd.
Warabi, Saitama, Japan
suzuki543@oki.com

ABSTRACT

In this study, we used a humanoid robot as a telepresence robot and compared with the basic telepresence robot which can only rotate the display in order to reveal the effect of embodiment. We also investigated the effect caused by changing the body size of the humanoid robot by using two different size of robots. Our experimental results revealed that the embodiment increases the remote person's social telepresence, familiarity, and directivity. The comparison between small and big humanoid robots showed no difference and both of them were effective.

Author Keywords

Social Telepresence; Embodiment; Humanoid; Size of robot

ACM Classification Keywords

H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interface-Computer-supported cooperative work.

INTRODUCTION

It is well known that non-verbal information plays an important role in face-to-face communication [2]. However, in remote communication using ordinary video conferencing system, it is difficult to transfer non-verbal information due to the discrepancy caused by communication media [3]. One of the solutions to this problem is a telepresence robot which acts as a surrogate of a remote participant [5, 12, 13]. In general, a telepresence robot has a mechanism to move a display which shows a remote person's face (hereafter, 'moving display telepresence robot'), and they found that communication became smoother and the remote person's sense of presence increases by reflecting a remote person's motion to the display. However, such moving display causes the cases for a local participant to watch the display from oblique angles which leads to the misunderstanding of gaze directions of remote participant and decrease the remote participant's presence [5, 6, 13]. To solve these problems, alternative methods are necessary. As extended forms of a moving display telepresence robots, telepresence robots that are extended with human-like body parts (e.g. arms) have been developed (here after, 'body-parts extended telepresence robot') to represent remote participants' bodily gestures [1, 14]. In addition to the studies with telepresence robots, effect of body gestures have also been investigated using communication robots [8, 15]. These studies show that presentation of the robot's body gestures, such as head, body and arm movement, improve the impression of the remote participant or the robot itself from the various aspects. Although various telepresence robots have been proposed, no comparison was made between a moving display telepresence robot and a body-parts extended telepresence robot, in terms of the sense of presence of the remote participant. Thus the purpose of our study is to compare the body-parts extended telepresence robot and the moving display telepresence robot to reveal the effect of embodiment. As the body-parts extended telepresence robot, we used a humanoid robot with physical body parts (head, body, and arms). We thought the humanoid robot with body parts similar to human is most effective because the movement of the robot's body parts is known to improve the impression of remote participant and robot itself [1, 8, 14, 15]. In addition, it is known that the size of the robot affect the impression that human receives [4, 11], so in this study we compared

Display rotate Small Humanoid robot Big Humanoid robot

Figure 1. Experimental conditions. These pictures show the appearance when each robot is explaining the document.

HAI '16, October 04-07, 2016, Biopolis, Singapore
ACM 978-1-4503-4508-8/16/10.
http://dx.doi.org/10.1145/2974804.2980498

humanoid robots in a two sizes (table top size, and close to human life-size).

EXPERIMENT

Referring to Nakanishi's previous study [10], we conducted a task which a remote participant gives explanation to the local participant using a telepresence robot. In addition, we assumed quasi-naturalistic situation of using the telepresence robot and included attentional guidance to the local document which required to transmit the gaze not only to the front person, but also to other direction. For evaluation, we used a questionnaire mainly on social telepresence and conducted a video analysis. An interview was also conducted after all the tasks were finished.

In this study, we compared three conditions. As Figure 1 shows, we used three types of robots corresponding to each condition.

Display Rotate Condition: This condition used a rotatable table which can rotate in horizontal direction with a 10-inch tablet (iPad, Apple Inc.) presenting the video of the remote participant's upper body. Attentional guidance to the local document was performed by rotation of the display and finger pointing gesture in the video. The directivity in horizontal plane was presented by display rotation, so the direction of the head and finger in the video was fixed to the frontal direction.

Small Humanoid Robot Condition: This condition used a humanoid robot with a total height of 60cm. We developed a humanoid robot which consists of a head, body, and two arms. 5 inch smartphone (305SH, SHARP Co.) was placed on the head to present the video of the remote participant. Only the face of the remote participant was shown in the display and the size of the face image is adjusted to the same size as the display rotate condition. Attentional guidance to the local document is performed by the head-body rotation and pointing movement using the arm. The directivity in horizontal direction is presented by the robot's head-body rotation, so the direction of the head in the video is fixed to the frontal direction.

Big Humanoid Robot Condition: A condition using a humanoid robot with a height of 120cm. The humanoid robot was developed utilizingTalkTorque2 [9]. 5-inch smartphone (305SH, SHARP Co.) was placed on its head to present the video of the remote participant. The video and body movement of the robot was same as the small humanoid.

Since we aimed to compare the effect of embodiment itself, we presented the remote participant's face image in all conditions. In conditions using a humanoid robot, only the remote participant's face was shown in the smartphone placed on the robots' head and the body movement of the robot presented remote participant's gesture, while in the display rotate condition, both remote participant's face and gestures were presented by the video shown in the display. Skype was used for video communication in all conditions.

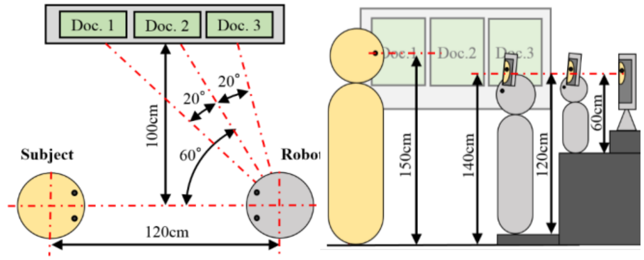

Figure 2. Overview of the environment of experiment.

Experimental setup and task

Figure 2 shows an overview of experimental environment. From the remote site, a remote participant (experimenter) communicated with a local participant (subject) through a telepresence robot and explained about the three documents placed on the whiteboard which is related to a fast food restaurant. This experiment used a within subject design, so we prepared three scripts for each trial. Order of the conditions was counter balanced to minimize the learning effect. In each script, three yes-or-no questions were embedded to have an easy conversation for the subject, because if there were no conversation in the script, subject may feel that the remote participant as recorded video. One description took about 2 minutes. Although we used three scripts, movement and duration of the remote participant/robots were the same between each script to eliminate the effect of script difference. We recorded the behavior of the subject and conducted video analysis. In video analysis, we focused the subject's gaze especially mutual gaze between the robot and the subject because it plays an important role in human communication [7] (e.g. regulate interaction, express intimacy, and facilitate service and task goals).After each task, we gave a questionnaire to the subjects. The questionnaire contained 8 questions, including questions on social telepresence (Q1-Q4), naturalness of the movement (Q5), directivity of the robot (Q6, Q7), and sense of affinity (Q8). The questions are shown on Figure 3. We used the 7 point Likert-scale for our questionnaire. An interview was also conducted after all the tasks were finished. We recruited 12 subjects from a local university (all of them were male). Their ages ranged from 20 to 24 years old.

Results

Questionnaire Results

To evaluate the difference of each condition, a one-way within subjects ANOVA was conducted to compare the effect of condition on questionnaire result in three conditions. The result of each question is shown in Figure 3. There were significant effects of conditions in the questions except Q3. As the post hoc test, multiple comparisons with Bonferroni correction was carried out on questions except Q3. Significant differences were shown in Figure 3.

Evaluation of Subject's Gaze Direction

For the statistical analysis we measured the ratio of duration that the subject was looking at the robot in each trial based on the video analysis (Figure 4(a)). We also measured the ratio of duration that mutual gaze was achieved between the robot and the subject (Figure 4(b)). A one-way within subjects ANOVA was conducted to compare the effect of condition on these two results. On the ratio of duration that the subject was looking at the robot, there were significant effects of conditions ($F(2,22) = 19.812$, $p < 0.001$). As the post hoc test, multiple comparisons with Bonferroni correction was carried out and there was significant trend between display rotate condition and big humanoid robot condition ($p = 0.087$). On the ratio of duration that mutual gaze was achieved, there was no significant effect of conditions ($F(2,22) = 2.064$, $p = 0.151$).

As the additional evaluation of mutual gaze, we measured delay before mutual gaze had achieved. In each task remote participant looked at the subject 4 times, so totally 48 samples were obtained for each condition. Table 1 show the result of measured delay until mutual gaze had achieved. We conducted Friedman test and there was significant effect of condition on the delay ($p<0.001$). As the post hoc test, Dunn-Bonferroni test was carried out and there was significant differences between display rotate condition and big humanoid robot condition ($p<0.001$), and there was significant trend between display rotate condition and small humanoid robot condition ($p = 0.057$).

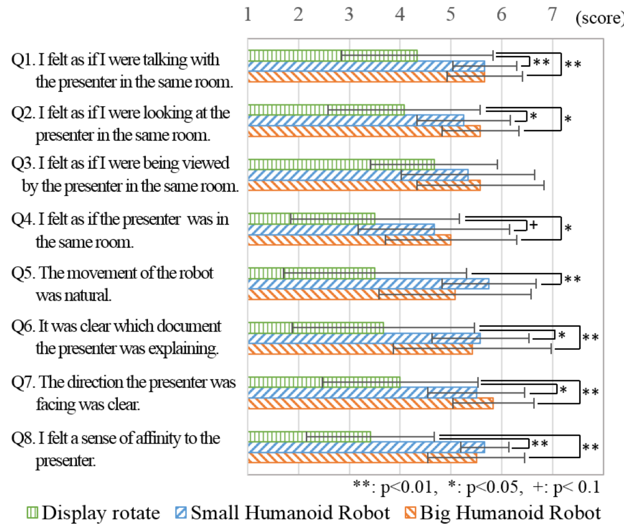

Figure 3. Results of the questionnaires

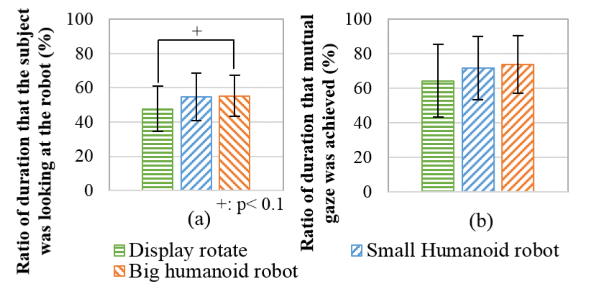

Figure 4. Result of video analysis of subject's gaze direction.

Table 1. Result of measured delay before mutual gaze had achieved.

Condition	N	Mean	SD	Min.	Max.	Percentiles 25th	50th	75th	Mean Rank
Display rotate	48	1.49	2.80	0	5.11	0	0.25	1.20	2.42
Small humanoid robot	48	0.67	1.57	0	8.24	0	0	0.73	1.94
Big humanoid robot	48	0.38	0.93	0	13.48	0	0	0.03	1.65

**: p<0.01, +: p<0.1

DISCUSSION

Effect of embodiment presentation

From the result of the questionnaire (Figure 3), two humanoid robot conditions were scored significantly higher than the display rotate condition in the questions except Q3. As shown in Figure 1, the difference of these conditions was medium (e.g. the video image, the small/big humanoid robot) to present the body gestures. So the differences between these conditions were considered to be the effect of the embodiment presentation by using humanoid robots.

For the questions about social telepresence (Q1-Q4), significant differences were seen in questions Q1, Q2 and Q4 between display rotate condition and small/big humanoid robot condition (In Q4, significant trend between display rotate condition and small humanoid condition). From these results, presenting physical gestures by humanoid robot increases the social telepresence of the remote participant. Meanwhile, there was no significant difference in Q3. In all conditions, the display faced straight toward the subject when the remote participant see the subject, so problems caused by viewing the display from an oblique angle [6, 13] did not occur. We are assuming that this is the reason why there was no significant difference between each condition for Q3.

For the questions about naturalness of the movement (Q5), significant difference was seen between display rotate condition and small humanoid condition. From the interview, subjects pointed out that Mona Lisa effect occurred in the display rotate condition when the display was rotated. Thus the mismatch in the direction of the display and the perceived gaze direction from the remote participant's face in the video might have decreased the score of how natural the movement was in the display rotate condition. In the conditions using humanoid robots, the display placed on the robot's head rotated in the same way as the display rotate condition. Thus, it is possible for the Mona Lisa effect to occur. However, the awareness of the robot's head motion seemed to be stronger than the effect of the display rotation, decreasing the Mona Lisa effect in the conditions using a humanoid robot.

For the questions about the directivity of the robot, significant differences were seen for both Q6 and Q7 between display rotate condition and small/big humanoid conditions. The difference between display rotate condition and small/big humanoid condition was caused by the difference of directivity between display rotation and humanoid robots' body gesture. This result indicates that physical movement using the humanoid robot is more effective than display rotation for presentation of directivity.

For Q8, a significant difference was seen between display rotate condition and small/big humanoid condition. It is conceivable that the sense of affinity to the remote participant increased by the increase of social telepresence.

From the result of the video analysis, mutual gaze was achieved significantly faster in big humanoid condition than display rotate condition. Small humanoid condition also had a tendency that subjects achieve mutual gaze faster than display rotate condition. Thus, we can assume that the physical motion of the humanoid robots increased awareness so subject could react to remote participant's gaze faster and achieved mutual gaze. The ratio of looking at the robot (Figure 4(a)) tend to increase in the big humanoid condition than the display rotate condition. We can assume that increased awareness by the physical motion of the humanoid robots invited more gaze from the subject.

Effect of the humanoid robot's size

In the experiment, there was no significant difference between small and big humanoid condition in all results. Based on these results, the effect of the embodiment presentation discussed in the previous section is valid regardless of the robot's size.

In the analysis of delay before mutual gaze had achieved significant difference was seen between big humanoid condition and display rotate condition but significant trend was seen between small humanoid condition and display rotate condition. In the analysis of ratio of duration that the subject was looking at the robot, significant difference was only seen between big humanoid condition and display rotate condition. Thus, the big humanoid robot seemed to be more effective to invite the gaze of subjects.

CONCLUSION

Overall, humanoid robot conditions increased social telepresence, directivity, and sense of affinity compare to display rotate condition. We could not find significant difference between small humanoid condition and big humanoid condition. Our results indicate that, regardless of its size, the humanoid robot can alleviate the problems of display moving telepresence robot [5, 6, 13], e.g., low telepresence, misunderstanding of gaze direction, low participation. Though we could not find significant difference between small and big humanoid condition, these two robots has different feature (e.g. Big humanoid seemed to be more effective to increasing the presence of remote participant and invites the gaze of subjects, small humanoid was perceived as more naturalistic.). These features may affect the subject differently according to the situation. Thus careful consideration for the size of a telepresence robot is still necessary.The limitation of our study is that the task we used was somewhat artificial and our assessment was based mainly on subjective evaluation. Thus, in our next step, quantitative assessment and observational analysis with more naturalistic task is necessary (i.e. multi-party meeting).

ACKNOWLEDGEMENTS

This work was supported by JSPS Grant-in-Aid for JSPS Research Fellow JP16J06932.

REFERENCES

1. Adalgeirsson, S. O., and Breazeal, C. Mebot: a robotic platform for socially embodied presence. In *Proc. HRI 2010*, IEEE Press (2010), 15-22.
2. Goodwin, C. Professional vision. *American anthropologist 96*, 3 (1994), 606-633.
3. Heath, C., and Luff, P.: Disembodied conduct: communication through video in a multi-media office environment. In *Proc. CHI 1991*, ACM Press (1991), 99-103.
4. Hiroi, Y., and Ito, A. Are bigger robots scary?-The relationship between robot size and psychological threat? In *Proc. AIM 2008*. IEEE Press (2008), 546-551.
5. Jacob, T.B., Daniel, A. and Anthony, D. Not Really There: Understanding Embodied Communication Affordances in Team Perception and Participation, In *Proc*. *CSCW 2015*, ACM Press (2015), 1567-1575.
6. Kawaguchi, I., Kuzuoka, H. and Suzuki, Y. Study on Gaze Direction Perception on Face Image Displayed on Rotatable Display, In *Proc. CHI 2015*, ACM Press (2015), 1729-1737.
7. Kleinke, C. L. Gaze and eye contact: a research review. *Psychological bulletin*, 100.1(1986), 78.
8. Kuzuoka, H., Kosaka, J., Yamazaki, K., Suga, Y., Suga, Y., Yamazaki, A., Luff, P. and Heath, C. Mediating dual ecologies. In *Proc. CSCW 2004*, ACM Press (2004), 477-486.
9. Madhumita, G., and Kuzuoka, H. An Ethnomethodological Study of a Museum Guide Robot's Attempt at Engagement and Disengagement. Journal of Robotics, Hindawi Publishing (2014).
10. Nakanishi, H., Kato, K. and Ishiguro, H. Zoom Cameras and Movable Displays Enhance Social Telepresence. In *Proc. CHI 2011*, ACM Press (2011), 63-72.
11. Rae, I., Takayama, L., and Mutlu, B. The influence of height in robot-mediated communication. In *Proc. HRI*2013, IEEE Press (2013). 1-8.
12. Revolve Robotics - Kubi. https://revolverobotics.com
13. Sirkin, D., Venolia, G., Tang, J., Robertson, G., Kim, T., Inkpen, K., and Sinclair, M. Motion and attention in a kinetic videoconferencing proxy. In *Human-Computer Interaction–INTERACT 2011*, Springer (2011), 162-180.
14. Tanaka, K., Nakanishi, H., and Ishiguro, H. Comparing video, avatar, and robot mediated communication: pros and cons of embodiment. Collaboration Technologies and Social Computing, Springer Berlin Heidelberg (2014), 96-110.
15. Xu, Q., Liyuan, L. and Gang W. Designing engagement-aware agents for multiparty conversations. In *Proc. CHI2013*, ACM Press (2013), 2233-2242.