Teleoperated or Autonomous?: How to Produce a Robot Operator's Pseudo Presence in HRI

Kazuaki Tanaka
Department of Adaptive Machine Systems,
Osaka University, CREST, JST
Suita, Osaka, Japan
tanaka@ams.eng.osaka-u.ac.jp

Naomi Yamashita
NTT Communication Science Laboratories
Souraku, Kyoto, Japan
naomiy@acm.org

Hideyuki Nakanishi
Department of Adaptive Machine Systems,
Osaka University
Suita, Osaka, Japan
nakanishi@ams.eng.osaka-u.ac.jp

Hiroshi Ishiguro
Department of Systems Innovation,
Osaka University
Toyonaka, Osaka, Japan
ishiguro@sys.es.osaka-u.ac.jp

***Abstract*—Previous research has made various efforts to produce human-like presence of autonomous social robots. However, such efforts often require costly equipment and complicated mechanisms. In this paper, we propose a new method that makes a user feel as if an autonomous robot is controlled by a remote operator, with virtually no cost. The basic idea is to manipulate people's knowledge about a robot by using priming technique. Through a series of experiments, we discovered that subjects tended to deduce the presence/absence of a remote operator based on their prior experience with that same remote operator. When they interacted with an autonomous robot after interacting with a teleoperated robot (i.e., a remote operator) whose appearance was identical as the autonomous robot, they tended to feel that they were still talking with the remote operator. The physically embodied talking behavior reminded the subjects of the remote operator's presence that was felt at the prior experience. Their deductions of the presence/absence of a remote operator were actually based on their "beliefs" that they had been interacting with a remote operator. Even if they had interacted with an autonomous robot under the guise of a remote operator, they tended to believe that they were interacting with a remote operator even when they subsequently interacted with an autonomous robot.**

Keywords—teleoperated robot; autonomous robot; Turing test; physical embodiment; telepresence; social presence; social response

I. Introduction

Previous studies on autonomous social robots have made various efforts to produce human-like presence. However, such efforts, e.g., reproducing realistic appearance and fine movements [5][18], often require costly equipment and complicated mechanisms. In this paper, we propose a method that improves the presence of an autonomous robot with virtually no cost - by manipulating user's knowledge. According to previous study in the social psychology field, people tend to make inferences about others (including robots) based on their prior experiences/knowledge [2][8]. We hence focused on the relationships between user's prior experience (i.e., priming) and their perception of the presence. If a user interacts with an autonomous robot after interacting with a teleoperated robot whose appearances are the same, a user may feel that the robot is controlled by a remote operator. In other words, the experience of interacting with a teleoperated robot may prime the user to recall the remote operator's presence when interacting with the autonomous robot if the appearances of the two robots are identical and their movements are similar.

In this study, we evaluate the presence through experiments resembling the Turing test. In a common Turing test, people decide whether an autonomous system's intelligence resembles that of a human. On the other hand, in our test called a "social telepresence test," people decide whether an autonomous system produces a remote operator's presence that we call a "pseudo presence" when this remote operator's presence is produced by an autonomous system.

Pseudo presence is related to social telepresence and social presence. Social telepresence is the degree to which people feel as if they are talking face-to-face with a remote partner [4] on the other side of a communication medium. Social presence, which is often used in the human-robot interaction field, is defined as the degree to which people treat a robot as a human partner [1][15]. In contrast, pseudo presence is the degree to which people feel as if an autonomous robot is actually controlled by a remote operator.

Pseudo presence could be valuable in many cases. For example, consider a case where a user receives care from a caregiver robot [17][21]. The caregiver robot is operated either in an autonomous mode or a tele-operated mode, depending on the situation so that remote caregivers can engage in other activities during the autonomous mode. In such a case, pseudo presence of a remote caregiver might reduce user feelings of loneliness even when the robot is actually operated in an autonomous mode. In case of a teacher robot [5], if students feel the remote teacher's pseudo presence in an autonomous lecture, the students might pay attention to autonomously played lectures.

II. Related Work

When talking through a teleoperated robot, although a user can see body motions that are controlled by a remote operator, they normally cannot see the operator's current appearance. Some studies reported the superiority of teleoperated robots over other communication media, such as videoconferencing [9][18][20]. Studies have shown that a teleoperated robot with a realistic human appearance enhances social telepresence

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

more than audio-only conferencing and videoconferencing [18]. Even a human-looking anonymous robot without a specific age or gender [14] can produce higher social telepresence than voice and avatar chats [20]. Since people can believe that an operator is controlling the robot without seeing the operator's appearance, perhaps an autonomous robot will also produce the remote operator's presence: pseudo presence.

In terms of how to improve telepresence, previous studies have suggested that physical embodiment is one factor that enhances social telepresence [20] and builds trust [16]. For an autonomous robot, several studies have indicated that physical embodiment produces higher social presence than on-screen agents [1][7]. Building on previous studies, we suspect that the physical embodiment of an autonomous robot might also contribute to produce the pseudo presence of a remote operator.

III. Research Question

To improve human-like presence of autonomous robots, we addressed the factors that produce pseudo presence. As described in Section I and II, prior experience of talking with a teleoperated robot (i.e., priming effect) and physical embodiment might be the factors.

Even though researchers have developed robots that can be controlled by a remote operator and an autonomous system [17], the effect of such a robot on producing the operator's presence has not been addressed. In this paper, we pursue two research questions: 1) whether presenting physically embodied motions effectively produces pseudo presence; and 2) whether the experience of talking with a remote operator through a teleoperated robot produces pseudo presence when interacting with an autonomous robot whose appearance is identical to the teleoperated robot.

IV. Current Study

In this paper, we first introduce a pre-experiment, which developed an autonomous system that generates natural talking behaviors and used it to examine our two research questions (Section III) in three experiments. Experiment 1 (Section V) compared the presence and absence of physical embodiment and priming to confirm whether these factors contributed to producing pseudo presence [20]. The findings of Experiment 1 led us to a set of hypotheses asking what experiences are effective for producing pseudo presence. Experiment 2 and 3 (Sections VI and VII) were conducted to test those hypotheses. Specifically, we compared different experiences in which the robot was listening or speaking. The following section describes our autonomous system and then shows that it successfully generated natural talking behaviors.

A. System Development: Our Autonomous System

In this study we used a humanoid robot with a human-like anonymous face [14], a three-degrees-of-freedom neck, and a one-degree-of-freedom mouth.

The roles of the interaction partner are listener and speaker whose behaviors are mainly nod and lip motions. We constructed a backchannel system that detects the appropriate timing of backchannel feedback from the user's speech and a

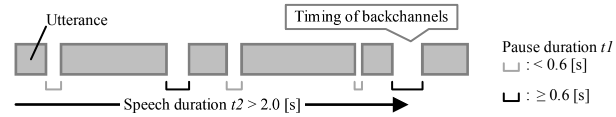

Fig. 1. Method that detects timing of backchannel responses.

lip-sync system that generates lip motions synchronized with pre-recorded speech. We simplified the system so that our findings would apply to more complicated systems.

To construct back-channel and lip-sync systems that generate natural talking behaviors, we conducted a series of preliminary experiments. The subjects in our pre-experiments spoke to a robot that gave backchannel responses generated by our autonomous system and judged whether the robot was teleoperated or autonomous. They also evaluated the degree of the naturalness of timing and the frequency of the robot's backchannel responses to adjust the parameters of our backchannel system. We repeated the procedure to refine the system until almost all the subjects judged that the robot was teleoperated.

Backchannel System: Many methods detect the best timing of backchannel responses during a user's speech. Most use prosodic information, including pause [13][19][22][24] and fundamental frequency [13][22][23]. Our method used only speech pause since it is good cue to identify sentence breaks or ends, which seem the appropriate timing of backchannel responses. One study also only used speech pause, although its algorithm is more complex than ours for estimating earlier timing [24]. The backchannel systems proposed by these previous works detected more appropriate backchannel timing, but our simple algorithm was adequate so that subjects accept the remote operator's presence at a one-turn interaction.

The timing rule for providing backchannel responses is shown in Fig. 1. Each box represents an utterance, and the distance between each box is pause duration $t1$. The utterance and pause parts correspond to higher and lower sound pressure. The system judges t1 to be a target pause if it exceeds 0.6 seconds. Speech duration $t2$ is the elapsed time from the start of the speech to the time at which the target pause was recognized. If $t2$ exceeds 2.0 seconds, the system judges the target pause as the timing of the backchannel response and reset $t2$ to zero. This means that the system reproduced backchannel responses when the pause continued for 0.6 seconds after the speech continued for more than 2.0 seconds.

The pre-experiment results implied that backchannels, which are repeated in less than 2.0 seconds, decrease naturalness. In addition, backchannels, which are done more than 0.6 seconds after sentence breaks or ends, tended to be felt later; pauses that are less than 0.6 seconds are insufficient to judge sentence breaks or ends. We therefore set pause durations $t1$ and $t2$ to 0.6 and 2.0 seconds.

In the backchannel response, the robot nodded and a pre-recorded voice said "hai" ("yes" in Japanese). When we used only one pattern of nodding motion and voice, the subjects pointed out that the robot's response seemed monotonous. We therefore prepared three nodding motions with different degrees of pitch and speed and two voices that slightly differed in their tone. Preliminary experiments showed that subjects felt naturalness the most when the three nodding motions and the two voices were randomly played in robot's backchannel response.

Lip-sync System: Some lip-sync methods generate lip motions from a human's voice to control a robot [6][24] and a computer graphic avatar [3][24]. Since our robot had only one-degree-of-freedom in mouth movement, we used a simpler method to produce the robot's mouth movement. Our lip-sync system measured the acoustic pressure of the human's voice and related the level to the angle of the robot's chin. In other words, the robot's mouth was synchronized with the waveform of the human's voice. Our preliminary experiments, which used pre-recorded speech to produce the robot's lip movements, showed that this method worked the best in terms of naturalness.

B. System Development: Our Autonomous System

Below, we explain the shared methods and the terminologies of Experiments 1, 2, and 3.

1) Modes

We controlled the robot in the following two modes: **1) Teleoperated mode:** the robot's head and mouth moved at thirty frames per second based on the sensor data from the face tracking software (faceAPI). The software ran on a remote terminal and captured the remote operator's facial movements by a web camera. **2) Autonomous mode:** the robot moved based on the backchannel and lip-sync systems.

2) Procedure

A member of our research group acted as the remote operator. Before conducting each experiment, he directly met each subject and introduced himself as the remote operator. The robot's speeches and acoustic backchannel responses were his pre-recorded voices.

Each experiment included the following two phases: **1) Priming phase:** we manipulated the subjects' prior experience of talking with the robot in the teleoperated/autonomous modes. This manipulation is called "priming." **2) Testing phase:** After the priming phase, the subject talked with the robot in the autonomous mode. Before each phase, we revealed to the subjects which mode was used to control the robot. The conversations in each phase took about two minutes.

After the testing phase, the subjects answered questionnaires about their estimations of the pseudo presence in the testing phase. In the pre-experiments in which the subjects were told that the robot moved autonomously before talking, we confirmed that the subjects moderately felt as if the remote operator was listening, even though they knew that the robot autonomously moved. We also collected open-ended responses to infer what determined their scores.

V. EXPERIMENT 1

This experiment addressed two questions (described in Section III): whether the experience of talking with a remote operator and physically embodied motions produced the pseudo presence.

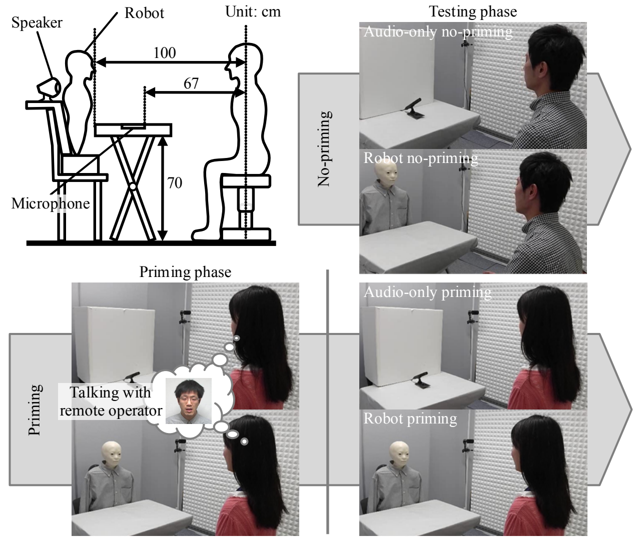

Fig. 2. Conditions of Experiment 1.

A. Conditions

As shown in Fig. 2, the subjects sat in front of a desk and faced the robot who was placed on the opposite side. A directional microphone was embedded in the desk to capture the subject's speech and hidden by a cloth. A speaker behind the robot produced the remote operator's speech.

We prepared the four conditions shown in Fig. 2: two audio-only and two robot conditions. To answer the first research question, we compared robot conditions with audio-only conditions, which do not present both physical embodiment and body motions. In the robot conditions, the subjects got an acoustic response with a nodding motion. In the audio-only conditions, no robot was used. Instead, we set a dummy microphone on the desk to suggest to the subjects that their speech was being listened to.

The experiment included priming and testing phases. Before the priming phase, the subjects were told that they would be talking with a remote operator in the teleoperated mode. Although the priming phase was actually conducted in the autonomous mode, the manipulation check (explained in Section D) confirmed that all the subjects believed that the remote operator was listening to their speech. The testing phase was conducted in the autonomous mode. Before it, the subjects were told that they would be talking with an autonomous system that autonomously gives backchannel responses and that their speech would be recorded instead of being listened to by the remote operator. When the subject stopped talking for five seconds, the system announced the end of the experiment in a pre-recorded voice. The two priming conditions included both the priming and testing phases, but the other two no-priming conditions only included the testing phase. To answer the first research question, we compared the presence/absence of the priming phase.

B. Subjects

Sixteen undergraduates participated in Experiment 1. Half (five females and three males) participated in both the priming and testing phases, and the other half (four females and four males) only participated in the testing phase. In each phase, they talked in both the audio-only and robot conditions. We counterbalanced the order of experiencing the audio-only and robot conditions.

C. Task

The subject was a speaker, and the robot or audio system was a listener who gave a backchannel response to his/her speech. This setting minimizes the time of playing the pre-recorded speech in the autonomous condition. If the subject is a listener, the autonomous system in the audio-only conditions only plays pre-recorded operator speeches unilaterally from the speaker, which would likely generate a disadvantage in the audio-only conditions over the robot conditions.

The subjects were asked to discuss the problems of various electronic devices and request new functions for them at the beginning of each conversation through the robot or the speaker. The topics in the priming and testing phases were portable audio players and robotic vacuum cleaners, and smartphones and 3D TVs, respectively. The order of the topics was counterbalanced.

D. Questionnaires

After talking about one topic, the subjects answered manipulation check questions to confirm whether they correctly understood our instructions. The manipulation check consisted of the following two sets of YES/NO statements:

- In the last conversation, a remote partner listened to your speech.
- In the last conversation, your speech was recorded instead of being listened to by a remote partner.

The following questionnaire statement estimated the pseudo presence:

- I felt as if the conversation partner was listening to me in the same room.

Since asking about feelings of being in the same room is useful to measure the remote partner's presence [10][11][20], we used the same representation to measure the pseudo presence. Answers were rated on a 7-point Likert scale: 1 = strongly disagree, 4 = neutral, 7 = strongly agree.

E. Result

According to the manipulation check, we confirmed that all the subjects believed that they had been talking to a remote operator in the priming phase. Experiment 1's result is shown in Fig. 3, where each box represents the mean scores of pseudo presence, and each bar represents the standard error of the mean. The figure compares the four conditions by a 2x2 mixed factorial ANOVA with 'embodiment' (audio-only vs robot) as a within-subjects factor and 'priming' (no-priming vs priming) as a between-subjects factor. We found an interaction between the factors ($F(1, 26)=5.561$, $p<.05$). We further performed a Tukey HSD test. Results indicated that the embodiment significantly increased pseudo presence when the subjects experienced priming ($p<.01$). Priming marginally increased pseudo presence when the subjects could see the physically

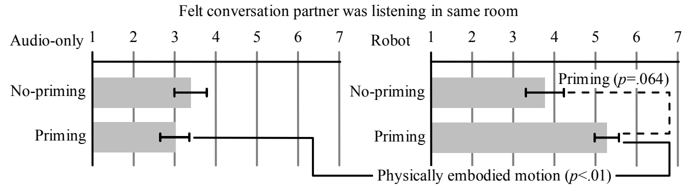

Fig. 3. Experiment 1 result.

embodied motion ($p=.064$). Therefore, both the embodiment and priming seem to be important factors to produce pseudo presence. This means that priming the subjects' beliefs that they were talking to the remote operator produced pseudo presence in the testing phase when they could see backchannel responses through the robot.

These results indicate that physically embodied motions and priming the subjects' beliefs are the factors that produce pseudo presence. However, the number of conversations might also have influenced the pseudo presence. In the priming conditions, the subjects had two conversations, but they only had one in the no-priming condition. Although the interaction between the number of conversations and the embodiment cannot be denied, we expected the priming to be the significant factor to produce pseudo presence since the effect of number of conversations was not seen in the audio-only conditions. Perhaps the physical movements lingered in the memories of the subjects and facilitated the priming effect. In Experiment 2, we controlled the number of conversations and tested the following hypothesis:

Hypothesis 1: Pseudo presence will be produced in subjects who believe that they are talking with a remote operator through a teleoperated robot that presents the operator's body motion.

In Experiment 1, even the experience where the robot gave only a backchannel response under the guise of a remote operator produced pseudo presence. We predicted that the experience of talking with a remote operator who is actually replying to the user's speech might produce a higher pseudo presence because various real operator responses depending on context can create a strong impression that the remote operator is listening. We hence set the following hypothesis and tested it in Experiment 2:

Hypothesis 2: Compared with the experience of talking with an autonomous robot under the guise of a remote operator, the experience of talking with a remote operator who can present interactive behaviors through a teleoperated robot will produce higher pseudo presence.

In Experiment 1, the subjects' open-ended responses, which explain the pseudo presence, suggested that all eight subjects estimated the pseudo presence based on the timing of the backchannel responses. There is a question whether the experience in which the robot unilaterally speaks to a subject produces pseudo presence. Because the robot is unilaterally talking and making pre-recorded speeches that resemble video messages, the user might feel less presence of a remote operator. In this case, it might be difficult to produce pseudo presence, since the information for estimating it (i.e., timing of backchannel responses) will be smaller. In Experiment 3, we addressed whether the priming by the experience of listening to a robot's speech can produce pseudo presence even with decreased interactive conversation (Section VII).

VI. EXPERIMENT 2

This experiment examined hypotheses 1 and 2 (described in Section V.E). Experiment 1 compared the presence/absence of priming, and Experiment 2 compared the difference of priming.

A. Conditions

We prepared three conditions (Fig. 4). The autonomous condition corresponded with the robot no-priming condition, but the subjects talked with the robot in the autonomous mode twice. The blur condition was identical to the robot-priming condition. Since the subjects who were assigned to the blur condition could not clearly recognize the border between the teleoperated and autonomous modes, we named it the "blur" condition. In these conditions, the robot was in the autonomous mode in both phases. Thus the only difference between the autonomous and blur conditions was the subjects' belief that they were talking with the autonomous or teleoperated robot. Comparing these conditions examined hypothesis 1.

We added a new teleoperated condition, in which the subjects received various responses to their speech. The remote operator repeated and rephrased the subject's opinions in addition to giving customary backchannel responses. Since such responses seemed difficult to automatize, the subjects easily believed that the remote operator was actually replying to their speech. The robot was in the teleoperated mode in the priming phase. Comparing the blur and teleoperated conditions examined hypothesis 2.

B. Subjects

Thirty undergraduate students participated in Experiment 2. None of the subjects in Experiment 1 participated in Experiment 2. Ten (six females and four males) participated in the autonomous condition. Ten (five females and five males) participated in the blur condition. Another group of five females and five males participated in the teleoperated condition.

C. Task

Basically, the task was the same as in Experiment 1. In the autonomous and blur conditions, the robot played pre-recorded instructions and acoustic backchannel responses. In the teleoperated condition, the remote operator actually instructed and replied to the subject's speech through the robot in the teleoperated mode. The topics in the priming and testing phases were 3D TVs and smartphones.

D. Measures

After the testing phase, the subjects answered manipulation check questions to confirm whether they correctly understood our instructions. The manipulation check consisted of the following two sets of YES/NO statements:

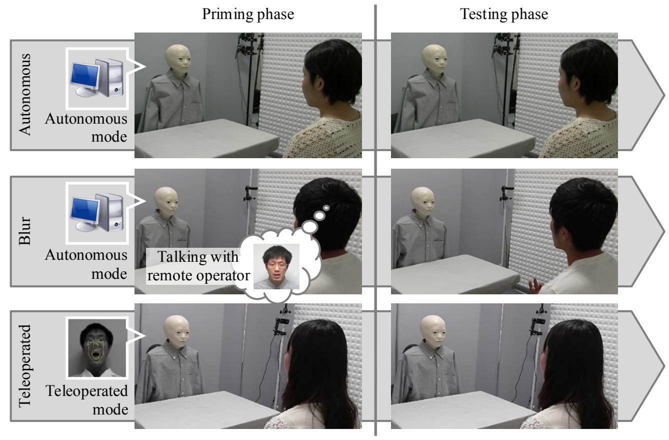

Fig. 4. Conditions of Experiments 2 and 3.

- In the first conversation (for the testing phase, "In the second conversation"), the robot was operated by the teleoperated mode.
- In the first conversation, the robot was operated by the autonomous mode.

In Experiment 1, since some subjects explained both why they felt that the robot automatically replied and why they felt the remote operator listened, we prepared two statements to separately evaluate these feelings and rated them on a 7-point Likert scale. We calculated the pseudo presence by subtracting the scores of the first statement from the second:

- I felt that the robot was automatically giving backchannel responses.
- I felt the robot was transmitting my remote partner's backchannel responses.

In this experiment, we also examined whether a subject felt the remote operator's pseudo presence by observing each subject's social response [12], e.g., whether subjects replied to the robot's greeting. At the end of the conversation in each phase, the subjects received a greeting from the operator: "Thank you for the conversation." If they felt that the remote operator had been listening, they might reply to the greeting; if they did not feel that way, they might ignore it.

E. Results

According to the manipulation check, we confirmed that all the subjects believed our instruction concerning which mode was used. Experiment 2's result is shown in Fig. 5, where each box represents the mean scores of the pseudo presence, and each bar represents the standard error of the mean. The figure compares the autonomous, blur, and teleoperated conditions by a one-way between-subjects ANOVA, followed by Bonferroni correction. We found a significant difference between these conditions ($F(2, 27)=4.881$, $p<.05$). Multiple comparisons showed that the blur condition was significantly higher than the autonomous condition ($p<.05$). This meant that priming the subjects' beliefs that they had talked to a remote operator produced pseudo presence. This result supports hypothesis 1 (described in Section V.E) and indicates that Experiment 1's result was caused by the priming regardless of the number of conversations. The differences between the teleoperated and autonomous conditions and the teleoperated and blur

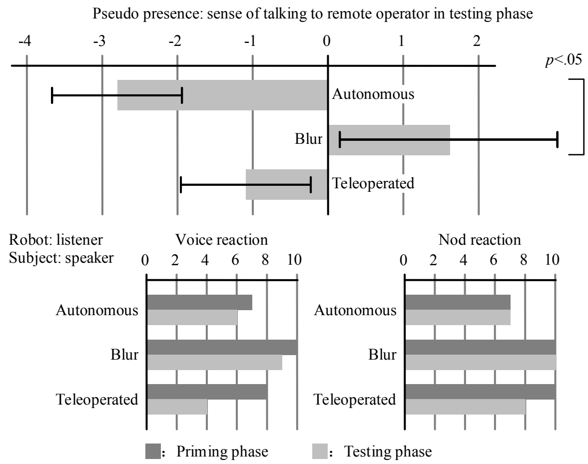

Fig. 6. Experiment 2 result.

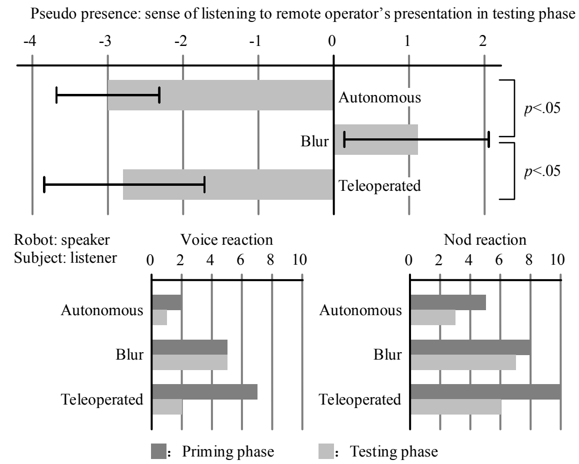

Fig. 5. Experiment 3 result.

conditions were not significant; hypothesis 2 (described in Section V.E) was not supported.

The result of observing the subject responses to the robot's greeting is also shown in Fig. 5. Most replied to the greeting by nodding and saying, "You're welcome"; several just nodded or just saying it. We counted these responses separately. In the teleoperated condition, the number of subjects who replied decreased between the phases. In the autonomous condition, the number of subjects who replied was less in both phases. These results indicate that the subjects tended to ignore the greeting from the autonomous robot, as we expected. On the other hand, in the blur condition, almost all the subjects replied, and the number did not decrease even after changing to the testing phase. Perhaps only the blur condition retained a higher presence through each phase. These tendencies greatly support our questionnaire results.

VII. Experiment 3

In Experiments 1 and 2, the subject was the speaker, and the robot was the listener. In Experiment 3, we examined hypotheses 1 and 2 in a task that reversed the subject and robot roles. This experiment confirmed whether priming can produce pseudo presence even in a less interactive conversation in which a user cannot get responses from the autonomous robot.

A. Conditions

The conditions were the same as in Experiment 2 (Fig. 4). In the autonomous and blur conditions, the autonomous mode only used the lip-sync system because it did not need to reply to the subject's speech. In the priming phase of the teleoperated condition, the remote operator used the teleoperated mode and asked the subject some questions to create an impression that the remote operator is actually talking. At that time, the remote operator simply replied "I see" to the subject's answer.

B. Subjects

Thirty undergraduate students participated in Experiment 3. No subjects from Experiments 1 and 2 participated in it. Ten subjects (five females and five males) participated in the autonomous, blur, and teleoperated conditions.

C. Task

In the autonomous and blur conditions, the robot presented pre-recorded speech to the subjects about a device's problem to which they only listened. In the teleoperated condition, the remote operator made a presentation and asked the subject three questions, e.g., "Have you ever watched a 3D movie?" The topics in the priming and testing phases were 3D TVs and smartphones. The speeches lasted about 1.5 minutes.

D. Measures

After the testing phase, the subjects answered identical manipulation check questions (described in Section VI.D) to confirm that they correctly understood our instructions.

The following two statements, which estimated the pseudo presence, were rated on a 7-point Likert scale. We calculated the pseudo presence by subtracting the scores of the first statement from the second:

- I felt that the robot was automatically talking.
- I felt the robot was transmitting the remote partner's talking behavior by teleoperation.

As with Experiment 2 (Section VI.D), we observed whether the subjects reply to the robot's greeting.

E. Results

Based on the manipulation check, we confirmed that all the subjects believed our instruction about which mode was used. Experiment 3's result is shown in Fig. 6, where each box represents the mean scores of the pseudo presence, and each bar represents the standard error of the mean. The figure compares the autonomous, blur, and teleoperated conditions by a one-way between-subjects ANOVA, followed by Bonferroni correction. We found a significant difference between these conditions ($F(2, 27)=5.806$, $p<.01$). Multiple comparisons showed that the blur condition was significantly higher than the autonomous condition ($p<.05$). This meant that priming the subjects' beliefs that they had listened to a remote operator's speech produced pseudo presence. Hypothesis 1 (described in Section V.E) was supported, even for conversations in which the robot unilaterally spoke to the subject. Additionally, the blur condition was significantly higher than the teleoperated condition (p<.05), and the difference between the teleoperated and autonomous conditions was not significant. The experience in which the remote operator was actually talking did not produce pseudo presence. This result was counter to hypothesis 2 (described in Section V.E).

The result of observing the subjects' responses to the robot's greeting is also shown in Fig. 6. The tendencies of the subject responses support the questionnaire results as well as Experiment 2 (Fig. 5). Overall, the number of responses in Experiment 3 is lower than in Experiment 2. This might be caused by the differences in the tasks. In Experiment 3, since the subjects were only listening to the remote operator's presentation, they did not need to reply except in the priming phase of the teleoperated condition. Because of less interaction, the subjects had difficulty feeling the operator's presence, and the number decreased. In contrast, in the teleoperated condition's priming phase of both Experiments 2 and 3, almost all of the subjects replied regardless of the task. Having more interaction with the operator seemed to increase his presence. In the next section, we discuss why the blur condition most effectively produced pseudo presence and why the teleoperated condition was ineffective.

VIII. Discussion

Our experiment results showed that users who believed that they had talked with a remote operator produced pseudo presence when they interacted with an autonomous robot even though they knew that a robot autonomously acted. Nevertheless, against our prediction, the experience in which users talked with an autonomous robot that behaved under the guise of a remote operator (blur condition) was more effective than the experience in which users actually talked with a remote operator (teleoperated condition). Open-ended responses suggest that the interaction gap between the priming and testing phases decreased the pseudo presence. In both Experiments 2 and 3, half of the ten subjects in the teleoperated condition commented that the degree of interaction, i.e., variety of talking behaviors and responses/questioning, decreased after changing to the testing phase.

In Experiment 2's result (Fig. 5), the pseudo presence of the teleoperated condition exceeded the autonomous condition but was lower than the blur condition, although the differences are not significant. This could mean that the positive effect of priming by the experience of talking with the remote operator was decreased by the negative effect of decreasing the degree of interaction. Four of ten subjects in the teleoperated condition mentioned that the backchannel timing in the testing phase was comparable to the priming phase, and so the appropriate backchannel responses produced by the autonomous system might have reduced the interaction gap. On the other hand, in Experiment 3's result (Fig. 6), the mean score of the teleoperated condition resembles the autonomous condition and is significantly lower than the blur condition. In the teleoperated condition, the interaction gap increased, since the subjects were only listening to the robot's presentation in the testing phase in contrast to the priming phase in which they answered questions from the remote operator. This large gap might have completely offset the positive effect of priming. In the observational data analysis result (Fig. 5 and Fig. 6), this gap also appeared as fewer subjects who replied to the remote operator's greeting. This tendency, which prominently appeared in Experiment 3, also supports the above discussion.

In the blur condition, most subjects replied to the robot's greeting in the priming phase; we observed no decrease in the testing phase. They could not recognize the interaction gap, since they all believed that they were talking with the remote operator in the priming phase, although the robot was in the autonomous mode. Concerning why they felt pseudo presence, five and six of the ten subjects (Experiments 2 and 3) in the blur condition commented that the autonomous mode behaved as naturally as the teleoperated mode. This subjective response suggests why blurring the interaction gap most effectively produces pseudo presence. The subject's assumption that the autonomous robot was behaving as naturally as the teleoperated robot evoked the remote operator's presence in the subject that was felt in the priming phase.

IX. Limitations and Future Work

Experiment 1 suggested that physically embodied motions facilitate reminding users of the operator's presence. However, it remains unclear whether the physical embodiment or the body motion produced pseudo presence since these factors were confounded in the experiment. Further investigation is needed to examine which factor contributes to producing pseudo presence.

Future work will examine how long the effect of priming would continue. In our experiments, the subjects interacted with the robot for about two minutes due to their limited knowledge of the topics. In our pre-experiment in which we tested the blur condition, one subject continued to talk to the robot that autonomously nodded for about thirteen minutes in the testing phase. After the experiment, this subject mentioned that he felt as if the remote operator was listening even in the testing phase. Although this case is special, it suggests that pseudo presence continues as long as an interaction continues.

The subjects in our experiments had never experienced talking with a teleoperated robot before participating our experiments. If the subjects had prior experience with teleoperated robots, results may have changed. Further exploration is needed to investigate how users' prior experience with a robot influences pseudo presence with a different robot.

In the testing phase, the robot used a simpler method to generate talking behaviors (Section IV.A). Technologies that generate more natural and various talking behaviors might fill the interaction gap. If interaction with an autonomous robot utilizes such technologies, it can approach the interaction level of a remote operator. We expect that such technologies will enable the experience of talking with a remote operator (teleoperated condition) to produce pseudo presence without decreasing the degree of interaction. This hypothesis can be experimentally tested by the Wizard of Oz method. A subject talks with a remote operator in both the priming and testing phases, but the subject is told that the robot will be changed to the autonomous mode before the testing phase. Future work will conduct this experiment.

X. Conclusion

This study proposed the method that produces the feeling of talking with a remote operator when the user is actually interacting with an autonomous robot. We conducted experiments based on the social telepresence test that evaluates whether an autonomous robot produces a remote operator's presence. From our experiments, we found that presenting a physically embodied motion and the user's belief that he/she had talking with a remote operator are factors for passing the social telepresence test. In fact, people decide the presence or absence of a remote operator based on their prior experience, and physically embodied talking behavior might remind them of the operator's presence. We also found that the interaction gap between prior and subsequent experiences reduces the chance of passing the social telepresence test. Prior experience in which a user talked with an autonomous robot under the guise of a remote operator blurred the gap and effectively produced the operator's presence even while interacting with an autonomous robot. Moreover, the improvement of technologies that produce natural and various talking behaviors will enable such autonomous robots to fill the interaction quality gap. We expect that this study will mutually facilitate telerobotics and intelligent robotics.

Acknowledgment

This study was supported by JST CREST "Studies on Cellphone-type Teleoperated Androids Transmitting Human Presence," JSPS Grants-in-Aid for Scientific Research No. 26280076 "Robot-Enhanced Displays for Social Telepresence," and KDDI Foundation Research Grant Program "Robotic Avatars for Human Cloud."

References

[1] Bainbridge, W.A., Hart, J., Kim, E.S. and Scassellati, B.: The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, Vol.3, No.1, pp.41-52, 2011.

[2] Baum, A., and Andersen, S. M.: Interpersonal roles in transference: Transient mood effects under the condition of significant-other resemblance. *Social cognition*, Vol.17, No.2, pp.161-185, 1999.

[3] Cao, Y., Tien, W.C., Faloutsos, P. and Pighin, F.: Expressive Speech-Driven Facial Animation. *ACM Transactions on Graphics*, Vol.24, No.4, pp.1283-1302, 2005.

[4] Finn, K.E., Sellen, A.J. and Wilbur, S.B.: Video-Mediated Communication. *Lawrence Erl-baum Associates*, 1997.

[5] Hashimoto, T., Kato, N. and Kobayashi, H.: Development of Educational System with the Android Robot SAYA and Evaluation. *International Journal of Advanced Robotic Systems*, Vol.8, No.3, pp.51-61, 2011.

[6] Ishi, C., Liu, C., Ishiguro, H. and Hagita, N.: Evaluation of formant-based lip motion generation in tele-operated humanoid robots. *Proc. IROS2012*, 2012.

[7] Lee, K.M., Jung, Y., Kim, J. and Kim, S.R.: Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. *International Journal of Human-Computer Studies*, Vol.64, No.10, pp.962-973, 2006.

[8] Lee, S.L., Lau, I.Y. M., Kiesler, S., and Chiu, C.Y.: Human mental models of humanoid robots. *Proc. ICRA2005*, pp. 2767-2772, 2005.

[9] Morita, T., Mase, K., Hirano, Y. and Kajita, S.: Reciprocal Attentive Communication in Remote Meeting with a Humanoid Robot. *Proc. ICMI2007*, pp.228-235, 2007.

[10] Nakanishi, H., Kato, K. and Ishiguro, H.: Zoom Cameras and Movable Displays Enhance Social Telepresence. *Proc. CHI 2011*, pp.63-72, 2011.

[11] Nakanishi, H., Tanaka, K. and Wada, Y.: Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence. *Proc. CHI2014*, pp.2143-2152, 2014.

[12] Nass, C. and Moon, Y.: Machines and mindlessness: Social responses to computers. *Journal of social issues*, Vol.56, No.1, pp.81-103, 2000.

[13] Noguchi, H. and Den, Y.: Prosody-Based Detection of the Context of Backchannel Responses. *Proc. ICSLP1998*, 1998.

[14] Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T. and Ishiguro, H.: Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.15, No.5, pp.592-597, 2011.

[15] Pereira, A., Prada, R. and Paiva, A.: Improving Social Presence in Human-Agent Interaction. *Proc. CHI2014*, pp.1449-1458, 2014.

[16] Rae, I., Takayama, L., & Mutlu, B.: In-body experiences: embodiment, control, and trust in robot-mediated communication. *Proc. CHI2014*, pp. 1921-1930, 2014.

[17] Ranatunga, I., Torres, N.A., Patterson, R.M., Bugnariu, N., Stevenson, M. and Popa, D.O.: RoDiCA: a Human-Robot Interaction System for Treatment of Childhood Autism Spectrum Disorders. *Proc. PETRA2012*, 2012.

[18] Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N.: Android as a Telecommuni-cation Medium with a Human-like Presence. *Proc. HRI2007*, pp.193-200, 2007.

[19] Takeuchi, M., Kitaoka, N. and Nakagawa, S.: Generation of Natural Response Timing Using Decision Tree Based on Prosodic and Linguistic Information. *Proc. Interspeech2003*, 2003.

[20] Tanaka, K., Nakanishi, H. and Ishiguro, H.: Physical Embodiment can Produce Robot Operator's Pseudo Presence. *Frontiers in ICT*, Vol.2, No.8, 2015.

[21] Tanaka, M., Ishii, A., Yamano, E., Ogikubo, H., Okazaki, M., Kamimura, K., Konishi, Y., Emoto, S. and Watanabe, Y.: Effect of a human-type communication robot on cognitive function in elderly women living alone. *Medical Science Monitor*, Vol.18, No.9, CR550-CR557, 2012.

[22] Truong, K.P. and Poppe, R and Heylen, D.: A rule-based backchannel prediction model using pitch and pause information. Proc. Interspeech2010, pp.26-30, 2010.

[23] Ward, N. and Tsukahara, W.: Prosodic Features which Cue Back-channel Responses in English and Japanese. *Journal of Pragmatics*, Vol.32, No.8, pp.1177-1207, 2000.

[24] Watanabe, T., Okubo, M., Nakashige, M. and Danbara, R.: InterActor: Speech-Driven Embodied Interactive Actor. *International Journal of Human-Computer Interaction*, Vol.17, No. 1, pp.43-60, 2010.