Enhancing Operator Awareness: Ambient Noise-Based Human Presence Sensing for Unobtrusive Notification in Cybernetic Avatar Interface

Atsushi Toyoda[1,2], Tomo Funayama[2], Kurima Sakai[2], Ryusuke Mikata[2], Takashi Minato[2,3], and Hiroshi Ishiguro[1,2]

***Abstract*— When an operator controls a Cybernetic Avatar (CA) while feeling as if it is another entity of themselves, she/he could feel that the CA's skills are directly under their control. In this case, the operator feels that her/his recognition capability is improved by using the CA, which leads them to perform more tasks with greater effectiveness and confidence in achieving the job. For this goal, this research designed a teleoperation interface of a receptionist CA which helps operators feel as if their recognition capabilities are enhanced. In the reception task, the operator can work on another task while no person is visiting the reception. Thus the operator needs to notice whether someone is coming to the reception or not. This paper proposes a novel method to make the operator realize a visitor has come by themselves. The proposed interface uses ambient background noise, known for its relevance to human concentration, in notification. The level of the noise decreases when the teleoperation system estimates the visitor comes. Experimental results show that diminishing ambient noise led participants to perceive themselves as having independently observed a visitor approaching the reception desk.**

I. INTRODUCTION

Working with CAs brings advantages such as operators being able to work remotely and engage in other tasks when manual teleoperation is not required. Additionally, CAs make it possible for operators to accomplish tasks that are impossible to achieve solely with their physical bodies. Furthermore, by incorporating the augmented ability, which makes the operator feel ownership of the CA's behavioral expressions and cognitive capabilities provided by sensors, the operator becomes more confident in their improved competencies and becomes less stressed when working with CA. A receptionist is a well suited job for CAs as it is not always required to interact with visitors, that is, the operator controls the CA only when a visitor arrives at the reception desk. Hence, the interface should notify the operator of the arrival of visitors even if the operator engages in other tasks. Given the study's objective of enhancing operators' job satisfaction through CA utilization, a notification interface is needed that does not impede operators' work and makes them feel that their ability to recognize visitor arrivals has been augmented.

However, a teleoperation system with a typical notification sound may be perceived as invasive, which the notification as told by the CA interface, which might disturb the operator's satisfaction. Therefore, this study explores the design of a notification interface that leads the operators to notice the arrival of visitors independently rather than being notified by the system.

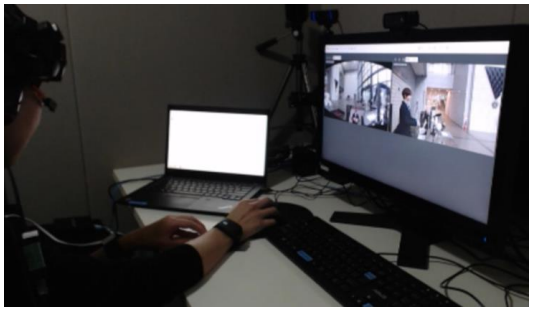

Fig. 1: Multi-task in action: A participant engaged in the typo-check task on the left laptop while monitoring visitor reception on the right display, facilitated by audio notifications through headphones.

II. METHOD

CA operators are typically engaged in other tasks and may not constantly observe the remote reception displayed on the interface screen. Thus, in lieu of displaying notifications on it, we designed the following three auditory notifications.

A. Ambient noise change (AnC)

Muranaka [1] explored a notification method that does not cause interference by altering the volume of ambient noise. It is known that ambient noise, background meaningless sound, and human concentration levels influence each other. According to Molloy et al.[2], when human concentration levels increase, ambient noise in the surroundings temporarily becomes inaudible. Additionally, moderate levels of ambient noise are known to enhance performance in creative tasks. Taking these into account, our interface regularly plays ambient noise and diminishes it when a visitor comes. This makes an operator feel that she/he notices the visitor since she/he concentrates on the visitor's approach.

B. Step change + Ambient noise (SC+An)

Next, designing an intuitive method for recognizing human presence is considered using footstep sound. Notification sounds often use sounds that evoke what is intended to be expressed to be intuitively understood for information transmission. Therefore, footsteps were chosen to represent people arriving at the reception. Additionally, considering the

[1]Graduation School of Engineering Science, Osaka University, Osaka, Japan
[2]Advanced Telecommunications Research Institute International, Kyoto, Japan
[3]Interactive Robot Research Team, Guardian Robot Project, RIKEN, Kyoto, Japan

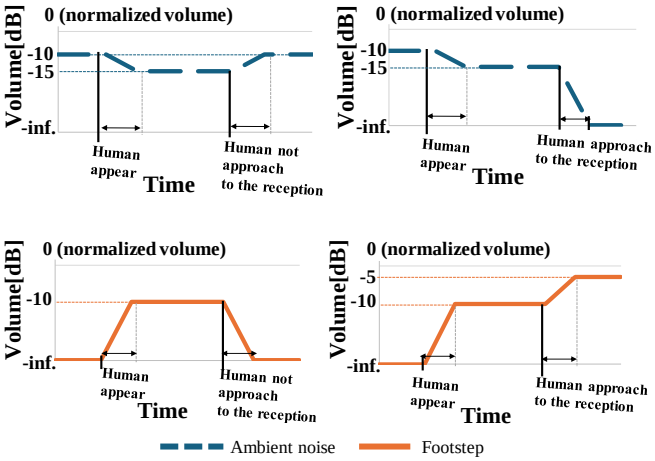

Fig. 2: The events and the corresponding volume changes in the notification of each sound. The upper and lower two represent the ambient noise and footstep sound respectively. The left side shows when humans appear in the reception area and go somewhere. The right side shows when humans appear in the reception hall and approach to the reception. The double arrows represent the duration time for sound to change, which is 0.5 seconds.

impact of ambient noise on task performance, ambient noise is continuously played.

C. Step change + Ambient noise change (SC+AnC)

As another intuitively understandable design, the signal-to-noise ratio (SNR) of human footsteps increases as individuals approach the reception area. Specifically, the ambient noise is reduced, and the volume of footsteps is increased.

In all three notification methods, the volume of sound changes as shown in Fig.2: when a person appears in the entrance space in front of the reception, the ambient noise decreases by one level, and the volume of footsteps increases by one level. When the operation system estimates the person will leave, the ambient noise returns to its original level (volume increases), and the footsteps disappear. Conversely, when it estimates the person will come to the reception, the ambient noise disappears, and the volume of footsteps becomes even louder.

III. EXPERIMENT

Subjects were asked to perform two tasks simultaneously during the experiment. In particular, they did the typo-check task while the CA greeted those walking past the reception. On the other hand, they did the reception task where they initiated the teleoperation by pressing a button when a visitor arrived at the reception. As soon as the button is pressed, the display blackouts, prompting to repeat the task. The participants performed the tasks under the three notification conditions (the order was counterbalanced). For each condition, they repeated the task 3 times, and lastly they answered a questionnaire.

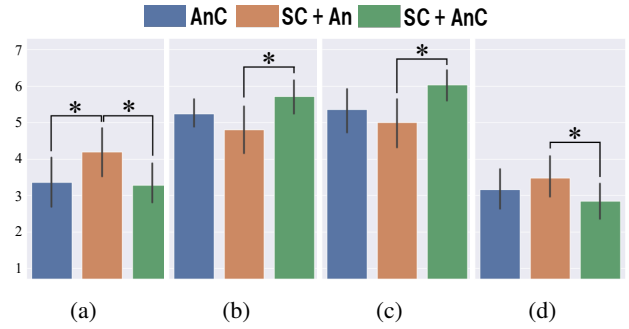

Fig. 3: The subjects' score of the following questionnaire. (a) Did you feel that the notification interfered with the task? (b) Did you initiate the teleoperation with no rush? (c) Were you able to sense the arraval of visitors? (d) Did you feel forced to initiate the teleoperation? (0: Strongly disagree - 7: Strongly agree, $^*p < .05$)

IV. RESULT

The Fig.3 shows the subjective evaluation results from the questionnaire that had significant differences between the conditions. ANOVA and following multiple comparison with Holm correction revealed significant differences between SC+AnC and AnC, and SC+An ($p < .05$) in question (a), between SC+AnC and SC+An ($p < .05$) in questions (b), (c), and (d).

V. DISCUSSION

We found that SC+AnC is the best design for our interface. The participants were more likely to sense human presence when the visitor came to the reception desk and did not feel forced to begin the reception task. In fact, the participants noticed the visitor came by the sound change but they had less feeling of compulsion to initiate the teleoperation. This design successfully led them think they independently switched their attention to the interface from the typo-check task. This might be because their subconscious mind to the human becomes clearer from background noise thanks to the cocktail party effect. In other words, the diminishing ambient sound makes the operator feel that they sensed the arrival of visitors independently.

ACKNOWLEDGMENT

This work was supported by JST Moonshot R&D Grant Number JPMJMS2011.

REFERENCES

[1] Y. Muranaka, M. Al-Sada and T. Nakajima, "Unobtrusive notification based on auditory changes in ambient music," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 2021, pp. 231-234

[2] K. Molloy, T. D. Griffiths, M. Chait, and N. Lavie, "Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses," Journal of Neuroscience, vol. 35, no. 49, pp. 16 046–16 054, 2015.