Dual Body: Method of Tele-Cooperative Avatar Robot with Passive Sensation Feedback to Reduce Latency Perception

Vibol Yem
Tokyo Metropolitan University
Hino, Tokyo, Japan
yem@vr.sd.tmu.ac.jp

Kentaro Yamaoka
Tokyo Metropolitan University
Hino, Tokyo, Japan
yamaoka@vr.sd.tmu.ac.jp

Gaku Sueta
Tokyo Metropolitan University
Hino, Tokyo, Japan
sueta@vr.sd.tmu.ac.jp

Yasushi Ikei
The University of Tokyo
Bunkyo, Tokyo, Japan
ikei@vr.u-tokyo.ac.jp

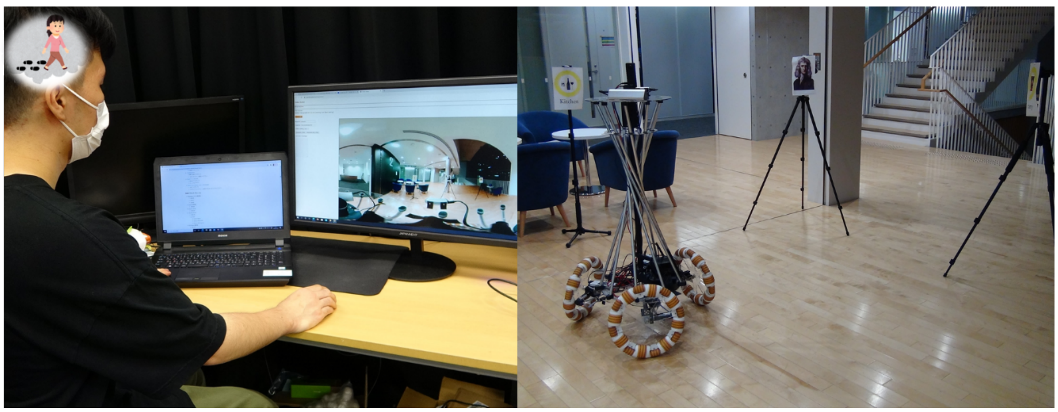

Figure 1: (Left) A user is sitting in front of a command browser window with speech recognition to control avatar robot actions (via voice commands) and to passively perceive sound and visual feedback from the robot. (Right) The avatar robot, with a 360° camera, a depth sensing camera, and image recognition algorithms.

ABSTRACT

Dual Body was developed to be a telexistence or telepresence system, one in which the user does not need to continuously operate an avatar robot but is still able to passively perceive feedback sensations when the robot performs actions. This system can recognize user speech commands, and the robot performs the task cooperatively. The system that we propose, in which passive sensation feedback and cooperation of the robot are used, highly reduces the perception of latency and the feeling of fatigue, which increases the quality of experience and task efficiency. In the demo experience, participants will be able to command the robot from individual rooms via a URL and RoomID, and they will perceive sound and visual feedback, such as images or landscapes of the campus of Tokyo Metropolitan University, from the robot as it travels.

SA '20 Emerging Technologies, December 01-09, 2020, Virtual Event, Republic of Korea

ACM ISBN 978-1-4503-8110-9/20/11.
https://doi.org/10.1145/3415255.3422893

CCS CONCEPTS

• **Human-centered computing → Virtual reality**.

KEYWORDS

Telecooperation, avatar robot, task recognition, virtual reality, artificial intelligence.

ACM Reference Format:
Vibol Yem, Kentaro Yamaoka, Gaku Sueta, and Yasushi Ikei. 2020. Dual Body: Method of Tele-Cooperative Avatar Robot with Passive Sensation Feedback to Reduce Latency Perception. In *SIGGRAPH Asia 2020 Emerging Technologies (SA '20 Emerging Technologies), December 01-09, 2020*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3415255.3422893

1 INTRODUCTION

Telexistence or *telepresence* is a kind of virtual reality technology that enables the user to immersively and remotely perform tasks [Tachi 2016]. The current issues surrounding the coronavirus disease 2019 (COVID-19) pandemic have made this type of technology very important, and the development of a system with high usability is urgently required.

Telepresence technology has been studied for decades to improve usability [Higuchi and Rekimoto 2012]. However, several limitations have not allowed systems that are developed to easily be put into practice. We believe that continuous operation of the robot by the user is the main issue that highly reduces system usability. Such a method of operation causes not only the feeling of fatigue but also the apparent perception of latency, since the user is required to wait the robot's reaction after each command input. Such a perception continually occurs when operating the robot to perform a task.

To increase the usability of the system, we propose combining passive sensation feedback and task recognition algorithms in the robot. This method requires command input at the beginning of a task, and subsequently, the robot cooperatively works to fulfill the commands. Thus, this method can reduce the perception of latency as well as the feeling of fatigue. Passive feedback creates the sensation of immersion while tasks are being performed. Figure 1 shows our system, named Dual Body, and Figure 2 shows the algorithm schematic for our system. In this demo experience, we use only visual and auditory feedback because they can easily be accessed by participants.

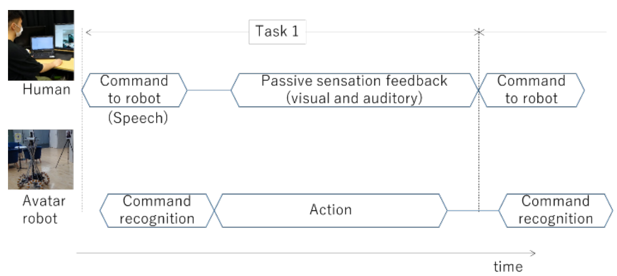

Figure 2: Algorithm schematics of the proposed method.

2 SYSTEM

The system is composed of two parts: the user interface and the avatar robot. We used WebRTC with Momo and Ayame software development kits to transmit sound and video between the two each side.

2.1 User Interface

A computer with microphone, speaker and web-camera are required for the user interface. We use a speech-to-text API (Google Inc) for speech recognition; therefore, Google Chrome is recommended as the interface because the API is suitable for this browser. When connecting to our robot via the URL[1], a unique RoomID is required to differentiate users from one another. Voice communication, auditory feedback, and live video from the robot are played via the browser. This is a browser-based system; therefore, users will not need to install any additional software.

2.2 Avatar Robot

We developed an avatar robot that has three omniwheels to enable the robot to move in three degrees of freedom: forward/backward, right/left, and clockwise/counterclockwise. These directions are sufficient to represent any direction of movement on a surface. Jetson Xavier AGX module (NVIDIA Corp.) was used for the main control board of the robot. One 360° camera with HD resolution (THETA R, Ricoh Co. Ltd.) was attached on the top of the robot to remotely capture live stereoscopic 3D images. A depth camera (RealSense D435i, Intel Corp.) was used to measure the distances between the robot and obstacles. RGB images from this camera were used for face and image detection. A speaker was placed on the top of the robot for interactive communications.

[1] https://yemvibol.github.io/dualbody/

2.3 Algorithm and Passive Sensation Feedback

When the user gives a command to the robot, the transmission latency is not large because voice data are typically small and are transmitted quickly. In contrast, latencies usually occur for visual data captured by a high-resolution 360° camera. In this system, we created footstep sounds that are sent synchronously with live video of the robot's motion, and the user only gives a command to initiate a task. Therefore, the user does not perceive any latency during robot actions.

Our previous study confirmed that multiple types of feedback sensations can increase the sense of agency even during passive motion [Yem et al. 2020]; therefore, we recommended including tactile and motion feedback to increase the quality of the user experience. For example, providing footstep vibrations to the foot would enhance walking-like sensations. However, these sensation feedback devices are difficult for most people to access at home. Therefore, we decided to use only auditory and visual feedback for our current system.

3 DEMO EXPERIENCE DURING VIRTUAL CONFERENCE

We will provide a RoomID to those who would like to participate in our demo experience. Participants, sitting in front of their own computers or tablets, will open a Chrome browser and enter the URL to connect to our avatar robot (when it is available) simply by clicking a button. Several images, such as images of Newton's face, a library, and a kitchen will be placed around the robot and participants will be able to command the robot to find them. Participants can also direct the robot to move to see the landscape of the Tokyo Metropolitan University campus. Participants will be able to simultaneously hear the sound of walking footsteps and remotely captured video while the robot is moving.

ACKNOWLEDGMENTS

This work was partially supported by the MIC/SCOPE #191603003, JSPS KAKENHI Grant Number 18H04118 and 18H03283.

REFERENCES

K. Higuchi and J. Rekimoto. 2012. Flying Head: Head-Synchronized Unmanned Aerial Vehicle Control for Flying Telepresence. *ACM SIGGRAPH Asia Emerging Technologies* (2012), 1–2. https://doi.org/10.1145/2407707.2407719

S. Tachi. 2016. Telexistence: Enabling Humans to Be Virtually Ubiquitous. *IEEE Computer Graphics and Applications* 36 (2016), 8–14. https://doi.org/10.1109/MCG.2016.6.

V. Yem, F. Miyashita, T. Amemiya, M. Kitazaki, and Y. Ikei. 2020. Sense of Agency in Drum Trainer with Multiple Sensation Feedback. *ACM SIGGRAPH 2020 Posters* (2020), 1–2. https://doi.org/10.1145/3388770.3407407