2nd Workshop Toward Robot Avatars, IEEE International Conference on Robotics and Automation (ICRA), London, UK, 2023.

I-Botics avatar system: Towards robotic embodiment

Jeanine van Bruggen[1], Christiaan Brekelmans[1], Robin Lieftink[1,2], Douwe Dresscher,[2] Jan van Erp[1,3]

***Abstract*— In this paper we present a fully integrated teleoperation system that enables the transport of the functional and social self. We strive to achieve embodiment, through maximum transparancy of all sensory modalities. This includes transporting vision, audio, eye-gaze and facial expressions; movement of the head, torso, arms, hands and fingers; force feedback on arms and fingers; tactile feedback, wind and heat. The system placed 5th in the international ANA Avatar Xprize competition, in which an unexperienced operator was able to do a variety of social and physical tasks. The system showed a high level of embodiment, however, contributions of different modalities on task performance should be further evaluated.**

Keywords—telepresence, teleoperation, embodiment, avatar, robotics

I. INTRODUCTION

Teleoperation is a rapidly growing topic in the field of robotics. The main goal of the first-generation teleoperation systems was to transport one's functional self to a remote location. One's functional self is related to spatial presence, task performance, and interacting with objects in the environment. However, there are also teleoperation applications that include an important social aspect. Therefore, we are also interested in transporting one's social self and the feeling of social presence, i.e. to interact with and feel connected to people in remote environments. [1]. In this paper we present a fully integrated teleoperation system that enables transporting the functional and social self. The system was evaluated during the ANA Avatar Xprize competition [2]. In section II, we present our vision on telepresence, in section III we give an overview of our system. Finally, in section IV and V we provide results and discussions.

II. OUR VISION

Our vision is that a universal telepresence system must support embodiment and social presence. To achieve embodiment, for full transparancy of all sensory modalities is required. The feeling that you are the telepresence robot implies that there is no longer a mediating device. This ultimate transparency may lead to lower cognitive workload and faster learning [3]. To achieve social presence, bidirectional exchange of social cues is required: people in the remote environment should feel socially connected to the remote user and vice versa. Considerations on sensory cues needed for a sense of embodiment to the operator[4], and social cues needed for a sense of connection with the recipient on the robot side[5] are more elaborately described by van Erp et al. [1]. These considerations led to the system in figure 1.

Figure 1: On the left: the operator control pod. On the right, the robotic avatar.

III. SYSTEM OVERVIEW

The components to achieve embodiment and social presence are described in Table 1. Three main elements can be identified. First, the robot morphology is humanlike, and the operator movements are mirrored through the head, torso, arms, hands, fingers, and crouching of the robot. This gives the ability to move intuitively, to perform tasks, and to convey social cues through body posture and movements. Secondly, to improve immersion, all senses are conveyed to the operator through the robot. Tactile and force feedback is given through haptic devices, depth vision by stereo cameras, directional audio for communication. Heat is provided from the sides and back, increasing situational awareness by sensing human presence by its body heat. Airflow and smell add to the illusion of embodiment. Thirdly, a facial representation of the operator is displayed on the robot enabling eye-contact with recipients and showing emotions through facial expressions.

IV. RESULTS

In November 2022, the system placed 5th in the finals of the ANA Avatar Xprize competition [2]. A video showing the integrated system and its immersive features is available at https://youtu.be/h6nvlPRlVzY. Through the integration of the described features, we achieved that an untrained operator can perform tasks such as a solving a children's puzzle or using an electric drill after only 30min of training, and can experience social interactions through the avatar system. This is also reflected in the ratings of the operator judges. Our system received the full score (5/5) on the following ratings from both operators:

- How much of the system were you able to control?
- I felt confident using the system.
- I found the system not very cumbersome to use
- Training was not cumbersome for this system.

---

1. Intelligent Imaging, Intelligent Autonomous Systems, Human Machine Teaming (TNO), email: name.surname@tno.nl
2. Robotics and Mechatronics group, University of Twente
3. Human Media Interaction, University of Twente.
:

A second to highest score (4/5) was given on the following ratings:

- How much of the system were you able to control?
- How quickly did you adjust to the remote environment experience?
- How proficient in moving and interacting with the remote environment did you feel at the end of the experience?
- I found the system not unnecessarily complex.
- I imagine most people would learn to use this system very quickly.

This shows that our system has a good sense of embodiment, with a high score in ease of use and ease of learning. A low score (1/5 or 2/5) was received by at least one of both operators on amongst others the following questions:

- How closely were you able to examine objects?
- Were you able to survey or search the environment using vision?
- Did you experience delays between your actions and expected outcomes?
- How much did the visual display quality interfere or distract you from performing test course tasks?
- How well could you localize sounds?

Lower scores were mainly received on audio and visual feedback, presumably due to the fact that we reduced audio and visual quality to meet the available bandwidth.

V. DISCUSSION

This paper describes a fully integrated robotic avatar system to transport both the functional and social self of an operator. In our system many modalities have been implemented. Although separate modalities are known to improve either the feeling of immersion or task performance [7]–[10], an extensive evaluation is needed on how the combination of all these modalities contribute to performance on a variety of tasks. Even though high-level embodiment is achieved [1], the performance is probably lower than that of an on-site operator. More work is required to develop robotic systems that match human dexterity and strength, especially of the hands. As technology improves, we foresee the applicability of robotic avatars in many domains where there is demand for expertise or social connection at a distance. Telepresence technology can affect many features of our lives, including economic, social, and cultural aspects, and it may change social systems, norms, and regulations. We believe that researchers and technology developers should actively seek the debate on ethical, legal and societal issues of telepresence technology.[1]

Table 1: System modality and implementation overview.

Modality	**Operator implementation**	**Robot implementation**
Vision	HMD (HTC Vive Pro Eye, HTC Corporation, Taiwan). The camera inputs displayed in VR on half domes matching the shape of the lens to compensate for distortion.	Stereo 180° FoV cameras placed in the head of robot. Scalable resolution for transferred images depending on available bandwidth. Resolution used in the ANA Avatar XPRIZE finals: 640x480 (per camera) of max. 1600x1200.
Bi-directional audio	HTC Vive Pro Eye microphone and headset.	Linear microphone array with echo cancellation and speaker system (XMOS Vocal fusion XVF-3500-L33).
Eye contact & facial expressions	HMD integrated Eye tracking and HTC Pro 2 Face Tracker.	Ready player me animated operator avatar (https://readyplayer.me/). Shown on the two face displays of the robot. An avatar matching the operator face can be generated using a picture.
Thermal sensing	Infrared heaters integrated in the walls of the control pod, on the back and side of the operator.	Two 90° FoV IR thermal sensors (Omron D6T-32L-01A) integrated in the lower back of the robot. Allowing for 180° temperature sensing on the back side of the robot.
Airflow	Two fans placed on the front left and front right side of the operator.	Robot wheel velocity input commands are used to control the fan speed.
Tactile feedback	DK2 Gloves (Haptx, USA), generating tactile feedback on the fingertips and palms of the operator using compressed air.	Ability hand (Psyonic, San Diego, USA) fingertip touch sensors, combined with custom thin-film pressure sensors (Holst centre, Netherlands) placed on the palms of the hands.
Finger force feedback	HaptX DK2 finger force feedback. Allowing proportional breaking force on each finger using a wire.	Psyonic Ability hand fingertip touch sensors.
Finger control	HaptX DK2 fingertip position measurement. Magnetic system giving the 6D position of each fingertip w.r.t. the palm.	The Psyonic Ability hand gives 1DoF control of each finger and 2DoF control of the thumb. The measured operator finger positions are mapped to the robot hand though a calibration procedure.
Arm control & force feedback	Virtuose 6D (Haption, France) devices are mechanically connected to the Haptx gloves and provide arm force feedback and position measurement.	6D Force sensors (SensOne, Bota Systems, Switzerland) are placed between the hands and arms of the robot to provide force feedback to the operator. The robot arm actuators are used for control and measurement. The teleoperation control loop is kept stable over the network using a passivity layer approach[6].
Robot head control	The HMD pitch angle measurement.	Pitch actuation of the robot head w.r.t. the torso.
Robot posture control	HTC puck. Strapped to the torso of the operator measures the rotations of the torso.	Hip and torso actuators are used to control the orientation of the torso.
Robot Crouching control	Crouching control is achieved through the position of the operator hands. If the hands are lowered below a threshold, the robot hands and hip will follow the vertical position of the operator hands.	Ankle, knee, and hip actuators are used to control the hip height of the robot.
Locomotion	A custom foot pedal is used, measuring the weight distribution (towards toes, or towards heels) and rotation of the operator feet.	The robot uses a wheeled differential drive for locomotion.

REFERENCES

[1] J. B. F. Van Erp *et al.*, “What Comes After Telepresence? Embodiment, Social Presence and Transporting One’s Functional and Social Self,” *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2022-Octob, pp. 2067–2072, 2022, doi: 10.1109/SMC53654.2022.9945544.

[2] XPRIZE, “ANA Avatar Xprize.” https://www.xprize.org/prizes/avatar (accessed Mar. 31, 2023).

[3] A. Toet, I. A. Kuling, B. N. Krom, and J. B. F. van Erp, “Toward Enhanced Teleoperation Through Embodiment,” *Front. Robot. AI*, vol. 7, no. February, pp. 1–22, 2020, doi: 10.3389/frobt.2020.00014.

[4] S. Falcone, A.-M. Brouwer, I. Cocu, K. Gijsbertse, D. Heylen, and J. van Erp, “The relative contribution of five key perceptual cues and their interaction to the sense of embodiment.,” *Technol. Mind, Behav.*, vol. 3, no. 1, pp. 1–29, 2022, doi: 10.1037/tmb0000068.

[5] N. N. Sharan, A. Toet, T. Mioch, O. Niamut, and J. B. F. van Erp, *The Relative Importance of Social Cues in Immersive Mediated Communication*, vol. 319, no. 1. Springer International Publishing, 2022.

[6] M. Franken, S. Stramigioli, S. Misra, C. Secchi, and A. MacChelli, “Bilateral telemanipulation with time delays: A two-layer approach combining passivity and transparency,” *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 741–756, 2011, doi: 10.1109/TRO.2011.2142430.

[7] B. De Jesus Jr. *et al.*, “Quantifying Multisensory Immersive Experiences using Wearables: Is (Stimulating) More (Senses) Always Merrier?,” pp. 1–8, 2022, doi: 10.5753/sensoryx.2022.20001.

[8] G. Goncalves, M. Melo, J. Vasconcelos-Raposo, and M. Bessa, “Impact of Different Sensory Stimuli on Presence in Credible Virtual Environments,” *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 11, pp. 3231–3240, 2020, doi: 10.1109/TVCG.2019.2926978.

[9] C. Flavián, S. Ibáñez-Sánchez, and C. Orús, “The influence of scent on virtual reality experiences: The role of aroma-content congruence,” *J. Bus. Res.*, vol. 123, no. September 2020, pp. 289–301, 2021, doi: 10.1016/j.jbusres.2020.09.036.

[10] N. Ranasinghe *et al.*, “A demonstration of season traveller: Multisensory narration for enhancing the virtual reality experience,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2018-April, pp. 1–13, 2018, doi: 10.1145/3170427.3186513.