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

Enhancement of Team JANUS' cybernetic avatar system for exploration and skill transfer

R. Cisneros[1,*], A. Dallard[1,2], M. Benallegue[1], K. Kaneko[1], H. Kaminaga[1], P. Gergondet[1], A. Tanguy[2], C. Fournier[2], R.P. Singh[1], Y. Chen[1,3], S. Chefchaouni[1], G. Lorthioir[1], Y. Osawa[4], M. Tsuru[1], L. Sun[1], M. Morisawa[1], G. Caron[1], M. Ono[5], K. Shirasaka[5], S. Wada[5], H. Wada[5], F. Kanehiro[1] and A. Kheddar[1,2]

***Abstract*— In this paper, we present the approach taken by Team JANUS to enhance our avatar system to use it not only for connectivity, but also for exploration and skill transfer: the three core domains of the ANA Avatar XPRIZE - Finals global competition. We briefly describe the improvements to the avatar system from the solution used at the Semifinals, which allowed us to qualify as finalists at the competition.**

I. Introduction

In [1], we described the cybernetic avatar system that we developed for the Semifinals of ANA Avatar XPRIZE and for which the main focus was the connectivity: An adult-sized humanoid avatar to embody [2] high fidelity human telepresence. As such, this avatar system was envisioned to mitigate the impact of travel restrictions on our daily life.

For the Finals the focus was not only the connectivity, but also the ability of the operator to explore a remote location and transfer own skills to the avatar wherever any specific know-how is needed. To test these capabilities, XPRIZE designed a course that required mobility over 30 m, as well as dexterous manipulation of objects and tools. Additionally, it had to be completed in 25 min or less. Meeting these requirements necessitated an untethered avatar with the ability to navigate its environment, perform precise grasping motions, and utilize an advanced haptic system.

Based on these requirements, we enhanced our avatar system (see Fig. 1) and qualified as finalists of the competition.

II. Avatar robot

For that semifinals, our avatar robot (HRP-4CR) had been enhanced with an additional dof at each arm to have 7 dof arms, force/torque sensors at the wrists and underactuated dexterous hands (developed by our partner Double R&D)[1]. However, it was still externally powered and its arm strength was just enough to lift an object of 1.3 kg.

The Finals imposed additional challenges as the avatar robot had to be untethered and capable of manipulating objects / operate tools up to 3 kg. This required us to design a new battery box for the robot, to improve the strength of the arms and chest joints, and to redesign the hands.

[1] CNRS-AIST JRL (Joint Robotics Laboratory), IRL, AIST, Japan.
[2] CNRS-UM, LIRMM, France.
[3] SIGMA, University of Tsukuba, Japan.
[4] Industrial Cyber-Physical Systems Research Center, AIST, Japan.
[5] Double R&D, Japan.
[*] Corresponding author E-mail: rafael.cisneros@aist.go.jp

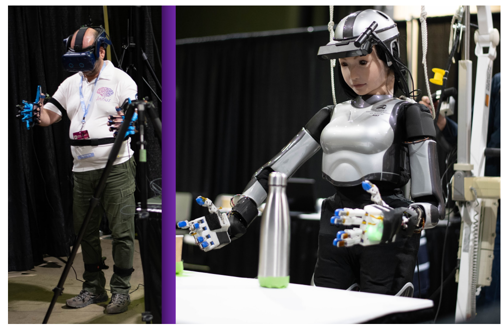

Fig. 1. HRP-4CR synchronized with the operator

The battery box was designed by considering the energy required for the Finals and the available space. We took into consideration some conditions: standing (302 W), performing manipulation (426 W), and walking (821 W), as well as an estimated time for each during the course of the competition. Based on this analysis, we needed at least 178.4 Wh (for 25 min). We developed a battery system with a minimum of 180 Wh using LiFe battery cells, and arranged in two boxes placed at the hips of the robot. The strength of the arms and chest joints was improved by increasing the rated torque, while keeping the appearance of the robot unchanged.

Finally, the dexterous hands that had been developed by Double R&D (D-Hands), although being capable of precision grasping, lacked the strength and repeatability required to manipulate the drill and turn it on. For this reason, they were redesigned to achieve this task (see Figs. 2, 3 and 4):

- Cam mechanism is used to sequentially actuate the abduction/adduction joint of the thumb (TM2), and then the flexion/extension joints of all fingers (TM1, MCPs) to realize non-prehensile grasping with only 1 DoF.
- Sheet metal links were replaced by machined links to improve strength against axial loads.
- The repeatability was improved by introducing ball bearings at the joints to reduce friction.

III. Haptics

At the Semifinals we were just able to display the weight of a manipulated object to the operator and give a proportional vibration through a pad on the hand-held controllers. However, the tasks at the Finals required a more elaborated haptics system capable of also transmitting the sensation of pressure and texture.

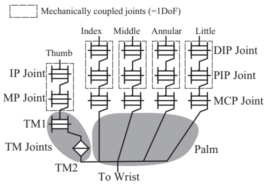

Fig. 2. Kinematic diagram of D-Hand v3

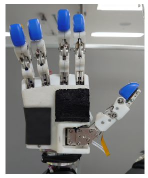

Fig. 3. View of D-Hand v3

Fig. 4. D-Hand v3 operating drill

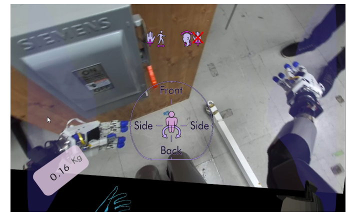

Fig. 5. Interface for walking, motion and field of view retargeting.

On the operator side we decided to use haptic gloves implemented as an exoskeleton for the hands (SenseGlove DK1), a solution also adopted by NimbRo [3]. These gloves can emulate the sensation of grasping (or pressure) to the operator by stopping the motion of the fingers. They also provide vibrotactile feedback on each fingertip to emulate the sensation of texture. However, contrary to [3], we are not using an exoskeleton for the arms as we prioritized simplicity. Because of that, we cannot provide the sensation of weight, but we can still display it overlaid on each hand.

To provide the haptics feedback, we attached miniature 6-dof F/T sensors to 3 of the fingertips (thumb, index, middle) and 1-dof FSRs to the remaining fingertips. The information about the force is acquired by a haptics controller PCB designed by us, and transmitted to the operator.

At the operator side the feedback is used in 3 ways:

- *Visually*. Force vectors are overlaid on each fingertip in the view of the operator.
- *As a grasping force*. The measured normal force is used to apply a pullback force on the fingers.
- *As vibration*. A high frequency component uses the measured forces to capture the roughness of the object and a low frequency one uses the variation of the measured normal forces and the speed of the hand to apprehend it's geometry.

IV. Operator interface

As our interface became button-less, we had to find another way to command the avatar. So, we used voice, gaze and head motion to achieve that purpose. See Fig. 5.

A. *Walking interface*

Given that we are using a humanoid avatar, we developed a walking interface that is triggered by stepping in place to improve the immersion. Walking is enabled by a voice command which displays a menu allowing to choose (also by voice) the appropriate stepping modality (forward, backward, or side steps). For turning, the operator must turn his head into the desired direction. A line is also drawn to show where the robot is expected to walk. Finally, the closer the head looks at the feet the shortest the steps are. Team iCub [4] followed a similar strategy, yet they used the first relative foot print as an indicator of the stepping modality.

B. *Motion retargeting*

To enable/disable the retargeting of the hands, we use shaded areas at both sides of the operator view (one per hand) which is activated by maintaining the gaze for 3 sec.

C. *Field of view retargeting*

As the motion of the head of the operator is different to the one of the robot (due to tracking speed and joint limits), we synchronized the field of view shown to the operator with the head of the robot for it to be spatially consistent [5]. This strategy improves the immersion and helps to understand the environment (for navigation). Additionally, a 3D model of the robot is visible outside of the field of view.

V. Safety

A. *Hierarchy inequality admittance*

Our framework cannot transmit a proper force feedback to the hands of the operator (only visual). Moreover, applying a significant force on the environment could lead to a loss of balance. We therefore needed a control scheme that limits the maximum force one can apply on the environment and whilst not altering or compromising the user control.

For the Semifinals, such a scheme was achieved using an admittance control [6] that was triggered only in specified conditions to override the user control. We enhanced this method by formulating a new admittance control scheme such that: if force constraints are violated, it will limit any motion that will increase the constrained force. In this way, we created a hierarchy between the maximum force one could apply and the desired motion. The key aspect of this scheme is that any motion getting away from the constraint is not limited.

B. *Wireless E-Stop*

We built a wireless emergency stop button (E-stop) that communicates to the robot using Zigbee 3.0 modules.

For a humanoid robot, turning off the actuators in an emergency could cause the robot to fall over. In addition, simply bringing the robot to a sudden stop during walking can also lead a similar outcome. Therefore, when receiving an emergency signal the real-time controller takes the robot to a stop stably with an appropriate timing.

VI. Experimental results

During the qualification day at the Finals, our Avatar fell down after successfully communicating with the mission commander. We got enough points to qualify, but although we repaired the robot it was not reliable and we were left with no other choice but to withdraw. Yet, we assessed our avatar afterward in-house realizing each of the skill-transfer tasks of the finals, as seen in the following video: `https://youtu.be/CaOOoSqWjCo`.

REFERENCES

[1] R. Cisneros, M. Benallegue, K. Kaneko, H. Kaminaga, G. Caron, A. Tanguy, R. Singh, L. Sun, A. Dallard, C. Fournier, M. Tsuru, C. Yang, Y. Osawa, G. Lorthioir, F. Kanehiro, and A. Kheddar. Team JANUS Humanoid Avatar: A cybernetic avatar to embody human telepresence. In *RSS 2022 Worshop on "Towards Robot Avatars: Perspectives on the ANA Avatar XPRIZE Competition"*, 2022.

[2] Laura Aymerich-Franch, Damien Petit, Gowrishankar Ganesh, and Abderrahmane Kheddar. The second me: Seeing the real body during humanoid robot embodiment produces an illusion of bi-location. *Consciousness and Cognition*, 46:99–109, 2016.

[3] M. Schwarz, C. Lenz, R. Memmesheimer, B. Pätzold, A. Rochow, M. Schreiber, and S. Behnke. Robust Immersive Telepresence and Mobile Telemanipulation: NimbRo wins ANA Avatar XPRIZE Finals, 2023.

[4] S. Dafarra, K. Darvish, R. Grieco, G. Milani, U. Pattacini, L. Rapetti, G. Romualdi, M. Salvi, A. Scalzo, I. Sorrentino, D. Tomè, S. Traversaro, E. Valli, P. Viceconte, G. Metta, M. Maggiali, and D. Pucci. iCub3 Avatar System, 2022. arXiv:2203.06972.

[5] Y. Chen, L. Sun, M. Benallegue, R. Cisneros, R.P. Singh, K. Kaneko, A. Tanguy, G. Caron, K. Suzuki, A. Kheddar, and F. Kanehiro. Enhanced Visual Feedback with Decoupled Viewpoint Control in Immersive Humanoid Robot Teleoperation using SLAM. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2022.

[6] Karim Bouyarmane, Kevin Chappellet, Joris Vaillant, and Abderrahmane Kheddar. Quadratic programming for multirobot and task-space force control. *IEEE Transactions on Robotics*, 35(1):64–77, 2019.