Mixed Reality-Based User Interaction Feedback for a Hand-Controlled Interface Targeted to Robot Teleoperation

Laura Cancedda, Alberto Cannavò, Giuseppe Garofalo, Fabrizio Lamberti, Paolo Montuschi, and Gianluca Paravati(✉)

Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy
{laura.cancedda,alberto.cannavo,s239732}@studenti.polito.it,
{fabrizio.lamberti,paolo.montuschi,gianluca.paravati}@polito.it
http://grains.polito.it

Abstract. The continuous progress in the field of robotics and the diffusion of its related application scenarios in today's modern world makes human interaction and communication with robots an aspect of fundamental importance. The development of interfaces based on natural interaction paradigms is getting an increasingly captivating topic in Human-Robot Interaction (HRI), due to their intrinsic capabilities in providing ever more intuitive and effective control modalities. Teleoperation systems require to handle a non-negligible amount of information coming from on-board sensors as well as input devices, thus increasing the workload of remote users. This paper presents the design of a 3D User Interface (3DUI) for the control of teleoperated robotic platforms aimed at increasing the interaction efficiency. A hand gesture driven controller is used as input modality to naturally map the position and gestures of the user's hand to suitable commands for controlling the platform components. The designed interface leverages on mixed reality to provide a visual feedback to the control commands issued by the user. The visualization of the 3DUI is superimposed to the video stream provided by an on-board camera. A user study confirmed that the proposed solution is able to improve the interaction efficiency by significantly reducing the completion time for tasks assigned in a remote *reach-and-pick* scenario.

Keywords: Human-robot interaction · Robot teleoperation · 3D user interface · Mixed reality · Visual feedback · Hand-based control

1 Introduction

Due to the recent advances in robotics, performing remote tasks with the support of robotic systems, possibly involving several interfaces to different robot functionalities, is becoming ordinary practice for a growing number of activities in several application fields, ranging from search and rescue in dangerous

L.T. De Paolis et al. (Eds.): AVR 2017, Part II, LNCS 10325, pp. 447–463, 2017.
DOI: 10.1007/978-3-319-60928-7_38

environments, to inspection of industrial plants and assistance for home-care, to name a few [1].

Human-robot interaction (HRI) plays a key role in teleoperation scenarios, since most of the developed system are still based on supervisory control by the human operator. In this scenario, it is of paramount importance to develop ever more effective and intuitive interaction modalities to let the operator focus on the relevant aspects of the task to be carried out, rather than getting distracted from the complexity of the interaction. Several factors influences the performance and the efficiency in carrying out teleoperation tasks. Along these lines, a careful design of the overall system is needed by considering, aside from technical aspects, interaction modalities, user interfaces as well as human factors.

The first factor influencing human-robot interaction is the input method. Several possibilities may be considered, ranging from the use of the most common user interfaces (i.e. keyboard, mouse, gamepad, etc.) to the adoption of the most recent technologies capable to elaborate and interpret human's voice and gestures. The latter category allows the operator to interact with the robot more naturally, by following the principles of Natural User Interaction (NUI) to increase the level of empathy and confidence [2].

Another important factor influencing the efficiency in robot teleoperation is the design of the Graphical User Interface (GUI), considered that the operator is remotely located and cannot directly check the maneuvering of a robot in response to issued commands. Furthermore, as argued in the remainder of the manuscript, despite their intuitiveness NUIs may introduce the need for additional information to be conveyed to the operator. Indeed, HRI relying on vision-based techniques lacks direct feedback concerning the working volume of the input device. In this scenario, mixed reality assumes a key role for enhancing the spatial awareness concerning the control of remote mobile robots by providing both virtual and physical information [3]. As a consequence, an essential and fundamental design principle is to provide the user with various forms of feedbacks. Indeed, feedback can be visual, tactile, auditory and, in general, may involve all senses together, but it is also tightly related to the selected input device.

Moved by the above considerations, this paper presents the design and development of a 3D user interface to control a multi-functional teleoperated system composed by a rover equipped with a robotic arm. The interface allows the user to interact with the system through the use of hand gestures captured by a Leap Motion controller, which has been selected due to its high accuracy [4]. A mixed reality interface is built by combining teleoperation information (e.g. concerning the surrounding environment, available functionalities, etc.) and visual feedback concerning user interaction in a single interface. Test results confirmed the ability of the designed interaction modality to reduce task completion times for telemanipulating a robotic arm with respect to common teleoperation approaches. Moreover, a subjective user study conducted on 12 participants revealed a preference for the proposed interface.

The remainder of this paper is organized as follows. Section 2 reviews related works concerning teleoperation systems by focusing on aspects related to interaction modalities as well as the design of the graphical user interface and motivates the design choices given in Sect. 3, which presents both architectural details and the design of the interaction modality. Section 4 reports both on the objective and subjective results obtained by testing the interface in a *reach-and-pick* scenario. Finally, conclusion are drawn in Sect. 5.

2 Background

In past years, many approaches have been proposed in the development of user interfaces for teleoperation systems [5,6]. Several interaction modalities have been investigated based on the specific task should be carried out by the robot itself. For instance, force-reflecting manual controllers are suitable for manipulating grippers and robotic arms, though they introduce time delays and suffer stability problems [7]. Datagloves have been widely used coupled with force feedback devices to give the remote user the perception of the situation faced by the controlled robot in the real environment as, for instance, in [8].

Despite the diversity of adopted interaction forms, control sticks and game controllers are definitely the most common interfaces used in human-robot interaction. The reasons behind their widespread diffusion in mass-market rest on their cost effectiveness, their accurate control capabilities and their general-purpose orientation. Complex teleoperation tasks involving robotic platforms with multiple degrees of freedom (DOF), like a robot manipulator, could be addressed by leveraging on the combination of multiple input devices. For instance, in [9] a couple of control sticks and a gamepad have been used to control a robotic arm by acting on different DOFs simultaneously, or by separately controlling each joint, respectively.

As widely known, the presence of a physical device for robot control permits to establish a feedback channel to be exploited for different purposes. A bidirectional system for teleoperation has been developed in [10], where a control stick has been used as input device to send commands to a mobile platform and information concerning the environment is sent back to the operator in the form of feedback forces through the control stick itself. Vibro-tactile feedback is also investigated in [11], where the interface shows a virtual space representing the robot and its adjacent objects in the real environment, superimposed to images coming from an on-board camera. The vibro-tactile feedback is used as complementary hint to visualization to improve spatial awareness by providing information about the distance from objects in the environment.

The functionalities of a teleoperation system may be difficult to be properly activated by means of control sticks and gamepads, in particular for inexperienced users. This is the situation when a large number of DOFs should be remotely controlled [12], e.g., for manipulating a robotic arm. Indeed, button- or stick-based input devices may require to operate simultaneously on numerous levers or buttons thus having a huge impact on the intuitiveness of the interface.

This shortcoming may be overcome by leveraging on intuitive gestures performed by the operator, which can be implemented in several ways from a technological standpoint, e.g., ranging from the use of wearable intertial sensors, to body tracking using vision based techniques, and so forth. Techniques able to capture user's poses and gestures are often referred to as Natural User Interfaces (NUIs). They have been proven to improve several aspects of human-robot interaction in teleoperation systems, e.g. by reducing the operator training time, cognitive load and enhancing situation awareness [13]. As a matter of example, in [13] gestures are defined to pick objects in a virtual environment by mapping the movement of users with a virtual representation of their hand. Virtual reality is exploited as well in [14] to let the user control a real robot through its virtual counterpart in a virtual setting. It is worth to outline that, although it is possible to virtually reconstruct the real environment, e.g., through simultaneous localization and mapping techniques, virtual reality decreases the user's sense of awareness with respect to a real image flow streamed by a camera [6].

As widely known, augmented reality techniques permit to retain real world vision by enlarging its level of knowledge. In [15], a robotic arm is controlled through an augmented reality interface exploiting an exocentric vision of the robot (i.e., slightly behind and above) built by means of a Kinect sensor and a head-mounted display. Hand tracking and gestures recognition relying on vision-based techniques are exploited to generate control commands. Similarly, a mixed reality user interface for teleoperation robots has been developed in [16] to visualize predicted scenarios. In this case, images coming from the onboard camera of a rover are augmented with a virtual representation of the wheeled platform itself, thus realizing an augmented reality exocentric view. The user directly operates by giving commands to drive the virtual robot, which is followed by the real one.

Despite the intuitiveness provided by NUIs through vision-based techniques, the main drawback with respect to handheld controllers is constituted by the impossibility to rely on vibro-tactile feedback to convey information to the operator. Indeed, NUI devices improve the situation awareness and naturalness of the interface, as noted by [17], but they need to provide an adequate visual feedback through additional information which could overload the screen. As a matter of example, in [18], the feedback has been provided in the form of visual information to develop an immersive interface to control a snake-like robot. The interaction is provided by a Leap Motion controller. The visual feedback is provided to the user by means of two separate devices: information concerning the environment is represented in the main screen of a workstation whereas control information relating to the field of view of the input device are provided through a tablet device. It is worth noticing that the decoupling of visualization information may increase the mental workload required to users due to repeated gaze switches among several devices.

Different sources of information may be provided in different viewports of the same interface, eventually by adapting contents and layout to the need of the user [19]. In [20], an application for remotely controlling an arm and grasp objects has been presented. In this case, the interface shows different views side by side: the real environment, the virtual corresponding scene, and the webcam stream from a ceiling camera above the robot; eventually in the virtual scene a hand skeleton mirrors the state of the tracked real hand. The dissociation among the hand representation, the arm virtual mapping and the image of the real arm could bring to a sense of unawareness [21] and does not improve the learnability of the interface. Similarly, in [22], the video coming from the on-board camera and the feedback concerning the field of view of the hand gesture driven controller are shown in distinct viewports. Furthermore, the field of view of the hand gesture driven controller is represented through two distinct two-dimensional views (front view and top view), which increases the complexity of the user interface.

By moving from the above considerations, we designed a system for interacting with remotely located robots which leverages on a Natural User Interface for effectively carrying out teleoperation tasks. The presented solution is able to overcome the above mentioned limitations of hands-free interaction modalities by means of a mixed reality-based feedback with the aim of integrating the robot egocentric vision with three dimensional information concerning the working volume of the hand gesture driven controller.

3 Teleoperation System

This section describes the main architectural components of the teleoperation system which has been used to test the mixed reality-based user interaction method. A multi-robotic platform composed by a rover (Lynxmotion A4WDI) and a robotic arm (Lynxmotion AL5D) can be remotely controlled.

Both of them are connected through a Wi-Fi network shield to communicate with a distributed controller software by receiving commands and sending information about sensors status. The rover moves reacting to directional and speed commands, while the arm is a 5-DoF manipulator as shown in Fig. 1a. The latter is controlled through an inverse kinematics solver by mapping the spatial coordinates of the user's wrist to the end effector of the robotic arm. Roll and pitch of the gripper are controlled through forward kinematics. A Web Cam Logitech C525 is mounted over the end effector of the robotic arm and connected to a Raspberry Pi to provide real-time visualization capabilities to the remote user, which can inspect the surrounding environment. The video is streamed through a Wi-Fi network.

The high-level architecture of the interaction system is illustrated in Fig. 1b. The user interacts with the system through a hand gesture based controller (in particular, a Leap Motion device has been used in this work), which is in charge to locate and track in real-time the user's hand within its working volume. Its interaction field is approximately shaped as an inverted pyramid centered at the device's center and extending upwards. The range of the tracking volume depends from the device's field of view, which is approximately 150 degree. Outside this region the tracking is lost.

An application is responsible for recognizing input commands for the rover and the arm and providing visual feedback, presented by its interface. Once recognized, numeric values associated to the gesture (i.e. spatial coordinates)are translated into actionable commands and then they are sent to the robot to properly activate its servomotors.

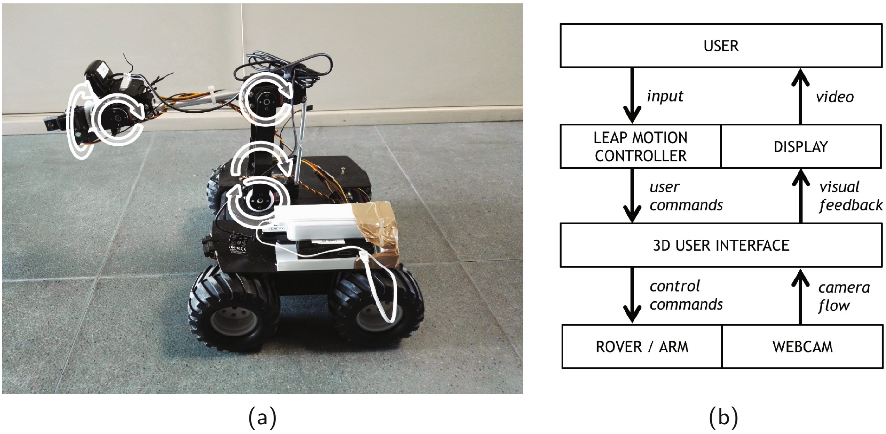


Fig. 1. The teleoperation system developed for testing the mixed reality user interface: (a) the multi-robotic platform is composed by a rover and a 5-DOF robotic arm, (b) high-level architecture of the designed interaction system.

Interfaces for teleoperation systems like the one developed for testing the presented user interface should permit a simple and direct interaction with the environment and objects. An interface is defined as effective when the user can reach his or her goals, when the important tasks can be done better, easier, or faster than by using another system, and when users are not frustrated or uncomfortable [20]. The user should be always in control of what is happening as a consequence of his/her commands; the less effort the user has to make to understand how the system is operating, the more attention he/she can pay to the task to be performed [21].

Any system based on gesture recognition requires an adequate feedback able to don't distract the user from the main task is being carried out. For this purpose we designed a mixed reality interface to overlap hints concerning the hand-gesture based controller (which intrinsically lacks feedback information otherwise available with other interfaces, e.g. handheld-based) to information aimed at providing situational awareness in a teleoperated system (in this case, the streaming flow provided by the front-facing camera mounted on board). Data related to the input controller are displayed by leveraging on a 3D visualization to enhance the user's perception of the controller's working limits. Dynamically adjusted 3D volumes represent the workable area where the user can effectively move his/her hands. A set of discrete commands are activated by the user by leveraging on some symbolic gestures previously defined (e.g., open hand to the left/right, roll the hand facing the palm upwards/downwards, close hand, etc.). It is worth to recall that interfaces for teleoperation benefit in naturalness and learnability by maintaining a coherence between the mapped gesture and the movement the user would do in real circumstances to reach a target, since real-world physical metaphors constitute a valid help for the user to remember associated actions [20,21].

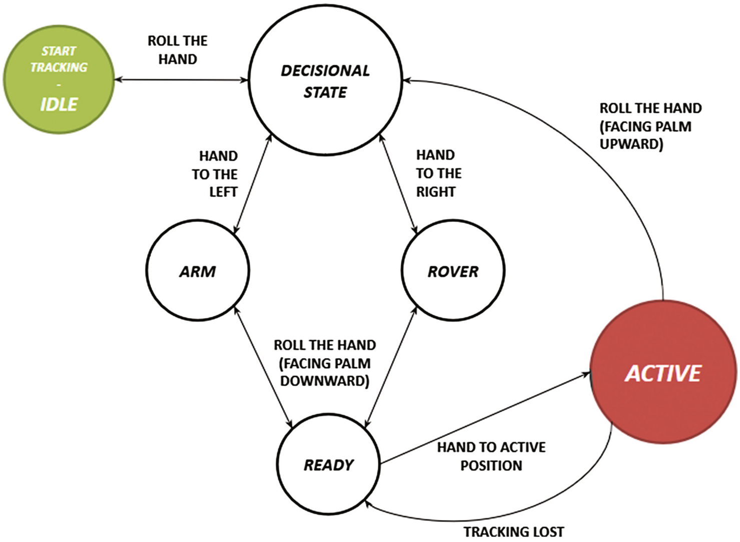


Fig. 2. State diagram of the devised interaction schema. Starting from the *Decisional State* it is possible to choose a control mode (*Arm/Rover*) to be selected (*Ready*); the interaction starts (*Active*) when a setup position is reached.

The interaction schema is based on the state diagram shown in Fig. 2. In the idle state, the hand tracker calculates the position and orientation of the user's hand. When the hand's palm is facing upwards, the application enters into a decisional state, in which the graphical user interface is divided into three main areas: two on the sides, which are labelled with the relative control mode that can be enabled (namely, "Arm" and "Rover" areas), and one centrally located, not associated to any control mode (i.e., a "Transition" area). Figure 3a shows a screenshot of the interface in the decisional state.

Once the user reaches one of the two sides, the specific operational states (i.e. arm- and rover-control modes) are activated through the rotation of the palm downwards. The reverse operation (i.e. rolling the hand's palm upwards) can be used to come back to the decisional state in any moment. For this reason, the latter represents a safe gesture always available. For safety reasons, the interaction does not start directly after the selection of the operational state. Indeed, although the hand position may be correct during the decisional state, abrupt or

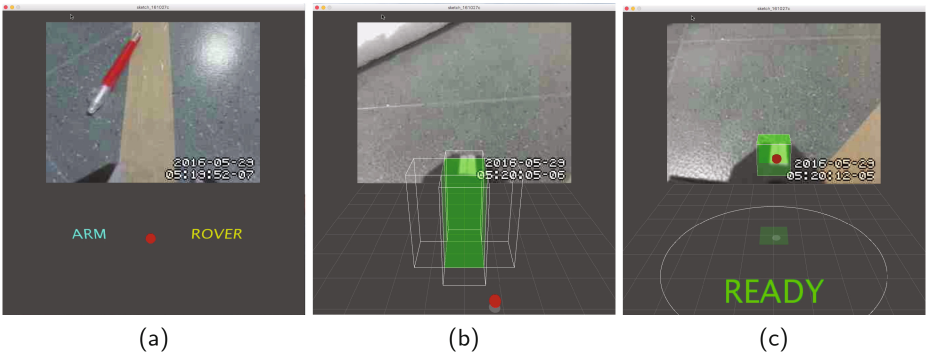
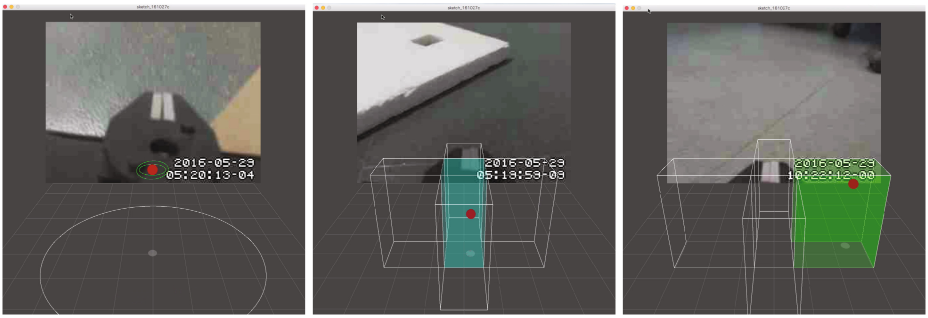


Fig. 3. Screenshots of the user interface for the selection of the robotic functionality to be controlled. (a) Decisional state with three areas, Arm area, Transition area and Rover area. (b) Setup position to be reached in rover driving mode. (c) Setup position to be reached in arm-control mode.

unwanted commands (e.g., sudden movements of the rover or the arm) may be produced when an operational state is activated. The user's hand – represented by a sphere in the interface – should reach a safe setup position within the 3D workspace of the hand gesture driven controller before the interaction with the selected robotic component begins. This way the user is driven to assume an initial pose which maximizes the available interaction volume. Figure 3 shows the setup position (highlighted in green) to be reached in the 3D user interface for each control mode. As soon as the user achieves the indicated volume, the interaction with the selected component begins.

When the arm functionality is selected, a grid – representative of the rover's base where the arm is mounted on – is overimposed to the streaming flow of the on-board camera. The relationship between the physical position of the hand and the underlying hand gesture driven controller is graphically represented together with the current tracking limits. In fact, a graphical indicator of the working area related to the hand tracking module is displayed on the screen in the form of an ellipsoid, which represents the projection of the tracking boundary on the surface grid. The border of this safe area should not be overrun by the sphere representing the hand of the user to avoid tracking failures. The tracking area is dependent from the height of the tracked hand with respect to the tracking hardware. For this reason, the size of the tracking area has been designed to be adaptive. In fact, the size of the projection reflects the width available for the movements on the horizontal plane at a certain height. As a result, the ellipsoid resizes as the hand moves in the vertical axis. To correctly track, the sphere's shadow should remain inside the indicated boundary. Otherwise, a tracking failure may occur, since the hand is located near the borders (or beyond) the working volume. In this case, the system detects the tracking failure and jumps to the idle state to avoid the triggering of unwanted commands.



(a) (b) (c)

Fig. 4. Workspace limits of the hand gesture driven controller. Their size is determined by the physical height of the tracked hand with respect to the input controller. (a) Interface for controlling the robotic arm and the tongs: the ellipsoid represents the boundary of the tracking volume. (b) Interface for driving the robot: a steady zone is used as a rest pose for the rover. (c) Interface for driving the robot: each (dynamically resized) volume represents a command.

Furthermore, to increase the spatial awareness, two concentric ellipses supplementing the sphere indicate the hand's roll (Fig. 4a). The user's hand position allows to configure in real-time all the robotic arm's joints through an inverse kinematics solver: it enables the movement of the robotic arm in all directions as well as the use of a tongs located in the arm's edge, which can rotate and further grab objects by opening and closing movements.

When the user selects the Rover area, the wheeled platform can be directly controlled. In this case, a custom shape represented by two perpendicular parallelepipeds in a cross configuration is overimposed to the camera stream flow (Fig. 4). The volume currently containing a sphere, representing the hand position, is highlighted by using different colours to give the user a prompt and immediate feedback about the state of the interaction tool: light blue for the central one corresponding to the steady position (Fig. 4b), green for forward, backward, left and right ones (Fig. 4c); when the sphere exits from one interaction shape, it returns to be transparent. As in the case of the arm control, graphical borders represents the physical boundaries of the tracking area. In fact, the size of the shape is dynamically changed according to the height reached by the hand with respect to the input controller in order to give to the user the perception that the tracked area becomes wider by increasing the distance between his/her hand and the controller itself. The right and the left sections are used to send rotational commands to the robot, while by moving forward or backward the hand, a user will trigger the movement of the wheels in the same direction of the moving hand.

Transitions between states (e.g. from Ready to Active state) as well as wrong movements (i.e. tracking losses) are associated to a visual feedback. For instance, tracking loss is particularly relevant in the interface due to the disorientation

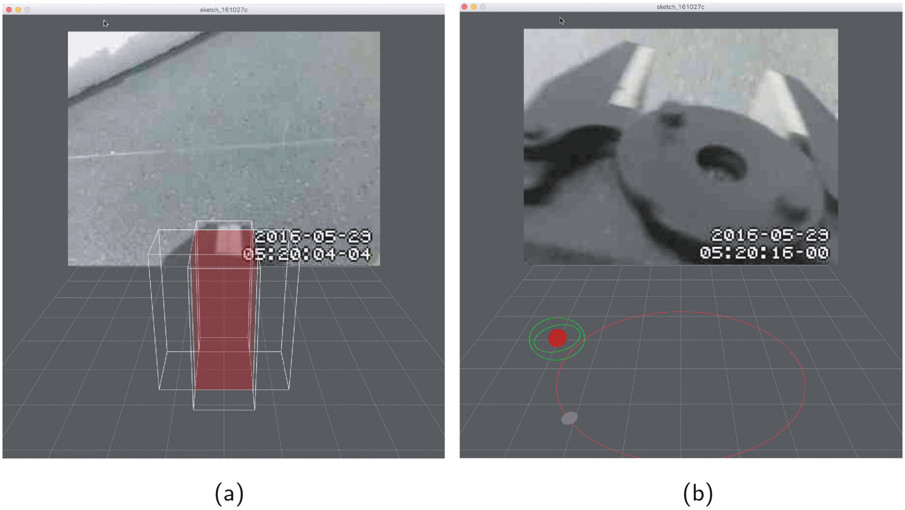


Fig. 5. Visual feedback for (a) tracking loss occurring in driving mode and (b) tracking loss occurring in arm-control mode.

generated to the user, thus slowing down the completion of the tasks. For this reason, as illustrated in Fig. 5, the tracking loss is clearly highlighted in both control modes by highlighting in red the interface elements. A video showing the interaction modalities with the robotic platform is available for download[1].

4 Results

To evaluate the usability and possibly identify the elements of the teleoperated system which can take more advantage from the designed 3D user interface, a user study has been conducted by involving the participation of 12 students at the University. The test consisted in carrying out a *reach-and-pick* task with three different user interfaces, with the aim of comparing the proposed one with a user interface based on a gamepad and another implementation based on a hand gesture driven controller [22] (in the following referred to as 3DUI, gamepad and 2DUI, respectively). The gamepad interaction method is related to the use of a long-established input device, where the teleoperation interface simply consisted in the real-time streaming flow of the on-board camera. The latter concerns the use of a Leap Motion controller to track user's hand position, orientation and status (open/closed), as in the current work, presented to the user by means of a 2D interface [22].

Participants were allowed to get accustomed with each interface by freely operating with the two components of the tele-manipulated platform, i.e., the

[1] Video: https://www.dropbox.com/sh/uxje5n18t41iyhu/AABBhfwCc1I1xfIIu6IYJoM4a?dl=0.

rover and the robotic arm, both by using the hand gesture driven controller and the gamepad. During this training session, participants were instructed about the interpretation of the information provided by the different user interfaces and they were invited to try to tele-control the rover and the arm by leveraging on the video stream received from the on-board camera.

The *reach-and-pick* scenario consisted in two consecutive sub-tasks, i.e., navigation and picking. The first one consisted in driving the rover along a predefined path drawn on the ground (Fig. 6). The first sub-task was considered as completed as soon as the rover reached a target point (i.e., on the top of a ramp). The latter consisted in grasping a small target object placed on the ground at the end of the path by using the robotic arm. In this case, the task was considered as concluded once the target object was grabbed. The two sub-tasks will be referred to as *Rover* and *Arm*. The overall scenario was designed to test the driving of the rover and the arm-control mode separately; each participant repeated three times the experiment, one for each interface. In order to reduce possible learning effects, the sequence of user interfaces to be used by each participant was randomly generated.

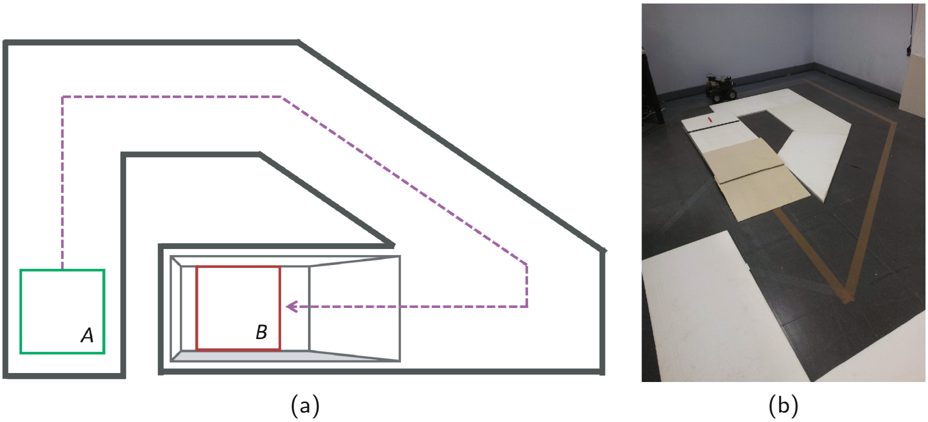


Fig. 6. *Reach-and-pick* scenario designed for the user study: A. start position of the rover for the navigation sub-task, B. target destination of the navigation sub-task and position of the picking sub-task.

Objective and subjective results were carried out to assess the designed user interface. The performance indicator selected for the objective evaluation was the task completion time, which has been measured for carrying out both sub-tasks. Subjective evaluations were collected to supplement the performance analysis with the users' opinions through a questionnaire filled in at the end of the experimental test.

Results obtained with the objective evaluation in terms of completion time for the two sub-tasks are reported in Fig. 7. On average, the completion time with the proposed 3DUI was lower both for the navigation (i.e., Rover) and picking (i.e., Arm) tasks. The statistical significance of the objective results was analyzed by running independent samples t-tests on collected data with significance level $\alpha = 0.05$. According to the statistical analysis, the proposed interface outperformed both the 2DUI and Gamepad ones in the picking scenario ($p = 0.0148$ and $p = 0.0460$, respectively). Concerning the navigation scenario, the average completion time was significantly lower by using the proposed interface with respect to the 2DUI ($p = 0.0189$).

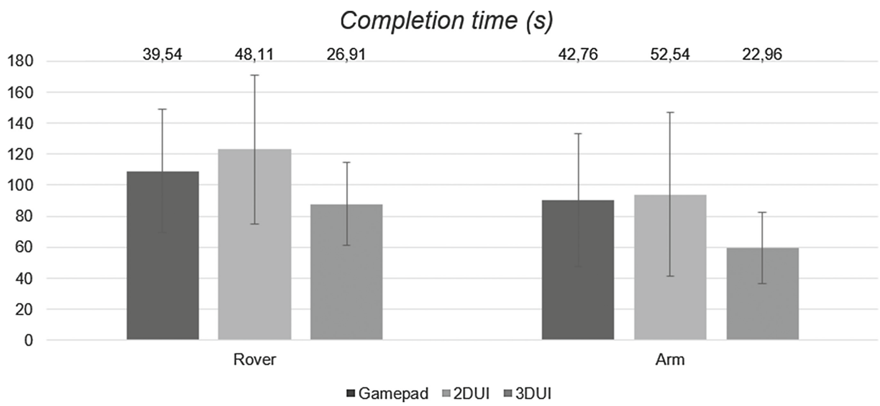


Fig. 7. Average completion time for the sub-tasks (carried out by tele-manipulating the rover and the robotic arm) by using the Gamepad, the 2DUI and the 3DUI. Bars height indicate the average value (lower is better) whereas labels indicate the standard deviation.

As said, after the conclusion of the experimental test, participants were asked to fill in a questionnaire for the evaluation of their experience by using the different interfaces. The survey was divided in three main sections. The first one was aimed at investigating usability aspects defined by Nielsen [23]. To this aim, five questions were included to determine whether a participant was satisfied in terms of learnability, memorability, efficiency of the user interface controls, pleasantness during the interaction, and number of errors made. The second section was created by considering the SASSI methodology [24], which was originally targeted to speech interfaces. In this work, the questionnaire was integrated with questions derived from the SASSI methodology to possibly evaluating usability factors not covered by Nielsen. In particular, participants were asked to evaluate the following five usability factors: System Response Accuracy (SRA), Cognitive Demand (CD), Annoyance (AN), Habitability (HAB) and Speed (SPE). Lastly, the preference for a particular interface was asked in the last section of the questionnaire. Overall, the survey had 10 statements to be evaluated on a five-point Likert scale from 0 (strong disagreement) to 4 (strong agreement) for each interface and for each task separately.

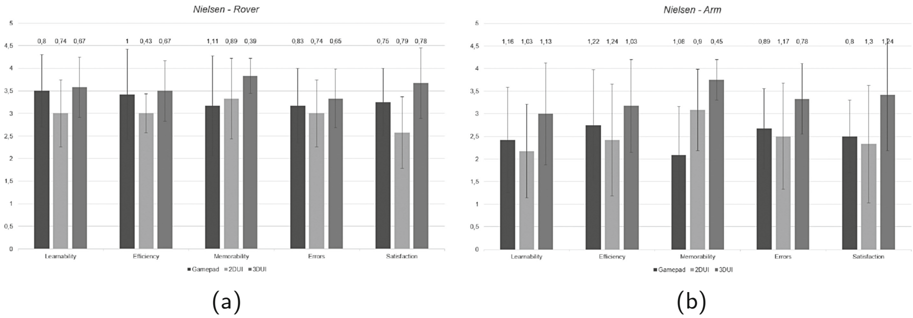


Fig. 8. Overview of the subjective results for the two sub-tasks using the three interfaces based on the Nielsen methodology and concerning Learnability, Efficiency, Memorability, Errors and Satisfaction. Bars height indicate the average value (higher is better), labels indicate standard deviation.

Figure 8 reports an overview of the average subjective results gathered with the Nielsen methodology. Results are plotted using a worse-to-better scale. Again, the statistical significance was checked by adopting the same procedure used for the objective evaluation. According to the outcomes of the statistical analysis for the Nielsen methodology, the proposed user interface always scored better than or equal to the other user interfaces. In particular, memorability reached the highest score for the 3DUI, demonstrating the easiness in gaining expertise and remembering commands in the use of the interface. Analyzing the picking sub-task more in the detail (Fig. 8b), the proposed interface proved to be robust from the point of view of the errors (i.e., perception of committed errors, easiness of recovery, and how severe they are) beside memorability. In fact, both the usability factors scored significantly better than Gamepad ($p = 0.0020$ for memorability, $p = 0.0389$ for errors) and 2DUI ($p = 0.0046$ for memorability, $p = 0.0437$ for errors). A significant preference over the 2DUI was observed for learnability ($p = 0.0021$), efficiency ($p = 0.0054$) and satisfaction ($p = 0.0053$). Results concerning the perception of efficiency are not statistically significant by comparing the 3DUI and the Gamepad, but the above objective results give evidence of how quickly users can perform tasks.

Figure 9 show the results for the second section of the questionnaire based on the SASSI methodology. Average results indicate that the proposed interface always scored better than or equal to the other ones. The Habitability usability factor was significantly higher than the Gamepad and 2DUI in both sub-tasks ($p = 0.0388$ and $p = 0.0015$, respectively, for the Rover; $p = 0.0116$ and $p = 0.0019$, respectively, for the Arm). The 3DUI outperformed the 2DUI also in terms of System Response Accuracy. In fact, statistical significance is achieved both for the navigation ($p = 0.0069$) and picking ($p = 0.0209$) interaction scenarios. Objective results are confirmed by the Speed usability factor. In this case, only results concerning the manipulation of the robotic arm are statistically significant (the 3DUI is perceived to provide a faster way to execute

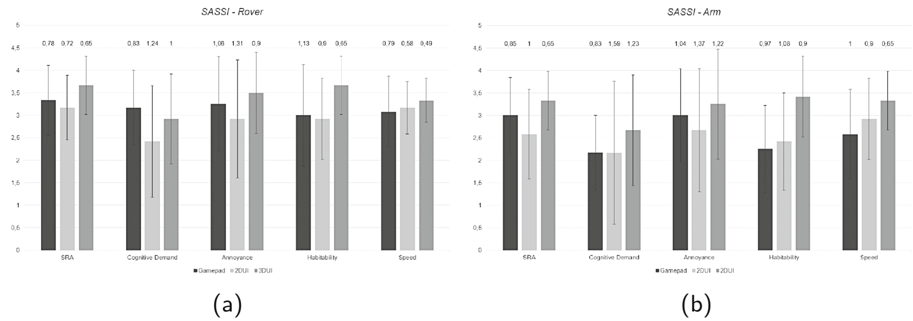


Fig. 9. Overview of the subjective results for the two sub-tasks using the three interfaces obtained based on the SASSI methodology and concerning System Response Accuracy (SRA), Cognitive Demand (CD), Annoyance (AN), Habitability (HAB) and Speed (SPE). Bars height indicate the average value (higher is better), whereas labels indicate standard deviation.

tasks than the Gamepad, $p = 0.0209$), totally in accordance with the previous measurements of the task completion time.

In the last part of the questionnaire, participants were asked to express their preference for a particular user interface among those analyzed. According to the scores summarized in Fig. 10, the overall preference for the whole experiment was 46% in favor of the proposed interface, 31% of participants indicated the Gamepad, and 24% preferred the 2DUI. To get more insights about the usability of the 3DUI for different tasks and to possibly identify the elements of the teleoperated system which can take more advantage from the proposed interaction modality, users were asked to express their preference for each of the different phases of the *reach-and-pick* scenario. For the first sub-task, i.e. navigation, neither of 3DUI and Gamepad emerged clearly as preferred interface (they were judged as first choice by 39% and 38% of participants, respectively). For the

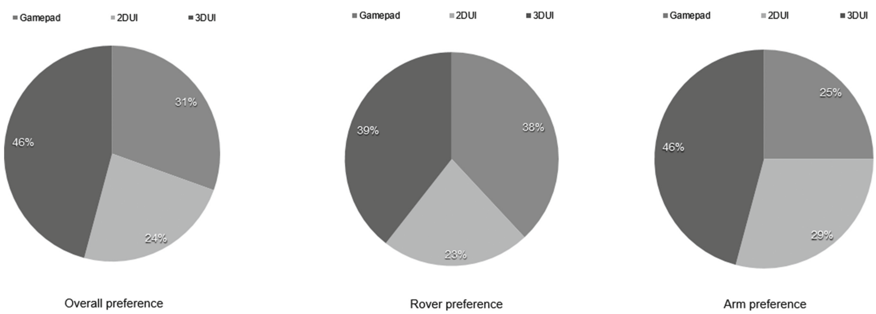


Fig. 10. Overall preference, Rover preference, Arm preference expressed in the questionnaires.

last phase, i.e. grabbing, 46% of participants preferred the proposed interface, 25% the gamepad, and 29% the 2DUI.

Based on the above results, the designed interface proved to be effective in particular when the teleoperation task involved the control of complex robotic components. Indeed, the "reaching" mode actually required to drive the rover by acting on a few commands. Conversely, "picking" mode required to concurrently act on several degrees of freedom.

5 Conclusion and Future Works

In this work, a mixed reality-based user interface aimed at improving the interaction efficiency for controlling multi-task mobile robots during teleoperation tasks is presented. The interaction is based on a hand gesture driven approach, which requires to provide users with additional status advices to effectively interact with the remote platform (e.g., concerning the physical workspace limits of the input controller), in addition to information related to situational awareness (i.e., data coming from on-board sensors, e.g., a camera). The proposed interface blends the ego-centric vision of the remote platform with a 3D representation of the interaction space for hand gesture based input controllers as well as the real-time position of the user's tracked hand. The combination of these elements into the same visualization area constitutes the strength of the proposed approach, as it allows to increase the usability of the interface as demonstrated by the experimental tests carried out in a *reach-and-pick* scenario, which involved the comparison with two different interfaces (based on a gamepad and on hand tracking with a 2D visualization of the same input data, respectively). The mobile platform was composed by a rover and a robotic arm, both controlled by the same user interface. Objective and subjective results confirm the ability of the proposed solution to make easier and efficient the interaction with the teleoperation platform, as well as pleasant to use. In particular, the designed mixed-reality based interface allowed participants to reduce execution times in the task involving the control of the robotic arm. Moreover, subjective evaluations pointed out that the presented solution outperforms the gamepad implementation in terms of memorability and habitability, thus suggesting that the determination of appropriate gestures and actions for reaching the task goal was easier than common approaches.

Given the encouraging results obtained, future works will be addressed to improve several features of the considered teleoperation interface. Several cameras and sensors (e.g. proximity) showing different points of view may be used to improve the spatial awareness and reduce the so-called "key-hole" effect [25]. In this case, additional information to be visualized on the interface should be taken into account. Other interaction possibilities will be considered as well, e.g., by adding additional gestures to move or change the current point of view.

References

1. Boboc, R.G., Moga, H., Talaba, D.: A review of current applications in teleoperation of mobile robots. Bull. Transilvania Univ. Brasov Ser. I Eng. Sci. **5**(54), 9–16 (2012)
2. Sriram, S., Wijnand, I.: Survey and classification of spatial object manipulation techniques. IPO, Center for User-System Interaction, Eindhoven University of Technology, pp. 1–2 (2000)
3. Ricks, B., Nielsen, C.W., Goodrich, M.A.: Ecological displays for robot interaction: a new perspective. In: Proceedings of the IROS, p. 2 (2004)
4. Coelho, J.C., Verbeek, F.J.: Pointing task evaluation of Leap Motion Controller in 3D virtual environment. In: Proceedings of the Chi Sparks 2014 Conference on Creating the Difference, pp. 78–85 (2014)
5. Nielsen, C.W., Goodrich, M.A., Ricks, R.W.: Ecological interfaces for improving mobile robot teleoperation. Trans. Robot. **23**(5), 927–941 (2007)
6. Chen, J.Y., Haas, E.C., Barnes, M.J.: Human performance issues and user interface design for teleoperated robots. IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. **37**(6), 1231–1245 (2007)
7. Cui, J., Tosunoglu, S., Roberts, R., Moore, C., Repperger, D.W.: A review of teleoperation system control. In: Proceedings of the Florida Conference on Recent Advances in Robotics (FCRAR), pp. 1–12 (2003)
8. Hu, H., Li, J., Xie, Z., Wang, B., Liu, H., Hirzinger, G.: A robot arm/hand teleoperation system with telepresence and shared control. In: Proceedings, 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Monterey, CA, pp. 1312–1317 (2005)
9. Crainic, M.F., Preitl, S.: Ergonomic operating mode for a robot arm using a game-pad with two joysticks. In: IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics, Timisoara, pp. 167–170 (2015)
10. Cho, S.K., Jin, H.Z., Lee, J.M., Yao, B.: Teleoperation of a mobile robot using a force-reflection joystick with sensing mechanism of rotating magnetic field. IEEE/ASME Trans. Mechatron. **15**(1), 17–26 (2010)
11. De Barros, P.G., Lindeman, R.W., Ward, M.O.: Enhancing robot teleoperator situation awareness and performance using vibro-tactile and graphical feedback. In: Proceedings of IEEE 2011 Symposium on 3D User Interfaces. 3DUI 2011, pp. 47–54 (2011)
12. Hashimoto, S., Ishida, A., Inami, M., Igarashi, T.: TouchMe: an augmented reality interface for remote robot control. J. Robot. Mechatron. **25**, 529–537 (2013)
13. Levine, S.J., Schaffert, S., Checka, N.: Natural User Interface for Robot Task Assignment, pp. 1–5 (2010)
14. Cheng-jun, D., Ping, D., Ming-lu, Z., Yan-fang, Z.: Design of mobile robot teleoperation system based on virtual reality. In: IEEE International Conference on Automation and Logistics, Shenyang, pp. 2024–2029 (2009)
15. Peppoloni, L., Brizzi, F., Avizzano, C.A., Ruffaldi, E.: Immersive ROS-integrated framework for robot teleoperation. In: IEEE Symposium on 3D User Interfaces 2015, pp. 23–24, March 2015
16. Sauer, M., Zeiger, F., Schilling, K.: Mixed-reality user interface for mobile robot teleoperation in ad-hoc networks. IFAC Proc. Volumes **43**(23), 77–82 (2010). ISSN 1474-6670
17. Hu, C., Meng, M.Q., Liu, P.X., Wang, X.: Visual gesture recognition for human-machine interface of robot teleoperation. In: Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robotics System, pp. 1560–1565 (2003)

18. Krupke, D., Lubos, P., Bruder, G., and Steinicke, F., Zhang, J.: Natural 3D interaction techniques for locomotion with modular robots. In: Mensch und Computer 2015: Gemeinsam Arbeit Erleben (2015)
19. Kawamura, K., Nilas, P., Muguruma, K., Adams, J.A., Zhou, C.: An agent-based architecture for an adaptive human- robot interface. In: Hawaii International Conference on System Sciences (HICSS 2003) (2003)
20. Krupke, D., Einig, L., Langbehn, E., Zhang, J., Steinicke, F.: Immersive remote grasping: realtime gripper control by a heterogenous robot control system. In: VRST 2016 Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, pp. 337–338 (2016)
21. Bowman, D., Koller, D., Hodges, L.: A methodology for the evaluation of travel techniques for immersive virtual environments. Virtual Reality Res. Develop. Applicat. **3**(2), 120–131 (1998)
22. Bonaiuto, S., Cannavò, A., Bazzano, F., Paravati, G., Lamberti, F.: Tele-operation of Robot Teams: A Comparison of Gamepad-, Mobile Device and Hand Tracking-based User Interfaces. Unpublished
23. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the ACM CHI 1990 Conference (Seattle, WA, 1–5 April), pp. 249–256 (1990)
24. Hone, K.S., Graham, R.: Towards a tool for the subjective assessment of speech system interfaces (SASSI). Nat. Lang. Eng. **6**, 287–303 (2000)
25. Plemmons, D., Mandel, P.: Designing intuitive interfaces. In: Introduction to Motion Control. https://developer-archive.leapmotion.com/articles/designing-intuitive-applications