STAG: Smart Tools and Applications in Graphics (2017)
A. Giachetti, P. Pingi and F. Stanco (Editors)

User Interaction Feedback in a Hand-Controlled Interface for Robot Team Tele-operation Using Wearable Augmented Reality

Alberto Cannavò and Fabrizio Lamberti

Politecnico di Torino, Dipartimento di Automatica e Informatica
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
alberto.cannavo@polito.it, fabrizio.lamberti@polito.it

Abstract

Continuous advancements in the field of robotics and its increasing spread across heterogeneous application scenarios make the development of ever more effective user interfaces for human-robot interaction (HRI) an extremely relevant research topic. In particular, Natural User Interfaces (NUIs), e.g., based on hand and body gestures, proved to be an interesting technology to be exploited for designing intuitive interaction paradigms in the field of HRI. However, the more sophisticated the HRI interfaces become, the more important is to provide users with an accurate feedback about the state of the robot as well as of the interface itself. In this work, an Augmented Reality (AR)-based interface is deployed on a head-mounted display to enable tele-operation of a remote robot team using hand movements and gestures. A user study is performed to assess the advantages of wearable AR compared to desktop-based AR in the execution of specific tasks.

CCS Concepts

•***Human-centered computing*** → *Mixed / augmented reality; Gestural input;* •***Computer systems organization*** → *Robotic control;*

---

1. Introduction

Due to rapid advancements in the field of robotics, supporting human activities with robotic systems is progressively becoming ordinary practice in a number of application fields, ranging from the inspection of industrial plants [MBGH14] to assistance in home-care settings [JWPD13], search & rescue in dangerous environments [EN16], etc. In many of the above situations, different forms of artificial intelligence are used to make the robots capable to perform activities independently from the human control. However, the supervision by a human operator is still needed, especially when robots operate in critical situations [SSF*00].

In these scenarios human-robot interaction (HRI) is exploited to combine the humans' reasoning skills and robots capabilities to create a tele-operated system. During tele-operation, effective user interfaces can help the operator to complete the assigned task in the most efficient way, without failures and damages to the robot [GS08]. At the same time, the development of intuitive interaction modalities is needed to let the operator focus only on the goals of the task to be carried out and not on the complexity of the interaction with the robot system, since complexity could break his or her cognitive involvement with a negative impact on the overall performance [FCC05].

Negative effects of poorly designed user interfaces on the execution of a given task are more evident when the system involves the collaboration of multiple robots. Taking advantages of maneuvering multiple robots at the same time can be helpful in many scenarios [SAB*07], as it gives the possibility to overcome the limitations of a single robot by activating the functionalities of another robot of the team [SSF*00], [SAB*07], [BCP*17].

According to [CCG*17], the two main factors that influence the execution of a tele-operation task are the input method(s) and the design of the Graphical User Interface (GUI). Concerning the first factor, several approaches for implementing efficient input methods have been considered in the literature, ranging from conventional user interfaces like keyboard, mouse, gamepad, etc. to new technologies referred as Natural User Interfaces (NUIs), e.g., based on speech recognition, hand and body gestures, etc.. The second factor is closely related to system's ability to provide a proper feedback to the operator about the actual state of the robot and the conditions of the surrounding environment. It is worth observing that, especially when NUIs are used, it is also important to provide the operator with a feedback about his or her own environment (e.g., regarding the physical space in which his or her body and hands are being tracked, if a gesture recognition technique is used).

To improve the user experience and enhance the spatial awareness of the operator, Augmented Reality (AR) -based approaches have been considered to provide visual feedback concerning both the robot's conditions and the working space of the input device

DOI: 10.2312/stag.20171227

[CCG*17]. The advantage provided by the use of AR is confirmed by the growing number of works experimenting with this technology in the considered domain [RLGSFL*11] [FK16] [HIII11] [SZS10] [PBAR15].

By moving from the above considerations, this paper presents the design and the development of a user interface based on hand gestures and wearable AR to tele-operate a robot team composed by a rover and a robotic arm. Interface design builds upon a previous work [CCG*17] where AR technology is implemented on a desktop computer screen. The input method select to control the multiple robot functionalities is a hand tracking system based on a Leap Motion controller. AR is used to visually represent the working space the operator has to move his or her hand into in order to tele-operate robot's functionalities.

With respect to [CCG*17], in this paper AR contents are displayed on a wearable video-see through device, which allows the operator to see augmented information overlapped to his or her own hand. The basic assumption is that affordances offered by this visualization can help the operator to have a clearer understanding of the functioning of the interface and a higher awareness of the space he or she is working into. This way, better performance could be possibly achieved. To evaluate the effectiveness of the proposed design, a comparison of the two interfaces was carried out through a user study.

The rest of the paper is organized as follows. In Section 2, a review of works concerning user interfaces exploited in tele-operation scenarios is provided. Section 3 describes the proposed interface in the context of a robot team tele-operation scenario. Section 4 illustrates the experimental setup and reports on the outcomes of the objective observations collected in the study. Lastly, Section 5 concludes the paper by providing possible directions for future research activities in the field.

2. Related works

This section describes the main approaches presented in the literature for HRI in robot tele-operation scenarios. In this domain, the robot is regarded as a slave, whereas the operator is the master who controls it. This approach is particularly helpful when the operator needs to manage a robot at distance, e.g., because humans cannot directly enter an hazardous area [AMP13]. In the literature, a large number of user interfaces for robot tele-operation have been presented, based on a number of heterogeneous technologies. For instance, in [UIA13], a combination of visual control, through electrooculography, and manual control, using a haptic manipulator is experimented. In [HPC*15], a brain-machine interface that is able to classify four mental tasks is used to control a robotic arm. In [SKS*06], a teddy bear-based robotic interface is proposed for entertainment.

Despite the richness of possibilities, joysticks and gamepads are by far the most common interfaces used in HRI applications, since they are rather cheap, easy to design and can provide accurate control. For instance, the joystick is considered as the *de facto* standard interface for most commercial manipulators, because it allows users to simply operate the end-effector of the robotic arm through directed selections [MLH*11]. In [CP15], a gamepad with 12 buttons and 2 joysticks in used to control the arm in two modalities: in X-Y-Z coordinates, with the left joystick commanding the X-Y movements and the right joystick controlling the Z axis, or in joint coordinates, with joysticks controlling each joint separately. In [CJLY10], a bidirectional tele-operation system is shown, where the operator sends command to the mobile robot through a joystick, and information regarding the environment is sent back to user in the form of feedback forces through the joystick itself.

Joysticks and gamepads could be difficult to use, especially by inexperienced users as well as when the robots have many functionalities or a large number of degrees of freedom (DoFs) to control [HIII11]. In fact, the need to operate, possibly simultaneously, on numerous buttons and levers could make the mapping between the user input and resulting robot movement not intuitive. In order to cope with this drawback, it is possible to replace, e.g., the click on a button or the manipulation of a lever with technologies that allow to directly track the operator's movements. A number of works in the literature have been reported implementing, e.g., body and hand tracking using vision based techniques, wearable inertial sensors, etc., often referred to as examples of NUIs, which improve different aspects of the HRI in tele-operation scenarios. For example, NUIs allow to reduce the operator training time and the cognitive load with an enhancement of the situation awareness [LSC10].

As a matter of example, in [NPM09], an industrial robot is controlled using a wearable system with a number of accelerometers mounted on operator's arm. In [AAK15], a first-person view- and body motion-based control method for a rover is presented. The interface is meant to boost sense of presence and spatial understanding, since it allows to directly map the user's movements captured by sensors on operator's hand onto the robot actions. In [KRO*14] a remote robotic manipulator is operated using an interface combining surface-electromyography with inertial measurements. In [YYDA15], a depth camera is used to detect human body joints and objects to be grasped. One hand of the operator is followed by a 7-DoF robotic arm, whereas the pose of the other hand controls the gripper. In this case, the Microsoft Kinect V2 is used.

The Leap Motion controller is another device that is getting quite commonplace in HRI applications, because it allows to track user's hands with a high accuracy [WBRF13]. Thus, in [VP15], the considered device is exploited to transfer the movement of a human hand onto a robotic one using inverse kinematics. In [BGG*14], an algorithm is proposed to achieve an optimum mapping between user's hand and a 6-DoF robotic arm with the goal to help users with upper limb problems to perform some daily activities.

It is worth saying that NUIs could be particularly effective in solving some of the issues of HRI, but could be affected by a limited working area and by occlusion phenomena. Furthermore, since the need for a physical connection between the user and the system he or she is operating onto is generally removed, a feedback, e.g., about the space he or she is moving his or her body or hands into needs to be provided (e.g., trough supplementary information conveyed on a screen).

For example, [BCP*17] presents an interface to manipulate a robot team composed by a rover, a robotic arm, and a drone trough the Leap Motion controller. In this case the interface shows on different viewports the two dimensional views (front and top) that describe the working volume of the hand gesture driven controller (together with the video streamed by the camera mounted on the robots). The repeated gaze switching among different viewports caused by the decoupling of information displayed may increase the complexity of the interface. For this reason, approaches based on Virtual Reality (VR) and AR have been proposed.

For instance, VR for tele-operation is exploited in [CjPMlYf09], where the operator controls a real robot by manipulating a virtual copy of it in a synthetic environment. In [LSC10] the picking gesture of a virtual object performed by a user sitting at a virtual reality control station is converted into an appropriate grasp and motion gesture executed by the robot. The main drawback of the above solutions is represented by the need to rely on a reconstruction of the robot's surrounding environment, which can reduce the user's sense of awareness with respect to the real image flow streamed by a camera mounted on the robot itself. In addition, although different solutions for 3D model reconstruction are available (e.g., photogrammetry, localization and mapping techniques, etc.), there are conditions that would prevent them to be used, e.g., in search & rescue tasks after a disaster event.

AR, in turn, has been widely exploited, because it easily allows tele-operation systems to improve the perception of the real word with digital contents that can be used to provide effective feedback.

For instance, in [YON17] it is shown how AR can enhance the operator's situational awareness and reduces his or cognitive load since he or she can be immersed in a helpful representation of the remote site. Another benefit of the considered technology is remarked in [RLGSFL*11], where authors demonstrated through a user study that AR can help inexperienced users to drive mobile robots through intuitive tele-operation interfaces. In [HIII11], a four-wheeled mobile robot equipped with a robotic arm is considered, and the user can choose the moving part to control by directly tapping on it in a third person view of the world that is shown on a tablet display with AR. In [PBAR15] an AR interface based on hand tracking and gesture recognition for controlling a robotic arm is presented. In this case, the interface shows to the operator an exocentric vision of the robot (slightly behind and above it). The mobile application discussed in [FK16] allows an operator to monitor and control a planar manipulator leveraging on an immersive interface executed on a tablet. The real-time video captured by the front facing camera of the tablet is exploited to track the position of the end-effector and the orientations of the four joints of the robot by applying computer vision algorithm. These measurements are enhanced with AR contents to provide visual feedbacks about robot's state. In [YON17], AR is used to perform remote maintenance with robots exploiting a wearable immersive display that is responsible to replicate the maintenance environment showing physical and virtual objects in the operator's workspace. An handheld manipulator allows the operator to manipulate a robotic arm whereas inverse kinematics solver and motion planning algorithms are used to visualize places of the environment that can be reached from the robot. In [SZS10] a mixed reality interface for robot tele-operation is presented, which combines the on-board camera of a rover with a virtual representation of the robot itself. A new approach to tele-operation of 7 DoF manipulators in rescue missions still based on mixed reality is presented in [LRRMRC*16]. The mixed reality interface executed on an Android device combines the virtual representation of the robotic arm, obtained through its embedded sensors, with camera images to help the operator to make decisions in real time on the correct command to be issued.

By moving from findings reported in the above works and in particular considering the benefits of the AR technology, a wearable AR-based interface for the control of a robot team is developed in this paper, by specifically building on the results of [CCG*17] and applying AR technology in a different way with the goal to provide the operator with an alternative visualization of the working space.

In [CCG*17], desktop AR is used to create a hand gesture-based interface for the control of robot team including a rover and a robotic arm. Hand tracking and hand gesture recognition are implemented using the Leap Motion controller. The rover can be tele-operated by moving the hand in a number of control boxes defined in the Leap Motion controller's working volume. Robotic arm is controlled by mapping the operator's hand on the arm's end-effector. The position and orientation of operator's hand as well as control boxes are displayed on a computer screen as AR contents, overlapped to the video stream transmitted by the robot's on-board camera.

With respect to [CCG*17], the current interface is implemented using video see-through AR on a head-mounted display to let the user see, at the same time, his or her hand and augmented contents. By leveraging the presence of the real hand in the field of view and its affordances, several changes were implemented in the way AR contents are displayed. In particular, the video stream received by the remote camera was linked to hand position, under the assumption that the operators looks at his or her own hand and to the tracking space defined by the interaction device. However, the control box-based interaction paradigm which was proved already to be significantly effective was not changed, with the objective to specifically focus on the impact that see-through visualization can have on the perception of the environment and of operations performed.

3. Tele-operation system and proposed interface

This section presents the tele-operation system which was exploited to control a robot team based on a rover carrying a robotic manipulator. The blocks that constitute the overall system are the input/output devices, the controller manager, and the robot team. In the following, blocks will be described in detail. Relations between then are illustrated in Figure 1.

3.1. Input/Output devices

The Leap Motion controller allows the user to interact with the system through a hand tracking-based interface. The device provides to the controller manager with the real-time position, orientation and status (open/closed) of the operator's hand. The Leap Motion controller embeds two cameras, which are used to capture the effect of three IR LEDs on operator's hand. Cameras' configuration defines a working volume that is shaped as an inverted pyramid

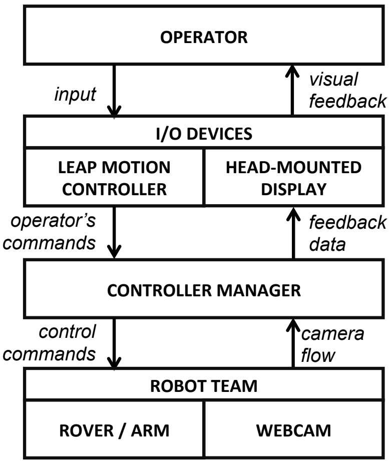


Figure 1: *Conceptual architecture of the tele-operation system.*

centered on the device and expands upward with a field of view of 150° wide on each side, 120° deep on each side and a range which extends from approximately 2.5 to 60 centimeters above the device's plane.

As in [CCG*17], the working volume of the Leap Motion controller is divided in several interactive regions, which are associated to different commands that can be sent to the robot (more details about regions' position and shape are provided in the following).

A head-mounted video-see-through display (based on a Samsung Galaxy S6 smartphone and a Samsung Gear VR device) is used to present the operator with AR contents defining the size and shape of the working volume, the currently active interactive regions as well as the video streamed by the robot's camera. Virtual contents are displayed on top of the Leap Motion controller in their real world position. Video stream, in turn, is hooked to the operator's hand, in order to bring it where the operator's is supposed to point his or her gaze and to have his or her hand constantly tracked.

3.2. Controller manager

This block is responsible to convert the operator's input (e.g., moving of the tracked hand forward, left, etc.) gathered by the Leap Motion controller into suitable commands for the active robot. For instance, if the active robot is the rover, the position of the operator's hand is translated into a value of velocity that is sent to the robot. Moreover, the block is in charge of sending the hand tracking information and the video stream to the head-mounted display, where augmented feedback enabling effective interaction is displayed.

The interaction schema is based on the state diagram shown in Figure 2. States names are given using capital letters. The transitions between two different states are activated by specific gestures (e.g., roll the hand, move the hand to the extreme left/right, etc.). The output shown to the operator on the head-mounted display is illustrated in the state box.

Initially, the system is in the IDLE state. In this phase the two robots are in a rest position and only the hand's position and orientation are tracked. The video stream of the remote camera follows the movement of the red point, that indicates the position of the operator's hand.

When the operator rolls the hand facing the palm upward, the system moves into the DECISIONAL state, in which the operator has the possibility to choose the robot to activate. In this state, the user interface is divided in two regions labeled with the name of the robot that can be activated (rover to the right, robotic arm to the left). If the user moves his or her hand to the left/right reaching one of the two side and rotates the palm downward, the system enters in the ready state for that specific robot (ROVER READY and ARM READY, respectively). In the ready state, interaction is not yet enabled for safety reasons. The interface show a region in the working volume colored in blue (different for the two robots). This region represents the safe position to be reached by the hand in order to begin the interaction with the particular robot. This mechanism forces the operator to assume a correct initial pose with his or her hand in order to avoid unwanted commands. A soon as the hand reaches the safe region, the interaction with the selected robot begins, and continues until the transition to a different state occurs (e.g., when tracking is lost).

When the rover is selected (ROVER ACTIVE), the operator is able to drive it by moving the hand inside the region represented by four boxes in a cross configuration. Each box represents a different command that can be issued to the rover. For example, when the hand enters the box to the left, as shown in the figure, the rover is turned left. When the hand enters the farthest box, the rover is moved forward. The region selected is highlighted in green to provide a visual feedback about the command sent. As said, the size of the interactive regions is dynamically changed depending on the height reached by the operator's hand in order to take into account the shape of the tracking space of the Leap Motion controller. The value of the speed and direction commends sent to the rover depends on the distance between the hand and the vertical axis of the Leap Motion controller.

When the robotic arm is selected (ARM ACTIVE), the joints of its kinematic chain can be manipulated through an inverse kinematics solver, which allows the operator to move the robot's end-effector in all the directions by simply moving his or her hand within the tracking domain. An ellipsoid indicates the boundaries of the working volume at any given height as shown in the figure. Boundaries become larger by moving the hand upward following the vertical axis of the Leap Motion controller, and vice versa. A grid on the bottom shows the position of the floor w.r.t. to the end-effector.

Additional visual feedbacks are associated to other relevant aspects of the interaction, like the transition between states or tracking lost.

As said, the interaction schema described above was obtained in [CCG*17] through the evaluation of different alternatives with a trial & error process. For instance the orientation of the tracked hand could be directly used to control rover direction. However, despite the intuitiveness, this approach proved to be far less accurate than the one that was ultimately adopted. Another choice that was made regarded the position and shape of the boxes that identify the control commands for the rover. With respect to a configuration where the entire interaction area is used, the cross configuration

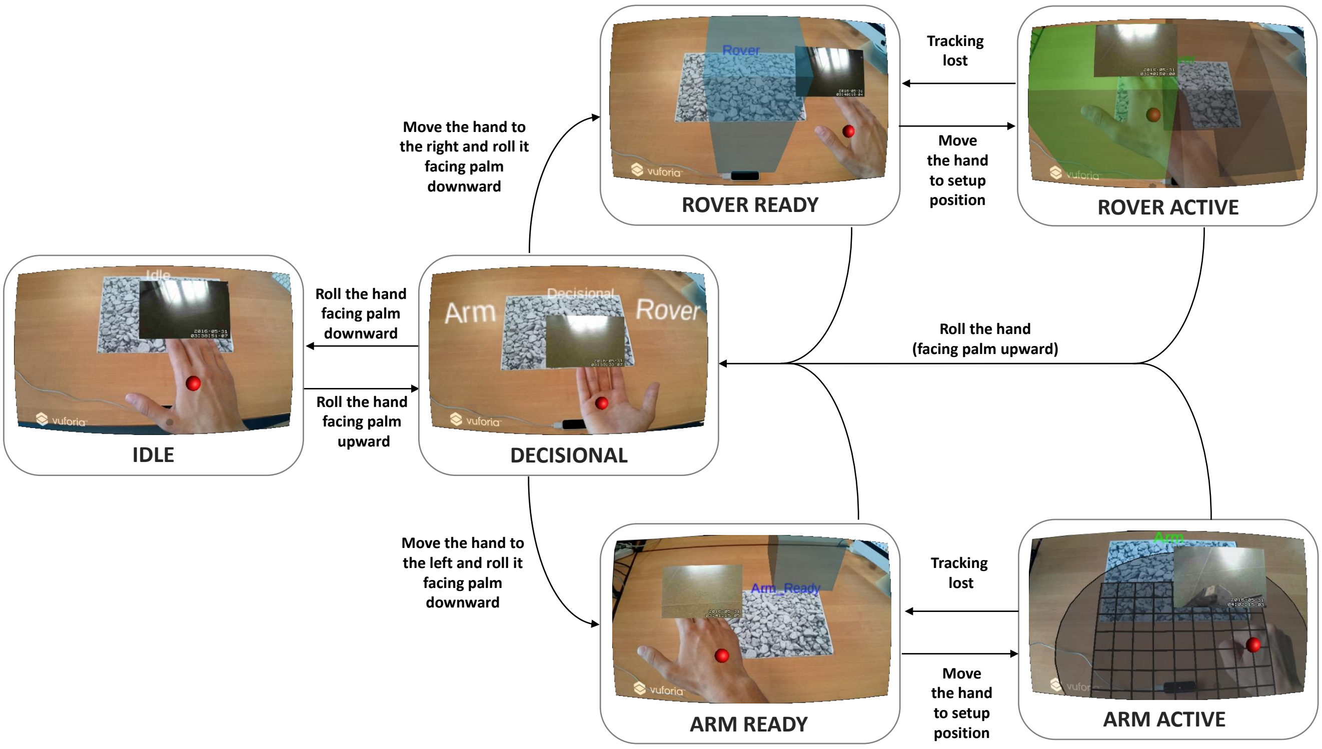


Figure 2: *State diagram of the interface for the control of the rover and the robotic arm and screenshots showing the states of the AR interface.*

that was ultimately chosen optimizes the separation between different commands. Moreover, this configuration provided a higher flexibility compared to an alternative configuration based on multiple control boxes, each designed to set a specific velocity. Concerning the control of the robotic arm, forward kinematics was experimented as well. In the end, inverse kinematics was selected, since tele-operation proved to be faster with it.

3.3. Robot team

As said, the robot team considered in this work is composed by a rover and a robotic arm.

The rover has been built by assembling the Lynxmotion Aluminum 4WDI Robot Kit. The operator can act on two control parameters, namely speed and direction (Figure 3), in order to drive it.

The robotic arm (Lynxmotion AL5D) is a 5-DoF manipulator which include a base, a shoulder, an elbow, and two wrists joints (Figure 3). In addition, a servo motor is employed to open and close a gripper mounted at the end of the kinematic chain. Hand's position tracked by the Leap Motion controller is exploited to define the angles of the first three joints (base, shoulder, elbow) through an inverse kinematics solver that maps the spatial coordinates of the hand's palm to the end-effector of the robotic arm. The orientation of the two wrist joints are controlled using forward kinematics by assigning the roll and the pitch of the operator's hand. Gripper status (open/close) are controlled by opening/closing the palm.

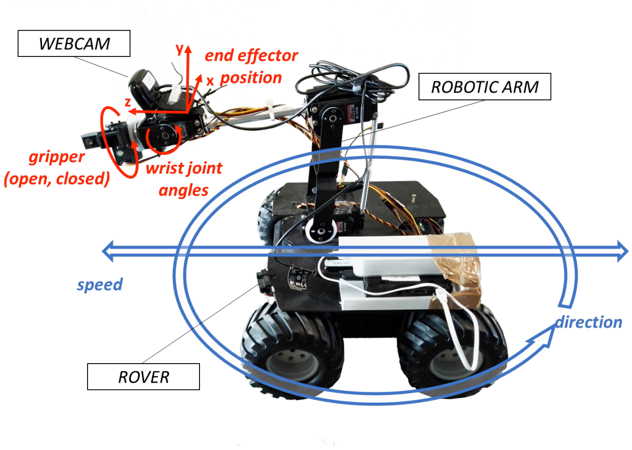


Figure 3: *Robot team controlled using the proposed interface.*

Rover and arm are connected to the network via an Arduino board. The remote environment the robot is moving into can be inspected through a webcam (Logitech C525) mounted on robotic arm, which streams the video to the remote operator on a separate network connection managed by a Raspberry Pi.

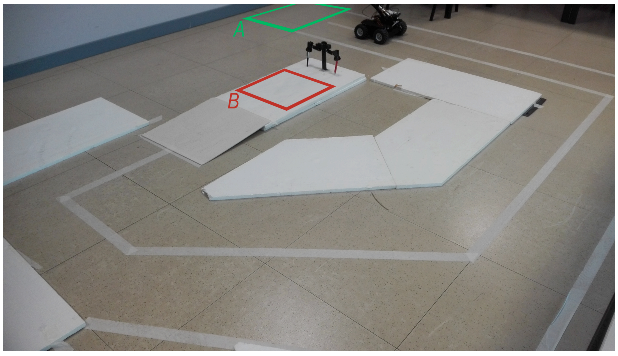


Figure 4: *Configuration of the environment for the experiments.*

4. Experimental results

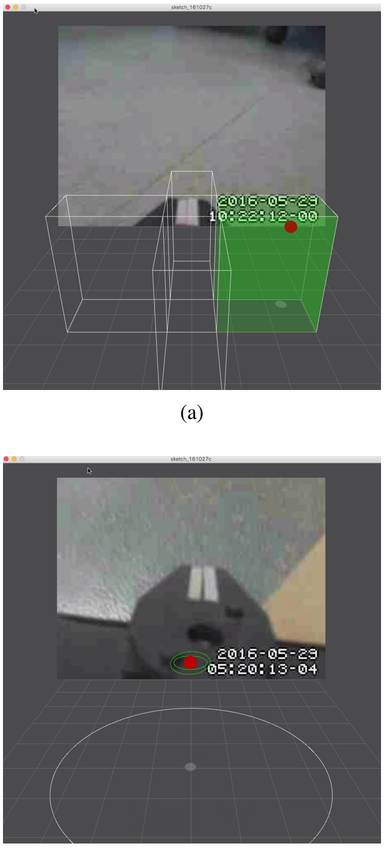
In order to test the effectiveness of the designed interface, a user study has been carried out by asking 15 volunteers (selected among University students) to perform a *reach-and-pick* task. The task included two sub-tasks, each involving a different robot of the team.

In the first sub-task, participants were requested to tele-operate the rover over a pre-defined path created by positioning physical elements on the ground (styrofoam boxes to define boundaries, paper tape to draw median strip, plastic ramp, etc.) as shown in Figure 4. The sub-task was considered as completed once the rover reached a specific position marked by a colored rectangle at the end of a ramp. This position represented the starting configuration for the second sub-task. In this sub-task, the robotic arm had to be used. Participants were requested to tele-operate the arm and explore the environment, looking for two colored objects located to the left and to the right. Afterwards, they were invited to use the gripper to pick up the first object found (this way, experience was comparable for all the participants). The second sub-task was considered as completed as soon as the object was grabbed with the gripper.

In the following, the two-subtasks will be referred to as *rover* and *arm*. Each participant was asked to carry out the overall task both with the desktop interface presented in [CCG*17] (later labeled *DI*) and the wearable interface (*WI*). The two screenshots in Figure 5a and Figure 5b represent the DI when the controlled robots are the rover and the arm, respectively. In order to the limit learning effect, the interface to start with was selected in a random way. Two videos showing the execution of the task with the two interfaces are available at `https://drive.google.com/open?id=0B27BuRM-44ZhanN5Ujk2LUZuLWM` (*DI*) and `https://drive.google.com/open?id=0B27BuRM-44ZhSmlrbUhlaFo2SGc` (*WI*).

Participants were given a certain amount of time before starting the experiment to familiarize with robot control, by letting them freely operate both the rover and the arm using the two interfaces. During this time, they were shown the functioning of the Leap Motion controller and gestures supported. Moreover, they were shown how to interpret the information provided by the two interfaces, and were asked to get accustomed with tele-operation by controlling robot's parts with the sole support of the video streamed by the on-board camera.

In order to compare performance with the two interfaces, both objective and subjective evaluations were used. Concerning objective evaluation, the time requested to complete each sub-task was measured. Moreover, during the execution of the task, the number of control commands issued was recorded. For the rover sub-task, control commands were easy to distinguish. For the arm sub-task, given the continuous nature of the inverse kinematics-based control mechanism used, a metric based on the number of attempts was adopted, considering the number of times a participant closed the gripper to grab the object (possibly missing it). The objective evaluation was supplemented by a subjective evaluation, which was based on feedback collected through a questionnaire delivered to each participant at the end of the experiment. Participants were also asked to express their preference for the *DI* or the *WI* for each sub-task.



(a)

(b)

Figure 5: *Screenshots showing the desktop interface in [CCG*17] when controlling (a) the rover and (b) the arm.*

Results achieved with the objective evaluation in terms of completion time and control commands are reported in Figure 6 and Figure 7, respectively. Statistical significance was analyzed by running paired samples t-tests with significance level $\alpha = 0.05$.

According to Figure 6, the difference in the completion time for the rover sub-task with the two interfaces is not significant (this result is confirmed by the statistical analysis). However, for the arm sub-task, participants were significantly faster ($P = 0.002$) with the

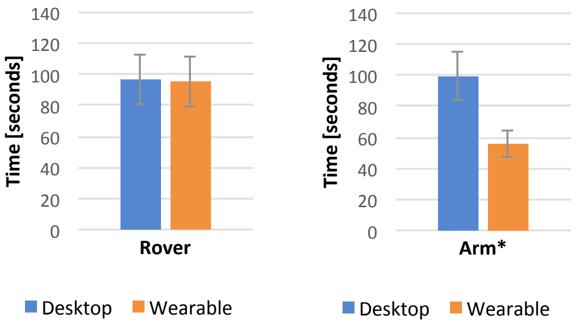


Figure 6: *Completion time for the two sub-tasks (performed by tele-operating the rover and the robotic arm) by using the DI and the WI. Bars height show the average value (lower is better). Standard deviation is also reported. The * symbol is used to indicate when differences are statistically significant.*

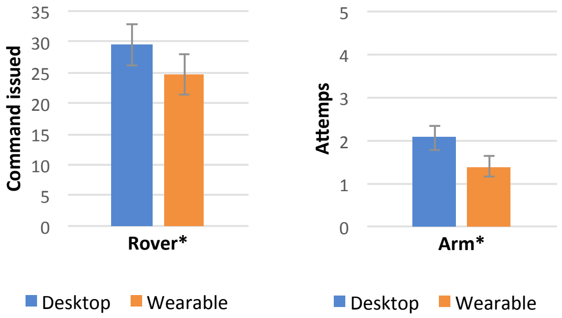


Figure 7: *Control commands for the two sub-tasks (performed by tele-operating the rover and the robotic arm) by using the DI and the WI. Bars height show the average value (lower is better). Standard deviation is also reported. The * symbol is used to indicate when differences are statistical significant.*

WI rather than with the *DI*. Average completion time was 55.73 and 99.53 seconds for the *WI* and *DI*, respectively. Standard deviation was 16.54 for the *WI* and 16.54 for the *DI*, meaning that with the proposed interface, differences among participants are smoothed.

When control commands are considered (Figure 7), it can be easily observed that the number of commands issued / the number of attempts made with the *WI* is significantly smaller than with the *DI* for both the rover and arm sub-tasks ($P = 0.030$ and $P = 0.001$, respectively).

In summary, based on objective observations it appears that when it comes to explore the surrounding environment and pick up objects using the robotic arm, the *WI* is superior to the *DI* both for what it concerns time and number of attempts required. When driving the rover, the *WI* appears to be capable to reduce the number of commands to be issued.

As said, after the execution of the whole task with both the interfaces, a subjective evaluation was performed using a questionnaire. Questionnaire was organized in five sections, encompassing a rather broad set of questions which were selected and adapted from different analysis tools to reach a comprehensive and multi-faceted understanding of participants' experience. Questions were expressed in the form of statements to be evaluated (for each sub-task and for each interface) on a five-point Likert scale from 0 (strong disagreement) to 4 (strong agreement). The first sections was aimed at investigating usability of the two interfaces based on attributes defined by Nielsen [NM90]. Thus, five questions were asked to evaluate learnability, efficiency, memorability, (recovery from) errors and satisfaction. The second and third sections was developed based on a questionnaire presented in [SPT*15], which is meant to assess usability for handheld Augmented Reality devices and applications. In particular, 10 questions investigating the comprehensibility (level of understanding of the information presented) of the interface were included. Moreover, 4 questions focusing on manipulability (ease of handling) of the interface were selected. The fourth section was developed based on the evaluation approach adopted in a work with a similar aim and specifically includes 8 questions on robot control [HIII11]. The last section consists of 9 questions based on usability heuristics defined in [Sch14, GMBP15]. For sake of readability, results concerning Nielsen's attributes of usability are illustrated in Figure 8. Results collected with the four remaining sections of the questionnaire are given in Table 1–4, in order to provide the reader with actual formulation for each statement. Statistical significance of data collected was checked by using the same procedure adopted for the objective evaluation. Statistically significant results are marked with a * symbol.

According to Figure 8 (upper part), for the rover sub-task performance appears to be slightly better in terms of learnability, efficiency and memorability when the *DI* interface is used, although results are not statistically significant. However, when the arm sub-task is considered, performance with the *WI* is largely better than with the *DI* for all the attributes (results are all significant, except for memorability). The above difference can be easily appreciated in Figure 8 (lower part).

With respect to comprehensibility (Table 1), an interesting result (which is also the only one that is statistically significant for this section) is that, for the arm sub-task, the *DI* required a higher mental effort (concentration) than the *WI* to be operated. This result could be a confirmation of the benefits brought by affordances associated with the visualization of augmented contents on the operator's hand that is possible in the *WI*. No significant difference was found between the two interfaces concerning, among others, appropriateness, readability, meaningfulness and responsiveness of information displayed.

Results concerning interface manipulability (Table 2) confirm findings obtained using Nielsen's methodology. In particular, statistical significance was verified only for the question concerning easiness of issuing robot control commands and easiness of the interface for the arm sub-task. Participants found it easier to issue commands with the *WI* rather than with the *DI*. Moreover, the *WI* was found to be simpler and less complicated than the *DI*.

Interesting findings come from the fourth section regarding robot control (Table 3). In fact, significantly different results are obtained

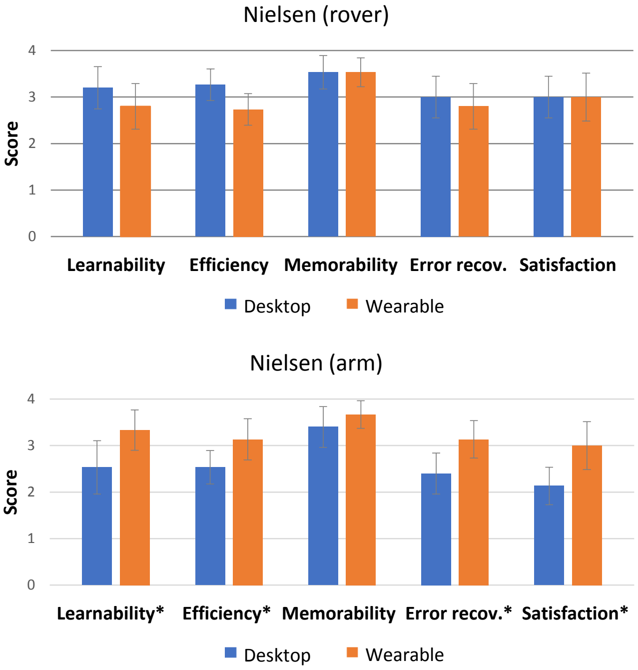


Figure 8: *Subjective results concerning Nielsen's attributes of usability for the two sub-tasks using the DI and the WI. Bars height show the average value (higher is better). Standard deviation is also reported. The * symbol is used to indicate when differences are statistically significant.*

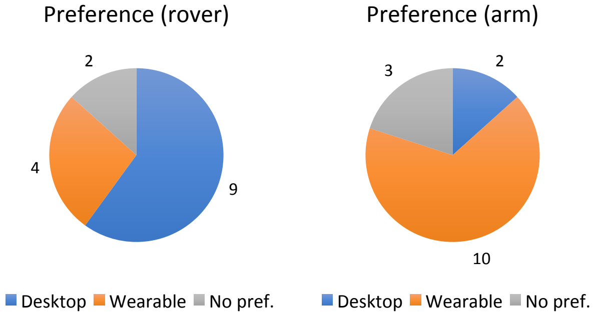


Figure 9: *Preference expressed by participants for a specific interface in the execution of the two sub-tasks.*

depending on the sub-task considered. For the rover-task, participants found it easier to control the robot with the *DI* than with the WI and stated that they would like to control the robot that way. Opposite results are obtained when considering the arm sub-task. Moreover, for the arm sub-task, statistical significance was verified also for statements concerning confidence in robot control, consistency and interface suitability for the specific operation, confirming the performance of the WI.

These opposite trend can be observed also for some statements in the last section (Table 4). In fact, when carrying out the rover sub-task, participants felt more confused with the WI than with the *DI*. Conversely, for the arm sub-task. Section five provides other interesting insights. First, it indicates that, in the rover task, participants judged the layout for the visualization of 3D contents more appropriate for the *DI* than for the WI. This could be due to the fact that, sometimes, the window showing the video streamed by the robot's camera overlaps the control regions in the interface. This phenomenon was far less frequent in the execution of the arm sub-task. Second, as expected, participants felt their eyes were much more tired after having operated the WI than the *DI*.

In summary, subjective results confirm the findings obtained through the objective evaluation, clearly showing a larger appreciation for the WI to control the robotic arm. Subjective observations gathered for the rover sub-task do not allow to identify a neat advantage in the use of one of the two interfaces. In fact, if the lower number of commands to be issued could make the WI more attractive for tele-operating the rover, eye fatigue could pull in the other direction.

However, helpful indications can be obtained by considering the specific feedback provided by participants concerning their preference for a specific interface to carry out the two sub-tasks. As shown in Figure 9, participants expressed a clear preference for the *DI* when it comes to tele-operate the rover. When the robotic arm needs to be controlled, preference is, conversely, almost completely for the WI.

5. Conclusions and future work

In this work, a wearable AR interface aimed to support an operator during hand tracking-based tele-operation of remote robot teams has been presented. AR is used to provide the operator with an immediate feedback about the position of his or her hand in the working space defined by the tracking system. The proposed interface has been compared with a similar solution in which AR contents are displayed on the screen of a desktop computer. Experimental results showed a marked preference for the proposed interface for arm manipulation operations. Conversely, desktop-based AR was preferred for driving operations. Given the encouraging results obtained, future works will be addressed to further investigate operators' preference for a specific interface by considering a wider set of tasks to be accomplished. Moreover, alternative ways to issue control commands as well as to visualize working volume, interaction regions, etc. will be experimented. Lastly, the applicability of AR to deliver an effective interaction feedback with other input mechanisms will be investigated.

References

[AAK15] Ahn J., Ahn E., Kim G. J.: Robot mediated tele-presence through body motion based control. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on* (2015), IEEE, pp. 463–468. 2

[AMP13] Arango C. A., Martinez J. R., Pérez V. Z.: Master-slave system using kinect and an industrial robot for teleoperations. In *Health*

Table 1: *Statements used in the subjective evaluation (second section of the questionnaire) and average results in a scale from 0 (strong disagreement) to 4 (strong agreement). The * symbol is used to indicate when differences are statistically significant.*

	Rover		**Arm**	
Comprehensibility of information	**Desktop**	**Wearable**	**Desktop**	**Wearable**
I think that interaction requires a lot of mental effort and concentration	1,40	1,67	2,00*	1,20*
I thought the amount of information displayed on screen was appropriate	3,40	3,07	3,27	3,33
I thought that the information displayed on screen was difficult to read and interpret	0,67	0,80	0,73	0,87
I felt that the information display was responding fast enough	2,67	2,67	2,40	2,67
I thought that the information displayed on screen was confusing	1,00	1,20	1,27	0,93
I thought that the meaning of information displayed on screen was self-explanatory	3,00	3,00	2,80	3,07
I thought the words and symbols on screen were easy to read	3,80	3,60	3,73	3,67
I felt that the information displayed on the screen was flickering too much	0,73	1,07	0,87	1,00
I thought that the information displayed on screen was consistent	3,27	3,27	3,20	3,20

Table 2: *Statements used in the subjective evaluation (third section of the questionnaire) and average results in a scale from 0 (strong disagreement) to 4 (strong agreement). The * symbol is used to indicate when differences are statistically significant.*

	Rover		**Arm**	
Manipulability of the interface	**Desktop**	**Wearable**	**Desktop**	**Wearable**
I thought that interaction requires a lot of body muscle effort	0,87	0,87	0,93	0,87
I felt that using the application was comfortable for my arm and hands	2,93	2,93	2,73	2,73
I found it easy to issue commands through the interface	3,13	2,87	2,60*	3,26*
I thought the operation of the interface is simple and uncomplicated	3,27	3,33	2,73*	3,33*

Table 3: *Statements used in the subjective evaluation (fourth section of the questionnaire) and average results in a scale from 0 (strong disagreement) to 4 (strong agreement). The * symbol is used to indicate when differences are statistically significant.*

	Rover		**Arm**	
Robot control	**Desktop**	**Wearable**	**Desktop**	**Wearable**
It was easy to control the robot this way	3,20*	2,60*	2,40*	3,13*
I controlled the robot with confidence	3,13	2,87	2,27*	3,13*
I would like to control the robot this way	3,20*	2,40*	2,40*	3,20*
A lot of training is necessary for using this interface	0,47	0,53	0,80	0,67
Many people could control the robot easily using this interface	3,20	3,00	2,73	2,93
I could control the robot as I expected	3,27	2,87	2,40*	3,07*
The task was very difficult to execute	0,80	1,13	1,20	1,00
The interface is suitable to control the robot and execute the task	3,33	2,87	2,93*	3,47*

Table 4: *Statements used in the subjective evaluation (fifth section of the questionnaire) and average results in a scale from 0 (strong disagreement) to 4 (strong agreement). The * symbol is used to indicate when differences are statistically significant.*

	Rover		**Arm**	
Other heuristics	**Desktop**	**Wearable**	**Desktop**	**Wearable**
I felt that my arm or hand became tired after using the interface	1,20	1,20	1,27	1,13
I felt that my eyes became tired after using the interface	0,27*	1,87*	0,27*	1,87*
Layout for the visualization of 3D contents is visually pleasant	3,13	2,60	3,00	2,93
Layout for the visualization of 3D contents is efficient	3,47*	2,80*	3,13	2,93
The system provided me with a proper feedback about what I was working on	3,07	2,93	2,67	3,00
I was not confused or lost while performing the task	3,40*	2,80*	2,80*	3,40*
I was not requested to memorize things unnecessarily	3,87	3,93	3,80	3,93
The first-time experience with the interface is encouraging	3,33	3,27	2,93	3,33
It was easy to find the desired control options at any time	3,40	3,27	3,13	3,33

Care Exchanges (PAHCE), 2013 Pan American (2013), IEEE, pp. 1–6. 2

[BCP*17] Bonaiuto S., Cannavò A., Piumatti G., Paravati G., Lamberti F.: Tele-operation of robot teams: a comparison of gamepad-, mobile device-and hand tracking-based user interfaces. 1–6. 1, 2

[BGG*14] Bassily D., Georgoulas C., Guettler J., Linner T., Bock T.: Intuitive and adaptive robotic arm manipulation using the leap motion controller. In *ISR/Robotik 2014; 41st International Symposium on Robotics; Proceedings of* (2014), VDE, pp. 1–7. 2

[CCG*17] Cancedda L., Cannavò A., Garofalo G., Lamberti F., Montuschi P., Paravati G.: Mixed reality-based user interaction feedback for a hand-controlled interface targeted to robot teleoperation. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (2017), Springer, pp. 447–463. 1, 2, 3, 4, 6

[CJLY10] 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 Transactions On Mechatronics 15*, 1 (2010), 17–26. 2

[CjPMlYf09] Cheng-jun D., Ping D., Ming-lu Z., Yan-fang Z.: Design of mobile robot teleoperation system based on virtual reality. In *Automation and Logistics, 2009. ICAL'09. IEEE International Conference on* (2009), IEEE, pp. 2024–2029. 3

[CP15] Crainic M.-F., Preitl S.: Ergonomic operating mode for a robot arm using a game-pad with two joysticks. In *Applied Computational Intelligence and Informatics (SACI), 2015 IEEE 10th Jubilee International Symposium on* (2015), IEEE, pp. 167–170. 2

[EN16] Erdelj M., Natalizio E.: Uav-assisted disaster management: Applications and open issues. In *Computing, Networking and Communications (ICNC), 2016 International Conference on* (2016), IEEE, pp. 1–5. 1

[FCC05] Faisal S., Cairns P., Craft B.: Infovis experience enhancement through mediated interaction. 1

[FK16] Frank J. A., Kapila V.: Towards teleoperation-based interactive learning of robot kinematics using a mobile augmented reality interface on a tablet. In *Control Conference (ICC), 2016 Indian* (2016), IEEE, pp. 385–392. 2, 3

[GMBP15] Gale N., Mirza-Babaei P., Pedersen I.: Heuristic guidelines for playful wearable augmented reality applications. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (2015), ACM, pp. 529–534. 7

[GS08] Guo C., Sharlin E.: Exploring the use of tangible user interfaces for human-robot interaction: a comparative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2008), ACM, pp. 121–130. 1

[HIII11] Hashimoto S., Ishida A., Inami M., Igarashi T.: Touchme: An augmented reality based remote robot manipulation. In *21st Int. Conf. on Artificial Reality and Telexistence, Proc. of ICAT2011* (2011). 2, 3, 7

[HPC*15] Hortal E., Planelles D., Costa A., Iáñez E., Úbeda A., Azorín J. M., Fernández E.: Svm-based brain–machine interface for controlling a robot arm through four mental tasks. *Neurocomputing 151* (2015), 116–121. 2

[JWPD13] Jiang H., Wachs J. P., Pendergast M., Duerstock B. S.: 3d joystick for robotic arm control by individuals with high level spinal cord injuries. In *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on* (2013), IEEE, pp. 1–5. 1

[KRO*14] Kim M. K., Ryu K., Oh Y., Oh S.-R., Kim K.: Implementation of real-time motion and force capturing system for tele-manipulation based on semg signals and imu motion data. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on* (2014), IEEE, pp. 5658–5664. 2

[LRRMRC*16] Lovon-Ramos P. W., Ripas-Mamani R., Rosas-Cuevas Y., Tejada-Begazo M., Mogrovejo R. M., Barrios-Aranibar D.: Mixed reality applied to the teleoperation of a 7-dof manipulator in rescue missions. In *Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR), 2016 XIII Latin American* (2016), IEEE, pp. 299–304. 3

[LSC10] Levine S. J., Schaffert S., Checka N.: Natural user interface for robot task assignment. 2, 3

[MBGH14] Mateo C., Brunete A., Gambao E., Hernando M.: Hammer: An android based application for end-user industrial robot programming. In *Mechatronic and Embedded Systems and Applications (MESA), 2014 IEEE/ASME 10th International Conference on* (2014), IEEE, pp. 1–6. 1

[MLH*11] Malkin J., Li X., Harada S., Landay J., Bilmes J.: The vocal joystick engine v1. 0. *Computer Speech & Language 25*, 3 (2011), 535–555. 2

[NM90] Nielsen J., Molich R.: Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1990), ACM, pp. 249–256. 7

[NPM09] Neto P., Pires J. N., Moreira A. P.: Accelerometer-based control of an industrial robotic arm. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (2009), IEEE, pp. 1192–1197. 2

[PBAR15] Peppoloni L., Brizzi F., Avizzano C. A., Ruffaldi E.: Immersive ros-integrated framework for robot teleoperation. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on* (2015), IEEE, pp. 177–178. 2, 3

[RLGSFL*11] Rodríguez Lera F. J., García Sierra J. F., Fernández Llamas C., Matellán Olivera V., et al.: Augmented reality to improve teleoperation of mobile robots. 2, 3

[SAB*07] Skubic M., Anderson D., Blisard S., Perzanowski D., Schultz A.: Using a hand-drawn sketch to control a team of robots. *Autonomous Robots 22*, 4 (2007), 399–410. 1

[Sch14] Schaeffer S. E.: Usability evaluation for augmented reality. 7

[SKS*06] Shimizu N., Koizumi N., Sugimoto M., Nii H., Sekiguchi D., Inami M.: A teddy-bear-based robotic user interface. *Computers in Entertainment (CIE) 4*, 3 (2006), 8. 2

[SPT*15] Santos M. E. C., Polvi J., Taketomi T., Yamamoto G., Sandor C., Kato H.: Toward standard usability questionnaires for handheld augmented reality. *IEEE computer graphics and applications 35*, 5 (2015), 66–75. 7

[SSF*00] Suzuki T., Sekine T., Fujii T., Asama H., Endo I.: Cooperative formation among multiple mobile robot teleoperation in inspection task. In *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on* (2000), vol. 1, IEEE, pp. 358–363. 1

[SZS10] Sauer M., Zeiger F., Schilling K.: Mixed-reality user interface for mobile robot teleoperation in ad-hoc networks. *IFAC Proceedings Volumes 43*, 23 (2010), 77–82. 2, 3

[UIA13] Ubeda A., Ianez E., Azorin J. M.: An integrated electrooculography and desktop input bimodal interface to support robotic arm control. *IEEE Transactions on Human-Machine Systems 43*, 3 (2013), 338–342. 2

[VP15] Venna T. V. S. N., Patel S.: Real-time robot control using leap motion a concept of human-robot interaction. ASEE. 2

[WBRF13] Weichert F., Bachmann D., Rudak B., Fisseler D.: Analysis of the accuracy and robustness of the leap motion controller. *Sensors 13*, 5 (2013), 6380–6393. 2

[YON17] Yew A., Ong S., Nee A.: Immersive augmented reality environment for the teleoperation of maintenance robots. *Procedia CIRP 61* (2017), 305–310. 3

[YYDA15] Yang Y., Yan H., Dehghan M., Ang M. H.: Real-time human-robot interaction in complex environment using kinect v2 image recognition. In *Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2015 IEEE 7th International Conference on* (2015), IEEE, pp. 112–117. 2