ORIGINAL ARTICLE

Application of immersive technologies and natural language to hyper-redundant robot teleoperation

Andrés Martín-Barrio[1] · Juan Jesús Roldán[1] · Silvia Terrile[1] · Jaime del Cerro[1] · Antonio Barrientos[1]

Abstract

This work presents an analysis of immersive realities and natural language applied to the teleoperation of hyper-redundant robots. Such devices have a large number of degrees of freedom, so they often exhibit complex configurations frustrating their spatial understanding. This work aims to contrast two hypotheses; first, if immersive interfaces enhance the telepresence and efficiency against conventional ones; and second, if *natural language* reduces workload and improves performance against other conventional tools. A total of 2 interfaces and 6 interaction tools have been tested by 50 people. As a result, immersive interfaces were more efficient, improved situational awareness and visual feedback, and were chosen by 94% of participants against conventional ones. On the other hand, participants performed better using natural language than conventional tools despite having less previous experience with the first ones. Additionally, according to 52% of the population, the preferred interaction tool was a mixed strategy that combined voice recognition and hand gestures. Therefore, it is concluded that immersive realities and natural language should play a very important role in the near future of hyper-redundant robots and their teleoperation.

1 Introduction

In every robotic device, the concept of *Degree of Freedom* (DoF) is very important. It can be defined as the number of movements or independent parameters that define the configuration of a robot (Barrientos et al. 2007). Traditionally, manipulator robots have been designed to have up to 6 DoF because it is the minimum number for which a robot can properly position and orient its end-effector in a three-dimensional (3D) space. However, robots with a higher number of DoF have some advantages against traditional ones: they have higher kinematic skills, and they can recover from a failure in some of their joints and have better abilities to actuate in complex environments with obstacles. Those robots with more than 6 DoF are called *redundant*, while those with more than 12 DoF can be named as *hyper-redundant* (Martín-Barrio et al. 2018b).

Traditionally, hyper-redundant robots have been conceived to be discrete, with a high number of distinguishable links. However, one of the most researched topics in the robotics scope is currently focused on *continuous* robots, especially those made by *soft* materials. These robots are characterized to continuously deform instead of being moved by the concatenation of a set of rigid links. Therefore, compared to traditional ones, they exhibit more flexibility and inherent safety to interact with their environment.

To control any robot, a prior knowledge of a *model* is usually required to extract its behaviour and predict its movements. At a kinematic level, two main problems are usually studied: the direct and inverse kinematics. The *direct kinematics* aims to determine the end-effector position and orientation when the joint values—the *configuration* or *pose*—are known. This problem is relatively easy to be solved for hyper-redundant robots using traditional techniques. On the other hand, the *inverse kinematics* problem is usually more

✉ Andrés Martín-Barrio
andres.mb@upm.es

1 Centre for Automation and Robotics (CAR), CSIC-UPM, Universidad Politécnica de Madrid, c/ José Gutiérrez Abascal, No. 2, 28006 Madrid, Spain

interesting since it aims to solve the configuration that leads to a determined position for its end-effector. This problem has more difficulties in the hyper-redundant scope since such devices offer an infinite number of configurations that lead to a single position and orientation for the end-effector (Chirikjian and Burdick 1994). Choosing the best solution can be a great challenge because it is very relative and dependant on every situation. Several works have focussed on this matter (Collins and Shen 2016; Espinoza et al. 2012; Marques et al. 2009; Martín-Barrio et al. 2018a).

However, commanding the end-effector location of a hyper-redundant robot offers a limited number of applications. Besides, methods aiming to solve the inverse kinematics problem usually do not exploit all the kinematic capabilities for such robots. Thus, *shape controls* emerge as an alternative. This work will focus on the application of this type of control scheme. Under this paradigm, the entire configuration of the robot is controlled to perform tasks such as local avoidance of obstacles, surrounding objects or follow the leader guidance. Therefore, multiple references are needed to control such robots and not just one.

Strategies to adequately define this set of references are usually studied in *telerobotics*. This outlook implies linking a human operator and a robot in order to execute a remote task. Thus, *teleoperation* of the shape of a hyper-redundant manipulator should study how to properly associate a set of user-defined references to move the robot pose (Fig. 1). This field is highly related to *telepresence*, which means that the information about the remote environment is naturally displayed to the operator, implying a feeling of presence at the remote site. A good degree of telepresence guarantees the feasibility of a determined manipulation task (Aracil et al. 2007). Since hyper-redundant robots often exhibit intricate shapes, hindering their spatial understanding, maybe emerging *immersive technologies* can be a good solution to solve this issue.

Fig. 1 Teleoperation of a continuous robot for space exploration using a *Virtual Reality* headset and a *Leap Motion* controller

According to the previous definition, *immersive realities* could be understood as those real or simulated environments in which a perceiver experiences telepresence (Rheingold 1991; Vertut 2013). For example, *Virtual Reality* (VR) provides interaction of virtual elements in virtual environments, *Augmented Reality* (AR) features interaction of virtual elements in real environments, and *Mixed Reality* (MR) offers interaction of real and virtual elements through real or virtual environments (Roldán et al. 2019b). Nowadays, all these realities are achieved using headsets, headphones, glasses or controllers among several other devices, providing stimulus to the human senses including, but not limited to, vision, audition and touch. This feedback can be used within a teleoperation system to properly manage the pose of hyper-redundant robots.

On the other hand, traditional telerobotic systems have often relied on a *master*, intended to command the movements of a remote robot, called *slave*. These master–slave architectures were conceived in this manner to be *transparent*, in other words, to exactly reproduce the remote environment (Aracil et al. 2007). However, they often implied the design and construction of a physical interface to reproduce the real-robot movements. Immersive technologies could replace those physical masters within the virtual world while enhancing the immersion and telepresence. But such transparency will highly rely on the *interaction tools* used to replace those physical architectures. In that context, *natural language* could be used as one possible tool. It is understood as a rich, intuitive mechanism by which humans can interact with systems around them, offering sufficient signal to support robot task planning (Matuszek et al. 2013).

Accordingly, this work aims to contrast two hypotheses: the first one states that immersive interfaces are better in terms of efficiency and situational awareness against conventional ones to teleoperate hyper-redundant robots; the second one states that *natural language* can improve the performance and reduce the workload against conventional tools for the same purpose.

The first main contribution of this work is the comparison of emergent immersive technologies against conventional ones to teleoperate the shape of hyper-redundant manipulators. Furthermore, an analysis of natural language and other tools used to interact within the immersive scenarios has been conducted and compared against conventional tools. Additionally, an experimentation strategy is proposed in order to measure both subjective and objective parameters, relevant to contrast the conjectures. In summary, this work is expected to clear the path to which the research of teleoperation of hyper-redundant, continuous and soft robots should be focused on.

This document will be structured as follows: Sect. 2 introduces the up-to-date state of the art in telerobotics, teleoperation, immersive technologies and natural language.

Section 3 introduces the system setup, in order to contextualize the following chapters. Section 4 explains of the developed interfaces to contrast the benefits of using immersive realities or conventional ones. Section 5 explains the interaction tools and strategies used to perform in immersive scenarios for further comparisons. Section 6 presents the experiments performed to contrast the hypotheses and the results, and Sect. 7 outlines the conclusions derived from this work.

2 State of the art

This section presents the up-to-date research in robot teleoperation, specifically in human–robot interfaces and interaction strategies, intended for many types of robots and purposes. Hopefully, this will help to clarify the current research, potential applications and limitations related to this work.

One of the major challenges in robot teleoperation is the difficulty of achieving a precise and reliable representation of the environment. Remote perception and manipulation are affected by factors such as limited *field of view* (FOV), orientation, camera viewpoint, depth perception, degraded video image, time delay and motion. Researchers have shown that enhancing teleoperator feedback plays an important role in decreasing task difficulty and creating a greater sense of operator immersion in a teleoperation environment (Chen et al. 2007). Accordingly, some authors have studied the strengths and weaknesses of conventional 2D interfaces against immersive 3D ones. Some of them predict that immersive interfaces can help users to achieve a higher level of understanding of the robot's point of view and situational awareness (Hedayati et al. 2018).

Several works related to mobile robotic applications agree at selecting virtual and augmented reality interfaces in order to improve performance (Cheng-jun et al. 2009; Kot and Novák 2018; Mostefa et al. 2015; Nielsen et al. 2007). One example of enhancing situational awareness is *Flying Frustum*, an AR interface aimed to control semi-autonomous *Unmanned Aerial Vehicles* (UAV) using a 3D interactive printout of the terrain. In this work, UAV operators could use pen-based interactions to input flight paths and commands to the UAVs by sketching directly on the physical topographical model of the terrain (Li et al. 2015).

Moreover, many works claim that VR technologies applied to robotics can reduce the cost of expensive specialized systems in manufacturing (Bugalia et al. 2015; Lipton et al. 2018; Roldán et al. 2019a).

Concerning the teleoperation of service robots, immersive realities can be used for surveillance and remote interventions in our homes. (Labonte et al. 2010) conducted an experiment in which two MR visualization modalities were compared with standard video-centric and map-centric perspectives. As a result, novice operators experienced a higher workload and were less efficient than expert ones. Regarding the interfaces, the performance was significantly inconclusive. According to the authors, this is due to a high complexity of different factors such as the operation environment, the mission's objectives, the robots' capabilities or the choice of performance metrics. In another work, teleoperation enabled operators to establish effective communications between humanoid robots and people in order to avoid the automatic recognition difficulties of spoken language (Koizumi et al. 2006).

In the hyper-redundant robotics scope, teleoperation schemes often follow traditional master–slave architectures. This is reasonable since these robots with a large number of DoF are usually perceived to need a replica with the same number of DoF in order to exploit all their kinematic skills. However, using masters and slaves with different number of DoF is a common practice. For example, a *Phantom Omni* haptic device (master) was used to accurately teleoperate the end-effector of an articulated snake with 12 DoF (slave) for minimally invasive surgical procedures. In these situations where the number of DoF of the master is lower than in the slave, it is necessary to obtain a mapping relation between each other (Ren et al. 2017). Another example uses the same master to teleoperate a tendon-driven continuous robot by using a VR interface in real time, also intended for surgical procedures. This research underlines that the kinematic structure of the master is very dissimilar to the kinematic model of the continuum robot, so the teleoperation strategy may not be the best (Bhattacherjee et al. 2018). Such dissimilarities were also studied trying to teleoperate a redundant continuous manipulator using a non-redundant rigid-link master. These differences were minimized using feedback linearizing task-space controllers (Kapadia et al. 2012). In all these works, the inverse kinematics is interrelated between the master and slave for the end-effector only. Then, only the inverse kinematics problem is solved, but shape controls are not managed. This work will study such whole-body movements to be the ones that fully take advantage of the kinematic capabilities of these robots for remote operations.

Different interaction strategies are used to teleoperate robots, varying from traditional and mechanic peripherals to multimodal ones. On the one hand, physical tools vary from game controllers, electronic devices, keyboards, mouse devices, joysticks or *Phantom* haptic devices, among others. Depending on the field of application and the robot kinematics, one can be chosen against another. For example, to fly a UAV, specific physical controllers with buttons and axis are commonly used (Hedayati et al. 2018).

Human gestures without physical intermediaries are another way of controlling robots in teleoperation. They

can be advantageous because are easy to use, robust, fast and can be used in a wide range within the field of view (Hu et al. 2003). Such interfaces are oriented to tasks that can be performed hands free. In fact, some authors claim a higher prominence of devices like the *Leap Motion* sensor for the future of human–machine interaction applications (Mantecón et al. 2016). Some works describe vision system interfaces that track human gestures for relatively simple robot commands such as up-down, stop, turn, approach and go (Frigola et al. 2003). Others developed a robotic wheelchair controlled using face movements (Kuno et al. 1999). Moreover, it is worth highlighting that these strategies are also used to control UAV by using body postures (Mashood et al. 2015). Usually, vocabulary rules or ad hoc methods are used to select hand gestures. Some authors affirm that analytical approaches are better to design gesture vocabularies for multiple objectives for psycho-physiological and gesture recognition factors (Stern et al. 2007). Accordingly, gesture-based techniques can be a valid alternative to physical master–slave architectures since they can be cheaper and more intuitive (Shirwalkar et al. 2013).

Voice commands are also useful especially in environments where manual control is difficult or when both hands of the operator are busy, as in the case of *ROBTET*, a robotic system for maintaining live power lines. In such a system, users use voice commands, while both hands perform other control tasks (Aracil et al. 2002). Other projects compared the utility of manual and speech inputs for several UAV control tasks. Results showed that using speech was significantly better than manual inputs in terms of task completion time, accuracy, flight and navigation measures and pilot ratings (Draper et al. 2003). Voice commands have been used in telerobotic surgical systems as well, as in *ARTEMIS*, an Advanced Robotics and Telemanipulator System for Minimally Invasive Surgery. The main components are two master–slave units guiding the surgical instruments and a remotely controlled endoscope guiding system. This last one can be guided by a simple joystick, voice control or automatic camera tracking (Rininsland 1999). In addition, voice can be a powerful tool to use in teleoperation of humanoid robots for being convenient and simple, with less workload and learning time involved (Lu et al. 2010; Poncela and Gallardo-Estrella 2015).

Natural-language interfaces can be applied to make the use of intelligent robots more flexible, such as in the case of the autonomous mobile two-arm robot *KAMRO* (Lueth et al. 1994). Combination of gesture and voice strategies gives place to multimodal interfaces, providing a large range of interactions, intuitive to humans and advantageous to service robots (Yoshizaki et al. 2001). There is evidence that multimodal displays and input controls have great potential towards improving the teleoperation performance (Chen et al. 2007). In fact, some authors claim that these approaches increase awareness of surroundings, cue visual attention and convey a variety of complex information, especially when the visual channel is heavily loaded (Shilling and Shinn-Cunningham 2002).

3 System setup

The system arrangement will rely on the information exchanged between the user and the robot in both directions. Thus, communication is named *bilateral* (Fig. 2). In other words, the user sends information to operate the robot in one direction, whereas the robot transmits information to the user in the other direction in order to report its current pose, both in real time. As a reminder, a *real-time* system can be understood as an information processing activity that responds to an externally generated stimulus in a previously determined frame of time (Young 1982). In the presented system, this bilateral communication performs a cycle with a frequency of 10 Hz, which depending on the application can be acceptable or not. In this work, it was high enough to test the capabilities of the interfaces that will be explained in Sect. 4.

The bilateral controller implemented in this work follows a position–position scheme. Thus, the robot will move to a determined position previously commanded by the user. Although this scheme is very simple, it is the best suited when the master–slave size relation is similar, there are no limitations in the slave speed, a high level of telepresence is required, and the task involves a high number of degrees of freedom (Ferre et al. 2007).

In this work, the hyper-redundant and soft manipulator *Kyma* will be used to conduct the experiments (Martín-Barrio et al. 2019) (Fig. 3). Its structure consists of a set of bellows that allow both flexion and prismatic movements. It is divided into 4 sections, each one actuated by 3 angularly equidistant tendons and motors. Therefore, teleoperating this robot will rely on 4 input positions. Moreover, the robot is also equipped with a gripper for pick and place operations.

Additionally, a computer is used to host the interfaces and control the robot movements. It uses different peripherals to transmit or capture information from the user
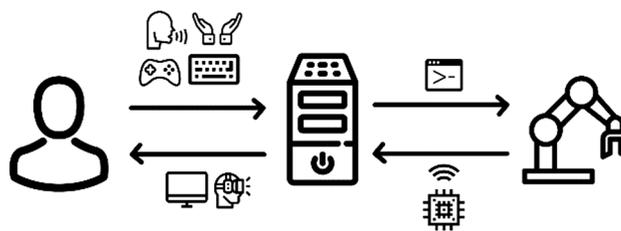

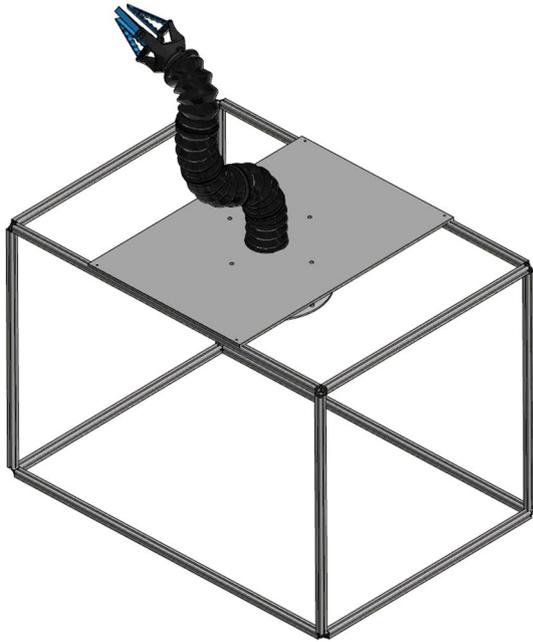**Fig. 2** Bilateral scheme for the robot teleoperation

Fig. 3 Kyma, a soft continuous manipulator robot actuated by tendons

such as a display monitor, a VR headset, a keyboard or a gaming controller. The robot feedback to the user will be visual, so it will rely on displaying its shape, obtained from the onboard sensory system based on *Inertial Measurement Units* (IMUs).

Therefore, the developed setup corresponds to *Mixed Reality* (MR), since interaction occurs with both real and virtual elements through real and virtual environments. This MR can be achieved from a *Virtual Reality* (VR) perspective, by introducing the user in a virtual world, or from an *Augmented Reality* (AR) one, by overlapping virtual elements over the real scene. In this work, the developments have been made in MR from both perspectives using the *HTC Vive* VR headset and the *Hololens* AR glasses, respectively, although the experiments have been entirely conducted on the first ones.

It is worth highlighting that virtual and augmented environments are not competitors, but rather different perspectives for different purposes and applications. Usually, AR is thought to be useful in local environments, enriching the real world, while VR is oriented to replicate remote or imaginary places by immersing the user in the environment. For example, teleoperating a robot in a nuclear reactor (Dutta et al. 2017), or in a remote place such as the moon (Britton et al. 2015), is an application better suited for VR. However, teleoperating a robot in a surgery room (Pessaux et al. 2015), or in a cooperative environment for in situ assemblies (Michalos et al. 2016), are perspectives more appropriate for AR.

4 Human–robot interfaces

This work aims to determine whether immersive interfaces are better or worse than conventional ones to teleoperate hyper-redundant robots. Hence, both interfaces have been created based on the same constitutive elements in *Unity3D*, a cross-platform real-time engine. A *human–robot interface* can be understood as a set of elements that enable the communication and interaction between a person and a robotic system. In this context, a good interface would be clear, concise, familiar, responsive, consistent, aesthetic, efficient and forgiving. Binding such features will give place to an efficient interface while also providing high situational awareness.

Teleoperation of hyper-redundant robots, similarly to multi-robot systems, usually requires a relatively high number of inputs to appropriately manage their shape (Roldán et al. 2017). And it is important to highlight that human beings perform better paying full attention to one thing at one time (Raskin 2000). Therefore, choosing the most convenient interface for this purpose is very important in order to maximize efficiency. In the context of teleoperating robots, e*fficiency* will be defined as the ability to maximize precision and speed by commanding a robot to perform a determined task.

In this context, the number of inputs will be highly dependent on the number of degrees of freedom, but not always strictly the same. For example, several hyper-redundant devices such as continuous or soft robots can be considered as under-actuated (Martín-Barrio et al. 2018b). Such systems have a higher number of degrees of freedom than actuators, so their teleoperation will usually rely on a *model* to estimate their pose when being teleoperated. In the case of *Kyma* (Fig. 3), it has 4 sections so the number of inputs will be the same, although the virtual robot will rely on model based on *Piecewise Constant Curvature* (PCC) to reconstruct its shape (Gravagne and Walker 2000). This is the main difference between commanding non-redundant robots and the shape of continuous or soft ones.

On the other hand, *situational awareness* can be defined as the perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future (Endsley 1988a). Under these circumstances, a high level of situational awareness is considered a result of a high level of telepresence. Thus, situational awareness will be measured later to infer the quality of telepresence.

Therefore, two interfaces were created in order to make further comparisons: a *conventional* one that shows the information in a regular monitor, and an *immersive* one that uses a virtual reality headset to display the same information.

The scenes are originally developed in a three-dimensional space, although in the conventional interface the information is shown in a two-dimensional (2D) display. As it will be explained in Sect. 6, in both interfaces the user is expected to move the input towards 5 predefined positions and, in case of making a selection, the robot will move accordingly. The pose of the robot will be updated in the interface in real time using the onboard sensory system explained in Sect. 3.

4.1 Conventional interface

In this interface, the user is sitting looking at the monitor and can use the keyboard to move in 3D a sphere that symbolizes the position where the robot's end-effector is commanded to move (Fig. 4). The real robot was located in the same space so the users could see the robot being commanded. However, in a real scenario, the robot could be located in a remote environment, so the current pose of the robot is adequately displayed in real time using the bilateral scheme explained in Sect. 3.

4.2 Immersive interface

In the immersive interface, the same information is displayed, but in this case an immersive headset with three-dimensional (3D) stereoscopic vision is used. For this practice, the user is asked to stand up and make small movements to command the robot position. Here, the user moves the 3D sphere that symbolizes the end-effector position by pushing it using the bare hands, tracked by the *Leap Motion* device (Fig. 5).

As a remainder, these interfaces were developed to contrast which one is better in terms of efficiency and situational awareness. In later experiments, only visual feedback will be further analysed to enlighten which one of such interfaces is
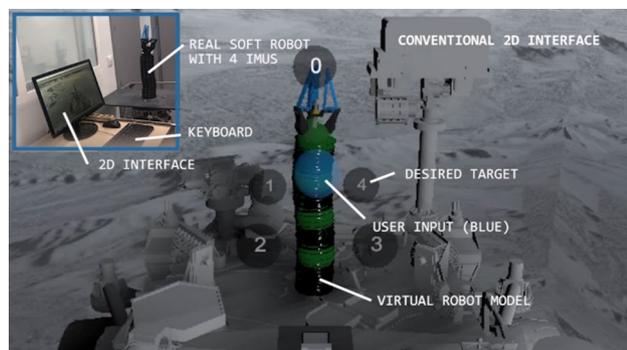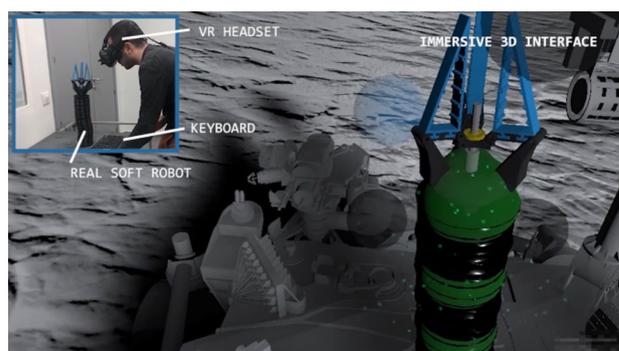

Fig. 5 Immersive interface. The setup is nearly the same than in the conventional interface but using a VR headset instead of a monitor

better suited to enhance telepresence and efficiency, regardless of the interaction strategy used.

Next, Sect. 5 will introduce different interaction tools to contrast if *natural language* can improve the performance and reduce the workload against conventional tools to teleoperate hyper-redundant robots. Specific details of the experiments and results for both hypotheses are explained in Sect. 6.

5 Interaction tools

One of the main purposes of this work aims at determining which strategy is more appropriate to teleoperate hyper-redundant robots in an immersive interface in terms of efficiency and workload.

In this context, *efficiency* will be defined in the same way as the previous Sect. 4, as the ability to maximize precision and speed. On the other hand, the *workload* can be defined as a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance (Hart and Staveland 1988).

As a result, 6 different approaches have been developed in *Unity3D*, making use of a computer, the *HTC Vive* VR headset and, in some cases, *Xbox* or *Leap Motion* controllers. These strategies are the following: controller, master–slave, local gestures, remote gestures, voice commands and a hybrid approach of gestures and voice recognition.

The strategy based on a *controller* was chosen to illustrate one of many possible conventional tools based on a physical device to command the movements. The *master–slave* was inspired by past efforts based on physical replicas of the robot; but in this case, the replica was designed in the virtual world, so it is easily scalable, the teleoperation transparency is preserved, and the construction and maintenance costs are minimized. The *local gestures* strategy was developed to be the most intuitive since it relied on shaping the real-size robot with the hands. While *remote gestures* were

Fig. 4 Conventional interface. Here, the scene is displayed in two dimensions by a monitor. The robot is commanded to move using a keyboard, and its pose is displayed in the interface in real time

also conceived to manipulate the robot with the hands, they promise better interactions when the robot is too big or too small. Additionally, *voice commands* were implemented to let the robot be moved when the operator hands are busy or their movement is hindered. Finally, the *gestures and voice* strategy combines two natural-language approaches to determine if they synergize and give place to better results.

In every method, the user has to exploit all the available resources to adequately manage the pose of a virtual robot, which has to be approached towards a predefined virtual reference. This reference was only conceived to measure the performance and conduct the experiments. A more detailed description of each interaction tool is given below. Basically, in all these interaction tools the user will be introduced in a *Virtual Reality* scenario and then will be requested to move a virtual robot (green) towards a virtual reference (blue). In all these experiments, the participants needed to complete the tasks as fast and precise as possible, as it will be explained in Sect. 6.

5.1 Controller

This strategy uses an *Xbox* gaming controller to teleoperate the robot, so it can be defined as a conventional technique against others developed in this work based on natural language. The user, provided with a *VR* headset and the controller, will rely on pressing buttons and axis to move the virtual robot towards the reference.

First, the desired joint must be selected using some buttons, and then, it can be moved along the Cartesian directions using the controller's axis. The implementation allows simultaneous inputs so the user can combine the three directions to perform 3D movements. The joints' speed can be adjusted by the quantity of displacement in the axis. In fact, these inputs have been filtered to minimize drifting or undesired sensibility issues (Fig. 6).
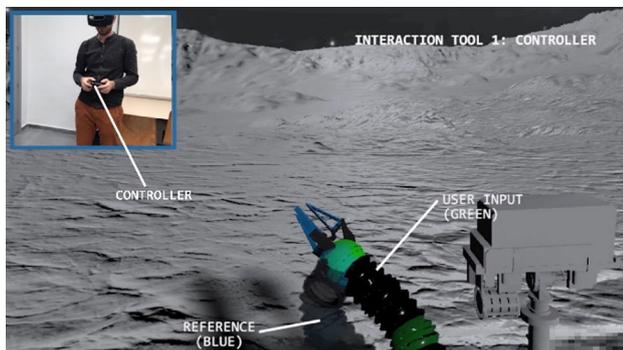
5.2 Master–slave

Traditionally, the master–slave architecture often implied the design and construction of one physical device (master) to reproduce the real-robot movements (slave). In this context, the system is more transparent if the similarity of the master and the slave is higher. However, creating a physical replica of the robot with onboard sensors usually involved additional cost, maintenance and construction time.

Currently, immersive realities offer the possibility of rapidly and easily creating a virtual master with the same features of the slave without any physical restriction on its construction, and any additional cost or maintenance. According to the specific literature, all the presented strategies in this chapter have a master and a slave. However, this strategy takes its name inspired from traditional master–slave physical models, since the robot relies on a virtual model of different size. Thus, this strategy could be most beneficial when the robot size is too big or too small to be adequately manipulated using regular gestures.

In this scenario, the user's hands were tracked using the *Leap Motion* controller, which uses a set of infrared *light-emitting diodes* (LED) and cameras to track the movements of hands and fingers. The sense of *proprioception* using this device is very high, so the users feel like touching virtual objects, although they are not.

Thus, when the user raises one hand, a smaller replica of the robot is created. Additionally, it will follow the hand movements until it closes so it can be positioned as desired. Then, the other hand can be raised to gently push the robot joints towards the desired reference by applying contact. By doing this, the slave will be moved the same way as the master is being manipulated. But in any case the user will enter in contact with the slave directly (Fig. 7).

Fig. 6 Controller strategy. In this interaction tool, a gaming controller is used to select and move the robot joints one by one using buttons and axis
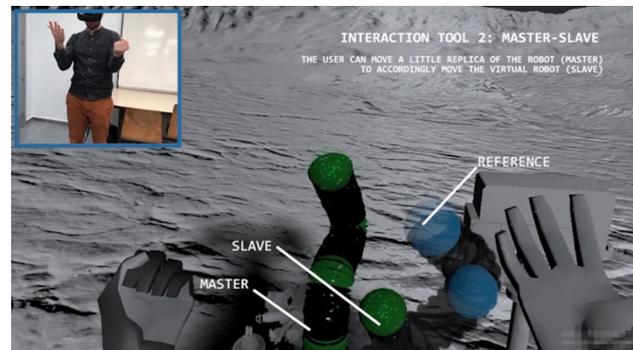

Fig. 7 Master–slave strategy. Here, a little replica of the virtual robot (master) is manipulated so the bigger one (slave) is identically moved towards the desired reference

5.3 Local gestures

This strategy also uses hand tracking to capture gestures. However, the main difference is that the user must approach the virtual real-size robot and use the hands to push the robot joints towards the desired reference by direct contact. It was found that pushing enhanced efficiency and reduced workload against grabbing, so the joints could not be grabbed but only pushed using this approach.

Here, the performance is highly dependent on the relative size of the robot and the user. If the virtual robot is too small, it will be difficult to be precise. On the other hand, if the virtual robot is too big, then the user will be very slow and will need to move more in the workspace. However, in *VR* is possible to scale all the environment so the perfect size of the robot can be chosen to the most adequate even if the real one has other dimensions (Fig. 8).

5.4 Remote gestures

Next, similarly to the master–slave strategy, remote gestures allow the robot to be moved without direct contact with the virtual robot model. In this case, the virtual robot model is moved without any intermediary using hand gestures as well.

Therefore, the user can open the index finger on the one hand, so a virtual laser appears as an extension to easily select the joint to move. Then, when the other hand is opened, it will move the selected joint accordingly to its own movement. When this hand closes, the joint does not move.

This implemented technique is called *reindexing*. It can be defined as the temporary removal of connection between master and slave when the range of motion is smaller for the first than for the second one (Sayers 1999). It is the same principle as when a computer mouse is raised when getting closer to the edge of a table.

Since this strategy relies on remote gestures, the operator is not expected to move a lot, so reindexing is very important to minimize the difference of size between master and the slave (Fig. 9).

5.5 Voice commands

This following tool is based on speech recognition to command the robot movements. This approach was created so the user can teleoperate the robot even if the hands are busy or their movement is hindered. Therefore, the user needed to remember some intuitive commands previously defined in a dictionary. Such commands could activate the movement of the robot by small increments or continuously by defining when to start and stop. This work opted for the second alternative because it was found to increase the precision and dramatically reduce the number of commands.

Therefore, the joint to move is first selected by pronouncing numbers, in this case, from 1 to 4. Then, the speed can be defined as slow, medium or fast. Only three levels of speed are taken to reduce the complexity of the strategy. However, they are tuned for the robot dimensions, so the slowest one is aimed to be more precise, while the fastest is more adequate to move long distances. Finally, the user can speak in which direction to move the selected joint among left, right, forward, backward, up and down.

These directions have not been allowed to be combined using this method because the result was unfamiliar to the users. Alternatively, this method could be implemented using voice intensity as a variable to select the movement speed (Fig. 10).

5.6 Gestures and voice

If the user can speak and move in the real environment, another approach is based on using both voice and gestures

Fig. 8 Local gestures strategy. The user's hands are tracked using a Leap Motion system so they can be used to directly interact with the virtual robot
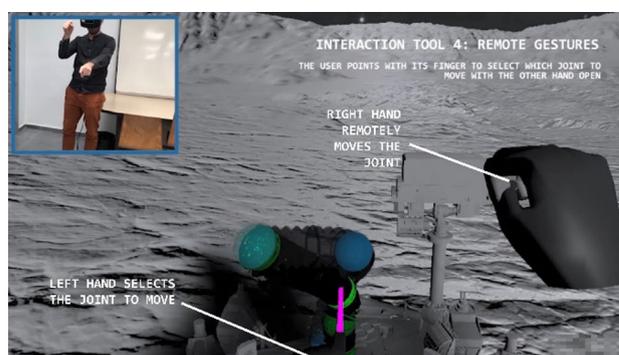

Fig. 9 Remote gestures strategy. In this setup, the user can select the joint to move by pointing with a virtual laser that emanates from the left index finger. Then, the selected joint remotely moves when the right hand is opened
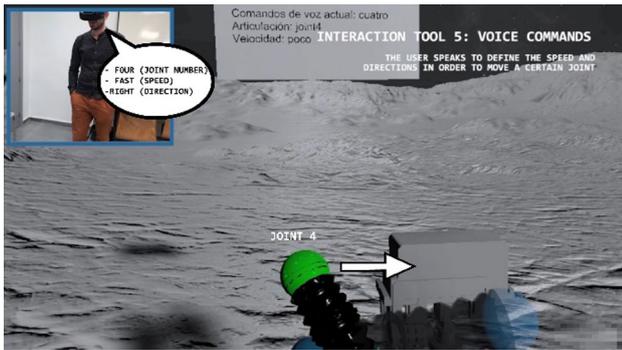
Fig. 10 Voice commands strategy. The VR headset records the user's voice to recognize simple commands in order to select the joint to move, its speed and the direction

to command the robot. Thus, it is the only multimodal approach in this work. Several methods could be proposed in this matter, but it was finally decided to combine what voice and gestures do better.

Therefore, and similarly to the previous method, *Voice Commands* are programmed to select the desired joint by pronouncing which joint to move from 1 to 4. Then, similarly to the *Remote Gestures* strategy, the user can raise the hand to move the joint with reindexing, as explained before. Hence, the voice is proposed as an alternative of pointing with the other hand to unequivocally select the joints, and gestures are selected to efficiently move the joint. This method can be understood as a hybrid method within the natural-language strategies (Fig. 11).

6 Experiments and results

Two experiments were conducted between February and March of 2019 to validate the proposed hypotheses of this work. As a reminder, the first one wonders whether immersive 3D interfaces or conventional 2D ones are better suited
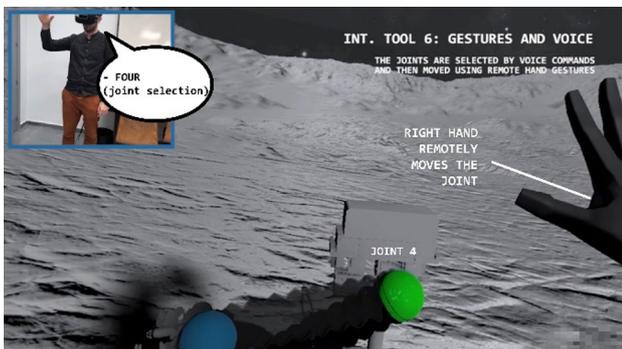

Fig. 11 Gestures and voice strategy. This strategy uses the voice to select a joint and hand gestures to remotely move it

to teleoperate hyper-redundant robots. On the other hand, the second one aims at selecting which interaction tools are more efficient and reduce workload for the same purpose.

For both experiments, a sample of 50 volunteers was selected in order to be heterogeneous enough to lead to generalizable results, while also homogeneous enough to minimize variances due to factors not relevant in the context of this work. For example, an engineering student with experience using computers or videogames is expected to perform differently compared to an elderly person. Since this work is not intended to compare people skills, but the interfaces and interaction tools instead, the population sample should exhibit a compromise between heterogeneity and homogeneity.

Therefore, the participants have different academic backgrounds but all of them come from a university context, they have different genders and ages but come from the same generation, and have different previous experience in immersive realities, robotics and videogames (Table 1).

Since this sample is just one sector of the potential population, the obtained data are first evaluated to measure the goodness of fit to a determined statistical distribution. For this purpose, the *Kolmogorov–Smirnov test* is applied to all the normalized data using the software *MATLAB* (Daniel 1990). As a result, it has been found that the data follow a *normal distribution*. Based on this assumption, determining if the means of two sets of data are significantly different from each other, require the use of a *Student's t test* or *one-way analysis of variance* (ANOVA), so they will be applied throughout this chapter. If not mentioned, the chosen significance level is $\alpha = 0.05$.

6.1 Human–robot interfaces

In order to test the objective efficiency of the developed interfaces in terms of precision and time, the participants were asked to move a blue sphere as a representation to where the end-effector of the robot is desired to move. They were warned to do it equally as fast and precise as possible. This input could be moved towards 5 numbered and predefined target positions, depicted by grey spheres of the same size. The participants could practice for a while until they

Table 1 Mean and variance values of the participant ages and their subjective previous experience (the higher, the more experience)

Personal data	Mean	Variance
Age (years)	24.2	11.5
VR experience (%)	35.6	58.4
AR experience (%)	28.1	50.2
Robotics experience (%)	50.1	79.2
Videogame experience (%)	66.9	81.5

felt ready, in order to minimize the learning factor between interfaces. They could press a begin button and move the input towards 3 of the 5 target positions, previously specified. Then, a timer began to run, and the precision was measured as the mean of the Euclidean distances between the input and the desired targets. Both time and precision variables were first rescaled—min–max normalization—(Eq. 1) among the data relative to both interfaces to allow further comparisons, then averaged (Eq. 2), and then rescaled again (Eq. 1).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2}$$

where x is a single data for each participant, and x' the scaled value relative to all the population for both interfaces, $\bar{x}$ the mean value between time and precision and n the number of participants. By this method, speed and precision are measured the same to test efficiency since both parameters can be considered as equally important to teleoperate a generic hyper-redundant robot. However, it must be remarked that one variable could be more important than the other depending on the robot and its purpose.

As a result, the efficiency is a 33.55% higher on average and the variance a 41.98% lower for the immersive interface than for the conventional one (Fig. 12a). The results between immersive and conventional interfaces are significantly different between each other based on both the *Student's t test* ($p = 1.9 \times 10^{-6}$) and the *ANOVA* ($p = 6.6 \times 10^{-7}$). Therefore, their comparison is statistically reliable.

The situational awareness (S_A) is measured in this document inspired by previous works (Endsley 1988b). Since such variable is subjective and depends on the environment setup, the questionnaire is usually custom-made. In this test, 3 questions were made to quantify the user perception of what was happening in the real world, compared to the images displayed on the interface. Therefore, the participants were able to see both the real robot and the interface in order to make an inference.

It is worth highlighting that the interface was programmed to hide the target number labels, so the participants were previously asked to remember them in order to test memory retention (s_1). Moreover, they were asked to focus on which of the 3 commanded movements they thought their precision was higher (s_2). To finish, they were asked to evaluate visual feedback from both interfaces (s_3). Parameters s_1 and s_2 were binary variables which could take value 0 if the answer was wrong or 1 if it was right. Variable s_3, the visual feedback, it was measured from 1 to 10. All these variables were normalized (Eq. 1) to allow further averaging (Eq. 3).

Furthermore, the influence of s_3 was increased to reduce the variance of the results a 10% (Eq. 3).

$$S_A = \frac{s_1 + s_2 + 3s_3}{5} \tag{3}$$

Afterwards, the resulting data are again rescaled (Eq. 1) in order to facilitate comparisons between interfaces. The results show that situational awareness is a 73.71% higher in the immersive interface than for the conventional one (Fig. 12b). Concerning the visual feedback evaluation, the participants were able to see the real robot and the model within the interface simultaneously and in real time, so they could make reliable comparisons between them. As a result, they perceived that the visual feedback was a 39.47% higher for the immersive interface than for the conventional one on average (Fig. 12c). These results are also significantly different based on the *Student's t test* ($p = 1.8 \times 10^{-10}$) and the *ANOVA* ($p = 4.3 \times 10^{-12}$).

To finish, and regardless of previous questions, the participants were asked to subjectively choose between both interfaces assuming they had to perform the same task. Consequently, 94% of the participants preferred to use the
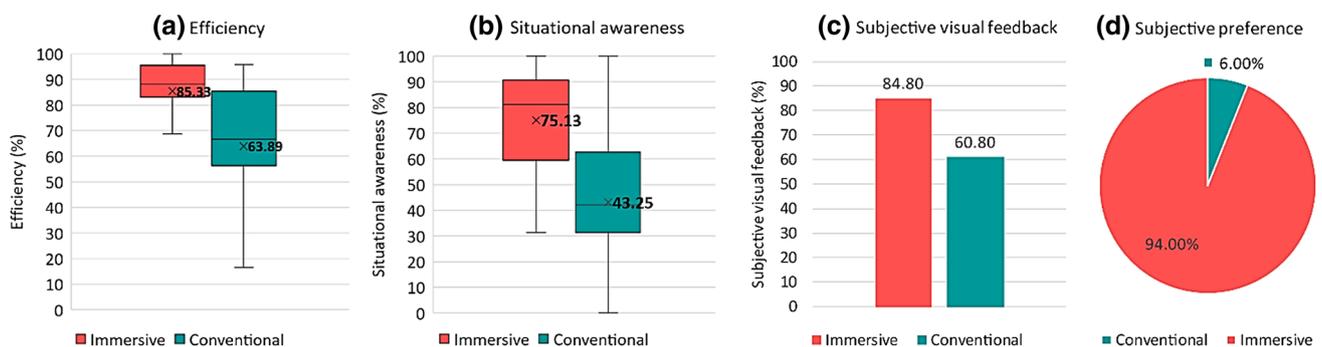

Fig. 12 Results for the human–robot interfaces test. **a** Efficiency in terms of precision and time, **b** situational awareness, **c** subjective visual feedback and **d** subjective preference

immersive interface (Fig. 12d). Therefore, both subjective and objective parameters agree to select the immersive 3D interface to teleoperate hyper-redundant robots as the best to enhance efficiency and telepresence.

6.2 Interaction tools

The human–robot interfaces test shows that immersive ones obtain better results to teleoperate hyper-redundant robots. The next inquiry is based on selecting which interaction strategy is better within this immersive reality for the same purpose. As a reminder, there are 6 developed strategies, as explained in Sect. 5: controller (CR), master–slave (MS), local gestures (LG), remote gestures (RG), voice commands (VC) and gestures and voice (GV). In this test, the participants used a VR headset to command the movements of a hypothetical virtual robot towards a randomly defined set of references by using the previously mentioned interaction approaches.

On the one hand, the efficiency was similarly measured using both elapsed time and precision. For each one of the approaches, the user could also practice before the timer began in order to reduce the learning factor. And the precision was measured and averaged as the Euclidean distance between every 4 sections of the virtual robot and the relative ones of the reference, differently coloured to ease comprehension.

In all the strategies, the user had to repeat the process 3 times, moving the virtual robot towards different references in order to increase the data normalization. Such references were randomly defined but bounded to a predefined and limited space so they would not differ too much from one participant to another. Therefore, the virtual robot joints were moved 3600 times in total in these experiments. Accordingly, the learning factor was minimized, whereas the randomness and repetitions standardize the data in order to achieve generalization.

Similarly to the previous experiment, the speed and time were first rescaled (Eq. 1), then averaged (Eq. 2), and then rescaled again (Eq. 1), giving place to the previously defined *efficiency*. In this case, the obtained *p values* using the *Student's t test* in pairs show that the interaction tools are significantly different except for two combinations (Table 2). In those cases, it is not possible to determine which one of them is statistically better or worse.

Therefore, as seen in Fig. 13a, the efficiency is indistinguishably the best for both *local gestures* and *gestures and voice*, with mean values of around 90%. They are closely followed by *remote gestures* and the *controller*, also indistinguishable between each other with efficiencies of around 85% on average. And finally, the *master–slave* and the *voice commands* are, respectively, the ones with less objective efficiency.

Table 2 Pair-wise *p* values to show the significance between the results of efficiency in the interaction tool test

	CR	MS	LG	RG	VC	GV
MS	10^{-11}					
LG	10^{-5}	10^{-15}				
RG	**0.18**	10^{-12}	10^{-5}			
VC	10^{-22}	10^{-5}	10^{-26}	10^{-20}		
GV	10^{-3}	10^{-17}	**0.13**	0.02	10^{-26}	

Relations in red do not reject the null hypothesis

CR controller, *MS* master–slave, *LG* local gestures, *RG* remote gestures, *VC* voice commands, *GV* gestures and voice

On the other hand, the most common method to determine the workload (W) of a mission is the *NASA Task Load Index* (NASA-TLX) (Hart 2006; Hart and Staveland 1988). It is based on rating 6 variables related to the influence of workload: mental demands, physical demands, temporal demands, performance, efforts and frustration. These parameters were first defined to ease their comprehension, and then, the users could evaluate them using natural numbers ($w_i = [0, 20]$). On the other hand, the participants ordered from lowest to highest the subjective relevance of such parameters in workload ($r_i = [0, 5]$), so the obtained scores were weighted and normalized according to (Eq. 4).

$$W = \frac{\sum_{i=1}^{6} w_i r_i}{\sum_{i=1}^{6} 20 r_i} \tag{4}$$

Afterwards, the resulting data are again rescaled (Eq. 1) in order to facilitate comparisons between strategies. Again, the *p values* using the *Student's t test* in pairs have been analysed to objectively make reliable comparisons from the results showed in Fig. 13b (Table 3).

Thus, local *gestures*, *controller*, *gestures and voice* and *remote gestures* have half of the workload compared to *master–slave* and *voice commands* with mean values of approximately 32% and 61%, respectively.

Besides, the participants selected the frustration as the most important factor to influence workload, followed by performance, mental demands, efforts, time demands and physical demands as shown in Fig. 13d. Furthermore, it is possible to compare the objective efficiency represented in Fig. 13a with the subjective performance obtained from the workload test. Surprisingly, the participants were very accurate evaluating their performance and it was found to be highly influenced by the level of frustration (Fig. 13e).

As explained in Table 1, the participants were asked about their previous experience in VR and AR, robotics and videogames. If that experience is averaged and plotted against the objective mean efficiency for both tests, the variables are found as directly proportional (Fig. 14). Therefore, previous

Fig. 13 Results for the interaction-tools test. **a** Efficiency in terms of precision and speed, **b** workload, **c** subjective preference, **d** subjective importance of variables in workload, **e** subjective and objective performance comparison and **f** previous experience influence in performance

Table 3 Pair-wise p values to show the significance between the results of the workload in the interaction tools test

	CR				
MS	10^{-10}	**MS**			
LG	**0.29**	10^{-13}	**LG**		
RG	**0.25**	10^{-10}	10^{-3}	**RG**	
VC	10^{-11}	**0.59**	10^{-13}	10^{-10}	**VC**
GV	**0.91**	10^{-8}	**0.26**	**0.16**	10^{-12}

Relations in red do not reject the null hypothesis

CR controller, *MS* master–slave, *LG* local gestures, *RG* remote gestures, *VC* voice commands, *GV* gestures and voice

Fig. 14 Previous experience influence in efficiency both for the human–robot interfaces and the interaction tools experiments

experience can be understood as a very important factor to achieve a high level of efficiency.

Although the *master–slave* strategy is controlled by gestures, it has been inspired by the physical architecture. Therefore, this work proposes to consider this tool as conventional, together with the *controller* one. In this context, the results show that despite having less experience with natural language than conventional tools, the participants

performed better using the first ones against the last ones (Fig. 13f).

Finally, and regardless of previous results, the participants were asked to choose among their preferred interaction

tool to teleoperate hyper-redundant robots. More than half of the participants, exactly 52%, chose the *gestures and voice* strategy, followed by *local gestures* with 28% and *remote gestures* with 14% (Fig. 13c). If the number of conventional and natural-language tools were the same, 92.31% of the participants would have preferred to use a natural-language tool against a conventional one.

7 Conclusions

This work aims to clear the path to focus the future research of hyper-redundant robot's teleoperation. This kind of robots has numerous degrees of freedom, so their spatial understanding is often difficult to be perceived. Teleoperation is a very important field in robotics since it is essential for several applications such as surgery, spatial missions, inspection or surveillance. A bilateral communication based on a position–position scheme has been developed to be the most adequate to teleoperate hyper-redundant robot applications.

On the one hand, this work aimed to determine whether immersive 3D interfaces are better or worse than conventional 2D ones in order to teleoperate hyper-redundant robots. Experiments with 50 participants show that immersive interfaces exhibit 33.55% mean higher efficiency in terms of speed and precision and 73.71% mean higher situational awareness than conventional ones. As a result, immersive interfaces were chosen to teleoperate this kind of robots by 94% of the participants.

On the other hand, some interaction tools were tested by the same population in order to determine which strategy is better suited for the same purpose within an immersive reality. A total of 6 methods were developed: *controller*, *master–slave*, *local gestures*, *remote gestures*, *voice commands* and *gestures and voice*. In this case, efficiency is indistinguishably the best for both *local gestures* and *gestures and voice*, with mean values of around 90%, closely followed by *remote gestures* and *controller*. Also, a workload test showed that *local gestures*, *controller*, *gestures and voice* and *remote gestures* have half of the workload than the other two approaches.

Besides, the participants were very accurate in evaluating their performance, which was found to be highly influenced by their level of frustration. Moreover, results showed that natural-language tools are more intuitive since exhibited higher performance despite the less previous experience. Finally, more than half of the users chose the *gestures and voice* strategy as the preferred one. This is understandable since we, as human beings, are used to communicate by simultaneously gesticulating and speaking. This is closely related to the *Principle of Least Astonishment* (*POLA*), which states that a user interface should behave in the way that most users will expect it to behave. Thus, it can be deduced that this combined strategy was the most intuitive, while, at the same time, the distribution of tasks to different stimulus did not increase the subjective workload to the users. Moreover, 92.31% selected an interaction tool based on natural language against a conventional one. Future work in this matter should be focused on proposing and classifying new interaction strategies in order to select the most appropriate, depending on the hyper-redundant robot specifications and its applications.

Therefore, this work concludes that immersive realities and natural-language tools should play an important role in the near future of hyper-redundant robots and their teleoperation.

References

Aracil R, Ferre M, Hernando M, Pinto E, Sebastian J (2002) Telerobotic system for live-power line maintenance: ROBTET. Control Eng Pract 10:1271–1281

Aracil R, Buss M, Cobos S, Ferre M, Hirche S, Kuschel M, Peer A (2007) The human role in telerobotics. In: Advances in telerobotics. Springer, pp 11–24

Barrientos A, Peñin LF, Balaguer C, Aracil R (2007) Fundamentos de robótica. McGrawHill, Madrid

Bhattacherjee S, Chattopadhayay S, Rao V, Seth S, Mukherjee S, Sengupta A, Bhaumik S (2018) Kinematics and teleoperation of tendon driven continuum robot. Procedia Comput Sci 133:879–886

Britton N, Yoshida K, Walker J, Nagatani K, Taylor G, Dauphin L (2015) Lunar micro rover design for exploration through virtual reality tele-operations. In: Field and service robotics. Springer, pp 259–272

Bugalia N, Sen A, Kalra P, Kumar S (2015) Immersive environment for robotic tele-operation. In: Proceedings of the 2015 conference on advances in robotics. ACM, p 49

Chen JY, Haas EC, Barnes MJ (2007) Human performance issues and user interface design for teleoperated robots. IEEE Trans Syst Man Cybern Part C (Appl Rev) 37:1231–1245

Cheng-jun D, Ping D, Ming-lu Z, Yan-fang Z (2009) Design of mobile robot teleoperation system based on virtual reality. In: IEEE international conference on automation and logistics. IEEE, pp 2024–2029

Chirikjian GS, Burdick JW (1994) A hyper-redundant manipulator. IEEE Robot Autom Mag 1:22–29. https://doi.org/10.1109/100.388263

Collins T, Shen W-M (2016) PASO: an integrated, scalable PSO-based optimization framework for hyper-redundant manipulator path planning and inverse kinematics. Information Sciences Institute Technical Report

Daniel WW (1990) Kolmogorov–Smirnov one-sample test. In: Applied nonparametric statistics, vol 2

Draper M, Calhoun G, Ruff H, Williamson D, Barry T (2003) Manual versus speech input for unmanned aerial vehicle control station operations. In: Proceedings of the human factors and ergonomics society annual meeting, vol 1. SAGE Publications Sage CA: Los Angeles, CA, pp 109–113

Dutta P, Rastogi N, Gotewal KK (2017) Virtual reality applications in remote handling development for tokamaks in India. Fusion Eng Des 118:73–80

Endsley MR (1988a) Design and evaluation for situation awareness enhancement. In: Proceedings of the Human Factors Society annual meeting, vol 2. SAGE Publications Sage CA: Los Angeles, CA, pp 97–101

Endsley MR (1988b) Situation awareness global assessment technique (SAGAT). In: Proceedings of the IEEE 1988 National Aerospace and Electronics Conference. IEEE, pp 789–795

Espinoza MS, Goncalves J, Leitao P, Sanchez JLG, Herreros A (2012) Inverse kinematics of a 10 DOF modular hyper-redundant robot resorting to exhaustive and error-optimization methods: a comparative study. In: Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), 2012 Brazilian. IEEE, pp 125–130

Ferre M, Aracil R, Balaguer C, Buss M, Melchiorri C (2007) Advances in telerobotics, vol 31. Springer, Madrid

Frigola M, Fernandez J, Aranda J (2003) Visual human machine interface by gestures. In: IEEE international conference on robotics and automation (Cat. No. 03CH37422). IEEE, pp 386–391

Gravagne IA, Walker ID (2000) Kinematic transformations for remotely-actuated planar continuum robots. In: IEEE international conference on robotics and automation. Proceedings. ICRA'00. IEEE, pp 19–26

Hart SG (2006) NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting, vol 9. Sage publications Sage CA: Los Angeles, CA, pp 904-908

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Advances in psychology, vol 52. Elsevier, pp 139–183

Hedayati H, Walker M, Szafir D (2018) Improving collocated robot teleoperation with augmented reality. In: Proceedings of the 2018 ACM/IEEE International conference on human–robot interaction. ACM, pp 78–86

Hu C, Meng MQ, Liu PX, Wang X (2003) Visual gesture recognition for human-machine interface of robot teleoperation. In: Proceedings 2003 IEEE/RSJ international conference on intelligent robots and systems (IROS 2003) (Cat. No. 03CH37453). IEEE, pp 1560–1565

Kapadia AD, Walker ID, Tatlicioglu E (2012) Teleoperation control of a redundant continuum manipulator using a non-redundant rigid-link master. In: IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 3105–3110

Koizumi S, Kanda T, Shiomi M, Ishiguro H, Hagita N (2006) Preliminary field trial for teleoperated communication robots. In: The 15th IEEE international symposium on robot and human interactive communication. IEEE, pp 145–150

Kot T, Novák P (2018) Application of virtual reality in teleoperation of the military mobile robotic system TAROS. Int J Adv Rob Syst 15:1729881417751545

Kuno Y, Nakanishi S, Murashima T, Shimada N, Shirai Y (1999) Robotic wheelchair with three control modes. In: Proceedings 1999 IEEE international conference on robotics and automation (Cat. No. 99CH36288C). IEEE, pp 2590–2595

Labonte D, Boissy P, Michaud F (2010) Comparative analysis of 3-D robot teleoperation interfaces with novice users. IEEE Trans Syst Man Cybern Part B (Cybern) 40:1331–1342

Li N, Cartwright S, Shekhar Nittala A, Sharlin E, Costa Sousa M (2015) Flying frustum: a spatial interface for enhancing human-UAV awareness. In: Proceedings of the 3rd international conference on human–agent interaction. ACM, pp 27–31

Lipton JI, Fay AJ, Rus D (2018) Baxter's homunculus: virtual reality spaces for teleoperation in manufacturing. IEEE Robot Autom Lett 3:179–186

Lu Y, Liu L, Chen S, Huang Q (2010) Voice based control for humanoid teleoperation. In: International conference on intelligent system design and engineering application. IEEE, pp 814–818

Lueth TC, Laengle T, Herzog G, Stopp E, Rembold U (1994) KANTRA-human–machine interaction for intelligent robots using natural language. In: Proceedings of 3rd IEEE international workshop on robot and human communication. IEEE, pp 106–111

Mantecón T, del-Blanco CR, Jauregizar F, García N (2016) Hand gesture recognition using infrared imagery provided by leap motion controller. In: International conference on advanced concepts for intelligent vision systems. Springer, pp 47–57

Marques L, Dinis J, Coimbra AP, Crisóstomo MM, Ferreira JP (2009) 3D hyper-redundant robot. Paper presented at the 11th Spanish Portuguese Conference on Electrical Engineering, Zaragoza, Spain

Martín-Barrio A, Barrientos A, del Cerro J (2018a) The Natural-CCD algorithm: a novel method to control hyper-redundant and soft robots. Soft Robot 5:242–257

Martín-Barrio A, Terrile S, Barrientos A, del Cerro J (2018b) Hyper-redundant robots: classification, state-of-the-art and issues. Revista Iberoamericana de Automática e Informática Industrial 15:351–362

Martín-Barrio A, Terrile S, Díaz-Carrasco M, Del Cerro J, Barrientos A (2019) Modeling the soft robot Kyma based on Real-Time Finite Element Method (under review)

Mashood A, Noura H, Jawhar I, Mohamed N (2015) A gesture based kinect for quadrotor control. In: International conference on information and communication technology research (ICTRC). IEEE, pp 298–301

Matuszek C, Herbst E, Zettlemoyer L, Fox D (2013) Learning to parse natural language commands to a robot control system. In: Experimental robotics. Springer, pp 403–415

Michalos G, Karagiannis P, Makris S, Tokçalar Ö, Chryssolouris G (2016) Augmented reality (AR) applications for supporting human-robot interactive cooperation. Procedia CIRP 41:370–375

Mostefa M, El Boudadi LK, Loukil A, Mohamed K, Amine D (2015) Design of mobile robot teleoperation system based on virtual reality. In: 3rd international conference on control, engineering & information technology (CEIT). IEEE, pp 1–6

Nielsen CW, Goodrich MA, Ricks RW (2007) Ecological interfaces for improving mobile robot teleoperation. IEEE Trans Robot 23:927–941

Pessaux P, Diana M, Soler L, Piardi T, Mutter D, Marescaux J (2015) Towards cybernetic surgery: robotic and augmented reality-assisted liver segmentectomy. Langenbeck's Arch Surg 400:381–385

Poncela A, Gallardo-Estrella L (2015) Command-based voice teleoperation of a mobile robot via a human-robot interface. Robotica 33:1–18

Raskin J (2000) The humane interface: new directions for designing interactive systems. Addison-Wesley Professional, Boston

Ren L, Omisore OM, Han S, Wang L (2017) A master-slave control system with workspaces isomerism for teleoperation of a snake robot. In: 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 4343–4346

Rheingold H (1991) Virtual reality: exploring the brave new technologies of artificial experience and interactive worlds-from cyberspace to teledildonics. Secker & Warburg, Portland

Rininsland H (1999) ARTEMIS. A telemanipulator for cardiac surgery. Eur J Cardiothorac Surg 16:S106–S111

Roldán J, Peña-Tapia E, Martín-Barrio A, Olivares-Méndez M, Del Cerro J, Barrientos A (2017) Multi-robot interfaces and operator situational awareness: study of the impact of immersion and prediction. Sensors 17:1720

Roldán JJ, Crespo E, Martín Barrio A, Peña-Tapia E, Barrientos A (2019a) A training system for Industry 4.0 operators in complex assemblies based on virtual reality and process mining. Robot Comput Integr Manuf 59:305–316

Roldán JJ, Peña-Tapia E, Garzón-Ramos D, de León J, Garzón M, del Cerro J, Barrientos A (2019b) Multi-robot systems, virtual reality and ROS: developing a new generation of operator interfaces. In: Robot Operating System (ROS). Springer, pp 29-64

Sayers C (1999) Remote control robotics. Springer, New York

Shilling RD, Shinn-Cunningham B (2002) Virtual auditory displays. Handbook of Virtual Environment Technology, pp 65–92

Shirwalkar S, Singh A, Sharma K, Singh N (2013) Telemanipulation of an industrial robotic arm using gesture recognition with Kinect. In: International conference on control, automation, robotics and embedded systems (CARE). IEEE, pp 1–6

Stern H, Wachs J, Edan Y (2007) A method for selection of optimal hand gesture vocabularies. In: International Gesture Workshop. Springer, pp 57–68

Vertut J (2013) Teleoperation and robotics: applications and technology, vol 3. Springer, Berlin

Yoshizaki M, Kuno Y, Nakamura A (2001) Human–robot interface based on the mutual assistance between speech and vision. In: Proceedings of the Workshop on Perceptive user Interfaces. ACM, pp 1–4

Young SJ (1982) Real time languages: design and development. Ellis Horwood Ltd, Chichester