8

Choosing the tools for Improving distant immersion and perception in a teleoperation context

Nicolas Mollet, Ryad Chellali, and Luca Brayda
TEleRobotics and Applications dept. Italian Institute of Technology
Italy

1. Introduction

The main problems we propose to address deal with the human-robots interaction and interface design, considering N teleoperators who have to control in a collaborative way M remote robots. Why is it so hard to synthetize commands from one space (humans) and to produce understandable feedbacks from another (robots) ?

Tele-operation is dealing with controlling robots to remotely intervene in unknown and/or hazardous environments. This topic is addressed since the 40s as a peer to peer (P2P) system: a single human or tele-operator controls distantly a single robot. From information exchanges point of view, classical tele-operation systems are one to one-based information streams: the human sends commands to a single robot while this last sends sensory feedbacks to a single user. The forward stream is constructed by capturing human commands and translated into robots controls. The backward stream is derived from the robots status and its sensing data to be displayed to the tele-operator. This scheme, e.g. one to one tele-operation, has evolved this last decade thanks to the advances and achievements in robotics, sensing and Virtual/Augmented Reality technologies: these last ones allow to create interfaces that manipulate information streams to synthesise artificial representations or stimulus to be displayed to users or to derive adapted controls to be sent to the robots. Following these new abilities, more complex systems having more combinations and configurations became possible. Mainly, systems supporting N tele-operators for M robots has been built to intervene after disasters or within hazardous environments. Needless to say that the consequent complexity in both interface design and interactions handling between the two groups and/or intra-groups has dramatically increased. Thus and as a fundamental consequence the one to one or old fashion teleoperation scheme must be reconsidered from both control and sensory feedback point of views: instead of having a unique bidirectional stream, we have to manage N * M bidirectional streams. One user may be able to control a set of robots, or, a group of users may share the control of a single robot or more generally, N users co-operate and share the control of M co-operating robots. To support the previous configurations, the N to M system must have strong capabilities enabling co-ordination and co-operation within three subsets: Humans, Robots, Human(s) and Robot(s).

The previous subdivision follows a homogeneity-based criteria: one use or develop the same tools to handle the aimed relationships and to carry out modern tele-operation. For instance, humans use verbal, gesture and written language to co-operate and to develop strategies and planning. This problem was largely addressed through Collaborative Environments (CE). Likely, robots use computational and numerical-based exchanges to co-operate and to co-ordinate their activities to achieve physical interactions within the remote world. For human(s)-robot(s) relationships, the problem is different: humans and robots belong to two separate sensory-motor spaces: humans issue commands in their motor space that robots must interpret, to execute the corresponding motor actions through actuators. Conversely, robots inform humans about their status, namely they produce sensing data sets to be displayed to users' sensory channels. Human-Machine Interfaces (HMI) could be seen here as spaces converters: from robot space to human space and vice versa. The key issue thus is to guarantee the bijection between the two spaces. This problem is expressed as a direct mapping for the one-to-one (1 * 1) systems. For the N * M systems, the direct mapping is inherently impossible. Indeed, when considering a 1 * M system for instance, any aim of the single user must be dispatched to the M robots. Likely, one needs to construct an understandable representation of M robots to be displayed to the single user. We can also think about the N * 1 systems: how to combine the aims of the users to derive actions the single robot must perform?

This book chapter is focused on the way we conducted researches, developments and experiments in our Lab to study bijective Humans-Robots interfaces design. We present our approach and a developed platform, with its capabilities to integrate and abstract any robot into Virtual and Augmented worlds. We then present our experiences for testing N*1, 1*M and N*M contexts, followed by two experiences which aims to measure human's visual feedback and perception, in order to design adaptative and objectively efficient N*M interfaces. Finally, we present an application of this work with a real N*M application, an actual deployment of the platform, which deals with remote artwork perception within a museum.

2. State of the art

Robots are entities being used increasingly to both extend the human senses and to perform particular tasks involving repetition, manipulation, precision. Particularly in the first case, the wide range of sensors available today allows a robot to collect several kinds of environmental data (images and sound at almost any spectral band, temperature, pressure...). Depending on the application, such data can be internally processed for achieving complete autonomy [WKGK95,LKB+07] or, in case a human intervention is required, the observed data can be analysed off-line (robots for medical imaging, [GTP+08]) or in real time (robots for surgical manipulations such as the Da Vinci Surgical System by Intuitive Surgical Inc., or [SBG+08]). An interesting characteristic of robots with real-time access is to be remotely managed by operators (Teleoperation), thus leading to the concept of Tele-robotics [UV03,EDP+06] anytime it is impossible or undesirable for the user to be where the robot is: this is the case when unaccessible or dangerous sites are to be explored, to avoid life threatening situations for humans (subterranean, submarine or space sites, buildings with excessive temperature or concentration of gas).

Research in Robotics, particularly in Teleoperation, is now considering cognitive approaches for the design of an intelligent interface between men and machines. This is because interacting with a robot or a (inherently complex) multi-robots system in a potentially unknown environment is a very high skills and concentration demanding task. Moreover, the increasing ability of robots to be equipped with many small - though useful - sensors, is demanding an effort to avoid any data flood towards a teleoperators, which would dramatically drawn the pertinent information. Clearly, sharing the tasks in a collaborative and cooperative way between all the N $*$ M participants (humans, machines) is preferable to a classical 1 $*$ 1 model.

Any teleoperation task is as much effective as an acceptable degree of immersion is achieved: if not, operators have distorted perception of distant world, potentially compromising the task with artefacts, such as the well know tunneling effect [Wer12]. Research has focused in making Teleoperation evolve into Telepresence [HMP00,KTBC98], where the user feels the distant environment as it would be local, up to Telexistence [Tac98], where the user is no more aware of the local environment and he is entirely projected in the distant location. For this projection to be feasible, immersion is the key feature. VR is used in a variety of disciplines and applications: its main advantage consists in providing immersive solutions to a given Human-Machine Interface (HMI): the use of 3D vision can be coupled with multi-dimensional audio and tactile or haptic feedback, thus fully exploiting the available external human senses.

A long history of common developments, where VR offers new tools for tele- operation, can be found in [ZM91][KTBC98][YC04][HMP00]. These works address techniques for better simulations, immersions, controls, simplifications, additional information, force feedbacks, abstractions and metaphors, etc. The use of VR has been strongly facilitated during the last ten years: techniques are mature, costs have been strongly reduced and computers and devices are powerful enough for real-time interactions with realistic environments. Collaborative tele-operation is also possible [MB02], because through VR more users can interact in Real-Time with the remote robots and between them. The relatively easy access to such interaction tool (generally no specific hardware/software knowledge are required), the possibility of integrating physics laws in the virtual model of objects and the interesting properties of abstracting reality make VR the optimal form of exploring imaginary or distant worlds. A proof is represented by the design of highly interactive computer games, involving more and more a VR-like interface and by VR-based simulation tools used for training in various professional fields (production, medical, military [GMG+08]).

3. A Virtual Environment as a mediator between Humans and Robots

We firstly describe an overview of our approach: the use of a Virtual Environment as an intermediate between humans and robots. Then we briefly present the platform developed in this context.

3.1 Concept

In our framework we first use a Collaborative Virtual Environment (CVE) for abstracting and standardising real robots. The CVE is a way to integrate in a standardised way of interaction heterogenous robots from different manufacturers in the same environment, with the same level of abstraction. We intend in fact to integrate robots being shipped with the related drivers and robots internally assembled together with their special-purpose operating system. By providing a unique way of interaction, any robot can be manipulated through standard interfaces and commands, and any communication can be done easily: heterogenous robots are thus standardised by the use of a CVE. An example of such an environment is depicted in Figure 1: a team of teleoperators N1;N is able to simultaneously act on a set of robots M1;M through the CVE. This implies that this environment provides a suitable interface for teleoperators, who are able to access a certain number of robots simultaneously, or on the other hand just one robot's sensor in function of the task.

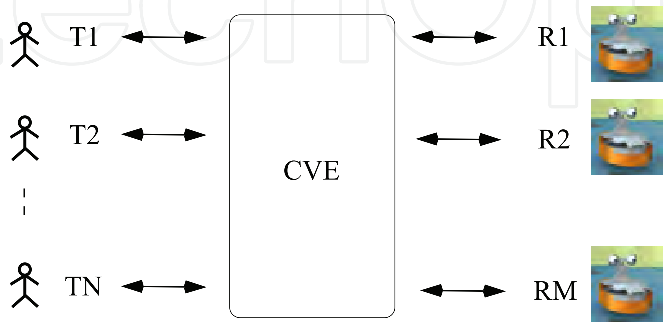


Fig. 1. Basic principle of a Virtual-Augmented Collaborative Environment: N teleoperators can interact with M robots.

3.2 Technical developments: the ViRAT platform

We are developing a multi-purposes platform, namely ViRAT (Virtual Reality for Advanced Teleoperation [MCB09][MBCF09][MBCK08]), the role of which is to allow several users to control in real time and in a collaborative and efficient way groups of heterogeneous robots from any manufacturer. We presented in the paper [MBCK08] different tools and platforms, and the choices we made to build this one. The ViRAT platform offers teleoperation tools in several contexts: VR, AR, Cognition, groups management. Virtual Reality, through its Virtual and Augmented Collaborative Environment, is used to abstract robots in a general way, from individual and simple robots to groups of complex and heterogeneous ones. Internal ViRAT's VR robots represent exactly the states and positions of the real robots, but VR offers in fact a total control on the interfaces and the representations depending on users, tasks and robots, thus innovative interfaces and metaphors have been developed. Basic group management is provided at the Group Manager Interface (GMI) Layer, through a first implementation of a Scenario Language engine[MBCF09]. The interaction with robots tends to be natural, while a form of inter-robots collaboration, and behavioral modelling, is implemented. The platform is continuously evolving to include more teleoperation modes and robots.

As we can see from the figure 2 ViRAT makes the transition between several users and groups of robots. It's designed as follows:

1. ViRAT Human Machine Interfaces provide high adaptive mechanisms to create personal and adapted interfaces. ViRAT interfaces support multiple users to operate at the same time even if the users are physically at different places. It offers innovative metaphors, GUI and integrated devices such as Joystick or HMD.

2. Set of Plug-in Modules. These modules include in particular:
   - Robot Management Module (RMM) gets information from the ViRAT interface and tracking module and then outputs simple commands to the control module.
   - Tracking Module (TM) is implemented to get current states of real environment and robots. This module also outputs current states to abstraction module.
   - Control Module (CM) gets simple or complex commands from the ViRAT interface and RMM. Then it would translates them into robots' language to send to the specific robot.
   - Advance Interaction Module (AIM) enables user to operate in the virtual environment directly and output commands to other module like RMM and CM.
3. ViRAT Engine Module is composed of a VR engine module, an abstraction module and a network module. VR engine module focuses on VR technologies such as: rendering, 3D interactions, device drivers, physics engines in VR world, etc. VR abstraction module gets the current state from the tracking module and then it abstracts the useful information, that are used by the RMM and VR Engine Module. Network Module handles communication protocols, both for users and robots.

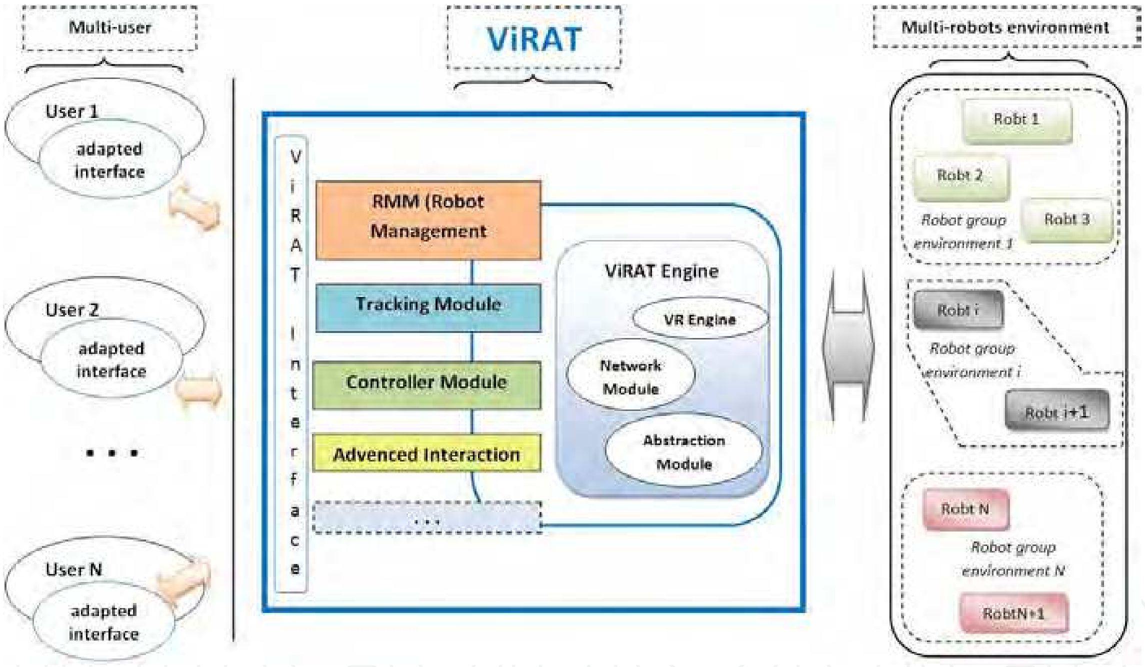


Fig. 2. ViRAT design

When a user gives some commands to ViRAT using his/her adapted interface, the standardised commands are sent to the RMM. Internal computations of this last module generate simple commands for the CM. During the running process, the TM gets the current state of the real environment and send it to the Abstraction Module, which abstracts the useful information in VIRAT's internal models of representation and abstraction. Considering this information, VR engine module updates the 3D environment presented to the user. RMM readapts its commands according to users' interactions and requests.

ViRAT project has many objectives, but if we focus on the HRI case there are two main objectives that interest us particularly for this paper:

Robot to Human

Abstract the real environment into the virtual environment: This will simplify the environment for the user. Ignorance of useless objects makes the operation process efficient. In the abstraction process, if we use a predefined virtual environment (Figure 5a), it will be initialised when the application starts running. Otherwise we construct the new virtual environment, which happens when we use ViRAT to explore an unknown area for example. After construction of a virtual environment in accordance with the real environment, we can reuse the virtual environment whenever needed. Thus the virtual environment must be adaptable to different applications. ViRAT has an independent subsystem to get the current state information from real environment termed as 'tracking module' in the previous section. The operator makes decisions based on the information perceived from the virtual environment. Because the operator does not need all the information from the tracking module, this abstraction module will optimise, abstract and represent the useful state information in real-time to user.

Human to Robot

The goal is to understand the human, and to transfer commands from the virtual environment into the real world. Several Teleoperators can interact simultaneously with 3 layers of abstraction, from the lowest to the highest (Figure 3) : the Control Layer, the Augmented Virtuality (AV) Layer, the Group Manager Interface (GMI) Layer. The Control layer is the lowest level of abstraction, where a teleoperator can take full and direct control of a robot. The purpose is to provide a precise control of sensors and actuators, including wheel motors, vision and audio system, distance estimators etc... The remaining operations, generally classified as simple, repetitive or already learnt by the robots, are executed by the Control Layer without human assistance; whether it is the case to perform them or not is delegated above, to the Augmented Virtuality Layer. Such layer offers a medium level of abstraction: teleoperators take advantage of the standardised abstracted level, can manipulate several robots with the same interface, which provides commands close to what an operator wants to do instead of how. This is achieved by presenting a Human-Machine Interface (HMI) with a purely virtual scene of the environment, where virtual robots move and act. Finally, the highest level of abstraction is offered by the Groups Manager Interface (GMI). Its role is to organise groups of robots according to a set of tasks, given a set of resources. Teleoperators communicate with the GMI, which in turns combines all the requests to adjust priorities and actions on robots through the RMM.

3.3 Goals of ViRAT

The design and tests of ViRAT allow us to claim that this platform achieves a certain number of goals:

- *Unification and Simplification*: there is a unified and simplified CVE, able to access to distant and independent rooms, which are potentially rich of details. Distant robots are parts of the same environment.

- *Standardisation*: we use a unified Virtual Environment to integrate heterogenous robots coming from different manufacturers: 3D visualisation, integration of physics laws into the 3D model, multiple devices for interaction are robot-independent.
- *Reusability*: behaviours and algorithms are robot-independent as well and built as services: their implementation is reusable on other robots.
- *Pertinence via Abstraction*: a robot can be teleoperated on three layers: it can be controlled directly (Control Layer), it can be abstracted for general commands (AV Layer), and groups of robots can be teleoperated through the GMI Layer.
- *Collaboration*: several, distant robots collaborate to achieve several tasks (exploration, video-surveillance, robot following) with one or several teleoperator(s) in real time.
- *Interactive Prototyping* can be achieved for the robots (conception, behaviours, etc.) and the simulation.
- *Advanced teleoperation interfaces*: we provided interfaces which start considering cognitive aspects (voice commands) and reach a certain degree of efficiency and time control.
- *Time and space navigation* are for the moment limited in the current version of ViRAT, but the platform is open for the next steps: teleoperators can already navigate freely in the virtual space at runtime, and will be able to replay what happened or to predict what will be (with for example trajectories planification and physics).
- *Scenario Languages applicability*. The first tests we made with our first and limited implementation of the Scenario Language for the GMI allow us to organise one whole demonstration which mixes real and virtual actors.

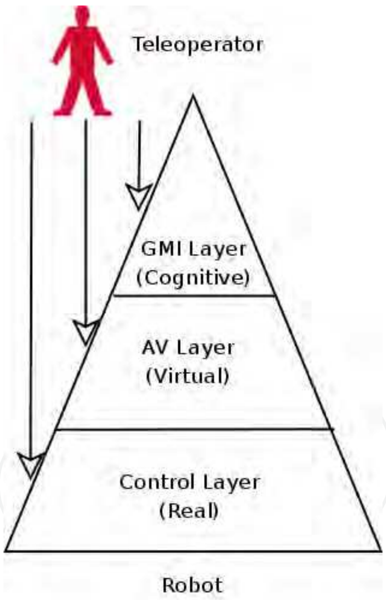


Fig. 3. In our CVE three abstraction layers (GMI, AV, Control) are available for teleoperation.

4. ViRAT's scenarios on the different actors and their interactions

As previously introduced, we aim to provide efficient N*M interfaces. To achieve such a goal, we divide the experiments in first, N*1 context, and second, 1*M context.

4.1 N*1: collaboration between humans

This basic demonstration is using one wheeled robot equipped with two cameras (figure 4). The camera video streams can be seen by a user who wear a Head Mounted Display (HMD). The movements of the HMD are tracked and transmitted to the robot's pan-tilt cameras. At any moment, this teleoperator can also see the VR world, synchronised with the real one. This VR environment is used by a second teleoperator who plan the robot's displacements. Following to this basic collaborative demonstration, we developed in the VR world a set of metaphors that for example allow to see the presence of another teleoperator in the environment, or also to understand which robot the other user is going to control, etc.

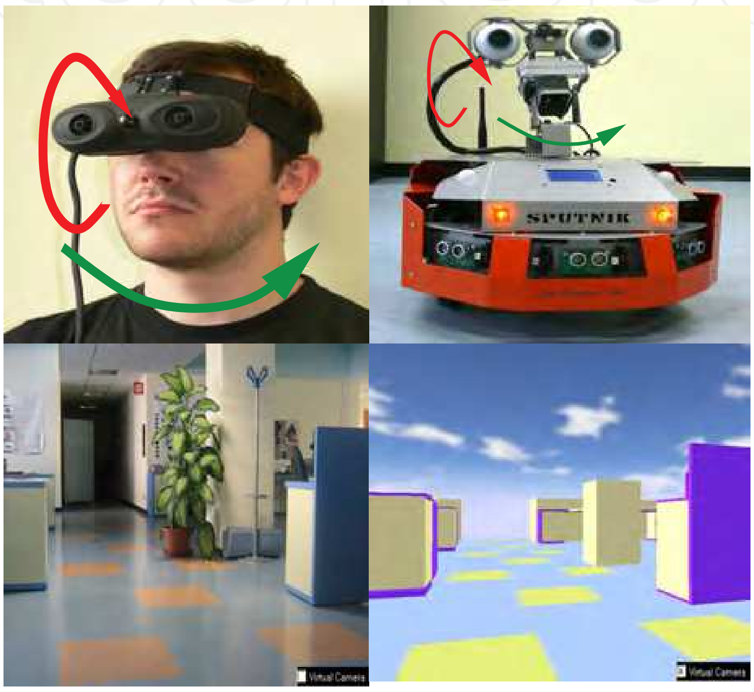


Fig. 4. One user controls robot camera, while another one uses the virtual world to choose the robot's displacements

4.2 1*M: collaboration between robots

The user supervise two robots that are collaborating to offer a real camera view to the user according to a target he pointed in the VR world (Figure 5b). One robot is a small humanoid with a camera, which moves slowly, while the second wheeled robot can go quickly to a chosen target. The user can give general commands to the robots through the Group Manager Interface, and then the RMM will generate the subtasks for this command, so it allows easily to ask to the wheeled robot to bring the humanoid one (Figure 6a), which can climb on the fast transportation robot. The TM provides the position and the orientation of the robots continuously. During the mission, user may interact with the group, showing the path and the targets to Sputnik (the wheeled robot) and redefining the requested viewpoint from the VR environment. Since HMD and Humanoid's head are synchronised, therefore user can move freely and naturally his/her head to feel immersed and present through the Humanoid's robot when this one is arrived at his final location. More details on this experiment can be found in [KZMC09].

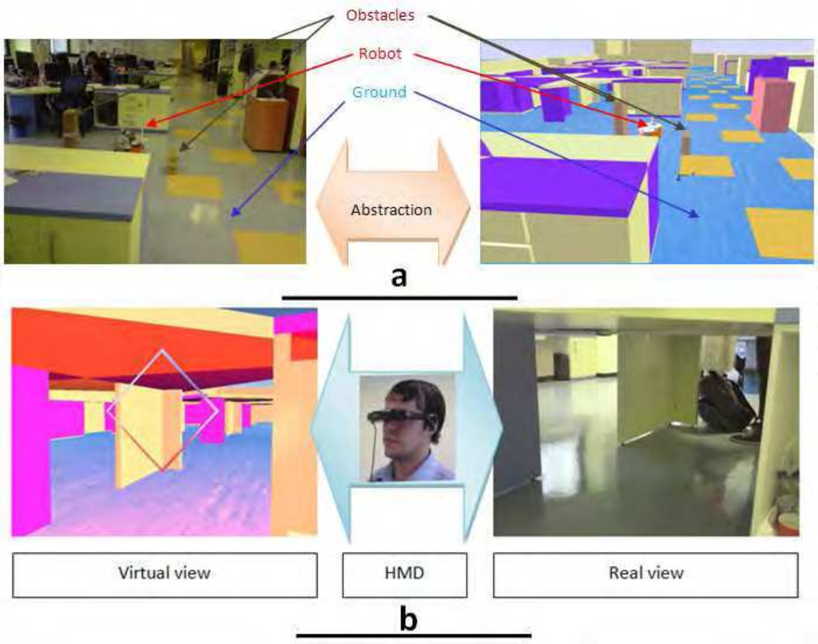


Fig. 5. VR abstraction, (a) Virtual environment interaction tool. (b) One example of some Metaphors given through VR tools

4.3 N*M: a simple case of multi-teleoperators interacting on multi-robots

We use for this demonstration two real and distant environments. The first one contains one robot, the second two robots. We considered two teleoperators who are also located in two different places. One operator acts through a PC, equipped with classical physical interfaces (mouse, keyboard, monitor), while the second one uses more immersive devices (head mounted display, joystick). The operators manage the three robots through the unified virtual world, without limitations due to site distribution. The Group Manager Interface (GMI) is responsible of scheduling and executing scenarios in function of the available resources (mainly robots' states), and to synchronise actions between the tele-operators. One of the advantages of using the VR world is that teleoperators can navigate freely in the two rooms, both when robots are moving or not. The real-time tracking of the position and velocity of the real robots (mirrored by the locations of avatars) is achieved thanks to a calibrated camera system, able to locate the real position of robots and input them in the AV Layer. The virtual avatars appear in the same virtual room, while real robots are separated in their real rooms. Thus, a distributed, complex space is represented via a unique, simple virtual room. This ViRAT's basic demonstration is a mixed of the two previous ones, so it includes all the combinations. Teleoperators can for example interact with the same robot (camera/displacements) while the GMI takes in charge the two other ones automatically.

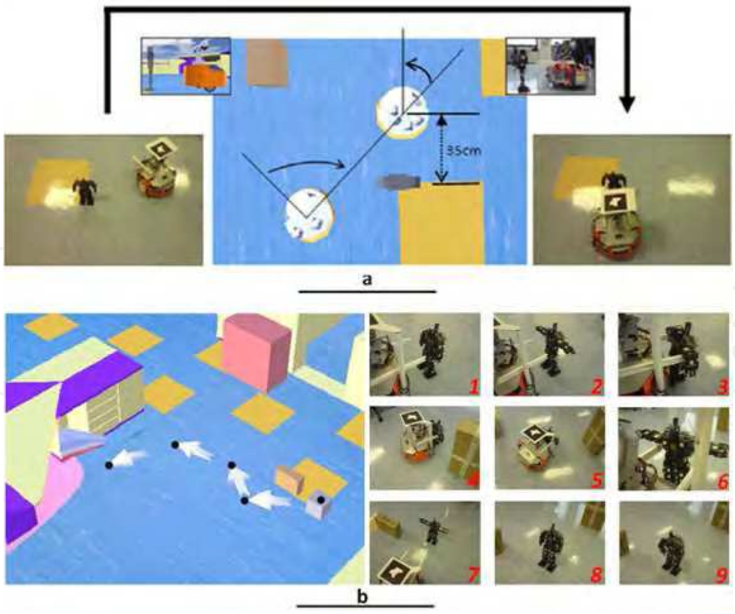


Fig. 6. Interaction through VR environment. (a) Go Near sub-task. (b) Collaboration scenario

5. Human Analysis

While we developed a set of tools for allowing the N*M general interactions pattern, we need precise analysis on human's perception in order to create adaptative and objective efficient interfaces. We present here two of our set of experiments currently in progress. We first present the influence of the height of the camera in a context of a path-to-follow task. Then, we evaluate the efficiency of a 3D map, compared to a 2D one, to allow self-localisation for the teleoperators according to a distant camera's video stream.

5.1 Height influence on task efficiency and context understanding

In this work we aim at finding ways to measure the capability of a teleoperator to achieve a simple task: a path following task that the operator must perform. The path is depicted on the ground and the user must drive the robot as close as possible to this path. The evaluation is done by comparing the path traced by the mobile robot and the original path. This allows us to drive some conclusions concerning the behaviour of the operator. Specifically, one way to measure the degree of accuracy of the task is to compute the surface between the theoretical (T) and the experimental (E) path. Each path is modelled as a curve, approximated by a piecewise linear segment joined by points in a 2D space: the approximation comes from the fact that the position and orientation of the robot is sampled, e.g by a camera acquisition system. By considering that the T and E frequently cross each other, the in-between surface S can be computed as:

$$S = \frac{1}{2}\sum_{i\in I}\sum_{p\in P_i}\begin{vmatrix} x_p x_{p+1} \\ y_p y_{p+1} \end{vmatrix} \tag{1}$$

where $I \in \{T \cap E\}$ is the set of points in which the two paths intersect, $Pi \in \{T \cup E\}$ is a subset of points between two consecutive intersections, p and p+1 are two consecutive points in each subset and x, y are the 2D coordinates of a point. The inner sum is the known Surveyor□s formula for the calculus of the area of a polygon. S can be interpreted as a surface-based error. Furthermore, because we make tests across different paths of different lengths, we can normalise by the theoretical path lengths by defining a Normalised Average Distance (NAD):

$$NAD = \frac{S}{\sum_{p \in T} \sqrt{\Delta x_p^2 + \Delta y_p^2}} \tag{2}$$

With such metric, the operators with a high/low NAD will be likely to have experienced a higher/lower deviation in following the main path. Such deviations are related to the degree of ability people have to change mentally their point of view (POV) or, on the contrary, it may represent the distortion the teleoperation system imposes to them. In other words, the deviation depends (at least partially) on the fidelity of the perception of space that each operator can feel. Figure 8(d) depicts an example of surface S, where the area is depicted in gray. The relationship is partial because other ingredients are missed such as the motor transformation between the hand actions and the robot rotations.

Experimental setup

In the experiments, the users had to follow as best as they could a stained path, by carrying out the teleoperation of an Unmanned Ground Vehicle (UGV) using a joystick for motor control output and a Head Tracking System (HTS) for perceptive control input. The users didn’t have any previous knowledge about the UGV or the path, and during the experiments they could rely on the sole subjective vision by teleoperating in a separated room. To reduce the experiment variability, the speed of the vehicle was fixed to 0.15 m/s (25% of the maximum speed). This way, the user only had to care about one degree of freedom of the UGV, i.e. the steering, and two degrees of freedom for the HTS (pan & tilt): this way the comparisons can be simpler and clearer.

The experiment was carried out by 7 people (3 women and 4 men), with an age range from 22 to 46 years old. Every user made a total number of 9 trials, i.e. 3 paths by 3 POV configurations. The use of the HTS was alternated between trials, so there is an average of 3.5 users for every possible combination of paths, POV and pan & tilt. The total amount of trials is then 63 (7 users times 9 trials).

To avoid the influence between experiments, the user never made two trials in a row (the user distribution is interlaced): rather, we tried to maximize the time between two successive trials.

The scene could be observed via three different POV, each of them corresponding to a different [tilt, height] pair (see Table 5.1(a)). The height is referred to the ground level and the tilt angle is referred to the horizon: the higher the value is, the more the camera is looking down. Note that the users could not perform “self-observation”, thus they were not able to develop any possible new proprioceptive model. After every trial, the users were asked to draw the shape of the path. Finally, once all trials were finished, the users filled a short form with questions regarding to the subjective perception of the experiment.

The UGV used during testing was a small vehicle (0.27m length x 0.32m width) which was built using a commercial platform. This base has four motored wheels without steering system. The speed control of each wheel is used to steer the vehicle. Figure 7(a) shows a picture of the UGV. The pan & tilt camera system was placed in a vertical guide to change the height of the camera. This system uses a manual configuration since the height was only changed between experiments and not during them. The webcam has a standard resolution of 640x480 pixels and a horizontal FOV of 36 degrees. For the experiments the frame capture was made at 15 frames per second.

The user interface is composed by three main elements:

- Head Mounted Display. The user watched the images acquired by the UGV's webcam through a HMD system (see figure 7(b)).
- Joystick. The user only controlled the steering of the UGV, since it travels at constant speed. To make the control as natural as possible, the vertical rotation axis of the joystick was chosen (see figure 7(b)). The joystick orientation was recorded during the experiments.
- Head Tracking System. To acquire the user's head movement when controlling the pan & tilt movement of the camera, a wireless inertial sensor system was used (see figure 7(b)). The head orientation was also recorded during the experiments.

During the experiments, the position and rotation of the UGV as well as the movements of the UGV's webcam were recorded at 50Hz using an optical motion capture system (Vicon). Such system acquires the position of seven markers placed on the UGV (see Figure 7(a)) by means of 10 infrared cameras (8 x 1.3Mpixel MX cameras and 2 x 2Mpixel F20 cameras). The raw data coming from this system was then properly reconstructed and filtered to extract the robot center. The user's input (joystick and HTS) was recorded with a frequency of 10Hz, since that is the rate of the UGV's commands. To analyse the data, this information was resampled to 50Hz with a linear interpolation.

Three different paths were used in the experiments because we intend to compare the results in different conditions and across different styles and path complexities. They were placed under the Vicon system, covering a surface of about 13 square meters. The first path (figure 8(a)) is characterised by merlon and sawtooth angles. The second path (figure 8(b)) has the same main shape of the first but is covered CCW by the robot and has rounded curves of different radius. The third (see figure 8(c)) is simpler with wider curves and with rounded and sharp curves. The table 5.1(b) shows a measure comparison between paths.

(a) Points of view

	1	2	3
Height (m)	0.073	0.276	0.472
Tilt angle (deg)	1.5	29.0	45.0

(b) Paths

	1	2	3
Length (m)	19.42	16.10	9.06
Width (m)	0.28	0.28	0.42

Table 1. Experimental constraints (a) and paths data (b)

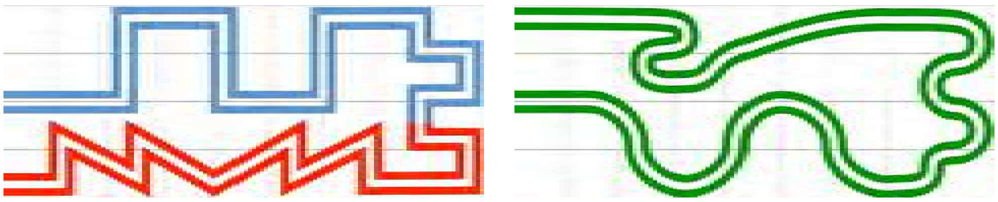
Results

All the detailed results and their analysis can be found in [BOM+09]. In this work we found that performances of a basic teleoperation task are influenced by the viewpoint of the video feedback. Future work will investigate how the height and the fixed tilt of the viewpoint can be studied separately, so that the relative contribution can be derived. The metric we used allows us to distinguish between a tightly and a loosely followed path, but one limitation is that we still know little about the degree of anticipation and the degree of integration of the theoretical path that an operator can develop.

(a) General view of the UGV, with Vicon markers

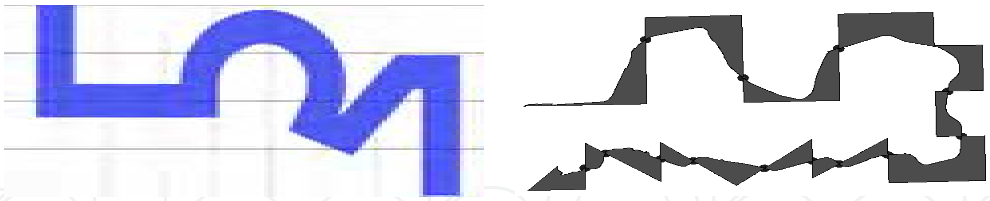
(b) Operator during the experiments

Fig. 7. Experimental setup



(a) Path 1

(b) Path 2



(c) Path 3

(d) S Metric

Fig. 8. Paths used for the experiments and the S metric applied to Path 1

Furthermore, we have shown that, non intuitively, the effects of a HTS were detrimental for performance: we speculate that results with an active HTS could be negative because we constrained velocity to be fixed. On one side, in fact, we added two degrees of freedom and approached a human-like behaviour, on the other side we forced the user to take decisions at an arbitrary, fixed, and as such unnatural speed, thus conflicting with the given freedom. However, we point out that the operators who positively judged an active HTS also spontaneously used the first seconds of the experiment to watch the global path, then concentrated on the requested task. The results coming from HTS could also be biased by the absence of an eye-tracking system, as the true direction of attention is not uniquely defined by the head orientation. From the questionnaire, the post-experiments drawings and from further oral comments, we can conclude that operators cannot concentrate both on following and remembering a path. This is a constraint and a precious hint for future considerations about possible multi-tasking activities. Globally speaking, our evaluations show that good performances imply that self-judgement about performance can be reliable, while the sole judgements are misleading and cannot be used as a measure of performance and no implications can be derived from them. This confirms the motivation of our study about the need of quantitative measures for teleoperation purposes.

5.2 Self representation of remote environment, localisation from 2D and 3D

The ability for teleoperators to localise remote robots is crucial: it allows them situation awareness and presence feeling and precedes any navigation or other higher level tasks. Knowing the robot's location is necessary for the operator to interact and decide about the actions to achieve safely. For some situations, mainly when the remote environment has changed or due to inherent localisation sensors uncertainty, the robot is unable to give its location neither his context. Thus, placing the robot within the tele-operator's map is meaningless. We compare here two video-based localisation methods. A tele-operator is wearing a helmet displaying a video stream coming from the robot. He can move the head freely to move the remote camera allowing him to discover the remote environment. Using a 2D map (a top view) or an interactive 3D partial model of the remote world (the user can move within the virtual representation), the tele-operator has to specify the supposed exact place of the robot.

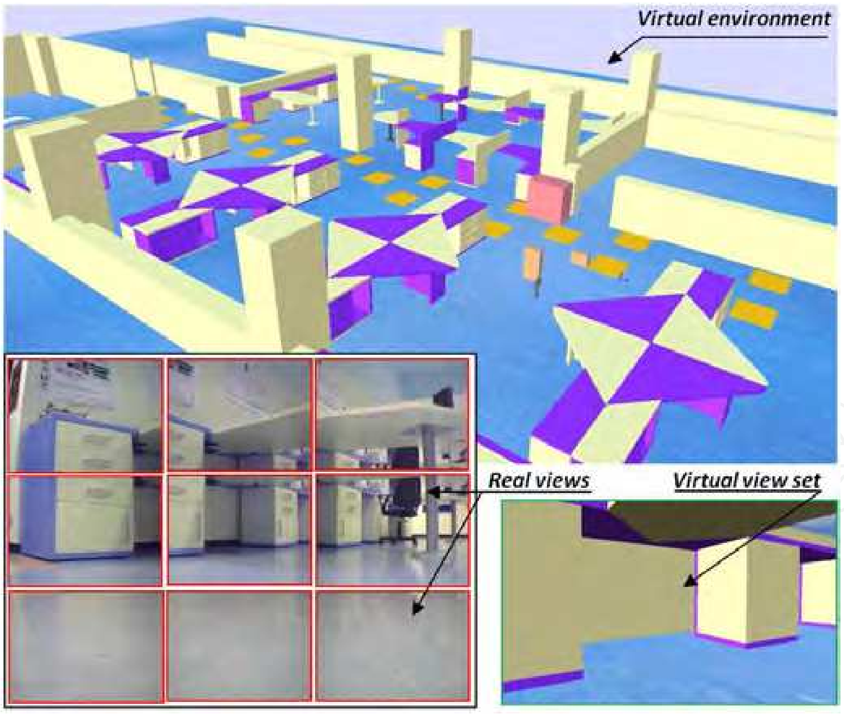


Fig. 9. 3D virtual environment used in the experiments, with a real example of robot's localization

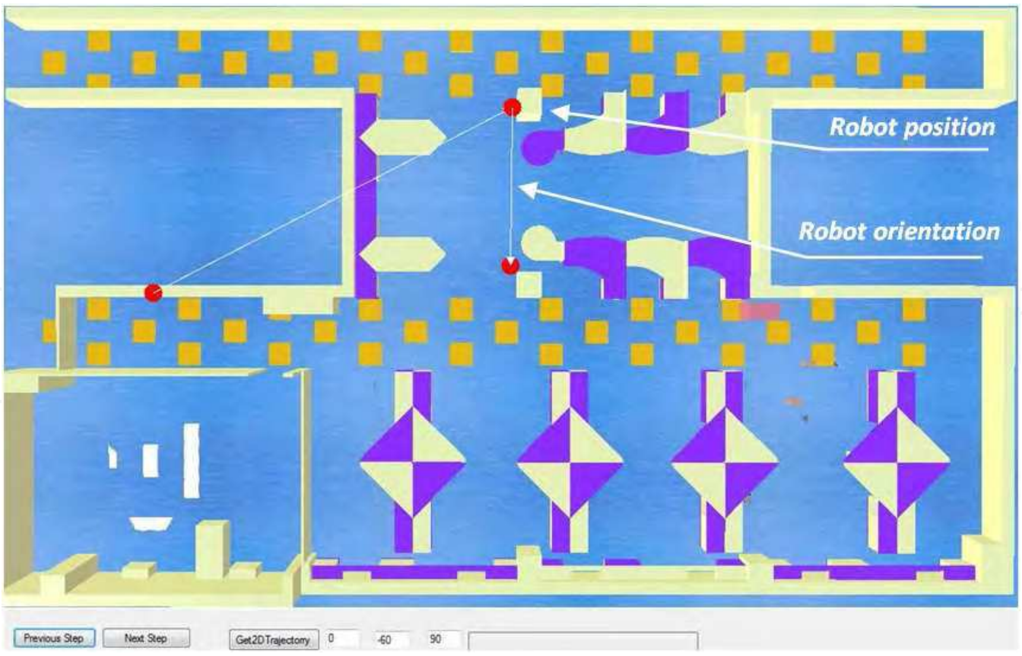


Fig. 10. 2D map used in the experiments for localization of remote robot

Experimental Setup

In 2D maps, subjects have a global view and can indicate their position and orientation on the map. In 3D display case, subjects navigate within the virtual environment till they reach their supposed position and orientation. Our laboratory was selected as the working environment for the experiments. This last contain objects of different dimensions, poses, colors and geometries such as office cabins, furniture, walls, windows, plants , etc. Practically, the robot provides the current video stream and users can move their heads with the HMD, consequently moving the robot's camera. Subjects were requested to explore the video in a minimum time and then to move to 2D maps or 3D virtual environment to localise the robot. The possibility of naturally moving the robot camera as their own head allows users to perceive information and to feel immersed in the distant location and thus to find out their location.

We have evaluated 10 subjects with 10 positions each in both 3D virtual environment and in 2D map. Thus in total the experimental scenario included 200 positions. A test session allowed the subjects to understand the meaning of the tasks and let them familiar with the 3D environment and interface.The following experimental data has been recorded: the time taken by the subject to find the robot's position, the difference between perceived and real positions in the 3D virtual environment as well as in 2D map, the perceived orientation errors with respect to the actual robot's orientation.

The first experiment aimed at finding the robot's location in the 3D environment. Subjects can navigate inside the virtual environment and then set the derived position of the robot. Figure 9 shows an example of real robot's views and the corresponding virtual view set in the 3D environment.

In the second experiment, there was only a 2D top view map. The subjects had to imagine the corresponding view and projection of the 2D points on the map to identify the view that they can see through the robot's camera. Then they pointed out the final chosen location and orientation on the 2D map (fig. 10).

Ten people of different laboratories (engineers, technicians and PhD students) have been selected as subjects for the two experiments. The subjects' age ranged from 23 to 40 years. The percentage of males was 80% and females was 20%. This variance of subjects provided a good sample space for this preliminary experimental study.

Results

Quantitative results corresponding to the 3D environment and 2D map localisation are presented in figures 11 and 12. The errors in position and orientation during the localisation of remote robot by the subjects have been noticed. As well, time (fig. 11) spent by different subjects to localise the robot has also been considered.

When using the 3D interactive environment, the average of the position error was of 48.5 cm with 2.5 degrees of orientation error. In the 2D map, the average value of position error was 100.85 cm with 5.7 degree of orientation error, so the position-orientation errors in 2D map is higher than 3D virtual environment. Possible reason for this fact is that 3D environment is richer in terms of features and landmarks subjects can rely on to derive more accurate robot position-orientation (fig. 12). In other words, the correlation between real (e.g. video data) and virtual environment representation is more effective. On the other hand, the average time consumed by all subjects in 3D was greater than 2D map. This could be due to two facts:

- the time spent in navigation in the 3D environment,
- the (quick) global view approach through the 2D top view map.

Another important result concerns personal variability: on one hand almost all subjects have the same observation concerning the time consumption and the position-orientation errors in the 3D compared to 2D. This observation could be more related to the nature of the two interfaces rather than subjects skills. On the other hand, there is a variability inter-persons: the execution time for each subject is different from others. For example, the subject number 1 has taken 117.8s to find the position in 3D and 108.5s in 2D while the subject number 7 has taken 59.8s in 3D and 30.1s in 2D.

When considering position-orientation errors, subjects' performances has been found significantly different (fig. 11b) and no correlation between 3D and 2D errors were found: subjects made big errors in 2D based localisation and perform well when using 3D environment and inversely.

The last point to notice is the distribution of the global performances in both 3D and 2D based localisation. The standard deviation is much smaller for the 3D case than for the 2D one. This suggests that the solution space in 3D is smaller than in the 2D case, and that subjects rely on the richness of the 3D environment to eliminate false estimations. This could be seen also when considering the ratio between navigation time taken by the subjects and the position error. This last in 3D is about (0.47197cm/s) is almost half of the 2D one (0.91681 cm/s). Similarly we observed that the ratio of orientation error and time consumption reduces significantly when the subjects navigate in 3D compared to 2D map.

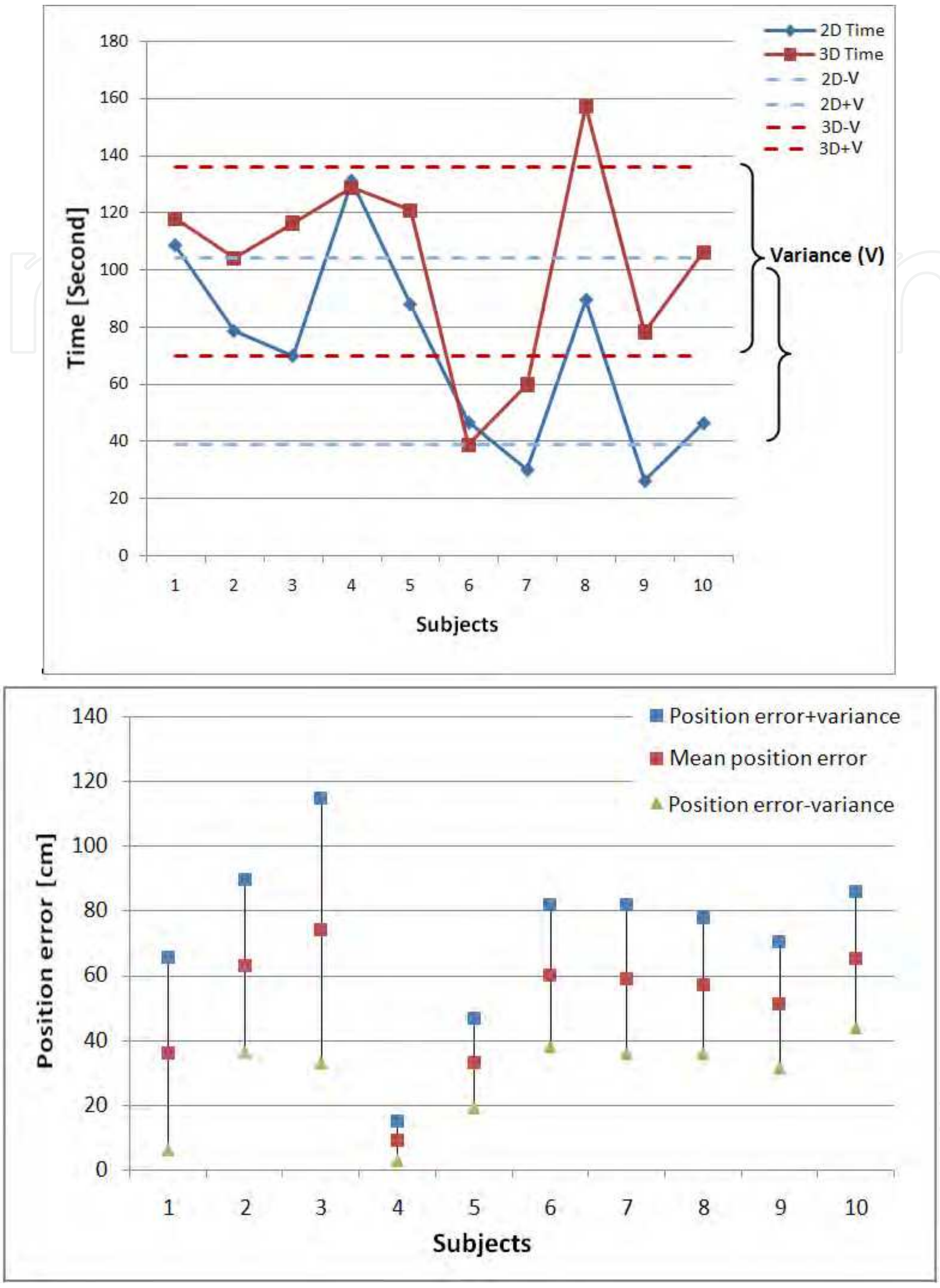


Fig. 11. (a)Average value of the time taken by subjects to find the robot. (b) Average value of position error and variance of each subject corresponding to the 3D environment interaction

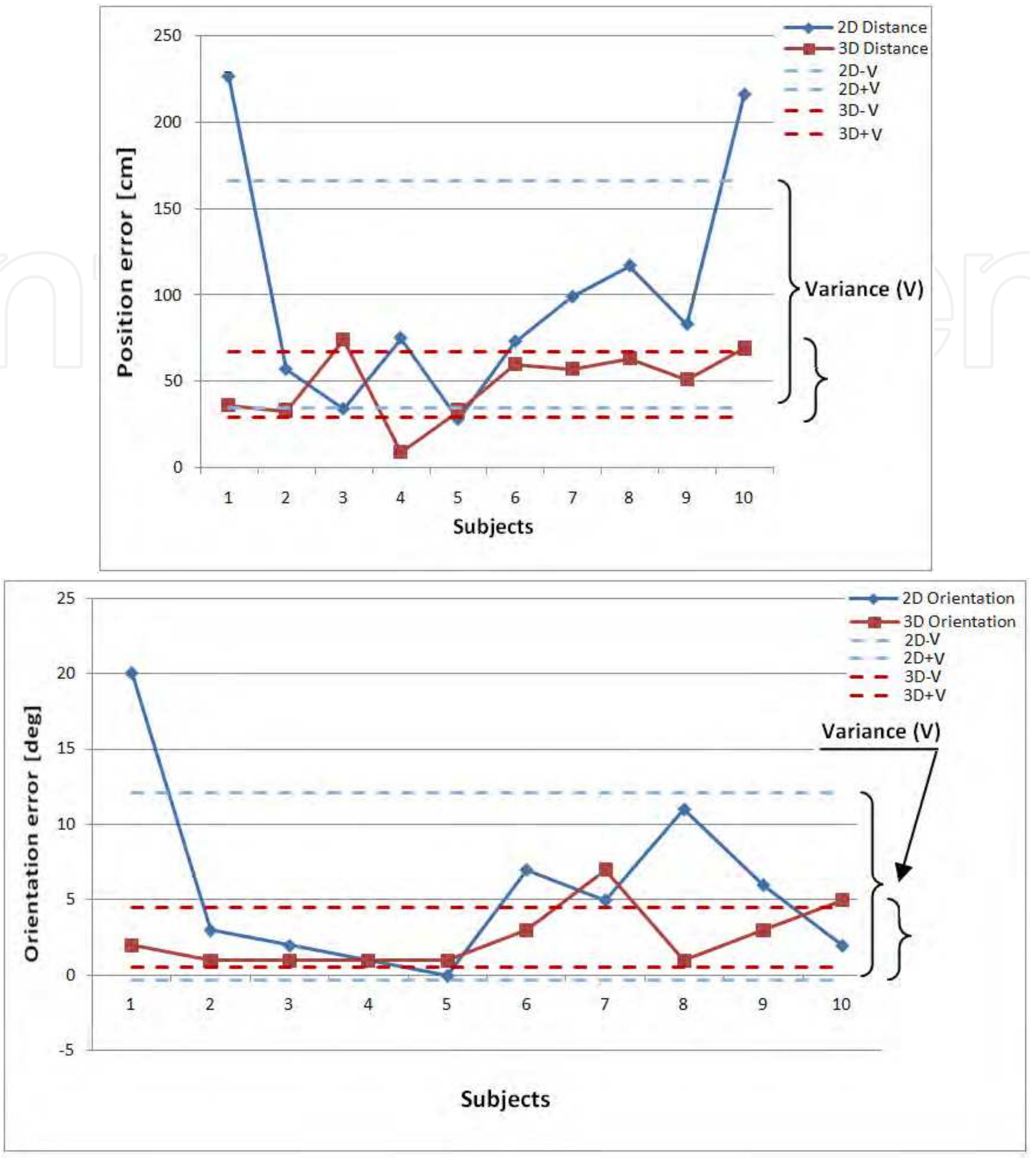


Fig. 12. (a) Average value of position error of each subject. (b) Average value of the orientation error of each subject

6. Example of a N*M real Application: improving immersion in artwork perception by mixing Telerobotics and VR

The ViRAT platform proposes now several demonstrations that are focused on interfaces or human analysis, and takes in account a set of experiments that aim to help the design of adaptative interfaces for teleoperators. We are also deploying some real-case projects with this platform. One of those is a collaboration with a museum, where the major goal is to offer the ability for distant people to visit a real museum. We'll see in this section that we are interesting in improving the sensation of immersion of real visits for virtual visitors, and that such a system may have different usages such as surveillance when the museum is closed.

The existing VR system for virtual visits of museum, like the excellent Musée du Louvre[BMCd], are still limited, with for example the lack of any natural light conditions in the Virtual Environment. Another interesting point is that the user is always alone in exploring such virtual worlds. The technologic effort to make an exploration more immersive should also take into account such human's factors: should navigation compromise with details when dealing with immersion? We believe this is the case. Does the precise observation of an artwork need the same precise observation during motion? Up to a certain degree, no. We propose a platform able to convey realistic sensation of visiting a room rich of artistic contents, while demanding the task of a more precise exploration to a virtual reality-based tool.

Fig. 13. A robot, controlled by distant users, is visiting the museum like other traditional visitors.

6.1 Deployment of the ViRAT platform

We deployed our platform according to the particularities of this application and the museum needs. Those particularities deal mainly with high-definition textures to acquire for VR, and new interfaces that are integrated to the platform. In this first deployment, consisting in a prototype which is used to test and adapt interfaces, we only had to install two wheeled robots with embedded cameras that we have developed internally (a more complete description of those robots can be found in [MC08]), and a set of cameras accessible from outside through internet (those cameras are used to track the robot, in order to match Virtual Robots locations and Real Robots locations). We modelled the 3D scene of the part of the museum where the robots are planned to evolve. A computer, where the ViRAT platform is installed, is used to control the local robots and cameras. It runs the platform, so the VR environment. From our lab, on a local computer, we launch the platform which uses internet to connect to the distant computer, robots and cameras. Once the system is ready, we can interact with the robots, and visit the museum, virtually or really.

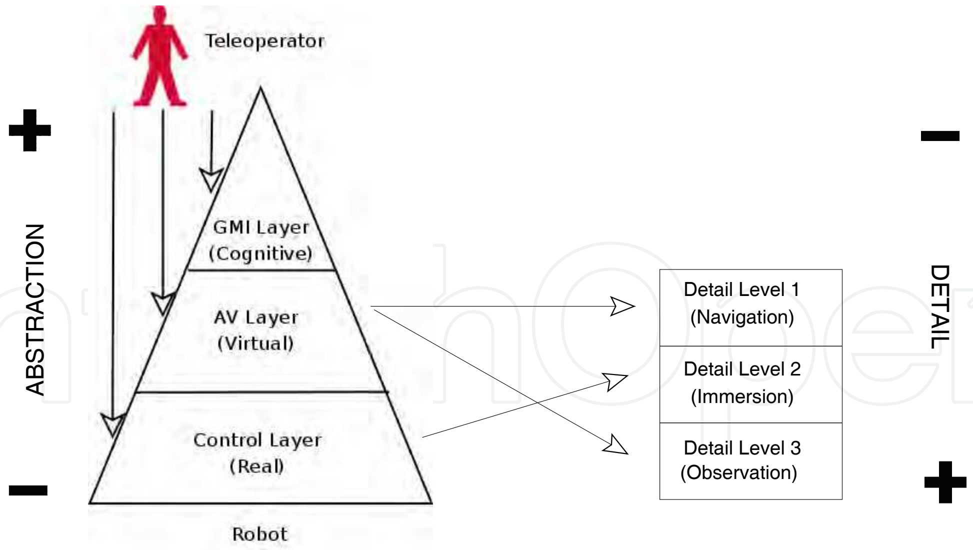


Fig. 14. Different levels of abstraction mapped into different levels of detail.

6.2 Usage of Telerobotics and VR for artwork perception

As presented in [BMCd], existing works with VR offer the ability to virtually visit a distant museum for example, but suffer from lacks of sensations: first, users are generally alone in the VR environment, and second, the degree and sensation of immersion is highly variable. The success of 3D games like «Second Life» comes from the ability to really feel the virtual world as a real world, where we can have numerous interactions, in particular in meeting other real people. Moreover, when we really visit a place, we have a certain atmosphere and ambience, which is fundamental in our perception and feeling. Visiting a very calm temple with people moving delicately, or visiting a noisy and very active market would be totally different without those feedbacks. So, populating the VR environment was one of the first main needs, especially with real humans behind those virtual entities. Secondly, even if such VR immersion gives a good sensation of presence, so of a visit, we're not really visiting the reality. Behind Second Life virtual characters, we have people sit down, in front of their computer. What about having a bijection between the reality and the virtuality ? Seeing virtual entities in the VR environment and knowing that behind those entities the reality is hidden, directly increases the feeling of really visiting, being in a place. Especially when we can switch between virtual world and real world.

Following those comments, the proposed system mixes VR and Reality in the same application. The figure 14 represents this mix, its usage, and the adaptation we made of our general framework.

On the left part, we have the degree of immersion, while on the right part, we have the level of details. The degree of immersion is made of the three levels[MBCK08]: Group Management Interface, Augmented Virtuality and Control Layer:

- First, the GMI layer, still gives the ability to control several robots. This level could be used by distant visitors, but in the actual design it is mainly used by people from the museum to take a global view on robots when needed, and to supervise what distant visitors are doing in the real museum.

- Second, the Augmented Virtuality layer, allows the user to freely navigate in the VR environment. It includes high-definition textures, coming from real high-definition photos of the art-paintings. This level offers different levels of interactions: precise control of the virtual robot and its camera (so as a consequence, the real robot will move in the same way), ability to define targets that the robot will reach autonomously, ability to fly through the 3D camera in the museum, etc.
- Third, the Control layer. At this levels, teleoperators can control directly the robots, in particular the camera previously presented. Users can see directly like if they were located at the robot's location. This level is the reality level, the users are immersed in the real distant world where they can act directly.

Fig. 15. Detail Level 1 is purely virtual, and is the equivalent of the reality (Detail Level 2)

Fig. 16. Detail Level 3 (high detail) is purely virtual, with high-resolution pictures as textures. This one is used in the scene of the figure 15

On another hand, on the right part of the figure 14, the level of details represents the precision the users perceive of the environment:

- Detail Level 1 represents mainly an overview of the site and robots for navigation. The figure 15 shows the bijection between virtual and real, so the usage that a distant visitor can have of the virtual world as an abstraction of the real word.
- Detail Level 2 represents the reality, seen through the robots cameras. At this level of details, users are limited by the reality, such as obstacles and cameras limitations. But they are physically immersed in the real distant world.
- Detail Level 3 is used when distant visitors want to see very fine details of the art-paintings for example, or any art-objects that have been digitalised in high-definition. We can see in figure 16 a high-definition texture, that a user can observe in the virtual world when he wants to focus his attention on parts of the art-painting of the figure 15, that could not be accessible with the controlled robots.

When distant visitors want to have an overview of the site, and want to move easily inside, or on the opposite when they want to make a very precise observation of one art-painting for example, they use the two Detail Levels 1 and 3, in the Virtual Environment. With this AV level, they can have the feeling of visiting a populated museum, as they can see other distant visitors represented by other virtual robots, but they do not have to fit with real problems like for example occlusions of the art-painting they want to see in details due to the crowd, or displacement problems due to the same reasons. On another hand, when visitors want to feel themselves more present in the real museum, they can use the Detail level 2. This is the point where we mix Telerobotics with Virtual Reality in order to improve the immersion.

7. Conclusion

We presented in this paper our approach for designing N*M interactions pattern, and especially our objective analysis of the Human to make the interface cope with users, rather than the opposite. We presented our innovative platform, ViRAT, for an efficient teleoperation between several teleoperators and groups of robots, through adaptative interfaces. We introduced in this system our vision and usage of different levels of interactions: GMI with a scenario language, AV and direct control. We briefly presented the CVE we developed to model the robots activities and states, an environment where teleoperators can have collaborative an intermediate level of interactions with the real distant robots by using the virtual ones. We then presented in details the current experiments that are conducted to make a precise evaluation of the human's perception, to design and choose adaptative interfaces that will be objectively adapted to each teleoperator, according to contexts of tasks. We finally presented one deployment of this platform for an innovative artwork perception proposed to distant visitors of a museum. Our project is currently very active and new results come frequently. As the technical environment is ready, our actual experiments are clearly turned on human's perception evaluation, aiming the case of complex interactions with groups of robots.

We would like to make some special acknowledgments to Delphine Lefebvre, Baizid Khelifa, Zhao Li, Jesus Ortiz, Laura Taverna, Lorenzo Rossi and Julien Jenvrin for their contributions in the project and the article. The locations for our platform in the museum application are kindly provided by Palazzo Ducale, Genova.

8. References

[BMCd] L. Brayda, N. Mollet, and R. Chellali. Mixing telerobotics and virtual reality for improving immersion in artwork perception. In *Edutainment - Banff - Canada*, 2009 - to be published-.

[BOM+09] L. Brayda, J. Ortiz, N. Mollet, R. Chellali, and J.G. Fontaine. Quantitative and qualitative evaluation of vision-based teleoperation of a mobile robot. In *ICIRA 2009*, 2009.

[EDP+06] Alberto Elfes, John Dolan, Gregg Podnar, Sandra Mau, and Marcel Bergerman. Safe and efficient robotic space exploration with tele-supervised autonomous robots. In *Proceedings of the AAAI Spring Symposium*, pages 104 – 113., March 2006. to appear.

[GMG+08] Stephanie Gerbaud, Nicolas Mollet, Franck Ganier, Bruno Arnaldi, and Jacques Tisseau. Gvt: a platform to create virtual environments for procedural training. In *IEEE VR 2008*, 2008.

[GTP+08] J.M. Glasgow, G. Thomas, E. Pudenz, N. Cabrol, D. Wettergreen, and P. Coppin. Optimizing information value: Improving rover sensor data collection. *Systems, Man and Cybernetics, Part A, IEEE Transactions on*, 38(3):593–604, May 2008.

[HMP00] S. Hickey, T. Manninen, and P. Pulli. Telereality - the next step for telepresence. In *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics (VOL 3) (SCI 2000), pp 65-70, Florida.*, 2000.

[KTBC98] A. Kheddar, C. Tzafestas, P. Blazevic, and Ph. Coiffet. Fitting teleoperation and virtual reality technologies towards teleworking. 1998.

[KZMC09] B. Khelifa, L. Zhao, N. Mollet, and R. Chellali. Human multi-robots interaction with high virtual reality abstraction level. In *ICIRA 2009*, 2009.

[LKB+07] G. Lidoris, K. Klasing, A. Bauer, Tingting Xu, K. Kuhnlenz, D. Wollherr, and M. Buss. The autonomous city explorer project: aims and system overview. *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ* International Conference on, pages 560–565, 29 2007-Nov. 2 2007.

[MB02] Alexandre Monferrer and David Bonyuet. Cooperative robot teleoperation through virtual reality interfaces. page 243, Los Alamitos, CA, USA, 2002. IEEE Computer Society.

[MBCF09] N. Mollet, L. Brayda, R. Chellali, and J.G. Fontaine. Virtual environments and scenario languages for advanced teleoperation of groups of real robots: Real case application. In *IARIA / ACHI 2009, Cancun,* 2009.

[MBCK08] N. Mollet, L. Brayda, R. Chellali, and B. Khelifa. Standardization and integration in robotics: case of virtual reality tools. In Cyberworlds – Hangzhou - China, 2008.

[MC08] N. Mollet and R. Chellali. Virtual and augmented reality with headtracking for efficient teleoperation of groups of robots. In *Cyberworlds - Hangzhou* - China, 2008.

[MCB09] N. Mollet, R. Chellali, and L. Brayda. Virtual and augmented reality tools for teleoperation: improving distant immersion and perception. *ToE Journal*, 5660:135–159, 2009.

[SBG+08] A. Saffiotti, M. Broxvall, M. Gritti, K. LeBlanc, R. Lundh, J. Rashid, B.S. Seo, and Y.J. Cho. The peis-ecology project: Vision and results. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 2329–2335, Sept. 2008.

[Tac98] S. Tachi. Real-time remote robotics-toward networked telexistence. *Computer Graphics and Applications, IEEE*, 18(6):6–9, Nov/Dec 1998.

[UV03] Tamas Urbancsek and Ferenc Vajda. Internet telerobotics for multi-agent mobile microrobot systems - a new approach. 2003.

[Wer12] M. Wertheimer. Experimentelle studien ber das sehen von bewegung,. *Zeitschrift fr Psychologie*, 61:161265, 1912.

[WKGK95] K. Warwick, I. Kelly, I. Goodhew, and D.A. Keating. Behaviour and learning in completely autonomous mobile robots. *Design and Development of Autonomous Agents, IEE Colloquium on*, pages 7/1–7/4, Nov 1995.

[YC04] Xiaoli Yang and Qing Chen. Virtual reality tools for internet-based robotic teleoperation. In DS-RT '04: *Proceedings of the 8th IEEE International Symposium on Distributed Simulation and Real-Time Applications*, pages 236–239, Washington, DC, USA, 2004. IEEE Computer Society.

[ZM91] S. Zhai and P. Milgram. A telerobotic virtual control system. In *Proceedings of SPIE, vol.1612, Cooperative Intelligent Robotics in Space II, Boston*, pages 311–320, 1991.

Remote and Telerobotics
Edited by Nicolas Mollet

ISBN 978-953-307-081-0
Hard cover, 220 pages
Publisher InTech
Published online 01, March, 2010
Published in print edition March, 2010

Any book which presents works about controlling distant robotics entities, namely the field of telerobotics, will propose advanced technics concerning time delay compensation, error handling, autonomous systems, secured and complex distant manipulations, etc. So does this new book, Remote and Telerobotics, which presents such state-of-the-art advanced solutions, allowing for instance to develop an open low-cost Robotics platform or to use very efficient prediction models to compensate latency. This edition is organized around eleven high-level chapters, presenting international research works coming from Japan, Korea, France, Italy, Spain, Greece and Netherlands.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Nicolas Mollet, Ryad Chellali, and Luca Brayda (2010). Choosing the Tools for Improving Distant Immersion and Perception in a Teleoperation Context, Remote and Telerobotics, Nicolas Mollet (Ed.), ISBN: 978-953-307-081-0, InTech, Available from: http://www.intechopen.com/books/remote-and-telerobotics/choosing-the-tools-for-improving-distant-immersion-and-perception-in-a-teleoperation-context

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821