Comparative Analysis of 3D Robot Teleoperation Interfaces with Novice Users

Daniel Labonté, *Member, IEEE,* Patrick Boissy, *Member, IEEE*, François Michaud, *Member, IEEE*

Abstract — **Being able to act remotely in our homes could be very useful in providing various services such as surveillance and remote interventions, which are key features for telehomecare applications. In addition to navigation and environmental challenges such that a telepresence robot would face in home settings, the system requires an appropriate teleoperation interface for safe and efficient usage by novice users. This paper describes the design criteria and characterizes visualization and control modalities of user interfaces with a real robot. By considering the user's needs along with the current state-of-the-art in teleoperation interfaces, two novel mixed reality visualization modalities are compared with standard video-centric and map-centric perspectives. We report teleoperation trials under six different task scenarios with a sample of 37 novice operators in home-like conditions. Results based on three quantitative and one qualitative metrics outline under which conditions the novel mixed reality visualization modalities significantly improve the performances of novice users.**

Index Terms— **User interface, teleassistive robotic, virtual reality, mixed perspective**

I. INTRODUCTION

Telepresence in a real or a virtual environment is defined as the experience involving the displacement of a remote operator's perception into a computer-mediated environment [1], [2]. In the context of telepresence with a mobile robot system, interaction between the remote operator and the environment (including people in the environment) is carried through the robot, which provides mobility to sensors (e.g., vision, audition) and actuators (e.g., manipulator, speakers) installed on the mobile platform. System performances are directly affected by the sensory, telecommunication, processing, visualization, control, decisional, interaction and operator capabilities, and the design challenge is to come up with the most appropriate configuration(s) for a given application and operating environment.

With teleoperated robotic systems natural settings, safety of people, of the environment and of the robot are primary concerns [3]. Teleoperation of the mobile robot can help alleviate these concerns and decrease the technological challenges of making a completely safe autonomous platform operating in uncertain and changing environments (e.g., small spaces, multiple obstacles, different surfaces, stairs, people, and animals). Having a human in the loop decreases the challenge of reaching these requirements compared to making a service robot completely autonomous. However, operation of teleoperated robotic systems is influenced greatly by operator experience and context of use of the mobile robot [4].

Visual information and control mechanisms implanted in the interface have great impact on performances during teleoperation [5], [6]. For example, in urban search and rescue activities, low situation awareness during teleoperation leads to inefficiency, confusion and human errors of the operator [6], [4]. Video user interfaces are common in telerobotic systems, and intelligent interfaces are becoming increasingly important as users face increasing system complexity and information overload. An optimal teleoperation user interface in such conditions must provide pertinent information about the system states and conditions (e.g., objects, persons, free space, data) in conjunction with an efficient command system to the operator, with a reasonable cognitive load for sustainable and adequate uses.

Following an iterative design methodology, with the ultimate objective of designing a complete telepresence robotic system for telehomecare applications [7], [8], [9], for the first design cycle we decided to study the different elements of such a system by conducting focus groups with home care practitioners and older adults to identify potential applications [11], and interviewing system users to derive the interfaces required for the health information architecture [12]. Safety, usability for untrained, infrequent and novice operators and efficiency came out as important design criteria for the system. We also conducted a pilot study with two telepresence robotic systems [10], confirming that the interface design is a limiting factor for efficiency and security of operation. This study identified that visual information and control mechanisms implanted in the interface have great impact on performances during teleoperation. It also provided us with key insights on the experimental methodology to adopt in conducting experimental trials in real conditions with novice users.

Manuscript received December 10, 2008. This work was supported in part by the Canada Institutes for Health Research – Musculoskeletal Health and Arthritis, Grant 120591.

D. Labonté was with the Department of Electrical Engineering and Computer Engineering, Université de Sherbrooke, Québec, Canada (819-829-3785; e-mail: daniel.labonte@usherbrooke.ca).

F. Michaud holds the Canada Research Chair (CRC) in Mobile Robotics and Autonomous Intelligent Systems. He is with the Department of Electrical Engineering and Computer Engineering, Université de Sherbrooke, Québec, Canada, (819 821-8000 x62107, e-mail: Francois.Michaud@usherbrooke.ca).

P. Boissy is with the Department of Kinesiology, Université de Sherbrooke, Québec, Canada (821-1170 x45628, e-mail: Patrick.Boissy@usherbrooke.ca).

In continuation with this pilot study, this paper presents our novel user interface designed specifically to address requirements from their use by novice operators. Based on important concepts and critical considerations in designing teleoperation interfaces, outlined in Section II, we made the hypothesis that mixed reality interfaces (i.e., interfaces combining virtual reality with real sensory information taken from the robot) would improve the performance of novice users. Two novel display modalities have been developed: one superimposes an ego-centric video display on a 3-dimensional (3D) virtual model of the world providing a wider perspective of the operating environment, and the other is an exo-centric 3D map representation mixed with video images coming from the robotic platform. These interfaces are presented in Section III. Section IV presents the experimental protocol to conduct a comparative evaluation of the interfaces with a statistically significant group of novice operators in a special set up reproducing home environment conditions, examining precision, efficiency, cognitive workload and usability. Results are presented in Section V, followed with a discussion in Section VI, conclusions and future work.

II. Design Guidelines for Teleoperation Interfaces

The design of teleoperation interfaces may be influenced by a broad set of factors ranging from human or technological factors, addressed in a variety of different fields such as human computer interaction (HCI) [8], [13], and human robot interaction (HRI) [14]. Providing a complete survey of these fields is outside the scope of this work. The objective of this section is to identify the key elements in deriving requirements for the teleoperation interfaces designed for our telepresence mobile robotic system for home care assistance.

The principal human factors found in the literature that should be taken into consideration in the design of a teleoperation interface for a mobile robot are presented in Table I. These concepts are:

- Level of cognition [20], which range from skill-based to rule-based and knowledge-based.
- Situation awareness (SA) [21], [22], [23] during telepresence, defined as being the perception of the robot's location, surroundings and status, the comprehension of their meaning, and the projection of how the robot will behave in the near future.
- Types of tasks, which can be navigating and situation understanding for the first two levels of cognition [24] and anticipating and planning for the third level.
- Cognitive memory model [17], starting with short-term memory for skill-based level, to long-term memory using a collection of mental models such as rule or knowledge. A cognitive memory model is defined as "… *an internal representation employed to encode, predict, and evaluate the consequences of perceived and intended changes to the operator's current state within a dynamic environment*." [17].

TABLE I
HUMAN FACTORS IN INTERFACE DESIGN

Level of cognition	Situation Awareness (SA)	Type of tasks	Cognitive model
Skill-based	Perception	Navigating	Working space / Short term memory
Rule-based	Comprehension	Situation understan-ding	Long term memory / Rules
Knowledge-based	Projection	Anticipa-ting and planning	Long term memory / Knowledge

TABLE II
GUIDELINES FOR TELEOPERATION INTERFACE DESIGN

Elements	Guidelines
PERCEP-TION	*Provide a frame of reference to determine the robot's position in its environment [8]. -Memorize in a map where the robot has been [9]. -Ability to self-inspect the robot's body for damage or entangled obstacles [8].
DISPLAY	*Information from multiple sensors presented in an integrated fashion [15], [8], [9]. *Complement video stream with feedback information from other sensors [8]. *Minimize the use of multiple windows [9]. -Automatic presentation of contextually-appropriate information, such as automatically switching to a rear camera view if the robot is backing up [8]. *Allow the user to adjust the perspective of the environment to match the task [15]. *Ground the information displayed with the reality [16]. -Indicators of robot status, including which camera is being used, the position(s) of camera(s), etc. [8]. *Provide more spatial information about the robot in the environment [9]. *Display the robot's body in the interface [17]. *Convey the orientation of the video stream with respects to robot orientation [18]. -Implicitly switch interfaces modalities and autonomy [17]. -Allow the user to manipulate the information displayed [17] and to store information [15]. *Help direct the operator's focus of attention [19].
CONTROL	*Provide assistance and autonomous modes [9].
MISCALLE-NOUS	-Let the robot use natural human cues [17]. -Manipulate relationship between robot and world [17]. -Learning mechanisms [17].

Therefore, an optimal design of a teleoperation interface must be able to lower the cognitive load of the operator during navigation tasks, while allowing easy interaction with the control modes of the robot. This is especially true with novice and untrained operators. Situation awareness (SA) is of prime concern [22], [23], [9]. As a general observation, it appears also that efficient perception of the task to accomplish must be easily (unconsciously) acquired at the skill-based level. Integration of precise sensory information is therefore required, presented in natural and usable fashion in order to compensate for limitations in the system's perception and telecommunication capabilities. If rule- or knowledge-based cognitive processes are required, information integration and grounding are key elements to a successful teleoperation interface design. Table II summarizes guidelines derived from our survey of the literature and found to be important for the design of teleoperation interfaces. Most guidelines reported are directed toward Perception and Display, indicating the importance of these elements in teleoperation interfaces. Guidelines identified by asterisks in Table II are those considered in the design of our new interface.

A. *Display Modalities*

The simplest form of a teleoperation interface is simply to provide the operator with a hand-controller (e.g., a joystick) and a video stream taken from vehicle mounted camera(s) [26]. Also known as video-centric (VC) interfaces, they are effective in providing good surroundings and activities awareness [27]. They do however have limitations in terms of SA, navigation, localization and spatial awareness [28]. Compared to direct observation of a scene by a human, simple video displays are limited by the camera(s) field of view, position, resolution, 2D projection with no head compensation during movement, etc. Human operators must compensate for such limitations to maintain SA by inferring a 3D mental model of the world and anticipating what to do for a specific task to accomplish. This requires the operator to use rule- and knowledge-based level, and long-term memory, which consequently increases cognitive load. If the operator is not able to compensate for these limitations, poor SA will be observed which will lead to confusion [9] and affect safe and efficient teleoperation of the robot. Using multiple or omnidirectional/panospheric cameras may improve SA [29], [30], [31], but their usability may be limited by telecommunication bandwidth [32].

Adding another window to the interface to provide position and range information coming from other sensors on the robot may help improve SA. Multimodal and multisensor interfaces try to compensate for such limitations by providing the operator with more information and control modes [26]. The most obvious solution is to use multiple windows presenting information from specific sensors [33]. The operator is able to receive more information about the environment and the robot's status, but most of the time at the cost of increased cognitive load (leading to fatigue, to stress and in extreme cases to incapacity of performing a task) [19]. This is because the operator has to process more low-level perception and integrate this information at a rule- or knowledge-based level to accomplish navigation tasks. It is also important to note that information comes from two separate windows and that each window uses its own information metaphors, which again requires more cognitive processing. This means that a large portion of working memory (short and long term) is dedicated to maintain a sense of what is going on in the world [34].

As another alternative, Virtual Reality (VR) displays have been shown to improve teleoperation performance with reduced telecommunication bandwidth [19], [35]. They however also have limitations [16], [36]. An interesting compromise is to combine real images with virtual reality representation, an approach known as Mixed Reality (MR) [37]. MR includes Augmented Reality (AR) (video displays enhanced by means of computer graphics) and Augmented Virtuality (AV) (enhancement of virtual displays by real images or objects) [37], [38], [39]. Variations on this theme go from a display combining a video stream in the upper part of the display with a 2½D virtual representation of obstacles derived from the robot's sensors [15], to superimposing the video stream onto the 2½D representation based on the position of the camera [17], [18]. It is also important to note the concept of an ecological interface where MR is used to combine different perspectives from video, map and robot view feeds [47]. A distinction is made between 2½D and 3D because range sensors are used to derive 2D information from which a volume (barrel or plane) is derived. Another dimension to consider is the perspective used to display the relationship between the robot and its environment [40]. Mainly, three types exist:

1. **Ego-centric view**, in which the world is presented from an inside-out perspective in relation to the robot (e.g., video stream);
2. **Exo-centric view**, where the world and the robot are presented from an outside-in perspective (e.g., map).

TABLE III
CLASSIFICATION OF DISPLAY MODALITIES

Modality	Type
Video-centric / Augmented Reality	AR2D [41], [42], [43], [27]
	AR3D
	AR3DS [26], [44]
Map-centric / Augmented Virtuality	ME2½D [15], [35], [27]
	ME3D
	ME3DS

Making an association between video-centric and ego-centric and between map-centric and exo-centric views, Table III present a classification of display modalities used in teleoperation interfaces in relation to their dimensionality (2D, 2½D, 3D, or 3DS for interfaces using stereoscopic projection and special screen, generally using goggles). Most of the interfaces are either based on AR in 2D or Mixed perspective Exo-centric (ME) displays in 2½D, but none have been validated using a 3D model of the world. Mixed perspective displays combine in the same window an exo-centric view of the world with ego-centric data (e.g., a video stream), making a connection between the two using a grounding mechanism. For instance, a mixed-perspective interface can position the robot in an exo-centric representation of the world, and make the scene turn and move around the robot while presenting ego-centric video information. Mixed perspective interfaces have demonstrated better performances in avoiding collisions and moving in obstacle-laden spaces [34], [40]. Also, a chase perspective (exo-centric type) is also reported to lower collision with objects better than with a egocentric perspective [46]. No study on AR3D or ME3D display modalities have yet been conducted.

III. 3D INTERFACES FOR NOVICE USERS

Based on the analysis presented in Section II, we decided to study the use of a 3D virtual reality model of the world to improve novice users' performances in teleoperating a mobile robot in home environments. We designed our own virtual reality 3D teleoperation interface with the following reconfigurable display modalities:

- *Video-Centric Display* (VC2D), as illustrated in Figure 1, showing a predominant video-centric view in 2D (center), a map view also in 2D (upper left).
- *Augmented Reality Display* (AR3D), superimposing the video stream on a 3D environment model, as shown in Figure 2, increasing by a factor of two the angle of view. The position of the video stream is derived from the robot's pose.
- *Mixed perspective Exo-centric Display* (ME3D), shown in Figure 3, reducing the size of the video stream by 25% and moving it down, and showing a 3D exo-centric aerial map view of the virtual environment. The robot body remains fixed (grounded) in the center of the window, facing forward, and the 3D model moves in order to correspond to video images position on a lowered and centered position. The video stream shows what is seen in front of the robot.

The rationale for the teleoperation display design was to provide mixed perspectives (ego/exo), mixed realities (real video images/3D virtual model) and grounded references without causing any information obstruction on the 3D model map in front of the robot. To do so we chose to provide simultaneous presentation of 2D video images with a virtual 3D map representation of the operation environment. The operator can use this combined information to perceive, analyze and extract geometric features of the environment. This facilitates comparison with its own mental space model construction. This is not possible with the VC2D modality, in which the direction of the robot moving in the environment must be taken into consideration by the operator (i.e., the left and right sides change when the robot is moving down the map), thus increasing operator's cognitive load. The ME3D combines an exo-centric representation (i.e., aerial display of a 3D virtual model) with ego-centric data (i.e., video stream) grounded by keeping static the robot's orientation in the 3D virtual model.

Figure 1. Video-centric Display (VC2D).

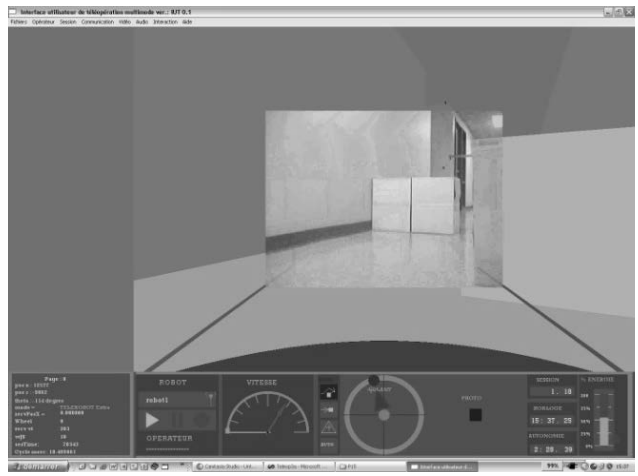

Figure 2. Augmented Reality ego-centric display (AR3D).

The interface is built in C++ language and structured like a game engine where the 3D model of the scene is rendered with the openGL library. The video feed is textured on a specific and well positioned surface. Information to move the robot by the operator through mouse and keyboard commands are transmitted at a minimum of ten times per second. That information is transmitted by TCP/IP protocol on the wireless Ethernet connection to the robot. Location, orientation and speed information from robot sensors are also transmitted at a minimum of five times per second to the interface in order to adjust the position of the robot in the 3D model. An extrapolation algorithm is used to calculate exact positioning of the robot in the 3D model for each rendering of the video, which is ten times more frequent than the message received from the robot. 3D model processing did not significantly influence system latency.

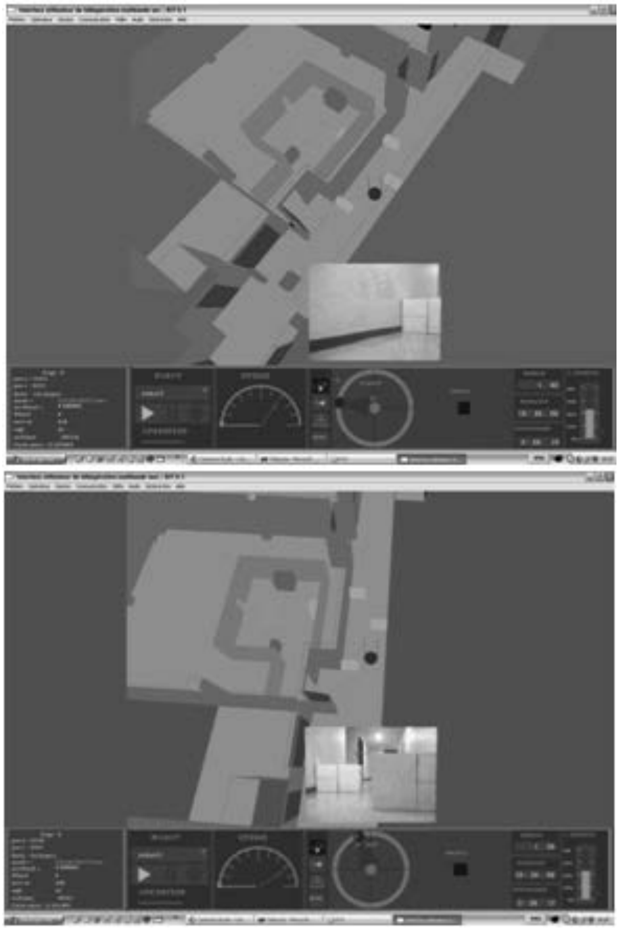

Figure 3. Mixed perspective exo-centric display (ME3D) shown on two consecutive (2 sec) screen shots while rotating the robot.

Even though the focus of our work is on the display modalities, the control interface was designed to integrate different modes and data found in teleoperation interfaces, to eventually analyze the influence of control modalities on teleoperation performances. Located at the bottom part of the interface, the control interface consists of eight panels (from left to right): debug window; robot control; speedometer; mode selection buttons; virtual joystick or drive wheel; photo button; clocks; battery status. Four control mechanisms are possible:

- *Joystick Control* (JC). This replicates the same functions of a physical joystick with one activation button, a common control interface in robot teleoperation. Figure 4 illustrates the mechanism. The virtual joystick is activated by clicking and holding the left mouse button in the drive wheel inner circle. Longitudinal velocity is proportional to the distance of the joystick with the center. Rotational velocity is done in the direction of the joystick. Stopping occurs when the left mouse button is released.
- *Drive Wheel and Speedometer Control* (DWSC). This control mechanism use the same portion of the JC control interface, but use a driving metaphor (decoupling velocity and direction) to navigate the robot The operator can just steer the robot by clicking on the drive wheel panel after or while adjusting the speed with the mouse wheel.
- *Go-Point Control* (GPC). Go-points are set simply by clicking on the map, the 3D model or the video stream. The robot is allowed to move by clicking on the play button. Multiple go-points can be specified and dynamically added or removed.
- *Autonomous Control* (AC). This mode allows the robot to return to a known location (e.g., its charging station).

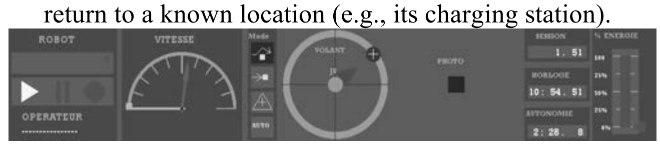

*Figure 4. Dashbo*ard JC mechanism: 50% longitudinal velocity and turning right.

Mode selection buttons allow the operator to use a temporary push mode instead of the default safe mode that simply stops the robot when an object is in the way. The push mode lets the robot come in contact with an object and push it using JC. For security reasons, this mode is automatically reset once the joystick is released.

IV. Experimental Protocol

The following section presents elements of our experimental methodology used to tests the interface designs.

A. *Robotic Platform*

Figure 5 shows our robot, named Telerobot [12], specifically designed for telehomecare applications. Telerobot is essentially a mobile robotic platform mounted with videoconferencing equipment and a pan-tilt-zoom (PTZ) camera on top. The mobile robot base is 0.29 m high and the height of the videoconferencing equipment can be manually adjusted from 0.65 to 0.95 m. To navigate through obstacles and narrow spaces in homes while providing stable and appropriate video feed for teleoperating the platform, we designed a round-shaped differential-drive robot with a rocker-bogie suspension. A specific web camera for navigation is incorporated in the robot base.

Fig. 5. Telerobot, the telepresence robot for home care assistance.

Locomotion is realized using two motorized wheels with their attachment axis going through the center of the robot, allowing it to turn on the spot. Four omnidirectional wheels (two at the front and two at the back of the robot) are attached to the rocker-bogie suspension, in order to stabilize the frame. This mechanism acts as a mechanical filter, preventing the robot to be influenced by punctual irregularities of the ground (compared to a robot with a rigid body, such influences is reduced by close to 50%). It minimizes disturbances of the video streams coming either from the bottom camera for teleoperation (placed underneath the laser range finder), or the cameras fixed on the top shelf, and increase mobility on irregular ground. It also allows the robot to go over large steps (up to 3 cm high).

Telecommunication between the robot and the TCP/IP communication protocol is done through a wireless communication link (802.11b) and a local area network (Ethernet 100 Mbps). A dedicated port is used to stream video from the front facing webcam (15 fps, resolution of 320 × 240 pixels) dedicated to the teleoperation of the robot. A mean delay of 0.4 sec is observed between the time a command is sent to the robot and the time where the robot reports its execution to the interface. The video stream is matched to the virtual environment based on localization information coming from a SICK LMS 200 laser range finder and odometry data. The robot uses CARMEN for localization [45]. A linear extrapolation algorithm is used to position the virtual model every 200 msec cycle (5 fps). The mean error for x or y coordinates of positioning of the robot is 5,7 cm (standard deviation of 2,8 cm). For angular position error of the robot, the mean error is 1,3° (standard deviation of 1,1°). By directly superimposing the video stream onto the 3D model, AR3D is the display most sensitive to the precision of the robot's pose. Misalignment between the two must be minimized to avoid disturbing the operator. Figure 6a present a minimal but still observable misalignment on the position of the real chair with its virtual representation. Figure 6b shows an extreme situation where the angular positioning is important.

Fig. 6. Angular position errors in the AR3D display; a) at left, small error and b) at right, maximum error.

Finally, as shown in Figure 7, a three-level hybrid deliberative-behavioral architecture is used to control the system, with behaviors making the robot (in reverse order of priority) go to a specific location (e.g., going back to a charging station), be teleoperated at a maximum speed of 25 cm/sec while avoiding obstacles all around itself or holes in the floor, or stop moving.

For our trials, the robot was programmed to stop in front of obstacles in order to only measure teleoperation influences on the robot.

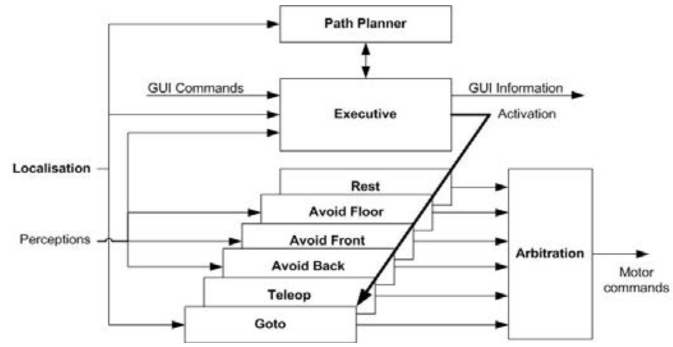

Fig. 7. Telerobot's computational architecture.

B. Experimental Scenario

An experimental scenario to evaluate performances of novice users using the three display modalities was constructed. We set the control mechanism to JC (making the assumption that this is the control mechanism that requires the highest cognitive load), and made the robot stop in front of obstacles (making the operator deal with the experienced situations).

Two teleoperation missions reproducing typical home-like situations, such as: navigating in a small area and requiring multiple rotations; navigating in a corridor and avoiding obstacles; pushing a small door; positioning the robot and passing a door frame were elaborated. The Kitchen setup (Figure 8) has furniture including a table, four chairs, two boxes with sign number and a metal cabinet with a door.

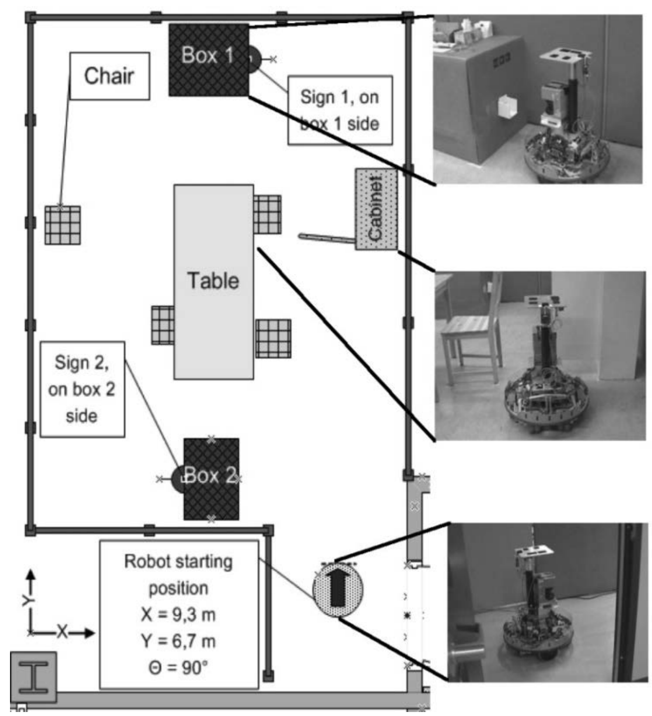

Fig. 8. Kitchen setup.

The Corridor setup (Figure 9) is composed of three sets of boxes placed along the walls for slalom navigation, and a sign on a post at the end of the course. In both missions, the robot started at the same initial position, and the operator had to execute tasks and return to the initial position.

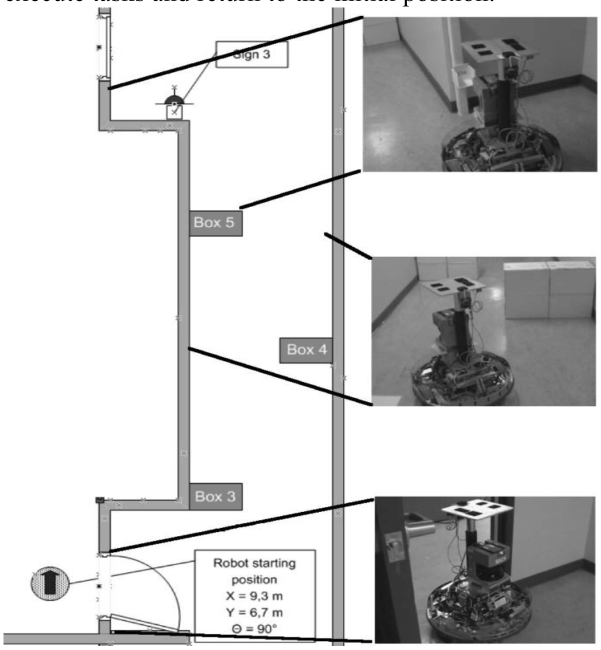

Fig. 9. Corridor setup.

As presented in Table IV, a total of six tasks were selected for the experiment using these setups, three in the kitchen and three in the corridor.

TABLE IV
TASK DESCRIPTIONS

Task	Place	Description
A	Kitchen	Move the robot, push the cabinet door and position the robot in front of sign 1 to take a picture.
B	Kitchen	Move the robot from sign 1 to sign 2, and position it in front of sign 2 to take a picture.
C	Kitchen	Move back the robot from sign 2 to the starting position.
D	Corridor	Move the robot in the corridor and position it in front of the sign 1 to take a picture.
E	Corridor	Move the robot from sign 1 to the first box near the door.
F	Corridor	Move the robot through the door to get back to the starting position.

C. *Experimental protocol*

We conducted standardized field trials of the experimental scenario with a group of untrained operators (n=37) to: 1- analyze the effects of the display modalities on teleoperation performances and 2- explore qualitatively (using a questionnaire) the usability of the system. Each participant received a 5 minute video presentation of the system after visiting the two setups. Following that presentation, they were allowed a maximum of 10 minutes to familiarize themselves with the interface to move the robot in a practice area. To characterize the effect of the display modalities, three quantitative metrics were chosen to evaluate the performances of operators with the system: completion time of each task (t), the distance (d) traveled by the robot and the number of commands (nc) executed by the operator for each task. The time metric gives indication of efficiency; distance approximates precision of movements; number of command relates to the operator's cognitive load.

Participants performed each task three times (one for each visual modality), for a total of 18 trials per participant. Sequencing of tasks and visual modalities during the experimentation were alternated across participants following a Latin square ordering, to prevent bias associated with a fixed order (i.e., the learning effect). Between each task, a small pause was allowed (with no influence on the metrics).

In order to compare performances in terms of a common reference, thus eliminating potential bias from the robotic platform and the environmental conditions, an expert operator was trained until three consecutive near-optimal performances were obtained. It took at least 10 trials of each task to reach that level. This process determined the limit of performances that can be obtained by a trained human operator with the system in the operation conditions for the chosen three metrics (in terms of completion time, distance and command quantity). Table V presents average expert operator performances. Performances' stability was validated by having the standard deviation divided by the mean of the last three trials be lower or equal to 5%, for each metric (as shown in parenthesis).

TABLE V
MEAN PERFORMANCES OF THE EXPERT OPERATOR

Setup	Modality	*t* (s)	*d* (m)	*nc*
Kitchen	VC2D	102.17 (5%)	19.64 (1.5%)	6.0 (0.0%)
Kitchen	AR3D	104.57 (2%)	20.12 (1%)	6.0 (0.0%)
Kitchen	ME3D	107.41 (3%)	20.56 (0.5%)	6.0 (0.0%)
Corridor	VC2D	101.62 (2%)	23.41 (3%)	5.0 (0.0%)
Corridor	AR3D	100.79 (2%)	22.68 (2%)	5.3 (1%)
Corridor	ME3D	100.61 (2%)	23.41 (2%)	5.0 (0.0%)

Expert performance cannot be used to extract significant differences between the display modalities, because the expert's skills are not only derived from his ability to use the interface but also from his knowledge about the tasks and the operating conditions. Only trends and insights about the tasks can be observed. The main difference observed is in terms of completion time. In the case of the Kitchen setup (in which the robot pathway is shorter with more curves), VC2D shows the best performance while ME3D shows the worst. For the Corridor setup, the inverse situation is observed, probably because it has a longer path with fewer rotations and that involves less positioning tasks that can be well executed simply using video. The expert operator reached a stable average velocity in the Kitchen setup of 19 cm/sec and 23 cm/sec in the Corridor setup, with the maximum velocity of the robot set at 25 cm/sec. These velocities reflect the inherent complexity of the Kitchen setup compare to the Corridor. The number of rotations done within a shorter distance between near objects seems to reduce the overall operational velocity. The optimal command quantity for the Kitchen setup is 6 and 5 for the Corridor (one trial resulted in one additional command).

These results served to confirm that our teleoperated robotic system can be exploited efficiently, through all of its display modalities, in typical operational scenarios. They also give the best performances that can be reached with the system (considering its limited perceptual, processing and action capabilities) and the environmental conditions.

V. Experimental results for Novice Operators

Participants were selected in order to have a heterogeneous sample of users. Table VI and Table VII present the professional occupations and socio-demographics characteristics of the group of 37 novice operators that participated in our study.

TABLE VI
PROFESSIONS OF THE NOVICE OPERATORS

Professions	Number
Teacher	5
Researcher	2
Administration	7
Technician	3
Programmer/analyst	1
Student	14
Health professional	2
Manual worker	3

TABLE VII
CHARACTERISTICS OF THE NOVICE OPERATORS

Characteristics	Mean	Standard deviation	Minimum	Maximum
Age	34.3	13.3	19	61
Education level (years)	16.1	2.5	13	22
Computer work (hr/week)	23.7	15.1	1	60
Video game (hr/week)	3.1	6.9	0	30
Car driving (hr/week)	5.5	5.6	0	25

A. Novice Performances Referenced to the Expert

Novice performances were referenced to performances of the expert using a Novice Expert Ratio (NER) expressed by (1):

$$NER_x(m) = \frac{PN_x(m)}{\left(PE_{VC}(m) + PE_{AR}(m) + PE_{ME}(m)\right)/3} \quad (1)$$

with PN being the average performance of novice operators for metric m and display x, and PE being the average performance of the expert (taken from Table V). A ratio of one indicates that novice operators show similar performance to that of the expert, and a value greater than one indicates a degradation of performance. This way of looking at operator performances removes bias and gives a more realistic target of human performance [10].

Table VIII present NER (mean μ and standard deviation σ) evaluated for the three metrics in relation to display modalities and tasks. Tasks are ordered in four groups: moving, pushing and positioning task (A); moving and positioning (B and D); moving in between objects (C and E) and moving through a doorway (F). Looking at the performances over all tasks, the novice operators only show similar performances to those of the expert in terms of distance (with means of 1.14, 1.11 and 1.16 for the display modalities), with an overall average of 1.14 (σ = 0,15). This can be explained by the fact that mission paths are tight, not leaving much room for variations. It gives an evaluation of how the novice operators moved the robot (left-right) from the optimal path. For completion time, novice operators are about 2.24 (σ = 1.1) times overall slower than the expert, and they use an average of 5.23 (σ = 4.37) times more commands than the expert (5.50, 5.72 and 4.47). In other words, these results suggest that the expert shows about the same precision as novice operators, is about twice as more efficient and uses 5 times fewer commands. The *nc* metric seems to be the most sensitive to operator experience and is related to the greater cognitive load experienced by novice operators.

TABLE VIII
NOVICE PERFORMANCES REFERENCED TO THE EXPERT

	VC2D		RA3D		ME3D	
Task	μ	σ	μ	σ	μ	σ
Completion time *(t)*						
A	2.37	0.95	2.24	1.12	1.99	0.60
B	2.77	1.66	2.51	1.40	2.64	1.06
D	2.30	1.11	2.05	0.98	2.05	0.75
C	2.01	0.78	1.89	0.86	1.95	0.57
E	1.81	0.94	2.06	1.02	1.63	0.53
F	2.42	1.30	3.07	2.96	2.57	1.33
All tasks	2.28	1.12	2.30	1.39	2.14	0.81
Distance *(d)*						
A	1.13	0.09	1.14	0.24	1.12	0,13
B	1.21	0.22	1.15	0.19	1.28	0.22
D	1.14	0.17	1.06	0.09	1.13	0.10
C	1.10	0.11	1.07	0.07	1.14	0.11
E	1.09	0.10	1.09	0.07	1.09	0.06
F	1.20	0.23	1.14	0.28	1.17	0.19
All tasks	1.14	0.15	1.11	0.16	1.16	0.14
Number of commands *(nc)*						
A	2.36	1.00	2.05	1.12	1.74	0.79
B	10.30	8.05	8.49	7.09	8.16	5.72
D	4.07	2.87	4.06	3.15	3.57	2.22
C	7.81	5.48	7.24	7.74	6.14	4.06
E	4.70	5.68	6.70	7.68	3.68	3.22
F	3.73	2.95	5.76	7.25	3.54	2.51
All tasks	5.50	4.34	5.72	5.67	4.47	3.09

To illustrate trends between display modalities over tasks, Figure 10 illustrates NER in relation to tasks for each display modalities and metrics. Task difficulty can be estimated by summing the NER for a particular task and metric. In terms of completion time, task F is the most difficult, followed by tasks B, A, D, C and E. Mixed performances are observed in relation to display modalities. In terms of distance, the most difficult task is B, followed by F, A, D, C and E. AR3D seems to perform well over all tasks. Finally, in terms of number of commands, task B is the most difficult, followed by C, E, F, D and A. ME3D seems to be performing well. It is however not possible to derive clear conclusions with this data, as the standard deviations are too large. It does however points out that for specific tasks and metrics, it may be possible to find statistically significant results. These are presented in Section V.B. This shows that special attention must be taken when determining tasks so that comparisons of interface modalities provide statistically significant results.

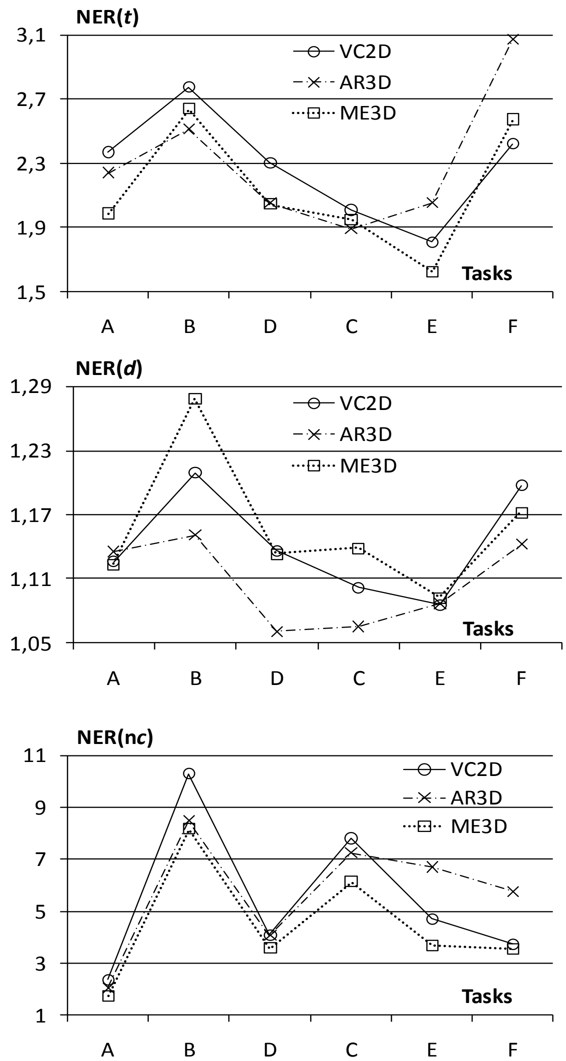

Figure 10. $NER_x(m)$ of novice operators referenced to the expert in relation to tasks, for each display modality.

B. Novice Performances Self-Referenced

For comparison purposes between display modalities, we realized an analysis to identify the best display modality for each metric. To perform that, Novice Ratio (NR) for a specific display modality x, is calculated using as the denominator the mean performance of this novice with the three display modalities..

$$NR_x(m) = \frac{PN_x(m)}{\left(PN_{VC}(m) + PN_{AR}(m) + PN_{ME}(m)\right)/3} \quad (2)$$

Analysis of variance (ANOVAs) revealed significant differences in mean performances across tasks and metrics for six cases, as summarized by Table IX:

- For time completion metric, significant differences in performances between display modalities were observed for task A (F=3.565, dl=35, p=0.039) and task E (F=5.233, dl=35, p=0.01). Contrast analyses between group means showed that ME3D is significantly better than VC2D (t=2.408, dl=36, p=0.21), and ME3D is also significantly better than AR3D (t=3.277, dl=36, p=0.002).
- For the distance metric, significant differences in performances between display modalities were observed for task B (F=6.871, dl=35, p=0.003), task D (F=7.388, dl=35, p=0.003) and task C (F=6.902, dl=35, p=0.002). Contrast analyses showed that AR3D is significantly better than ME3D for task B (t=3.414 dl=36, p=0.002) and C (t=3.888 dl=36, p=0.001) and for task D (t=3.607, dl=36, p=0.001), and AR3D is also better than VC2D for task D (t=2.584, dl=36, p=0.014).
- For the number of commands metric, significant differences in performances between display modalities were observed for task A (F=8.951, dl=35, p=0.001). For this task, the contrast analysis indicates that ME3D is significantly better than VC2D (t=4.1, dl=36, p=0.001).

TABLE IX
STATISTICALLY SIGNIFICANT RESULTS BY METRICS AND TASKS

Metric	Task	Best modality
t	A	ME3D > VC2D
	E	ME3D > AR3D
	B	AR3D > ME3D
d	D	AR3D > VC2D
	C	AR3D > ME3D
nc	A	ME3D > VC2D

Therefore, for completion time and number of commands, the display modality with the best performance is ME3D, followed by AR3D and VC2D. This suggests that ME3D and AR3D improve efficiency and cognitive load (and therefore security) for novice operators, compared to VC2D. For distance, AR3D provides the best performance, followed by VC2D and by ME3D. AR3D may provide more precise teleoperation because of its wider representation of the world (superimposing the video stream on the 3D model, facilitating navigation in tight spaces).

In order to control for individual covariates that could have influenced the observed performances (i.e., gender, age and computing work), multi-factors variance analyses were computed for each metric, task and modality. The following observations were made:

- **Gender**: When controlling for gender (men, women), significant differences in performances between display modalities were observed for the metrics of time (F=3.624, dl=34, p=0.037) and distance (F=3.825, dl=34, p=0.032) in task C. Within-Subject contrast analyses on those results for time (F=6.639, dl=35, p=0.014) and for distance (F=7.663, dl=35, p=0.009) indicate that women perform better with ME3D and worse with VC2D, while men do the opposite and perform better with VC2D and worse with ME3D.
- **Age**: When controlling for age (30 years or less, more than

30 years), significant differences in performances between display modalities were observed for the metric distance for task E (F=4.59, dl=34, p=0.017). Within-Subject contrast analyses show that novice operators 30 years of age or higher are better with the ME3D than with the VC2D (F=8.501, Dl=35, p=0.006), while the opposite is observed for novice operators younger than 30 years of age.
- **Computing work**: When controlling for computer use (22 hours and less, more than 22 hours), significant differences in performances between display modalities were observed for the metrics of time (F=4.561, dl=34, p=0.018) and the commands quantity (F=5.935, dl=34, p=0.006). Within-Subject contrast analyses indicate for task E that novice operators working less than 22 hours per week showed a better performance with ME3D (F=8.923, dl=35, p=0.005) against the VC2D. Those working more did the opposite (F=11.538, dl=35, p=0.002).

Without inferring strong conclusions about gender and age, these results suggest that comparison of gender or age for spatial orientation and localization tasks might be worth investigating with a larger population doing specific tasks.

C. Usability Study

Each participant gave their perception on the usability of the different display modalities by answering a questionnaire just after their trials. Two aspects of usability were explored in this questionnaire: ease of use for each visual modality (What was the system facility of use when teleoperating the robot doing task *T*?), user perception of his performance (When executing the task *T,* how do you qualify your performance?). Ease of use and perceived performance were scored on a nine levels Likert scale, with 1 corresponding to very difficult use or inadequate performance, and 9 corresponding to very easy use or very inadequate performance. Group means on a total possible score of 9 for ease of use were respectively 5.6 (standard deviation $\sigma = 1.9$) for VC2D, 6.1 ($\sigma = 1.9$) for AR3D and of 6.7 ($\sigma = 1.9$) for ME3D. Significant differences in ease of use between display modalities were found (F=10.977, dl35, p=0.001), and ME3D is perceived as the most usable modality while VC2D is the least usable one. For perceived performance, mean scores (mean and standard deviation) for each display modalities are the following: VC2D (5.8, 1.8), AR3D (6.3, 1.9), ME3D (6.8, 1.8). This indicates that operators believe that they were reaching better performances with the ME3D modality, followed with AR3D and at last VC2D.

D. Training of Novice Operators

We also conducted a pilot study with three randomly selected participants to observe the effect of time on the training of novice operators, having them repeat the same scenario two to three weeks afterwards. The following observations were made (by comparing the mean performances according to the three metrics for each trial and expressing this difference in %): the performances of the first operator improved globally by 41%, the second by 40% and the third by 38%. So, after two hours of using the system in the first set of trials, operator's performances improve by about 40% ($\sigma = 1$). We also observed that the operator's performance improved by 11% over difficult tasks, and are more regular across tasks.

VI. Discussion

To summarize our results, novice operators experience higher cognitive load (5 times more when considering the amount of commands needed to performed a specific task) and are less efficient (by 50% when considering the time to complete a specific task) than an expert operator. We did not expect to observe statistically significant differences between the display modalities for all metrics, tasks and their combination: the additional information brought by the 3D display modalities may or may not bring significant influences for the metrics used and the tasks put in place. We believe this to be an interesting observation that can contribute in future work to either enhance the teleoperation tasks or focus on specific metrics, or at the very least consider that evaluation of robotic teleoperation interfaces is a complex problem affected by the operation environment, the mission's objectives, the robots capabilities, the choice of performance metrics, etc. Also, in terms of precision, because of the limited variability in the mission paths, measuring variations in relation to optimal paths may provide more insights (but requires an external tracking system). In general, for the cases with statistically significant differences, the overall performances of novice operators improved when using teleoperation 3D virtual reality interfaces in comparison to conventional video-centric and map-centric display. ME3D seems to be the most effective interface in terms of completion time and command quantity, especially for woman, people older than 30 years old, or people working on computers less than 22 hours per week executing moving tasks. Improved efficiency and ease of use provided by ME3D may come from the use of exo-centric perspective (present in map-centric representation) connected to video-centric data. AR3D seems to be preferable in situations requiring precise navigation, enlarging the video stream viewpoint by superimposing it on a 3D model of the world. However, AR3D is sensitive to the robot's localization accuracy, which influences the alignment of the 3D model with object features in the video stream. In addition, the network delay of 0.4 sec between the time a command is issued and the time it is executed is significant. This causes small disturbances on the display operator without causing significant problem on navigating and positioning the robot, but more visible on AR3D than on ME3D. Therefore, rendering quality probably explains why the perceived performance with AR3D is not as good as with ME3D.

Table X compares our results with related studies on teleoperation user interfaces. Baker et al. [41] provide a qualitative assessment indicating improvement in ease of use with an AR2D interface. This is corroborated by our results. Keyes et al. [43] noted a decrease in cognitive load with an AR2D interface compared to VC2D, which is also confirmed (using *nc*) by our results. Such improvements are even better with the ME3D interface. Ricks et al. [15], Bruemmer et al. [35] and Drury et al. [27] indicate that ME2½D interfaces are more efficient and require less cognitive load than VC2D, as corroborated by our results. Drury et al. [27] also identified conditions in which VC2D interfaces provide better performances. In our case, we observed that men, people younger than 30 years old or people working more than 22 hours per week on a computer perform better with VC2D in navigation tasks, suggesting that for these groups visualizing the local surroundings is preferred (compared to having a global perspective of the environment). Similar tendencies can be observed with the expert operator or in results from our preliminary study conducted with video-centric and map-centric interfaces [10].

TABLE X
RELATED STUDIES ON TELEOPERATION USER INTERFACES

Study	Display	Operators	Population	Metrics
Baker et al. [41]	AR2D	Expert	5	Usability
Keyes et al. [43]	AR2D		19 (9 men, 10 women)	*t*, Number of collisions
Ricks et al. [15]	ME2½D, VC2D	Minimal experience	32	*t* Usability
Bruemmer et al. [35]	ME2½D, VC2D	Novice	64	Usability Cognitive load
Drury et al. [27]	ME2½D, VC2D		8 (7 men, 1 women)	Usability
Labonté et al.	ME3D, AR3D, VC2D	Novice	37 (17 men, 20 women)	*t, d, nc* Usability

VII. CONCLUSION

Two novel display modalities targeting novice operators were designed for a mobile robot teleoperation interface. These modalities make use of 3D representations of the world by either providing an ego-centric (AR) or a mixed exo/ego-centric (ME) perspective. In the ME display, the ego-centric perspective is grounded with the exo-centric perspective by keeping the robot and the video stream fixed in the interface, and by making the 3D model move according to the displacement of the robot in the world. These two display modalities were implemented and tested on a real robot with a population of 37 novice operators having to accomplish six tasks. Looking at four metrics and analyzing performances in relation to an expert operator and to the group, we found interesting conditions in which these new interfaces improve novice operator performances.

Based on these promising results, we are currently developing an improved user interface that allows operators to dynamically change viewpoints, going from ego- to exo-centric, to easily adapt to user's preferences [48]. This interface combines 3D reconstruction of laser range finder readings with video projection methods, making it possible to derive on-line a 3D model of the environment. In addition, we are conducting a usability study involving ten rehabilitation professionals to evaluate how the telepresence system can be used to perform clinical evaluations in a home environment. We also plan to study the influence of control modalities (e.g., physical versus virtual joystick, drive wheel metaphor, go-points, semi-autonomous control, audible feedback) on the cognitive load of novice operators.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contribution of the CIHR – Musculoskeletal Health and Arthritis, Grant 120591, the Canada Research Chair (held by F. Michaud), the Natural Sciences and Engineering Research Council of Canada, the Canadian Foundation for Innovation and the Université de Sherbrooke, in support of this work.

REFERENCES

[1] J. V. Draper, D. B. Kraper, J. M. Usher, "Speculations on the value of presence," *Cyber Psychology and Behaviour*, vol. 2, no. 4, 349-362, 1999.

[2] M. Minsky, "Telepresence," *Omni*, pp. 45-51, 1980.

[3] M. Nokata, N. Tejima, "A safety strategy for rehabilitation robots," *Advances in Rehabilitation Robotics*, Springler-Verlag, LNCIS 306, pp. 177-185, 2004.

[4] J. Casper and R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center," *IEEE Transactions on Systems, Man, and Cybernetics*, part B, Vol. 33, pp. 367-385, 2003.

[5] J. Scholtz, J. Young, H.A. Yanco, J.L. Drury, "Evaluation of human-robot interaction awareness in search and rescue," in *Proceedings of IEEE Conference of Robotics and Automation,* vol 3, pp. 2327 – 2332, Apr 26-May 1, 2004.

[6] H. A. Yanco, J. Drury. ""Where Am I?" Acquiring situation awareness using a remote robot platform," in *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, October 2004.

[7] Health Canada, *Better Health: Strategic Issues*, Interim Report, 1998.

[8] S. M. Finkelstein, S. M. Speedie, S. Potthoff, "Home telehealth improves clinical outcomes at lower cost for home healthcare," *Telemedicine and E-Health Journal*, vol. 12, no. 2, pp. 128-136, 2006.

[9] B. Johnston, L. Wheeler, J. Deuser, K. H. Sousa, "Outcomes of the Kaiser Permanente tele-home health research project," *Archives of Family Medicine*, vol. 9, no. 1, pp. 40-45, 2000.

[10] D. Labonté, F. Michaud, P. Boissy, H. Corriveau, R. Cloutier, M.A. Roux, "A pilot study on teleoperated mobile robots in home environments," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* Beijing, China, pp. 4466–4471, 2006.

[11] P. Boissy, H. Corriveau, F. Michaud, D. Labonté, M.P. Royer, "Exploring the potential use of home mobile telepresence for telehomecare: A qualitative study with healthcare professionals and community living elderly," *Journal of Telemedicine and Telecare*, vol. 13, no. 2, pp. 78-84, 2007.

[12] F. Michaud, P. Boissy, D. Labonté, H. Corriveau, A. Grant, M. Lauria, R. Cloutier, M.-A. Roux, M.-P. Royer, D. Iannuzzi, "Telepresence robot for home care assistance," *Technology and Aging*, Assistive Technology Research Series, 2008.

[13] J. Raskin, *The Human Interface*. Addison Wesley, ACM press, 2000.

[14] T. Fong, I. Nourbakhsh, K. Dautenhahn, "A survey of socially interactive robots," *Robotic and Autonomous Systems*, vol. 42, pp. 143-166, 2003.

[15] B. W. Ricks, C. W. Nielsen, M. A. Goodrich, "Ecological displays for robot interaction: A new perspective," in *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, 2004.

[16] J. L. Drury, B. Keyes, H. A. Yanco, "LASSOing HRI: Analyzing situation awareness in map-centric and video-centric interfaces," in *Proceedings of International Conference on Human Robot Interaction*, pp. 279-286, 2007.

[17] M. A. Goodrich, "Using models of cognition in HRI evaluation and design," In *Proceedings of the AAAI 2004 Fall Symposium Series: The Intersection of Cognitive Science and Robotics: From Interfaces to Intelligence*, Arlington, Virginia, October 21-24, 2004.

[18] C. W. Nielsen, M. A. Goodrich, R. J. Rupper, "Towards facilitating the use of a pan-tilt camera on a mobile robot," In *Proceedings of IEEE International Workshop on Robots and Human Interactive Communication*, pp. 568-573, 2005.

[19] T. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*, MIT Press, Cambridge, 1992.

[20] J. Rasmussen, “Skills, rules, knowledge, signals, signs, and symbols, and other distinctions in human performance models,” in *IEEE Transactions on System, Man and Cybernetics*, SMC-13 (3), pp. 257-266, 1983.
[21] H. A. Yanco, J. Drury, “Classifying human-robot interaction: An updated taxonomy,” in *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, October 2004.
[22] M. R. Endsley, “Design and evaluation for situation awareness enhancement,” in *Proceedings of the Human Factors Society 32th Annual Meeting*, Santa Monica, CA, Vol. 1, pp. 97-101, 1988.
[23] J. C. Scholtz, B. Antonishek, J. D. Young, “Implementation of a situation awareness assessment tool for evaluation of human-robot interfaces,” in *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol. 35, No. 4, pp. 450-459, July 2005.
[24] C. D. Wickens, J. G. Hollands. *Engineering Psychology and Human Performance,* Third Edition, Prentice Hall, Upper Saddle River NJ, 2000.
[25] A. Steinfeld, “Interface lessons for fully and semi-autonomous mobile robots,” in *Proceedings of International Conference on Robotics and Automation*, New Orleans, LA, pp. 2752-2756, April 2004.
[26] T. Fong, C. Thorpe, “Vehicle teleoperation interfaces,” *Autonomous Robots*, vol. 11, pp. 9-18, 2001.
[27] J. L. Drury, J. Scholtz, H. Yanco, “Awareness in human robot interactions,” in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, Washington D.C., pp. 568-573, October 2003.
[28] D. McGovern, “Experiences and results in teleoperation of land vehicles,” Sandia National Laboratories, Albuquerque, NM, Technical Report SAND 90-0299, 1990.
[29] D. D. Woods, J. Tittle, M. Feil, A. Roesler, “Envisioning human-robot coordination in future operations,” in *IEEE Transactions on Systems, Man and Cybernetics*, Part A, Vol. 34, No. 6, pp. 749-756, November 2006.
[30] M. Voshell, D. D. Woods, F. Willips. “Overcoming the keyhole in human-robot coordination: simulation and evaluation,”, in *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Orlando, September 2005.
[31] S. Hughes, J. Manojlovich, J. Gennari, “Camera control and decoupled motion for teleoperation,” in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, Vol. 2, pp. 1339-1344, October 2003.
[32] G. Thomas, W. D. Robinson, S. Dow, “Improving the visual experience for mobile robotics,” in *Seventh Annual Iowa Space Grant Proceedings*, Drake University, pp. 10-20, November 1997.
[33] C. A. Johnson, J. A. Adams, K. Kawamura, “Evaluation of an enhanced human-robot interface,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Washington, DC, pp. 900-905, October 5-8, 2003.
[34] M. A. Goodrich, R. J. Rupper, C. W. Nielsen, “ Perceiving head, shoulders, eyes and toes in augmented virtuality interfaces for mobile robots,” In *Proceedings of IEEE International Workshop on Robots and Human Interactive Communication*, pp. 574- 579, 2005.
[35] D. J. Bruemmer, A. F. Douglas, M. C. Walton, R. L. Boring, J. L. Marble, C. W. Nielsen, J. Garmer, “Turn off the television!: Real-world robotic exploration experiments with a virtual 3-D display,” in *Proceedings of the 38th Hawaii Conference on System Sciences*, 2005.
[36] W. C. T. Wang, “The design of multimodal human-machine interface for teleoperation,” in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3187-3192, October 2001.
[37] P. Milgram, D. Drascic, J. J. Grodski, A. Restogi, S. Zhai, C. Zhou, “Merging real and virtual worlds,” in *Proceedings of IMAGINA’95*, pp. 218-230, February 1995.
[38] H. K. Keskinpala, J. A. Adams , “Objective data analysis for PDA-based human-robot interaction,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics,* 2004, The Hague, The Netherlands.
[39] C. W. Nielsen, B. Ricks, M. A. Goodrich, D. Bruemmer, D. Few, M. Walton, “Snapshots for semantic maps,” in *Proceedings IEEE International Conference on Systems, Man and Cybernetics*, The Netherlands, October 2004.
[40] M. Lewis, J. Wang, “Gravity-referenced attitude display for mobile robots: Making sense of what we see,” in *IEEE Transaction on Systems, Man and Cybernetics*, part A, vol. 37, No. 1, pp. 94-105, January 2007.
[41] D. B. Baker, R. Zhou, D. Song, “Design and prototyping of an economical teleoperations test-bed for human factors research: Cost, resource requirements and capability assessment,” in *Proceedings of the 25th International Conference on Computers and Industrial Engineering*, pp. 526-529, New Orleans, 1999.
[42] R. Ott, M. Gutiérrez, D. Thalmann, F. Vexo, “VR haptic interfaces for teleoperation : an evaluation study,” pp. 789-794, 2005.
[43] B. Keyes, R. Casey, H.A. Yanco, B.A. Maxwell, Y. Georgiev, “Camera placement and multi-camera fusion for remote robot operation,” in *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics*, National Institute of Standards and Technology (NIST), Gaithersburg, MD, 2006.
[44] A. Montferrer, D. Bonyuet, “Cooperative robot teleoperation through virtual reality interfaces,” in *Proceedings of Sixth International Conference on Information Visualization*, pp. 243-248, 2002.
[45] M. Montemerlo, N. Roy, S. Thrun, “Perspectives on standardization in mobile robot programming: The Carnegie Mellon Navigation (CARMEN) toolkit,” in *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2436-2441, 2003.
[46] J. Cooper, M.A., Goodrich, “Towards combining UAV and sensor operator roles in UAV-Enable Visual Search,” in *Proceedings ACM/IEEE International Conference on Human-Robot Interaction*, pp. 351-358, 2008.
[47] C.W., Nielsen, M.A., Goodrich, R.W., Ricks, “Ecological interfaces for improving mobile robot teleoperation,” in *IEEE Transactions on Robotics*, vol. 23, no. 5, October 2007.
[48] F. Ferland, F. Pomerleau, F. Michaud, “Egocentric and exocentric teleoperation interface using real-time, 3D video projection,” in *Proceedings ACM/IEEE International Conference on Human-Robot Interaction*, pp. 37-44, 2009.

Daniel Labonté (M’94) received a bachelor degree in electrical engineering from Polytechnic school of Montréal, Québec, Canada in 1973 and received a Ph.D. degree in 2008 from the Université de Sherbrooke, Québec, Canada. His research activities are on telepresence, immersive system, embedded system and user interfaces for robot teleoperation.

He is a professional consultant in embedded systems and software development and before joining the LABORIUS group, he was working for the Materials and Intelligent Systems Institute IMSI of the Université de Sherbrooke on a digital imagery center project. He also has 27 years of experience as an engineer in software and electronic product development. He is a member of IEEE and OIQ (Ordre des ingénieurs du Québec).

Patrick Boissy received his bachelor’s degree (’91) and Master’s degree (‘93) in kinesiology from Université de Sherbrooke, Québec , Canada. He graduated from the Universite de Montreal, Québec, Canada in 1999 with a Ph.D. degree (’99) in biomedical sciences (rehabilitation) After postdoctoral training at Boston University’s Neuromuscular Research Centre, he joined the department of kinesiology, at Université de Sherbrooke in 2002 where he is currently an associate professor. He holds appointments as a researcher at the Research Centre on Aging of the Health and Social Service Centre, University Institute of Geriatrics of Sherbrooke, and at the Center of Excellence in Information Engineering of the Université de Sherbrooke. His research interests include technological and clinical evaluation of telehealth applications and the study of the dose-response relationship in geriatric rehabilitation. He is currently funded as a Research Scholar Junior II by the Fonds de la recherche en santé du Quebec.

François Michaud (M’90) received his bachelor’s degree (’92), Master’s degree (‘93) and Ph.D. degree (‘96) in electrical engineering from the Université de Sherbrooke, Québec Canada.

After completing postdoctoral work at Brandeis University, Waltham MA (’97), he became a faculty member in the Department of Electrical Engineering and Computer Engineering of the Université de Sherbrooke, and founded LABORIUS, a research laboratory working on designing intelligent autonomous systems that can assist humans in living environments. His research interests are in architectural methodologies for intelligent decision-making, design of autonomous mobile robotics, social robotics, robot for children with autism, robot learning and intelligent systems.

Prof. Michaud is the Canada Research Chairholder in Mobile Robots and Autonomous Intelligent Systems, and the Director of the Centre of Excellence on Information Engineering. He is a member of IEEE, AAAI and OIQ (Ordre

des ingénieurs du Québec). In 2003, he received the Young Engineer Achievement Award from the Canadian Council of Professional Engineers.