A context-aware immersive interface for teleoperation of mobile robots

João Quintas
Instituto Pedro Nunes
Coimbra
jquintas@ipn.pt

Luís Almeida
Instituto Politécnico de Tomar
Tomar
laa@ipt.pt

Elísio Sousa Paulo Menezes
Universidade de Coimbra
Coimbra
pm@deec.uc.pt

Abstract

In this paper we present a context-aware immersive teleoperation interface to assist operators during navigation tasks. This new interface strategy aims to address the problems associated with mental overload, often experienced by operators of teleoperated devices. Our approach simplifies the high complexity of information displayed in control rooms. Our approach includes a context-based human-robot interaction framework that detects relevant information and automatically adapts the displayed interface in virtual windshield. Results showed that the proposed approach enhances user immersion while maximizes task performances and minimizes the operator physical and cognitive workload.

Keywords

Human-Robot Interaction, telepresence, Embodiment, Teleoperation, Context-Awareness

---

1 Introduction

A telepresence robot [Minsky 80] presents a solution for search and rescue, remote reconnaissance, space exploration or maintenance in contaminated areas. In these scenarios, operators often intervene in the robot control loop when the robot is deployed in remote and unstructured environments [Sheridan 93].

Figure 1. A typical ROV Control Room. Courtesy of Monterey Bay Aquarium Research Institute.

An example of a typical control room for Remote Operated Vehicles (ROV) is depicted in figure 1. In spite of operator's skills and expertise, human decisions in teleoperation rely on diverse remote information sources. While teleoperating, operators must be fully focused in their task, which requires processing all inputs and filtering relevant information in order to execute the appropriate action. This intense use of perceptual and cognitive skills may lead to mental and physical strain, which may cause catastrophic hazards. This fact was addressed in [Wickens 08], where the authors studied how humans capabilities vary while performing tasks that require processing information from multiple resources. The studies concluded that multiple sources of information contribute to a high mental workload, causing negative implications on task performance.

In order to reduce the difficulties and stress of teleoperation, several authors propose solutions that allow the user to have a better understanding of the remote environment without the need to keep a mental record of the same.

The sensation of being (inside) the robot improves operator's performance of driving it. Thus, by combining the concepts of telepresence and physical embodiment we are able to create tele-embodiment [Paulos 97]. As result, the operator feels the remote robot body as his own and he/she acts more naturally, minimizing the physical and cognitive workload.

In the other hand, studies showed that operators quite often are not sufficient aware of robot location and surroundings, resulting in most operator decisions are based on remote video information, which forces the operator to try understanding the remote environment through a ”keyhole” [Woods 04].

In [García 15], a virtual cockpit was proposed for intervention underwater robots that simplifies the high complexity of information displayed through specifically designed Graphical User Interface (GUI).

To tackle this challenge, we propose an approach to create a context-aware immersive interface for teleoperation of mobile robots that extends typical teleoperation functionalities, allowing human operators to benefit from an improved user experience. Figure 2 illustrates the expected outcome for this approach, which aim to improve over typical control rooms as depicted in figure 1.

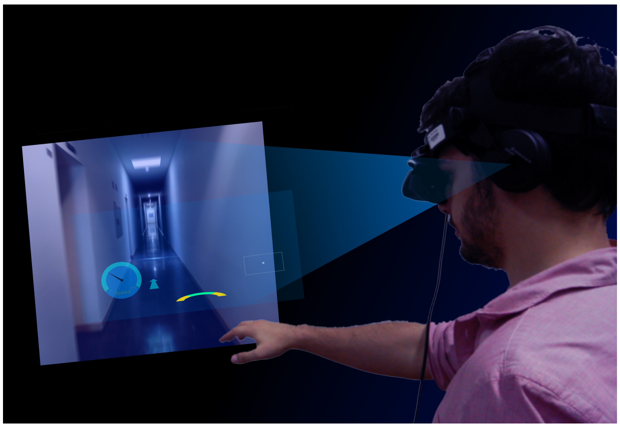


Figure 2. Context-aware immersive interface for teleoperation of mobile robots

Consider a teleoperation scenario where an operator is using our immersive interface. The operator is performing a navigation task, of a mobile robot, in remote environment with good weather conditions (e.g. partially sunny and low humidity). For this task, the operator could be more interested in paying attention in bearing and speed of the robot. The immersive interface would display a simplified control panel with widgets relevant only for that given context (i.e. navigating the robot with good weather). Suppose now, that during teleoperation, the weather changed and the robot's environment is rainy with wind gusts. Our context-aware immersive teleoperation interface will now adapt and display widgets related with wind direction and indication of applied torque in the wheels, as a muddy floor requires a more skillful driving to avoid getting stuck.

1.1 Context-awareness

Schmidt in [Schmidt 00] regarded situational context, such as location, surrounding environment or state of the device, as implicit input to the system. This extended the concept of context beyond the informational context into real world environments. The authors used this concept to define the term "implicit human-interaction" as *"... an action performed by the user that is not primarily aimed to interact with a computerised system but which such a system understands as input ..."*

Röning and Riekki in [Röning 01] proposed a context-aware mobile system, which included mobile personal robots. They proposed the "Genie of the Net" architecture as an ever expanding system providing helpful information and guidance when human capabilities are exceeded. Their proposed approach also aimed to be a technique to handle several individual robots so that they can co-operate with each other and human beings. Their initial application tests selected the approach of first building a teleoperated robot and then gradually shifting tasks from the human to the robot.

Celikkanat, et.al. in [Celikkanat 15] demonstrated on the iCub platform that using context resulted in an adaptive, online and robust approach for executing two important tasks: object recognition and planning.

In our approach we will define context as:

> *Context is the set of information that is relevant, affects or constrains how some action is taken, without being the center of interest of the action.*

1.2 Contributions and structure

Based on these principles, we present an approach that aims to improve the telepresence experience for the operator when remotely operating a mobile robots. The Augmented Reality based user interface (UI) proposed in [García 15] is now coupled with a context-aware module and will automatically adapt operator's UI to changing conditions that are relevant for the task being performed, resulting in a context-aware immersive interface. This auto-adaptation consists in providing cues to the operator that aim to simplify the teleoperation interface.

The paper proceeds as follows. Section 2 presents design aspects, including teleoperation mechanism and the role of context information in the interaction process of telepresence and teleoperation. Section 3 describes experimental and comparative results of different interface styles. As an application example, we address the scenario of an operator remotely controlling a robot while his context aware user interface adapts to help him during the navigation task, providing the necessary information for the given context while hiding the irrelevant one to avoid distracting or overloading the user. Section 4 summarize the conclusions.

2 Designing the context-aware immersive interface for teleoperation of mobile robots

In this section we will address some design aspects for our approach referring to teleoperation architectural details and necessary adaptations to achieve a context-aware immersive system.

2.1 Teleoperation architecture

Literature proposes teleoperation models with the human operator inside the control loop. Usually, the robot control commands are transmitted through a delayed transmission channel [Islam 14][Sheridan 93][Almeida 14] and, the action feedback is also affected by a transmission delay. The model purpose is to integrate these delays and keep the robot controllable.

In our research, we explore the relationship between the human and the interface used to control the remote robot. We propose a simplified model composed by an outer teleoperation control loop that uses an inner perception control loop, see figure 3.

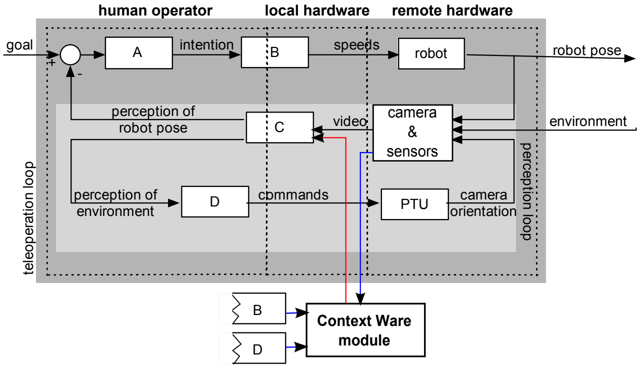


Figure 3. Teleoperation and perception as control loops

teleoperation loop – the robot teleoperation process can be modeled as a standard control loop. Basically, the human operator compares a given goal with the robot's position in the remote environment. The operator perceives the difference and develops an intention to compensate it, which is later translated into robot's commands by some interaction system. Figure 3 depicts this loop, where block A represents the perception of the error and the intention generation. This intention is converted into commands through block B, which models the human action into an interface that produces proper robot commands. This control loop will be closed only if the user can perceive the pose of the robot in the remote place.

Perception Loop – this research in teleoperation systems considers a camera point of view, as the operator being inside the robot. The camera's purpose is to enable user to perceive both the robot motion and the surrounding environment as being driving inside the mobile robot. This perception process can similarly be described as a control loop. In this case, the human operator controls the robot's camera orientation and utilizes the visual feedback to compensate the scanning process required for a task (ex: track objects, look around, inspect, or navigate). As in the control loop, the camera acquires images and sends them through a channel to the user. This visual information enables the operator to perceive the relative pose of the robot in the remote environment and, the environment itself. Block C represents this process.

The Context Aware module recognizes an activity based on robot's sensors information, operator's positional intention of the robot (block B) and operators visual point of view (block D) and, with it selects the useful information to be presented in the operator's windshield or UI (block C).

2.2 From teleoperation to remote embodied operation

Using the presented model lets map the different perception and control mechanisms into blocks A, B, C and D. We demonstrate how to evolve from a traditional teleoperation concept to new and more immersive approaches. In traditional teleoperation systems, block B represents the robot motion control using a joystick and, block D, represent the control of the pan-and-tilt camera unit using also a joystick. Block C provides the remote images to the user through standard screen, while block A, enables him to convert the positional perceived error into an intention to move the robot.

To create a more immersive interface we propose a viewpoint transfer. To solve the challenge of controlling the remote viewing camera, we suggest the use of a head mounted display (HMD) in which the operator can move his head and almost simultaneously control a pan-and-tilt unit (PTU) that supports the robot's camera. Block C provides the visual information that enables the user to perceive the difference between the visual goal, and the means to compensate. The human, through block D, acts into camera PTU to gather new point of views.

This type of camera control provides an egocentric view, as the camera movements are synchronous with the operator's head movements. It enables the user to have an egocentric perception of the remote environment just as if the human was at the robot position and orientation. The described process is a crucial step to give the user the sensation of being physically embodied in the remote robot, which means that "the user will see what the robot can see".

2.3 Creating a context-aware immersive interface for teleoperating mobile robots

In our approach we consider context recognition to be a periodic process that operates in the background of the system, while interacting with a user. This process is illustrated in figure 4.

It plays the role of detecting changes in the context and control the adaptation in the user interface during teleoperation.

To incorporate context-awareness into our architecture we designed a Context-based Human-Robot Interaction Framework (CB-HRI). Figure 5 illustrates the CB-HRI framework, which conceptually extends teleoperation architecture.

This framework acts as a middleware to integrate contextual information in the overall system and control the workflow related with human-robot interaction.

The main components of the framework are:

1. the *Message Bus* module that includes the interfaces with other components in the architecture.
2. the *Recognition algorithms* module includes the algorithms to match perceived information with contextual information.

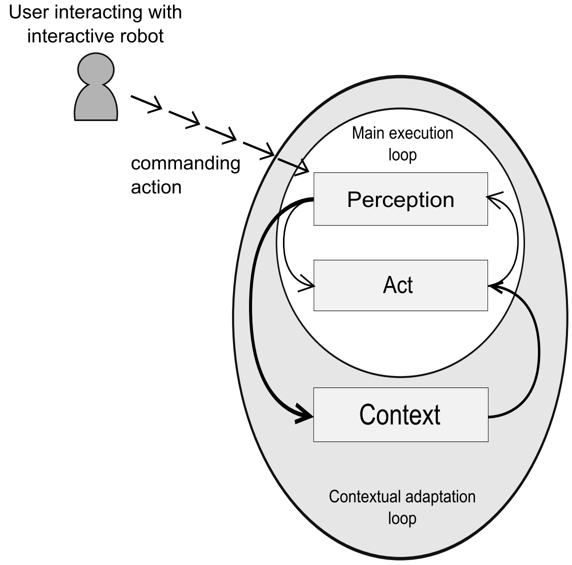


Figure 4. Context verification process

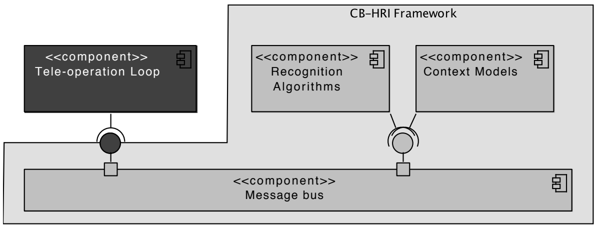


Figure 5. Context-based Human-Robot Interaction Framework extending existing architectures

3. the *Context models* is a repository with apriori context data models.

In our approach we propose contextual information as an integration mechanism between a variety of available algorithms and other resources that are known to perform well under a certain conditions.

Therefore, the CB-HRI framework must be integrated with the components dealing with perception, reasoning, data storage and actuation.

3 Implementation and Results

In our experiment the objective was to navigate as quickly as possible, without colliding against walls or obstacles, as illustrated in figure 6. The operator could make the robot move forward, move backward, turn 360° on itself, and control the robot's camera point of view. The robot on-board sensors could provide the following information: movement speed, movement direction, 360° proximity information, camera's pose and battery levels.

Choosing only the most relevant information at any given time, provides uncluttered field of view and decreases the user's mental workload. Furthermore, for the information that is always present, this can be slightly transparent as not to block the user's view. The graphical elements representing the information should not be too big and placed near the user's view centre, as not to strain the user's eye and focus.

Figure 6. Navigation task, comparison of different teleoperation interaction styles designed to enhance embodiments sensations

3.1 Modeling and recognizing Contexts

In our approach we represent Context as a vector, where each element is a numerical representation of a feature (i.e. context features).

In order to select coherent context features we take as baseline the same measured information in previous works, where we explored immersive teleoperation interaction styles and their application in a virtual cockpit in a navigation task. Thus, we enumerate the context features used as:

- Task (e.g. reserved for future use)
- User proficiency (e.g. 0 = beginner, 1 = amateur, 2 = professional)
- Distance to obstacle (e.g. 0 = close, 1 = near, 2 = far)
- Safe speed limit (e.g. 0 = slow, 1 = fast)
- Bearing to obstacle (e.g. 0 = no adjustment, 1 = adjust right, 2 = adjust left)

In this set of features we can neglect Task, as this information will be irrelevant because we are only performing navigation.

We create a set of rules based on the previous features to define Context classes, as follow:

- Context0: Navigating in open space
- Context1: Narrow space / Close proximity to obstacle

- Context2: Too much speed to avoid obstacle without colliding
- Context3: Wrong bearing to overcome the obstacle

The result of the classification is then used by the user interface, which loads the appropriate widgets that provide relevant information to the user while navigating in a specific environment. Specifically, taking into account our experiment, we have the following adaptations:

- No widgets are displayed in Context0;
- Display the speedometer widget in Context1;
- Display the bearing indicator widget in Context2.

3.2 Implementing the immersive interface

To implement the graphical interface, a combination of OpenGL and OpenCV was used. Setting the video stream as background with OpenCV, the 2D elements are created with OpenGL and then placed on top of the stream. The robot movement speed was designed according to modern speedometers to offer a degree of familiarity to the user, since it is one of the most common and intuitive ways to display an object's speed. As such, a speedometer background containing a gauge was designed and a needle (with a transparent background) was created and placed on top of the background, with the lower end of the needle aligned with its center. By applying rotation to the needle, we can make it move and indicate the robot speed. The robot's direction is simply an arrow indicating forwards or backwards. It is designed to look like it's pointing along the Z-axis (depth) for better intuitiveness (see figure 7). Proximity information is represented as a circle with as many sections as there are sensors. Each section changes color depending on the distance from the robot to the nearest object in the corresponding direction. The circle is designed to look like it's aligned with the Z-axis to facilitate the user's perception of which sensors are displaying information. To create the sections, various circles with single sections are stacked upon each other and controlled independently. The camera pose is represented by a circle inside a square. Pan is represented through the X-axis and tilt is represented through the Y-axis. The square represents the minimum and maximum pan and tilt limits. The battery level is represented by a numerical percentage inside a drawing of a common battery.

Other widgets, like battery level will be displayed if context awareness module considers that is important for the task.

3.3 Validating user interaction

To validate our approach we compared experimentally four interaction styles, which included from traditional joystick approaches to more innovative based on deictic gestures and natural body postures. We carried out a quantitative and subjective task performance analysis involving 13 participants. All the participants had to teleoperate a mobile robot and navigate through a predefined obstacle course.

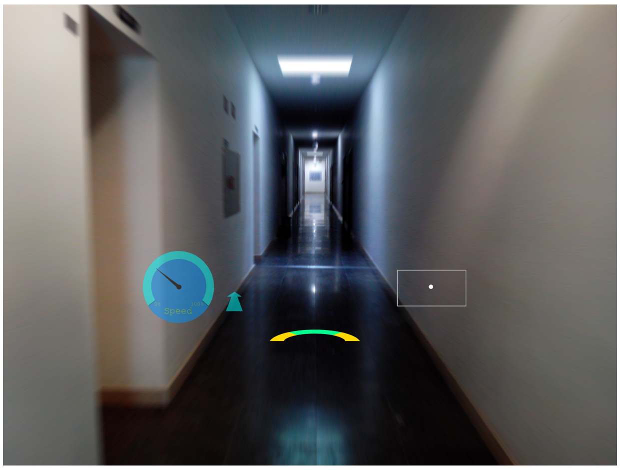


Figure 7. Operator's immersive windshield with smart widgets: semi transparent speedometer, robot direction motion arrow, proximity sensor and camera pose.

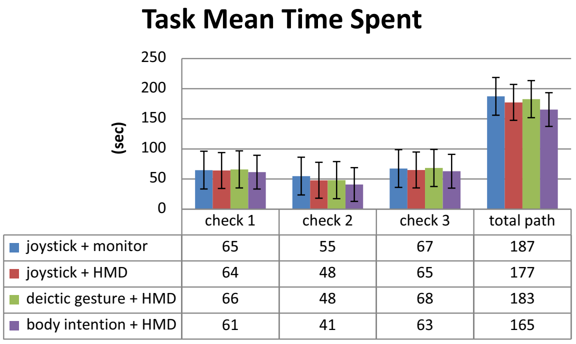
The experiment goals were to maximize the task performance and minimize the operator's physical and cognitive workload. We induced in the operator the sensation of being at the remote environment; to generate the remote physical embodiment feeling, the approach consisted in letting user perceive the robot's structure as his/her own body.

To evolve from teleoperation to embodied operation we explored 3 approaches:

1) view transfer using an HMD (i.e. with an egocentric controlled view, the user will see what robot can see);
2) pointing gestures to control the robot (i.e. user sees him as being the robot, or inside of it, and his pointing gestures are used to control his own motions);
3) body posture to control the robot.

To understand the influence of the four different interaction styles on the teleoperation of a mobile robot; and to assess how natural can a user interact and perceive the remote robot structure as his own body, a quantitative and subjective task performance analysis were carry out (13 participants in driving tasks)(figure 6).

Results demonstrated that visual feedback through an HMD improved significantly users task performance (figure 8). The introduction of natural deictic gestures based robot control presented some gain in task performance when compared with joystick. Body intention-based robot control was the operator's choice in all subjective questionnaires, and was confirmed by time performance measures in path driving. As conclusions of the introduced gesture, postures and view control mechanisms improves the physical embodiment sensation. Sensation of controlling the robot from inside reduces mental workload of the operator. There is a positive effect on user satisfaction and task performance.



	check 1	check 2	check 3	total path
🟦 joystick + monitor	65	55	67	187
🟥 joystick + HMD	64	48	65	177
🟩 deictic gesture + HMD	66	48	68	183
🟪 body intention + HMD	61	41	63	165

Figure 8. Navigation task performance time comparison while using different teleoperation interaction styles

4 Conclusion and future work

The paper addressed the challenges that enhances telepresence and teleoperation. We considered the importance of integrating contextual information in the interaction process with the objective to improve user experience while performing a teleoperated task.

To understand the influence of 4 different interaction styles on the teleoperation of a mobile robot; and to assess how natural can a user interact and perceive the remote robot structure as his own body, a quantitative and subjective task performance analysis were carry out demonstrating that visual feedback through an HMD improved significantly users task performance.

Moreover, we addressed the navigation task as an example task where the integration of context information can improve usability of the system by providing pro-active cues that help the operator to perform the task. An rational management of the information and a egocentric point of view maximizes task performances and minimizes the operator physical and cognitive workload.

References

[Almeida 14] L. Almeida, B. Patrao, P. Menezes, and J. Dias. Be the robot: Human embodiment in tele-operation driving tasks. In RO-MAN, 2012 IEEE, Aug 2014.

[Celikkanat 15] H. Celikkanat, G. Orhan, N. Pugeault, F. Guerin, E. Sahin, and S. Kalkan. Learning Context on a Humanoid Robot using Incremental Latent Dirichlet Allocation. IEEE Transactions on Autonomous Mental Development, 2015.

[García 15] Juan Carlos García, Bruno Patrão, Javier Perez, Joaão Seabra, Paulo Menezes, Jorge Dias, and Pedro J. Sanz. Towards an immersive and natural gesture controlled interface for intervention underwater robots. In Oceans'15 MTS/IEEE Genova, Genova (Italy), 05/2015 2015.

[Islam 14] Shafiqul Islam, XiaopingP. Liu, AbdulmotalebEl Saddik, Lakmal Seneviratne, and Jorge Dias. Control schemes for passive teleoperation systems over wide area communication networks with time varying delay. International Journal of Automation and Computing, 11(1):100–108, 2014.

[Minsky 80] Marvin Minsky. Telepresence. Omni, pages 45–51, 1980.

[Paulos 97] Eric Paulos and John Canny. Ubiquitous tele-embodiment: Applications and implications. International Journal of Human-Computer Studies, 46:861–877, 1997.

[Röning 01] Juha Röning and Jukka Riekki. Context-Aware Mobile Robots: Part of Smart Environment. In HRI-workshop at KTH, 2001.

[Schmidt 00] Albrecht Schmidt. Implicit human computer interaction through context. Personal Technologies, 4(2-3):191–199, June 2000.

[Sheridan 93] T.B. Sheridan. Space teleoperation through time delay: review and prognosis. Robotics and Automation, IEEE Transactions on, 9(5):592–606, Oct 1993.

[Wickens 08] Christopher D. Wickens. Multiple Resources and Mental Workload. Human Factors: The Journal of the Human Factors and Ergonomics Society, 50(3):449–455, June 2008.

[Woods 04] D.D. Woods, J. Tittle, M. Feil, and A. Roesler. Envisioning human-robot coordination in future operations. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 34(2):210–218, May 2004.