

■ Visual Representation of Interaction Force and | Sound Source in a Teleoperation User Interface | for a Mobile Robot

Aurélien Reveleau, François Ferland, Mathieu Labbé, Dominic Létourneau, François Michaud

Department of Electrical Engineering and Computer Engineering, Université de Sherbrooke, Québec, Canada

Commercial telepresence robots provide video, audio, and proximity data to remote operators through a teleoperation user interface running on standard computing devices. As new modalities such as force sensing and sound localization are being developed and tested on advanced robotic platforms, ways to integrate such information on a teleoperation interface are required. This paper demonstrates the use of visual representations of forces and sound localization in a 3D teleoperation interface. Forces are represented using colors, size, bar graphs and arrows, while speech or ring bubbles are used to represent sound positions and types. Validation of these modalities is done with 31 participants using IRL-1/TR, a humanoid platform equipped with differential elastic actuators to provide compliance and force control of its arms and capable of sound source localization. Results suggest that visual representations of interaction force and sound source can provide appropriately useful information to remote operators.

Keywords: Telepresence robotics, visual representation, sound source, interaction force

1. Introduction

An increasing number of commercial telepresence robotic platforms are available (Guizzo, 2010) (e.g., VGo from VGo Communications, QB from Anybot (Desai, Tsui, Yanco, & Uhlik, 2011), TiLR from RoboDynamics, Giraff from HeadThere (Gonzalez-Jimenez, Galindo, & Ruiz-Sarmiento, 2012), beam+ from Sutable Technologies, TeleMe from Mantarobot, Double from Double Robotics, RP7i and RP Vita from InTouch Health). These platforms usually consist of a mobile base, a camera, proximity sensors, a screen, and microphones, making them mobile videoconference systems, commonly referred to as “Skype on wheels” (Dahl & Boulos, 2013). Their intended usage goes from teleworkers virtually present at meetings and remote sites (Lee & Takayama, 2011) to telehealth applications in hospitals and homes (Mendez, Jong, Keays-White, & Turner, 2013; Michaud et al., 2010).

The graphical user interface (GUI) of telepresence robots plays a central role in maxi-

Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work’s authorship and initial publication in this journal.

mizing situation awareness (Endsley, 1988; Scholtz, Young, Yanco, & Drury, 2004) while minimizing cognitive load (Yanco, Drury, & Scholtz, 2004) for operators using these systems. As perceptual and localization modalities were added to teleoperation interfaces over the years, research demonstrated the benefits of minimizing the number of windows by directing information on the main display (Keyes, Micire, Drury, & Yanco, 2010). Accordingly, the ecological interface paradigm uses augmented reality displays to represent the telepresence robot in virtual environments and by integrating video stream directly in these virtual environments (Labonté, Boissy, & Michaud, 2010; Ferland, Pomerleau, Dinh, & Michaud, 2009; Nielsen, Goodrich, & Ricks, 2007; Ricks, Nielsen, & Goodrich, 2004) to facilitate navigation of the platform.

Recent robots bring additional interaction capabilities by being equipped with compliant manipulators, for example, Rollin' Justin (Borst et al., 2009), PR2 (Wyrobek, Berger, der Loos Van, & Salisbury, 2008), Cody (Chen & Kemp, 2011) and Care-O-bot (Graf, Parlitz, & Hagele, 2009). Providing mobile platforms with sound source localization, tracking, and separation capabilities is now possible with systems such as HARK (HRI-JP Audition for Robots) (Nakadai et al., 2010; Mizumoto et al., 2011) and ManyEars (Grondin, Létourneau, Ferland, Rousseau, & Michaud, 2013). Using such capabilities on a telepresence robot requires communicating information from these sensing capabilities to the remote operator. Assuming that teleoperation happens only through a standard computer without the use of specific peripherals, such as haptic devices (Bar-Cohen, 2003), the objective of this paper is to study the use of interaction force and sound source visual representation in an ecological teleoperation interface. More specifically, we have tested the use of colors, size, bar graphs, and arrows for the visualization of forces, and the use of a speech bubble or a ring to position and identify sound types. Use of IRL-1/TR (Ferland et al., 2012), a compliant humanoid platform capable of force interaction and sound source localization, validated these modalities.

The paper is organized as follows: Section 2 identifies related work on visual representation of forces and sounds; Section 3 presents the IRL-1/TR robot system used; Section 4 describes visual representation modalities for force and sound implementation; Section 5 presents the experimental methodology; and Section 6 describes the results.

2. Related Work

The sense of touch appears to be as important as vision and audition in real and virtual environment perception (Robles-De-La-Torre, 2006). Remote touch sensing could be accomplished through a haptic interface (Hu et al., 2005; Pamungkas & Ward, 2013), such as the use of either joysticks with force feedback or data gloves. Providing visual cues related to touch is also possible. For instance, Lindeman (2003) experimented with various visual haptic feedback in immersive virtual environments, such as illuminating the virtual contact point with a binary color change, having the contact location and intensity represented by a change of color, texture, or shape distortion, or rendering force intensity and direction by vector glyphs. Haptic substitution by visual cues has also been studied with teleoperated surgical robots. Contact forces are represented using a bar indicator proportional to force intensity, overlaid on top of the video stream (Kitagawa, Dokko, Okamura, & Yuh, 2005; Tavakoli, Aziminejad, Patel, & Moallem, 2006), or by superimposing circles over instrument tips changing from green to red according to measured forces on the real device (Reiley et al., 2008). To our knowledge, visual representation of interaction force for teleoperation of a mobile robot has not yet been studied.

According to Lombard and Ditton (1997), mediated sounds are also important in gen-

erating presence. In addition to hearing remote sounds through a headphone, visually representing sounds on a GUI can be beneficial. For instance, the difference in amplitude between left and right microphones on a urban search and rescue (USAR) platform was displayed using an icon sliding on a horizontal bar placed in the upper part of the video stream (Hestand & Yanco, 2004). This provided a one-dimensional sound location cue but no indication of the sound source's type, intensity, or the definitive location. Mizumoto et al. (2011) represent direction and intensity of sound source, but not the sound source localization, in relation to the position of the robot on the video stream. Sound visualization has also been studied to assist the hearing impaired. Matthews, Fong, Ho-Ching, and Mankoff (2007) compared the use of icons (to represent a phone ring or another sound), spectrograms, and a bar graph of sound intensities over time to display information such as the type, localization, amplitude and frequency of sounds. Results suggest that participants preferred simple and attractive means of sound visualization, from which the intensity can be evaluated at a glance. Azar, Saleh, and Al-Alaoui (2009) examined the representation of raw FFT (Fast Fourier Transform) data, spectrograms, and pitch vector displays, and all were found to be difficult to understand. Finally, Janicke, Borgo, Mason, and Chen (2010) used visual metaphors (i.e., icons for chimes, door being opened or closed, aircraft flying by, twittering, growling animal, and lightsaber) to illustrate film sound tracking in real time. These illustrations were useful for situational awareness when visual cues are absent (i.e., a person entering a room without being visible on the screen).

3. IRL-1/TR Robot System

As shown in Fig. 1, IRL-1/TR consists of a humanoid torso (IRL-1) and a mobile base (TR, for Telerobot) (Michaud et al., 2007), measuring 1.45 m in height and 0.60 m in both width and depth. IRL-1 is equipped with an expressive face; a pan-tilt head; a Kinect motion sensor (providing 640×480 RGB-D images with an angular field of view of 57° horizontally and 43° vertically, up to 5 m, 30 frames per second); a Hokuyo UTM-30LX laser range finder (30 m range, 270° angle, and 25 ms/scan, 0.4 m from the ground); an eight-microphone array; two compliant arms with four degrees of freedom (DOF) each; and grippers (Ferland et al., 2012). Images data are compressed using the image transport package from ROS (Robot Operating System) (Quigley et al., 2009), in which we added a libx264-based plugin to compress the RGB image. The libx264 is a free software library used to encode video streams into the H.264/MPEG-4 AVC compression format and allows for low latency video encoding, which is ideal for telepresence. The depth image is encoded using the lossless PNG compression plugin provided by the image transport package from ROS.

We used the TR base, a wheeled platform, for the robot's locomotion. Consisting of two triplets of wheels (one on the left and one on the right) and a rear rocker arm, the TR can transport the robot forward and backward and rotate it in place. A triplet of wheels is made of two omnidirectional wheels (at the extremities) and one motorized wheel (at the center), which are linked together by a bogie-like mechanism. The rear rocker arm connects the two triplets of wheels to another pivot attached to the robot's chassis. This rocker-bogie suspension minimizes displacement of the robot's chassis when moving on uneven surfaces, helping to provide stable video feed to the user.

Differential Elastic Actuators (DEAs) (Lauria, Legault, Lavoie, & Michaud, 2008) were used to provide force control and feedback to IRL-1/TR manipulators. Each arm on IRL-1/TR has four DOFs: three in the shoulder and one in the elbow, actuated using DEAs. The shoulder roll is responsible for end-effector orientation around an axis

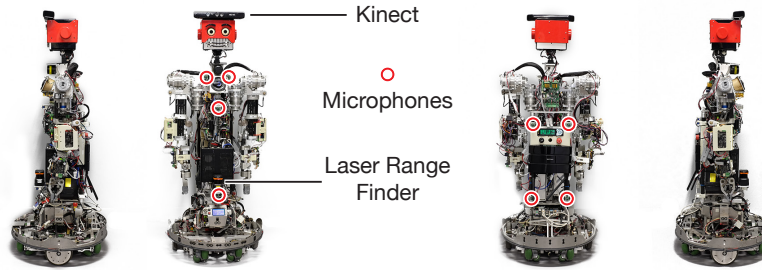


Figure 1. IRL-1/TR

that goes through the upper arm. A DEA acts as an active elastic element that can inherently absorb shocks, perceive the forces from the environment on the robot, and control the forces applied back. DEAs are conceptually similar to Series Elastic Actuators (SEAs) (Williamson, 1995; Robinson, 2000), which are used on the Meka M1, Cody (Jain & Kemp, 2009), Nexi-MDS platforms, and Baxter (Fitzgerald, 2013), but instead use a differential coupling (harmonic drive) rather than a serial coupling of a high-impedance mechanical speed source (an electrical DC brushless motor) and a low-impedance mechanical spring (a passive torsion spring). A non-turning joint sensor connected in series with the spring measures the torque output of the actuator.

ManyEars (Grondin et al., 2013) is the sound source localization, tracking and separation algorithm used with the eight-microphone array on IRL-1/TR. It performs real-time beamforming for localization, particle filtering for tracking, and Geometric Source Separation (GSS) for providing distinct audio streams of each sound source. It can localize and track up to four distinct sound sources and separate the content of up to three sources in real time and in reverberant environments. The algorithm locates the sound source within a 1 m unit sphere around the robot, with its origin at the bottom IRL-1/TR's neck, level with the top two front microphones, and provides the direction of the sound source.

Fig. 2 illustrates the software architecture implemented to teleoperate IRL-1/TR. A workstation and a gamepad controller are used to teleoperate IRL-1/TR. Wireless Ethernet (802.11N, 5 GHz WiFi network) is the medium used to interface the robot with the remote workstation. IRL-1/TR had no autonomous functionalities and was always teleoperated during the trials for this project to only observe the influences of the teleoperation user interface on the given tasks. Specific software modules used in this implementation are described in the following sections.

3.1 Arm Force Estimation

This module estimates the force applied on the end effectors. It provides joint-space impedance control of the DEAs of the arms and uses an algorithm inspired from Ferland, Aumont, Létourneau, and Michaud (2013). Fig. 3 illustrates a model of IRL-1/TR's 4 DOF arm, with θ_1 and θ_2 representing respectively the shoulder pan and tilt angles and θ_3 representing the elbow tilt angle. The fourth joint, the shoulder roll, is not represented, because this joint is responsible for the end-effector orientation and is not used in the force estimator algorithm. Axis Z_C is perpendicular to the page. Joints $\{1\}$, $\{2\}$, and $\{3\}$ are set with high stiffness in a way that each arm remains in a fixed position to perceive applied forces on these joints. $\{H_N\}$ and $\{H_C\}$ represent the end-effector reference frames when the arms

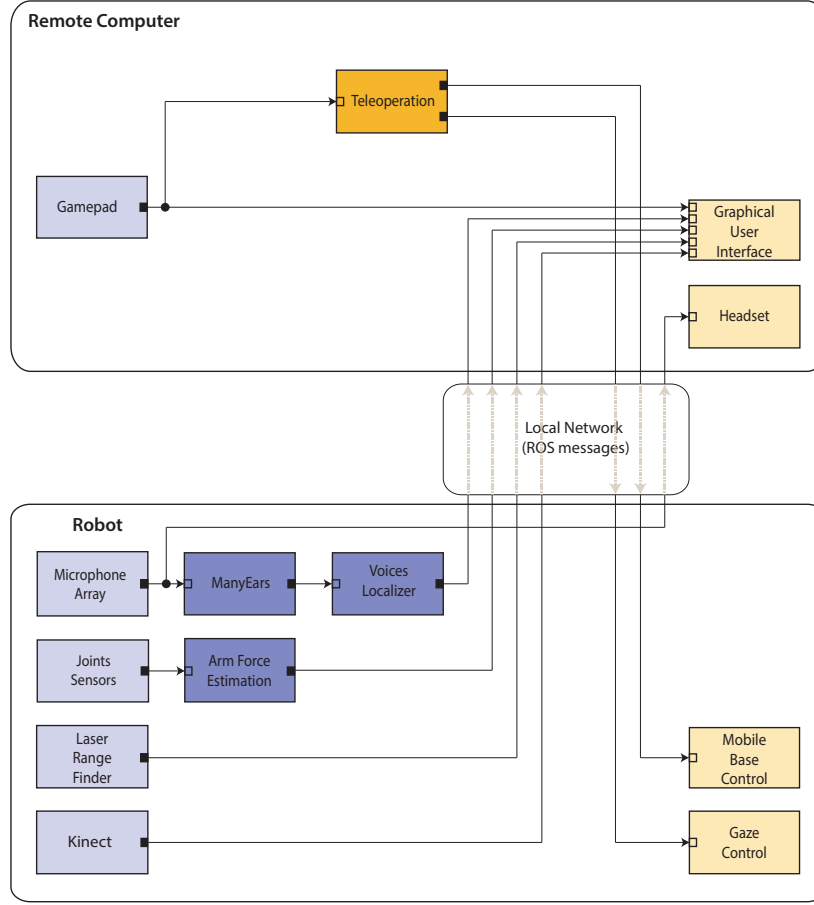


Figure 2. Teleoperation control architecture of IRL-1/TR. Empty squares represent inputs, and black squares represent outputs. Sensors are represented in light blue, perception modules in deep blue, the remote control module in gold, and control modules and devices in light yellow.

are in the neutral position and the current arm position, respectively. The neutral position was set empirically to have the arms high enough to be visible by the Kinect camera but also low enough to keep the center of gravity at a reasonable height and thus improve dynamic stability when maneuvering at high speed. Similarly, the flex angle at joint $\{3\}$ was selected to place the end effectors far enough from the robot without fully extending the elbow, as this would prevent measuring forces applied directly in line from the end effector to the elbow of the robot. The arms move when a force \vec{f}_A is applied on them. This force can be derived by modelling it as a virtual spring attached to the end-effector neutral position and the current position, according to Eq.1:

$$\vec{f}_A = k_A l_A \frac{\vec{r}_{H_N H_C}}{\|\vec{r}_{H_N H_C}\|} \quad (1)$$

where k_A is the virtual spring constant, l_A the virtual spring length and $\vec{r}_{H_N H_C}$ the vector between the reference frames. Table 1 presents the parameters used in our implementation with IRL-1/TR. These parameters were set empirically to ensure safety and to provide high sensitivity when small forces are applied on the end effectors.

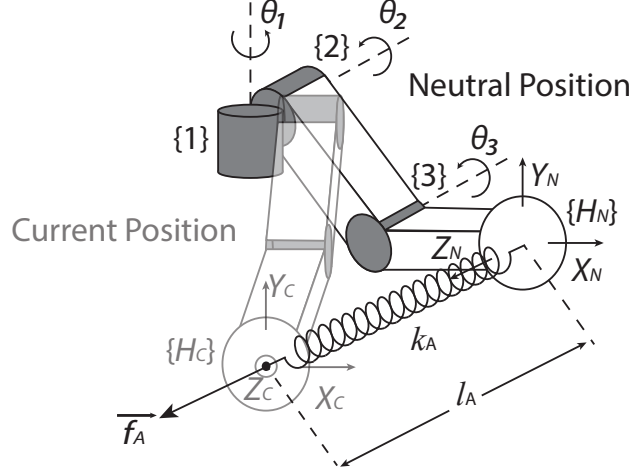


Figure 3. Representation of the model used for the arms of IRL-1/TR.

Table 1: Parameters of IRL-1/TR arm model

Parameter	Value	Unit
k_1	30.0	Nm/rad
k_2	30.0	Nm/rad
k_3	10.0	Nm/rad
θ_1	0.00	rad
θ_2	-0.90	rad
θ_3	-0.25	rad
k_A	140.00	N/m

3.2 Voice Localizer and ManyEars

This module is used to distinguish between speech and non-speech in the separated audio streams generated by ManyEars. The process consisted of performing pitch extraction and classification using the approach described in Sasaki, Kaneyoshi, Kagami, Mizoguchi, and Enomoto (2009).

First, a database of known speech and non-speech sounds was created and then encoded using vector quantization. Sounds in the database are depicted by a fixed number K of representative pitch clusters. Each sound signal is segmented using the short-time Fourier transform algorithm and binarized by comparison with a dynamic computed threshold to obtain a binarized pitch spectrum. By applying a K -means clustering algorithm to the pitch spectrum, only K representative clusters are computed. To visually differentiate speech

and non-speech sounds, Sasaki et al. (2009) recommend using $K = 26$ clusters. Using this database, input audio frames to classify either speech or non-speech are segmented, binarized, and categorized according to their probability of belonging to one or none of these clusters. The same approach could also be used to identify additional daily sounds such as water running, keys jingling, or doors slamming, if this would be required by the system.

3.3 Teleoperation

Teleoperation of IRL-1/TR happens by using the wireless gamepad, as shown in Fig. 4. The right thumb joystick controls motion of the mobile base: linear velocity is applied when it is moved up or down, angular velocity is generated by moving it left or right, and coupled translation and rotation are also possible. A deadzone of 5% is used. Linear velocity is limited to 0.45 m/s and acceleration to 1 m/s². Angular velocity is limited to 0.50 rad/s and acceleration to 4 rad/s². Commands can be generated only when the Live-Man button (bottom left shoulder button) is pressed. In addition, left thumb joystick controls head orientation (pan left and right, and tilt up or down), but this functionality was deactivated for the participant trials. Arm positions were not controlled using the gamepad. Head and arm positioning were achieved by scripts launched remotely through a text console for each experimental trial.



Figure 4. Gamepad used to control IRL-1/TR.

4. Graphical User Interface (GUI) With Visual Representations of Sound Source and Interaction Force

To explore the use of interaction force and sound source visual representations, we designed an ecological teleoperation interface for IRL-1/TR. As shown in Fig. 5, the basis of our teleoperation user interface consisted of an adjustable ego/exo-centric 3D display built from the color point clouds provided by the Kinect RGB-D images (Ferland et al., 2012). 3D rendering is implemented using the Visualization Toolkit library (VTK) (Schroeder, Martin, & Lorensen, 2006). The exocentric viewpoint (elevation angle = 30° and distance = 1.3 m from the head) is kept fixed for this study. This was chosen to have the operator see a virtual representation of the robot and what is seen from the Kinect, complemented by the laser range finder data over a wider but planar area. Keeping this viewpoint constant also removes a variable not critical for our study, that could, instead, influence observations during the trials. The point cloud generated by and visible from the Kinect camera is shown in middle of the display. An articulated rendering of IRL-1/TR is provided to visualize the orientation of the head and the position of its arms. A 3 m × 3 m square grid fixed to the world frame corresponding to the robot booting position and a circular grid with a 1.5 m

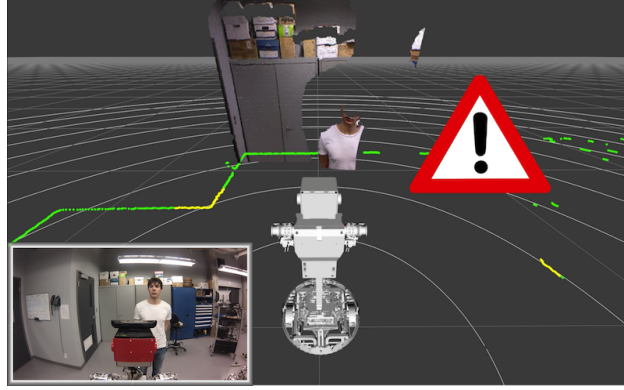


Figure 5. General view of the teleoperation user interface with laser range finder data, the image from the Kinect, and the navigation warning icon. The external view of the scene is presented on the bottom left portion of the image for reference and is not part of the GUI.

increment fixed to the mobile base are used to represent the floor and to display motion by moving the circular grid in relation to the square grid. Laser range readings are displayed as points with colors, from green (> 2.50 m) to yellow (≤ 2.50 m) and red (≤ 1.50 m), based on the distance of the perceived objects relative to the robot's current position. Distance is computed relative the end-effector of IRL-1/TR to provide precise feedback when forces are applied on the arms. When an obstacle is detected at less than 0.1 m from the end-effectors and $\pm 45^\circ$ in front of IRL-1/TR based on the laser range finder data, an exclamation point icon is displayed to warn the operator, to cover the blind spot caused by the rendering of the robot in the display. The icon is placed not to obstruct the view from the Kinect but can block some laser range finder data not critical for danger perception. This revealed to be an essential feature for communicating proximity information in such situations.

4.1 Visualization of Interaction Force

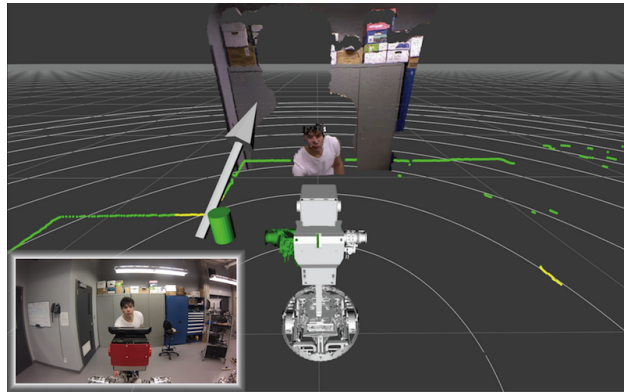
Visual cues used to represent the interaction force perceived on the arms are illustrated by Fig. 6 and consist of:

- Color and size of the arms. Based on the force intensity perceived for each arm, color goes from green (low intensity) to red (high intensity for f_{max} , the maximum force). The RGB components are defined according to Eq. 2. Arm opacity α is set to 50% to make the arm display transparent (and avoid obstructing a potential obstacle in front). When no force is perceived, the arm is represented in gray, as shown in Fig. 5 and Fig. 7.

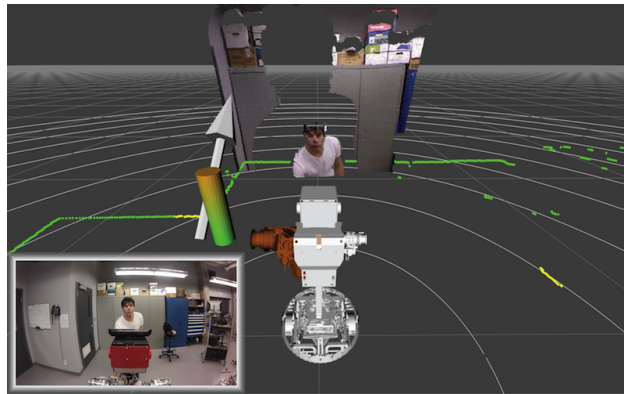
$$\begin{cases} R = \frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \\ G = 1 - R \\ B = 0 \end{cases} \quad (2)$$

The size of the arm is scaled from regular (low intensity) to large (high intensity) size according to Eq. 3:

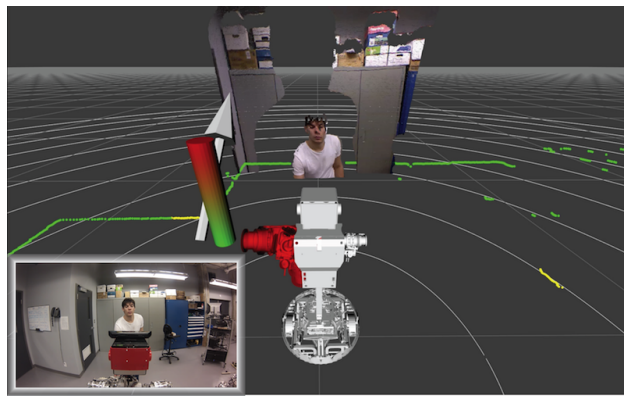
$$scale = 1 + \left(\frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \right) \quad (3)$$



(a) Low intensity



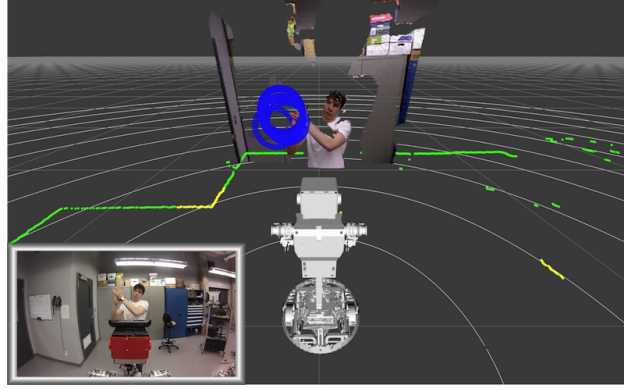
(b) Medium intensity



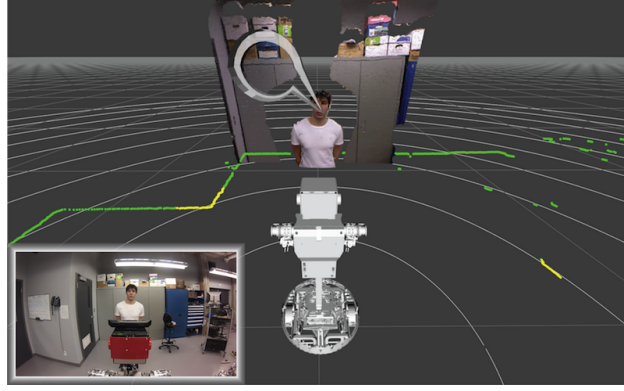
(c) High intensity

Figure 6. Display of interaction force applied on the left arm of IRL-1/TR when someone is pulling its end effector.

- Intensity Bar. The intensity bar position is set 0.35 m next to the axe of the shoulder tilt joint, to always remain visible to the operator, and grows in the +Z axis of the robot



(a) Five hand-clapping sounds



(b) Speech utterance

Figure 7. Examples of sound visualization.

reference frame. The bar is rendered by a colored cylinder with constant radius of 0.1 m and a length l_b computed by Eq. 4 to reach 1 m at the maximum authorized force. The top bar extremity color components are computed as in Eq. 2. Bar color is interpolated from green to the computed extremity color.

$$l_b = \frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \quad (4)$$

- Arrow. A gray arrow located next to the intensity bar represents the direction of the force sensed on the arm. The arrow is a unit vector oriented in the direction of the force applied on the end effector. The origin is placed 0.5 m next to the associated shoulder joint to be visible to the operator. Arrow length is constant and set to be 1 m. By looking at the arrow, the operator can see the force direction applied to the end-effector in relation to its neutral position.

4.2 Visualization of a Sound Source

For sound visualization, we chose to experiment with two icons to represent sound source location and type, as shown by Fig. 7: a blue ring, representing the location of an unidenti-

fied type or non-speech sounds, and a speech bubble, to localize where speech is identified. Since it takes at least 100 ms of audio stream to initiate the sound identification algorithm, speech is first displayed with a blue ring, to then change to a speech bubble. Once the sound has ceased, icon opacity is set to decrease linearly over a period of p , the persistence parameter. To position the icon of the sound source on the image, it is assumed that the sound source is located on the closest obstacle found in the direction of the sound, based on the Kinect depth image. The sound source location is estimated to be at the intersection between the straight line starting from the ManyEars' unit sphere origin toward the localized point and the Kinect depth image point. However, if an intersection is not found, the sound source location is set to be on a 1 m unit sphere as used by ManyEars. Sound source outside the field of view of the Kinect are also displayed on a 1 m unit sphere around IRL-1/TR.

5. Experimental Methodology

The underlying objective of our study is to explore how visual representations of interaction force and sound source can provide useful information for robot teleoperation. Considering that it would be difficult to measure the impact of combined modalities on participant situational awareness and cognitive load, and that this is a first integration of these modalities in a telepresence user interface, we decided to conduct small scale pilot studies to examine how people respond to these visual representations of interaction force and sound source individually and then test how well these modalities work in realistic conditions.

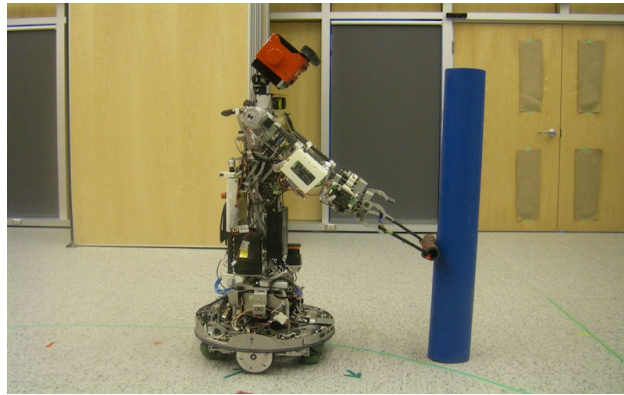
5.1 Force Visualization

For force visualization, the task we chose to experiment with consists of having the arms of IRL-1/TR touch a vertical pole without making it fall. We wanted to validate if the participants could achieve the task better and more precisely using force visualization. Fig. 8 illustrates the experimental setup. The vertical pole was 1.25 m in height and 0.15 m in diameter, weighing 9 kg and would fall when tilt was greater than 10° . This task was accomplished without sound perception or sound visualization.

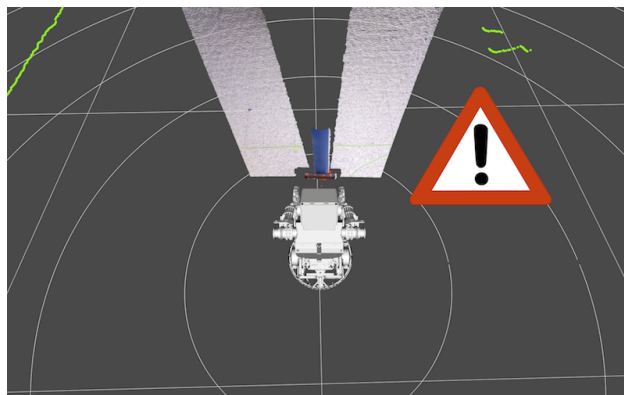
To make IRL-1/TR's arm visible by the Kinect (which has a minimal perceptual range of 0.5 m), we installed a U-shaped attachment to the grippers. The U-shaped attachment was made with a 0.3 m metal tube and a foam cover, with each end attached to the robot's grippers. The position of the arms remained fixed during the trials. IRL-1/TR's head was kept slightly tilted to make the U-shaped attachment visible in the field of view provided by the Kinect. The maximum linear velocity was set to 0.25 m/s while the maximum angular velocity was 0.30 m/s for precise motion control of the base. The constant f_{max} value was set empirically to have the arm displayed in red when $l_A > 0.05$ m to allow the operator to get precise feedback from the visual representation of interaction force on IRL-1/TR's arms. Each participant repeated the task ten times, five with and five without force visualization, randomly sampled. For each trial, IRL-1/TR started in front of the vertical pole at a distance of 1 m between the pole and the U-shaped attachment. The head was tilted downward at a 16° angle, and the exocentric viewpoint was situated 1.30 m away from the robot at a 30° elevation angle. The participant had to drive IRL-1/TR forward and stop when he or she believed that the U-shaped attachment was touching the vertical pole. Task completion time was measured, along with the number of times contacts were made with the pole, the tilt angle of the pole, and the number of times the pole fell. At the end of each trial, no indication was given to the participants about whether or not the U-shaped attachment touched the vertical pole. However, the participant was able to see if the pole fell, through the GUI.

When the trials for each participant were completed, participants had to answer the following question:

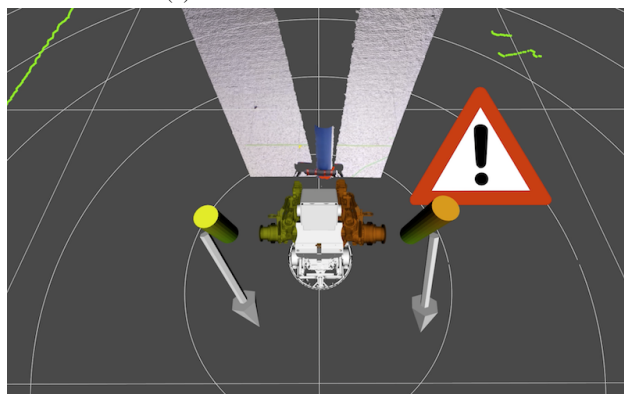
- Q1: Did you prefer achieving the task with or without force visualization?



(a) Experimental setup



(b) GUI without force visualization



(c) GUI with force visualization

Figure 8. Force visualization trials. The vertical pole was blue, 1.25 m in height, and 0.15 m in diameter. The U-shaped attachment was 0.36 m by 0.25 m.

5.2 Sound Visualization

Our goal was to evaluate the impact and the usefulness of visual sound source representations, assuming that sound identification icons would help localize sound source around IRL-1/TR. Two scenarios were tested: one without the robot moving and one in realistic conditions with the robot mobile.

5.2.1 Sound Visualization Under Static Conditions.

To isolate the influence of sounds from visual data (such as lip motion when people talk) or voice/face recognition, we decided to have IRL-1/TR stand still 1.5 m in front of four loudspeakers, each separated by 0.4 m and at a height of 1.2 m, visible in the field of view of the Kinect. Fig. 9 illustrates the experimental setup. Identification (ID) numbers were added to the teleoperation interface display to identify the sound source. Six sound tracks of male speech, each lasting 6 sec, were played on a randomly selected loudspeaker, with 2 sec of silence between each. To avoid recreating the same experimental scenario for each participant, we pre-recorded trials with the following two conditions:

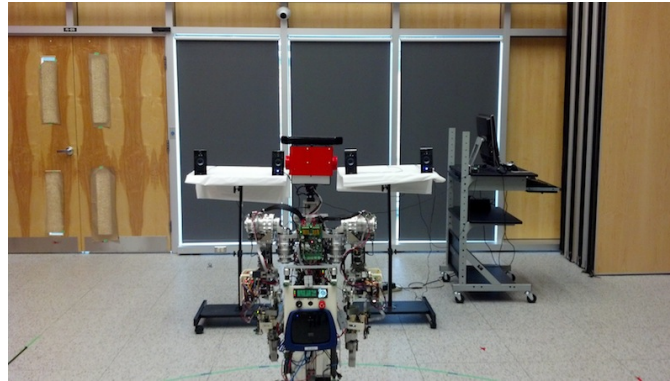
1. A sequence of six speech utterances without sound visualization, and a sequence of six speech utterances with sound visualization using the persistence parameter p set to 2 sec. The trial started by presenting either the sequence without or with sound visualization, chosen randomly. The resulting sequence of 12 speech utterances was used to evaluate the participant's ability to localize where the six sound sources are located over time. p was set to be equal to the silence time. Participants were asked to write down sound source ID numbers.
2. Sound visualization with the persistence parameter p set to 0.03, 0.5, 1, 2, 3, 4 and 5 sec, in that order. This generated a fixed sequence of 42 speech utterances to visualize for each participant. Participants were asked to focus only on the influence of the persistence parameter p and not on sound source localization.

Since p could last longer than the period of silence for Condition 2, the next randomly activated loudspeaker had to be different from the last. Each participant watched the recording while wearing a headphone. The headphone allowed the participant to hear the audio streams coming from the top-front left and top right microphones on IRL-1/TR through its left and right channels. Then, participants had to answer the following questions:

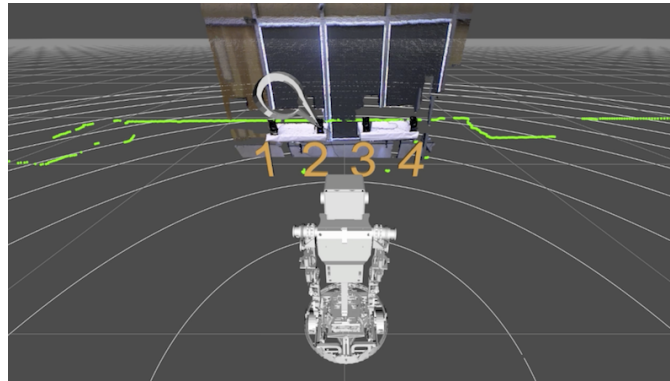
- Q2: For Condition 1, did you answer randomly or were you able to distinguish at least the right from the left?
- Q3: For Condition 1, did you prefer the use of sound visualization to accomplish the requested task, and why?
- Q4: For Condition 2, did you find icon persistence to be short, suitable, or long?

5.2.2 Sound Visualization Under Realistic Conditions.

In real conditions, teleoperating a robot requires significant attention of the operator, directed toward the interpretation of the visual scene. Therefore, for the sound visualization trials under realistic conditions, we decided to conduct a between-subject study with participants having IRL-1/TR push a trolley carrying boxes in a narrow room from one end to the other, where one of four loudspeakers had to be identified. Figs. 10 and 11 illustrate the experimental setup. From the IRL-1/TR's initial position, the loudspeakers were located 90° left in the opposite side of the room. The head of IRL-1/TR was tilted 16° downward with an exocentric viewpoint at 1.30 m and a 30° elevation angle (same as for



(a) Experimental setup



(b) Sound visualization

Figure 9. Sound visualization trials.

the force visualization task) so that the boxes remained in the field of view of the Kinect. Participants were asked to bring the trolley to the speakers at the edge of the table, keeping it at a distance of about 0.3 m from the speakers to avoid making contact with the table. At this distance, it was still possible to make the base of IRL-1/TR rotate on itself by at least 30° in each direction to help locate the sound source. As IRL-1/TR came within 2 m of the loudspeakers, one loudspeaker (randomly selected for the trial) started to emit a periodical sound consisting of 2 sec of white noise followed by a 5 sec pause. Speakers were located in an arc of a circle, each separated by 0.8 m. When confident, the participant had to identify the active loudspeaker. If a participant stopped the robot as soon as the sound was heard, the instruction to reach the table was repeated. To compensate for the difficulty of discriminating closely located sound sources, each participant was given two attempts to successfully localize the sound source. Force visualization was also activated during these trials, to help navigate safely in the environment and to have the operator do more than just focus on the navigation and sound localization, which also created more realistic conditions.

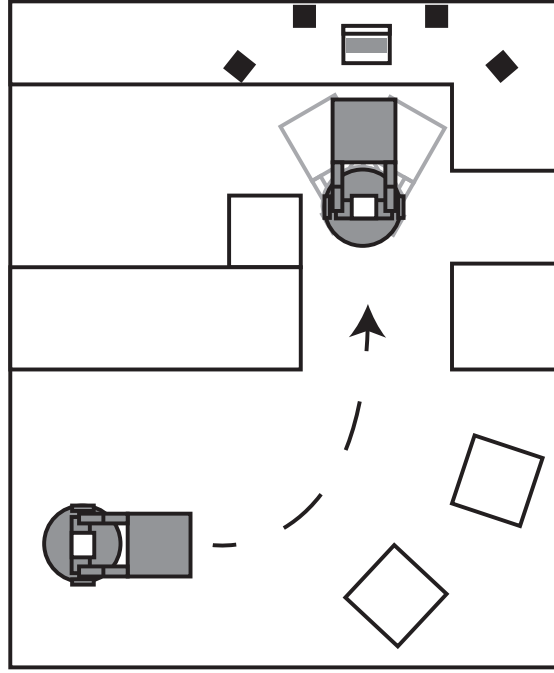


Figure 10. Top-view diagram of the typical path taken by users. Small black squares represent the loudspeakers, and light-gray lines represent the robot's range of motion when it reached its destination in front of the laptop computer.

5.3 Test Population, Familiarization and Settings

For the trials testing visual representations of interaction force and of sound source individually, our test population consisted of a convenience sample of 31 volunteers (26 males and 5 females), aged from 21 to 69, that had never used the teleoperation user interface or the gamepad controller as configured for the trials.

Before conducting the experiments, each participant was first brought in the experimentation room with IRL-1/TR and listened to a brief presentation on its hardware and perceptual capabilities, which included presenting the GUI's visualization modalities. Then, the participant took part in a two-step training session with the GUI and IRL-1/TR, all in the same room:

1. The participant controlled IRL-1/TR freely using the gamepad, interacted by pulling its arms or making sounds around the robot to see how these were displayed on the GUI. This step lasted approximately five minutes.
2. The participant trained to have IRL-1/TR touch the vertical pole at least two times with and without force visualization and until the participant felt ready to do the experiment. The participant was asked not to face the robot and the pole during the attempts but could take a look afterward to see the results. This step took from 10 to 20 min for each participant.

The experimental trials were conducted by having the participant teleoperate IRL-1/TR in a separate room. The force visualization trials were done before the sound visualiza-

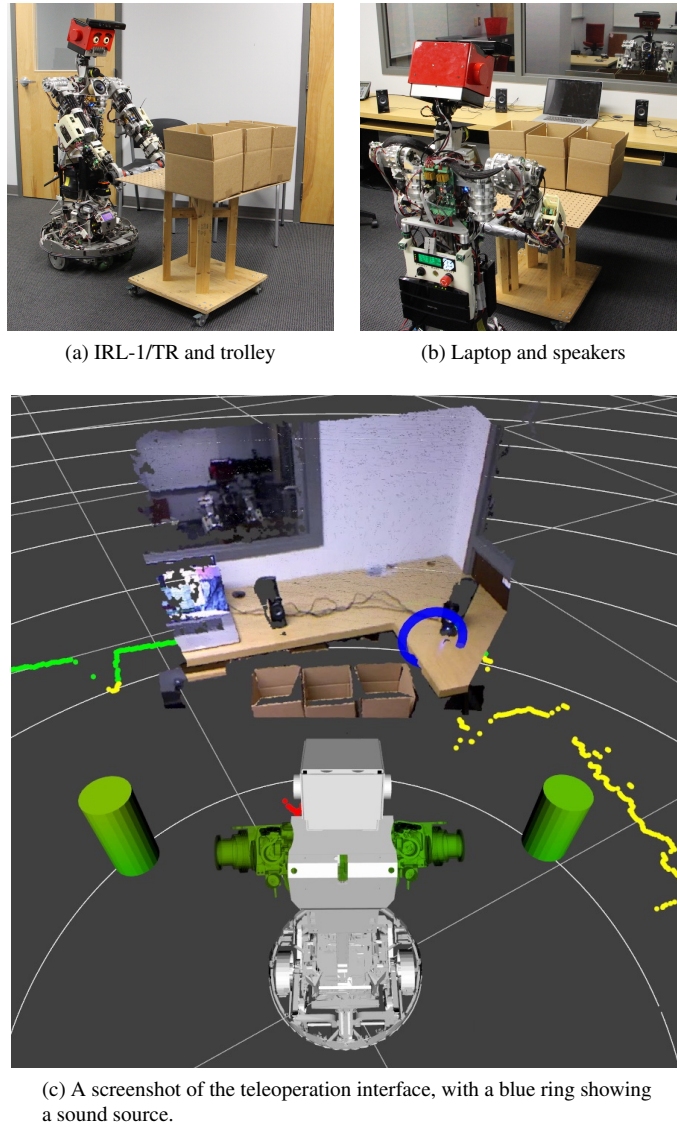


Figure 11. Experimental setup for the sound visualization trials under realistic conditions.

tion trials. It took about one hour per participant to complete the trials and to fill out the questionnaires.

For the sound visualization trials under realistic conditions, twenty participants (19 male, 1 female, age ranging from 22 to 38 with an average of 26.2) completed two trials. The first trial was for familiarization with the setup and the task but not with the sound condition (without or with sound visualization) to be tested. The second trial was to measure the time taken to identify the sound source once triggered and the number of attempts to successfully localize the active loudspeaker (with a maximum of two attempts) in the sound

condition to be tested. Each participant received instructions about the tasks, the GUI, the headphone, and the gamepad.

The GUI ran on a standard desktop computer with a 3.40 GHz 4th generation Core i7 quad core processor, 8 GB of RAM, a NVIDIA GPU with 1 GB of dedicated memory, and a 27 inch screen. IRL-1/TR had two on-board computers: a 2.67 GHz 1st generation Core i7 dual core processor with 3 GB of RAM located on the mobile base for the motion controller and a 2.10 GHz 3rd generation Core i7 quad core processor with 4 GB of RAM located in front of the torso and used for sensor processing, data compression and transmission, force estimation, and arm control. Communication and data exchange between the remote computer and IRL-1/TR were done using ROS (Quigley et al., 2009). Audio streams from the two upper microphones in front of IRL-1/TR were compressed using the Opus codec (Valin, Vos, & Terriberry, 2011) and sent to the headphone connected to the remote computer. This made it possible for the operator to get stereo inputs from the robot.

6. Results

6.1 Force Visualization

Regarding the force visualization trials, from the participants' performance and their answers to the questionnaire, we wanted to verify the following hypotheses:

- H1: The participants touch the pole (making it fall or not) more often with force visualization.
- H2: The participants make the pole fall less often with force visualization.

With H1, the goal is to validate whether or not force visualization makes it easier to reach a specific physical target in space (because visually, the participant may think the robot is touching the pole while it may not), while H2 verifies if the use of force visualization modalities result in more precise control of the platform. Results are analyzed using a one-tailed paired t-test (within subject study) to identify significant differences between metrics measured with and without the visualization modalities.

Data for 27 of the 31 participants were considered here because we initially started the trials with a longer vertical pole that was damaged after the fourth participant. We then had to alter the pole and make it smaller, which changed the experimental conditions.

Table 2 summarizes the results. The average time taken to complete the task was 17.9 sec without force visualization and 19 sec with force visualization, with no significant differences between the two conditions ($t = 1.53, dl = 26, p = 0.069$). The average number of cases that ended without touching the vertical pole is 1.8/5 ($\sigma = 1.4$) without force visualization, and 0.3/5 ($\sigma = 0.6$) with force visualization. Significant difference between the two conditions was observed ($t = 5.62, dl = 26, p < 0.001$). The average tilt angle was 2.4° ($\sigma = 1.9$) without force visualization and 1.8° ($\sigma = 1.2$) with force visualization. Significant difference between the two conditions was observed ($t = 1.745, dl = 26, p = 0.046$). Finally, the average number of fallen poles is 0.5/5 ($\sigma = 0.6$) without force visualization and 0.3/5 ($\sigma = 0.5$) with force visualization. In this case however, no significant difference is observed ($t = 1.54, dl = 26, p = 0.068$). These results suggest that H1 is validated in relation to the ability to touch the vertical pole with precision expressed in terms of the tilt angle, but H2 is not validated in relation to the number of times the participant made the vertical pole fall.

As for Q1, all participants answered that they preferred doing the task with force visualization because they were confident that the U-shaped attachment was touching the vertical pole. Some participants noticed that it was difficult to achieve the task using only

Table 2: Performance for the force visualization trials

		No contact (number)	Tilt angle (°)	Pole fall (number)
Without force visualization	Mean	1.8	2.4	0.5
	σ	1.4	1.9	0.6
With force visualization	Mean	0.3	1.8	0.3
	σ	0.6	1.2	0.5

the 3D point cloud of the Kinect because of the shadow created by the U-shaped attachment. Another observation is that applying high acceleration or deceleration to the mobile base of IRL-1/TR made the arms move, which generated forces that were displayed on the GUI. Some participants indicated annoyance by these displays of force. As a solution, force visualization could be activated by the user only when necessary or an arm position compensation algorithm could be used.

6.2 Sound Visualization

The hypothesis to validate in both of the experimental conditions involving sound visualization is as follows:

- H3: The participants localize sound source near IRL-1/TR more often with sound visualization.

6.2.1 Sound Visualization Under Static Conditions.

For Condition 1, Table 3 summarizes the average error observed when having participants list the sound source ID number. Without sound visualization, the average error rate is 2.64/6 (44%), with a standard deviation of $\sigma = 1.8/6$ (30%). Five participants found all the right answers, but only one of them, an experienced musician, was confident about the answers provided. Looking to see if the participants were at least able to distinguish between the left or right sound source, the error rate without sound visualization decreases to 1.2/6 (20%), with $\sigma = 1.62/6$ (27%). Thirty of the thirty-one participants indicated that they were able to distinguish between right and left source just from the headphone but revealed to have difficulty when it came to distinguishing loudspeakers from the same side. Humans can localize sound source coming from the side in the horizontal plane with an error from 8° to 20° according to source angle (Blauert, 1996). In our trials, loudspeakers on the same side were separated by 14.2° from each other, making it hard for participants to discriminate using the headphone. With sound visualization however, the average error rate is 0.06/6 (1%), with $\sigma = 0.18/6$ (3%). Only one participant made a mistake by waiting until the end of the trial to list the sound source ID number, and forgot one answer. Significant difference ($t = 8.076, dl = 30, p < 0.001$) was observed between error rates without and with sound visualization: The participants provided more correct identifications of the speaker's location with the sound source visualization than without, validating H3.

In relation to Q2, only one participant admitted to have answered randomly. Regarding Q3, 30 of the 31 participants found sound visualization useful because it provided precise and simple indications, allowing them to concentrate on the task. One participant in the set of 30 only focused on the visual display and not on the audio stream: The participant

mentioned that the visual display overloaded unnecessarily the teleoperation interface, and found this feature useful only when necessary. However, without the use of sound visualization, this participant made two errors of the six conditions, with $p = 2$.

Table 3: Number of identification errors for the sound visualization trials, with $p = 2$ sec

		ID	Right/Left
Without sound visualization	Mean	2.64/6	1.20/6
	σ	1.80/6	1.62/6
With sound visualization	Mean	0.06/6	0.06/6
	σ	0.18/6	0.18/6

For Condition 2, Table 4 summarizes results regarding Q4. For a persistence of $p = 0.03$ sec, 22 of the 31 participants found icon persistence to be too short. Icons disappeared between two localizations, resulting in noticeable flickering which was considered to be annoying. Persistence of 0.5 sec and 1 sec were considered suitable by 26 of the 31 participants. Finally, p values of 2 sec or greater were considered as being too long for at least 21 participants: The previous icon had not disappeared when another loudspeaker was activated, making a second icon appear on the GUI. As an alternative, it was suggested to use different colors for different sources.

Table 4: Perception of icon persistence for sound visualization (Q4).

p (sec)	0.03	0.5	1	2	3	4	5
Short	22	4	1	0	0	0	0
Suitable	9	26	26	10	2	0	0
Long	0	1	4	21	29	31	31

6.2.2 Sound Visualization Under Realistic Conditions.

Table 5 presents the results of the sound visualization trials under realistic conditions, expressed in terms of the time taken to either successfully identify to the sound source or to reach the maximum of two attempts and the number of unsuccessful trials (Error Count). Welch's t -tests for both the times taken ($p = 0.015$, $df = 10.562$) and the Error Count ($p = 0.026$, $df = 9$) suggests that participants performed better with sound visualization than without. No errors were observed with sound visualization, clearly validating H3. Also, without sound visualization, two participants made an error on their first attempt but found the correct source on the second one, and two other participants reached the maximum two attempts without localizing the sound source. Furthermore, sound source localization was accomplished more than two times faster with sound visualization.

To get some insights about how performance can be affected by the experience of the operator, we conducted trials with three participants that repeated each condition (without and with sound visualization) five times. We observed that the number of errors decreases as the operator gain experience with task, for instance, by moving IRL-1/TR to help discriminate between the sound sources. However, in all cases, the average time taken to localize the sound source is shorter with visualization rather than without.

Table 5: Performance for the sound visualization trials Under realistic conditions

		Average	σ
Without sound visualization	Time (sec)	46.8	30.1
	Error Count	0.6	0.8
With sound visualization	Time (sec)	22.1	8.9
	Error Count	0	0

While force visualization was not the focus of this experiment, we still observed its use when two participants collided with a table before reaching the end of the course. These two participants mentioned seeing that something went wrong when the left arm of the robot briefly turned red. Also, one participant associated this visual event to hearing a faint sound to the left of the robot, confirming that a collision with the left arm occurred.

7. Conclusion and Future Work

Sound and force sensing capabilities of mobile platforms bring additional information to be displayed on telepresence interfaces. This paper demonstrates how force and sound visualization can be used to provide information on an ecological 3D telepresence user interface. Results suggest that sound visualization can provide useful information if persistence is appropriately set. Force visualization enables users to achieve more accurate and precise manipulation. Based on the observations made from the pilot studies conducted, we plan to conduct additional trials examining additional modalities, such as changing the size of the sound icons according to the amplitude of the sound, and study the influence of different force visualization modalities (e.g., combine or not the intensity bar with the arrow, compare the color/size of the arms with the intensity/arrow modalities, and allowing operators to change viewpoints) in relation to the task, and comparing their uses by novice and experienced operators. In future work, these modalities will also be compared and integrated with the use of additional feedback modes (e.g., sound log, vibration, haptic) to create an integrated telepresence mobile robotic system for homecare assistance.

Acknowledgements

This work was supported by the Fonds de Recherche du Québec - Nature et Technologies (FRQ-NT). The authors want to thank Ronan Chauvin and all the volunteers who took part in the trials.

References

- Azar, J., Saleh, H. A., & Al-Alaoui, M. (2009). Sound visualization for the hearing impaired. *International Journal of Emerging Technologies in Learning*, 2(1), 1–7.
- Bar-Cohen, Y. (2003). Haptic devices for virtual reality, telepresence, and human-assistive robotics. *Biologically Inspired Intelligent Robots*, 73.
- Blauert, J. (1996). *Spatial hearing Revised edition: The psychophysics of human sound localization*. Cambridge, MA: The MIT Press.
- Borst, C., Wimbrock, T., Schmidt, F., Fuchs, M., Brunner, B., Zacharias, F., . . . Hirzinger, G. (2009). Rollin’ Justin Mobile platform with variable base. In *Proceedings of the IEEE International Conference on Robotics and Automation* (p. 1597–1598). doi:10.1109/ROBOT.2009.5152586
- Chen, T. L., & Kemp, C. (2011). A direct physical interface for navigation and positioning of a robotic nursing assistant. *Advanced Robotics*, 25(5), 605–627.

- Dahl, T. S., & Boulos, M. N. K. (2013). Robots in health and social care: A complementary technology to home care and telehealthcare? *Robotics*, 3(1), 1–21.
- Desai, M., Tsui, K. M., Yanco, H. A., & Uhlik, C. (2011). Essential features of telepresence robots. In *Proceedings of the IEEE International Conference on Technologies for Practical Robot Applications* (p. 15–20). doi:10.1109/TEPRA.2011.5753474
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 97–101. doi: 10.1177/154193128803200221
- Ferland, F., Aumont, A., Létourneau, D., & Michaud, F. (2013). Taking your robot for a walk: Force-guiding a mobile robot using compliant arms. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (p. 309–316). doi: 10.1109/HRI.2013.6483604
- Ferland, F., Létourneau, D., Aumont, A., Frémy, J., Legault, M.-A., Lauria, M., & Michaud, F. (2012). Natural interaction design of a humanoid robot. *Journal of Human-Robot Interaction, Special Issue on HRI Perspectives and Projects from Around the Globe*, 1(2), 14–29.
- Ferland, F., Pomerleau, F., Dinh, C. T. L., & Michaud, F. (2009). Egocentric and exocentric teleoperation interface using real-time, 3D video projection. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction* (p. 37–44). La Jolla, CA.
- Fitzgerald, C. (2013). Developing Baxter. In *Proceedings of the IEEE International Conference on Technologies for Practical Robot Applications* (p. 1–6). doi:10.1109/TePRA.2013.6556344
- Gonzalez-Jimenez, J., Galindo, C., & Ruiz-Sarmiento, J. (2012). Technical improvements of the Giraff telepresence robot based on users’ evaluation. In *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication* (p. 827–832). doi: 10.1109/ROMAN.2012.6343854
- Graf, B., Parlitz, C., & Hagele, M. (2009). Robotic home assistant Care-O-bot 3 product vision and innovation platform. In *Human-Computer Interaction Novel Interaction Methods and Technique Lecture Notes in Computer Science* (Vol. 5611, p. 312–320). doi:10.1007/978-3-642-02577-8-34
- Grondin, F., Létourneau, D., Ferland, F., Rousseau, V., & Michaud, F. (2013). The ManyEars open framework. *Autonomous Robots*, 34(3), 217–232.
- Guizzo, E. (2010). When my avatar went to work. *IEEE Spectrum*, 47(9), 26–50.
- Hestand, D., & Yanco, H. A. (2004). Layered sensor modalities for improved human-robot interaction. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (p. 2966–2970). doi:10.1109/ICSMC.2004.1400784
- Hu, H., Li, J., Xie, Z., Wang, B., Liu, H., & Hirzinger, G. (2005). A robot arm/hand teleoperation system with telepresence and shared control. In *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (p. 1312–1317). doi: 10.1109/AIM.2005.1511192
- Jain, A., & Kemp, C. C. (2009). Pulling open novel doors and drawers with equilibrium point control. In *Proceedings of the IEEE-RAS International Conference on Humanoid Robotics*. Paris, France.
- Janicke, H., Borgo, R., Mason, J. S. D., & Chen, M. (2010). Sounddriver: Semantically-rich sound illustration. *Computer Graphics Forum*, 29(2), 357–366.
- Keyes, B., Micire, M., Drury, J. L., & Yanco, H. A. (2010). Improving human-robot interaction through interface evolution. In *Human-Robot Interaction* (p. 183–202). InTech. doi: 10.5772/8140
- Kitagawa, M., Dokko, D., Okamura, A. M., & Yuh, D. D. (2005). Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *The Journal of Thoracic and Cardiovascular Surgery*, 129(1), 151–158.
- Labonté, D., Boissy, P., & Michaud, F. (2010). Comparative analysis of 3-D robot teleoperation interfaces with novice users. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 40(5), 1331–1342.

- Lauria, M., Legault, M. A., Lavoie, M. A., & Michaud, F. (2008). Differential elastic actuator for robotic interaction tasks. In *Proceedings of the IEEE International Conference on Robotics and Automation* (p. 3606–3611). doi:10.1109/ROBOT.2008.4543763
- Lee, M. K., & Takayama, L. (2011). "Now, I have a body": Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 33–42). doi:10.1145/1978942.1978950
- Lindeman, R. W. (2003). Virtual contact : The continuum from purely visual to purely physical. In *Proceedings of the 47th Annual Meeting of the Human Factor and Ergonomics Society* (p. 2103–2107). Santa Monica, CA.
- Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2).
- Matthews, T., Fong, J., Ho-Ching, W.-L., & Mankoff, J. (2007). Evaluating non-speech sound visualizations for the deaf. *Behavior and Information Technology*, 25(4), 333–351.
- Mendez, I., Jong, M., Keays-White, D., & Turner, G. (2013). The use of remote presence for health care delivery in a northern Inuit community: A feasibility study. *International Journal of Circumpolar Health*, 72.
- Michaud, F., Boissy, P., Labonté, D., Brière, S., Perreault, K., Corriveau, H., ... Létourneau, D. (2010). Exploratory design and evaluation of a homecare teleassistive mobile robotic system. *Mechatronics*, 20(7), 751–766.
- Michaud, F., Boissy, P., Labonté, D., Corriveau, H., Granty, A., Lauria, M., ... Royer, M.-P. (2007). Telepresence robot for home care assistance. In *Proceedings of Multidisciplinary Collaboration for Socially Assistive Robotics, AAAI Spring Symposium, Technical Report SS-07-07* (p. 50–55).
- Mizumoto, T., Nakadai, K., Yoshida, T., Takeda, R., Otsuka, T., Takahashi, T., & Okuno, H. (2011). Design and implementation of selectable sound separation on the Texai telepresence system using HARK. In *Proceedings of the IEEE International Conference on Robotics and Automation* (p. 2130–2137). doi:10.1109/ICRA.2011.5979849
- Nakadai, K., Takahashi, T., Okuno, H., Nakajima, H., Hasegawa, Y., & Tsujino, H. (2010). Design and implementation of robot audition system 'HARK' open source software for listening to three simultaneous speakers. *Advanced Robotics*, 5(6), 739–761.
- Nielsen, C. W., Goodrich, M. A., & Ricks, R. W. (2007). Ecological interfaces for improving mobile robot teleoperation. *IEEE Transactions on Robotics*, 23(5), 927–941.
- Pamungkas, D., & Ward, K. (2013). Tele-operation of a robot arm with electro tactile feedback. In *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (p. 704–709). doi:10.1109/AIM.2013.6584175
- Quigley, M., Conley, K., Gerkey, B. P., Faust, J., Foote, T., Leibs, J., ... Ng, A. Y. (2009). ROS: An open-source Robot Operating System. In *Proceedings of the ICRA Workshop on Open Source Software*. Kobe, Japan.
- Reiley, C., Akinbiyi, T., Burschka, D., Chang, D., Okamura, A., & Yuh, D. (2008). Effects of visual force feedback on robot-assisted surgical task performance. *The Journal of Thoracic and Cardiovascular Surgery*, 135(1), 196–202.
- Ricks, B., Nielsen, C. W., & Goodrich, M. A. (2004). Ecological displays for robot interaction: A new perspective. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 2855–2860). doi:10.1109/IROS.2004.1389842
- Robinson, D. (2000). *Design and analysis of series elasticity in closed loop actuator force control* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.
- Robles-De-La-Torre, G. (2006). The importance of the sense of touch in virtual and real environments. *IEEE MultiMedia*, 13(3), 24–30.
- Sasaki, Y., Kaneyoshi, M., Kagami, S., Mizoguchi, H., & Enomoto, T. (2009). Daily sound recognition using pitch-cluster-maps for mobile robot audition. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 2724–2729). doi:10.1109/IROS.2009.5354241

- Scholtz, J., Young, J., Yanco, H., & Drury, J. (2004). Where am I? Acquiring situation awareness using a remote robot platform. In *Proceedings of the IEEE Conference on Systems, Man and Cybernetics* (p. 2835–2840). doi:10.1109/ICSMC.2004.1400762
- Schroeder, W., Martin, K., & Lorensen, B. (2006). *Visualization Toolkit: An Object-Oriented Approach to 3D Graphics* (4th ed.). USA: Kitware.
- Tavakoli, M., Aziminejad, A., Patel, R. V., & Moallem, M. (2006). Methods and mechanisms for contact feedback in a robot-assisted minimally invasive environment. *Surgical Endoscopy And Other Interventional Techniques*, 20(10), 1570–1579.
- Valin, J., Vos, K., & Terriberry, T. (2011). *Definition of the Opus Audio Codec*. Retrieved from <http://tools.ietf.org/html/rfc6716>
- Williamson, M. M. (1995). *Series elastic actuators* (Master's thesis). Massachusetts Institute of Technology, Cambridge, MA.
- Wyrobek, K., Berger, E., der Loos Van, H., & Salisbury, K. (2008). Towards a personal robotics development platform: Rationale and design of an intrinsically safe personal robot. In *Proceedings of the IEEE International Conference on Robotics and Automation*. doi: 10.1109/ROBOT.2008.4543527
- Yanco, H. A., Drury, J. L., & Scholtz, J. (2004). Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Human-Computer Interaction*, 19(1-2), 117–149.

Authors' names and contact information: Aurélien Reveleau (aurelien.reveleau@usherbrooke.ca), François Ferland (francois.ferland@usherbrooke.ca), Mathieu Labbé (mathieu.m.labbe@usherbrooke.ca), Dominic Létourneau (dominic.letourneau@usherbrooke.ca), François Michaud (francois.michaud@usherbrooke.ca), Interdisciplinary Institute for Technological Innovation, Université de Sherbrooke, 3000, boul. de l'Université. Sherbrooke (QC) Canada, J1K 0A5.