

UNIVERSITÉ DE SHERBROOKE
Faculté de génie
Département de génie électrique et de génie informatique

REPRÉSENTATION VISUELLE DES
SOURCES SONORES, DES FORCES
D'INTERACTION ET INTÉGRATION
D'UNE REPRÉSENTATION 3D DE
L'ENVIRONNEMENT DANS UNE
INTERFACE DE TÉLÉOPÉRATION POUR
ROBOT MOBILE

Mémoire de maîtrise
Spécialité : génie électrique

Aurélien REVELEAU

Jury : François Michaud (directeur)
Daniel Labonté
Ahmed Khoumsi

Sherbrooke (Québec) Canada

Janvier 2015

À mon père, Noël.

RÉSUMÉ

Ces dernières années, de nombreuses compagnies ont mis sur le marché des robots de téléprésence pour des usages de vidéoconférence mobile. Ces plateformes robotiques transmettent généralement la vidéo, l'audio et les données de capteurs de proximité à l'opérateur distant par l'intermédiaire d'une interface de téléopération s'exécutant sur un ordinateur de bureau standard. De nouvelles modalités d'interaction comme la localisation et le suivi de sources sonores, la capacité de ressentir les forces appliquées et la visualisation d'une représentation 3D de l'environnement sont développées et testées sur des plateformes robotique avancées. Il devient donc nécessaire d'explorer de nouvelles techniques pour intégrer ces informations supplémentaires au sein de l'interface.

Ce projet de maîtrise propose une solution pour représenter visuellement les sources sonores et les forces d'interaction du robot à travers une interface de téléopération 3D. Cette interface inclut sous une seule vue les fonctionnalités de base comme la représentation du modèle du robot, la reconstruction 3D du nuage de points transmis par le capteur Kinect, la présentation des données des capteurs de proximité ainsi que la possibilité de passer d'une vue égocentrique à une vue exocentrique. L'affichage des sources sonores est réalisé avec des bulles de bandes dessinées pour la parole et des anneaux bleus pour tous les autres types de sons. L'intensité des forces d'interaction sur les bras du robot est affichée par changement de couleur et de taille du modèle 3D et avec des jauges. La direction des forces exercées est matérialisée par des flèches. Enfin, le champ de vision limité de la caméra est augmenté par l'affichage d'une représentation 3D de l'environnement construit dynamiquement par le robot. La phase de validation a été effectuée avec la plateforme robotique IRL-1 qui dispose des modalités perceptuelles nécessaires.

Les performances de 31 participants ont été comparées dans l'exécution de trois sous-tâches réalisées avec et sans l'ajout des modalités supplémentaires. La première épreuve d'écoute visait à connaître l'impact de l'affichage de la position des sources sonores. La tâche de manipulation évaluait l'impact de l'affichage des forces ressenties sur les bras. Enfin, le pilotage du robot à travers un slalom avait pour but de tester l'impact de la représentation 3D de l'environnement. L'étude démontre que l'affichage visuel des modalités d'interaction sonores et de force du robot soulage et améliore les performances des utilisateurs dans la réalisation de certaines tâches précises d'écoute ou de manipulation. Cependant, les tests de pilotage du robot à travers un slalom autour de trois cylindres ne permettent pas d'affirmer que l'ajout d'une représentation 3D de l'environnement déjà visité par le robot améliore les performances de navigation.

Mots-clés : Interface-opérateur, robotique de téléprésence, robots mobiles, représentation visuelle des sources sonores et des forces d'interaction

REMERCIEMENTS

Tout d'abord, j'aimerais remercier François Michaud pour son soutien et ses précieux conseils tout au long du projet.

Je remercie aussi tous les membres du laboratoire IntrRoLab, et plus particulièrement François Ferland pour sa patience et son aide précieuse sur les aspects techniques du projet, ainsi que Dominic Létourneau pour sa disponibilité et sa réactivité dans la maintenance du robot.

Merci également à Ronan Chauvin, ami de longue date qui m'a aidé dans la réalisation des expériences.

Enfin, je tiens à remercier le Fond de Recherche du Québec, Nature et Technologie (FRQ-NT) pour son soutien financier.

TABLE DES MATIÈRES

1	INTRODUCTION	1
2	INTERFACES DE TÉLÉOPÉRATION POUR ROBOTS MOBILES	3
2.1	Standards de conception pour les interfaces de téléopération	3
2.2	Approches par interfaces écologiques	5
2.3	Interfaces de téléopération et réalité augmentée	8
2.4	Métriques pour les interactions humain-robot	10
2.4.1	Conscience de la situation	11
2.4.2	Charge cognitive	13
2.5	Retours d'informations	14
2.5.1	Retours visuels d'informations de forces	15
2.5.2	Retours visuels d'informations de sons	15
3	REPRÉSENTATION VISUELLE DES FORCES D'INTERACTION ET DES SOURCES SONORES DANS UNE INTERFACE DE TÉLÉOPÉ- RATION	19
3.1	Avant-propos	19
3.2	Abstract	21
3.3	Introduction	21
3.4	Related Work	22
3.5	Visual Representations of Interaction Forces and Sound Sources Using IRL- 1/TR	23
3.5.1	Arm Force Estimation	25
3.5.2	Voice Localizer and ManyEars	27
3.5.3	Teleoperation	28
3.5.4	Graphical User Interface	28
3.6	Experimental Methodology	31
3.6.1	Force Visualization	33
3.6.2	Sound Visualization	34
3.6.3	Test Population, Familiarization and Settings	35
3.7	Results	36
3.7.1	Force Visualization	37
3.7.2	Sound Visualization	38
3.8	Conclusion and Future Work	39
4	INTÉGRATION DE LA REPRÉSENTATION 3D DE L'ENVIRONNE- MENT DANS UNE INTERFACE DE TÉLÉOPÉRATION	43
4.1	Expérimentation	44
4.2	Résultats	46
5	CONCLUSION	49

A PROFIL DES PARTICIPANTS	51
B ATTESTATION DE CONFORMITÉ DU COMITÉ D'ÉTHIQUE	53
C FORMULAIRE DE DEMANDE D'ÉVALUATION ÉTHIQUE	55
LISTE DES RÉFÉRENCES	71

LISTE DES FIGURES

2.1	Vue similaire affichée avec deux interfaces différentes	6
2.2	Interface à modalité vidéo centrée 2D (VC2D)	7
2.3	Interface à modalité réalité augmentée 3D (RA3D)	7
2.4	Interface à modalité mixte exocentrique 3D (ME3D)	7
2.5	Interface graphique tridimensionnelle développée par Ferland <i>et al.</i>	8
2.6	Interface de téléopération avec affichage des directions par réalité augmentée	9
2.7	Interface avec représentation des sons par curseur	16
2.8	Historique des sons émis dans le temps en fonction de leur intensité et de leur contexte	16
2.9	Interface de visualisation des sons pour malentendants	17
2.10	Représentation des sons par des icônes	18
3.1	IRL-1/TR	24
3.2	Teleoperation control architecture of IRL-1/TR. Empty squares represent inputs, and filled squares represent outputs.	26
3.3	Representation of the model used for the arms of IRL-1/TR.	27
3.4	Gamepad used to control IRL-1/TR.	28
3.5	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.	30
3.6	Force	32
3.7	Sound visualization	33
3.8	Force visualization trials. The vertical pole was blue and 1.25 m in height and 0.15 m in diameter. The U-shaped attachment was 0.36 m by 0.25 m.	40
3.9	Sound visualization trials	41
4.1	Architecture logicielle du module de cartographie 3D	44
4.2	Représentation 3D de l'environnement affichée sur l'interface.	45
4.3	Vue externe du parcours à accomplir	47
4.4	Vue de l'interface pendant les expérimentations de navigation	47

LISTE DES TABLEAUX

3.1	Parameters of IRL-1/TR arm model	27
3.2	Performance for the force visualization trials	37
3.3	Number of identification errors for the sound visualization trials, with $p = 2$ sec	38
3.4	Perception of icon persistence for sound visualization (Q4).	39
A.1	Profil des participants	52

CHAPITRE 1

INTRODUCTION

De nos jours, de nombreuses plateformes robotiques de téléprésence sont commercialisées par diverses compagnies [Guizzo, 2010] (e.g., VGo par VGo Communications, QB par Anybot [Desai *et al.*, 2011], TiLR par RoboDynamics, Giraff par HeadThere [Gonzalez-Jimenez *et al.*, 2012], beam+ par Suitable Technologies, TeleMe par Mantarobot, Double par Double Robotics, RP7i et RP Vita par InTouch Health). Ces robots sont généralement constitués d'une base mobile, d'une caméra, de capteurs de proximité, d'un écran et de microphones, faisant d'eux des systèmes de vidéoconférences mobiles, parfois aussi qualifiés de "Skype sur roues" [Dahl et Boulos, 2013]. Leurs usages s'étendent du télétravailleur voulant assister virtuellement à une réunion de travail [Lee et Takayama, 2011] à des usages de télésanté dans les hôpitaux ou à domicile [Mendez *et al.*, 2013; Michaud *et al.*, 2010].

Pour contrôler ces diverses plateformes, les informations de senseurs et la conception de l'interface de téléopération jouent un rôle clé pour maximiser la conscience de la situation (*Situation Awareness*) [Endsley, 1988; Scholtz *et al.*, 2004b] tout en minimisant la charge cognitive imposée aux opérateurs (*Cognitive Load*) [Yanco *et al.*, 2004]. L'ajout de nouvelles modalités perceptuelles se fit via la présentation de diverses fenêtres et utilisant diverses métaphores, ce qui avait pour tendance d'augmenter la charge cognitive des opérateurs. De récents travaux de recherche ont démontré la nécessité de minimiser le nombre de fenêtres afin d'afficher toute l'information sur la vue principale [Keyes *et al.*, 2010]. À cette fin, le paradigme d'interfaces écologiques utilise la réalité augmentée afin de représenter le robot de téléprésence dans son environnement virtuel tout en y intégrant le flux vidéo et des informations de carte [Ferland *et al.*, 2009; Labonté *et al.*, 2010; Nielsen *et al.*, 2007; Ricks *et al.*, 2004].

Les robots récents comme Rollin' Justin [Borst *et al.*, 2009], PR2 [Wyrobek *et al.*, 2008], Cody [Chen et Kemp, 2011] et Care-O-bot [Graf *et al.*, 2009] offrent des modalités d'interactions supplémentaires en étant équipés de bras manipulateurs compliants. Il est aussi possible d'équiper les plateformes mobiles de capacités de localisation, de suivi et de séparation des sources sonores avec des systèmes comme HARK (HRI-JP Audition for Robots with Kyoto University) [Nakadai *et al.*, 2010] et ManyEars [Grondin *et al.*, 2013]. L'ajout de ces capacités sensorielles sur un robot de téléprésence requiert de transmettre ces informations supplémentaires d'une quelconque manière à l'opérateur. En supposant que la

téléopération est effectuée avec un ordinateur de bureau standard sans utilisation d'appareils spécifiques comme des contrôleurs haptiques [Bar-Cohen, 2003], l'objectif de ce projet est d'étudier l'utilisation de représentations visuelles des sources sonores et des forces d'interactions dans une interface écologique de téléopération. Plus particulièrement, l'idée de base est d'expérimenter l'utilisation de bulles de bande dessinée et d'anneaux bleus pour l'affichage des sources sonores. Les forces appliquées sur les bras sont affichées par changement de couleur et de taille du modèle 3D du robot ainsi que par l'affichage de jauges de forces et de flèches. Le champs de vision restreint de la caméra Kinect embarquée sur le robot est augmenté par l'affichage d'une représentation 3D de l'environnement matérialisé par un nuage de point 3D sous échantillonné et construit dynamiquement pendant les déplacements du robot. La validation de ces nouvelles modalités a été faite par des utilisateurs non familiers avec l'utilisation du système de téléopération et en utilisant IRL-1 [Ferland *et al.*, 2012], un robot humanoïde muni d'actionneurs compliants (contrôlés en couple [Ferland *et al.*, 2013]) afin de capter les forces appliquées sur lui, et disposant du système de localisation de sources sonores ManyEars.

Le présent document est organisé de la façon suivante. Le chapitre 2 présente la revue de littérature concernant les approches utilisées pour la conception d'interfaces pour robots mobiles. Les méthodes et les bonnes pratiques de conception pour favoriser une bonne conscience de la situation tout en minimisant la charge cognitive y sont décrites. Des techniques de superposition d'informations sur le flux vidéo sont ensuite présentées. Le chapitre 3 correspond à un article de journal, soumis à *Journal of Human-Robot Interaction*, qui décrit l'interface développée, le déroulement des expériences et les résultats obtenus lors de l'utilisation de la représentation visuelle des sources sonores et des forces à travers une interface 3D. Enfin, le chapitre 4 explique les expériences réalisées pour l'évaluation de l'ajout d'une représentation 3D de l'environnement sur la navigation. Les résultats obtenus n'ont pas été présentés dans l'article car aucune différence significative n'a été observée dans l'exécution des tâches de navigation avec et sans l'ajout de cette modalité.

CHAPITRE 2

INTERFACES DE TÉLÉOPERATION POUR ROBOTS MOBILES

Cette section expose les notions couramment utilisées dans la conception d'interfaces de téléopération pour les robots mobiles. Les recherches déjà effectuées dans ce domaine donnent des indications sur les méthodes de conception d'interfaces qui permettent à l'utilisateur d'avoir une bonne conscience de la situation tout en limitant sa charge cognitive lors de la téléopération. Des techniques de présentation intuitive et de fusion de l'information par superposition d'informations sur le flux vidéo sont aussi étudiées.

2.1 Standards de conception pour les interfaces de téléopération

De nombreuses recherches ont été effectuées sur les interfaces de téléopération de robots mobiles, surtout dans les domaines de recherche et sauvetage urbain et dans la téléprésence à domicile. Dans ces deux domaines, l'interface joue un rôle prépondérant dans la qualité des opérations effectuées. Elle doit être conçue de manière à ce que l'utilisateur ait la meilleure perception possible de l'environnement lointain orientée vers l'efficacité des gestes et des mouvements, ainsi que la sécurité des personnes, des lieux et du robot.

Chen *et al.* [Chen *et al.*, 2007] ont recensé les facteurs critiques qui affectent le plus négativement la perception à distance, tels que : champ de vision restreint, orientation et attitude du robot, présence de multiples vues de caméras pouvant provoquer des difficultés pour l'opérateur à maintenir son attention, angle de vue des caméras et cadre de référence, manque de perception de profondeur, taux d'affichage d'images vidéo (moins de 10 Hz), temps de latence de l'affichage, mouvement du robot pouvant provoquer la dégradation de la précision et de la latence d'affichage.

Keyes *et al.* [Keyes *et al.*, 2010] ont dernièrement essayé de concevoir une interface centrée sur la vidéo en rappelant les lignes directrices déjà énoncées par divers groupes de recherche [Scholtz *et al.*, 2004a; Yanco *et al.*, 2004]. Une interface conçue pour des applications de type recherche et sauvetage urbain devrait inclure les fonctionnalités suivantes :

4 CHAPITRE 2. INTERFACES DE TÉLÉOPERATION POUR ROBOTS MOBILES

- Disposer d'une carte pour savoir où le robot est situé dans son environnement.
- Fusionner les informations pour réduire la charge cognitive de l'utilisateur.
- Minimiser le nombre de fenêtres.
- Donner des informations spatiales sur le robot dans son environnement.
- Aider à la décision du mode d'autonomie le plus adapté.
- Disposer d'une fenêtre de référence pour déterminer la position relative du robot dans son environnement ainsi que permettre l'inspection des dommages et des obstacles enchevêtrés.
- Indiquer les informations sur l'état du robot (état, caméra utilisée, position de la/les caméras, informations de tangage « *pitch* » et de roulis « *roll* »).

Ils se basent aussi sur les indications de Nielsen [Nielsen, 1993] sur la conception d'interfaces utilisateurs tout en les adaptant à la robotique :

- Assurer la cohérence entre ce qui est affiché sur l'interface et l'action que le robot est en train d'effectuer.
- Donner des retours d'informations (« *feedbacks* »).
- Utiliser un design simple et clair.
- S'assurer que l'interface aide à prévenir et corriger les erreurs engendrées par l'utilisateur ou le robot.
- Suivre les conventions du monde réel pour la présentation des messages d'erreurs.
- Permettre des actions réversibles de la part de l'utilisateur.
- S'assurer que l'interface montre bien à l'utilisateur quelles actions sont disponibles.
- Assurer un contrôle efficace.

Enfin, l'interface a été réalisée dans le but de maintenir la conscience de la situation de l'opérateur dans cinq dimensions :

- L'opérateur doit garder une bonne connaissance de la position du robot dans son environnement global.
- L'opérateur doit avoir connaissance des activités courantes du robot, c'est-à-dire, de son état.

- L'utilisateur doit garder une bonne compréhension de l'environnement immédiat du robot.
- L'opérateur doit avoir un aperçu de la mission générale à réaliser et on doit lui donner un « retour en temps réel » de son avancée dans la mission globale.
- Les informations importantes doivent être affichées le plus près possible de la vue vidéo.
- L'interface doit inclure un système automatique d'inversion de direction dans le cas où le robot dispose d'une caméra à l'arrière.

Il semble très difficile de concevoir en une seule fois une interface qui respecterait tous ces standards. C'est pour cela que Keyes *et al.* [Keyes *et al.*, 2010] ont décidé de réaliser une interface en suivant un processus itératif et en faisant tester l'interface à chaque itération par un groupe d'utilisateurs.

2.2 Approches par interfaces écologiques

L'ajout d'informations provenant des différents capteurs du robot peut créer une surcharge d'information ou distraire l'utilisateur lorsqu'on augmente le nombre de fenêtres d'affichage, ce qui n'est pas souhaitable conformément aux standards de conception listés à la section 2.1. C'est pour cela que Nielsen *et al.* [Nielsen *et al.*, 2007] ont proposé une alternative aux interfaces centrées sur la vidéo en introduisant le paradigme d'interface écologique. Une interface écologique fusionne différentes informations provenant du robot en créant un environnement 3D virtuel augmenté par la vue vidéo. Cette représentation de l'environnement lointain en réalité mixte combine la vidéo, la carte et le robot de manière intégrée dans une seule vue. Le but de leur interface est de présenter les informations d'une manière intuitive à l'utilisateur. Elle produit à l'utilisateur l'impression de profondeur en affichant le robot virtuellement dans son environnement, parmi les obstacles détectés par les différents capteurs. La vidéo en 2D est affichée dans la même fenêtre. Le robot se déplace en même temps que la vue, et l'utilisateur possède toujours un point de vue exocentrique par rapport au modèle du robot. Par exemple, la figure 2.1 montre deux interfaces qui présentent les mêmes informations de deux manières différentes : l'image en a) présente l'information de chaque capteur dans des fenêtres séparées alors que l'image en b) la présente de manière fusionnée dans une interface écologique. Les obstacles et le robot y sont représentés par des cylindres. L'étude comparative entre ces deux types d'interfaces a montré que les participants préféraient l'interface de type écologique. De plus, il a été

remarqué qu'il y avait moins de collisions accidentelles parce que l'opérateur possédait une meilleure conscience de l'environnement à proximité du robot [Nielsen *et al.*, 2007].

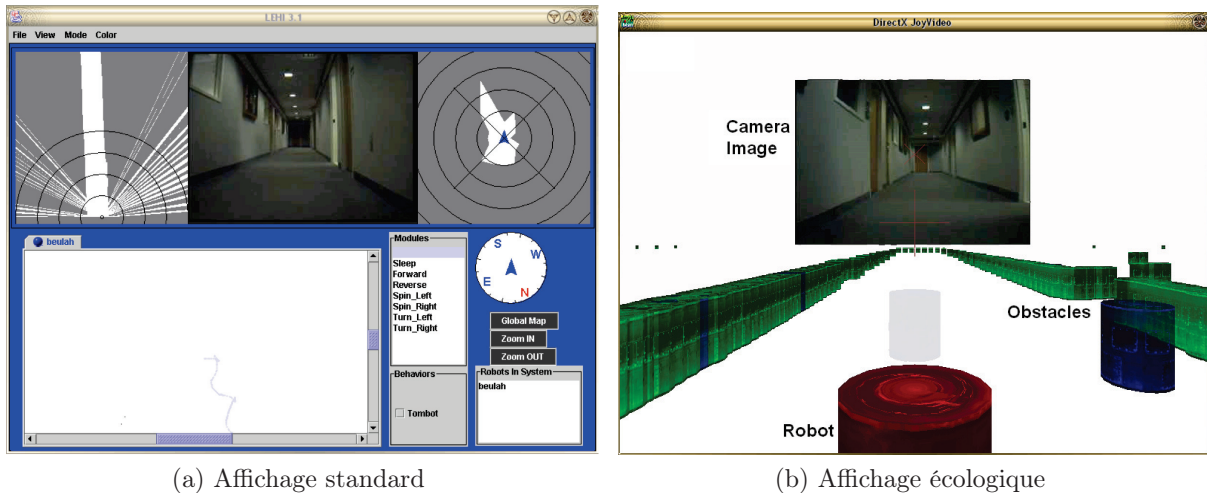


Figure 2.1 Vue similaire affichée avec deux interfaces différentes [Ricks *et al.*, 2004]

Par ailleurs, Labonté [Labonté, 2008] a réalisé une étude comparant trois types d'interfaces de téléopération pour robots de téléprésence. Le but de cette étude est d'identifier le type d'interface permettant de réaliser des opérations de navigation sécuritaires avec des opérateurs novices. Pour cela, une comparaison des résultats entre un opérateur expert et 37 opérateurs novices a été menée. Le premier type d'interface étudiée, montré à la figure 2.2, est une interface à modalité de visualisation vidéo centrée 2D (VC2D). Cet affichage offre de meilleures possibilités de visualisation de l'activité locale ainsi que de l'environnement immédiat, et s'est révélé plus performant avec les hommes de moins de 30 ans et les personnes qui travaillent plus de 22 heures par semaine sur ordinateur. La deuxième interface à modalité réalité augmentée 3D (RA3D), correspondant à la figure 2.3, ajoute une représentation 3D de l'environnement à la vue vidéo, ce qui élargit le champ de vision. Cette modalité s'est révélée meilleure en termes de facilité d'utilisation et de précision des déplacements. Le principal inconvénient du modèle 3D est qu'il doit être construit à l'avance et qu'il ne prend pas en compte les obstacles dynamiques. Enfin, la dernière interface à modalité mixte exocentrique 3D (ME3D), montrée à la figure 2.4, combine une vue principale exocentrique du modèle 3D de l'environnement et une vue secondaire égocentrique des images vidéos de la caméra du robot. Les tests ont montré que ce dernier type d'interface est plus performant en termes d'efficacité et diminue la charge cognitive sur les novices. Il assure une meilleure conscience de la localisation car il permet une localisation globale du robot dans l'environnement. Une meilleure performance avec les

femmes, les plus de 30 ans et les personnes qui travaillent moins de 22 heures par semaine sur ordinateur a aussi été observée. Globalement, les performances sont améliorées avec les modalités RA3D et ME3D avec les novices par rapport à la modalité VC2D. L'angle de vue limité semble donc être un problème pour les opérateurs novices puisque l'intégration de la vue vidéo avec un modèle 3D de l'environnement améliore substantiellement les performances de téléopération chez les opérateurs novices.



Figure 2.2 Interface à modalité vidéo centrée 2D (VC2D) [Labonté, 2008]



Figure 2.3 Interface à modalité réalité augmentée 3D (RA3D) [Labonté, 2008]

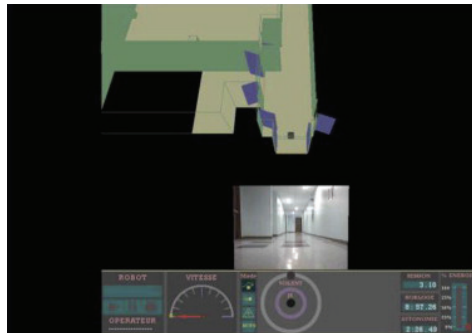


Figure 2.4 Interface à modalité mixte exocentrique 3D (ME3D) [Labonté, 2008]

Suite à cette étude, Ferland *et al.* [Ferland *et al.*, 2009] ont réalisé une interface de type écologique, mais en insérant un modèle virtuel 3D fidèle du robot ainsi que la possibilité pour l'utilisateur de changer de vue selon les tâches qu'il devait réaliser. Des informations d'un capteur de proximité laser et la vidéo 3D y sont projetées. La carte 2D construite avec les lectures du capteur laser augmente la vue vidéo. Cette interface est représentée à la figure 2.5. Les observations ont montré qu'il était particulièrement utile pour les utilisateurs de pouvoir changer de vue et de passer à une vue exocentrique, en particulier pour connaître l'environnement proche du robot et pour l'évitement d'obstacles.

Enfin, plus récemment, avec la démocratisation des caméras 3D à bas coûts comme la Kinect, il est devenu possible de modéliser l'environnement grâce à un nuage de point

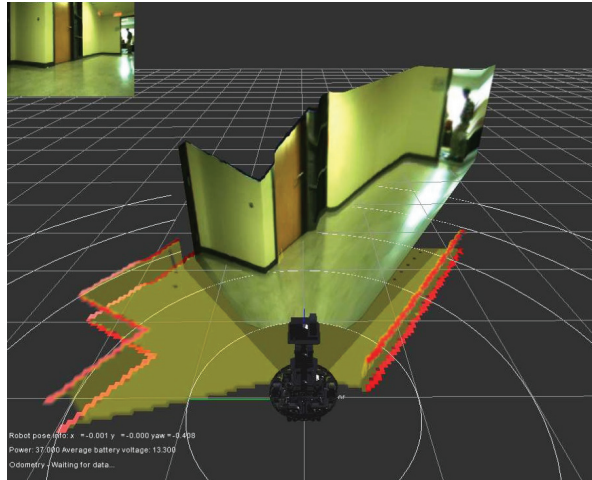


Figure 2.5 Interface graphique tridimensionnelle développée par Ferland *et al.* [Ferland *et al.*, 2009]

3D [Woodall et Bevly, 2012]. L'utilisation de ce type de carte s'est révélé utile pour la recherche d'objet et pour la navigation autour d'obstacles situés en hauteur [Mast *et al.*, 2013]. Dans ces deux cas, la carte 3D augmente la conscience de la situation de l'opérateur tout en diminuant sa charge cognitive.

2.3 Interfaces de téléopération et réalité augmentée

Milgram *et al.* [Milgram *et al.*, 1993] ont développé un système pour la téléopération de bras robotiques nommé ARGOS (« *Augmented Reality through Graphic Overlays on Stereovideo* »). Ce projet avait pour but de créer une interface permettant d'améliorer les jugements perceptuels de l'opérateur par rapport au site distant. L'interface intègre une vue vidéo stéréoscopique, ce qui permet à l'opérateur de disposer des informations de profondeur. De plus, l'interface offre la possibilité d'afficher des informations superposées au flux vidéo par réalité augmentée. Parmi elles figurent un pointeur virtuel, un mètre à ruban virtuel, une attache virtuelle et la possibilité de superposer un modèle virtuel 3D de l'objet sur le véritable objet de la vidéo. Avec ces indicateurs supplémentaires, l'opérateur peut connaître plus facilement la localisation réelle de l'objet et ainsi diminuer ses erreurs de jugement des mesures. Le but d'ARGOS est d'augmenter la facilité d'utilisation et la précision du contrôle d'un bras robotique à distance. La possibilité d'ajouter des consignes particulières directement sur l'outil en cas de contact imminent permet d'améliorer la sécurité et de décroître l'attention portée par l'utilisateur sur la tâche à effectuer. Ainsi, les tests effectués avec l'attache virtuelle, entre l'outil et le point d'arrivée, ont montré une diminution de l'erreur de placement de l'outil de 39% par rapport à la même opération

sans attache virtuelle. Dans ce cas, l'utilisation de la réalité augmentée apparaît donc comme un soutien pour l'opérateur distant.

Plus récemment, Green *et al.* [Green *et al.*, 2008] ont étudié l'impact de l'utilisation de la réalité augmentée sur la collaboration humain-robot. L'usage de la réalité augmentée peut permettre aussi bien une collaboration locale entre l'humain et le robot qu'une collaboration à distance. L'usage de la réalité augmentée peut améliorer la collaboration puisqu'elle permet de véhiculer des repères visuels et d'améliorer la communication. L'humain peut ainsi disposer d'une meilleure compréhension des actions du robot et de ses intentions. La réalité augmentée peut aussi améliorer les perspectives de collaboration entre de multiples plateformes robotiques et de multiples humains.

Dans le domaine de la robotique mobile, Lera *et al.* [Lera *et al.*, 2011] ont conçu une interface de téléopération pour des opérateurs novices. Pour cela, ils ont essayé d'améliorer leur interface par affichage d'icônes par réalité augmentée. Des marqueurs sont disposés dans l'espace et des flèches indiquant la direction dans laquelle l'utilisateur doit diriger le robot sont affichées dynamiquement. Cette interface est présentée à la figure 2.6. Des tests de comparaison de l'interface avec et sans réalité augmentée ont été menés. Les opérateurs disposaient d'une carte de l'environnement avec le point d'arrivée. Dans un cas, ils devaient se déplacer uniquement avec la carte et la vidéo alors que dans le second cas, ils étaient guidés par des flèches. Le temps de complétion de la tâche s'est révélé plus faible avec la seconde modalité. Dans ce cas, l'ajout d'icônes sur le flux vidéo est perçu comme un avantage pour les opérateurs novices, surtout dans la connaissance de la mission globale et pour la localisation dans l'environnement.

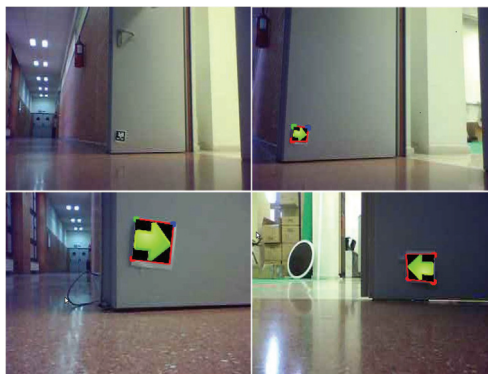


Figure 2.6 Interface de téléopération avec affichage des directions par réalité augmentée [Lera *et al.*, 2011]

Enfin, Vozar et Tilbury [Vozar et Tilbury, 2011] ont développé une interface de téléopération à réalité augmentée pour un robot mobile disposant d'un bras manipulateur. Cette

interface comprend deux fenêtres : une avec la vue vidéo égocentrique disposant d'un réticule superposé à la vidéo, et l'autre disposant d'une vue virtuelle exocentrique du robot et de la scène. Le réticule peut être utilisé pour indiquer des relations géométriques, manipuler le bras ou encore conduire le robot. Quand le réticule est placé sur un objet, il passe du noir au vert si l'objet est atteignable par le bras, ou au rouge sinon. De plus, l'opérateur peut disposer d'une boîte de dialogue superposée, lui permettant de connaître les distances séparant le châssis du robot d'un objet. Un bras virtuel peut être superposé sur les deux vues afin de prévisualiser la trajectoire du bras avant d'effectuer les opérations avec la plateforme physique. Cette prévisualisation permet d'alerter l'opérateur sur les éventuels problèmes et de garantir la sécurité des manipulations. Cette interface n'a pas fait l'objet d'une évaluation par des opérateurs pour déterminer son efficacité dans des tâches réelles de manipulations. Cependant, il a été montré que, pour des opérations de manipulations et d'inspections, l'opérateur doit nécessairement disposer d'informations de distance et d'orientation pour réaliser des opérations plus précises et plus sécuritaires sans pour autant augmenter le temps d'exécution de la tâche [Witus *et al.*, 2007]. Ces informations peuvent être présentées en utilisant diverses méthodes, comme de la réalité augmentée.

2.4 Métriques pour les interactions humain-robot

Steinfeld *et al.* [Steinfeld *et al.*, 2006] ont essayé de formaliser les métriques communes pour évaluer la qualité des interactions humain-robot. Les auteurs dénoncent la difficulté d'établir des métriques généralisables, puisqu'il existe une grande variété d'applications humain-robot. Ils ont choisi de séparer ces interactions en trois aspects : l'humain, le robot et le système. Dès lors qu'une action est effectuée par une personne en collaboration avec un robot, les performances de téléopération ne dépendent pas seulement de l'interface, mais aussi des capacités de l'opérateur et de la plateforme mobile. La qualité de la communication entre l'interface et le robot et le temps de réponse de la plateforme aux commandes sont par exemple des éléments critiques. Il est donc difficile d'évaluer un élément en particulier et ainsi de connaître l'impact exact de l'interface utilisateur uniquement sur les performances de téléopération.

Pour évaluer la qualité d'une interface de téléopération, il est nécessaire de mesurer son impact sur la conscience de la situation et la charge cognitive de l'utilisateur. L'interface doit maximiser la conscience de la situation de l'opérateur tout en minimisant sa charge cognitive, c'est-à-dire le travail mental qu'il effectue lors des opérations de contrôle du dispositif mobile distant. De plus, lorsque l'évaluation porte sur des participants humains, elle

doit porter autant sur leurs performances que sur des mesures de leur attitude [Takayama, 2009].

De nombreux auteurs évaluent la qualité de leur interface en comparant les performances des utilisateurs en fonction des nouvelles fonctionnalités intégrées. Labonté *et al.* [Labonté *et al.*, 2010] introduisent la comparaison de performances d'un opérateur expert versus des opérateurs novices afin de faire ressortir les limites de performances pouvant être attendues avec un système donné de téléopération. Les corrélations entre les performances et les fonctionnalités sont quantifiées par des études statistiques. Ces métriques concernent souvent le temps de complétion de la tâche, le nombre de collisions entre le robot et l'environnement, le nombre d'objets trouvés, la qualité de la carte dessinée par l'opérateur après les opérations [de Barros *et al.*, 2011; Keyes *et al.*, 2010; Yanco *et al.*, 2006], et aussi la distance et le nombre de commandes faites par l'opérateur pour effectuer des tâches [Labonté, 2008].

2.4.1 Conscience de la situation

La notion de conscience de la situation (« *Situation Awareness* ») pour caractériser les performances de l'utilisateur a été introduite par Endsley pour le domaine de l'aviation [Endsley, 1988]. Elle était définie comme ceci : « La conscience de la situation est la perception des éléments de l'environnement dans un volume de temps et d'espace, ainsi que la compréhension de leur signification et la projection de leur statut dans un avenir proche. »

Trois niveaux de conscience de la situation ont été établis :

- Perception des éléments dans la situation courante (ex. : affichage d'un témoin d'alerte lumineux à l'écran).
- Compréhension de la situation courante (ex. : l'apparition de ce témoin informe l'utilisateur d'une collision imminente).
- Projection dans le futur (ex. : les manœuvres qu'il faut effectuer pour éviter cette collision).

Cette notion a été ensuite adaptée au domaine de la robotique de recherche et sauvetage urbain et défini comme la « perception de la position, de l'environnement proche et de l'état du robot, la compréhension de leur signification ainsi que la projection dans un avenir proche du comportement du robot » [Scholtz *et al.*, 2004b]. Elle peut se décomposer en cinq catégories pour définir les rapports humain-robot [Drury *et al.*, 2003] :

Humain-robot : La connaissance que les humains possèdent sur la position, l'identité du robot piloté, les activités, le statut et l'environnement des robots. Aussi, l'exactitude avec laquelle les humains connaissent ces informations.

Humain-humain : La connaissance dont les humains disposent sur leurs positions, leurs identités et les activités de leurs collaborateurs humains.

Robot-humain : La connaissance par les robots des commandes dont les humains ont besoin pour diriger les activités ou toute autre contrainte qui nécessite d'ignorer un commandement ou de modifier le cours normal des opérations.

Robot-robot : La connaissance que les robots possèdent sur les commandes qui leurs sont attribués, le cas échéant, par d'autres robots, les plans tactiques des autres robots et la coordination robot à robot pour réallouer les tâches entre robots si nécessaire.

La conscience générale des humains de la mission globale : La connaissance des humains sur leurs objectifs généraux, les activités communes entre humains et robots et de la mesure étape par étape du progrès par rapport aux objectifs.

Une technique adaptée au domaine des secours urbains a été mise en place par Drury *et al.* [Drury *et al.*, 2007] pour l'évaluation de la conscience de la situation pour des interfaces centrées sur la carte ou bien sur la vidéo . Les participants sont d'abord soumis à un questionnaire pré-expérimental afin de connaître leurs expériences précédentes dans le domaine de la téléopération. Ensuite, ces derniers disposent de quelques minutes pour prendre en main le robot en s'exerçant dans un lieu différent de celui choisi pour les tests finaux. Ils ont pour objectif de trouver le plus de victimes possible en un temps donné, dans un lieu clos. Enfin, il leur est demandé de faire une évaluation des éléments de l'interface en remplissant un questionnaire. Durant l'expérimentation, les opérateurs sont soumis à la technique d'évaluation de leur conscience de la situation par la méthode LASSO (« *Location, Activities, Surroundings, Status and Overall mission* »). On demande aux utilisateurs de donner leurs commentaires spontanément et à voix haute. Ces commentaires sont ensuite classés selon cinq catégories de compréhension de la situation comme positif, neutre ou négatif :

- Connaissance du positionnement (« *Location Awareness* ») du robot en se basant sur des repères.
- Connaissance des tâches (« *Activity Awareness* ») que le robot est en train d'effectuer.

- Connaissance de l'environnement proche du robot (« *Surroundings Awareness* »).
- Connaissance de l'état du robot, comme son niveau de batterie (« *Status Awareness* »).
- Connaissance de la mission globale et du progrès réalisé dans l'atteinte du but (« *Overall Mission Awareness* »).

Pour résoudre le problème de classement d'un commentaire dans l'une des catégories, celui-ci est évalué pour toutes les catégories à la fois comme positif, négatif ou neutre. Le principal avantage de disposer des commentaires au fur et à mesure de l'expérience plutôt que les recueillir seulement à la fin réside dans le fait que l'utilisateur pourrait être influencé par ce qu'il a vécu seulement dans les dernières minutes. Il pourrait par exemple donner des commentaires négatifs en étant influencé par ses dernières expériences, et ce même si la téléopération s'est bien déroulée dans les premières minutes.

Il est à noter qu'il y a seulement un rapport probabiliste entre la performance et la conscience de la situation de l'opérateur. En effet, plus la conscience de la situation est bonne, plus la probabilité de réaliser une bonne performance est élevée, et inversement. Il se peut très bien que la performance soit très mauvaise alors que l'utilisateur disposait d'une bonne conscience de l'environnement lointain [Endsley, 2000].

2.4.2 Charge cognitive

En robotique, la charge cognitive imposée par les interfaces est souvent mesurée avec un questionnaire de fin de test. Un des tests les plus utilisés est un test adapté du « *NASA-TASK LOAD INDEX* ». Il se base sur une évaluation de la charge de travail selon six sous-échelles en demandant à l'utilisateur de quantifier les exigences mentales, physiques et temporelles, sa frustration ainsi que l'effort et la performance occasionnée par la tâche. Il est effectué pendant ou immédiatement à la fin des phases de tests [Hart, 2006].

Jong [Jong, 2010] a recensé trois moyens de mesures de la charge cognitive dans le domaine de l'éducation :

Mesure de la charge cognitive par auto-déclaration ou par mesure subjective :

Pour ce type de mesures, les problèmes rencontrés sont la subjectivité des questions et le nombre de tests différents qui rend souvent impossibles les comparaisons.

Mesures physiologiques sur l'individu : Des études ont essayé de quantifier la charge cognitive en mesurant les variations du rythme cardiaque seul, ou en l'associant avec

des mesures de pression sanguine, ou encore des mesures de dilatation des pupilles. Le principal inconvénient de ce type de mesure est qu'il est très intrusif pour la personne.

Mesure par tâche duale : Dans ce cas, une seconde tâche très simple et peu coûteuse par rapport à la première est demandée à l'opérateur. On considère dans cette approche que si le temps de réaction sur la deuxième tâche ou la performance de celle-ci est mauvaise, ceci est dû à une trop grande consommation des ressources venant de la première tâche.

Un exemple de tâche duale peut correspondre à l'affichage d'une lettre dans une petite fenêtre en haut de l'écran. Le temps de réaction au changement de couleur de la lettre est mesuré. L'opérateur doit appuyer le plus rapidement possible sur une touche du clavier dès le changement de couleur [Brünken *et al.*, 2003].

2.5 Retours d'informations

Plusieurs concepteurs d'interfaces s'accordent sur le fait qu'il est important de donner un retour d'informations approprié pour que l'utilisateur reste en permanence informé des actions et de l'état du robot [Lathan et Tracey, 2002; Nielsen, 1993].

De multiples moyens de retours d'informations ont été testés, qu'ils soient haptiques, visuels, sonores, ou encore la combinaison de plusieurs d'entre eux. Une méta-analyse concernant 43 études a été menée par Burke *et al.* [Burke *et al.*, 2006] pour examiner les effets des retours d'informations multimodaux sur les performances des utilisateurs. Il ressort de l'étude que l'ajout de modalités supplémentaires au retour visuel diminue le temps de réaction et augmente le score de performance de l'utilisateur. Le problème d'une méta-analyse est qu'elle ne prend pas en compte le type de source sonore utilisée ou bien le type de dispositif haptique.

Il est important de fournir les informations nécessaires et exactes à l'opérateur [Desai *et al.*, 2011]. Il faut éviter de surcharger l'opérateur avec un trop grand nombre d'informations mal transmises. En effet, celles-ci sont efficaces seulement si elles ont une réelle utilité et si elles sont délivrées au moment opportun. Plusieurs exemples crédibilisent cette affirmation. Johnson *et al.* [Johnson *et al.*, 2003] ont testé l'ajout d'une égosphère autour du modèle du robot censé diminuer la charge cognitive de l'utilisateur. Les tests ont montré que ce type d'affichage augmentait la charge cognitive et le stress provoqué à l'utilisateur. Plus récemment, de Barros *et al.* [de Barros *et al.*, 2011] a obtenu de moins bons résultats avec l'affichage d'anneaux de couleurs autour du modèle du robot afin de prévenir le risque de

contact par rapport à l'utilisation de l'interface sans cette fonctionnalité. Cette contre-performance a été expliquée par le fait que les anneaux autour du robot masquaient les lignes de reconstruction de l'environnement.

2.5.1 Retours visuels d'informations de forces

Lindeman [Lindeman, 2003] a effectué des recherches dans le domaine de la représentation des interactions entre objets et humains dans les environnements virtuels. Pour donner à l'utilisateur un retour intuitif et efficace, des formes de retours visuels, sonores et haptiques peuvent être utilisées. Les modalités visuelles étudiées étaient le changement de couleur binaire de l'objet, l'illumination du point de contact, l'affichage de vecteurs de force, la distorsion de texture ainsi que la distorsion de la forme.

Il est difficile de donner un aperçu de la force exercée par un humain à l'aide d'un robot ou d'un appareil sur l'environnement distant. Des possibilités de substituer les retours haptiques par des retours visuels et audios ont été étudiées sur des systèmes de robots chirurgicaux téléopérés. En effet, ce type de robot permet aux chirurgiens de disposer d'une plus grande précision, dextérité et visualisation de la tâche à accomplir. Ces systèmes permettent de donner un retour visuel de la force appliquée par le chirurgien en temps réel par représentation des forces avec des jauges superposées à la vue vidéo et dont la dimension et la couleur changent en fonction de la force mesurée [Kitagawa *et al.*, 2005; Tavakoli *et al.*, 2006]. Une autre technique consiste à superposer des cercles dont la couleur varie du vert au rouge en fonction de la force mesurée sur le dispositif réel directement sur les pinces [Reiley *et al.*, 2008].

2.5.2 Retours visuels d'informations de sons

L'affichage de représentations sensorielles comme les sources sonores n'est pas très développé dans la robotique. Parmi les études menées, une interface a été développée dans le domaine de la recherche et des secours urbains. Elle permet la représentation d'une source sonore par un curseur horizontal qui se dirige du côté du microphone qui capte la source sonore avec la plus grande amplitude [Hestand et Yanco, 2004]. La figure 2.7 montre cette interface. Cette représentation peut ainsi permettre à l'opérateur de s'orienter dans la direction d'un appel au secours qui serait émis par une victime. Cependant, avec ce type d'affichage, il n'est pas possible de connaître ni la provenance, ni l'intensité du son émis. De plus, si aucune distinction n'est faite entre le bruit ambiant et les bruits utiles comme les voix de personnes, ce type d'affichage possède une utilité restreinte.



Figure 2.7 Interface avec représentation des sons par curseur [Baker *et al.*, 2004]

En ce qui concerne les retours visuels de sources sonores, des travaux de recherches ont été menés à des fins d'amélioration des dispositifs pour les malentendants. Une étude menée par Matthews *et al.* [Matthews *et al.*, 2005] avait pour objectif de permettre aux personnes malentendantes d'être conscientes des bruits émis dans leur domicile. Ainsi, les informations sonores recueillies à leur domicile peuvent être présentées sur un écran disposé sur un mur ou bien sur un dispositif portatif de type PDA (« *Personal Digital Assistant* »). Les types d'informations dont les personnes malentendantes souhaitent disposer sont le type de son, la localisation du son émis, le volume du son émis et le ton de la parole.

Trois types d'affichage ont été implémentés et comparés : la représentation par icônes, la représentation par spectrophotographe avec une icône associée, et la représentation d'un historique sous forme de diagramme en bâtons avec l'information d'intensité, montré à la figure 2.8. En ce qui concerne la technique d'affichage par icônes, les sons non reconnus sont affichés sous forme d'anneaux.

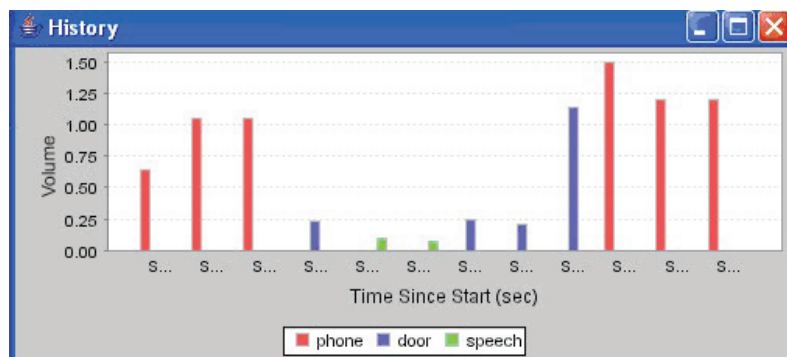


Figure 2.8 Historique des sons émis dans le temps en fonction de leur intensité et de leur contexte [Matthews *et al.*, 2005]

Il est ressorti de l'étude que les participants préfèrent des moyens d'affichages qui rendent les sons faciles à interpréter et dont ils peuvent déterminer l'intensité en un coup d'œil et avec un visuel attractif. Les éléments importants recensés par les auteurs sont :

- Identifier quel son a été émis.
- Voir un historique des sons émis.
- Pouvoir adapter l'information affichée en fonction des besoins.
- Donner un indice de confiance des informations affichées.

Toujours pour améliorer les perceptions de sources sonores pour les élèves malentendants d'une université, Azar *et al.* [Azar *et al.*, 2009] ont conçu un logiciel capable de les afficher par divers moyens : affichages spectraux, spectrographes, icônes pour le contexte, traduction en langage des signes et affichage textuel si un mot a été reconnu. Les inconvénients majeurs de ce système se trouvent être le manque d'intuitivité des types d'affichage pour l'utilisateur et l'obligation d'avoir une fenêtre de grande taille pour pouvoir les afficher, comme illustré à la figure 2.9.

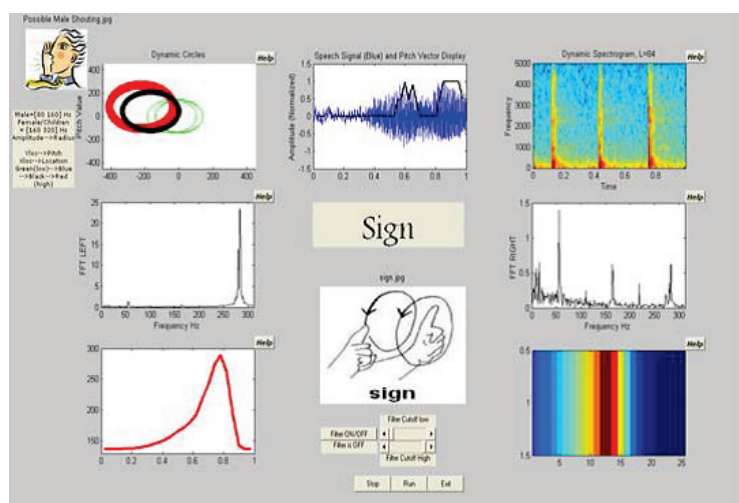


Figure 2.9 Interface de visualisation des sons pour malentendants [Azar *et al.*, 2009]

Enfin, Janicke *et al.* [Janicke *et al.*, 2010] ont conçu un système capable de donner un aperçu en temps réel de la bande sonore d'un film pour les malentendants en matérialisant les sons par des métaphores. Les métaphores utilisées sont montrées à la figure 2.10. Ces représentations sont principalement appréciées lorsque la perception de certains sons est primordiale pour la compréhension de la situation. Par exemple, si une personne entre dans la pièce et n'est pas visible à l'écran, le claquement de la porte permet au téléspectateur

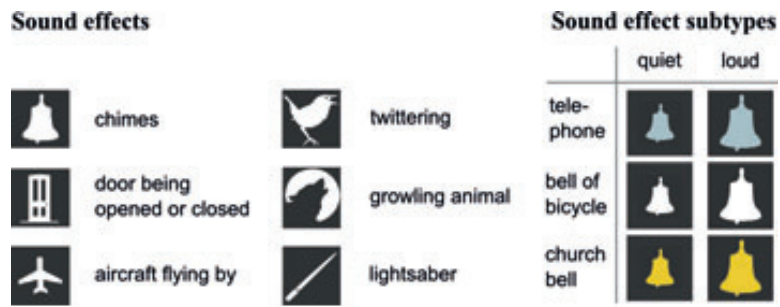


Figure 2.10 Représentation des sons par des icônes [Janicke *et al.*, 2010]

d'être informé de sa venue. Ce type de représentations pallie dans ce cas à l'absence de perception du son par la personne.

CHAPITRE 3

REPRÉSENTATION VISUELLE DES FORCES D'INTERACTION ET DES SOURCES SONORES DANS UNE INTERFACE DE TÉLÉ-OPÉRATION

Visual Representation of Interaction Forces and Sound Sources in a Teleoperation Interface for a Mobile Robot

3.1 Avant-propos

Auteurs et affiliation :

A. Reveleau : étudiant à la maîtrise, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

F. Ferland : étudiant au doctorat, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

M. Labbé : étudiant au doctorat, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

D. Létourneau : professionnel de recherche, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

F. Michaud : professeur, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

Date de soumission : 31 juillet 2014

Revue : *Journal of Human-Robot Interaction*

Titre français : Représentation visuelle des forces d'interaction et des sources sonores dans une interface de téléopération pour un robot mobile

Contribution au document :

Le chapitre 2 met en évidence l'état de l'art sur la conception des interfaces opérateurs

pour le contrôle de robots mobiles ainsi que les moyens de représentations visuelles des sources sonores et des forces exercées sur des plateformes mobiles. À partir des informations relevées, une interface graphique permettant de visualiser les modalités d'interaction du robot IRL-1/TR a été réalisée. Cet article contribue au mémoire en décrivant les fonctionnalités de l'interface développée pour la représentation visuelle des forces d'interaction et des sources sonores, l'intégration par rapport à l'architecture du robot ainsi que les expériences et les résultats menées avec 31 opérateurs pour évaluer l'interface. La version de l'article présent dans le mémoire a été modifiée à la suite des corrections demandées par les membres du jury.

Résumé français :

Les robots de téléprésence actuellement commercialisés transmettent la vidéo, l'audio et les informations de proximité aux opérateurs distants à travers une interface utilisateur graphique fonctionnant sur des ordinateurs standards de bureau. Alors que de nouvelles modalités d'interactions comme la perception des forces ressenties ou bien la localisation des sources sonores sont développées sur des plateformes robotique avancées, de nouveaux moyens d'intégration de ces informations sont nécessaires. Cet article présente l'utilisation de la représentation visuelle des forces et des sources sonores à travers une interface 3D. Les forces sont matérialisées par les changements de couleur, de taille du modèle 3D du robot et par des jauges de forces et des flèches tandis que les sources sonores sont représentées à l'aide de bulles de bande dessinée ou d'anneaux bleus. La validation de ces modalités est réalisée avec 31 participants en utilisant IRL-1/TR, une plateforme humanoïde équipée avec des actionneurs différentiels élastiques permettant la souplesse et le contrôle en force de ses bras et capable de localiser les sources sonores. Les résultats montrent que la représentation visuelle des forces et des sources sonores peut fournir des informations utiles et appropriées à la réalisation de tâches par des opérateurs distants.

3.2 Abstract

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 like force sensing and sound localization are being developed and tested on advanced robotic platforms, ways to integrate such information in a teleoperation user interface are required. This paper demonstrates the use of visual representations of forces and sound localization on a 3D teleoperation interface. Forces are represented using colors, size, bar graphs and arrows, and 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 forces and sound sources can provide useful and appropriate information to remote operators.

3.3 Introduction

Nowadays, an increasing number of commercial telepresence robotic platforms are available [Guizzo, 2010] (e.g., VGo from VGo Communications, QB from Anybot [Desai *et al.*, 2011], TiLR from RoboDynamics, Giraff from HeadThere [Gonzalez-Jimenez *et al.*, 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 video-conference systems, commonly referred to as "Skype on wheels" [Dahl et Boulos, 2013]. Their intended usage goes from teleworkers virtually present at meetings or on remote sites [Lee et Takayama, 2011], to telehealth applications in hospitals and homes [Mendez *et al.*, 2013; Michaud *et al.*, 2010].

The graphical user interface (GUI) of telepresence robots plays a central role in maximizing situation awareness [Endsley, 1988; Scholtz *et al.*, 2004b] while minimizing cognitive load [Yanco *et al.*, 2004] for operators using these systems. As perceptual and localization modalities were added to teleoperation user interfaces over the years, research demonstrated the benefits of minimizing the number of windows by directing the information on the main display [Keyes *et al.*, 2010]. Accordingly, the ecological interface paradigm uses augmented reality displays to represent the telepresence robot in virtual environments and by integrating the video stream directly in these virtual environments [Ferland *et al.*, 2009;

Labonté *et al.*, 2010; Nielsen *et al.*, 2007; Ricks *et al.*, 2004], to facilitate navigation of the platform.

Recent robots like Rollin' Justin [Borst *et al.*, 2009], PR2 [Wyrobek *et al.*, 2008], Cody [Chen et Kemp, 2011] and Care-O-bot [Graf *et al.*, 2009] bring additional interaction capabilities by being equipped with compliant manipulators. It is also now possible to provide mobile platforms with sound source localization, tracking and separation capabilities, with systems such as HARK (HRI-JP Audition for Robots with Kyoto University) [Mizumoto *et al.*, 2011; Nakadai *et al.*, 2010] and ManyEars [Grondin *et al.*, 2013]. Using such capabilities on a telepresence robot requires to communicate somehow information from these sensing capabilities to the remote operator. Assuming that the teleoperation is done 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 visual representation of interaction forces and sound sources in an ecological teleoperation interface. More specifically, the basic idea is to experiment 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 the type of sounds. Validation of these modalities is done using IRL-1/TR [Ferland *et al.*, 2012], a compliant humanoid platform capable of force interaction and sound source localization.

The paper is organized as follows. Section 3.4 situates related work on visual representation of forces and sounds. Section 3.5 describes the visual representation modalities for forces and sounds implemented using IRL-1/TR. Section 3.6 presents the experimental methodology, followed by Section 3.7 with the results.

3.4 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 et Ward, 2013], like joysticks with force feedback or data gloves. But it is also possible to provide visual cues related to touch. For instance, Lindeman [Lindeman, 2003] studied with various visual haptic feedbacks in immersive virtual environments, such as illuminating the virtual contact point with a binary color change, having the contact location and the intensity represented by a change of color or texture or by shape distortion, or rendering force intensity and direction by vector glyphs. Haptic substitution by visual cues have also been studied with teleoperated surgical robots. Contact forces are represented using a bar indicator

proportional to the force intensity, overlaid on top of the video stream [Kitagawa *et al.*, 2005; Tavakoli *et al.*, 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 forces for teleoperation of a mobile robot has not yet been studied.

According to Lombard and Ditton [Lombard et Ditton, 1997], mediated sounds are important in generating presence. In addition to hearing remote sounds through a headset, it can be beneficial to visually represent sounds on a GUI. 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 on the upper part of the video stream [Hestand et Yanco, 2004]. This provided a 1-dimensional cue of where the sounds were coming from, but no indication on the type, the intensity or the absolute location of the sound source. Mizumoto et al. [Mizumoto *et al.*, 2011] represent direction and intensity of sound sources in relation to the position of the robot, but not the localization of the sound sources on the video stream. However, sound visualization has been studied to assist the hearing-impaired. Matthews et al. [Matthews *et al.*, 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 et al. [Azar *et al.*, 2009] examined the representation of raw FFT (Fast Fourier Transform) data, spectrograms or pitch vector displays, and all were found to be difficult to understand. Finally, Janicke et al. [Janicke *et al.*, 2010] used visual metaphors (i.e., icons for chimes, door being opened or closed, aircraft flying by, twittering, growling animal, lightsaber) to illustrate in real-time the sound track of a film. These illustrations revealed to be useful for situational awareness when visual cues are absent (i.e., a person entering a room without being visible on the screen).

3.5 Visual Representations of Interaction Forces and Sound Sources Using IRL-1/TR

To explore the use of visual representations of interaction forces and sound sources, we designed an ecological teleoperation interface for IRL-1/TR. Shown in Figure 3.1, IRL-1/TR consists of a humanoid torso, IRL-1, and a mobile base, TR. IRL-1/TR is 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), a 8-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 [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. It allows low latency video encoding, ideal for telepresence. The depth image is encoded using the lossless PNG compression plugin provided by the image transport package from ROS.

The TR base, for Telerobot [Michaud *et al.*, 2007] is the wheeled platform used for the locomotion of the robot. It can translate the robot forward and backward, and rotate in place. It consists of two triplets of wheels (one on the left and one on the right) and a rear rocker arm. 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 mechanism acting like a bogie. The rear rocker arm connects the two triplets of wheels to another pivot attached to the robot's chassis. This rocker-bogie suspension minimizes the displacement of the robot's chassis when moving on uneven surfaces, helping to provide stable video feed to the user.

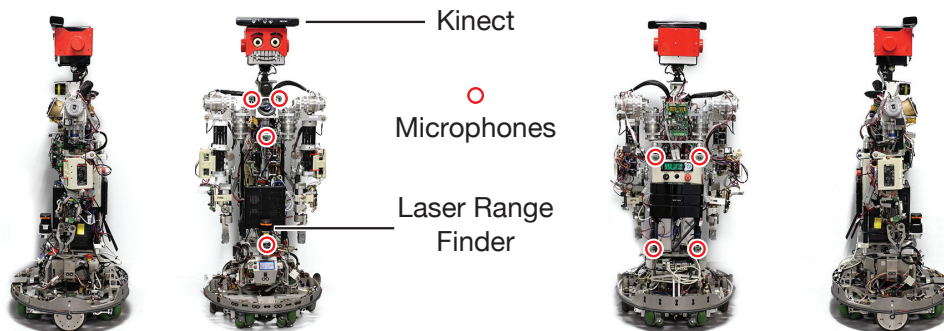


Figure 3.1 IRL-1/TR

Differential Elastic Actuators (DEAs) [Lauria *et al.*, 2008] are used to provide force control and feedback to the manipulators of IRL-1/TR. Each arm on IRL-1/TR has four degrees of freedom (DOFs) : three in the shoulder and one in the elbow, actuated using DEAs. The shoulder roll is responsible for the end effector orientation around an axis 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) [Robinson, 2000;

Williamson, 1995], which are used on the Meka M1, Cody [Jain et 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 8-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 sound sources in real-time and in reverberant environments. The algorithm locates the sound sources on 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, providing the direction of the sound sources.

Figure 3.2 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.

3.5.1 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 *et al.*, 2013]. Figure 3.3 illustrates a model the 4 DOF arm of IRL-1/TR, with θ_1 and θ_2 representing respectively the shoulder pan and tilt angles and θ_3 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. 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 are in the neutral position and the current arm position, respectively. 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 equation (3.1) :

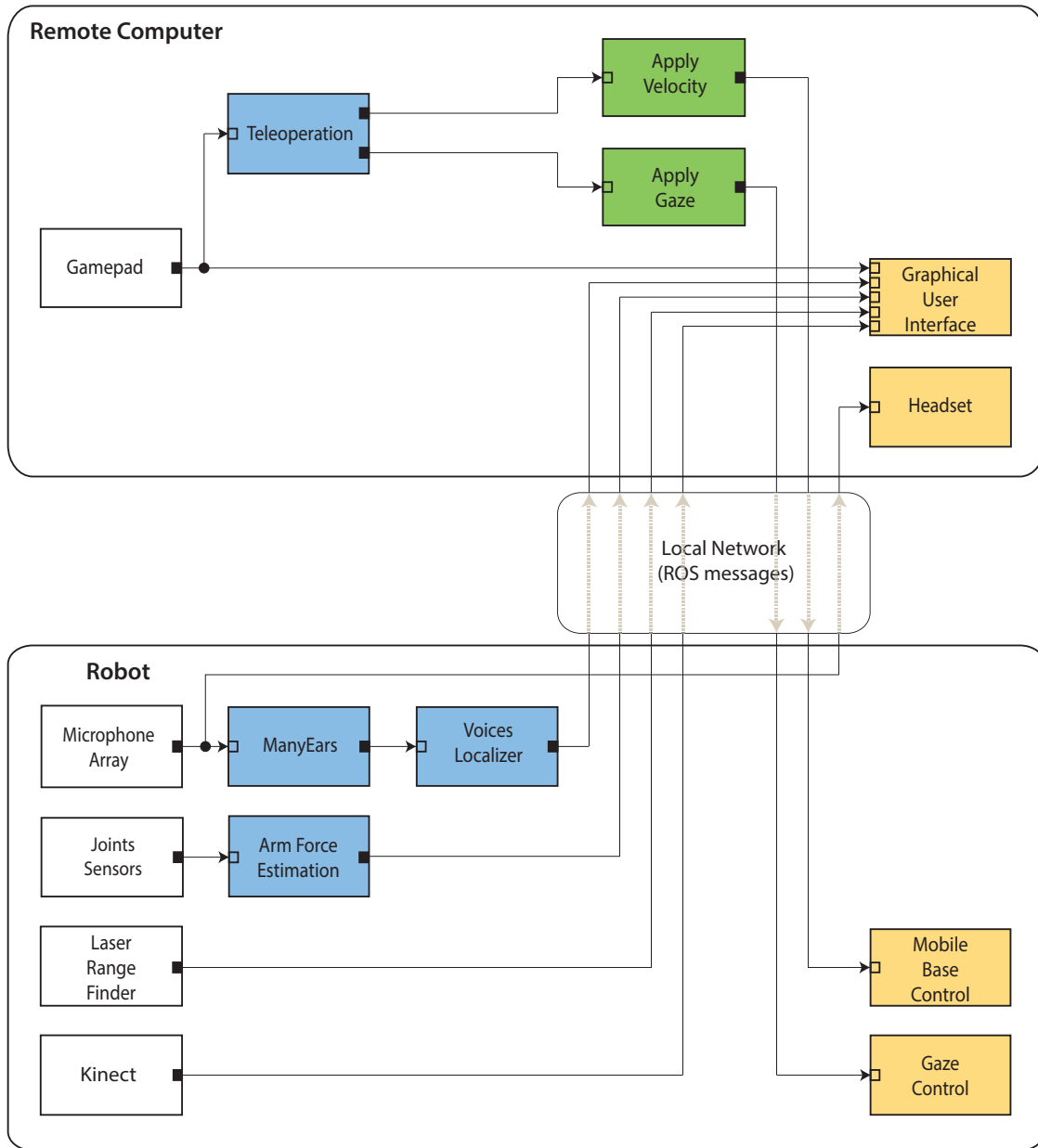


Figure 3.2 Teleoperation control architecture of IRL-1/TR. Empty squares represent inputs, and filled squares represent outputs.

$$\vec{f}_A = k_A l_A \frac{\vec{r}_{H_N H_C}}{\|\vec{r}_{H_N H_C}\|} \quad (3.1)$$

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 3.1 present 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.

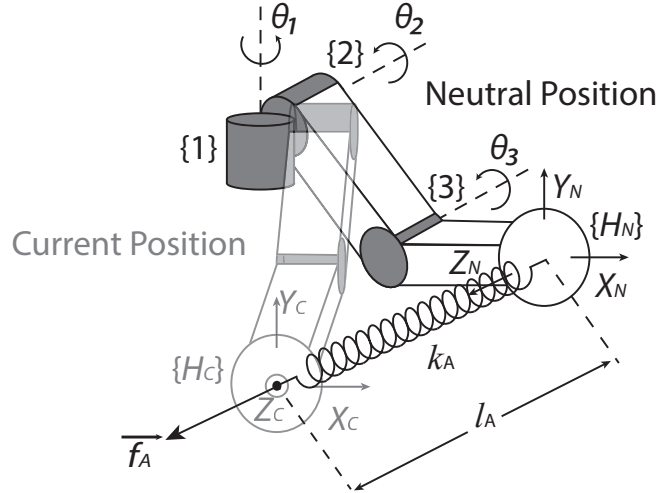


Figure 3.3 Representation of the model used for the arms of IRL-1/TR.

Tableau 3.1 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

3.5.2 Voice Localizer and ManyEars

This module is used to distinguish between speech and non-speech for the separated audio streams generated by ManyEars. The process consists of doing pitch extraction and classification using the approach described in Sasaki et al. [Sasaki *et al.*, 2009].

First, a database of known speech and non-speech sounds was created, 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. [Sasaki *et al.*, 2009] recommend to use $K = 26$

clusters. Using this database, input audio frames to classify as either speech or non-speech are segmented, binarized and categorized according to their probability of belonging to none or one of these clusters. The same approach could also be used to identify additional daily sounds such as water running, key jingling or door slamming, if this would be required by the system.

3.5.3 Teleoperation

Teleoperation of IRL-1/TR is done using the wireless gamepad shown in Figure 3.4. 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 is 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 trials with participants. 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 3.4 Gamepad used to control IRL-1/TR.

3.5.4 Graphical User Interface

As shown in Figure 3.5, the basis of our teleoperation user interface consists 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 *et al.*, 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, using the laser range finder stat to detect what is not visible. Keeping this viewpoint constant also remove a variable not critical for our study and that could influence the observations during the trials. The point cloud generated by and visible from the Kinect camera is shown in the 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 $3\text{ m} \times 3\text{ m}$ square grid fixed to the world frame corresponding to the robot booting position, and a circular grid with 1.5 m increment and fixed to the mobile base are used to represent the floor and to simulate motion by moving the circular grid in relation to the square grid. Laser range readings are displayed as points with colors, from green ($> 2.50\text{ m}$) to yellow ($\leq 2.50\text{ m}$) and red ($\leq 1.50\text{ m}$), based on the distance of the perceived objects relative to the robot current position. The 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 the danger perceived. This revealed to be an essential feature to communicate proximity information in such situations.

The visual cues used to represent the interaction forces perceived on the arms are illustrated by Figure 3.6 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 equation (3.2). 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 being perceived, the arm is represented in grey.

$$\begin{cases} R = \frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \\ G = 1 - R \\ B = 0 \end{cases} \quad (3.2)$$

The size of the arm is scaled from regular (low intensity) to large (high intensity) size according to equation (3.3) :

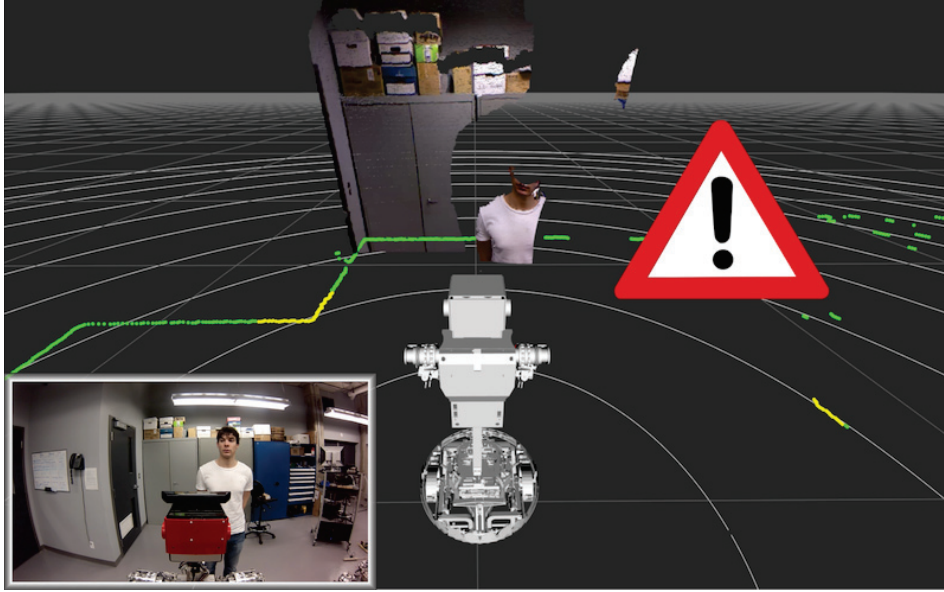


Figure 3.5 General view of the teleoperation user interface with laser range finder data, the image from the Kinect and the navigation warning icon. The external view of the scene is presented on the bottom left portion of the image for reference, and is not part of the GUI.

$$scale = 1 + \left(\frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \right) \quad (3.3)$$

- 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 reference frame. The bar is rendered by a colored cylinder with constant radius of 0.1 m and a length l_b computed by equation (3.4) to reach 1 m at the maximum authorized force. The top bar extremity color components are computed as in equation (3.2). Bar color is interpolated from green to the computed extremity color.

$$l_b = \frac{1}{f_{max}} \times \min(\|\vec{f}_A\|, f_{max}) \quad (3.4)$$

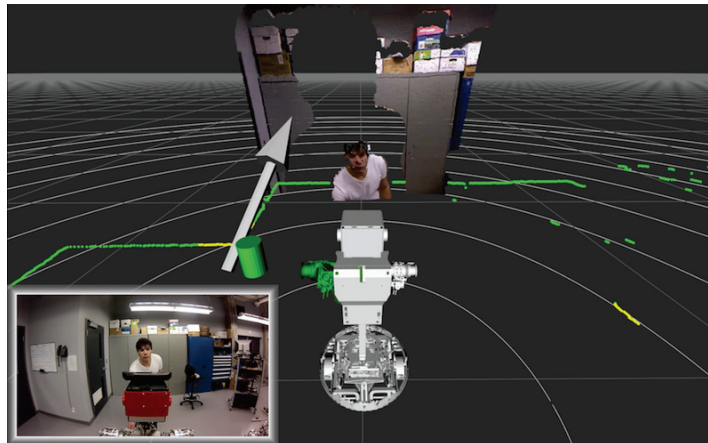
- Arrow. A grey arrow located next to the intensity bar represents the direction of the force sensed on the arm. The arrow is a unit vector oriented in the direction of the force applied on the end effector. The origin is placed 0.5 m next to the associated shoulder joint to be visible to the operator. Arrow length is constant and set to be

1 m. By looking at the arrow, the operator can see the direction of the force applied on the end effector in relation to its neutral position.

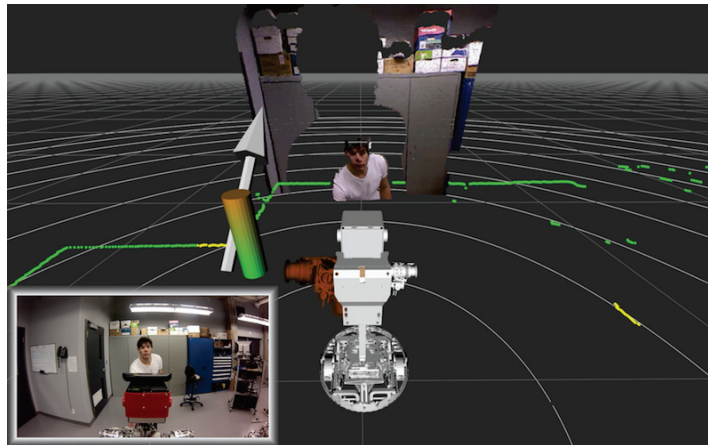
For sound visualization, we chose to experiment with two icons to represent sound source location and type, as shown by Figure 3.7 : a blue ring, representing the location of an unidentified type or non-speech sounds ; a speech bubble, to localize where speech is identified. Since it takes at least 100 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 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 sources outside the field of view of the Kinect are also displayed on a 1 m unit sphere around IRL-1/TR.

3.6 Experimental Methodology

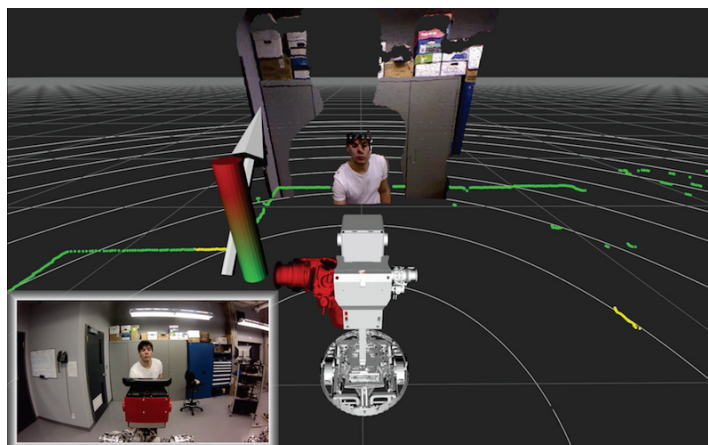
The underlying objectives of our study is to explore how visual representations of sound sources and interaction forces can provide useful information for robot teleoperation. Initially, our intent was to simultaneously use the sound and force visualization modalities in realistic HRI scenarios, comparing trials with and without the visual representation modalities. For instance, we envisioned teleoperating IRL-1/TR to find somebody located in another room, having IRL-1/TR open the door using its compliant arms. Door opening however requires extensive trained to safely manipulate the arms. Also, finding the desired person would be influenced by how IRL-1/TR would entered the room. Another scenario involved having IRL-1/TR come behind a group of seating people and identify a person to help stand up, but having repeatable force interaction with the person is difficult. Moreover, with an integrated scenario, it would be hard to find metrics that evaluate the influence of each modality. Therefore, considering that it would be difficult to measure the impact of combined modalities on the participant situation 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 forces and sound sources individually.



(a) Low intensity

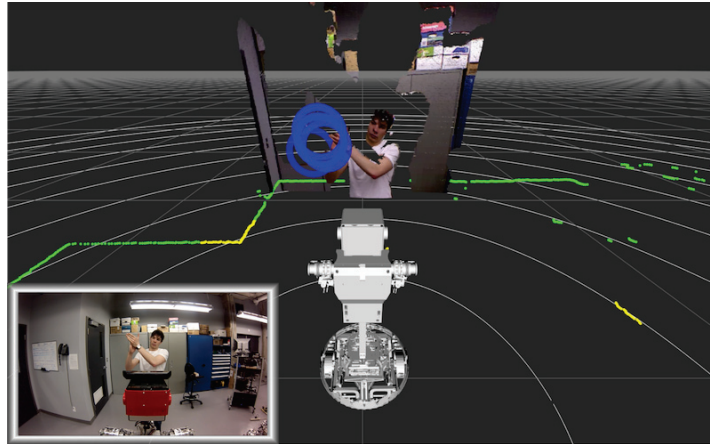


(b) Medium Intensity

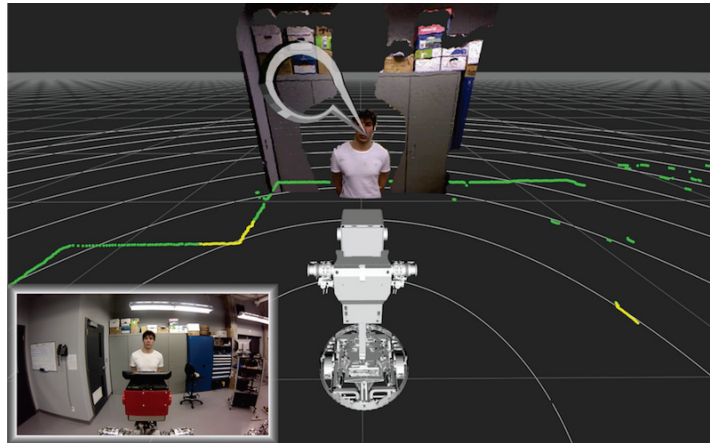


(c) High intensity

Figure 3.6 Display of interaction forces applied on the left arm of IRL-1/TR when someone is pulling its end effector.



(a) Five hand clapping sounds



(b) Speech

Figure 3.7 Sound visualization

3.6.1 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 achieved better and more precisely the task using force visualization. Figure 3.8 illustrates the experimental setup. The vertical pole was 1.25 m in height and 0.15 m in diameter, weighted 9 kg and fell when tilt was greater than 10° . This task was accomplished without sound perception or visualization.

To make the arm of IRL-1/TR 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. The head of IRL-1/TR 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 forces on the arms of IRL-1/TR. Each participant repeated the task ten times, five with and five without force visualization, randomly sampled. For every 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 forward IRL-1/TR and stop when he or she believed that the U-shaped attachment was touching the vertical pole. Time taken to complete the task 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. If the pole fell, the participant was able to see it through the GUI. 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.

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

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

3.6.2 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 sources around IRL-1/TR. 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. Figure 3.9 illustrates the experimental setup. Identification (ID) numbers were added to the display of the teleoperation user interface to identify the sound sources. Six sound tracks of male speech, each lasting 6 sec, were played on a randomly selected loudspeaker, with 2 sec of silence between each. The volume was empirically set to allow ManyEars to correctly localize sound sources. To avoid having to recreate the same experimental scenario for each participant, we pre-recorded trials with the following two conditions :

1. No sound visualization, and sound visualization with the persistence parameter p set to 2 sec. The trial started by presenting either the no sound or sound visualization

sequence of six speech utterances, chosen randomly. This sequence of 12 speech utterances was used to evaluate the ability of the participant to localize where the six sound sources are located over time, without and with sound visualization. p was set to be equal to the silence time. Participants were asked to write down the sound source ID numbers for this sequence.

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 loudspeaker to be randomly activated 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, they 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 ?

3.6.3 Test Population, Familiarization and Settings

Our test population consisted of a convenience sample of 31 volunteers (26 males and 5 females) aged from 21 to 69 and 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, including all the visualization modalities of the GUI. 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 made sounds around the robot to see how these were displayed on the GUI. This step lasted approximately 5 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 first, and then the sound visualization trials. It took about 1 hour per participant to complete the trials and to fill out the questionnaires.

The GUI ran on a standard desktop computer with a 3.40 GHz 4th generation Core i7 quad core processor with 8 GB of RAM and 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 *et al.*, 2011] and sent to the headset connected to the remote computer. This made it possible for the operator to get stereo inputs from the robot.

3.7 Results

From the participants' performance in the various trials and their answers to the questionnaire, our objective was to verify the following hypotheses :

- H1 : The participants make contact with the pole more often with force visualization.
- H2 : The participants make the pole fall less often with force visualization.
- H3 : Sound icons help participants to localize sound sources around IRL-1/TR.

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.

3.7.1 Force Visualization

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

Table 3.2 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 is observed ($t = 5.62, dl = 26, p < 0.001$). The average tilt angle is 2.4° ($\sigma = 1.9$) without force visualization, and 1.8° ($\sigma = 1.2$) with force visualization. Significant difference between the two conditions is 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 and the precision expressed in terms of the tilt angle, but not H2 related to the number of times the participant made the vertical pole fall.

Tableau 3.2 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

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 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 to be annoyed 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.

3.7.2 Sound Visualization

For Condition 1, Table 3.3 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 left or right sound sources, the error rate without sound visualization decreases to 1.2/6 (20%), with $\sigma = 1.62/6$ (27%). 30 out of the 31 participants indicated that they were able to distinguish between right and left sources just from the headphone, but revealed to have difficulty when it came to distinguish loudspeakers from the same side. Humans can localize sound sources 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 at the end of the trial to list the sound sources ID number, and forgot one answer. Significant difference ($t = 8.076, dl = 30, p < 0.001$) is observed between error rates without and with sound visualization, validating H3.

In relation to Q2, only one participant admitted to have answered randomly. Regarding Q3, 30 out 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 indicated to only focus on the visual display and not on the audio stream : the participant mentioned that it overloaded unnecessarily the teleoperation interface, and would use this feature only when necessary. However, without the use of sound visualization, this participant made two errors out of the six conditions with $p = 2$.

Tableau 3.3 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 3.4 summarizes results regarding Q4. For a persistence of $p = 0.03$ sec, 22 out 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 out 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.

Tableau 3.4 Perception of icon persistence for sound visualization (Q4).

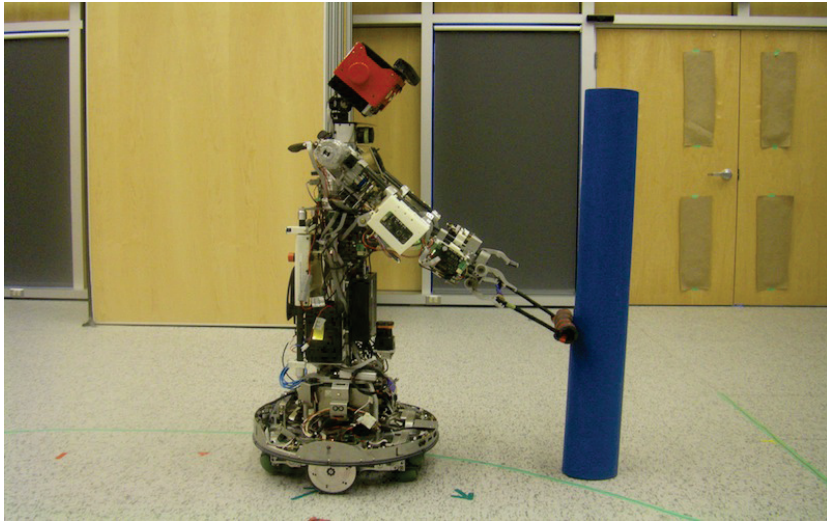
p (sec)	0.03	0.5	1	2	3	4	5
Short	22	4	1	0	0	0	0
Suitable	9	26	26	10	2	0	0
Long	0	1	4	21	29	31	31

3.8 Conclusion and Future Work

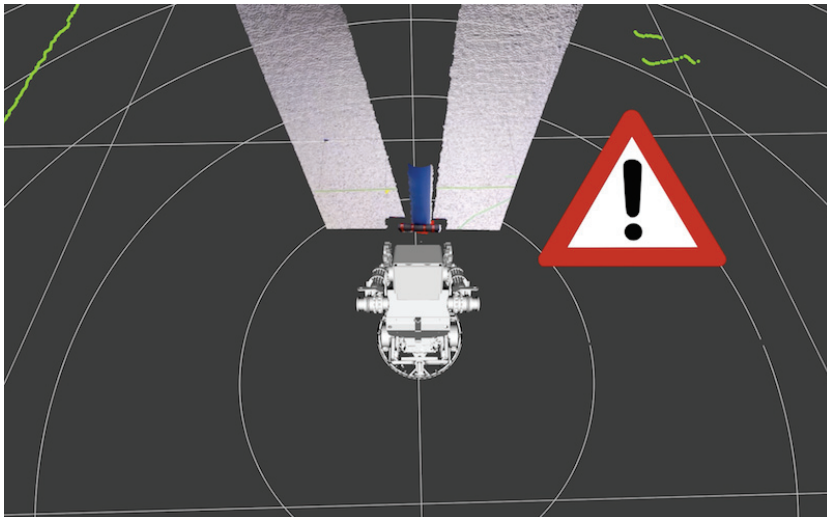
Sound and force sensing capabilities of mobile platforms bring additional information to be displayed on telepresence interfaces. This paper demonstrates how force and sound visualization can be used to provide information on an ecological 3D telepresence user interface. Results suggest that sound visualization 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

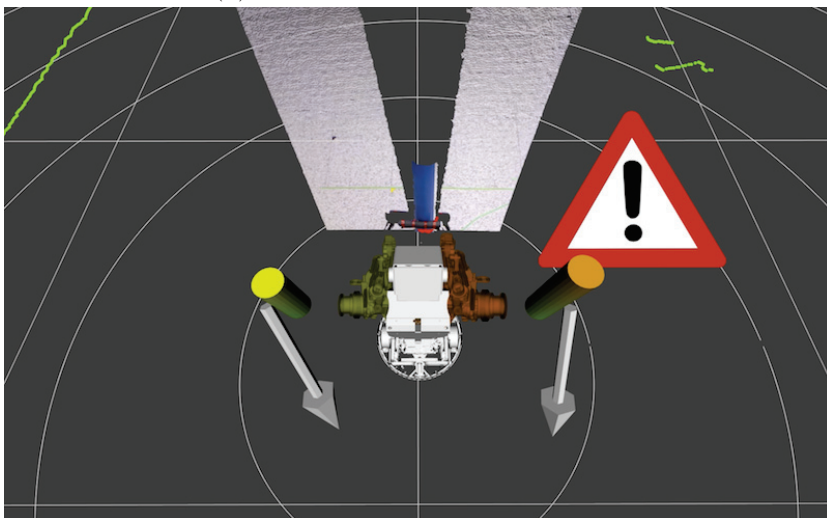
This work was supported by the Fonds de Recherche du Québec - Nature et Technologies (FRQ-NT). The authors want to thank Ronan Chauvin and all the volunteers who took part in the trials.



(a) Experimental setup

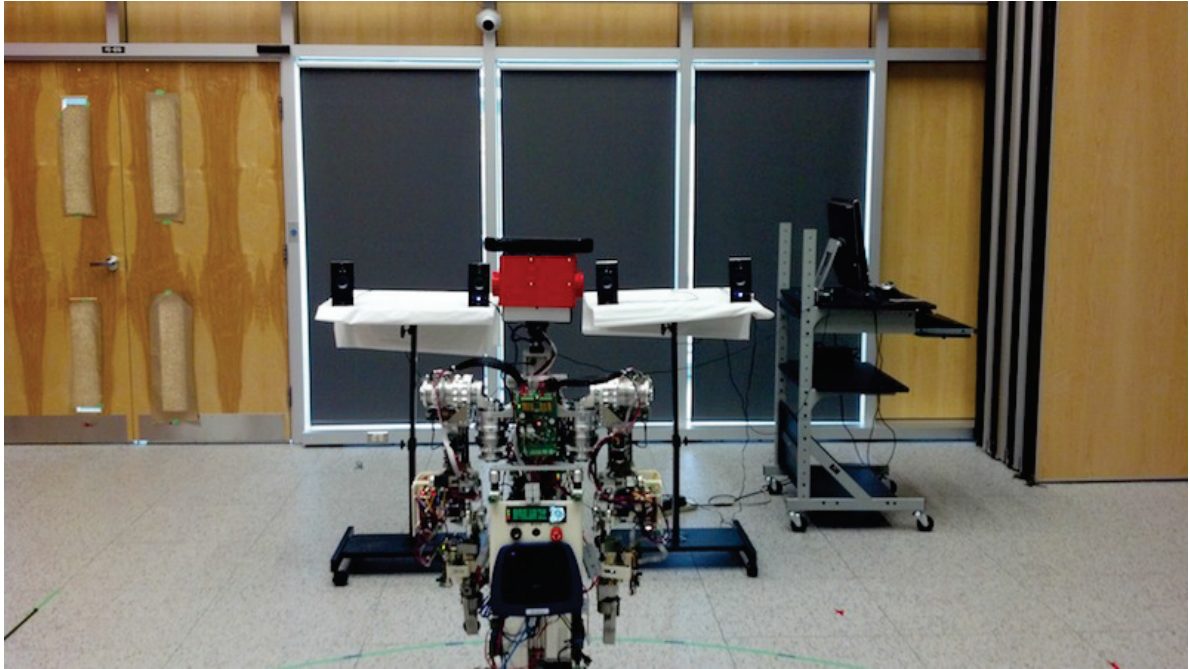


(b) GUI without force visualization

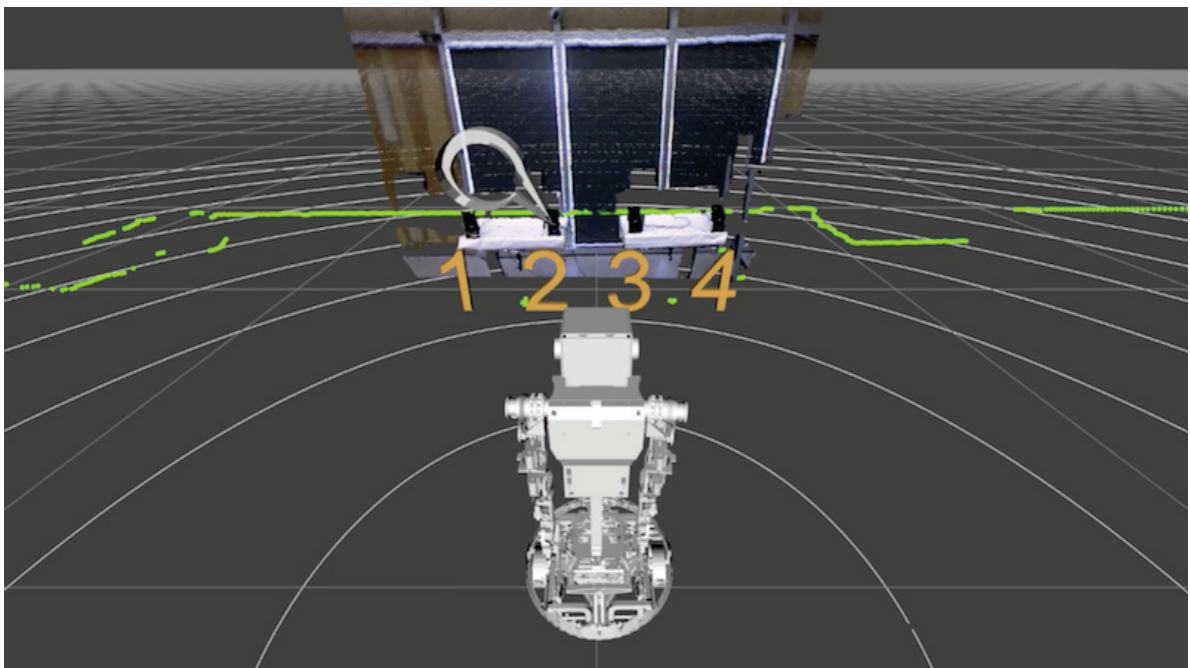


(c) GUI with force visualization

Figure 3.8 Force visualization trials. The vertical pole was blue and 1.25 m in height and 0.15 m in diameter. The U-shaped attachment was 0.36 m by 0.25 m.



(a) Experimental setup



(b) Sound visualization

Figure 3.9 Sound visualization trials

CHAPITRE 4

INTÉGRATION DE LA REPRÉSENTATION 3D DE L'ENVIRONNEMENT DANS UNE INTERFACE DE TÉLÉOPÉRATION

Le champs de vision limité de la Kinect (57° horizontalement et 43° verticalement jusqu'à 5 m) est apparu comme une limitation importante dans la compréhension de l'environnement distant par l'opérateur. En particulier, le nuage de points occupe seulement une petite partie de l'interface, provoquant une impression de vide. Ceci est renforcé lorsque la caméra virtuelle permettant d'ajuster le point de vue de l'utilisateur est située loin derrière la représentation modélisée du robot dans l'interface de téléopération. Dans un premier temps, nous avons envisagé de fusionner l'image provenant d'une caméra munie d'une lentille panoramique permettant d'obtenir une vue grand angle avec le nuage de point 3D de la Kinect. Seulement, par soucis d'afficher la modalité visuelle en 3D, nous avons choisi d'élargir le champs de vision du robot par affichage du nuage de points de la représentation de l'environnement générée dynamiquement par le module de cartographie 3D s'exécutant sur l'ordinateur du robot, module aussi appelé SLAM (« *Simultaneous Localization and Mapping* »).

L'architecture logicielle du module de cartographie est montrée à la figure 4.1. Les images de couleur et de profondeur provenant de la Kinect et les données d'odométrie du robot sont collectées pour produire une représentation 3D de l'environnement à chaque déplacement d'IRL-1. Le module de cartographie est capable de déterminer les endroits déjà visités par le robot par détection de boucle. À chaque détection de boucle, les erreurs d'alignements provoquées par le bruit et l'imprécision des capteurs peuvent être corrigées. En utilisant RTAB-Map (« *Real-Time Appearance-Based Mapping* ») [Labbe et Michaud, 2013], le module de cartographie génère un graphe dans lequel chaque noeud correspond à un nuage de points et chaque lien correspond à une transformation 6D entre deux nuages. Dès qu'une fermeture de boucle est détectée, un lien est ajouté entre le nouvel endroit visité et l'ancien. Avec cette nouvelle contrainte, TORO [Grisetti *et al.*, 2007] (« *Tree-based netwORk Optimizer* ») optimise le graphe de la carte utilisée pour générer la représentation 3D, afin de réduire l'erreur entre le nouvel et l'ancien endroit visité à chaque fermeture de

boucle. Les nuages de points de la carte sont ainsi utilisés pour augmenter le champs de vision de la Kinect et localiser IRL-1 dans son environnement.

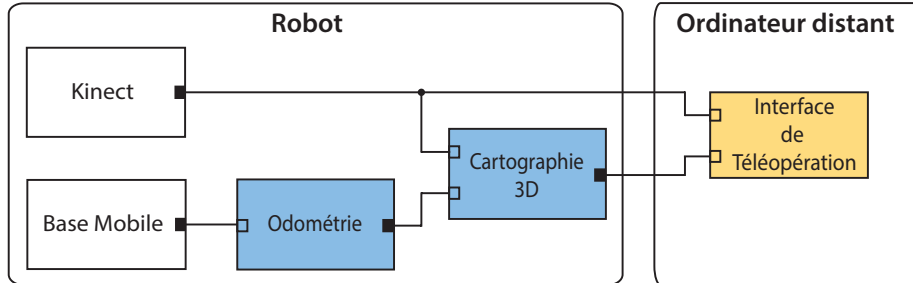


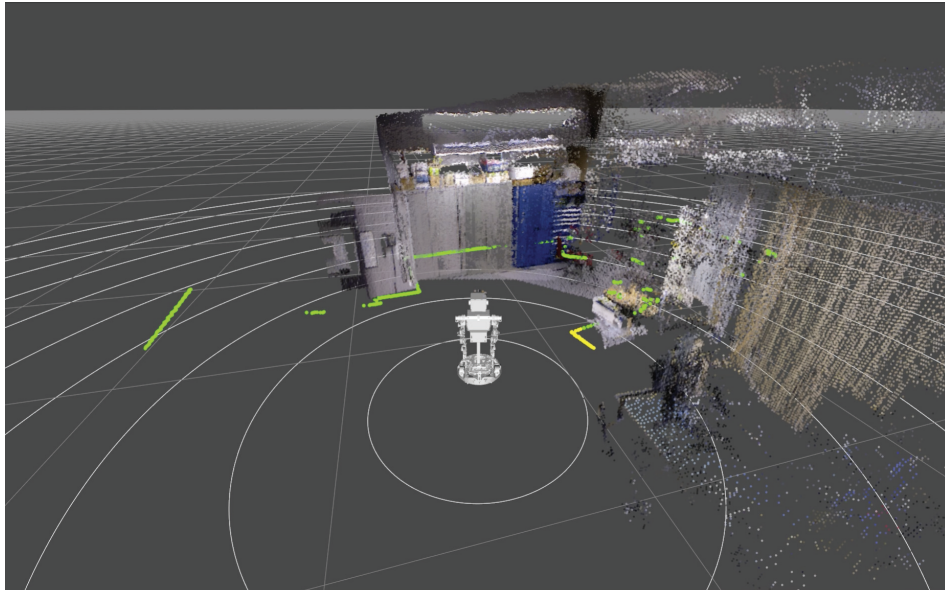
Figure 4.1 Architecture logicielle du module de cartographie 3D

La figure 4.2 illustre une vue typique générée par le module de cartographie et affichée sur l'interface graphique. La vue élargie de l'environnement peut être utile dans une perspective exocentrique pour donner une vue d'ensemble de l'environnement du robot comme montré à la figure 4.2a. Elle peut aussi augmenter le champs de vision aux abords du flux vidéo renvoyé par la Kinect dans le cas d'une perspective égocentrique comme illustré par la figure 4.2b.

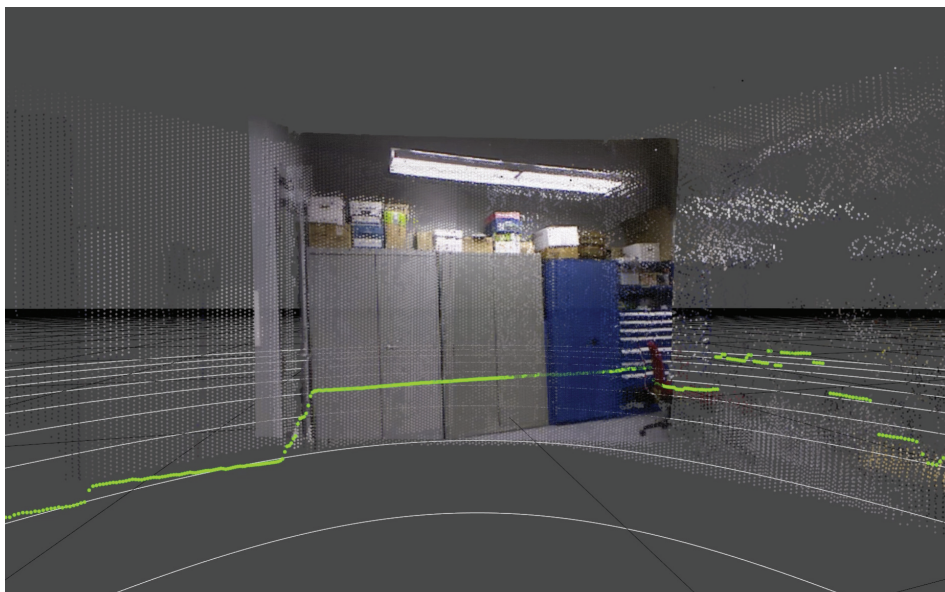
Le nombre de points contenu dans les nuages de points de la vue est sous-échantillonnés par un rapport de 25 : 1 pour permettre un taux de rafraichissement adéquat. Pour être différencié facilement du nuage de points courant provenant de la Kinect et affiché en pleine résolution, chaque point de la vue est grossi et affiché en transparence.

4.1 Expérimentation

Afin de mesurer l'impact de l'ajout de la vue 3D sur les performances de navigation des opérateurs, une troisième sous-tâche fut ajoutée aux expérimentations décrites à la section 3. Les participants devaient piloter IRL-1 à travers un slalom autour de trois cylindres amovibles espacés de 1.80 m les uns des autres. Le parcours à réaliser est montré à la figure 4.3. Les lignes de départ et d'arrivée tracées en vert au sol étaient situées entre deux chaises rouges visibles de loin par l'opérateur. Pour des raisons de sécurité, la vitesse linéaire du robot était limitée à 0.25 m/s. Le slalom était accompli quatre fois, deux avec la vue augmentée par la représentation 3D de l'environnement et 2 autres sans la vue augmentée. L'ordre de passage était déterminé aléatoirement tout en permettant d'intervertir les modalités. Ainsi, deux passages consécutifs ne pouvaient pas s'exercer avec les mêmes modalités. Le son ambiant était transmis en stéréo à l'opérateur par l'intermédiaire d'un



(a) Perspective exocentrique



(b) Perspective égocentrique

Figure 4.2 Représentation 3D de l'environnement affichée sur l'interface.

casque afin de lui permettre d'entendre le bruit des roues lors des déplacements du robot et de confirmer ce qui est affiché dans l'interface. Les fonctionnalités de visualisation des forces et des sources sonores étaient désactivées. Le temps d'exécution de chaque parcours et les consignes de vitesses transmises par la manette étaient enregistrés. Les performances mesurées lors de cette expérience visaient à vérifier l'hypothèse suivante :

- H4 : Les participants réalisent plus rapidement le slalom avec le champs de vision augmenté par la représentation 3D.

À la fin de l'expérience, chaque participant devait répondre à la question suivante :

- Q5 : Avez vous préféré la navigation avec la représentation 3D de l'environnement ?

4.2 Résultats

Les temps d'exécutions mesurés avec et sans la représentation 3D sont équivalents, 44.41s ($\sigma = 11.58s$) avec et 44.30s ($\sigma = 11.52s$) sans. En ce qui concerne le pourcentage du temps où les opérateurs avaient utilisé la consigne maximum, les résultats montrent qu'en moyenne 81.82% du temps est passé à pleine vitesse avec la représentation 3D et 84.62% sans, ce qui ne permet pas de d'observer une différence significative. L'étude qualitative révèle que 15 participants sur les 31 ont préféré réaliser le parcours avec la représentation 3D alors que 16 ont préféré sans. L'hypothèse H4 de départ n'est donc pas vérifiée. Plusieurs participants ont indiqué s'être fiés uniquement aux lectures du capteur laser, beaucoup plus précises et moins difficiles à interpréter et à utiliser que la représentation 3D pour la navigation. Dans notre cas, tous les obstacles à éviter étaient posés au sol et donc visibles par le capteur laser, ce qui montre que la simplicité du parcours choisi ne permettait pas de faire ressortir une différence appréciable. Cependant, les observations ont montré que pour certains participants qui n'arrivaient pas à effectuer une trajectoire continue et qui marquaient des pauses pendant le trajet, la représentation 3D permettait de se rendre compte du chemin parcouru et de la présence d'un cylindre, en particulier lorsque l'orientation du robot ne permettait plus au capteur laser de le voir. Les mêmes conclusions ont été obtenues par Mast *et al.* [Mast *et al.*, 2013] qui remarqua qu'une représentation 3D était utile pour naviguer entre des objets élevés ne jonchant pas le sol, mais pas pour naviguer autour d'objets déjà visibles par le capteur laser. Dans une prochaine expérience, il pourrait être intéressant de tester l'utilité de la représentation 3D dans une épreuve d'orientation. Par exemple, le robot pourrait être perdu dans un labyrinthe et le participant devrait retrouver la sortie en un minimum de temps.

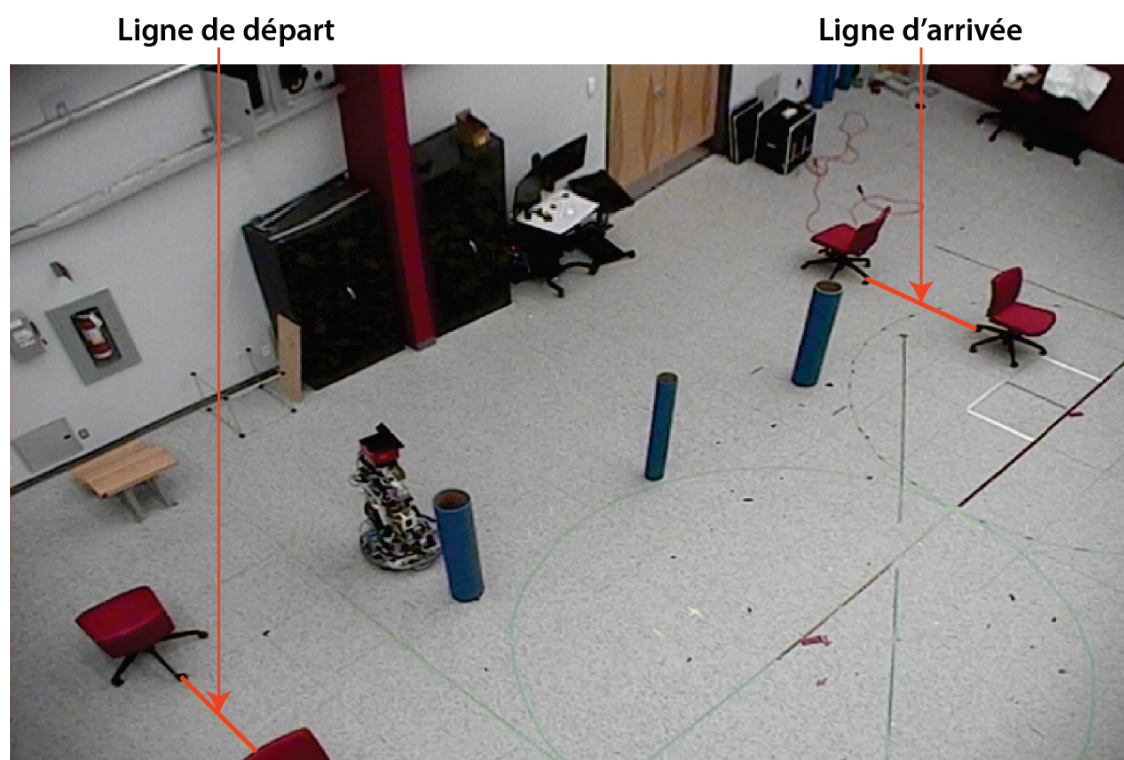
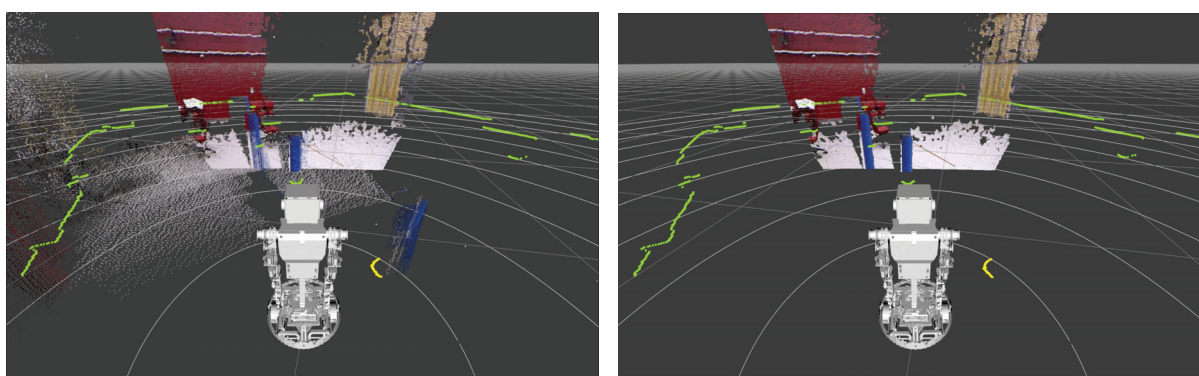


Figure 4.3 Vue externe du parcours à accomplir



(a) Interface augmentée avec la représentation 3D de l'environnement

(b) Interface sans affichage de la représentation 3D de l'environnement

Figure 4.4 Vue de l'interface pendant les expérimentations de navigation

CHAPITRE 5

CONCLUSION

Lors de cette étude, divers moyens de représentation intuitifs des modalités perceptuelles du robot IRL-1 ont été étudiés au sein d'une interface de téléopération écologique. En particulier, cette interface présente visuellement les sources sonores et les forces d'interaction perçues par le robot. Les sources sonores sont représentées par des bulles de bande dessinée ou des anneaux bleus selon le type de son. L'intensité des forces d'interaction du robot est matérialisée par la modification de la couleur et de la taille du rendu 3D du robot et avec des jauges de force, alors que la direction des forces est indiquée par des flèches. Enfin, le champ de vision restreint de la caméra 3D Kinect est augmenté par l'affichage de la représentation 3D de l'environnement construite lors des déplacements. Conformément aux recommandations relevées dans la revue de littérature conseillant de limiter le nombre de vues pour maximiser la conscience de la situation, l'affichage des nouvelles modalités a été intégré à la vue principale. Les expérimentations réalisées avec un groupe de 31 participants ont montré que la représentation visuelle des sons et des forces perçus par la plateforme améliore les performances des opérateurs pour des tâches spécifiques d'écoute et de manipulation. Les observations confirment que l'ajout d'informations visuelles facilement interprétables comme des icônes ou bien des modifications de taille et de couleur est utile aux opérateurs. L'utilité de l'affichage de la représentation 3D de l'environnement n'a toutefois pas pu être démontrée par le type de scénario de navigation envisagé. Pour valider cette modalité, le scénario choisi aurait dû comporter plus de difficultés et l'affichage des lectures lasers permettant de se repérer auraient dû être désactivées.

En fait, il reste encore difficile de concevoir un scénario de tests permettant d'évaluer l'interface de manière complète. Le scénario doit comporter des tâches suffisamment difficiles pour les opérateurs et permettre de mesurer l'apport de l'interface dans la diminution de la charge cognitive. La charge cognitive imposée par notre interface n'a notamment pas pu être mesurée pendant les expériences. Dans un premier temps, nous avons considéré l'idée de la mesurer par exécution d'une tâche secondaire simple, comme résoudre de petites opérations mathématiques. Les mains des participants étant prises par la manette, cela était difficile à mettre en œuvre. Dans le futur, la charge cognitive pourrait être calculée en mesurant le temps entre l'affichage de l'opération mathématique à l'écran et la réponse orale de l'opérateur. Ceci nécessite d'enregistrer les réponses des opérateurs et

de synchroniser la bande sonore avec l’affichage. Il pourrait alors être utile de se pencher exclusivement sur la normalisation des scénarios de tests dans une prochaine étude. La difficulté du scénario pourrait être déterminée par comparaison des performances d’experts par rapport aux novices, comme démontrée par Labonté [Labonté, 2008].

De plus, pour nos travaux, le débit moyen des données échangées entre l’interface et la plateforme était de 2.2 MB/s avec la compression des flux audio et vidéo. Nous disposions d’un réseau performant limitant la latence dans le transfert des données. Malgré l’amélioration continue de la performance des réseaux de communication, la latence reste un problème important car les robots mobiles disposent de plus en plus de capacités perceptuelles et de capteurs toujours plus performant. Par exemple, la version 2 de la Kinect est désormais en haute définition (1920×1080 pixels), ce qui augmentera considérablement le flux de donnée à transmettre.

Enfin, nous croyons qu’à moyen terme, l’opérateur distant partagera le contrôle avec le robot. Le robot pourra par exemple se déplacer de manière autonome dans l’environnement distant et l’opérateur prendra le contrôle seulement pour certaines tâches nécessitant encore l’expertise de l’humain. Il sera alors nécessaire de permettre à l’opérateur d’activer les modalités visuelles utiles selon ses besoins pour éviter toute surcharge de l’interface. La modularité de l’interface pouvant s’adapter à tous les types d’utilisateurs paraît donc être un requis essentiel.

ANNEXE A

PROFIL DES PARTICIPANTS

Tableau A.1 Profil des participants

No	Âge	Sexe	Occupation	Nbre d'heures de jeux vidéos par semaine (heures)	Expérience en robotique
1	23	M	Étudiant	1	Développement et Téléopération
2	24	M	Étudiant	1	Développement et Téléopération
3	23	M	Étudiant	1	Aucune
4	25	M	Étudiant	0	Aucune
5	25	M	Étudiant	1	Développement
6	24	M	Étudiant	Environ 7	Aucune
7	35	M	Étudiant	1	Développement
8	27	M	Pilote d'avion, Instructeur	30	Téléopération
9	25	M	Étudiant	5	Développement
10	69	M	Professeur	4	Aucune
11	22	M	Étudiant	3	Aucune
12	28	M	Étudiant	Environ 2	Aucune
13	34	F	Étudiante	0	Aucune
14	26	M	Étudiant	1	Aucune
15	25	M	Étudiant	2	Développement
16	27	M	Étudiant	1	Développement
17	24	M	Étudiant	2	Aucune
18	55	F	Professeur	Environ 5	Aucune
19	24	M	Étudiant	Environ 28	Aucune
20	21	F	Étudiante	0	Aucune
21	23	M	Étudiant	4	Aucune
22	25	M	Étudiant	4	Aucune
23	26	M	Étudiant	Environ 2	Téléopération
24	22	F	Étudiante	0	Aucune
25	28	M	Étudiant	1	Téléopération
26	34	M	Professionnel de recherche, ing. méc.	0 à 1	Téléopération
27	24	F	Étudiante	0	Aucune
28	23	M	Étudiant	0	Aucune
29	21	M	Étudiant	Environ 15	Aucune
30	24	M	Étudiant	3	Téléopération
31	27	M	Étudiant	7	Aucune

ANNEXE B

ATTESTATION DE CONFORMITÉ DU COMITÉ D'ÉTHIQUE



UNIVERSITÉ DE
SHERBROOKE

Comité d'éthique de la recherche
Éducation et sciences sociales

Attestation de conformité

Le comité d'éthique de la recherche Éducation et sciences sociales de l'Université de Sherbrooke certifie avoir examiné la proposition de recherche suivante :

Interface opérateur multimodale a réalité augmentée pour la robotique de téléprésence

Aurélien Reveleau

Étudiant, Maîtrise en génie électrique Faculté de génie

Projet financé par le FRQ-NT (Fonds de recherche du Québec - Nature et technologies)

Le comité estime que la recherche proposée est conforme aux principes éthiques énoncés dans la *Politique en matière d'éthique de la recherche avec des êtres humains (2500-028)*.

Membres du comité

Serge Striganuk, président du comité, professeur à la Faculté d'éducation, Département de gestion de l'éducation et de la formation

Chantale Beaucher, professeure à la Faculté d'éducation, Département de pédagogie

Mélanie Lapalme, professeure à la Faculté d'éducation, Département de psychoéducation

Mirela Moldoveanu, professeure à la Faculté d'éducation, Département d'enseignement au préscolaire et primaire

Gerardo Restrepo, professeur à la Faculté d'éducation, Département d'adaptation scolaire et sociale

Carlo Spallanzani, professeur à la Faculté d'éducation physique et sportive

Christina St-Onge, professeure à la Faculté de médecine et des sciences de la santé, Département de médecine

Vincent Beaucher, membre versé en éthique

France Dupuis, membre de la collectivité

Le présent certificat est valide pour la durée de la recherche, à condition que la personne responsable du projet fournisse au comité un rapport de suivi annuel, faute de quoi le certificat peut être révoqué.

Le président du comité,

Serge Striganuk, 21 janvier 2013

ANNEXE C

FORMULAIRE DE DEMANDE D'ÉVALUATION ÉTHIQUE

Comité d'éthique de la recherche – Éducation et sciences sociales Formulaire de demande d'évaluation éthique

Consignes pour remplir ce formulaire : le texte en italique comprend des précisions, des consignes particulières, des indications du type de réponse attendu, etc. Le texte en italique peut être effacé afin de laisser seulement les questions et les réponses. Dans le cas des choix de réponses (oui /non en général), effacez les réponses qui ne s'appliquent pas et laissez seulement la réponse de votre choix.

1. INFORMATIONS GÉNÉRALES

Chercheuse ou chercheur principal (étudiante ou étudiant si c'est le cas) :

Nom : Aurélien REVELEAU

Fonction : Étudiant à la maîtrise en génie électrique

Cochercheuses et cochercheurs (s'il y a lieu): Non

Équipe de direction dans le cas d'un projet d'étudiante ou d'étudiant :

François Michaud, Professeur à l'Université de Sherbrooke et directeur de l'IntRoLab

L'équipe de direction a approuvé cette demande d'évaluation éthique : Oui

Titre du projet : INTERFACE OPÉRATEUR MULTIMODALE A RÉALITÉ AUGMENTÉE POUR LA ROBOTIQUE DE TÉLÉPRÉSENCE.

Le projet est-il réalisé dans le cadre d'un programme d'étude?

Oui

Si oui, spécifier de quel programme il s'agit : Maîtrise en génie électrique.

Le projet de recherche a-t-il été évalué et accepté par un jury? Non

S'il s'agit d'un autre programme, expliquer l'état d'avancement du projet.

Nous débutons actuellement la deuxième année du projet. La première partie consistait à définir le projet. Nous sommes dans la phase de développement qui consiste à développer une interface de téléopération. La dernière étape du projet sera la phase des expérimentations.

Est-ce que ce projet est subventionné?

Oui

Si oui, spécifiez l'organisme subventionnaire : FRQ-NT (Fond de recherche du Québec - Nature et technologies).

Dans le cas des étudiants, recevez-vous une bourse : Oui (15 000 CAD/an).

2. PRÉSENTATION DU PROJET DE RECHERCHE

Ce projet est un nouveau projet.

S'agit-il d'un projet interfacultaire ou interuniversitaire?

Oui. L'équipe rattachée à la demande de subvention implique des professeurs de l'École de réadaptation (Michel Tousignant, Patrick Boissy, Hélène Corriveau), et de la Faculté de médecine (Andrew Grant).

À quelle date souhaitez-vous commencer la collecte de données?

Mars-avril 2013.

Résumer le projet de recherche

Ce résumé fait en général trois pages environ, ou plus dans le cas d'un projet plus complexe. Évitez toutefois de dépasser huit pages. Exposez brièvement la problématique. Appuyez vos propos de quelques références. Exposez clairement les objectifs et décrivez la méthodologie, en précisant le nombre de participantes et participants, la nature des instruments de collecte, etc. S.V.P. gardez à l'esprit que votre demande sera relue par des personnes qui ne proviennent pas de votre département : donc, assurez-vous que le projet soit compréhensible même pour des personnes non spécialistes dans votre domaine.

Problématique

De nos jours, l'humain se tourne de plus en plus vers les robots mobiles pour l'aider à réaliser certaines tâches difficiles. L'utilisation de robots s'est avérée particulièrement utile dans les domaines comme la recherche et le sauvetage urbain (« Urban Search And Rescue », ou USAR) [Baker *et al.*, 2004]. En effet, lors de catastrophes naturelles ou d'attentats volontaires, les robots mobiles peuvent permettre de retrouver des victimes sous les décombres en se faufilant sous les gravats, chose dont l'humain est incapable. De plus, le développement de la robotique de téléprésence pourrait soulager les hôpitaux et permettre aux personnes en perte de mobilité de rester chez elles plus longtemps [Michaud *et al.*, 2010]. Ces robots pourraient donc occuper une place de plus en plus importante dans la société. Des chercheurs ont récemment recensé les tâches domestiques pour lesquelles les foyers aimeraient disposer d'un robot [Sung *et al.*, 2009]. Des besoins dans l'assistance, la réalisation de tâches quotidiennes et l'aide à la prise de médicaments sont ressortis. Comme les robots ne sont pas encore totalement autonomes et que certaines situations nécessitent l'expertise humaine, ils doivent pouvoir être contrôlés à distance par un opérateur. Cette téléopération ne peut s'effectuer que par l'intermédiaire d'une interface qui permet à l'utilisateur de percevoir l'environnement distant dans lequel évolue la plateforme mobile. Cette interface doit retranscrire au mieux l'environnement à l'opérateur pour lui donner l'impression qu'il y est physiquement présent. Ces contraintes nécessitent donc de disposer de plateformes robotiques non seulement capables de naviguer, mais aussi disposant de modalités d'interactions supplémentaires, comme la vision, l'ouïe et le toucher. Le laboratoire IntRolab dispose justement de la plateforme robotique IRL1 (voir figure 1) intégrant ces modalités d'interaction.

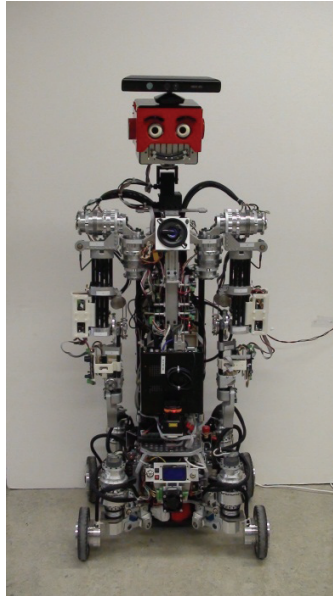


Figure 1 : Robot du laboratoire IntRoLab « IRL1 ».

Dans le contexte de l'étude, le robot est destiné à des fins de télésanté, et donc à être utilisé par des opérateurs inexpérimentés en robotique, comme des personnels soignants ou des praticiens par exemple. La revue de littérature montre qu'il est difficile de concevoir des interfaces qui permettent à l'opérateur de disposer d'une bonne conscience de la situation, c'est-à-dire de se faire une représentation la plus fidèle possible de l'environnement distant. De plus, l'ajout d'informations peut aboutir à une surcharge cognitive de l'utilisateur, c'est-à-dire qu'il va subir une charge de travail mental conséquente pour pouvoir exploiter correctement ces informations.

Des travaux sur les interfaces de téléopération ont déjà été effectués au sein du laboratoire au cours des neuf dernières années. Daniel Labonté a réalisé son doctorat en comparant trois types d'interfaces de téléopération pour robots de téléprésence. Le but de cette étude était d'identifier le type d'interface permettant de réaliser des opérations de navigation sécuritaires avec des opérateurs novices [Labonté, 2008]. Ensuite, le mémoire de François Ferland [Ferland *et al.*, 2008] porta sur une interface intégrant un modèle virtuel 3D fidèle du robot ainsi que la possibilité pour l'utilisateur de changer de vue selon les tâches qu'il devait réaliser. Dans la continuité de ces travaux, nous émettons l'hypothèse qu'en transmettant les modalités sensorielles de notre robot IRL1 à l'opérateur distant par l'intermédiaire de l'interface, celui-ci pourra se faire une représentation plus fidèle de l'environnement dans lequel évolue la plateforme. Pour ne pas surcharger cognitivement l'opérateur, nous allons adapter la méthode de transmission des informations en utilisant de la réalité augmentée, c'est-à-dire en superposant des icônes directement sur le flux vidéo transmis par la caméra du robot. La figure 2 montre l'interface retranscrivant le son de la voix d'une personne perçu par le robot sous forme d'une bulle superposée au flux vidéo.

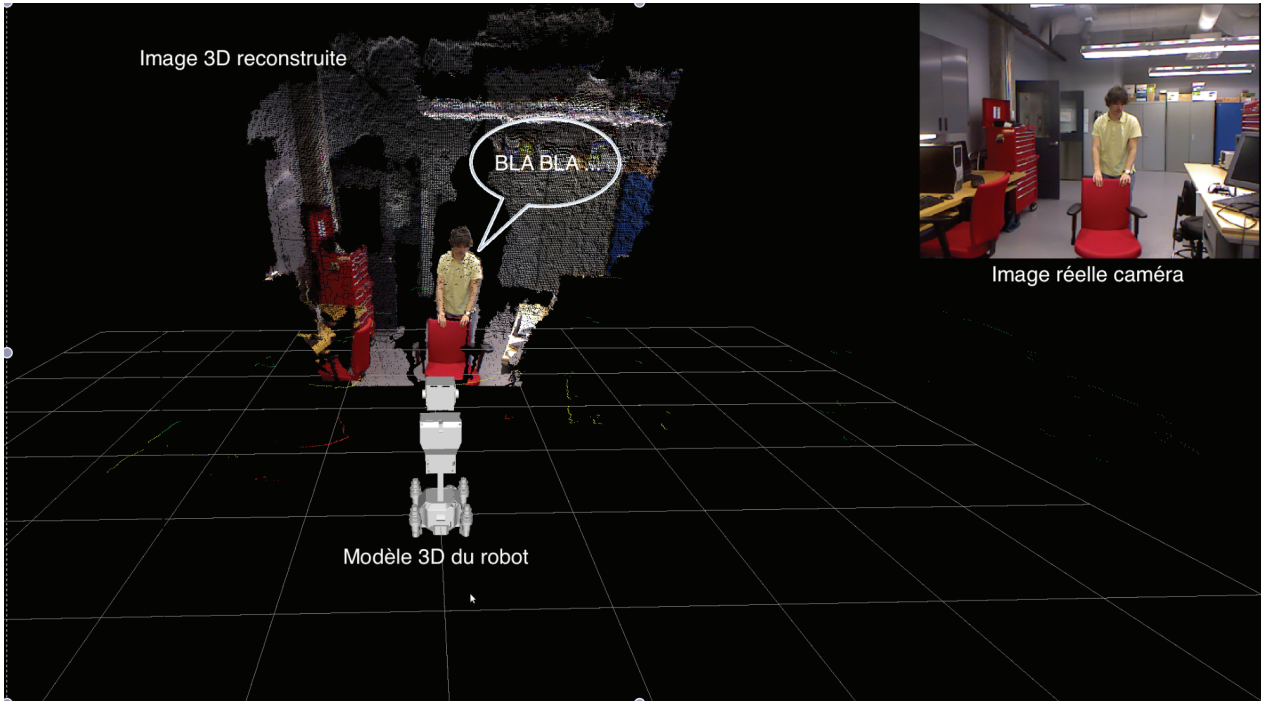


Figure 2 : Interface représentant les sons perçus par le robot par réalité augmentée.

Le but du projet est de concevoir et tester une interface de téléopération retranscrivant les modalités sensorielles de notre robot IRL1 qui soit adaptée à des opérateurs novices.

Objectifs

L'objectif principal est d'étudier l'impact de la fusion et de l'affichage des modalités sensorielles par icônes superposées sur le flux vidéo sur la conscience de la situation et sur la charge cognitive de l'opérateur. Il nous paraît aussi important de tester plusieurs types d'icônes et ainsi de connaître les préférences des utilisateurs. Nous pensons que ce type d'interface, en plus d'être préférée par les opérateurs débutants, améliorera leurs performances.

Méthodologie

Le dispositif de recherche est un robot humanoïde (IRL1) piloté à distance par des opérateurs novices par l'intermédiaire de l'interface développée pour simuler un dispositif de téléprésence.

Position des opérateurs et matériel de mesure à leur disposition

Les opérateurs seront situés dans un local à partir duquel ils n'auront aucun contact physique ou visuel avec le robot. Les personnes seront assises devant un poste informatique par l'intermédiaire duquel elles devront contrôler le robot.

Ce poste sera équipé de FaceLABTM, un instrument de mesure permettant de suivre le regard des opérateurs, et ainsi de connaître les endroits de l'écran les plus scrutés¹. En synchronisant FaceLABTM avec l'apparition des représentations des informations sensorielles, il sera possible de connaître le temps de réaction d'un opérateur à chaque affichage. Nous pensons que le type d'affichage préféré concordera avec celui dont la moyenne du temps de réaction sera le plus faible.

¹ <http://www.seeingmachines.com/product/faceLAB/>

Position du robot

Le robot sera quant à lui situé dans le laboratoire d'évaluation des usages et d'ergonomie (LEUE) de l'Institut interdisciplinaire d'innovation technologique (3IT). Ce laboratoire est conçu spécialement pour réaliser des tests dans des conditions réalistes d'utilisation (avec des obstacles et des espaces similaires à un domicile dans le cas du présent projet) pour des applications de télésanté.

Déroulement des expérimentations

Deux tâches de téléopération sur deux parcours de difficulté similaire seront réalisées par le même groupe de test. Une tâche sera réalisée avec l'interface sans présentation des informations de sons et de forces par icône, alors que l'autre sera réalisée avec la nouvelle interface développée. Pour ne pas subir un effet d'entraînement, la moitié du groupe devra commencer avec la première interface tandis que l'autre moitié commencera avec la deuxième. Les opérateurs devront faire déplacer le robot dans les pièces du domicile reconstitué en suivant un parcours prédéfini. La région d'expérimentation sera divisée en plusieurs zones et il y aura un point d'intérêt émettant une source sonore par zone. Les opérateurs devront trouver et identifier ces points d'intérêt en s'approchant de la zone d'émission sonore avec le robot. Il y aura trois types d'icônes de représentation des sons et ceux-ci changeront durant les opérations. L'idéal serait qu'il y ait donc six sources sonores pour avoir deux présentations de sources sonores par le même type d'icônes. Au moins un corridor et un passage étroit (passage de porte) seront reconstitués dans l'appartement. La présentation des sons devrait permettre aux opérateurs de se diriger rapidement vers les points d'animation de la zone d'expérimentation. La présentation des forces devrait quant à elle permettre aux utilisateurs de bien connaître la situation dans laquelle se trouve le robot, s'il roule sur un obstacle ou si ses bras entrent en collision avec un obstacle.

Variables de mesure

La performance des opérateurs sera évaluée selon les trois métriques suivantes :

- Temps de complétion de la tâche
- Nombre de collisions avec l'environnement
- Temps de perception des événements pour comparer les moyens de présentation des informations

Évaluation de la conscience de la situation

L'évaluation de la conscience de la situation sera faite à l'aide de la méthode LASSO (« *Location, Activities, Surroundings, Status and Overall mission* »). Cette technique développée par Drury et al. [Drury et al., 2007] dans le domaine des secours urbains consiste à séparer la conscience de la situation en cinq catégories qui sont la connaissance du positionnement du robot (« *Location Awareness* »), des tâches qu'il est en train d'effectuer (« *Activity Awareness* »), de son environnement proche (« *Surroundings Awareness* »), de son état physique (« *Status Awareness* ») et de la mission globale qu'il est en train d'effectuer. L'évaluation se fait en incitant les opérateurs à exprimer leurs pensées à voix hautes. On catégorise ensuite les différents commentaires comme positifs, négatifs ou neutres pour les catégories énoncées précédemment.

Évaluation de la charge cognitive

L'évaluation de la charge cognitive sera faite par ajout d'une tâche duale nécessitant peu de ressource. Cette seconde tâche prendra la forme d'une addition très simple à effectuer pendant les opérations de contrôle. Nous considérons que les performances sur la seconde tâche dépendront directement du niveau de charge cognitive imposée par la première.

Évaluation qualitative

Enfin, nous souhaitons connaître l'avis subjectif des participants par rapport aux fonctionnalités de l'interface. Pour cela, une liste d'affirmations leur sera soumise à la fin des opérations. Les participants

devront exprimer leur degré d'accord ou de désaccord avec ces affirmations. Les choix de réponses seront répartis sur une échelle de Likert² à 5 niveaux. Un exemple d'affirmation pourrait être la suivante : « L'interface avec présentation des sources sonores me permet de mieux percevoir l'environnement dans lequel le robot évolue ». Les réponses seront les suivantes : Je suis « tout à fait d'accord », « d'accord », « ni d'accord, ni en désaccord », « en désaccord » ou « en total désaccord » avec l'affirmation précédente.

Participants à l'étude et taille de l'échantillon

Conformément aux analyses statistiques déjà réalisées au laboratoire par Daniel Labonté pour l'évaluation d'interfaces de téléopération [Labonté, 2008], nous avons estimé la taille de l'échantillon afin d'être en mesure de détecter avec une puissance de 80% et un alpha de 0,05 lors d'un test-t unilatéral les performances dans les métriques choisies correspondant à une taille d'effet de 0,5. La taille minimum de l'échantillon obtenue à l'aide du logiciel G*Power³ est de 27 participants. Nous souhaitons donc avoir un nombre de 30 participants.

Ces participants devront être inexpérimentés en robotique, c'est à dire ne pas être familier du contrôle et des interactions avec des robots. Bien que les participants de l'étude seront sélectionnés parmi les professeurs, étudiants et membres du personnel du 3IT qui seront sans doute à l'aise avec les technologies, ils n'auront probablement jamais piloté de robots en conditions réelles et seront totalement novices dans le domaine de la téléprésence. De plus, la connaissance des participants par rapport à la technologie peut permettre de faire ressortir des remarques techniques pertinentes associées à l'interface et non pas associées au matériel utilisé.

Analyse des données

Pour observer si des différences statistiquement significatives existent pour les métriques concernant le temps de complétions de la tâche, le nombre de collisions avec l'environnement, la conscience de la situation ainsi que la charge cognitive, nous choisissons de réaliser des tests d'hypothèses. L'hypothèse nulle sera : « il n'existe pas de différence statistiquement significative entre les interfaces ». L'hypothèse alternative sera : « l'interface augmentée améliore les performances des utilisateurs ». Nous choisissons donc d'utiliser un test-t puisqu'il permet de savoir si la différence entre deux moyennes est statistiquement significative. Les tests seront effectués avec deux interfaces différentes sur le même groupe de personnes. Nous disposons donc d'un échantillon apparié. Nous envisageons donc d'effectuer cinq test-t pour échantillons appariés sur les métriques suivantes :

- Moyenne des temps de complétion de la tâche.
- Moyenne du nombre de collisions.
- Moyenne du nombre de commentaires positifs pour les cinq catégories de conscience de la situation.
- Moyenne du nombre de commentaires négatifs pour les cinq catégories de conscience de la situation.
- Moyenne du temps de réaction dans l'exécution de la tâche duale.

Pour la comparaison du temps de réaction des opérateurs en fonction du type d'icône, une analyse de variance ANOVA a été choisie puisque ce type d'analyse statistique permet de comparer entre elles plusieurs moyennes observées. Nous disposons de trois types d'icônes différentes. Chaque icône sera présentée deux fois aux participants (deux sources sonores différentes avec le même type d'icône). Les moyennes des temps de réactions pour ces trois types d'icônes seront ainsi comparées et permettront de savoir si un type particulier d'icône diminue le temps de réaction de l'opérateur. Le nombre

² http://fr.wikipedia.org/wiki/Échelle_de_Likert

³ <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>

d'échantillons des temps de réactions pour chaque type d'icône correspondra à deux fois le nombre de participants (deux mesures par participant pour chaque type d'icône).

Justifier le recours à des êtres humains et le choix du groupe de personnes participantes

Le but est de concevoir une interface de téléopération utilisée par des opérateurs humains pour contrôler un robot distant. L'interface doit se révéler plus intuitive d'utilisation pour des opérateurs novices que les interfaces actuellement développées. Il est donc nécessaire de tester la nouvelle interface développée avec un échantillon de personnes répondant aux exigences énoncées précédemment. Il est important d'étendre les tests à des personnes externes au laboratoire pour disposer d'un échantillon de personnes le moins familier possible du pilotage de robots.

Retombées scientifiques et sociales attendues

Cette recherche doit mener à des progrès dans les spécifications de conception d'interface adaptée aux personnes débutantes, n'étant pas familière du pilotage et du contrôle de robots mobiles. Elle doit aussi permettre de voir si la réalité augmentée et la transmission d'informations sensorielles permettent aux opérateurs d'avoir une meilleure représentation de l'environnement.

Références bibliographiques

- Baker, M., Casey, R., Keyes, B. et Yanco, H. A. (2004). Improved interfaces for human-robot interaction in urban search and rescue. Dans *Proceedings IEEE International Conference on Systems, Man and Cybernetics*. p. 2960-2965.
- Drury, J. L., Keyes, B. et Yanco, H. A. (2007). LASSOing HRI : Analyzing situation awareness in map-centric and video-centric interfaces. Dans *Proceedings ACM/IEEE Conference on Human-Robot Interaction*. p. 279-286.
- Ferland, F., Pomerleau, F., Dinh, C. T. L. et Michaud, F. (2008). Egocentric and exocentric teleoperation interface using real-time, 3D video projection. Dans *Proceedings 4th ACM/IEEE International Conference on Human-Robot Interaction*. p. 37-44.
- Labonté, D. (2008). *Interface-opérateur à réalité et perspective mixtes de téléopération de robot mobile pour une application de téléassistance*. Thèse de doctorat, Université de Sherbrooke, Sherbrooke, Québec, Canada, 190 p.
- Michaud, F., Boissy, P., Labonte, D., Briere, S., Perreault, K., Corriveau, H., Grant, A., Lauria, M., Cloutier, R., Roux, M. -, Iannuzzi, D., Royer, M. -, Ferland, F., Pomerleau, F. et Letourneau, D. (2010). Exploratory design and evaluation of a homecare teleassistive mobile robotic system. *Mechatronics*, volume 20, numéro 7, p. 751-766.
- Sung, J., Christensen, H. I. et Grinter, R. E. (2009). Sketching the future: Assessing user needs for domestic robots. Dans *Proceedings 18th IEEE International Symposium on Robot and Human Interactive Communication*. p. 153-158.

3. ASPECTS ÉTHIQUES

Équilibre entre les risques et les bénéfices

Quels sont les **risques encourus** par les personnes participantes?

Il n'y a aucun risque encouru.

Tout risque physique sera inexistant puisque la personne qui pilotera le robot sera éloignée de celui-ci et aucun contact physique ne sera possible.

Il n'y aura pas non plus de risque psychologique puisque les personnes ciblées ne sont pas de nature vulnérable.

Enfin le risque social sera inexistant puisque les performances de chaque personne ne seront pas divulguées. Les participants ne pourront donc pas être jugés par rapport à celles-ci.

Combien de **temps** demandera la participation des personnes participantes?

Le temps total de participation devrait se situer autour de 20 minutes séparées en trois parties :

- La phase de familiarisation avec le robot sera d'une durée d'environ 5 minutes.
- La phase de l'expérience d'une durée d'environ 10 minutes (5 minutes par parcours).
- La phase d'évaluation qualitative (remplissage du questionnaire) d'environ 5 minutes.

Quels sont les **bénéfices** qu'ils pourront retirer?

Ils vont participer à l'avancement des connaissances en robotique. La robotique de téléprésence tend à se développer et le fait de piloter un robot permettra de leur faire découvrir ce qui fera peut être partie de leur quotidien dans quelques années.

Est-ce qu'une **compensation**, monétaire ou autre, sera remise pour le temps, les déplacements, etc. occasionnés par la participation au projet?

Non.

Est-ce que ce projet de recherche se situe sous le **seuil de risque minimal**?

Oui.

Consentement libre et éclairé

Est-ce que la recherche est de nature **consensuelle**?

Oui.

Quelles seront les **mesures prises pour s'assurer du consentement libre et éclairé des personnes participantes**?

Concrètement, comment seront recrutés les participants et participantes à la recherche?

Les personnes seront des étudiants, des professeurs et des membres du personnel travaillant au sein de l'Institut interdisciplinaire d'innovations (3IT) de l'Université de Sherbrooke.

Un courriel (copie en annexe) sera envoyé par le chercheur responsable du projet à tous les membres permanents (étudiants, professeurs, professionnels de recherche et personnel administratif) du 3IT pour solliciter leur participation. Les personnes ayant des questions sur le déroulement du projet seront invitées à venir dialoguer avec le chercheur responsable.

Pour les personnes qui auront accepté de participer, est-ce qu'un formulaire de consentement sera signé? Oui.

Est-ce que la recherche implique des **personnes mineures ou légalement inaptes**?

Non.

Confidentialité des données

Quelles seront les mesures prises pour assurer le **caractère confidentiel et anonyme des données**?
Utilisation de codes numériques pour protéger l'identité des personnes.

Où seront conservées les données? Seront-elles conservées sous clé?

Les fichiers seront conservés par le Laboratoire de robotique de l'Université de Sherbrooke, IntRoLab situé au 3^{ème} étage du bâtiment accessible seulement par les étudiants, professeurs, professionnels de recherche et membres du personnel du 3IT autorisés et ayant signé un contrat de non divulgation. Les fichiers électroniques seront sur le serveur du laboratoire, protégé par un mot de passe.

Qui aura accès aux données? (Chercheuses et chercheurs, assistantes et assistants de recherche,...)

Ces personnes doivent être conscientisées notamment à l'importance de respecter la confidentialité des données.

Les données seront accessibles uniquement par le chercheur, son directeur de recherche et les membres de l'équipe de recherche.

Quand seront détruites les données brutes (questionnaires papier, cassettes d'entrevues, etc.)?

Les données brutes seront détruites cinq ans après les publications associées au projet pour respecter les exigences des éditeurs.

Quels sont les **moyens de diffusion** prévus?

Les deux listes qui suivent présentent les cibles de conférences et de journaux où une publication d'article serait possible.

- Conférences

HRI Human Robot Interaction

- Soumission : fin septembre
- Déroulement : début mars
- Lieu : Bielefeld, Allemagne en 2014

ICRA IEEE International Conference on Robotics and Automation :

- Soumission : mi-septembre
- Déroulement : fin mai
- Lieu : Hong Kong, Chine en 2014

SMC IEEE International Conference on Systems, Man, and Cybernetics :

- Soumission : avril
- Déroulement : mi-octobre
- Lieu : Manchester, Royaume-Unis en 2103

RO-MAN IEEE International Symposium on Robots and Human Interactive Communications :

- Soumission : mi-février
- Déroulement : fin août
- Lieu : Gyeonjgu, Corée en 2013

IROS IEEE/RSJ International Conference on Intelligent Robots and Systems :

- Soumission : fin juillet

- Déroulement : novembre 2013
- Lieu : Tokyo, Japon

- Journaux

- Journal of Human-Robot Interaction
- International Journal of Advanced Robotic Systems
- IEEE Transactions on Systems, Man, and Cybernetics

Est-ce que les moyens de diffusion des résultats risquent de permettre **l'identification** des personnes ayant participé à la recherche?

Non.

Est-ce qu'une **utilisation secondaire des données** est prévue?

Non.

En remettant votre demande au comité, vous vous engagez à faire parvenir au comité un rapport annuel et un avis à la fin du projet. Vous pouvez à cette fin remplir le formulaire de suivi disponible sur le site Internet du comité d'éthique de la recherche Éducation et Sciences sociales et nous l'acheminer par courriel.

Également en cas de changements au projet, vous devez aviser le comité à l'aide de ce même formulaire.

**Invitation à participer et formulaire de consentement pour le projet de recherche
INTERFACE DE CONTRÔLE POUR UN ROBOT MOBILE DE TÉLÉPRÉSENCE**
Aurélien Reveleau étudiant à la maîtrise en génie électrique au Laboratoire de Robotique Intelligente
(IntRoLab)

Sous la direction de François Michaud, Professeur à l'Université de Sherbrooke et directeur de l'IntRoLab. Le chercheur responsable du projet a reçu un financement du Fond de recherche du Québec - Nature et technologies (FRQ-NT) pour mener à bien ce projet de recherche.

Madame,
Monsieur,

Nous vous invitons à participer à la recherche en titre. Ce projet de recherche a pour but de participer à l'amélioration des interfaces de téléopération de robots mobiles dans des contextes de téléprésence. Pour cela, nous vous proposons de contrôler à distance notre robot humanoïde IRL-1 que vous devrez piloter dans un environnement s'apparentant à un appartement ou un domicile. L'objectif est de savoir si la transmission des modalités d'interactions du robot (vision, ouïe et forces) améliore les performances de téléopération d'opérateurs inexpérimentés en robotique.

En quoi consiste la participation au projet?

Votre participation à ce projet de recherche consiste à contrôler le robot à distance. Le robot sera situé dans un local éloigné de vous et vous devrez le contrôler par l'intermédiaire d'un poste informatique. **Ce poste informatique se trouvera dans un local de télésanté possédant 4 caméras permettant de filmer l'expérience.**

L'expérience se déroule en trois phases :

- Vous vous familiariserez avec l'interface et le contrôle du robot en lui faisant faire quelques mouvements libres pendant environ 20 minutes.
- Vous effectuerez trois petites tâches ou parcours différents (environ 35 minutes). Chacune d'entre elles sera réalisée plusieurs fois, soit à l'aide d'une interface avec représentation des sources sonores, des forces ainsi que de la carte du robot ou bien avec une interface n'affichant pas ces modalités.
- Vous répondrez à un questionnaire (environ 5 minutes) sur vos préférences.

Le seul inconvénient lié à votre participation est le temps consacré à la recherche, soit environ une heure.

Qu'est-ce que le chercheur fera avec les données recueillies?

Pour éviter votre identification comme personne participante à cette recherche, les données recueillies par cette étude seront traitées de manière **entièrement confidentielle**. La confidentialité sera assurée par l'utilisation de codes numériques pour remplacer vos noms. Les résultats de la recherche ne permettront pas d'identifier les personnes participantes. Les résultats seront diffusés dans le mémoire de maîtrise, dans des articles scientifiques et dans des congrès internationaux dans le domaine de la robotique. Les données recueillies seront conservés par le Laboratoire de robotique de l'Université de Sherbrooke, IntRoLab situé au 3^{ème} étage du bâtiment accessible seulement par les étudiants, professeurs, professionnels de recherche et membres du personnel du 3IT autorisés et ayant signé un contrat de non divulgation. Les fichiers électroniques seront sur le serveur du laboratoire, protégé par un mot de passe. Les données seront détruites en 2018, soit 5 ans après leur collecte pour respecter les exigences des éditeurs et ne seront pas utilisées à d'autres fins que celles décrites dans le présent document.

Est-il obligatoire de participer?

Non. La participation à cette étude se fait sur une base volontaire. Vous êtes entièrement **libre de participer ou non**, et de vous retirer en tout temps sans avoir à motiver votre décision ni à subir de préjudice de quelque nature que ce soit.

Y a-t-il des risques, inconvéniens ou bénéfiques?

Il n'y a aucun inconvéniens direct pouvant découler de votre participation à ce projet. Le temps requis pour participer aux expérimentations et remplir le questionnaire est d'environ 60 minutes. La contribution à l'avancement des connaissances dans le domaine de la robotique sont les bénéfiques prévus. Aucune compensation d'ordre monétaire n'est accordée.

Que faire si j'ai des questions concernant le projet?

Si vous avez des questions concernant ce projet de recherche, n'hésitez pas à communiquer avec moi aux coordonnées indiquées ci-dessous.

[Signature de la chercheuse ou du chercheur] [Date de la signature]

Aurélien Reveleau, étudiant à la maîtrise en génie électrique
Chercheur responsable du projet de recherche sous la direction de François Michaud (Directeur)
Faculté de génie

*J'ai lu et compris le document d'information au sujet du projet **Interface opérateur multimodale a réalité augmentée pour la robotique de téléprésence**. J'ai compris les conditions, les risques et les bienfaits de ma participation. J'ai obtenu des réponses aux questions que je me posais au sujet de ce projet. J'accepte librement de participer à ce projet de recherche.*

Participante ou participant :

Signature :

Nom :

Date :

**S.V.P., signez les deux copies.
Conservez une copie et remettez l'autre au chercheur.**

Ce projet a été revu et approuvé par le comité d'éthique de la recherche Éducation et sciences sociales, de l'Université de Sherbrooke. Cette démarche vise à assurer la protection des participantes et participants. Si vous avez des questions sur les aspects éthiques de ce projet (consentement à participer, confidentialité, etc.), n'hésitez pas à communiquer avec M. Serge Striganuk, président de ce comité, par l'intermédiaire de son secrétariat.

Copie du courriel d'invitation à la participation aux expérimentations

Invitation à participer aux expérimentations pour le projet de recherche **INTERFACE DE CONTRÔLE POUR UN ROBOT MOBILE DE TÉLÉPRÉSENCE**

Aurélien Reveleau étudiant à la maîtrise en génie électrique au Laboratoire de Robotique Intelligente (IntRoLab)

Sous la direction de François Michaud, Professeur à l'Université de Sherbrooke et directeur de l'IntRoLab. Le chercheur responsable du projet a reçu un financement du Fond de recherche du Québec - Nature et technologies (FRQ-NT) pour mener à bien ce projet de recherche.

Madame,
Monsieur,

Nous vous invitons à participer à la recherche en titre. Ce projet de recherche a pour but de participer à l'amélioration des interfaces de téléopération de robots mobiles dans des contextes de téléprésence. Pour cela, nous vous proposons de contrôler à distance notre robot humanoïde IRL1 que vous devrez piloter dans un environnement s'apparentant à un appartement ou un domicile. L'objectif est de savoir si la transmission des modalités d'interactions du robot (vision, ouïe et forces) améliore les performances de téléopération d'opérateurs inexpérimentés en robotique.

En quoi consiste la participation au projet?

Votre participation à ce projet de recherche consiste à contrôler le robot à distance. Le robot sera situé dans un local éloigné de vous et vous devrez le contrôler par l'intermédiaire d'un poste informatique. **Ce poste informatique se trouvera dans un local de télésanté possédant 4 caméras permettant de filmer l'expérience.**

L'expérience se déroule en trois phases :

- Vous vous familiariserez avec l'interface et le contrôle du robot en lui faisant faire quelques mouvements libres pendant environ 20 minutes.
- Vous effectuerez trois petites tâches ou parcours différents (environ 35 minutes). Chacune d'entre elles sera réalisée plusieurs fois, soit à l'aide d'une interface avec représentation des sources sonores, des forces ainsi que de la carte du robot ou bien avec une interface n'affichant pas ces modalités.
- Vous répondrez à un questionnaire (environ 5 minutes) sur vos préférences.

Le seul inconvénient lié à votre participation est le temps consacré à la recherche, soit environ une heure.

Est-il obligatoire de participer?

Non. La participation à cette étude se fait sur une base volontaire. Vous êtes entièrement **libre de participer ou non**, et de vous retirer en tout temps sans avoir à motiver votre décision ni à subir de préjudice de quelque nature que ce soit.

Y a-t-il des risques, inconvénients ou bénéfices?

Il n'y a aucun inconvénient direct pouvant découler de votre participation à ce projet. Le temps requis pour participer aux expérimentations et remplir le questionnaire est d'environ 60 minutes. La contribution à l'avancement des connaissances dans le domaine de la robotique sont les bénéfices prévus. Aucune compensation d'ordre monétaire n'est accordée.

Si vous désirez participer aux expérimentations ou si vous avez des questions concernant ce projet de recherche, n'hésitez pas à communiquer avec moi aux coordonnées indiquées ci-dessous.

Aurélien Reveleau, étudiant à la maîtrise en génie électrique
Chercheur responsable du projet de recherche sous la direction de François Michaud (Directeur)
Faculté de génie

LISTE DES RÉFÉRENCES

- Azar, J., Saleh, H. A. et Al-Alaoui, M. (2009). Sound visualization for the hearing impaired. *International Journal of Emerging Technologies in Learning*, volume 2, numéro 1, p. 1–7.
- Baker, M., Casey, R., Keyes, B. et Yanco, H. A. (2004). Improved interfaces for human-robot interaction in urban search and rescue. Dans *Proceedings IEEE International Conference on Systems, Man and Cybernetics*. p. 2960–2965.
- Bar-Cohen, Y. (2003). Haptic devices for virtual reality, telepresence, and human-assistive robotics. *Biologically Inspired Intelligent Robots*, p. 73.
- Blauert, J. (1996). *Spatial Hearing - Revised Edition : The Psychophysics of Human Sound Localization*. The MIT Press.
- Borst, C., Wimbrock, T., Schmidt, F., Fuchs, M., Brunner, B., Zacharias, F., Giordano, P. R., Konietzschke, R., Sepp, W., Fuchs, S., Rink, C., Albu-Schaffer, A. et Hirzinger, G. (2009). Rollin’ Justin – Mobile platform with variable base. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 1597–1598.
- Brünken, R., Plass, J. L. et Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, volume 38, numéro 1, p. 53–61.
- Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R. B., Covert, M. D., Elliot, L. R. et Redden, E. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance : A meta-analysis. Dans *Proceedings 8th International Conference on Multimodal Interfaces*. p. 108–117.
- Chen, J. Y. C., Haas, E. C. et Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics-Part C (Applications and Reviews)*, volume 37, numéro 6, p. 1231–1245.
- Chen, T. L. et Kemp, C. (2011). A direct physical interface for navigation and positioning of a robotic nursing assistant. *Advanced Robotics*, volume 25, numéro 5, p. 605–627.
- Dahl, T. S. et Boulos, M. N. K. (2013). Robots in health and social care : A complementary technology to home care and telehealthcare ? *Robotics*, volume 3, numéro 1, p. 1–21.
- de Barros, P. G., Lindeman, R. W. et Ward, M. O. (2011). Enhancing robot teleoperator situation awareness and performance using vibro-tactile and graphical feedback. Dans *Proceedings IEEE Symposium on 3D User Interfaces*. p. 47–54.
- Desai, M., Tsui, K. M., Yanco, H. A. et Uhlik, C. (2011). Essential features of telepresence robots. Dans *Proceedings IEEE International Conference on Technologies for Practical Robot Applications*. p. 15–20.
- Drury, J. L., Keyes, B. et Yanco, H. A. (2007). LASSOing HRI : Analyzing situation awareness in map-centric and video-centric interfaces. Dans *Proceedings ACM/IEEE Conference on Human-Robot Interaction*. p. 279–286.

- Drury, J. L., Scholtz, J. et Yanco, H. A. (2003). Awareness in human-robot interactions. Dans *Proceedings System Security and Assurance*. p. 912–918.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, p. 97–101.
- Endsley, M. R. (2000). Theoretical underpinnings of situation awareness : A critical review. Dans Endsley, M. R. et Garland, D. J., *Situation Awareness Analysis and Measurement*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, p. 1–24.
- Ferland, F., Aumont, A., Létourneau, D. et Michaud, F. (2013). Taking your robot for a walk : Force-guiding a mobile robot using compliant arms. Dans *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*. p. 309–316.
- Ferland, F., Létourneau, D., Aumont, A., Frémy, J., Legault, M.-A., Lauria, M. et 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*, volume 1, numéro 2, p. 14–29.
- Ferland, F., Pomerleau, F., Dinh, C. T. L. et Michaud, F. (2009). Egocentric and exocentric teleoperation interface using real-time, 3D video projection. Dans *Proceedings 4th ACM/IEEE International Conference on Human-Robot Interaction*. p. 37–44.
- Fitzgerald, C. (2013). Developing Baxter. Dans *Proceedings IEEE International Conference on Technologies for Practical Robot Applications*. p. 1–6.
- Gonzalez-Jimenez, J., Galindo, C. et Ruiz-Sarmiento, J. (2012). Technical improvements of the Giraff telepresence robot based on users’ evaluation. Dans *The 21st IEEE International Symposium on Robot and Human Interactive Communication*. p. 827 – 32.
- Graf, B., Parlitz, C. et Hagele, M. (2009). Robotic home assistant Care-O-bot 3 product vision and innovation platform. Dans *Proceedings Human-Computer Interaction*. p. 312–320.
- Green, S., Billinghamurst, M., Chen, X. et Chase, J. (2008). Human-robot collaboration : A literature review and augmented reality approach in design. *International Journal of Advanced Robotic Systems*, volume 5, p. 1 – 18.
- Grisetti, G., Stachniss, C. et Burgard, W. (2007). Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Transactions on Robotics*, volume 23, numéro 1, p. 34–46.
- Grondin, F., Létourneau, D., Ferland, F., Rousseau, V. et Michaud, F. (2013). The ManyEars open framework. *Autonomous Robots*, volume 34, numéro 3, p. 217–232.
- Guizzo, E. (2010). When my avatar went to work. *IEEE Spectrum*, volume 47, numéro 9, p. 26–50.
- Hart, S. G. (2006). NASA-task load index (NASA-TLX) : 20 years later. Dans *Proceedings 50th Annual Meeting of the Human Factors and Ergonomics Society*. p. 904–908.

- Hestand, D. et Yanco, H. A. (2004). Layered sensor modalities for improved human-robot interaction. Dans *Proceedings IEEE International Conference on Systems, Man and Cybernetics*. p. 2966–2970.
- Hu, H., Li, J., Xie, Z., Wang, B., Liu, H. et Hirzinger, G. (2005). A robot arm/hand teleoperation system with telepresence and shared control. Dans *Proceedings IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. p. 1312–1317.
- Jain, A. et Kemp, C. C. (2009). Pulling open novel doors and drawers with equilibrium point control. Dans *Proceedings of the IEEE-RAS International Conference on Humanoid Robotics*.
- Janicke, H., Borgo, R., Mason, J. S. D. et Chen, M. (2010). Soundriver : Semantically-rich sound illustration. *Computer Graphics Forum*, volume 29, numéro 2, p. 357–366.
- Johnson, C. A., Adams, J. A. et Kawamura, K. (2003). Evaluation of an enhanced human-robot interface. Dans *Proceedings System Security and Assurance*. p. 900–905.
- Jong, T. (2010). Cognitive load theory, educational research, and instructional design : Some food for thought. *Instructional Science*, volume 38, p. 105–134.
- Keyes, B., Micire, M., Drury, J. L. et Yanco, H. A. (2010). Improving human-robot interaction through interface evolution. Dans *Human-Robot Interaction*. InTech, p. 183–202.
- Kitagawa, M., Dokko, D., Okamura, A. M. et Yuh, D. D. (2005). Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *The Journal of Thoracic and Cardiovascular Surgery*, volume 129, numéro 1, p. 151 – 158.
- Labbe, M. et Michaud, F. (2013). Appearance-based loop closure detection for online large-scale and long-term operation. *IEEE Transactions on Robotics*, volume 29, numéro 3, p. 734 – 745.
- Labonté, D. (2008). *Interface-opérateur à réalité et perspective mixte de téléopération de robot mobile pour une application en téléassistance*. Thèse de doctorat, Université de Sherbrooke, Sherbrooke, Québec, Canada, 190 p.
- Labonté, D., Boissy, P. et Michaud, F. (2010). Comparative analysis of 3D robot teleoperation interfaces with novice users. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, volume 40, numéro 5.
- Lathan, C. E. et Tracey, M. (2002). The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence : Teleoperators and Virtual Environments*, volume 11, numéro 4, p. 368–377.
- Lauria, M., Legault, M. A., Lavoie, M. A. et Michaud, F. (2008). Differential elastic actuator for robotic interaction tasks. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 3606–3611.

- Lee, M. K. et Takayama, L. (2011). "Now, I have a body" : Uses and social norms for mobile remote presence in the workplace. Dans *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. p. 33 – 42.
- Lera, F. J. R., Sierra, J. F. G. et Matellán, C. F. V. (2011). Augmented reality for improving teleoperation of mobile robots. Dans *Xi Workshop en Agentes Físicos (WAF)*.
- Lindeman, R. W. (2003). Virtual contact : The continuum from purely visual to purely physical. Dans *Proceedings of the 47th Annual Meeting of the Human Factor and Ergonomics Society*. p. 2103–2107.
- Lombard, M. et Ditton, T. (1997). At the heart of it all : The concept of presence. *Journal of Computer-Mediated Communication*, volume 3, numéro 2.
- Mast, M., Španěl, M., Arbeiter, G., Štancl, V., Materna, Z., Weisshardt, F., Burmester, M., Smrž, P. et Graf, B. (2013). Teleoperation of domestic service robots : Effects of global 3D environment maps in the user interface on operators' cognitive and performance metrics. Dans *Proceedings 5th International Conference on Social Robotics*. p. 392–401.
- Matthews, T., Fong, J., Ho-Ching, W.-L. et Mankoff, J. (2007). Evaluating non-speech sound visualizations for the deaf. *Behavior and Information Technology*, volume 25, numéro 4, p. 333–351.
- Matthews, T., Fong, J. et Mankoff, J. (2005). Visualizing non-speech sounds for the deaf. Dans *Proceedings The Seventh International ACM SIGACCESS Conference on Computers and Accessibility*. p. 52–59.
- Mendez, I., Jong, M., Keays-White, D. et 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*, volume 72.
- Michaud, F., Boissy, P., Labonté, D., Brière, S., Perreault, K., Corriveau, H., Grant, A., Lauria, M., Cloutier, R., Roux, M. A., Iannuzzi, D., Royer, M. P., Ferland, F., Pomerleau, F. et Létourneau, D. (2010). Exploratory design and evaluation of a homecare teleassistive mobile robotic system. *Mechatronics*, volume 20, numéro 7, p. 751–766.
- Michaud, F., Boissy, P., Labonté, D., Corriveau, H., Granty, A., Lauria, M., Cloutier, R., Roux, M.-A., Lannuzzi, D. et Royer, M.-P. (2007). *Telepresence Robot for Home Care Assistance* (Rapport technique SS-07-07). AAAI Spring Symposium, 50 - 55 p.
- Milgram, P., Zhai, S., Drascic, D. et Grodski, J. (1993). Applications of augmented reality for human-robot communication. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 1467–1472.
- Mizumoto, T., Nakadai, K., Yoshida, T., Takeda, R., Otsuka, T., Takahashi, T. et Okuno, H. (2011). Design and implementation of selectable sound separation on the Texai telepresence system using HARK. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 2130–2137.
- Nakadai, K., Takahashi, T., Okuno, H., Nakajima, H., Hasegawa, Y. et Tsujino, H. (2010). Design and implementation of robot audition system 'HARK' open source software for

- listening to three simultaneous speakers. *Advanced Robotics*, volume 5, numéro 6, p. 739–761.
- Nielsen, C. W., Goodrich, M. A. et Ricks, R. W. (2007). Ecological interfaces for improving mobile robot teleoperation. *IEEE Transactions on Robotics*, volume 23, numéro 5, p. 927–941.
- Nielsen, J. (1993). Usability heuristics. Dans *Usability Engineering*. Academic Press, Boston, p. 115–163.
- Pamungkas, D. et Ward, K. (2013). Tele-operation of a robot arm with electro tactile feedback. Dans *Proceedings IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. p. 704–709.
- Quigley, M., Conley, K., Gerkey, B. P., Faust, J., Foote, T., Leibs, J., Wheeler, R. et Ng, A. Y. (2009). ROS : An open-source Robot Operating System. Dans *ICRA Workshop on Open Source Software*.
- Reiley, C., Akinbiyi, T., Burschka, D., Chang, D., Okamura, A. et Yuh, D. (2008). Effects of visual force feedback on robot-assisted surgical task performance. *The Journal of Thoracic and Cardiovascular Surgery*, volume 135, numéro 1, p. 196–202.
- Ricks, B., Nielsen, C. W. et Goodrich, M. A. (2004). Ecological displays for robot interaction : A new perspective. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 2855–2860.
- Robinson, D. (2000). *Design and Analysis of Series Elasticity in Closed Loop Actuator Force Control*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Boston.
- Robles-De-La-Torre, G. (2006). The importance of the sense of touch in virtual and real environments. *IEEE MultiMedia*, volume 13, numéro 3, p. 24–30.
- Sasaki, Y., Kaneyoshi, M., Kagami, S., Mizoguchi, H. et Enomoto, T. (2009). Daily sound recognition using pitch-cluster-maps for mobile robot audition. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 2724–2729.
- Scholtz, J., Young, J., Drury, J. L. et Yanco, H. A. (2004a). Evaluation of human-robot interaction awareness in search and rescue. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 2327–2332.
- Scholtz, J., Young, J., Yanco, H. et Drury, J. (2004b). Where am I? Acquiring situation awareness using a remote robot platform. Dans *Proceedings IEEE Conference on Systems, Man and Cybernetics*. p. 2835–2840.
- Schroeder, W., Martin, K. et Lorensen, B. (2006). *Visualization Toolkit : An Object-Oriented Approach to 3D Graphics, 4th Edition*, 4^e édition. Kitware.
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. et Goodrich, M. (2006). Common metrics for human-robot interaction. Dans *Proceedings ACM Conference on Human-Robot Interaction*. p. 33–40.

- Takayama, L. (2009). Toward a science of robotics : Goals and standards for experimental research. Dans *Robotics : Science and Systems (RSS) Workshop on Good Experimental Methodology in Robotics*.
- Tavakoli, M., Aziminejad, A., Patel, R. V. et Moallem, M. (2006). Methods and mechanisms for contact feedback in a robot-assisted minimally invasive environment. *Surgical Endoscopy*, volume 20, numéro 10, p. 1570–9.
- Valin, J., Vos, K. et Terriberry, T. (2011). Definition of the opus audio codec. <http://tools.ietf.org/html/rfc6716>.
- Vozar, S. et Tilbury, D. M. (2011). Augmented reality user interface for mobile ground robots with manipulator arms. Dans *Proceedings of the SPIE - The International Society for Optical Engineering*. p. 78780M–78780M–9.
- Williamson, M. (1995). *Series Elastic Actuators*. Master’s thesis, Massachusetts Institute of Technology, Cambridge, Boston.
- Witus, G., Hunt, S. et Ellis, R. (2007). Experiments in augmented teleoperation for mobile robots-I. Dans *Proceedings of the SPIE - The International Society for Optical Engineering*. p. 65610 – 1.
- Woodall, W. et Bevly, D. (2012). Using the Microsoft Kinect for 3D map building and teleoperation. Dans *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*. p. 1054–1061.
- Wyrobek, K., Berger, E., der Loos Van, H. et Salisbury, K. (2008). Towards a personal robotics development platform : Rationale and design of an intrinsically safe personal robot. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 2165–2170.
- Yanco, H. A., Baker, M., Casey, R., Keyes, B., Thoren, P., Drury, J. L., Few, D., Nielsen, C. et Bruemmer, D. (2006). Analysis of human-robot interaction for urban search and rescue. *Symposium A Quarterly Journal In Modern Foreign Literatures*, volume 23, numéro 5, p. 1001–1012.
- Yanco, H. A., Drury, J. L. et Scholtz, J. (2004). Beyond usability evaluation : Analysis of human-robot interaction at a major robotics competition. *Human-Computer Interaction*, volume 19, numéro 1-2, p. 117–149.

