Workshop:
Towards Standardized Experiments in HRI

Organisers: Nicole Mirnig, Paolo Barattini, Dimitris Chrysostomou, Lars Dalgaard, Maria Elena Giannaccini, Manuel Giuliani, Tamas Haidegger, Adriana Tapus, Gurvinder Singh Virk

http://clawar.org/towards-standarised-experiments-in-human-robot-interactions/

This workshop is supported by the EU Robotics Topical Group Standardization

Objectives

This workshop aims to advance the topic of standardization of robot experiments in Human-Robot Interaction (HRI) scenarios. The workshop follows up on previous workshops focusing on international robot standardization and benchmarking in the areas of industrial, medical, and personal care robots (ICRA 2013, HRI 2014, IROS 2014 and ERF 2015). While the R&D community produces great amounts of scientific outputs on HRI, the results are scattered in a myriad of different approaches and ways of performing and testing the interaction; metrics which have been used include efficacy, effectiveness, user satisfaction, emotional impact and social components. The main consequence is that results are not comparable and benchmarking of the various approaches proposed is not possible. The community is still missing consensus tools to benchmark robot products (robot producer/industrial perspective) and robot applications (research/academic perspective). Modes are required for the standardized assessment of robot products and applications in use in terms of safety, performance, user experience, and ergonomics. The benefit of agreed approaches and methods to the assessment of HRI is the production of results, so called “normative” data in the standardization community, meaning that they have been formulated via wide consultation in an open and transparent manner. In this way, the results become widely acceptable, and can be exploited for the creation of international quality norms and standards which in turn would mean measurable robot performances in terms of HRI. We would like to draw from a wide set of experts from the industry, academy and standardization to focus on the key areas of industrial, personal care and medical robots. Together, we will work on establishing benchmarking scenarios and identifying suitable metrics common to HRI in these central and related robotics domains. As a result we aim for providing metrics and scenarios for robot producers and HRI researchers to evaluate their robots and robot systems and setups on a comparable level. Reproducible and comparable results and interoperable systems should be a long-term goal will be a valuable contribution to our community. Within the WS, there will be paper presentations, posters and hands-on panel discussion sessions in the WS and we invite interested WS participants to present their inputs as appropriate.

Topics of interest

- Benchmarking robots
- Benchmarking HRI

- Safety and performance for robots & HRI
- Performance indicators for HRI
- Industrial, medical, and personal care robots
- Human-robot collaboration

Invited Speakers:

1. Paolo Barattini, Kontor 46, Italy
   Update from previous workshops from the TG Standardization
2. Gurvinder Virk, University of Gävle, Sweden
   Update on international standardization projects and relevance in Europe
3. Sven Wachsmuth, University of Bielefeld, Germany
   Standardization in Robocup@home
4. Björn Matthias, ABB
   The role of collision experiments in safety standardization and in the characterization of collaborative robots, systems and applications
5. Claudia Pagliari, University of Edinburgh, UK
   Taxonomies and operational definitions of interactive robot applications
6. Andrea Bonarini, Politecnico di Milano
   Benchmarking medical HRI, robot therapy for the disabled
7. Reinhard Lafrenz, TUM, Germany
   Standards and experiments in Echord++
8. Stefan Profanter, Fortiss, Germany
   Industrial HRI experiments in the SMErobotics project
9. Hooman Samani, National Taipei University, Taiwan
   Possible standard evaluation method for HRI
10. Laurence Devillers, Paris-Sorbonne University, France
    Affective and social spoken dialog in robotics: evaluation of user engagement
11. Agnieszka Wykowska – Technische Universität München (TUM), Germany
    The method and objective measures of social cognitive neuroscience for reliability of results in HRI research

Program

08:30 – 09:00	Opening
09:00 – 10:00	Talks 1
10:00 - 10:30	Coffee Break
10:30 – 12:00	Hands-on, panel experiment drafting session
12:00 – 12:30	Poster session
12:30 - 14:00	Lunch Break
14:00 – 15:30	Talks 2
15:30 - 16:00	Coffee Break
16:00 – 17:30	Hands-on, panel benchmarking session
17:30 – 18:00	Wrap up
18:00	End

Organisers

Nicole Mirnig
Center for Human-Computer Interaction, University of Salzburg, Austria
nicole.mirnig@sbg.ac.at

Paolo Barattini
Kontor 46, Torino, Italy
paolo.barattini@kontor46.eu

Dimitris Chrysostomou
Aalborg University, Denmark
dimi@m-tech.aau.dk

Lars Dalgaard
Danish Technological Institute, Denmark
ldd@teknologisk.dk

Maria Elena Giannaccini
University of Salford, Manchester, United Kingdom
M.E.Giannaccini@salford.ac.uk

Manuel Giuliani
Center for Human-Computer Interaction, University of Salzburg, Austria
manuel.giuliani@sbg.ac.at

Tamas Haidegger
Obuda University, Hungary
haidegger@uni-obuda.hu

Adriana Tapus
Robotics and Computer Vision Lab, ENSTA-ParisTech, Paris, France
adriana.tapus@ensta-paristech.fr

Gurvinder Singh Virk
University of Gävle, Sweden
gurvinder.virk@hig.se

Contact

Nicole Mirnig
Center for Human-Computer Interaction
University of Salzburg, Austria
Sigmund-Haffner-Gasse 18
5020 Salzburg, Austria
nicole.mirnig@sbg.ac.at

Abstract Submissions - Table of Contents

Functional and Non-Functional Expressive Dimensions: Classification of the expressiveness of the face of humanoid robots

François Ferland and Adriana Tapus
Robotics and Computer Vision Laboratory
ENSTA-ParisTech
Palaiseau, France
Email: {firstname.lastname}@ensta-paristech.fr

***Abstract*— In Human-Robot Interaction (HRI), an important quantity of work has been done in investigating the reaction of people toward robots that can express emotions. However, the large variability of expression modalities (e.g., gaze, gestures, speech modulation) available can make comparisons between experimental results obtained on different robots difficult. We believe that developing a common taxonomy to describe these modalities would contribute to the standardization of HRI experiments. Starting with facial expressions from humanoid robots, this short paper aims at encouraging discussion toward a classification system for expression modalities of robots commonly used in HRI.**

I. Introduction

Motivated by the ageing population in many countries around the world, there have been multiple projects for the development of assistive robots and other intelligent devices, ranging from small assistant-like software for mobile phones to mobile robots embedded in elder care living facilities. One of the objectives pursued by the Human-Robot Interaction (HRI) and social robotics community is the development of natural, human-like behaviors for intelligent autonomous systems. These systems are often embodied by robots or virtual agents on a screen, both usually designed to have a human-like appearance. This enables system designers to imitate humans' behaviors so as to create expressive modalities, such as displaying a smile for joy or nodding in agreement of a statement.

Many studies have been conducted on the effectiveness of expression modalities in various interaction settings. For instance, proxemics have an important impact in human-robot interaction. A robot located too close to a person might induce stress [1], and people who dislike a robot tend to keep a greater distance from them, especially to compensate for an increasing eye contact from the robot gaze [2]. Having a directed or averted gaze also has an influence on the minimal comfortable interaction distance, increasing or decreasing depending on the gender of the person [3]. Similarly, a robot with motion-oriented gaze behavior can be perceived as more engaging and human-like [4], and it has been shown that a robot matching the personality of its users by adopting its gaze behavior can have a positive impact in a puzzle-solving task [5]. Furthermore, the appearance of the robot also has an impact on its perceived effectiveness, as it has been observed that people systematically preferred robots for jobs when the human-likeness of the robot matched the sociability required in those jobs [6]. In the same vein, it has been observed that the height of a telepresence robot has an impact on persuasion and dominance, which matches human-human interactions [7].

This illustrates how different robots, even with the same objectives, can have a different impact depending on not only their behavior, but also their capabilities and physical shape and appearance. To achieve standardization in HRI experiments, using identical robots would avoid introducing unwanted factors. Obviously, this is not possible in a practical sense. Except for a few popular robots like NAO from Aldeberan Robotics or PR2 from Willow Garage, there are not many other interactive robots that achieved the kind of commercial success necessary to make this feasible. Furthermore, research groups that are more interested in the design aspect of interactive robots understandably prefer to conduct experiments with their own unique systems.

However, we posit that there is an alternative to having researchers use identical robots. To facilitate comparisons between different robots used in similar HRI experiments, we propose the development of a classification system for the expressiveness of humanoid robots. As a start, this short paper illustrates an example of classification system that could be developed for face expressiveness, based on a small selection of robots that are used in HRI research. In this work, face expressiveness refers to animated features of the face such as eyebrows or the mouth, but also to gaze and general head motion.

This paper is structured as follows: Section II describes a selection of robots that have been used in HRI research, focusing on their capabilities for facial expressions; Section III proposes a classification system that can be used as a shorthand for describing these capabilities; finally, Section IV concludes the paper with suggestions on how this classification could be extended to other features of interactive robots.

II. Robots commonly used in HRI

Table I lists the features of the robots that were selected for this paper. This selection is not meant to be exhaustive. Instead, robots were selected to show sufficiently different ways of reproducing human features and behaviors, and thus help in the development of a classification system. These robots are Baxter from Rethink Robotics, Wakamaru from

Mitsubishi Heavy Industries, NAO from Aldebaran Robotics, IRL-1 from Université de Sherbrooke [8], M-1 from Meka Robotics, and EMYS developed by the Wroclaw University of Technology for the EU FP7 LIREC Project [9]. From this set of robots, we noted the list of features such as having an articulated mouth that are found on at least two robots. The following sub-sections discuss the differences between robots for each of those features.

A. Eye gaze

For eye gaze, we observe a wide variety of solutions, from completely virtual (Baxter) to unique features such as an extra degree of freedom (DOF) for eye popping (EMYS). Some of them have fixed orientation (Wakamuru, NAO), and rely entirely on the head to direct the gaze. An advantage of having eyes that can be rotated separately from the head is a faster reaction time to new gaze targets. Except for one (Wakamaru), all have a pupil distinguished from the rest of the eye, with one that are actually made from cameras (Meka M-1), which indicates that the robot actually sees from its eyes. This is not the case for all robots. For instance, hiding the eyes of the NAO does not block its cameras located at the top of its head and its mouth.

B. Head orientation

Except for one (Baxter), all robots can rotate their head. Only two (Wakamaru, Meka M-1) has an extra DOF for roll angle.

C. Eyebrows

Only one robot has physical eyebrows in this selection (IRL-1). However, EMYS can use the tilt angle of its top plate to represent the brow on a wide range of motion. Furthermore, the Baxter has them on its display.

D. Eyelids

Only two robots have mechanical eyelids (Meka M-1, EMYS), while one can display them on a screen (Baxter). While EMYS only has top eyelids, they can be rotated in addition to being closed. The rotation of the eyelids can play the role of frowning eyebrows. The Baxter also includes eyelids in its display.

E. Mouth

Only one robot has a mechanical mouth (IRL-1), represented by two flexible tubes moved by the rotation of four mouth corners. However, EMYS has a bottom plate that can act as a jaw. As with other features, the Baxter could be programmed to display a mouth.

III. Classification of Face Expressiveness

From the set of features described in Section II, we can attempt to generalize them and construct a classification system. First, one dimension can be extracted: functional or non-functional, purely expressive features. For instance, the gaze can be used in a neutral fashion to change the orientation of sensors and indicate where the robot is looking. However, eyebrows usually do not have any other function than expressing an emotion. We can select at least two categories of features that have functional features: gaze and mouth. Table II summarize the classification system for these features.

A. Gaze

Gaze, achieved by rotating the head and/or the eyes, serve both functional and expressive requirements. On robots such as Meka M-1, it is necessary to direct its cameras. For gaze, we propose to classify it on a spectrum: G0 for robots without gaze, G1 for a fixed gaze, G2 for an oriented gaze, and G2+ for a gaze that can be oriented independently from the head. For robots that use a display for the gaze, we add the ”V” suffix. So, a robot with a head that can have an orientation but using a display for its eyes would receive the G2+V classification. In Sec. II, robots are either G2 (Wakamaru, NAO, EMYS), G2V (Baxter), or G2+ (IRL-1, Meka M-1).

B. Mouth

The mouth can be also seen as a functional element. If its motion is synchronized with speech generation, it can be used as a visual cue, for instance to identify which robot is speaking in a close group, or even detect a speaker that is malfunctioning or with a volume set too low. For the mouth, we propose a spectrum similar to the one used for the gaze: M0 for robots without a mouth (Standard Baxter, Wakamaru), M1 for a fixed mouth (NAO), and M2 for mouths with one or more DOF (EMYS, IRL-1). For robots using a display for their mouth, we also add the V suffix (Meka M-1).

C. Non-functional expressive features

Purely non-functional expressive features are harder to classify from a DOF standpoint. For instance, if we try to classify the rotation of the eyelids of EMYS, it can be associated to the angle of the eyebrows. Since most expressions are largely inspired from human ones, we propose to use Action Units (AUs) of the Facial Action Coding System (FACS) [10], which are used for emotion recognition. For instance, happiness can be recognized as a combination of AU6 (Cheek Raiser) and AU12 (Lip Corner Puller), while surprise is a combination of AU1 (Inner Brow Raiser), AU2 (Outer Brow Raiser) AU5 (Upper Lid Raiser), and AU26 (Jaw Drop). Thus, a robot face expressiveness could be described with respect to the set of AUs that it can reproduce. For instance, while AU1 and AU2 can be performed by IRL-1, they cannot be combined, which is a feature achieved by the eyelids and the top plate of EMYS. Similarly, while AU26 is hard to achieve by all robots except EMYS, AU12 can only be achieved by IRL-1 and Meka M-1.

D. Non-human features

Finally, there are some features that do not have a direct human equivalent, but can still successfully be used to convey emotions. This is the case of the eye popping DOF of EMYS, which can effectively express a carton-inspired

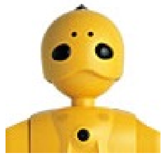
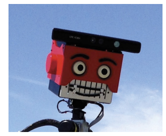
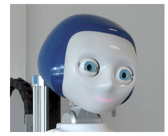
	 Baxter	 Wakamaru	 NAO	 IRL-1	 Meka M-1	 EMYS
Eye gaze	Virtual	Fixed	Fixed	Pan, tilt	Pan, tilt	Fixed
Head orientation	Fixed	Pan, tilt, roll	Pan, tilt	Pan, tilt	Pan, tilt (1), roll	Pan, tilt (1)
Eyebrows	Virtual	None	None	Yes	None	None (2)
Eyelids	Virtual	None	None	None	Yes	Yes, with orientation
Mouth	None (3)	None	None (4)	Yes	LED Matrix	None (2)
Gaze classification	G2V	G2	G2	G2+	G2+	G2
Mouth classification	M0 (5)	M0	M1	M2	M2V	M2

TABLE I: Description of facial expression features for six robots and their classification. (1) Both Meka M-1 and EMYS feature an additional tilt angle on the neck. (2) While EMYS does not explicitly have eyebrows or a mouth, the top and bottom plate can act as them. (3) While Baxter does not display a mouth on its screen, it could be programmed to. (4) NAO has a hole for a camera where one would expect a mouth. (5) Baxter could be programmed to display a mouth, which would receive a M2V classification.

Gaze	None	Fixed	Oriented	Head-independent
	G0	G1[V]	G2[V]	G2+[V]
Mouth	None	Fixed	Articulated	
	M0	M1[V]	M2[V]	

TABLE II: Classification of functional features. [V] is an optional suffix noting that the feature is rendered by a display.

surprise, or variations of Meka M-1 with ears that can change their orientation, such as Simon of the Socially Intelligent Machines Lab at Georgia Tech. Since these features are fairly unique in humanoid robots and thus not generalized, it is not currently possible to extract a specific classification for them.

IV. Conclusion

This short paper presents the first steps in a proposal for classifying the face expressiveness of robots found in the HRI community. We believe that having a common set of terms for describing the expressive features of robot faces will help the description of standardized HRI experiments.

The underlying goal of this work is to arrive at a full classification of whole-body expressiveness, as having a face is only one modality out of many that an autonomous intelligent system can use to interact with humans. For instance, if proxemics depend on the distance between a robot and a person interacting with it, the way a robot moves to change this distance can have an impact on the perception of the person. A wheeled robot, omnidirectional or not, does not move the same way as a legged one. Furthermore, the spoken expression of emotions, whether by content (e.g., stating "I am happy") or by the modulation of speech is another important component of expressiveness.

By generalizing functional and non-functional expressive features of each subsystem of an autonomous robot and evaluating them on a scale of their human-likeness, we will have a complete taxonomy for the expressiveness of interactive robots. As a result, HRI experiment results involving expressiveness will be easier to understand and compare between each other.

Acknowledgment

This work has been partially funded by EU Horizon2020 ENRICHME project grant agreement no. 643691C.

References

[1] M. Sorostinean, F. Ferland, and A. Tapus, "Reliable stress measurement using face temperature variation with a thermal camera in human-robot interaction," *Submitted to the IEEE-RAS Conference on Humanoid Robotics*, 2015.

[2] J. Mumm and B. Mutlu, "Human-robot proxemics: physical and psychological distancing in human-robot interaction," in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 331–338.

[3] L. Takayama and C. Pantofaru, "Influences on proxemic behaviors in human-robot interaction," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 5495–5502.

[4] M. Sorostinean, F. Ferland, T.-H.-H. Dang, and A. Tapus, "Motion-oriented attention for a social gaze robot behavior," in *Proceedings of the International Conference on Social Robotics*. Springer, 2014, pp. 310–319.

[5] S. Andrist, B. Mutlu, and A. Tapus, "Look like me: Matching robot personality via gaze to increase motivation," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 3603–3612.

[6] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in *Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on*. IEEE, 2003, pp. 55–60.

[7] I. Rae, L. Takayama, and B. Mutlu, "The influence of height in robot-mediated communication," in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2013, pp. 1–8.

[8] F. Ferland, D. Létourneau, A. Aumont, J. Frémy, M.-A. Legault, M. Lauria, and F. Michaud, "Natural interaction design of a humanoid robot," *J. Human-Robot Interaction*, vol. 1, no. 2, pp. 14–29, 2012.

[9] J. Kedzierski, R. Muszynski, C. Zoll, A. Oleksy, and M. Frontkiewicz, "Emys—emotive head of a social robot," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 237–249, 2013.

[10] P. Ekman and W. V. Friesen, *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press, 1978.

Using Standard Test Methods for Response Robots To Evaluate Remote Human-Robot Interaction

Adam Norton
New England Robotics Validation and Experimentation (NERVE) Center
University of Massachusetts Lowell
Lowell, Massachusetts 01854
anorton@cs.uml.edu

Holly A. Yanco
Department of Computer Science
University of Massachusetts Lowell
Lowell, Massachusetts 01854
holly@cs.uml.edu

***Abstract*—Test beds for evaluating human-robot interaction (HRI) are generally developed to fit a particular experiment, lacking a common set of tasks. The standard test methods for response robots specified through ASTM E54.08.01 are used to evaluate robot mobility, manipulation, sensors, and operator proficiency. There are four test methods that focus on proper situation awareness (SA): Line Following, Center in Alleys, Align Edges, and Pan Tilt Zoom. These tests serve as candidates for standardized HRI experimental set ups as they simply and effectively capture many characteristics of the robot, interface, and operator. We discuss an example data set of test method performance and how they can be used to evaluate HRI.**

I. Introduction

There currently exists no standard experimental set-ups for evaluating human-robot interaction (HRI). In order for such experiments to be standardized, they must be able to be applied to many different robots and interfaces of varying capabilities. The test metrics must be broadly applicable such that the performance of the robot, interface, and operator are captured.

The standard test methods specified through ASTM E54.08.01 Committee on Homeland Security Applications; Operational Equipment; Robots [1] are used to evaluate response robot capabilities and operator proficiency. They have been used to evaluate many teleoperated robots and to train end-users in urban search and rescue (USAR) and explosive ordinance disposal (EOD) domains. The test methods are performed without line-of-sight, meaning the operator must rely on the interface as they would in a real scenario. The settings for each test method are malleable such that they can be tuned to the characteristics of the robot. All aspects of the system (e.g., robot configuration, interface modalities, operator knowledge of system) have in impact on performance.

The test methods in the Maneuvering suite highlight the capabilities of the robot, interface, and operator at maintaining situation awareness (SA). Due to the test methods' malleability and holistic nature, they are good examples of potential standardized HRI experiments. To this end, we discuss a path forward for using them as such.

II. Related Work

During the early development of the E54.08.01 standard test methods, physical and virtual implementations of test arenas were used [2] [3]. The arenas used versions of the test methods in an operational scenario, combining elements to form the challenges of each arena. These experiments have also been used to perform iterative HRI designs.

A test bed for evaluating HRI with EOD robots [4] distilled a set of tasks based on existing law enforcement training programs and real world incidents. The number of button presses and mode changes needed to perform each task were used to evaluate the HRI. From this it is suggested that the information required from the interface to perform each task be defined to determine if an operator can access it properly.

One of the four lessons learned from a longitudinal study of real world USAR events and training exercises highlighted that SA was a very prevalent issue with the HRI [5], citing that half of the operation time is spent gaining SA. These interactions could be evaluated using many of the common metrics for HRI [6], such as assessing the accuracy of mental models, had they occurred in a more controlled scenario.

III. Standard Test Methods

There are four test methods in the Maneuvering suite:

- **LF**: Line Following (ASTM E2829)
- **CA**: Center in Alleys
- **AE**: Align Edges
- **PTZ**: Pan Tilt Zoom (ASTM WK33261)

An image of each test method can be seen in Figure 1. Each test method can be adjusted based on the characteristics of the robot and interface to allow for equally difficult challenges between platforms. Any teleoperated or semi-autonomous mobile robot with at least one camera can be used.

The performance metric used for the test methods is the number of tasks completed per minute, or rate of advance. For LF, CA, and AE, one task refers to traversing from one end of the test apparatus to the other and back, with some type of obstacle to be negotiated in between. The first half of a task is performed while traversing forward and the second half in reverse. For PTZ, one task refers to viewing near and far acuity targets in the apparatus. Each test has its own rules for fault conditions; if a fault occurs then that task is not counted and must be repeated. The time on task continues to increase, resulting in a decreased rate in advance.

A. Line Following (LF)

The operator drives the robot over a figure-8 line on the ground, of which each end falls in one of the apparatus' end zones. The line must remain underneath the robot's body while traversing, meaning the operator must maintain a view of the

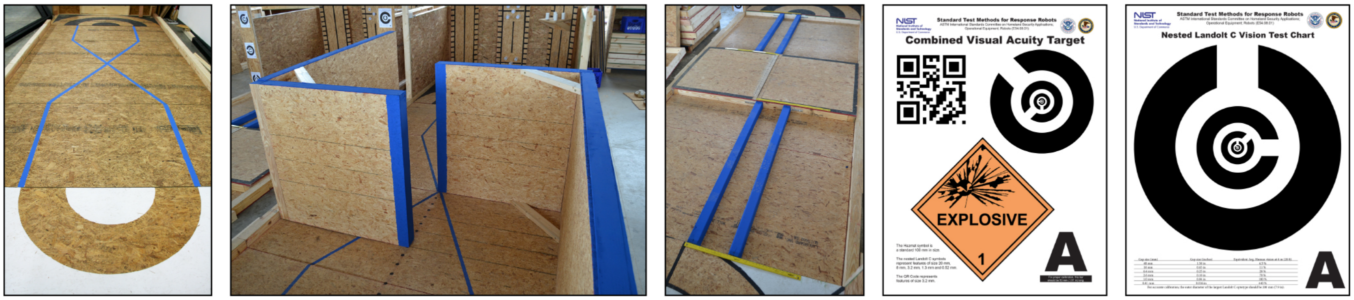


Fig. 1. Images of the standard test methods. Left to right: Line Following, Center in Alleys, Align Edges, and Pan Tilt Zoom near and field targets.

line with respect to the robot. The figure-8 shape forces the operator to match the direction of the line with the orientation of the robot. If the robot drives off of the line such that it is visible outside of the robot's body then a fault is incurred.

B. Center in Alleys (CA)

The operator drives the robot between two walls that form a confined passageway sized to the turning diameter of the robot. The passageway is perpendicular to a straight path between the end zones, requiring the operator to turn the robot while traversing through. The walls are attached to the apparatus using vertical barrel bolts into the floor such that if they are bumped slightly by the robot then they will remain in place, but hard collisions will move them. If the walls are moved due to a robot collision then a fault is incurred.

C. Align Edges (AE)

The operator drives the robot over two rails parallel to a straight path between the end zones. The distance between the outside edges of the rails is set to match that of the robot's wheels or tracks. Two sets of rails must be traversed; one on the right side of the lane and the other on the left. A platform in between each set allows the operator to orient the robot's approach to the next set. If the robot falls off of the rails, then a fault is incurred.

D. Pan Tilt Zoom (PTZ)

While staying in a fixed location, the operator points the robot's camera(s) at targets with visual acuity artifacts. The near and far field targets are each labeled A-J; the operator alternates between looking at a near field target and then its corresponding far field target. The visual acuity artifacts used are Landolt C eye charts, which have concentric circles with cuts in them at varying orientations. Based on the robot's camera(s) resolution combined with the interface's display resolution, the operator is able to identify the orientation of the Landolt C eye charts to the level of acuity that is achievable. The robot is allowed to rotate in place if necessary.

IV. Situation Awareness Factors

Each test method is defined by the level 1, 2, and 3 SA (as defined in [7]) it requires to be performed in Table I. The ability of an operator to acquire and maintain these SA elements is influenced by a variety of characteristics of the robot, its interface, and the operator.

A. Robot Characteristics

For CA and AE, the dimensions of the footprint of the robot will affect how the walls and rails are positioned, respectively. Many robots in this domain use manipulators (generally on top of the base) and articulators (on the front and/or back of the base) that increase their overall size profile. If these components are able to be moved such that the footprint of the robot is made smaller (i.e., closer to the center of the robot's volume), it will aid in completing CA. The robot may also be tethered, most commonly on a motorized spool. These components introduce additional SA of the robot's status that the operator must maintain.

Teleoperated robots require at least one or two cameras that provide forward and rear facing views, but the number of cameras, where they are located on the robot, and their individual qualities (e.g., fixed or dynamic, field of view) can vary. Exocentric cameras located above the robot's body and provide an outside view of it have also been shown to increase spatial reasoning [8]. These cameras may be located on the robot's manipulator or a vertical pole referred to as a mast. The ability to pan, tilt, and zoom these cameras around to view the body of the robot (CA and AE) and the environment around it (PTZ) is also beneficial. Alternatively, for PTZ, one or more fixed cameras can used, but the robot will have to rotate in place in lieu of a rotational degree of freedom.

B. Interface Characteristics

The interface used by the operator to control the robot (specifically the information it provides and the operator's knowledge of it) is the largest contributor to performance.

Test	Level 1 SA	Level 2 SA	Level 3 SA
Line Following	Local environment (line) underneath the front and back of the robot	Alignment deviation of the line from underneath the robot	How to adjust the robot's position to maintain alignment while traversing
Center in Alleys	Local environment (walls) around the outside of the robot's volume	Distance from the walls to the robot	How to adjust the robot's positioning to avoid collisions while traversing
Align Edges	Local environment (rails) underneath the outside edges of the robot	Alignment deviation from the robot's position to the edge of the rails	How to adjust the robot's positioning to maintain alignment while traversing
Pan Tilt Zoom	Local (near field targets) and global (far field targets) environment around the robot	Proper positioning of the robot's position and its camera, and the pan, tilt, and zoom settings of its camera(s)	How to adjust the robot's camera and/or positioning to decipher the environment

TABLE I. The specific level 1, 2, and 3 SA (as defined by [7]) that must be maintained to perform each test method.

Robot	Exocentric Camera	Panning DOF	Manipulator	Articulators	Tether Option	Camera Presentation	Interface Pose Info	LF	CA	AE	PTZ
A	n/a	Body rotate	n/a	Rear	No	Single, multiple	Side	*	*	*	0.62, 0.82
B	Mast, manipulator	Body rotate	5 DOF	n/a	No	Single	Side	1.3	2.5, 2.35, 2.45	0.8	1.1
C	Manipulator	Body rotate	4 DOF	Front	No	Single, multiple	Side	1.4	2.0	0.3	0.83
D	Manipulator	Body rotate	5 DOF	Front	No	Single, multiple	Isometric	2.1	3.3	0.7	1.7
E	Mast, manipulator	Body rotate	5 DOF	n/a	No	Single	Side	*	*	*	*
F	Mast, manipulator	Camera control	8 DOF	Front	Yes	Single, multiple	Isometric	*	0.13	0.1, 0.14	*
G	Manipulator	Camera control	7 DOF, telescoping limbs	Front, rear	No	Single, multiple	Isometric	*	*	*	*
H	Mast, manipulator	Camera control	4 DOF	n/a	No	Single, multiple	Isometric	*	0.5, 0.65	0.25	2.0, 1.82
I	Mast, manipulator	Camera control	6 DOF; telescoping limbs	Front, rear	Yes	Single, multiple	Side	*	0.46, 0.22, 0.5, 0.27	*	*

TABLE II. ROBOT AND INTERFACE CHARACTERISTICS THAT ARE PERTINENT TO HRI AND EXAMPLES OF PERFORMANCE DATA IN EACH OF THE TEST METHODS. EACH PERFORMANCE METRIC IS A RATE OF ADVANCE, MEANING THE NUMBER OF TASKS COMPLETED PER MINUTE. * INDICATES THAT THE PERFORMANCE DATA FOR THAT ROBOT'S PERFORMANCE IN THE TEST METHOD WAS NOT AVAILABLE AT THE TIME OF PUBLICATION.

Input devices generally use at least one joystick of some kind in addition to a series of buttons or switches. The sensitivity of the joystick, complexity in changing control modes (e.g., navigating a series of menu screens vs. flipping a switch), and latency between command and feedback can affect the operator's ability to perform the test. If a system employed automatic direction reversal (ADR), which has the system maintain the orientation of its control when driving in reverse, it could improve performance [9] in LF, CA, and AE.

To perform CA, the robot's footprint must be reduced by adjusting the its manipulators and/or articulators. Some interfaces offer predefined poses that can be selected, which is beneficial for robots with many degrees of freedom. A visualization of the robot's pose is also common, either as a side profile or isometric representation, which can increase the operator's SA of the robot's status [10].

The presentation of camera views can vary greatly between interfaces. Most systems are able to display full screen views of a single camera if desired. If multiple cameras are used, different options for picture-in-picture are generally available, such as a smaller display overlaid in the corner of a larger display (referred to as the "rear view mirror"), or displaying many views at once. Cameras may also be displayed in panoramic to provide a very wide field of view, which may aid in performing CA and AE. Local distance sensors could reduce collisions [11] while performing CA, although not many have been exhibited on deployed response robots.

C. Operator Characteristics

The operator of the robot must be able to acquire and maintain proper SA of the robot's surroundings and status. Continued use of the test methods is intended to increase an operator's understanding of the robot's capabilities and knowledge of how to control it. Given that poor exhibition of HRI with response robots has been observed during real world scenarios such as during the Fukushima Daiichi disaster response [12], the development of this work is pertinent.

The operator must have an accurate mental model of the robot, particularly if it has a manipulator and/or articulators, when performing CA. Some interfaces do not provide such information, so the operator must mentally update their mental model every time they move the robot. Operators may prefer a system that uses inverse kinematics to control a high degree of freedom manipulator, while others may prefer to control each joint individually. The operator must also have proper spatial awareness of the relationship between the environment features and the robot for all test methods.

Some of these characteristics can be aided or hindered by the interface being used. For instance, if an exocentric camera with a more angled view is used then the distance between the robot and local obstacles can be visually estimated. If one is not available, then the operator must interpret the distances using depth perception, which may be more difficult.

V. EXAMPLE PERFORMANCE DATA

Nine commonly used response robots and their pertinent characteristics are detailed in Table II, along with available test method performance data culled together from a variety of public and private test events. Some test methods have multiple measures with the same robot due to being performed by many operators. The robots have been anonymized by granting each an alphabetic identifier, ordered approximately from smallest (A) to largest (I) to give the letters some meaning.

A. Discussion

A comprehensive HRI study of these test methods has not yet been performed. Due to the settings of each test being tuned to the robot, all robot performance could theoretically be equivalent if they were all operated at the same pace. However, the difference in performance is due to the many varying characteristics of the robots, interfaces, and operators.

In general, smaller robots (A-D) tend to exhibit a higher rate of advance in CA, most likely due to faster traversing speeds. Operators may be more cautious with a larger robot. However, it also may be due to each robot's manipulator if it exceeds the base footprint even when positioned in the smallest form factor possible. It should be noted that some operators have been observed performing CA with the manipulator obstructing their camera view while operating robot I. This may be indicative of a lack of knowledge of how to use the system or is indeed the optimal way of performing the test with that particular system.

The differences between exocentric camera views from a robot's manipulator or mast are captured by the performance in AE, as both views can be used to approach the rails properly and maintain a constant direction of traversal. A manipulator camera generally requires more positioning (4-8 DOF) to swap between views for forward and reverse traversals than that of a mast (1-3 DOF), which may take more time. Some exocentric cameras are only placed high enough to provide a view of one side of the robot's body, which was observed for robots A, E, and H. This may provide enough SA for the operator, but could be insufficient for supporting their mental model.

For PTZ, some robot/operators rely solely on a high resolution camera that is able to pan and tilt without moving the body of the robot at all (robots F-I). Others rotate the robot's body and/or tilt the robot's body with their articulators to make up for missing degrees of freedom (robots A-E). This results in varying control schemes for different robots, and sometimes within the same robot.

In the test methods' current design, the number of faults is generally not reported. If they were, it would give more insight into the exhibited HRI, as many faults could be indicative of poor camera presentation on the interface, the operator's lack of skill at maneuvering the joystick properly, etc.

These test methods have not yet been exercised using robots with autonomous navigation capabilities. Increasing robot autonomy would allow for the evaluation of HRI with respect to sharing SA between the robot and the operator through the interface, including aspects like varying operator interaction frequency (potentially exhibiting "out-of-the-loop" performance [13]) and the communication of failures.

VI. Conclusion and Future Work

These test methods are examples of designs that can be used towards standardized HRI experiments, particularly for robots that operate remotely. The many reasons for performance discussed in the previous section highlight the type of information that can be extracted from using the test methods. The next step is to conduct more comprehensive testing and analyzation that investigates the exhibited HRI, such as:

- Interface settings (e.g., camera presentation, robot speed) used during each test method performance;
- Specific robot actions performed (e.g., change of robot pose, rotation of manipulator or mast camera views);
- Faults incurred on each test method caused by incorrect SA of the robot, its environment, and/or understanding of the interface;
- Varying operator experience levels and how performance changes over increased use of the test method;
- Additional sensors to provide SA (e.g., LIDAR); and
- Trends in the previous points to determine optimal HRI for maintaining each type of SA.

From the proposed testing, a set of common robot, operator, and interface characteristics that correlate with specific faults (e.g., backing into a CA wall due to not switching to the rear-facing camera) can be distilled from the performance data, categorized by the type of SA loss. These can be used to form guidelines for determining broader HRI-specific faults.

Evaluating SA in terms of spatial awareness is an important aspect of HRI with response robots that calls for a confined space apparatus. Test methods for HRI in other domains that use more autonomy or tasks in more structured environments may not require this type of SA and will need additional considerations. For instance, HRI of an autonomous space rover may depend more on proper alert techniques and treatment of historical information. The current test method designs do not necessarily highlight those types of capabilities.

Further experimental set-ups will need to be structured to capture these aspects, while still using the same design principles of the test methods discussed in this paper: malleable settings that enable fairness across many solutions, simple performance metrics that can capture many types of errors, and using baseline tasks to evaluate foundational HRI.

Acknowledgments

This research has been supported in part by the National Institute of Standards and Technology (60NANB14D286). The authors would like to thank Adam Jacoff and Ann Virts from NIST for their development of the test methods and assistance in gathering performance data.

References

[1] ASTM, "Subcommittee E54 on Operational Equipment," 2015, http://www.astm.org/COMMIT/SUBCOMMIT/E5408.htm.

[2] M. Lewis, K. Sycara, and I. Nourbakhsh, "Developing a Testbed For Studying Human-Robot Interaction in Urban Search And Rescue," in *Proceedings of the 10th International Conference On Human Computer Interaction*, 2003, pp. 22–27.

[3] S. Carpin, M. Lewis, J. Wang, S. Balakirsky, and C. Scrapper, "USARSim: A Robot Simulator for Research and Education," in *2007 IEEE International Conference on Robotics and Automation*. IEEE, 2007, pp. 1400–1405.

[4] J. Scholtz, M. Theofanos, and B. Antonishek, "Development of a Test Bed for Evaluating Human-Robot Performance for Explosive Ordnance Disposal Robots," in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. ACM, 2006, pp. 10–17.

[5] R. R. Murphy and J. L. Burke, "Up From the Rubble: Lessons Learned about HRI from Search and Rescue," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 49, no. 3. SAGE Publications, 2005, pp. 437–441.

[6] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common Metrics for Human-Robot Interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*.

[7] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[8] B. Keyes, R. Casey, H. A. Yanco, B. A. Maxwell, and Y. Georgiev, "Camera Placement and Multi-camera Fusion for Remote Robot Operation," in *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics*, 2006, pp. 22–24.

[9] B. Keyes, M. Micire, J. L. Drury, and H. A. Yanco, "Improving Human-Robot Interaction Through Interface Evolution," *Human-Robot Interaction*, 2010.

[10] J. L. Drury, H. A. Yanco, W. Howell, B. Minten, and J. Casper, "Changing Shape: Improving Situation Awareness for a Polymorphic Robot," in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. ACM, 2006, pp. 72–79.

[11] H. A. Yanco, B. Keyes, J. L. Drury, C. W. Nielsen, D. A. Few, and D. J. Bruemmer, "Evolving Interface Design for Robot Search Tasks," *Journal of Field Robotics*, vol. 24, no. 8, p. 779, 2007.

[12] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima *et al.*, "Emergency Response to the Nuclear Accident at the Fukushima Daiichi Nuclear Power Plants Using Mobile Rescue Robots," *Journal of Field Robotics*, vol. 30, no. 1, pp. 44–63, 2013.

[13] M. R. Endsley, "Automation and Situation Awareness," *(Automation and Human Performance: Theory and Applications*.

The IMHG dataset: A Multi-View Hand Gesture RGB-D Dataset for Human-Robot Interaction

Dadhichi Shukla[1], Özgür Erkent[1] and Justus Piater[1]

***Abstract*— Hand gestures are one of the natural forms of communication in human-robot interaction scenarios. They can be used to delegate tasks from a human to a robot. To facilitate human-like interaction with robots, a major requirement for advancing in this direction is the availability of a hand gesture dataset for judging the performance of the proposed algorithms. We present details of the Innsbruck Multi-View Hand Gesture (IMHG) dataset recorded with two RGB-D cameras (Kinect). The dataset includes two types of referencing (pointing) gestures with the ground truth location of the target pointed at, two symbolic gestures, two manipulative gestures, and two interactional gestures. The dataset was recorded with 22 participants performing all eight hand gestures.**

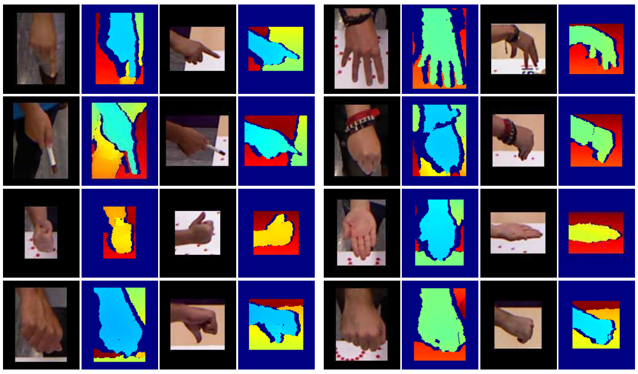


Fig. 1: Multi-view sample images from the IMHG dataset (*images cropped for visualization*). *Top-Bottom*: Finger pointing, Tool pointing, Thumb up (approve), Thumb down (disapprove), Grasp open, Grasp close, Receive, Fist (stop).

I. INTRODUCTION

As robots are moving closer to popular deployment in our daily lives, there is increasing activity in Human-Robot Interaction (HRI) research. Amongst various forms of communications, hand gestures are a highly effective, general-purpose tool for interaction, thanks to the flexibility of the hands. Despite the advancements in HRI methodologies, lack of standardized evaluation hinders their application in sectors such as manufacturing, healthcare, and domestic helper.

Gesture recognition has recently received much attention in the HRI community [1]. We provide a multi-view hand gesture RGB-D dataset for quantitative evaluation of gesture recognition systems in a HRI context. A detailed description of the different types of recorded gestures is given in section II. Some example gestures collected in this dataset are shown in Fig. 1.

A. Motivation

In the early years of human-robot interaction, Dautenhahn [2] indicated that people prefer to interact with robots in a "natural" way. In human-robot interaction, camera-based vision serves as a natural, unencumbered, non-contact, and prop-free mode of interaction. Adopting hand gestures as an interface in HRI opens up new frontiers of research in a wide range of applications (e.g., surgical robotic nurse [3], swarm robots [4], assistive robotics [5], HRI in industrial scenarios [6], etc.). Some insights from human-computer interaction (HCI) may also prove to be valuable to HRI. These notwithstanding, the field of HRI needs to develop its own methods for interaction between a human and a robot. For example, in pointing gestures, a robot has to recognize the gesture as well as to estimate the pointing direction i.e. the pose of the hand. Moreover, it is difficult to perform quantitative evaluation for such *referencing gestures* because of a lack of ground truth.

Many prominent methodologies [7], [8] use gesture recognition systems as an interface to interact with the robot. However, these gesture recognition systems mainly target robot guidance. The gestures are not conceived as commands for the robot to manipulate objects in the environment. This requires capabilities of recognizing human gestures and inferring intent (defined as the manipulative action expected from the robot). To address such scenarios of collaborative manipulation, we propose the novel, publicly-available IMHG dataset.

B. Related work

Human gesture recognition has been studied extensively to interact with robots. Several studies have addressed the topic to create a hand gesture dataset, although most of them are either American Sign Language (ASL) [9], [10] or human-computer interaction [11], [12], [13]. Other datasets [14], [15] capture full-body or upper-body of the participant. The intended domain of applicability of the dataset varies depending on the type of ground truth and chosen set of gestures.

A recent survey by Ruffieux et al. [16] provides a detailed overview of the most recent and/or popular publicly available vision based hand, upper-body, and full-body gesture datasets. Here, we briefly describe previously proposed hand gesture datasets. Kim et al. [11] released an RGB dataset with 9 classes of hand gesture with 900 image sequences. The target task of this dataset was to classify *shapes* as well

[1]The authors are with the Intelligent and Interactive Systems group, Institute of Computer Science, University of Innsbruck, Austria. Corresponding author: Dadhichi.Shukla@uibk.ac.at

as different *motions* at a time. The RGB-D dataset proposed by Liu et al. [13] consists of 10 categories of hand gestures representing shapes like circle, triangle, 'Z', etc. The gestures are performed with three hand postures with different backgrounds and varying illumination conditions. Kurakin et al. [9] proposed a RGB-D dataset of 12 dynamic American sign language (ASL) performed by 10 participants. More recently, Molina et al. [17] released a hand gesture dataset composed of English alphabets, Spanish sign language, and several miscellaneous annotated gestures captured by 11 participants and also, generated synthetically.

The RGB-D dataset collected by Ren et al. [12] can be related to our work. It consists of 10 types of gestures captured in cluttered background. Like the previous works it also addresses hand shape detection. Though the dataset is targeted for HCI applications, it can be applicable to HRI scenarios. One important difference with the former is that the hand gestures in the IMHG dataset are closely related to the semantic content of verbal language. A robot interprets these gestures as the command to be executed to interact with the environment.

We briefly summarize the previous datasets and the proposed IMHG dataset in Table I reviewing various characteristics.

C. Contribution

The main contributions of this dataset are:

- A multi-view RGB-D dataset with 22 participants performing 8 hand gestures.
- Two types of referencing gestures: (1) finger pointing, and (2) tool pointing with an elongated object in hand, are recorded.
- A corpus of 836 test images (704 referencing gestures with ground truth, and 132 other gestures).
- The data acquisition setup can be easily recreated to add new hand gestures in the future.

The IMHG dataset is dedicated to measuring the performance of recognition systems to do gesture understanding – that is, the interpretation of an indicative *robot manipulation* a user wishes to take place.

II. IMHG Dataset Description

There exist many semantic gestures in our daily lives. However, many of them are unsuitable for direct use in human-robot interaction. Nehaniv et al. [18] asserted five major categories of hand gestures in the context of human-robot interaction. One of the five categories, *expressive behaviour* that includes motions of hands, arms, and face, is excluded from the dataset. These types of gestures occur as part of the overall communicative behaviour, but without any specific interactive role of a robot.

According to the study conducted by Nehaniv et al. the eight hand gestures in the IMHG dataset can be grouped into the following four categories:

1. *Referencing hand gestures:* These gestures are used to refer to or to indicate objects (or loci) of interest.We record two types of pointing: (i) finger pointing, and (ii) tool pointing.
2. *Symbolic hand gestures:* The gestures in this category are defined by a prescribed set of interpretations. The static symbolic gestures are analogous to discrete actions on a user interface like Yes/No, Agree/Disagree, etc. In this dataset we capture two types of gestures: (i) Thumb up (approve), and (ii) Thumb down (disapprove), to indicate whether the task was understood/ performed correctly by a robot.
3. *Manipulative hand gestures:* These gestures involve displacement of or interaction with objects (e.g., pushing a box). We record two types of manipulative gestures: (i) *Grasp open* – The robot is to open the hand to grasp an object, (ii) *Grasp close* – The robot is to grasp the object of interest.
4. *Interactional hand gestures:* The gestures in this category are used to regulate interaction with a partner. They can be used to initiate, synchronize, or terminate an interaction. The emphasis on this category of gestures is neither reference nor communication, but for a cooperative action. The dataset includes two types of interactional gestures: (i) *Receive* – The robot is to hand the grasped object to the human, and (ii) *Fist (stop)* – The robot is to stop interacting with the environment. We include the *Fist* gesture because it is easily performed by a human within the camera view of our setup.[1]

Gesture	# Instances	Ground truth
Finger pointing	352	✓
Tool pointing	352	✓
Thumb up	22	×
Thumb down	22	×
Grasp open	22	×
Grasp close	22	×
Receive	22	✓
Fist (stop)	22	×

TABLE II: Summary of IMHG dataset. The ground truth of 'pointing' gesture and 'receive' gesture is the location of the pointed target and the location of the hand, respectively.

Some example images from the dataset are shown in Fig. 2. Each image in the dataset is labelled with its corresponding gesture. To evaluate *referencing* gestures we provide the location of the target pointed at as the ground truth. In the case of *receive* gesture the centroid of the palm is given as the ground truth. We summarize the IMHG dataset in Table II.

III. Dataset Acquisition Scenario

Camera setup

The IMHG dataset was captured using two RGB-D cameras (Kinect) placed orthogonally to record maximum information of the hand gesture. Figure 3a illustrates the multi-

[1] A raised, flat palm vertically facing the camera would certainly constitute a more intuitive *stop* gesture. However, since all other gestures in our scenario are performed low above the workspace, this would require a dedicated camera and eliminate the need for its explicit recognition.

Methods	#Classes	Views	RGB	Depth	Resolution	Pose of finger joints	Available	Application to HRI
Kim et al. [11]	9	T	✓	×	320×240	×	✓	×
Ren et al. [12]	10	F	✓	✓	640×480	×	✓	✓
Kurakin et al. [9]*	12	F	✓	✓	130×130	×	✓	×
Liu et al. [13]	10	T	✓	✓	320×240	×	✓	×
Molina et al. [17]*	55	F	×	✓	176×144	✓	✓	×
IMHG dataset	8	F, S	✓	✓	640×480	×	✓	✓

TABLE I: Summary of hand gesture datasets. Previous work is summarized based on the following characteristics: number of hand gesture classes; number of views (T - Top view, F - Front view, S - Side view); RGB data; depth data; resolution of images; pose of finger joints; availability of the dataset; application to HRI. *Sign language gestures.

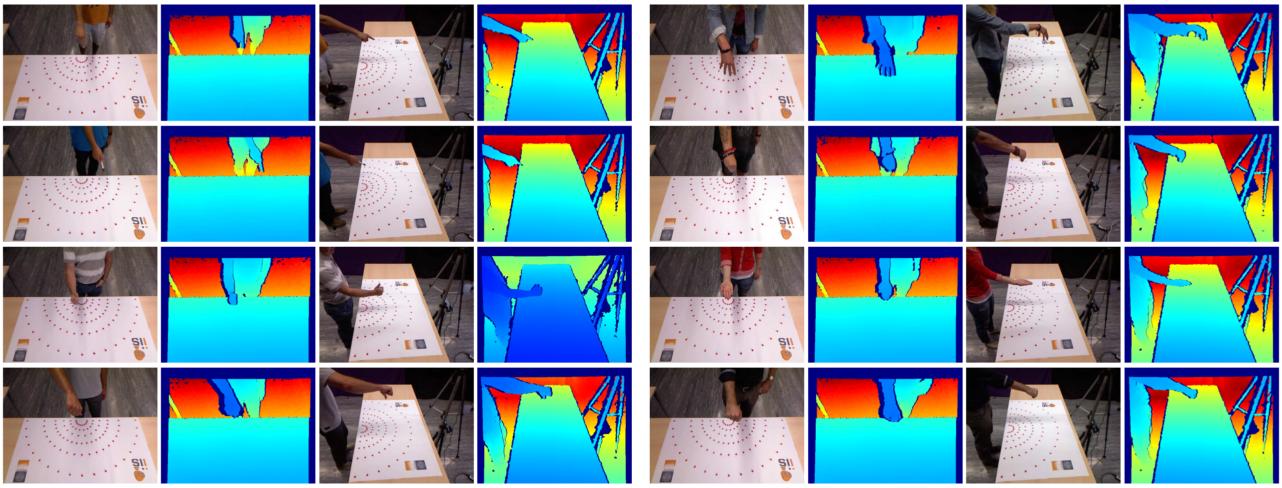


Fig. 2: IMHG dataset sample images.

view RGB-D camera setup. We captured 640×480 RGB images and `uint16` depth images. The depth sensing within Kinect is based on a structured infrared (IR) pattern. The simultaneous use of multiple depth cameras can extend the coverage of the vision system to a great extent. However, when multiple infrared patterns are projected at the same scene, the received depth signal degrades severely. To overcome this challenge Butler et al. [19] proposed the *Shake 'n' Sense* technique. The Kinect is minimally vibrated using an offset-weight vibration motion and thereby artificially introduces motion blur.

We address the depth interference problem in a different way. Instead of modifying the Kinect sensor we use the open-source *freenect* library to control the depth streaming of the Kinects. Using the freenect driver library it is possible to toggle the reading of the infrared pattern by controlling the flags of Kinects, thereby allowing multi-view RGB-D data to be captured. The extrinsic camera matrix between Kinects is estimated using ROS multiple camera calibration package. The calibration error can be up to 2 cm in 3D space.

Participants and Workspace

Participants from both genders were involved in the data acquisition process. They were asked to stand at a distance of approximately 1.3 m away from both the cameras, i.e. front view and side view, to perform eight classes of hand gestures. To avoid confusion, participants were shown different types of gestures prior to the recording, but no specific instructions were given to the participants on how to recreate the gesture, allowing their gestures to be recorded in a natural fashion.

We designed a polar coordinate system as shown in Fig. 3b with numbers marked at each *red* dot to capture the ground truth of the referencing gestures. The participants were asked to point at 16 randomly selected numbers. For the remaining six gestures we recorded only single instances, since they are not correlated with the location of an object. The workspace was configured such that hand gestures were visible from both cameras.

Dataset availability

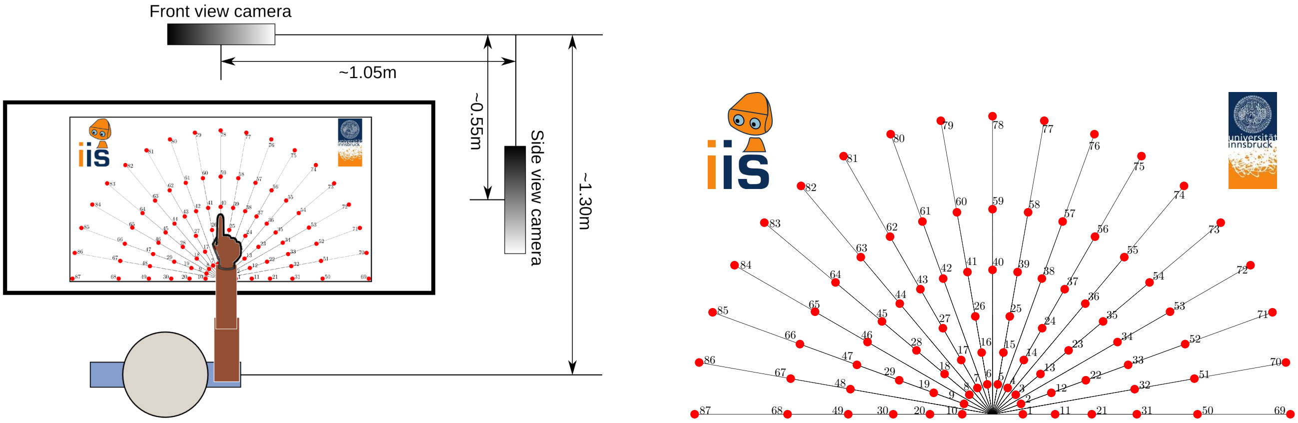
The IMHG dataset is available at this link[2]. New gestures can be added to the dataset by researchers, provided images are captured in a calibrated setup. Researchers can reproduce the data acquisition setup following the instructions given on the IMHG dataset page. It is to be noted that gestures should be recorded as static RGB-D images.

Researchers can contribute their work to the current dataset. The contributed set of images will be tested for calibration errors. Once accepted, they will be added on the page with an acknowledgement. It is encouraged to submit the hand gestures associated with a semantic content.

IV. Conclusions and Future Work

For a human-robot interaction to take place at close proximity it is necessary to measure the performance of

[2] `https://iis.uibk.ac.at/public/3rdHand/IMHG_dataset/`



(a) Illustration of the IMHG dataset acquisition setup.

(b) Polar coordinate system to record the ground truth location of the target pointed at.

Fig. 3: IMHG dataset acquisition setup.

gesture recognition systems independent of human body pose. We described a novel IMHG dataset from two RGB-D cameras with ground truth. The dataset comprises 8 classes of hand gestures with semantic meaning. The dataset mainly focuses on HRI scenarios. The data acquisition setup is easily reproducible for extension of the dataset with additional hand gestures. We are currently working on a baseline evaluation to detect hand gestures using a probabilistic framework.

Acknowledgement

The research leading to this work has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 (Specific Programme Cooperation, Theme 3, Information and Communication Technologies) under grant agreement no. 610878, 3rd HAND.

References

[1] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications," *Communications of the ACM*, vol. 54, no. 2, pp. 60–71, 2011.

[2] K. Dautenhahn, "The art of designing socially intelligent agents: Science, fiction, and the human in the loop," *Applied artificial intelligence*, vol. 12, no. 7-8, pp. 573–617, 1998.

[3] J. P. Wachs, M. Jacob, Y.-T. Li, and G. Akingba, "Does a robotic scrub nurse improve economy of movements?" in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83 160E–83 160E.

[4] J. Alonso-Mora, S. Haegeli Lohaus, P. Leemann, R. Siegwart, and P. Beardsley, "Gesture based human-multi-robot swarm interaction and its application to an interactive display," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5948–5953.

[5] M. Lopes, J. Peters, J. Piater, M. Toussaint, A. Baisero, B. Busch, O. Erkent, O. Kroemer, R. Lioutikov, G. Maeda, Y. Mollard, T. Munzer, and D. Shukla, "Semi-Autonomous 3rd-Hand Robot," in *Robotics in future manufacturing scenarios*, 3 2015, workshop at the European Robotics Forum, Vienna, Austria. [Online]. Available: https://iis.uibk.ac.at/public/papers/Lopes-2015-CogRobFoF.pdf

[6] P. Barattini, C. Morand, and N. M. Robertson, "A proposed gesture set for the control of industrial collaborative robots," in *RO-MAN, 2012 IEEE*. IEEE, 2012, pp. 132–137.

[7] C.-B. Park and S.-W. Lee, "Real-time 3d pointing gesture recognition for mobile robots with cascade hmm and particle filter," *Image and Vision Computing*, vol. 29, no. 1, pp. 51–63, 2011.

[8] M. Pateraki, H. Baltzakis, and P. Trahanias, "Visual estimation of pointed targets for robot guidance via fusion of face pose and hand orientation," *Computer Vision and Image Understanding*, vol. 120, pp. 1–13, 2014.

[9] A. Kurakin, Z. Zhang, and Z. Liu, "A real time system for dynamic hand gesture recognition with a depth sensor," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, pp. 1975–1979.

[10] A. Barczak, N. Reyes, M. Abastillas, A. Piccio, and T. Susnjak, "A new 2D static hand gesture colour image dataset for ASL gestures," *Letters in the Information and Mathematical Sciences*, vol. 15, pp. 12–20, 2011.

[11] T.-K. Kim and R. Cipolla, "Canonical correlation analysis of video volume tensors for action categorization and detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 8, pp. 1415–1428, 2009.

[12] Z. Ren, J. Yuan, and Z. Zhang, "Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera," in *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 2011, pp. 1093–1096.

[13] L. Liu and L. Shao, "Learning discriminative representations from rgb-d video data," in *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*. AAAI Press, 2013, pp. 1493–1500.

[14] I. Guyon, V. Athitsos, P. Jangyodsuk, and H. J. Escalante, "The chalearn gesture dataset (cgd 2011)," *Machine Vision and Applications*, vol. 25, no. 8, pp. 1929–1951, 2014.

[15] A. Sadeghipour, L. philippe Morency, and S. Kopp, "Gesture-based object recognition using histograms of guiding strokes," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2012, pp. 44.1–44.11.

[16] S. Ruffieux, D. Lalanne, E. Mugellini, and O. A. Khaled, "Gesture recognition corpora and tools: A scripted ground truthing method," *Computer Vision and Image Understanding*, vol. 131, pp. 72–87, 2015.

[17] J. Molina, J. A. Pajuelo, M. Escudero-Viñolo, J. Bescós, and J. M. Martínez, "A natural and synthetic corpus for benchmarking of hand gesture recognition systems," *Machine Vision and Applications*, vol. 25, no. 4, pp. 943–954, 2014.

[18] C. L. Nehaniv, K. Dautenhahn, J. Kubacki, M. Haegele, C. Parlitz, and R. Alami, "A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction," in *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*. IEEE, 2005, pp. 371–377.

[19] D. A. Butler, S. Izadi, O. Hilliges, D. Molyneaux, S. Hodges, and D. Kim, "Shake'n'sense: reducing interference for overlapping structured light depth cameras," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1933–1936.

Building a Schema for the Description of HRI Experiments

Minsu Jang[1], Cheonshu Park[1], Jaehong Kim[1]

***Abstract*— We present an intermediary work on building a schema for describing HRI experiments in a formal way. A set of properties that can describe characteristics and metrics of experiments are extracted, and a simple data description schema based on the properties is introduced with an exemplary sample table-like descriptions. We plan to formalize the schema into an ontology so that researchers can reference for designing new HRI experiments and to come up with a set of standardized experiment processes and elicit benchmark measures.**

I. INTRODUCTION

Standardizing evaluation metrics and experimental scenarios for human-robot interaction is a seemingly impossible task to undertake as the styles and steps of interaction as well as evaluation criteria are widely different across various application areas. The first step of tackling the hard problem is to collect as much information as possible from previous works done in the field, analyze them, and build a formal database that can easily be queried for reference.

We present our preliminary work on building a formal description framework for the representation of the multiple aspects of HRI experiments. In section II, we enumerate the first batch of properties that can well characterize experiments. Then in section III, a set of sample descriptions in the tabular form is presented. Finally, future work is suggested in the last section.

II. A SCHEME FOR DESCRIBING HRI EXPERIMENTS

HRI experiments can be described by a set of properties that might be used to classify them into a number of categories. The categories can be referenced later for designing new HRI experiments and reuse related metrics. We succinctly enumerate and describe some of the essential and frequent properties in the following subsections.

A. Scenario Properties

An HRI experiment can basically be characterized and categorized by the properties of the domain scenario or task.

1) Application Area: A list of keywords that define specific domain is specified. Some examples include tele-presence, entertainment, guidance, medical, education, elderly-care, health-care, physical assistant, manipulation etc. Yanco et al. suggested a similar category called *TASK* with several example values e.g. *urban search and rescue*, *walking aid for the blind* and *delivery* [1].

[1]Minsu Jang, Cheonshu Park and Jaehong Kim are with Human-Robot Interaction Research Section, Electronics and Telecommunications Research Institute, Gajeong-ro 218, Daejeon-si, South Korea {minsu,bettle,jhkim504}@etri.re.kr

2) No of Interacting Partners: The number of participants a robot handles in an interaction session is an important indicator of cognitive capability. The interaction can be classified as one-to-one or multi-party. For the latter case, an integer value indicating the maximum number of participants can additionally be specified. Yanco et al. introduced a more general property called *HUMAN-ROBOT-RATIO* that denotes a non-reduced fraction of the number of humans over that of robots [1]. In our scheme, we consider more socially situated interactions where one robot interacts with one or more human users.

B. System Properties

Robots and accompanied system components may have different capabilities and be controlled in various ways.

1) Level of Autonomy: A robot system can be operated in different levels of autonomy. The most common method is called the WoZ(Wizard-of-Oz). It employs human operators to control robot systems. Fully automated operation is at the other extreme, and variable autonomy in between. This property is usually specified by a number from 1 to 10, in which the smaller value indicates the lower level of autonomy.

2) Interaction Modality: Robots interact in multi-modality in most scenarios. But in some cases limited modalities are employed e.g. either verbal only or non-verbal only. The value of this property is specified by a set of modalities.

3) Robot Platform: The name of robot platforms employed in an experiment are specified. Additional information such as a link to the specification of the platform can be accompanied.

C. Demographic Information

Demographic data is important as it often suggest the level of objectivity and reliability of experiments.

1) No of Participants: The number of participants is a major indicator of the reliability of the experimental results.

2) Age Distribution: This is usually specified by the mean age and the standard deviation of age distribution.

3) Gender Distribution: This indicates the gender specificity of the results.

D. Metrics

A number of metrics for evaluating robot performances have been introduced in HRI literatures, and a comprehensive review of them exists [2]. Some metrics are general enough to be widely employed, but many others have been invented for specific task-oriented experiments. Most popular examples of the former are the metrics defined in the Godspeed questionnaire such as anthropomorphism, animacy and likeability [3], while those for the latter include task efficiency, interaction fluency etc. Steinfeld et al. divided HRI metrics into two categories of *task* and *common*, and suggested a lengthy list of metrics with extensive references [2].

In this subsection, we introduce a scheme for specifying properties of a metric. Specific metrics collected from previous works are listed in Table I.

1) ID: A universally uniquely identifiable name is assigned to each metric e.g. uniform resource identifier.

2) Name: The literal name of the metric.

3) Measuring Method: Two representative methods of measuring a metric in HRI experiments are by user surveys or by automated or manual observations and quantification.

4) Subjectivity: A metric is either subjective or objective. The former includes those measured by survey, and the latter by observation.

5) Instrumentation: Some metrics are designed elaborately and validated in a number of studies that they are established as pseudo-standards. Some examples include Godspeed questionnaire [3] and PARADISE framework [4].

E. Experiment Annotation Properties

Several annotations can be accompanied to an experiment description for better understanding of its various contexts.

1) ID: A universally uniquely identifiable name is assigned to each experiment e.g. uniform resource identifier.

2) Year: The year of the experimentation.

3) Datetime: The date and time of the experimentation.

4) Duration: The duration of the experimentation.

5) Organization: The list of organizations involved in the experimentation.

6) DOI: A digital object identifier of the reference.

III. DESCRIBING HRI EXPERIMENTS

We have so far surveyed 20+ recent HRI papers that include user studies, and built a tabular description using a subset of the properties introduced in the previous section. The first step of building descriptions is to collect metrics and their properties. Table I shows a list of descriptions of subjective and objective metrics. Then, HRI experiments are described by specifying the properties and their values, as shown in Table II.

By building experiment descriptions, a researcher can effectively search for viable evaluation metrics to employ for her or his experiments. By querying the description database, one can easily extract major metrics that are most widely adopted for HRI system evaluation. Also, by the analysis of the meaning of each metric, we can identify the metrics that are specified in a different words but have the same semantics, which might allow us to standardize the terminologies for specifying various aspects of HRI experiments including evaluation criteria and metrics.

The final product of our work shall be a formal description language or an ontology that can be used to specify HRI experiments, which can be referenced and queried by human users as well as machines. We plan to design our description scheme in W3C's OWL Web Ontology Language [5].

IV. FUTURE WORK

a) Expanding the experiment database: Incorporating as many research papers as possible into the description database would be the most crucial task in the process of building a widely acceptable description framework for HRI experiments. We are trying to review research papers presented or published in the HRI-related conferences and journals in recent 5 years.

b) Expanding the description scheme: The current scheme has several limitations, one of which is that it does not differentiate between *experiment* and *experimentation*. Suppose that $EX01$ describes an experiment to evaluate the social presence of a tele-presence robot with or without a non-verbal expression capability. If the description includes sufficient information to replicate the experiment, multitudes of researchers might be able to conduct $EX01$ and report results, which might be a great step forward to standardized experiments. Defining an *experimentation* as *an act of conducting an experiment*, there might be many experimentations for an experiment. We plan to elaborate our design so that an experimentation and its results can be described independent of experiment.

Another crucial part of this work is to incorporate metrics and taxonomies studied in the seminal previous works [2], [1].

ACKNOWLEDGMENT

This work was supported by the Robot R&D Program (10041659) funded by the Ministry of Trade, Industry and Energy (MOTIE, South Korea).

REFERENCES

[1] H. A. Yanco, J. L. Drury, *et al.*, "Classifying human-robot interaction: an updated taxonomy." in *SMC (3)*, 2004, pp. 2841–2846.
[2] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006, pp. 33–40.
[3] C. Bartneck, E. Croft, and D. Kulic, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots," in *Metrics for HRI Workshop, Technical Report*, vol. 471. Citeseer, 2008, pp. 37–44.
[4] M. A. Walker, D. J. Litman, C. A. Kamm, and A. Abella, "Paradise: A framework for evaluating spoken dialogue agents," in *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 1997, pp. 271–280.
[5] D. L. McGuinness, F. Van Harmelen, *et al.*, "Owl web ontology language overview," *W3C recommendation*, vol. 10, no. 10, p. 2004, 2004.
[6] S. O. Adalgeirsson and C. Breazeal, "Mebot: a robotic platform for socially embodied presence," in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2010, pp. 15–22.
[7] J. Fasola and M. J. Matarić, "Using socially assistive human–robot interaction to motivate physical exercise for older adults," *Proceedings of the IEEE*, vol. 100, no. 8, pp. 2512–2526, 2012.
[8] J. Hall, T. Tritton, A. Rowe, A. Pipe, C. Melhuish, and U. Leonards, "Perception of own and robot engagement in human–robot interactions and their dependence on robotics knowledge," *Robotics and Autonomous Systems*, vol. 62, no. 3, pp. 392–399, 2014.

TABLE I

DESCRIPTIONS OF HRI EVALUATION METRICS

ID	Name	Method	Subjectivity	Relevant Instrumentation	Reference
SM01	Social Presence	Survey	Subjective	NA	[6],[7]
SM02	Trust	Survey	Subjective	Grotz's Individualized Trust Scale	[6]
SM03	Cooperation	Survey	Subjective	Takayama questions	[6]
SM04	Engagement	Survey	Subjective	The Temple Presence Inventory	[6],[8]
SM05	Enjoyment	Survey	Subjective	NA	[9],[7]
SM06	Understandability	Survey	Subjective	NA	[9]
SM07	Convenience of Conversation	Survey	Subjective	NA	[9]
SM08	Compliance	Survey	Subjective	NA	[10]
SM09	Competency	Survey	Subjective	NA	[10],[11]
SM10	Persuasiveness	Survey	Subjective	NA	[10]
SM11	Sociability	Survey	Subjective	NA	[10]
SM12	Trustworthiness	Survey	Subjective	NA	[10]
SM13	Usefulness	Survey	Subjective	NA	[7]
SM14	Companionship	Survey	Subjective	NA	[7]
SM15	Capabilities as an Exercise Coach	Survey	Subjective	NA	[7]
SM16	Anthropomorphism	Survey	Subjective	GODSPEED Questionnaire	[12]
SM17	Animacy	Survey	Subjective	GODSPEED Questionnaire	[12]
SM18	Likeability	Survey	Subjective	GODSPEED Questionnaire	[9],[11],[12],[8]
SM19	Perceived Intelligence	Survey	Subjective	GODSPEED Questionnaire	[12],[7]
SM20	Perceived Safety	Survey	Subjective	GODSPEED Questionnaire	[12]
SM21	Overall Impression (Pos,Neg,Neutral)	Survey	Subjective	NA	[13]
SM22	Intention to Use	Survey	Subjective	NA	[13]
SM23	Level of Robot Understanding	Survey	Subjective	NA	[8]
SM24	Naturalness of Robot Behaviors	Survey	Subjective	NA	[11]
SM25	Perceived Agreeableness	Survey	Subjective	NA	[14]
SM26	Perceived Similarity	Survey	Subjective	NA	[14]
OM01	Mean Interaction Time Per Session	Observation	Objective	NA	[15],[7]
OM02	The Duration of Subject's Gaze Toward Robot	Observation	Objective	NA	[15]
OM03	The Number of Subject's Looks Toward Robot	Observation	Objective	NA	[15]
OM04	The Level of Subject's Knowledge Gain	Survey	Objective	Evaluation by Quiz	[15]
OM05	Task Success Rate	Observation	Objective	NA	[16]
OM06	Average Time for Task Completion	Observation	Objective	NA	[7]
OM07	Robot's Feedback Percentage	Observation	Objective	NA	[7]
OM08	Task Success	Observation	Objective	PARADISE Framework	[12]
OM09	Dialog Quality	Observation	Objective	PARADISE Framework	[12]
OM10	Dialog Efficiency	Observation	Objective	PARADISE Framework	[12]
OM11	Information Recall Correctness	Observation	Objective	Evaluation by Quiz	[11]
OM12	No of Interactions Per Day	Observation	Objective	NA	[17]
OM13	Accuracy of Friendship Estimation	Observation	Objective	NA	[17]
OM14	No of Subject's Decision Changes Caused by Robot	Observation	Objective	NA	[14]

TABLE II

AN HRI EXPERIMENT DATABASE (ID: THE ID OF AN EXPERIMENT, APP. AREA: APPLICATION AREA, LOA: LEVEL OF AUTONOMY, NOP: NO OF PARTICIPANTS, AD: AGE DISTRIBUTION, GD: GENDER DISTRIBUTION(M:MALE,F:FEMALE,U:UNKNOWN))

ID	DOI	Year	App. Area	LoA	Modality	Platform	NoP	AD[$\mu(\sigma)$]	GD	Metrics
EX01	[6]	2010	Telepresence	1	Non-Verbal	MeBot	48	23.21(8.92)	24F/18M	SM01,02,03,04
EX02	[9]	2012	Guidance	10	Mixed	Furhat	86	35	39F/46M/1U	SM05,06,07,18
EX03	[10]	2013	Guidance	10	Verbal	Mindstorm	48	23.69(7.83)	24F/24M	SM08,09,10,11,12
EX04	[15]	2013	Child-care	1	Mixed	Nao	20	9.5	NA	OM01
EX05	[15]	2013	Child-care	1	Mixed	Nao	10	9.5	NA	OM02,OM03
EX06	[16]	2013	Guidance	10	Verbal	Mindstorm	20	NA(18∼65)	11F/9M	OM05
EX07	[7]	2012	Elderly-care	10	Mixed	Bandit	13	83(77∼92)	12F/1M	SM01,13,14,15,OM06,07,01
EX08	[7]	2012	Elderly-care	10	Mixed	Bandit	24	77(68∼89)	19F/5M	SM13,19,OM06,07,01
EX09	[12]	2012	Serving	10	Verbal	iCat	31	27.9(21∼50)	9F/22M	SM16,17,18,19,20,OM08,09,10
EX10	[13]	2013	Guidance	5	Mixed	Robovie	40	NA	NA	SM21,22
EX11	[8]	2013	Conversation	10	Non-Verbal	Birt	63	20.4(18∼32)	48F/15M	SM04,18
EX12	[11]	2013	Conversation	10	Mixed	Wakamaru	32	24.9(18∼61)	NA	SM09,18,24,OM11
EX13	[17]	2012	Education	5	Mixed	Robovie	NA	NA	NA	OM12,13
EX14	[14]	2012	Conversation	1	Mixed	Robosapien	40	NA	20F/20M	SM25,26,OM14

[9] S. Al Moubayed, J. Beskow, B. Granström, J. Gustafson, N. Mirning, G. Skantze, and M. Tscheligi, “Furhat goes to robotville: a large-scale multiparty human-robot interaction data collection in a public space,” in *Proc of LREC Workshop on Multimodal Corpora*, 2012.

[10] S. Andrist, E. Spannan, and B. Mutlu, “Rhetorical robots: making robots more effective speakers using linguistic cues of expertise,” in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2013, pp. 341–348.

[11] C.-M. Huang and B. Mutlu, “The repertoire of robot behavior: Enabling robots to achieve interaction goals through social behavior,” *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 80–102, 2013.

[12] M. E. Foster, A. Gaschler, M. Giuliani, A. Isard, M. Pateraki, and R. Petrick, “Two people walk into a bar: Dynamic multi-party social interaction with a robot agent,” in *Proceedings of the 14th ACM international conference on Multimodal interaction*. ACM, 2012, pp. 3–10.

[13] D. F. Glas, S. Satake, F. Ferreri, T. Kanda, N. Hagita, and H. Ishiguro, “The network robot system: enabling social human-robot interaction in public spaces,” *Journal of Human-Robot Interaction*, vol. 1, no. 2, pp. 5–32, 2012.

[14] L. Takayama, V. Groom, and C. Nass, “I’m sorry, dave: i’m afraid i won’t do that: social aspects of human-agent conflict,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 2099–2108.

[15] T. Belpaeme, P. E. Baxter, R. Read, R. Wood, H. Cuayáhuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayová, G. Athanasopoulos, V. Enescu, *et al.*, “Multimodal child-robot interaction: Building social bonds,” *Journal of Human-Robot Interaction*, vol. 1, no. 2, pp. 33–53, 2012.

[16] C. Breazeal, J. Gray, and M. Berlin, “An embodied cognition approach to mindreading skills for socially intelligent robots,” *The International Journal of Robotics Research*, vol. 28, no. 5, pp. 656–680, 2009.

[17] T. Kanda, R. Sato, N. Saiwaki, and H. Ishiguro, “A two-month field trial in an elementary school for long-term human–robot interaction,” *Robotics, IEEE Transactions on*, vol. 23, no. 5, pp. 962–971, 2007.

Measuring Interaction Between People with Dementia and Socially Assistive Robot in Australian Home-Based Care

Rajiv Khosla, Khanh Nguyen, Mei-Tai Chu

Abstract—**This research aims at measuring the interaction between seniors with dementia and a socially assistive robot (named Lucy) in home-based care. The first ever longitudinal field trials have been conducted in Australian homes of people with dementia. The result analyses based on several interactional metrics show that socially assistive robots like Lucy have the ability of breaking technology barriers, positively engaging with its human partner for remarkable frequency and duration, demonstrating a potential to reduce demands on caring time and provide respite to the carers. This research also provides an evidence base to enable the selection of the robot services that are perceived most positively by people with dementia in home-based care.**

I. INTRODUCTION

The primary driving force behind this research is the predicted severe shortage of the human element and engagement in aged care in the coming decades. Like most of the developed countries, Australia's population is ageing. Over the next several decades, population ageing is projected to have the need for aged care services is growing at the rate of 68 percent but supply of health care workers is only growing at the rate of 14.8 percent labour [1].

Recently, health care researchers have shown the need for promoting person-centred care, self-identity and personhood for older persons and people with dementia [2-4]. Given the importance of pursuing this path, our research involves marrying personhood [3, 4] in dementia care with socially assistive robotics embodiment of care concepts [5] underpinned in personhood and facilitated by context sensitive cloud computing techniques involving artificial intelligence, soft computing and computer vision techniques to realise a symbiotic robotic system.

To measure the human robot interaction with social robots, [6] has recommended common metrics for standardization. In this paper, we employ several metrics to measure the interactional level, positive engagement, service preference and quality of robot experience between the human partners with the socially assistive robots in Australian home-based care. These employed measures are consistent with the engagement metrics in [6]. The results analysed from multi-modal data collection have shown that the socially assistive robot like Lucy have the ability of breaking technology barriers with human partner to achieve noticeable frequency and duration of interaction with the human partners. The emotional responses analysed using Observed Emotion Scales [7] indicate positive engagement of the human partners. In addition, the analysis from activity logs also reveals the most preferred services being accessed by the human partners in home-based context.

The rest of this article is organised as follows. Section 2 presents the theoretical foundation of our human-centred service design. Section 3 describes field trials, measurements and trial results. Finally, Section 4 concludes the paper.

II. THEORETICAL FOUNDATIONS OF HUMAN-CENTRED SERVICE DESIGN

Tobin [8] suggests that one of essential components of care that helps to support and maintain personhood is that care must be individualised to meet the unique needs of individuals. Indeed, most of the research shows that successful interventions are those that are tailored to the elders' preference [9, 10]. Research has demonstrated that individualised care whether in the context of person-centred approaches to care or leisure activities can significantly affect the elders' wellbeing such as improvements in the level of involvement in activities [11]; reductions in agitated and anxious behaviours [12, 13]; increased interest, pleasure and well-being [14]; engagement in meaningful activity, continuation of skills and hobbies and improved social interaction [15].

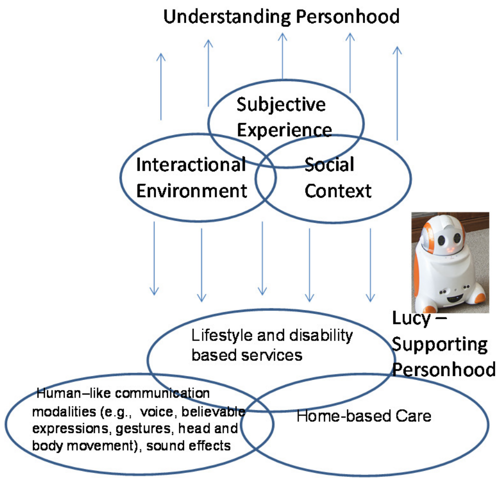


Figure 1. Mapping personhood [3] in Lucy

Personhood has been defined as 'the standing or status that is bestowed upon one human being, by others, in the context of relationship and social being' [16]. It includes three fundamental components, namely, interactional environment, subjective experience and social context. Figure 1 shows mapping of concepts related to these three components in Lucy.

The embodiment of interactional environment in Lucy involves modelling of person-centred services as well as human characteristics like gesture, emotional expressions, voice, motion, dancing, and dialog adaptation to deliver those services.

The subjective experience in a care context involves design of services personalised around the lifestyle of person with dementia. These lifestyle based services which reflect

Prof. Rajiv Khosla, Dr. Khanh Nguyen and Dr. Mei-Tai Chu are with the Research Centre for Computers, Communication and Social Innovation, La Trobe University, Australia (corresponding author's e-mail: r.khosla@latrobe.edu.au).

their personhood should enable a positive and engaging relationship between Lucy and person with dementia, It can consequently make them more resilient and useful [17]. The need for engagement, reciprocal relationship and usefulness is relevant in a home-based care environment where assistive technologies need to support the carers by providing the persons with dementia or other disabilities partial independence for limited periods. The interactional environment needs to employ human-like communication modalities like voice, emotive expressions, head and body movement, and gestures in an emotionally engaging manner to facilitate a reciprocal relationship [17].

III. Field Trials and Results

A. Field Trials

The longitudinal study (between 6 weeks and 8 months) was conducted in five Australian households. All participants are older people (ages 52-89) having dementia living in Victoria, Australia. Each participant has had a robot deployed at their home (figure 2). The robot has human attributes include baby face like appearance, voice vocalisation, face recognition, face registration and face tracking, facial expressions, gestures, body motion sensors, dance movements, touch sensors, emotion recognition and speech acoustics recognition. The deployed services include singing and dancing, quiz, weather forecast, storytelling, phone, news reader and reminder. The services were personalised to each participant based on their own preferences (i.e., favourite songs, books, daily calendar, etc). These preferences helped to augment their good memories.

Figure 2. Snapshots of home-based trials.

B. Measures, Indicators and Data Instruments

This research employed the common metrics for standardization for measuring the engagement [6] of the seniors with dementia to a social robot in home-based setting. Specifically, the outcomes of the trials were evaluated using following evaluation metrics: interactional level (frequency and duration), service preference, emotional responses and quality of robot experience. The mapping between these measures and their indicators & data instruments are summarised in Table I.

TABLE I. Mapping Between Measures And Indicators & Data Instruments

Measure	Indicators	Data Instruments
Interactional level	- Frequency of interaction - Duration of interaction	- Activity logs
Service preference	- Service interaction - Duration of service interaction	- Activity logs
Emotional engagement	- Alertness - Pleasure - Displeasure	- Videos
Quality of robot experience	- Five-point likert scales of survey responses	- Surveys

C. Results

1) Interactional level

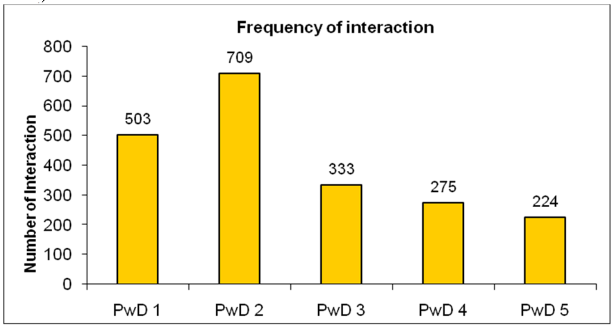


Figure 3. Interaction per participant

The total number of interaction between each participant to the robot is illustrated in Figure 3. The figure shows that the participants have interacted with their robot with significance of times, in which the participants 1 & 2 have 503 and 709 times of interactions respectively.

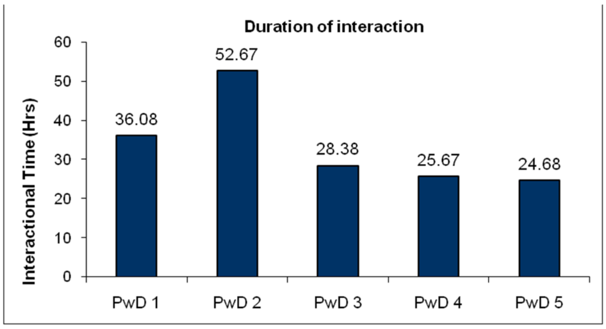


Figure 4. Interactional duration per participant

We analysed the interactional activity data to obtain the total duration time of interaction (Figure 4). The figure shows that five participants have spent 24 to 52 hours to interact with the robots. This not only gives family carers some respire but also potentially reduces social cost of caring for persons with dementia.

2) Service preference

The activity log data have been analysed to measure service preference. The word cloud being visualised from the activity data (Figure 5) for quickly perceiving the most prominent services accessed by the trial participants.

Figure 5. Wordcloud from activity logs

The frequency of interaction of the service determines its relative prominence. This shows that the participants listened to songs most. Other services like quiz, weather, and news are the next most accessed services respectively. This is consistent with the statistics (Fig. 6) from activity logs. The result implies that sensory enrichment service (singing & dancing) and cognitive support service (quiz) are most engaging for people with dementia in their home environment.

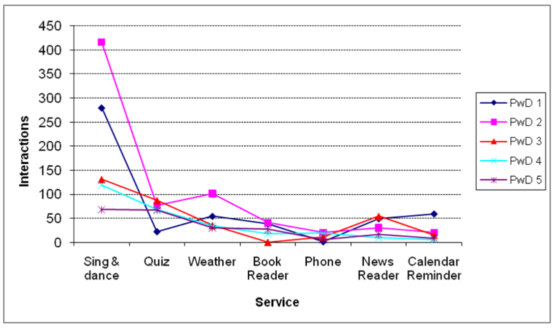


Figure 6. Pariticipants' interaction per service

3) *Emotional response*

Videos recorded during the interactions between Lucy and the participants are analysed based on OERS (The Observed Emotion Rating Scale) [7] which is an observational tool for rating two positive emotions (pleasure and general alertness) and three negative emotions (anger, anxiety or fear, and sadness). The observed emotion of the participant in each recorded video was annotated according to the observed emotion signs (Table II). For each video, the emotional response of the participant is annotated as a string of binary values which value '1' annotated to an emotion sign if that sign is detected. Finally, the emotion state is annotated if at least one of its signs is detected (annotated with value '1').

TABLE II. OBSERVED EMOTION SIGNS

Pleasure		Smiling
		Singing
		Kissing (robot)
		Flapping
Displeasure	Anger	Physical aggression
		Yelling
		Cursing
		Making distancing gesture
	Anxiety	Shrieking
		Hand wringing
		Wincing/grimacing
	Sadness	Crying
		Frowning
		Head turn down
General alertness		Participating in task
		Maintaining eye contact
		Following object (robot) or person
		Responding by moving or saying
		Turing body
		Moving toward person or object

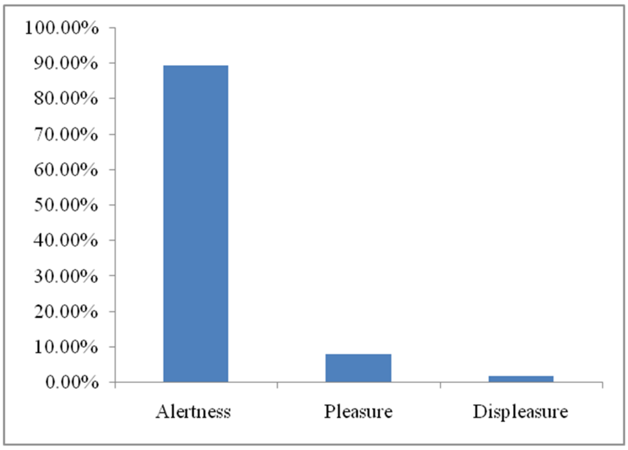


Figure 7. Observed emotion ratings

The statistics of all annotated data entries are summarised in Figure 7. The results show that the alertness rate is over 89%, and about 9% of the reaction are pleasure evolving smiling, touching or kissing actions. Only less than 2% of the reactions are classified as displeasure (most are observed as anxiety).

4) *Robot experience*

The quality of robot experience survey (Table III) has been conducted at the end of the trials using a standard five-point Likert scale (Strongly Disagree=1, Disagree=2, Neutral=3, Agree=4, Strongly Agree=5). Figure 7 shows the robot experience comparison amongst the participants and the mean. The figure shows that on average the responses are positive (above 3.0). The visual distribution of survey responses in Figure 8 shows that the median of responses to most of survey questions is around 4 (positively agree).

TABLE III. QUALITY OF ROBOT EXPERIENCE

Quality of Robot Experience	P1	P2	P3	P4	P5	M	SD
1. I enjoy the contact with the social robot.	3	3	5	5	4	4	0.894
2. I enjoy one to one activity with the social robot.	3	4	5	4	4	4	0.632
3. Betty makes me feel better.	3	3	4	5	4	3.8	0.748
4. Betty makes me smile.	4	4	5	5	4	4.4	0.490
5. Betty doesn't worry me.	4	4	5	5	4	4.4	0.490
6. Betty is a friend.	3	4	5	4	4	4	0.632
7. I like the way Betty responds to me.	3	4	5	3	3	3.6	0.800
8. Betty helps me to revisit happy memories	3	4	4	4	2	3.4	0.800
9. Betty helps me to become more independent.	4	3	4	3	2	3.2	0.748
10. I like to touch Betty.	4	4	4	1	2	3	1.265
11. I like Betty's blushing.	2	4	5	4	4	3.8	0.980
12. I like Betty's dancing.	2	4	5	5	4	4	1.095
13. Betty reacts appropriately to my actions.	3	3	5	3	3	3.4	0.800
14. I want you to come back with Betty.	3	3	5	5	2	3.6	1.200

15. I would like to have a robot Betty for myself.	3	3	5	4	2	3.4	1.020
*Note: M= mean, SD = Standard Deviation							

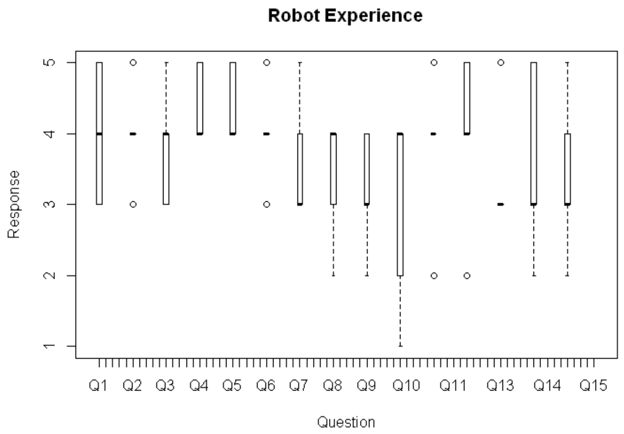


Figure 8. Distribution (box plot) of quality of robot experience responses

IV. CONCLUSION

This paper has studied the interaction between the socially assistive robots and seniors with dementia in home-based settings. The personhood has been embodied in service design for service personalization. Several metrics for standardization have been employed for interaction measures and the results show that the deployed robots have successfully eliminated the technology barriers of the human partners, positively engaged them in noticeable frequency and duration. This research demonstrates the potential of socially assistive robot for breaking technology barriers, positively engaging, providing sensory enrichment and cognitive support, augment their good memories and making older persons more resilient with dementia in home-based care as well as reducing the social cost of caring for persons with dementia.

REFERENCES

[1] Australian-Bureau-of-Statistics. (2010). *Population by Age and Sex, Regions of Australia (cat no. 3235.0).*

[2] J. Tinney, B. Dow, M. Fearn, L. Dowson, C. Hempton, K. Hill*, et al.*, "Best Practice in Person-centred Health Care for Older Victorians " National Ageing Research Institute2007.

[3] D. O'Connor, A. Phinney, A. Smith, J. Small, B. Purves, J. Perry*, et al.*, "Personhood in dementia care: Developing a research agenda for broadening the vision," *Dementia,* vol. 6, pp. 121-142, February 1, 2007 2007.

[4] J. Cohen-Mansfield, A. Parpura-Gill, and H. Golander, "Utilization of Self-Identity Roles for Designing Interventions for Persons With Dementia," *Journals of Gerontology: Series B,* vol. 61, pp. 202-212, 2006.

[5] A. Tapus, M. J. Mataric, and B. Scasselati, "Socially assistive robotics [Grand Challenges of Robotics]," *Robotics & Automation Magazine, IEEE,* vol. 14, pp. 35-42, 2007.

[6] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz*, et al.*, "Common metrics for human-robot interaction," presented at the Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, Salt Lake City, Utah, USA, 2006.

[7] M. P. Lawton, K. V. Haitsma, and J. Klapper, "Observed Affect In Nursing Home Residents," *Journal of Gerontology,* vol. 51, 1996.

[8] S. S. Robin, *Preservation of the Self in the Oldest Years: With Implications for Practive*: Springer, 1999.

[9] J. Cohen-Mansfield, A. Parpura-Gill, and H. Golander, "Utilization of Self-Identity Roles for Designing Interventions for Persons With Dementia," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 61, pp. P202-P212, July 1, 2006 2006.

[10] D. W. O'Connor, D. Ames, B. Gardner, and M. King, "Psychosocial treatments of behavior symptoms in dementia: a systematic review of reports meeting quality standards," *International Psychogeriatrics,* vol. 21, pp. 225-240, 2009.

[11] J. C. Rogers, M. B. Holm, L. D. Burgio, C. Hsu, J. M. Hardin, and B. J. Mcdowell, "Excess Disability During Morning Care in Nursing Home Residents With Dementia," *International Psychogeriatrics,* vol. 12, pp. 267-282, 2000.

[12] L. Chenoweth, M. T. King, Y.-H. Jeon, H. Brodaty, J. Stein-Parbury, R. Norman*, et al.*, "Caring for Aged Dementia Care Resident Study (CADRES) of person-centred care, dementia-care mapping, and usual care in dementia: a cluster-randomised trial," *The Lancet Neurology,* vol. 8, pp. 317-325, 4// 2009.

[13] J. Cohen-Mansfield, A. Libin, and M. S. Marx, "Nonpharmacological Treatment of Agitation: A Controlled Trial of Systematic Individualized Intervention," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences,* vol. 62, pp. 908-916, August 1, 2007 2007.

[14] G. Gibson, G. E. Chalfont, P. D. Clarke, J. M. Torrington, and A. J. Sixsmith, "Housing and Connection to Nature for People with Dementia," *Journal of Housing For the Elderly,* vol. 21, pp. 55-72, 2007/12/17 2007.

[15] M. B. Detweiler, P. F. Murphy, L. C. Myers, and K. Y. Kim, "Does a Wander Garden Influence Inappropriate Behaviors in Dementia Residents?," *American Journal of Alzheimer's Disease and Other Dementias,* vol. 23, pp. 31-45, February 1, 2008 2008.

[16] L. Flicker, "Dementia Reconsidered: the Person Comes First," *BMJ,* vol. 318, p. 880, 1999-03-27 00:00:00 1999.

[17] J. Goetz and S. Kiesler, "Cooperation with a robotic assistant," presented at the CHI '02 Extended Abstracts on Human Factors in Computing Systems, Minneapolis, Minnesota, USA, 2002.

Immersive underwater robot control: HRI benchmarking with UWSim

Juan C. García[1], Bruno Patrão[2], Javier Pérez[1], Luis Almeida[2,3], Paulo Menezes[2], David Fornas[1], Jorge Dias[4], Pedro J. Sanz [1]

Abstract— **During underwater intervention missions, one of the critical issues from the human-machine interaction perspective concerns the operators stress, which should be able to control the intervention system, mainly due to the high complexity of information displayed through specifically designed Graphical User Interface, and the usually available master-slave control architecture. This work introduces and evaluates a new approach with the aim to minimize the aforementioned drawbacks and increasing the users immersive feeling and ergonomics of the teleoperation systems. The UWSim benchmarking platform is used to analyze the user experience of this HRI interfaces. Preliminary results are discussed highlighting the pros and cons of this novel procedure.**

I. INTRODUCTION

Underwater robot intervention missions are becoming more frequent, due to the large number of applications that can be carried out: oil and gas industry (i.e. operating submerged infraestructures), search and recovery missions (i.e. grasping a crashed airplane blackbox) or deep water archeology. Usually, the typical robot used in these missions is Remotely Operated Vehicles (ROV), which has high economical and logistic requirements. In addition, these robots uses a master/slave architecture and relays all the responsibility in the pilot, who suffers cognitive fatigue and stress[1]. The evolution of this kind of robot, the Autonomous Underwater Vehicle (AUV), removes the human in the control loop but reduces the intervention capabilities. Nevertheless, both kinds of intervention systems (ROV and AUV) still have problems related with the control of the robot and the way the user interacts with the system. The ROV pilot uses different joysticks with several buttons to control the robot, which represents another drawback in sense of Human-Robot Interaction (HRI). In both cases, the Graphical User Interface (GUI) to control the robot is focused to an expert user, who should be an expert in both the robot and mission to perform, due to the complexity of the different screens. Thus, the ROV pilot should pay attention to different screens and control panels to get all the information about the robot, sensors and cameras. In the case of AUV robots, the user interfaces are also complex, but the main problem is the lack of feedback to the user, due to the lack of physical connection between the robot and the pilot. These challenges have been explored in the recent projects TRIDENT [2] and TRITON [3]. In the former, the HRI was developed taking into account that it could be used by a non-expert user [4] and does not rely on a specific robot platform.

We develop and benchmark a new approach to control the robot and to interact with the system (i.e. robot and computers). These improvements can be divided in different aspects:

- the information to be shown to the user, reducing the data depending on the mission and context.
- an immersive system, where the user would get the feeling to be in the robot.
- a natural gesture control interface to control robot improving traditional ways (i.e. joystick).
- an abstraction layer, which will manage all these improvements.

II. UWSIM AND INTERFACE DEVELOPMENT

It is worth to mention that all these functionalities will be integrated as a new GUI using the UWSim simulator [5] as base. This simulator is an open source software tool for visualization and simulation of underwater robotic missions and is under development at the Universitat Jaume I. Bearing in mind that both, the simulator and the GUI are using the same middleware, Robot Operating System (ROS), all the aforementioned progress should be ROS compatible, making easier the integration in the current architecture. Moreover, the simulator has a benchmarking architecture with which the HRI interfaces can be evaluated. The datasets used for the experiments are publicly available, making the experiments repeatable. This benchmarking platform has been used to compare and evaluate template tracking algorithms under different visibillity conditions and different PID vehicle controllers under different water currents, among other experiments [6]. This tools is able to measure aspects such as execution time or number of control actions.

Usually, the UWSim uses a Linux command line to show all the information to the user. That means the user should pay attention to different screens to monitor the robots direction and velocity, the distance to the seafloor or the surface, or sensors data. This generates a lot of information and it is quite common to cause stress to the user. So, related to the GUI improvements, we design a virtual windshield, where only the most relevant information according to the mission is being carried out, will be shown graphically to the user. For instance, when the robot gets close to the seafloor, a progress bar will be displayed indicating the distance. When this progress bar is full, it indicates the robot is on

[1] Department of Computer Science and Engineering, Universitat Jaume I, Castelln de la Plana, Spain. Email: {garciaju, japerez, dfornas, sanzp}@uji.es

[2] Institute of Systems and Robotics, University of Coimbra, Coimbra, Portugal. Email: {bpatrao, paulo}@isr.uc.pt

[3] Institute Polytechnic of Tomar, Tomar, Portugal. Email: laa@ipt.pt

[4] Khalifa University of Science and Technology and Research, Abu Dhabi, UAE. Email: jorge@deec.uc.pt

the seafloor and a warning icon will be displayed in the windshields. In order to achieve a more realistic and get a better user experience, some 3D techniques has been applied to the UWSim, using a head-up mounted display (HMD). This will allow the user to focus his attention into the scene and get an immersive feeling.

As we mentioned before, the ROV pilot uses a joystick with several buttons to control the robot. This could be useful for an expert user, but it will become quite difficult for non-expert user. Our proposal in terms of HRI is the use of natural user interface. For instance, using a Leap Motion controller, the user hand movements (i.e. move the hand forward) are transformed as robot movements (i.e. move the vehicle forward). The HMD sensor will help the user modifying the camera point-of-view instead of use a mouse.

Finally, in order to manage all the aforementioned features, an User Interface Abstraction Layer (UIAL) has been developed. This abstraction layer can be integrated easily in most of the architectures and has several tasks: receive data from the different controllers (keyboard, hand trackers, joysticks) and select the most suitable depending on the task carried out, receive data from sensors (virtual or real) or adapt and filter the information to be shown to the user.

III. EXPERIMENT EVALUATION AND RESULTS

In order to compare and evaluate our proposal with the traditional setup, we have created a new scenario in the UWSim, where there are located five rings not collinear with different orientation. The mission objective it to control the robot, making it to pass through each ring. The 13 participants were between 20-50 years-old with a mean age of 26, reporting normal or corrected to normal vision. Participation was voluntary and experiments are in accordance with ethical principles of research. Each participant received a short description about the mission with a non-detailed map (the position of the simulated robot and the waypoints in a top view representation) and the pertinent information, before they started the experiment.

The main objective in this experiment is evaluate our proposal, measuring the time needed by each user to finish the mission, the number of collision between the robot and the objects placed in the scenario, and the steering commands used to control the robot. In order to compare the results, we also measured the time needed to gain confidence with the robot controller, the time needed to arrive at each waypoint and the overall time needed to perform the whole mission Additionally, each participant has to answer a questionnaire to get suggestions and opinions.

We consider three different setups:

- Setup 1 (traditional): the user needs to pay attention to two different screens and control the camera POV control using the mouse. The robot will be controlled using a gamepad.
- Setup 2 (immersive traditional): the user will use the HMD to get the immersive experience and to control the camera POV. The robot will be controlled using a gamepad.
- Setup 3 (immersive with Leap Motion): the user will the HMD to get the immersive experience and to control the camera POV. The robot will be controlled using a Leap Motion.

We divided the participants in different groups, so they started the experiment running a different setup since the beginning. Due to the fact that the scene was the same in all the setups, the learning effect would be reduced. It is obvious that the users improve their results after each test because they know where are the waypoints.

After analyzing the time needed by the users, the vast majority of the users improved their overall time when using the immersive system. This is due to the camera's POV is intuitively connected to the participant's head movements, so it is not needed to manually switch between controlling the robot and camera's POV. All the participants reported that the head movement camera control is 'natural and intuitive' and required shorter adaptation times. According to the comparison about the gamepad and the Leap Motion, the people with experience in video games prefer the gamepad because they are get used. On the other hand, people without experience in video games preferred the Leap Motion because 'the movements to control the robot are easy to remember and they do not need time to learn'. Nevertheless, all the participants agree that the use of both Leap Motion and the Oculus Rift is the most difficult, because the HMD blocks the view between the user's hand and the Leap Motion, losing the reference. After the authors noticed about this problem, the setup was modified adding a fan close to the Leap Motion, so the user would be able to feel the airflow and localize the device.

Regarding the user experience questionnaire, the results can be summarized as follows: 92% of the participants preferred using Oculus Rift instead of the traditional system with the double screen setup; 77% of users considered Oculus Rift very important or essential for improving the mission performance; and 85% of users considered the virtual interface (windshield) very important or essential.

IV. CONCLUSIONS

We present a new approach for a GUI, focused in underwater robots, and we evaluate it with the UWSim benchmarking tools. The improvements are divided in the information to be displayed, the use of an immersive device for 3D experience, the natural gesture robot control (HRI) and the development of an abstraction layer.

In order to validate our proposal, an experiment with users have been carried out. Almost all the participants agree with the idea of the virtual windshield and the use of the HMD for an immersive experience, getting only the most relevant information. In the other hand, the use of a novelty device such as Leap Motion was unsuccessful, due to the lack of feedback and because the user loses the reference between his/her hand and the device. Nowadays, the authors are developing and testing new approaches to solve this issue.

ACKNOWLEDGMENT

This work was partly supported by Spanish Ministry of Research and Innovation DPI2011-27977-C03 (TRITON Project), by Universitat Jaume I grant PID2010-12 and PhD grants PREDOC/2012/47 and PREDOC/2013/46.

REFERENCES

[1] T.B. Sheridan, *Telerobotics, Automation and Human Supervisory Control*. Massachusetts, USA: MIT Press, 1992.

[2] P.J. Sanz, P. Ridao, G. Oliver, G. Casalino, Y. Petillot, C. Silvestre, C. Melchiorri, A.Turetta, *TRIDENT: An european project targeted to increase the autonomy levels for underwater intervention missions* in OCEANS'13 MTS/IEEE conference. Proceedings. San Diego, CA (2013)

[3] A. Peñalver, J. Pérez, J.J. Fernández, J. Sales, P.J. Sanz, J.C. García, D. Fornas, R. Marín, *Autonomous Intervention on an Underwater Panel mockup by using Visually-guided Manipulation Techniques*. in IFAC World Congress, Ciudad del Cabo, Africa, 2014.

[4] J.C. García, J.J. Fernández, P.J. Sanz, R. Marín, *Increasing Autonomy within Underwater Intervention Scenarios: The User Interface Approach*, in: 2010 IEEE International Systems Conference. 2010 IEEE International Systems Conference. San Diego, CA, USA; 2010:71-75.

[5] M. Prats, J. Pérez, J.J. Fernández, P.J. Sanz, *An open source tool for simulation and supervision of underwater intervention missions*, in 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2577-2582, 7-12 Oct. 2012

[6] J. Pérez, J. Sales, M. Prats, J. V. Martí, D. Fornas, R. Marín, and P. J. Sanz, *The underwater simulator UWSim: Benchmarking capabilities on autonomous grasping*, in 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2013.

GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

DPI2011-27977-C03

INSTITUTE OF SYSTEMS AND ROBOTICS
UNIVERSITY OF COIMBRA

IRS Lab

UNIVERSITAT JAUME•I

IMMERSIVE UNDERWATER ROBOT CONTROL: HRI BENCHMARKING WITH UWSIM

Juan C. García, Bruno Patrão, Javier Pérez, Luís Almeida, Paulo Menezes, David Fornas, Jorge Dias and Pedro J. Sanz

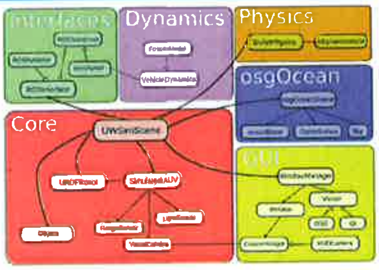
UW ROBOT CONTROL PROBLEMS

Some common problems:
- Master-slave control.
- Tethering cable problem
- Multiple monitors to pay attention.
- Different input devices to control.

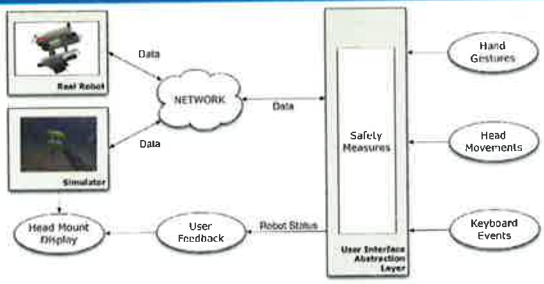
ARCHITECTURE

Archictecture details:
- ROS is used as middleware.
- UWSim relies on OSG and osgOcean.

UWSim capabilites:
- Monitor robot movements.
- User training.
- Mission specification.
- Benchmarking: algorithms, HRI, mission metrics...



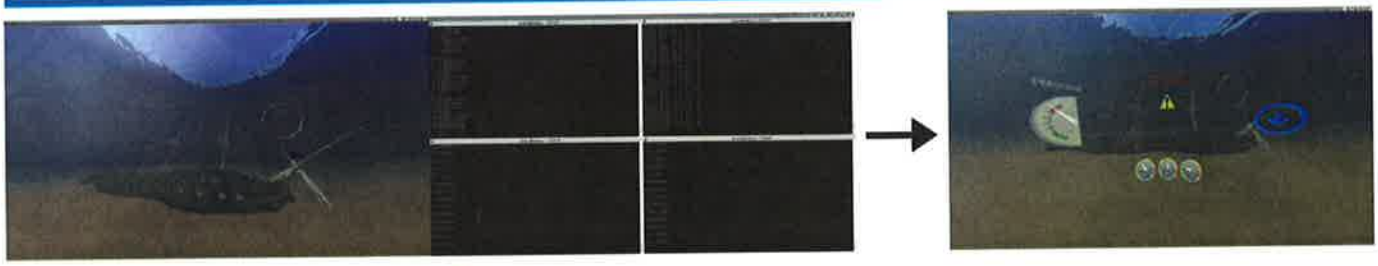
USER EXPERIENCE IMPROVEMENTS



UIAL: User Interface Abstraction Layer
- Connects different input devices with the architecture.
- Receive data from sensors (virtual or real).
- Adapts & filters the info. to be shown to the user.
- Applies safety measures to prevent robot damages.
- Reconfigure the use of both the input and output devices according to the mission or the task.

User immersion:
- A HMD is used to obtain a better 3D experience.
- The HMD IMU is used to control the camera POV.

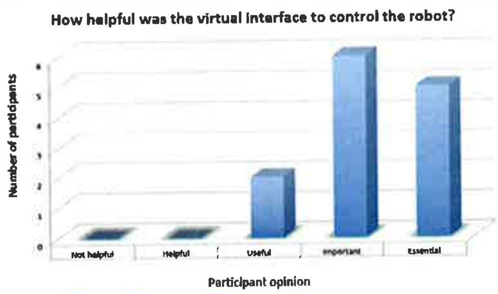
New input devices (apart of the gamepad and keyboad):
- Leap Motion, representing the Natural UI approach.
- Joystick, represeting the common ROV control device.



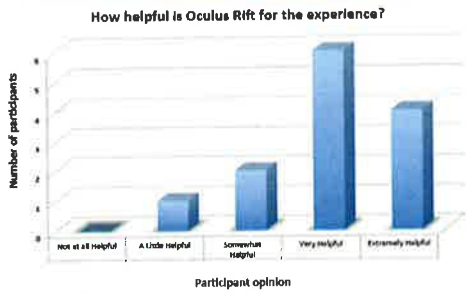
Improving user feedback:
- New UWSim GUI adding to the UWSim the previous improvements.
- Graphical icons shown information, instead of a command line terminal.

HRI BENCHMARKING USING THE UWSIM SIMULATOR

User performing a test: the monitor splitted view is used to create the Oculus Rift HMD image.



Two different setups were considered:
- Setup 1 (traditional): the user needs to pay attention to two different screens and control the camera POV control using the mouse. The robot will be controlled using a gamepad.
- Setup 2 (immersive traditional): the user will use the HMD to get the immersive experience and to control the camera POV. The robot will be controlled using a gamepad.
- Setup 3 (immersive with LeapMotion): the user will the HMD. The robot will be controlled using a LeapMotion.



Results:
- Setup 1 played better on challenges involving orientation and altitude complexities.
- Setup 2 gets the best mean time and minimum steering orders.
- Setup 3 robot controlled was preferred for non-expert users, altough they become tired because of the arm position.

Identification of Hazards in Invasive/Surgical Robotics

Péter Pausits, Gábor Szögi, Dénes Á. Nagy, Marsel Nallbani, Imre J. Rudas and Tamás Haidegger*
Antal Bejczy Center for Intelligent Robotics; Óbuda University Robotics Special College
Óbuda University, Budapest, Hungary
{peter.pausits, gabor.szogi, denes.nagy, marsel.nallbani, rudas, haidegger}@irob.uni-obuda.hu

Abstract—**Service robotics receives more and more attention in the developed world beside industrial applications. While industrial robotics conquered the factories, it was important that researches develop a number of principles and guidelines to help minimizing the risk of human accidents. Today's safety standards of industrial robotics almost completely exclude the possibility of physical interaction between the human operator and the robotic device. Just recently, a new paradigm, the divided workspace has prevailed, and as a consequence, a number of new and critical safety issues have emerged. Service robots have become even more complicated, as we cannot erect a fence around domestic robots, and in the case of medical robotics, the human–machine interaction is inevitable. The goal of this research was to explore and quantify human–machine interactions, and classify them based on their hazard level. The focus is on surgical robotic devices and their current applications, as this situation presents one of the most complex form of interaction. It is necessary to make service robots complying with safety standards, based on a unified and generally accepted methodology.**

Index Terms—invasive/surgical robots, human–machine interaction, hazards, medical device standard

I. Introduction

As a result of the continuous development and research, the field of medicine uses a great number of devices for diagnostics, surgery, treatment and rehabilitation [1-2]. The human body is so complex that every single procedure poses a different challenge to the inventors of medical devices and to the medical team. The growing autonomy of medical devices poses possible new hazards for their operators and the patients [3]. However, the recent development and targeted application of new medical devices greatly contributed to the recovery of patients and helps reducing the risks of complications. For hazard management and reduction of risks, a number of standardization bodies have been dealing with the safe operation of such devices. Due to their efforts, sophisticated standards exist in most fields, such as industrial robotics, nevertheless, the development of new applications keep requiring the revision of standards, as well as the creation of new ones.

In this work, the landscape of international standards (ISO and IEC) was explored, concerning single groups of devices and connections between the different overlapping fields were identified. Medical devices—including medical robots (robotic devices for medical use)—and numerous sub-domains within. Fig. 1 presents the various areas in which certain sub-domains are still hard to be defined accurately. Primarily, this study focuses on invasive/surgical robotic devices. While by the IEC definition of surgery refers to "procedures performed through a skin incision", practically, natural orifice access (NOTES) and minimally invasive procedures (MIS) are also considered in the same category. Today's surgical robots typically perform procedures in a master–slave (teleoperation) mode, based on image guidance, or in a cooperatively controlled way [1]. By examining the safety standards applied for these devices, shortcomings in the guidelines can be identified, especially regarding the autonomic functions of the robotic devices.

For this reason, the new IEC JWG35 standard focuses on the "Particular requirements for the basic safety and essential performance of medical robots for surgery". Most of the risks of these machines are very similar to those of industrial robots, which are included in the ISO 10218 (Robots and robotic devices – Safety requirements for industrial robots) international standard and the current FDIS ISO 15066 (Robots and robotic devices – Safety requirements for industrial robots – Collaborative operation). These standards served as a guideline when new sources of risk of invasive/surgical robots were explored.

II. Overview of Existing Standards

Due to the particularity of the field, one single safety standard for all medical devices is impossible. Some areas have seen new methods and proposals for safety validation [4], but no unified guiding standard has been proposed yet. Surgical robots (to some extent) are similar to industrial robots functioning in divided workspace, which have seen some recent improvement in standardizaton due to the ISO 13482:2014 (Robots and robotic devices – Safety requirements for personal care robots).

Certain domains, such as IGRT (image-guided radiotherapy) [5-6] and some other therapeutic procedures, including the application of CT (computed tomography) [7], MRI (magnetic resonance imaging) [8], US (ultrasound) [9] and X-ray [10] are already regulated via particular standards. Furthermore, single medical devices already exist for diagnostic measurement of the patient's state, for physiotherapy treatments, and for surgical procedures. Ultrasound is a widely used medical device, and its particular safety standards are the following: [9], [11-15]. The guidelines for the applications of laser devices in the medical field are also given.

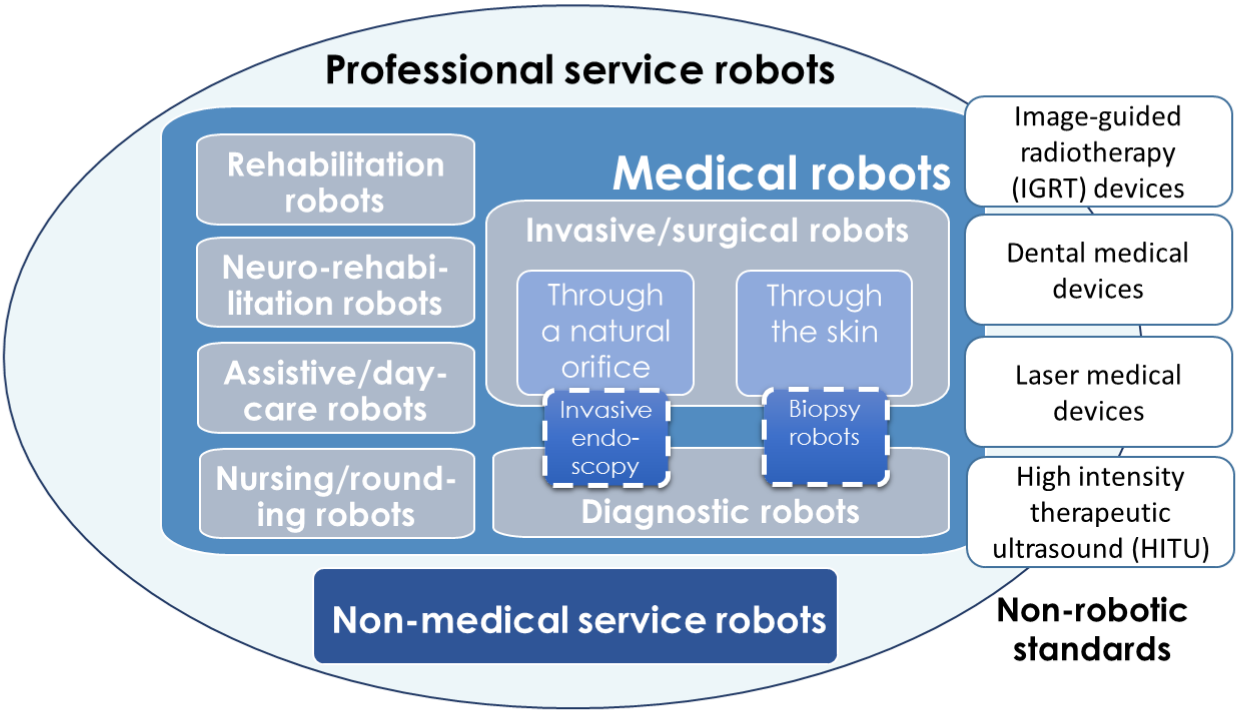


Figure 1. Various types of medical robotic devices fitting into the ISO 8373 and IEC 60601 context.

Medical devices used for intra-corporeal diagnostics are primarily endoscopes [16] and capsule endoscopes [17-18]. Endoscopy used in diagnostics contributes greatly to the success of surgical procedures and to the treatments [19] (IEC 60601-2-18). It is important to note that currently there is no safety standard for capsule endoscopy. Dental medical machines are a special field in which single diagnostic, surgical, rehabilitation devices and their safety standards have already been developed (IEC 60601-2-63, -65, -60).

III. Analysis of Hazards Posed by Invasive/Surgical Robots

Different ways of applications of invasive/surgical devices require further categorization. The two main groups of invasive/surgical robots are those that enter the body through a natural orifice and those that penetrate the skin. The two fields cannot be isolated completely, and the definition of surgery is still underway [20]. In this paper invasive surgical devices were also examined from the aspect of procedures executed by robots functioning in divided workspace. In this setup the safety technology of industrial robots have an important role in defining the threats posed by the robotic equipment.

Except for the end-effector, threats are similar in the case of all robots, so the starting point for analysis is the safety standards for industrial robots [21]. In Table I, the hazards of industrial robots are presented, whether they pose a threat in the case of surgical robots.

In the case of surgical robots, contact with the patient is necessary, thus we conducted further research to identify the hazards that can occur during single interactions. Automation of these functions may impose these hazards:

Lack of force feedback: Certain squeeze motions, incisions, drillings and all procedures that require adequate clamping force may not send feedback to the surgeon.

Failure of the force transmission device: In the case of wired force transmission, the wires or their mechanics can be damaged and result in unintended movement.

Blockage of the device in the body: Surgical devices that enter the body require continuous control and conduct due to the particularities of certain channels in the body. For this, they have intermittently placed joints [22] that can – in case of a failure – fix and get stuck.

Inflation of bodily cavities: Laparoscopic procedures require the inflation of the abdominal cavity, which can damage the patient's body if overpressure occurs [23], and may alter the function of the robotic device.

Unintended cessation of the power source: Both the procedure in process and the movements occurring during restart are in danger if a power outage happens during a procedure. In case of a restart, not adequately prescribed protocol movements and operations can pose significant threats.

Speed of signal transformation: If robots functioning in teleoperation mode are not adequately operated, the delay of visual and other feedbacks can be a hazard. If the time delay increases, a hazard of malfunctioning can occur [24].

Unintended movement from the operator: In the case of master–slave systems, movements of the operator that are opposite to the expected operation can pose a threat. A sneeze or a cough can provoke unintended movements of the hand that the robot might read as a command.

Application of high frequency: High frequency surgical procedures are primarily used when cutting soft tissues, and pose a number of new threats to the patient and the operator, unlike to the ones when traditional incisions are used [25].

Usage of materials damaging the living organism: Chemicals applied during the procedure can potentially be poisonous for the body.

Misplacement of the disconnected tissue: Disconnected tissue has to be continuously removed, as if they enter an

TABLE I

ROBOT HAZARDS BASED ON ISO 10218

Type of group	Hazard	Surgical robot
Mechanical hazards	movements (normal or unexpected) of any part of the robot arm (including back)	x
	movements (normal or unexpected) of external axis	x
	materials and products falling or ejection	
	manipulation of products and materials, including ejection	x
	impossibility to go out robot cell (via cell door) for a trapped operator in automatic mode	
	between fixtures (falling in); between shuttles, utilities	
Electrical hazards	process using high voltage or high frequency, i.e. electrostatic painting, inductive heating	x
	welding applications using high voltage	
Thermal hazards	cold surfaces or objects	x
	explosive atmosphere caused by the process, i.e., paint (atomized particles, powder painting), flammable solvents, grinding and milling dust	
	exposure to temperature extremes required to support the process	
Noise hazards	loss of balance, disorientation in working area of robot cell	
	inability of two persons assigned to a task to coordinate their actions through normal conversation	x
	long-term exposure to elevated noise levels	x
Vibration hazards	loosening of connections, fasteners, components resulting in unexpected stopping or expulsion of parts	x
Radiation hazards	EMF interference with proper operation of the robot system	
	exposed to process-related radiation, i.e. arc welding, laser	
Hazards associated with the environment	misidentification of real problem and compound problem by making incorrect or unnecessary actions	
Comb. of hazards	unexpected movements of robot or end-effectors or associated machine	x
	misinterpretation of collaborating robots or simultaneous motion	x
	high-speed rotational parts breaking or disengaging from part retention equipment	x
	contacted by process-related expulsion (i.e., spot welding)	
	part retention device fails	
	unrestrained robot or associated machine part (maintained in position by gravity) falls or overturns	x

undesired place they can cause complications during and after the procedure.

Collision of the robotic arms: Collision of the end-effectors is indispensable during a procedure, but as a consequence of a wrong command the arms can have an unintended contact, resulting in a negative effect.

Positioning of the end effector during its replacement: Repeated positioning of the end-effectors is necessary in the case of a laparoscopic procedure due to the restricted area of penetration. During the replacement of the equipment, the position can change leading to the damage of the tissue.

Unintended clamping force: Every soft and hard tissue has a different degree of tolerance regarding clamping force before they suffer permanent damage. Poorly chosen clamping force can be a threat to the tissue.

Unintended clamping time interval: Certain tissues can only be squeezed for only a certain amount of time with a certain amount of force before it potentially becomes damaged.

Speed of rotation: When drilling, the speed of rotation has to be well chosen to avoid a possible threat.

Unintended movement of the parts of the body: Soft and hard tissues have neurological particularities that can cause movement as a result of an unexpected effect. This is a potential hazard during surgical procedures.

Obviously, not every source of hazards appears in the case of a unique surgical robot; therefore, separate particular standards are necessary for invasive/surgical robots that enter through a natural orifice or through the skin. The threats included in the ISO 10218 standard provide the basis for the list of new threats in the case of surgical robots, as a special for of human–machine interaction. These hazards are presented in Table II. Tissue motions for example is a group of threats where adequate regulation of autonomous functions is required, however this issue only surfaced in the field of medical robotics and such was not present in the case of industrial applications. Further, when fixing tissue, taking the possible clamping forces and positions into consideration is necessary, as this also means a new threat to the success

TABLE II

NEW HAZARDS APPLYING TO INVASIVE/SURGICAL ROBOTS

Type of group	Hazard	Invasive/surgical robot	
		Natural orifice	Through skin
Mechanical hazards	Lack of force feedback	x	x
	Failure of the force transmission device	x	x
	Blockage of device in the body	x	x
	Pressure due to inflation	x	x
Electrical hazards	Unintended cessation of the power source	x	x
	Speed of signal transformation	x	x
Vibration hazards	Unintended movement from the part of the operator	x	x
Radiation hazards	Application of high frequency	x	x
Material/ substance hazards	Usage of materials damaging to the body	x	x
Hazards associated with the environment	Misplacement of the disconnected tissue	x	x
Comb. of hazards	Collision of the robotic arms		x
	Positioning of the end effector during its repl.	x	x
	Unintended clamping force	x	x
	Unintended clamping time interval	x	x
	Speed of rotation		x
Tissue motion hazards	Unintended movement of the parts of the body	x	x

of the operation and to the human body. An unexpected and unintended movement caused by the reflexes of the human body can also lead to serious injury. Consequently, adequate fixing also requires special attention.

IV. Conclusion

Medical robotics involves a special form of human–machine interaction, especially in the case of invasive applications. During our research, we concluded that as of today, there is no adequate standard for invasive/surgical robotic devices. Manufacturers and users have to apply certain related standards that do not completely address the risks of this field, especially in the case of autonomous operation. The creation and application of a safety standard is indispensable in the future, as the field of medical robotics is rapidly expanding, and complications during procedures can occur from previously unidentified sources. Furthermore, it can be concluded that current standards are obsolete and need an update due to the extensive pace of system development. This work has begun in the working groups of the major standardization bodies. In the future, definitions of invasive/surgical robots will be necessary with the further grouping along specific procedures that pose different threats to patients and operators.

Acknowledgment

T. Haidegger is a Bolyai Fellow of the Hungarian Academy of Sciences. The authors acknowledge the support of the Austrian Center for Medical Innovation and Technology (ACMIT, *www.acmit.at*) and Óbuda University Robotics Special College (ROSZ, *www.rosz.uni-obuda.hu*).

References

[1] M. Höckelmann, I. J. Rudas, P. Fiorini, F. Kirchner and T. Haidegger, "Current Capabilities and Development Potential in Surgical Robotics." *Int. J. Adv. Robot Syst.* 12(61), 2015.

[2] Haidegger T. and Rudas I. J., “From Concept to Market: Surgical Robot Development”, *Handbook of Research on Advancements in Robotics and Mechatronics*, pp. 242, 2014.

[3] Virk, G. S., and Haidegger T., “Classification guidelines for personal care robots–medical and non-medical applications”, *In Proc. of the 2012 IEEE IROS Workshop on Safety in Human–Robot Coexistence & Interaction*, 2012, pp. 33-36.

[4] Haddadin S., Khoury A., Rokahr T., Parusel S., Burgkart R., Bicchi A. and Albu-Schaffer A., ”A truly safely moving robot has to know what injury it may cause”, *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, Oct. 2012., pp. 5406-5413.

[5] (2007). Medical electrical equipment – Medical electron accelerators – Functional performance characteristics. IEC-60976. Available: http://www.iec.ch

[6] RADIOLOGYINFO, (Aug. 2015), Image-guided Radiation Therapy (IGRT). Available: http://www.radiologyinfo.org

[7] (2004). Evaluation and routine testing in medical imaging departments – Acceptance tests – Imaging performance of computed tomography X-ray equipment. IEC-61223-3-5.

[8] IEC 60601-2-44 Medical electrical equipment – Particular requirements for the basic safety and essential performance of X-ray equipment for computed tomography. Available: http://www.iec.ch

[9] (2010). Medical electrical equipment – Particular requirements for the basic safety and essential performance of magnetic resonance equipment for medical diagnosis. IEC-60601-2-33. Available: http://www.iec.ch

[10] (2007). Medical electrical equipment – Particular requirements for the basic safety and essential performance of ultrasonic medical diagnostic and monitoring equipment. IEC-60601-2-37. Available: http://www.iec.ch

[11] (2011). Nuclear instrumentation – Portable X-ray fluorescence analysis equipment utilizing a miniature X-ray tube. IEC-62495. Available: http://www.iec.ch

[12] (2009). Medical electrical equipment – Particular requirements for the basic safety and essential performance of high frequency surgical equipment and high frequency surgical accessories. IEC-60601-2-2. Available: http://www.iec.ch

[13] (2009). Medical electrical equipment – Particular requirements for the basic safety and essential performance of ultrasonic physiotherapy equipment. IEC-60601-2-5. Available: http://www.iec.ch

[14] ISO 11553-1 (Safety of machinery – Laser processing machines – General safety requirements); IEC 60825-8 (Safety of laser products – Guidelines for the safe use of laser beams on humans) standard. (2014). Medical electrical equipment – Particular requirements for the basic safety and essential performance of non-laser light source equipment intended for therapeutic, diagnostic, monitoring and cosmetic/aesthetic use. IEC-60601-2-57. Available: http://www.iec.ch

[15] (2013). Medical electrical equipment – Particular requirements for the basic safety and essential performance of high intensity therapeutic ultrasound (HITU) equipment. IEC-60601-2-62. Available: http://www.iec.ch

[16] Luo R.C., Jing Wen Chen and Yi Wen Perng, ”Robotic endoscope system with compliance effect including adaptive impedance and velocity control for assistive laparoscopic surgery”, *3rd IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, Sept. 2010, pp. 100-105.

[17] Wadge E., Boulougoura M. and Kodogiannis V., ”Computer-assisted diagnosis of wireless-capsule endoscopic images using neural network based techniques”, *2005 IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications*, July 2005, pp. 328-333.

[18] Ciuti G., Menciassi A. and Dario P., ”Capsule Endoscopy: From Current Achievements to Open Challenges”, *IEEE Reviews in Biomedical Engineering*, vol. 4, pp. 59-72, Oct. 2011.

[19] S. D. Ladas, B. Novis, et al., ”Ethical issues in endoscopy: patient satisfaction, safety in elderly patients, palliation, and relations with industry”, *2th European Symposium on Ethics in Gastroenterology and Digestive Endoscopy*, July 2006, pp. 556-565.

[20] T. Haidegger, “Defining Surgery survey”; surgrob.blogspot.hu/2015/07/we-need-your-help.html (2015).

[21] (2011). Robots and robotic devices – Safety requirements for industrial robots. International Standard Organization-10218. Available: http://www.iso.org

[22] Yanagida T., Adachi K. and Nakamura T., ”Development of endoscopic device to veer out a latex tube with jamming by granular materials”, *IEEE Int. Conf. on Robotics and Biomimetics (ROBIO)*, Dec. 2013, pp. 1474-1479.

[23] Rosenblatt A., Bollens R. and Espinoza Cohen B., General Laparoscopic Information (Serias Springer Tracts in Manual of Laparoscopic Urology), Spriger-Verlag, 2008, pp. 1-17.

[24] Haidegger T., Sándor J. and Benyó Z., “Surgery in space: the future of robotic telesurgery”, *Surgical endoscopy*, vol. 25, pp. 681-690, 2011.

[25] M. Höchtl and P. Pöcksteiner, ”High Frequency Surgery: Mechanism of Action, Risks and Risk Minimization”, *Sempermed*, vol. 5, 2004.

Calibration and Evaluation of Multispectral Visual System for Reconnaissance

P. Kocmanova and L. Zalud*

***Abstract*— The paper deals with appropriate calibration of multispectral vision systems and evaluation of the calibration and data-fusion quality in real-world indoor and outdoor conditions. Checkerboard calibration pattern developed by our team for multispectral calibration of intrinsic and extrinsic parameters is described in detail. Circular object for multispectral fusion evaluation is described as well. The objects were used by our team for calibration and evaluation of advanced visual system of Orpheus-X3 robot that is taken as a demonstrator, but their use is much wider and we suggest to use them as testbed for visual and optical measurement systems of mobile robots.**

I. Introduction

Reconnaissance mobile robotics gains importance during the last years. Visual and space measurement subsystem is typically the most important sensory equipment with most significant impact to mission success. There are many missions in today's society that may require expendable robots to perform exploration in inaccessible or dangerous environments instead of indispensable people, e.g. CBRNE (Chemical, biological, radio-logical, nuclear, explosive), counter-terrorist fight, US&R (Urban Search and Rescue), etc.

Since the missions take place in real world, the robots have to be equipped for most, if not all, possible conditions that may happen. During both military and non-military search and rescue missions the robot can meet such a conditions like complete darkness, smoke, fog, rain, etc. For these conditions, purely visual spectrum is not sufficient to provide valuable data. One of the most promising approach for wide spectrum of situations is combination of data from visual spectrum, near infrared spectrum and far infrared spectrum. In visual spectrum (using standard tricolor cameras) the operator has the best overview of the situation, since he/she gets signal that is most similar to what he knows. By using thermal imagers working in far infrared spectrum he/she can perfectly perceive even slight changes in temperatures and this spectrum very well penetrates through water particles (fog, rain) plus it is not affected by visible light conditions. Most TOF (time-of-flight) proximity scanners and cameras work in near-infrared spectrum.

Data from abovementioned sources may be used for two main tasks – building of multispectral digital maps (e.g. evidence grids [1]) and real-time operator control of the robot (e.g. using visual telepresence [2]). So using the mentioned sensors we have multispectral data and to get the most information from them for both named cases, it is necessary to align them. The first necessary step is to make geometrical calibration of each of the sensors independently, than it is necessary to find translation and orientation of coordinate system of each camera/sensor.

In this paper we deal with design of calibration objects for intrinsic and extrinsic parameters in that way that only one calibration object is necessary for all mentioned spectrums. We also deal with design of evaluation targets for evaluation of abovementioned multispectral calibration as well as data fusion – both in real-world indoor and conditions.

The described techniques were applied on Orpheus-X3 robot developed in our laboratory, but they are universal. The algorithms and techniques may be used for different combinations of sensors in more spectrum ranges. The proposed multispectral calibration pattern may be used for any subset of mentioned sensors, the evaluation targets can be coped as general tool for assessment of visual system fusion quality.

A. *Orpheus-X3*

The Orpheus-X3 is an experimental reconnaissance robot based on the Orpheus-AC2 model made by our team to facilitate the measurement of chemical and biological contamination or radioactivity for military. The Orpheus-X3 offers the same drive configuration as its predecessor which makes the robot very effective in hard terrain and enables it to achieve the maximum speed of 15 km/h. The robot markedly more versatile than Orpheus-AC2, and this is a very important aspect in a robot made primarily for research activities. Furthermore, the device is equipped with a 3DOF manipulator for the sensor head. The Orpheus robots are described in more details in our previous papers, such as [3].

Sensor Head

The sensor head containing five optical sensing elements is shown in Fig.1. The sensors are as follows:

- Two identical tricolor CCD cameras (see 1 in Fig.1): TheImagingSource DFK23G445 with the resolution of 1280x960 pixels.
- Two identical thermal imagers Flir Tau 640 with the resolution 640x480, temperature resolution 0.05K.
- One TOF camera (see 3 in Fig. 1): A Mesa Imaging SR4000 with the range of 10m, resolution of 176x144.

* P. Kocmanova is with LTR s.r.o., 616 00, Brno, Czech republic and with Brno University of Technology, CEITEC, 616 00, Brno, Czech republic (e-mail: Petra.Kocmanova@ceitec.vubtr.cz).

L. Zalud is with LTR s.r.o., 616 00, Brno, Czech republic (phone: +420 541 143 641, fax: +420 541 141 123, e-mail: zalud@feec.vutbr.cz).

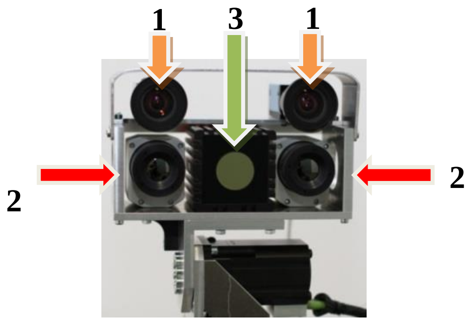


Figure 1. The sensor head. 1 – tricolor CCD cameras; 2 – thermal imagers; 3 – TOF camera

It is obvious from the preceding text that the fields of view (FOVs) of the sensors are similar. The largest FOV captures the TOF camera, which is required for the simultaneous use of stereovision and thermal stereovision. The main disadvantage of the applied TOF camera is its low number of pixels (spatial resolution). Compared to the CCD cameras, it is about 10 times lower in one axis, and in relation to thermal imagers it is twice lower.

II. Sensor Calibration

Calibration of the sensory head comprises the following stages:

- range calibration (TOF camera),
- temperature calibration (thermal imageres),
- calibration of intrinsic parameters,
- mutual calibration of extrinsic parameters.

SwissRanger manufacturer guarantees absolute accuracy of measured distance ±15 mm only for 11 x 11central pixels [4] and other image regions don't achieve this accuracy. Range calibration of TOF camera is described in detail in [5].

Temperature calibration is appropriate because of quite poor absolute precision measured temperatures (2K according to manufacturer specifications) and is described in [6]. Calibration of intrinsic and extrinsic parameters will be described hereinafter.

A. Calibration of Intrinsic and Extrinsic

We proposed 3 calibration plate based on checkerboard pattern. At first sufficient contrast of the calibration pattern should be achieved only by different materials. This version comprised an aluminum panel (low emissivity; high reflectivity) and a self-adhesive foil (high emissivity; low reflectivity). The main problem related to this initial board consisted in the high reflectivity of the aluminum base in cases that images are acquire under non-perpendicular angle.

The second version consisted of an aluminum panel with a laser-cut, anodized pattern and a chipboard covered by a black, matt foil. Anodizing of aluminum panel reduces high reflectivity. Good contrast of checkerboard pattern for thermal imagers was achieved by heating of aluminum part at 50°C.

The final version included a 2 mm laser-cut aluminum plate with active heating. This version is more comfortable and shortens time needed to prepare calibration. (see Fig.2).

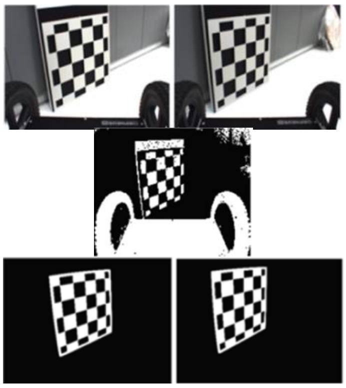


Figure 2. The final calibration plate: the left and right CCD cameras (up); the TOF camera intensity image (center). the left and right thermal imager cameras (down).

The calibration of the intrinsic and extrinsic parameters comprises the following stages:

- images acquisition,
- corner extraction based on automatic corner extraction from Omnidirectional Camera Calibration Toolbox for Matlab [7],
- homography from extracted corners,
- intrinsic and extrinsic parameters are computed from homography according to [8],
- nonlinear optimization that minimizes the sum of the squares of the re-projection.

Optimal image configuration that leads to reliable calibration parameters was determined. It is appropriate to use 9 images in configuration according to Fig. 3. Testing of optimal configuration is described in detail in [9].

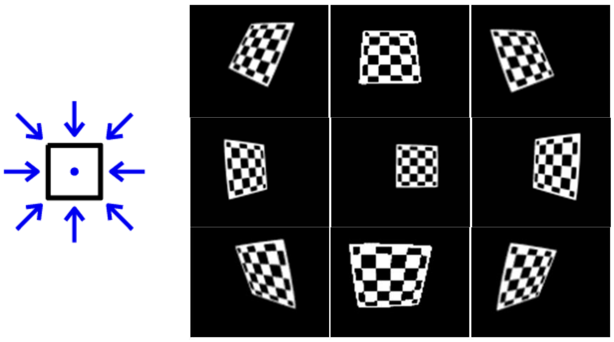


Figure 3. Image configuration for sensor head calibration: blue arrows indicate direction of image acquisition, blue dot shows image in normal position; examples of images for optimal configurations (down)

Camera calibration model is described in detail in [5]. Table 1. shows values of re-projection errors from calibration used for data fusion evaluation

TABLE I. REPROJECTION ERROR FOR MUTUAL CAMERA CALIBRATION

Results for mutual camera calibration	
Camera	***Re-projection error [pixel]***
TOF camera	0.34
CCD left camera	0.45
CCD right camera	0.44
Thermal imager left	0.57
Thermal imager right	0.42

III. DATA FUSION

Data fusion is performed by means of image transformations. The range measurements of the TOF camera can be displayed into images of the CCD cameras and thermal imagers using spatial coordinates. According to identical points (ID) of the TOF camera transformed into frames of the CCD camera and the thermal imager, the thermal image can be displayed into the CCD image and vice versa (see Fig. 4).

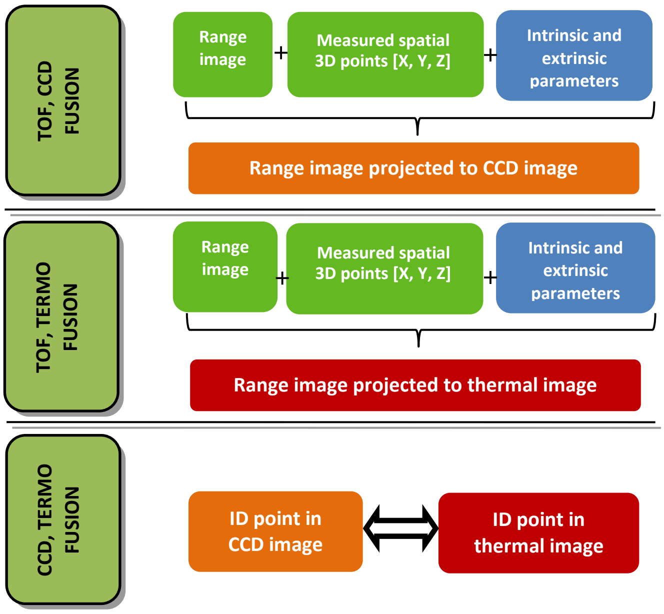


Figure 4. Scheme of data fusion: up – TOF and CCD data fusion; centre - TOF and thermal data fusion; down - CCD and thermal data fusion).

The input data for data fusion include the range measurement, the image coordinates of all sensors, and the results of previous calibration. The procedure comprises the following stages:

- Computation of spatial coordinates measured by TOF camera.
- Homogeneous transformation to determine measured spatial coordinates in frames of other cameras.
- Perspective projection to determine image coordinates in frames of other cameras.
- Correction of recalculated image coordinates to the calibrated position of the principal point.

The spatial coordinates *X*, *Y*, and *Z* are computed according to Eq. 1 and 4, where *x*, *y* are image coordinates of TOF camera, *f* focal length and d_0 is measured distance projected on optical axis. Calculation of spatial coordinate Z is simplified by substitution of cyclometric function Eq.3

$$X = \frac{d_0 x}{f}, Y = \frac{d_0 y}{f}. \qquad (1)$$

$$Z = d_0 = d \cos\left(\arctan\left(\frac{y}{\sqrt{f^2 + x^2}}\right)\right) \cos\left(\arctan\left(\frac{x}{f}\right)\right) \qquad (2)$$

$$\cos(\tan^{-1} a) = \frac{1}{\sqrt{1 + a^2}} \qquad (3)$$

$$Z = \frac{df}{\sqrt{x^2 + y^2 + f^2}} \qquad (4)$$

The homogeneous transformation is determined by Eq. 5, where $R_{[3\times3]}$ is the rotational matrix, $t_{[3\times1]}$ is the translation vector, and *X'*, *Y'*, *Z'* are the spatial coordinates of the second sensor. The image coordinates of the TOF camera in the next frame x_c', y_c' are computed using perspective projection (Eq. 6), where *f'* is the focal length of the second sensor.

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (5)$$

$$x_c' = \frac{f' X'}{Z'}, y_c' = \frac{f' Y'}{Z'} \qquad (6)$$

IV. INDEPENDENT EVALUATION OF DATA FUSION

The principle of this evaluation is comparison of identical objects directly extracted from images from CCD cameras and thermal imagers with objects extracted from images from TOF camera and projected to CCD cameras and thermal imagers frames using data fusion algorithm.

Final design of target clearly identifiable in images of all cameras was aluminum circle covered with black paper in the middle and with 3M red reflective tape on the edge with active heating. Reflective tape is used for easier identification of targets in images of TOF camera, but significant disadvantages of this reflectivity is missing measured distances, since too big portion of light is returned unidirectionally. The matte paper in the middle of the circle was used to overcome this problem – it is easy-to-be-identified by the TOF camera. We used 3 aluminum circles with 20 cm and 30 cm diameters. The targets are well identifiable on images of all 3 camera types (see Fig. 5).

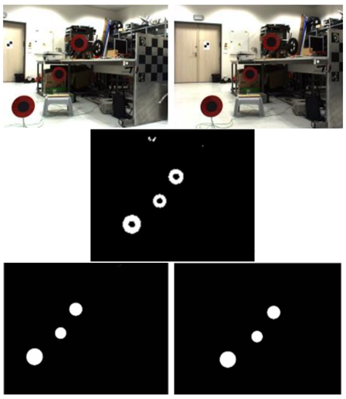


Figure 5. The images of final target for the verification of the data fusion accuracy: the left and right CCD cameras (up); the TOF camera intensity image (center); the left and right thermal imager cameras (down).

87 images were obtained in the experiment under real indoor conditions from the free ride of robot. 211 extracted objects were used for data fusion evaluation, TOF camera image radial distance for these objects was in range from 1 - 67 pixels, range for measured distance was from 1.1 to 5.7 m.

Table 2. shows standard deviations σ_x, σ_y of image coordinates *x*, *y* projected by proposed data fusion algorithm. Values of standard deviation are given in pixels of CCD cameras and thermal imagers. Standard deviation for CCD cameras reach values about 3 CCD pixels (around 0.3 TOF camera pixels) and standard deviation for thermal imagers about 1 thermal pixel (around 0.5 TOF camera pixels).

TABLE II. IMAGE COORDINATES X,Y AND IMAGE STANDARD DEVIATION OF DATA FUSION

Camera	**Standard deviation of coordinates *x, y* [pixels of investigated camera]**	
	σ_x	σ_y
CCDl	2.4	3.2
CCDr	3.0	3.3
THERMOl	1.0	1.2
THERMOr	1.2	1.1

V. CONCLUSION

As it is apparent from evaluation experiment described in Chapter 4, the fusion described in Chapter 3 is possible, but has its limits. The main problems come from the fact, the cameras used in the described case have significantly different spatial pixel resolution. It has to be said, the cameras were carefully selected to have parameters appropriate for Orpheus-X3 robot's main mission – real-time telepresence with augmented reality containing thermal information. The cameras had to be small, lightweight, but they had also offer unusually wide field-of-view. We can suppose for bigger robots sensors with considerably higher resolution might be used. The sensor resolution will also evolve in time (thermal cameras, 3D proximity cameras).

Numerical evaluation of data fusion algorithm is as follows: standard deviation for x, y image coordinates is around 3 pixels for CCD cameras (0.3 Pixel of TOF camera) and around 1 pixel for thermal imagers (around 0.5 TOF camera pixels).

The presented calibration process and evaluation may be used for visual and optical measurement systems of mobile robots in general, so its use is much wider than on presented Orpheus-X3 robot demonstrator.

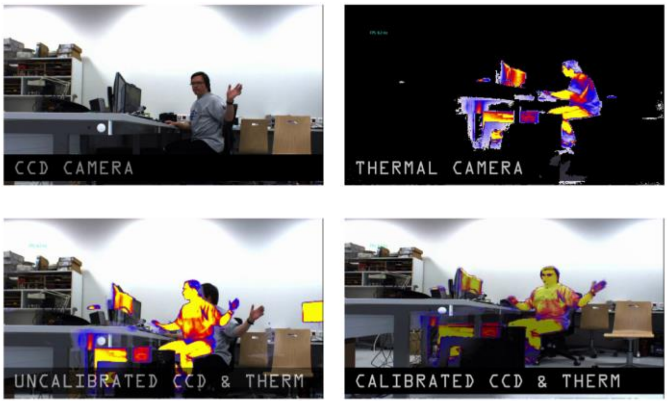


Figure 6. Image of CCD camera (upper left), image of thermal imager (upper right), uncalibrated data fusion (bottom left), calibrated data fusion (bottom right).

ACKNOWLEDGMENT

This work was supported by VG 2012 2015 096 grant named Cooperative Robotic Exploration of Dangerous Areas by Ministery of Interior, Czech Republic, program BV II/2-VS.

This work was supported by Technology Agency of the Czech Republic under the project TE01020197 "Centre for Applied Cybernetics 3".

REFERENCES

[1] Martin C. M., Moravec H. P., Robot Evidence Grids, Carnegie Mellon University, Pittsburg, 1996

[2] Zalud L., ORPHEUS – Reconniaissance Teleoperated Robotic System. In: 16th IFAC World Congress. 2005. pp. 1-6, Prague, Czech Republic.

[3] Zalud L., Burian F., Kopecny L., Kocmanova P., Remote Robotic Exploration of Contaminated and Dangerous Areas, International Conference on Military Technologies, pp 525-532, Brno, Czech Republic, ISBN 978-80-7231-917-6. (2013)

[4] SR4000 Data Sheet, MESA Imaging AG. Rev. 5.1, 2011.

[5] Kocmanova P. and Zalud L., Spatial Calibration of TOF Camera, Thermal Imager and CCD Camera. In: Mendel 2013: 19th International Conference on Soft Computing. Brno: Brno University of Technology, Fakulty of Mechanical Engineering, 2013, s. 343-348. ISBN 978-80-214-4755-4.

[6] Zalud L. and Kocmanova P., Fusion of Thermal Imaging and CCD Camera-based Data for Stereovision Visual Telepresence. In: SSRR 2013: 11th IEEE International Symposium on Safety Security and Rescue Robotics, pp.1-6

[7] Scaramuzza D.: OCamCalib: Omnidirectional Camera Calibration Toolbox for Matlab, https://sites.google.com/site/scarabotix/ocamcalib-toolbox

[8] Zhang Z., Flexible camera calibration by viewing a plane from unknown orientations. Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999, vol. 1, pp. 666-673

[9] Kocmanova, P.. Sensor Calibration for Multispectral Data Fusion in Mobile Robotics. Brno, 2015. Ph.D. thesis. Brno University of Technology.

[10] El khrachy, I. A. El H. M. Towards an automatic registration for terrestrial laser scanner data. Braunschweig: Technische Universität Carolo-Wilhelmina (2007).