Proceedings of the 2012

Ro-man 2012 Workshop

http://www.aass.oru.se/~ali/srt2012/

Social Robotic Telepresence

September 9th, 2012 - Paris, France

Co-Chairs

Silvia Coradeschi, Örebro University, Sweden
Gabriella Cortellessa, CNR – Italian National Research Council, ISTC, Italy
Annica Kristoffersson, Örebro University, Sweden
Amy Loutfi, Örebro University, Sweden

Preface

The aim of Social Robotic Telepresence is to increase presence via embodiment in a robotic platform for the particular purpose of enabling social interaction between humans. Following the success of the first edition of the workshop held in Lausanne, Switzerland in conjunction with HRI'11, this 2nd edition of the workshop continues to examine important research questions in the design, development and evaluation of mobile robotic telepresence systems.

In particular, the 2012 workshop on SRT is characterized by a variety of social robotic telepresence systems whose design ranges from anthropomorphically inspired system to low-cost and light weight systems that are market ready. Evaluation of SRT systems and evaluation methodologies also vary but is a continued topic of interest in this field. This year's workshop edition features suggestions for evaluation methodologies that are particularly application motivated, with focus in applications for the elderly. Further, evaluation methodologies address the study of interaction between a human actor and the robot per se, and the interaction between the remote user and the interface used to connect to the robot. Evaluation techniques are inspired by adjacent fields in psychology, and human computer interaction and motivated by the desire to perform longitudinal analysis in real life settings.

Finally, an evitable trend in the design of robotic telepresence is the inclusion of add-on features made possible by new technological development. Technologies such as android telephones are enabling new interfaces to be used with the social robotic telepresence systems in a variety of ways. This year, another featured enhancement is the effect of added expressive movement to enhance interaction and communication.

The contributions of this year's workshop edition are organized according to three categories: Methodologies in Evaluation of Social Robotic Telepresence; Design of Social Robotic Telepresence Systems; Evaluation of Social Robotic Telepresence.

We thank the Ro-MAN 2012 workshop chairs and program committee for their support. We thank the invited speakers and moderator who agreed to lead the discussion during the workshop.

Co-chairs, Social Robotic Telepresence Workshop 2012

Proceedings of the Ro-man 2012 Workshop on Social Robotics Telepresence

Editors:

Silvia Coradeschi
Örebro University, Sweden
Center for Applied Autonomous Sensor Systems
School of Science and Technology, Örebro University
Fakultetsgatan 1, 701 82, Örebro Sweden
Email: silvia.coradeschi@oru.se

Gabriella Cortellessa,
CNR – National Research Council of Italy
Istituto di Scienze e Tecnologie della Cognizione
Via S. Martino della Battaglia, 44, Rome Italy
Email: gabriella.cortellessa@istc.cnr.it

Annica Kristoffersson,
Örebro University, Sweden
Center for Applied Autonomous Sensor Systems
School of Science and Technology, Örebro University
Fakultetsgatan 1, 701 82, Örebro Sweden
Email: annica.kristoffersson@oru.se

Amy Loutfi,
Örebro University, Sweden
Center for Applied Autonomous Sensor Systems
School of Science and Technology, Örebro University
Fakultetsgatan 1, 701 82, Örebro Sweden
Email: amy.loutfi@oru.se

The Workshop is supported by the EU Ambient Assisted Living Joint Program through the project ExCITE "Enabling SoCial Interaction Through Embodiment" (http://excite-project.eu/)

Consiglio Nazionale delle Ricerche
Istituto di Scienze e Tecnologie della Cognizione
Via S. Martino della Battaglia, 44 – 00185 Rome, Italy

ISBN 978 – 88 – 85059 – 29 - 0

Table of Contents

Producing Expressive Movement for Telepresence Robotics

David Sirkin and Wendy Ju

***Abstract*—Much as people perceive body language during a face-to-face conversation, they perceive embodied cues—such as gaze, orientation, subtle movements and proximity—when interacting with telepresence robots. But while people have learned to ignore the incidental behaviors enacted by other people, they interpret such behaviors from telepresence robots as intentionally communicative. Knowledgeable remote pilots can *explicitly* control these cues, but performing these motions distracts them from actively participating in the topic of their conversations. On the other hand, *implicitly* controlled cues (such as provided by motion tracking) reproduce most, if not all, of the pilot's incidental behavior, causing distraction and confusion among the other people interacting with the robot. We believe that motion tracking is a starting point in the design of interfaces for expressive movement, but on its own is an incomplete approach.**

I. Introduction

To explore how people interpret local and remote action in geographically distributed, telepresent work teams, we have been running a series of studies focusing on motion, interaction and collaboration. Through building and testing several of our own telepresence robot platforms [1] and interactive devices [2], we discovered that a great, under-studied challenge is how to invest telepresence robots with expressive, communicative behaviors, and how to afford control over these behaviors to their remote pilots. Relying on direct control through, say, a joystick, leads to infrequent use of expressive (as opposed to pragmatic) motion, while naïve automatic replication of a remote participant's gestures leads to a wide range of distracting and misleading cues for the local collaborators [3].

II. Dilemmas in Telepresence Robot Control

To date, the primary purpose of motion for robotic telepresence has been to provide mobility through and around physical space. Most commercial and many research systems have adopted a form of *explicit* control, using either an on-screen or physical joystick [4, 5, 6] to enable pilots to move their robots around. But while this approach is familiar and easy to learn, it also distracts pilots from their robots' intended purpose—to immerse remote workgroup members in the meetings and activities at the local hub location. Lee and Takayama [4] found that when navigating, pilots "could not focus on the social conversation as their attention was divided between driving the system and carrying on a conversation." And our own studies found that explicit control "incurred a delay in operating the interface" [3] which influenced interaction patterns. In these ways, explicit control

David Sirkin is with Stanford University, Stanford, CA 94305 USA (email: sirkin@stanford.edu).

Wendy Ju conducts research at Stanford University, Stanford CA 94305

Wendy Ju conducts research at Stanford University, Stanford CA 94305 USA and Willow Garage, Menlo Park, CA 94025 USA (email: wendyju@stanford.edu).

Figure 1: This telepresence robot prototype supports live audio and video of a remote collaborator on a remotely-controlled, movable screen. It has a robotic arm, allowing the collaborator to point and gesture expressively.

actually increases the cognitive load on pilots, who find themselves shifting focus between the semantics of controlling their robot's movement, and the content of their coworkers' conversations. Any time they wish to move their robotic proxy—or even adjust its gaze—pilots have to look away from their view of these coworkers to find controls and click buttons. To the local conversational partners, the robot's movements appear stilted, and don't coincide with the pilots' on-screen gestures and behaviors. A physical body, whether human or robotic, communicates information about that person's state and intentions, and we have found that perceptions suffer when verbal and nonverbal cues conflict, or when robots are not able to fluidly express the embodied cues that people exhibit in everyday exchanges.

In an attempt to provide more expressive behavior, we and others in the research community have explored a form of *implicit*, gesture-based control [3, 7]. In this approach, a camera tracks the pilot's head and/or body motions and mirrors those on the robot. These systems do improve how robots project their pilots' focus-of-attention cues and expressive states, and can result in improved interaction quality, but at the cost of increased cognitive load on the conversational partners. While these robots do enact their pilots' intentionally communicative motions (nodding, for example), they also reproduce their pilots' incidental motions. People move *a lot*, often without being aware of doing so. A robot may look downward as its pilot scratches her ankle, or its head may turn to the side as its pilot faces someone in her own physical space. This leaves hub coworkers wondering which one of them the robot is looking at now, or what is so interesting in that corner of the room where the robot's screen just turned. Our experience shows that people typically interpret such unintentional motion as communicative.

III. Distinguishing Navigation and Expression

Navigative motion that provides room-scale mobility is quite different from gestural motion that expresses near-scale proxemics. When a pilot is speaking with several coworkers, rotating the robot's torso around its base and leaning forward provides an engagement cue, turning the robot's head from side to side projects immediate changes in focus of attention, and quickly shifting vertical posture from slumped to upright suggests alertness. Such conversational behaviors require different mechanical degrees-of-freedom, programming and (therefore) interfaces than those needed for driving around.

These interfaces may have to mediate between the potentially conflicting representations of the pilot's actions visible on-screen and the robot's physical motions—even if that conflict is merely a lack of coordination between them [2]. For example, the pilot may turn his head to the side, or nod, but the robot may not mirror that action. As long as telepresence robots rely on flat display screens, the Mona Lisa effect makes it likely that such conflicts between on-screen and physical orientations cannot be completely resolved [3]. However, it does raise the issue of how such cues can be mitigated, or overridden, so that pilots can present themselves as they intend.

IV. Implications for Interface Design

We recommend that telepresence roboticists explore the design of interfaces that neither interrupt the flow of natural conversation nor reproduce every action that a pilot takes—specifically, algorithms that intermediate the explicit (manual) and implicit (gesture-tracked) forms of control. We thought of a few example approaches:

- ***Account for context:*** Adapt gesture-tracking so that only pilot motions that cross some *activity threshold* map to robotic motion. The threshold may be determined by motion scale, direction or time, as well as the body movements involved. For motion scale, smaller gestures could be ignored while larger gestures would be mirrored, as the former are more likely to represent unintentional motion, while the latter are more likely to convey usable information. For direction or time, turns of the pilot's head to either side could be ignored for a brief period—the *dwell time*—after which they most likely do represent shifts in focus of attention, and should be reflected in the robot's actions.

- ***Provide a manual override:*** In many instances, pilots may be aware that they are about to behave in a way that won't make sense to hub coworkers, such as speaking to someone nearby at their remote location. Interfaces could provide an override (a clutch) that disengages the link between motion tracking and the robot's actions. Another advantage is that someone who knows that she fidgets a lot might prefer to leave tracking disengaged most of the time, and engage it only when she wants to be particularly expressive. Similarly, someone who remains relatively still most of the time might leave tracking engaged, and disengage it only when he wants to shift position.

- ***Infer focus of attention:*** Alternatively, it may be possible to programmatically determine whether pilots are engaging with their local environment or interacting with distant colleagues based on their gaze. This knowledge can be used to determine which of the pilot's actions should be reproduced by the robot, whether to present an alternate view of the pilot, or even whether to switch the robot into a standby state that indicates that the pilot has temporarily disengaged from the interaction.

- ***Moderate or amplify motion:*** Mirroring the full range of the pilot's movements may be unnecessary. A small, leading gesture—just enough to hint at a change in attention, or a deictic reference—may be all that's required to communicate intent. Moving the robot at a lower speed, or with smaller range of motion, relative to the pilot's movements can reduce motor noise as well as visual distraction. It also permits the pilot to adjust how the robot responds on-the-fly as a conversation progresses. If the group becomes more animated, the pilot can change a setting to produce faster or larger scale movements. What otherwise would have been a soft tap against the table can become a loud knock.

V. Toward Expressive Movement

The idea of the implicit interface is so appealing: that you could just *be yourself* and the robot would express that. But interface designers are faced with a dilemma: explicit control imposes cognitive effort on the pilot and takes him out of the moment, while implicit control shifts the cognitive load onto local collaborators, who try to interpret the meaning behind each of the robot's movements. Still, we believe that implicit interfaces should be part of any solution, only tempered by some form of manual or adaptive control. The interaction design community can contribute to our exploration by including such adaptive control conditions in telepresence studies. In addition, we encourage evaluation of whether local participants continue to perceive unintentional robot motion as communicative, after repeated interactions.

References

[1] Sirkin, D. and Ju, W. Consistency in physical and on-screen action improves perceptions of telepresence robots. *Proc. HRI 2012*, ACM Press (2012), 57-64.

[2] Ju, W. and Sirkin, D. Animate objects: How physical motion encourages public interaction. *Proc. PERSUASIVE 2010*, Springer (2010), 40-51.

[3] Sirkin, D., Robertson, G., Venolia, G., Tang, J., Kim, T., Inkpen, K., Sedlins, M., Lee, B. and Sinclair, M. Motion and attention in a kinetic videoconferencing proxy. *Proc. INTERACT 2011*, Springer (2011), 162-180.

[4] Lee, M. and Takayama, L. Now, I have a body: Uses and social norms for mobile remote presence in the workplace. *Proc. CHI 2012*, ACM Press (2012), 33-42.

[5] Yankelovich, N., Simpson, N., Kaplan, J. and Provino, J. Porta-Person: Telepresence for the connected conference room. *Ext. Abstracts CHI 2007*, ACM Press (2007), 2789-2794.

[6] Paulos, E. and Canny, J. PRoP: personal roving presence. *Proc. CHI 1998*, ACM Press (1998), 296-303.

[7] Adalgeirsson, S. and Breazeal, C. MeBot: A robotic platform for socially embodied telepresence. *Proc. HRI 2010*, ACM Press (2010), 15-22.

Using a Mental Workload Index as a Measure of Usability of a User Interface for Social Robotic Telepresence

Andrey Kiselev[1] and Amy Loutfi[2]
School of Science and Technology, Örebro University
SE-701 82 Örebro, Sweden
[1]andrey.kiselev@oru.se
[2]amy.loutfi@oru.se

***Abstract*— This position paper reports on the use of mental workload analysis to measure the usability of a remote user's interface in the context of social robotic telepresence. The paper discusses the importance of remote/pilot user's interfaces for successful interaction and presents a study whereby a set of tools for evaluation are proposed. Preliminary experimental analysis is provided when evaluating a specific teleprence robot, called the Giraff.**

I. INTRODUCTION

Various interactions take place simultaneously when humans are communicating through a robotic telepresence system. These interactions include human-robot interaction between local user and communication device (i.e. remotely controlled robot), human-human interaction between two or more users, and not least human-computer interaction between the remote user and robot's remote interface [1]. This paper is part of a study series which focuses particularly on the latter interaction and presents an ongoing project on design, implementation and evaluation of the user interface for the Giraff robotic telepresence system [8], [9], [11].
The approach taken in this work combines different techniques for measuring interface usability. Some methods used in this work are standard for evaluating usability of "office" applications, but others are normally used for drivers' and pilots' productivity assessment. The rationale for taking these measures into account comes from the fact that driving the robot is a secondary task for remote users while the primary is communication between remote and local users. Thus driving a robot should remain mentally and physically non-demanding. In this light, performance or quality of interaction will not be classified as a good result if the overall mental workload of the subject was also high.

II. BACKGROUND

Giraff pilot's interface allows remote users to establish connection with a Giraff robot, drive it and interact with local user(s) through the embedded video-conferencing system. Example of the driving screen can be found on Fig. 1.
Driving the robot can be done by using mouse, touchpad or any other standard pointing devices. The approximate tentative trajectory is drawn as a red line on a video panel. When left mouse button is pressed and held, the line transitions to green and the robot starts driving. The robot's direction and speed are controlled by the orientation and length of the line respectively. The robot's head tilt is controlled by dragging mouse pointer to the upper or lower parts of the video panel. Tilt can be adjusted at any point during driving.
It is also possible to drive the robot backwards by either using mouse or special button on the left panel. Alongside with it another button can be found which is used to rotate the robot 180 degrees counter-clockwise. Rotation can also be done by double-clicking a left mouse button on the left or right parts of the video panel. In this case the rotation angle is calculated according to the position of the mouse pointer on the panel. The left panel also contains controls (slide bars) for remote and local volume adjustment, call management button, battery information and local image.
The Giraff pilot's interface is a combination of normal "office" application look-and-feel, which target users are supposed to be familiar with, and robot remote control functionalities. Thus a combination of usability assessment methods must be used in order to comprehensively evaluate

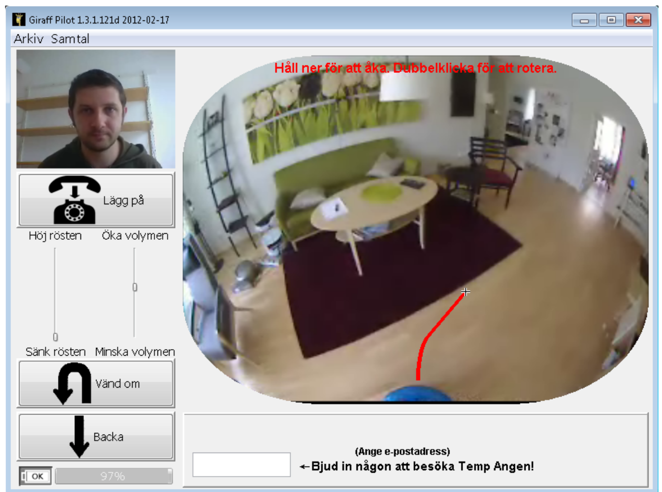


Fig. 1. A screenshot of the application's main driving window.

[1]By "local user" authors assume one who is physically located in the same environment with a robot. Thus, "remote user" is one who controls the robot from a remote location.

this interface. Particularly, in this work authors partially follow [1] in order to assess the Giraff pilot's interface usability from the perspectives of efficiency and effectiveness of achieving goals and conduct mental workload analysis as a joint reflection of users' satisfaction and performance.

III. METHOD

The main application of the Giraff robot within the context of the ExCITE project [10] is serving as a movable communication device between elderly and remote visitors. The two tasks remote users usually perform are interacting with local user (e.g. elderly) and controlling robot's behaviour.

Complete typical real interaction scenario can be divided into several stages and some typical pilots' actions can be extracted. For instance, such typical actions include undocking, driving a robot (following a person, following a path), finding objects, docking. Some of the actions might be quite challenging depending on the pilots' experience and technical limitations[2] of the platform.

One of the situations when pilots are usually faced difficulties is when they have to avoid collisions with some objects in the environment. This requires them to feel size of the robot and distance to obstacle. This problem comes from the mechanical design of the robot, which uses wide angle lens for capturing bigger scene. Another typical task which makes problems for pilots is connecting the robot to the docking station. Although the docking station is designed in such way that it is tolerant to some degree of robot misalignment, this still require precise controlling and good feeling of size and distance.

In the current experiment the performance of novice users' in performing some typical tasks was measured. The measurements were done by analysing time spent by subjects to drive the robot between checkpoint (please, see the detailed description of the experiment in Section IV) and number of collisions made on each part of the path. The performance measurements were supplemented with the mental workload analysis, which was measured with the NASA TLX test [2], [3]. Although interactions with local users is a typical task for a Giraff system, pure driving performance is vital for pilots to successfully accomplish more sophisticated interaction tasks.

Additionally authors use a profiling questionnaire that collects demographical data such as age, gender, education, and usage experience with communication and electronic products (phone, computer, DVD, Skype, video games, cameras, and other). The education level was obtained according to the ISCED 2011 [4] in order to allow conducting further comparative experiments in other countries with different standards of education levels.

The questionnaire which is used in this experiment for UI evaluation is the USE Questionnaire [5]. This questionnaire is selected among others (such as Computer System Usability Questionnaire [6]) because it allows to comprehensively evaluate the target system in terms of its usefulness, ease of use, ease of learning, and satisfaction. The Use Questionnaire is a widely used tool for UI evaluation and results can be easily correlated with other studies. One of the known issues with the USE Questionnaire (and with a number of others well known questionnaires) is that it suffers from "acquiescence bias" [7], which must be considered when analysing final results of the experiment.

Not all the questions of the USE Questionnaire are applicable to the experiment. For instance, the "Usefulness" section cannot be considered as a valuable measure since subjects did not have any strong demand for using the Giraff in their daily life. The sixth question in the "Satisfaction" section ("I feel I need to have it.") can not be used for final results for the same reason. All other questions are applicable to the experiment.

IV. EXPERIMENT

The experiment was conducted in the "Ängen intelligent home" for elderly between 3-d and 4-th of May 2012.

Ten subjects participated in the experiment, six males and four females, average age is 40.7, SD 15.2. Subjects represent different user groups, have different exposure to technology, but none of them have prior experience with using Giraff pilot's interface.

The 35-meter path was drawn in the apartment with bright blue dashed line with arrows. The path had several key points: docking station (DS), bedroom checkpoint (B), kitchen checkpoint (K), fridge checkpoint (F), goal (G). The scheme and a photo of the path can be found in Fig. 2. Subject start from the docking station, then they visit bedroom checkpoint and kitchen checkpoint. At the fridge checkpoint they have to read a task. The task for this experiment is to find a circle with number 1 inside somewhere on the floor in the living room. This is the goal checkpoint. Its position is the same for all subjects and its main role is to be a reference point for docking performance measurements.

The complete procedure of the experiment for each participant consists of several stages. First, each participant was shown a short introductory film about the Giraff system and pilot's GUI. The total length of the film is 2:12 sec. Then each participant was given verbal instruction supplemented by a screenshot about how to drive Giraff, which controls should they use. After that the driving section began.

During the driving section each subject had to drive through all checkpoints until they reach the fridge checkpoint. There they had to read their task ("Go to 1") written on the fridge. Then they had to find the G checkpoint in the living room and dock robot from that point back to the docking station. Driving sections were filmed for further analysis. At each part time to approach checkpoint and number of collisions

[2]By "technical limitations" we assume those that are derived from the robot's design, such as robot's physical dimensions or camera resolutions, and also those which come from environment. For instance, low video quality or temporal inoperability might be caused by poor internet connection between the robot and remote users.

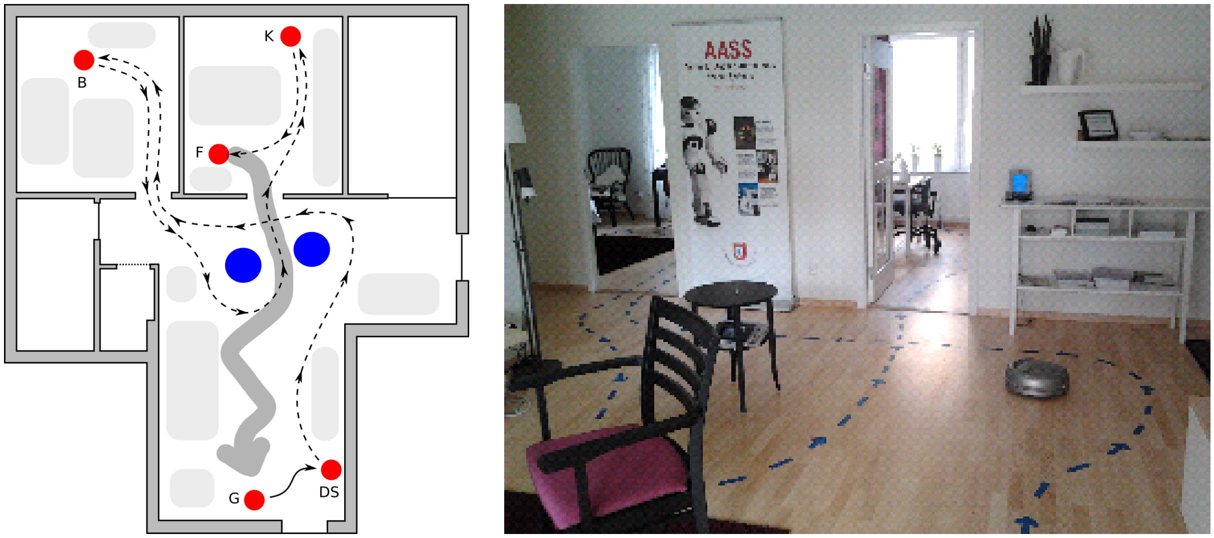


Fig. 2. *Left:* Outline of the Ängen apartment. Dashed line shows a path on the floor which subjects had to follow. Medium gray path - free driving which searching the object on the living room. Red circles - checkpoints: DS - docking station; B - bedroom checkpoint; K - kitchen checkpoint; F - fridge checkpoint; G - goal. Blue circles - artificial obstacles (coffee table, iRobot Roomba); light gray - other obstacles in the environment. *Right:* An example of the real environment.

were calculated.
At the final part of the experiment subject were asked to fill questionnaires: first NASA TLX, then profiling and finally the USE Questionnaire. All the questionnaires were administered through a web-page [12].

V. RESULTS

A. Performance

Results of the performance analysis along with NASA TLX score and average results of the USE questionnaire can be found in Tab. I[3]. Authors would like to refrain from providing any final conclusions based the results of this initial experiment.

B. Observations and user reports

Observing user behaviours along with collecting user reports and opinions is an important step in UI evaluation. This subsection summarizes our findings, derived from video analysis and conversations with the participants. It is important to remember that these observations only derived from the current experiment, reported here. Correlating user reports across several studies is a subject for further investigation.

1) Video resolution / quality: It was clearly seen while setting up the experiment that when our task is written by pen or pencil it is simply can not be recognized by remote users. Authors had to use more contrast black marker and large font in order to make the task visible.

[3]The column “Confusion” shows whether or not subject was confused with undesirable tilt or moving-backwards robot’s behaviours. TLX stands for the NASA TLX test score. USEQ stands for the average score of USE questionnaire (applicable questions only)

2) Control over robots behaviour: Two subjects, who have experience with computer games, reported that they would want to have more control over robot’s behaviour and using keyboard seems to be more convenient for them. At the same time other participants reported that they are happy with current mouse-based control as it does not require any specific skills to control the robot.

3) Pointing at objects of interest: One of the most important observations shows that all subject tend to at least initially click at the point of interest (e.g. docking station or checkpoint) by mouse pointer when they start driving.

VI. FUTURE WORKS

The main objectives of this initial experiment, which are a) to establish a general procedure for the Giraff pilot interface evaluation, and b) to provide a reference point for future interface evaluations, are achieved. The method provides useful information for interface refinement and will be used for further evaluations. For instance, it was clearly observed, that screen tilt functionality should be implemented in a different way. Nevertheless, the current user interface is easy to learn and use, which is clearly seen from the results of the USE questionnaire and supported by mental workload analysis and users’ reports. Although the proposed method looks promising in principle, authors are interested in adding objective mental workload or either user satisfaction measurements into the current procedure if such measurement techniques are considered usable within the scope of the entire project. Also the USE questionnaire must be refined as well to overcome its known bias problem and applicability.

ACKNOWLEDGMENT

This work has been done within the context of the ExCITE project which is supported by the EU under the

TABLE I
PERFORMANCE MEASUREMENTS AND RESULTS OF THE NASA TLX.

Subject ID	Age	Gender	Performance, seconds						Collisions	Confusion	TLX	USEQ
			DS - B	B - K	K - F	F - G	G - DS	Overall				
1	21	male	86	92	54	49	60	341	1	No	37	95
2	20	male	100	86	38	43	19	286	0	No	16	86,8
3	27	female	81	74	20	126	36	337	5	Yes	55	64,1
4	50	male	211	99	34	138	25	507	3	Yes	61	84,9
5	42	male	123	70	27	51	28	299	2	No	48	94,2
6	47	male	74	74	30	58	37	273	2	No	47	73,1
7	67	female	183	127	55	62	168	595	4	Yes	62	78,5
8	55	female	108	88	124	44	30	394	1	Yes	21	59,3
9	35	male	137	90	32	66	53	378	1	No	49	78,2
10	43	female	114	110	35	79	72	410	1	Yes	68	64,9

Ambient Assisted Living Joint Programme (AAL-2009-2-125)

REFERENCES

[1] ISO9241-11:1998(E) Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability.

[2] Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(9), 904-908. Human Factors and Ergonomics Society.

[3] Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. Applied Psychology, 53(1), 61-86. Wiley Online Library.

[4] International Standard Classification of Education 2011, UNESCO

[5] Lund, A. M. (2001). Measuring Usability with the USE Questionnaire. Usability Interface, 8(2).

[6] Lewis, J. R. (1995) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction, 7:1, 57-78.

[7] Messick, Samuel; Jackson, Douglas N (1961). Acquiescence and the factorial interpretation of the MMPI. Psychological Bulletin, Vol 58(4), Jul 1961, 299-304.

[8] A. Kristoffersson, S. Coradeschi, K. Severinson Eklundh, and A. Loutfi (2011). Sense of Presence in a Robotic Telepresence Domain. Universal Access in Human-Computer Interaction. Users Diversity, pp. 479–487, 2011.

[9] A. Kristoffersson, K. Severinson Eklundh, and A. Loutfi (2012). Measuring the Quality of Interaction in Mobile Robotic Telepresence: A Pilot's Perspective. International Journal of Social Robotics, 2012.

[10] S. Coradeschi, A. Loutfi, A. Kristoffersson, S. Von Rump, A. Cesta, and G. Cortellessa. Towards a Methodology for Longitudinal Evaluation of Social Robotic Telepresence for Elderly. In Proc. of Human-Robot Interaction Wksp. on Social Robotic Telepresence, 2011.

[11] A. Kristoffersson, S. Coradeschi, A. Loutfi, A. and K. Severinson-Eklundh (2011). An Exploratory Study of Health Professionals' Attitudes about Robotic Telepresence Technology. Journal of Technology in Human Services, 29:4, 263-283, Taylor & Francis.

[12] Giraff Pilot User Interface Evaluation Questionnaire. Temporary located at: http://hej-hej.tw1.ru/

Into the Wild: Pushing a Telepresence Robot Outside the Lab*

Amedeo Cesta[1], Gabriella Cortellessa [1], Andrea Orlandini [1] and Lorenza Tiberio[1]

***Abstract*— Most robotic systems are usually used and evaluated in laboratory setting for a limited period of time. The limitation of lab evaluation is that it does not take into account the different challenges imposed by the fielding of robotic solutions into real contexts. Our current work evaluates a robotic telepresence platform to be used with elderly people. This paper describes our progressive effort toward a *comprehensive*, *ecological* and *longitudinal* evaluation of such robots outside the lab. It first discusses some results from a twofold short term evaluation performed in Italy. Specifically we report results from both a *usability* assessment in laboratory and a subsequent study obtained by interviewing 44 healthcare workers as possible secondary users (people connecting to the robot) and 10 older adults as possible primary users (people receiving visits through the robot). It then describes a complete evaluation plan designed for a long term assessment to be applied "outside the lab" dwelling on the initial application of such methodology to test sites in Italy.**

I. Introduction

The area of social robotics is receiving increasing attention and the task of "robot as companion" has received attention at research level [1]. Several projects have also proposed different types of solutions with robots that both interact with humans and are connected to heterogeneous technology to build innovative living environments (e.g., [2], [3], [4]). This paper aims at underscoring one aspect connected to such a line of innovation that deserves special attention: the study of *attitude and perceptions of people who share the environments in which the robot operates* over *long periods of time*.

It is also worth noting how in robotics there is a deep-rooted tradition in developing technology usually shown in sporadic events and for short periods, i.e., for demos or live show cases. These demonstrations usually aims to present the "enhanced" characteristics of a prototype, making them attractive while "hiding" or at least "containing" the technical problems connected with any long term use within a comprehensive application. Indeed, a key requirement for social companions (e.g., robots assisting older adults at home) is their continuous operation, their robustness and the continuous interaction with humans over time. Such continuity of use has significant implications on the technology development but it also highlights the need to design a methodology for assessing human reactions with respect to prolonged use of the proposed solutions. The challenges for the Intelligent Technology and the Human Robot Interaction researchers are numerous and mainly related to two aspects: (a) in terms of *users perspective*, robots must adhere to user requirements and be acceptable in the long term, (b) in terms of *technology*, the need exists to create usable, robust, efficient and secure solutions. More specifically, the transition from a use in the laboratory to an actual deployment into real contexts, highlights the need for a shift from short term to long term experiments. In particular we underscore how *long-term use* and *evaluation* are key points to be addressed to ensure that robotic technology can make a leap forward and be used in real environments.

In the framework of the EU Ambient Assisted Living (AAL) Joint Program[1] we are part of a project called ExCITE[2], which is performing a wide program of evaluation in the field of an industrial mobile telepresence platform called Giraff produced by Giraff Technologies AB[3], Sweden. More specifically, we take part in an evaluation spanning three different EU countries – Italy, Spain and Sweden. The evaluation takes social and psychological factors into account to study users attitude and reaction, but also analyzes the emergence of "undesired behaviors" like technological weaknesses in continuous operation, possibly leading to human rejection. In this work, we present the results gathered in Italy after the short term evaluation phase and, then, we present and discuss the general long term evaluation methodology, showing its current application to real test sites. The paper[4] introduces the context of work (Section II), then analyzes and reasons about the work both to realize short term experiments with real users and to develop a methodology for addressing long term evaluation (Section III, Section IV, Section V); finally it describes the status of the first test sites in Italy where the long-term evaluation methodology is being applied (Section VI).

II. Context of work

Telepresence robots have been increasingly proposed to be used in workplace and Mobile Remote Presence (MRP) systems have been studied as a means to enable remote collaboration among co-workers [5], [6]. Furthermore, MRPs are also being used to provide support to elderly people. In this respect, some research exists which aims to understand the acceptance of older adults, their concerns and attitude toward the adoption of MRPs [7], [8], [9]. Our work is motivated by the participation to the ExCITE project, aiming at promoting the use of MRPs to foster interaction and social

*Authors are partially supported by the EU under the Ambient Assisted Living Joint Program – ExCITE Project (AAL-2009-2-125) – and under the ICT Call 7 – *GiraffPlus* Project (GA 288173).

[1]Authors are with CNR, Italian National Research Council, ISTC, Rome, Italy. http://www.istc.cnr.it/

[1]http://www.aal-europe.eu/

[2]http://www.excite-project.eu/

[3]http://www.giraff.org

[4]This is a revised version of this work: A. Cesta, G. Cortellessa, A. Orlandini, L. Tiberio. Evaluating Telepresence Robots in the Field. To appear in a special issue of Springer's LNCS TCCI Journal.

participation of older adults as well as to provide an easy means to possible caregivers to visit and interact with their assisted persons in their living environment. GIRAFF is a remotely controlled mobile, human-height physical avatar integrated with a videoconferencing system (including a camera, display, speaker and microphone). It is powered by motors that can propel and turn the device in any direction. An LCD panel is incorporated into the head unit. The robotic platform is accessed and controlled via a standard computer/laptop using a software application. From a remote location the *client*, or *secondary user* (member of family or healthcare professionals) with limited prior computer training teleoperates the robotic platform while older adults (*end users, or primary user*) living in their own home (where the robot is placed) can receive their visit through the MRP. The remote user can charge the robot batteries by driving it onto a docking station.

Key pursued ideas

The EXCITE project aims at assessing the validity of an MRP in the field of elderly support in different European countries. The project fundamental concepts are the following:

- *User centered product refinement*. This approach is based on the idea of obtaining users feedback during the time they use the robot and cyclically refine the prototype in order to address specific needs;
- *User tests outside labs*, rather than testing the system in laboratory setting, the MRP is placed in a real context of use. This approach is in line with several research that highlights how systems that work well in the lab are often less successful in real world settings [10]. The evaluation of robots made in a laboratory environment does not favor the emergence of robotic aid suitability to support elderly who are able to stay in their own homes. For this reason an essential step is to assess the technology in the specific contexts in which the technology is supposed to be used [11];
- *Use on a time period long enough*, to allow habituation and possible rejection to appear. Indeed, interviews and survey conducted after a short period of time, though useful and valuable can not be the only way to assess techonology since they can be limited and can prevent other effects to emerge. On the contrary, a key aspect of relationship is that it is a persistent construct, spanning multiple interactions [12]. In this light, in order to assess the human-robot interaction it is important to investigate how people interact with robots over long periods of time.
- *Analysis of cultural and societal differences*, an interesting part of our project stems from the idea of comparing the long term deployment of the telepresence platform in different countries so as to allow an analysis of cultural and societal differences over European countries.

Different GIRAFF prototypes are being deployed for long periods of time (at least three months, and possibly 1 year) in real context of use. Feedback obtained from the users (both primary and secondary) is used to improve the robot. In what follows, we describe our progressive work toward a long-term human-robot interaction assessment showing how we combing short term evaluation sessions with long term efforts.

III. THE EVALUATION APPROACH

We have conceived a twofold path for evaluating the human-robot interaction gathering both feedback from short interactions between potential users and the GIRAFF robot and also focusing on a long term assessment plan. More specifically we identified two tracks for our effort:

- *Short Term Evaluation*, which consists of a collection of immediate users feedback (i.e., after a short interaction with the robot) on the telepresence robot, connected to different aspects of the interaction mainly related to the usability, willingness to adopt it, possible domains of applications, advantages and disadvantages.
- *Long Term Evaluation*, which relates to the study of the long-term impact of GIRAFF's social and physical presence on elderly users using the system both to communicate with their relatives and friends and to receive visits from healthcare providers and in general caregivers.

The short term evaluation effort, though not sufficient alone, still provides immediate feedback that can be used to quickly improve the technological development, to possibly add functionalities to the system or to simply confirm the validity of some technological choices. In addition it can give valuable guidance to the long-term assessment. For this reason we adopted a combined approach involving participants representative of both types of users: the *secondary* and *primary* users.

Following this schema, we first present results for the short term evaluation performed in Italy, then we introduce our complete design for a methodology to assess the long-term impact of the GIRAFF in EXCITE also reporting the status of the Italian long-term test sites that are currently running according to this methodology.

IV. SHORT TERM EVALUATION

For the short term evaluation effort we first realized some usability experiments in laboratory, so as to identify possible problems in the user interaction with the system. Subsequently we organized user evaluation sessions with real potential users of the system to investigate other complementary aspects.

A. Usability evaluation

The usability assessment has been made by using both an observational technique and a usability questionnaire. Specifically, we relied on the *Thinking Aloud* evaluation technique [13], which consists of asking the users to verbalize their thoughts while performing certain tasks and interacting with the system. The experimenter observes silently the interaction session, and records user's actions and thoughts, focusing on the difficulties and problems encountered. In addition, the System Usability Scale (SUS) [14] was administered as an additional measure[5].

1) Participants and procedure: five participants took part in our usability experiment (see Figure 1). Four of them were male students (with a mean age of 18,4) and one was their teacher (male, age 54)[6]. All the participants had experience in software and computer and received training prior to the test consisting of a tutorial presentation of 20 minutes and a practical session. After the tutorial each participant received written instructions on specific tasks and how to carry them out. Four main tasks have been considered that can be grouped as the following: (a) *make a video call*; (b) *navigate in the environment*; (c) *read a text through the robot*; (d) *perform the docking*.

During the sessions participants were encouraged to "think aloud" to verbalize their opinions while completing the assigned tasks. The sessions were recorded and the experimenter took notes during the session.

At the end of the test, the SUS questionnaire was administered and a final interview was conducted to understand opinions with respect to the telepresence system experience and to discover further problems and take note of additional advices. Also this interview was recorded. The recordings have been analyzed and experiment results have been written in the form of Usability Aspect Reports (UARs)[7].

2) Results: Overall the interface was judged usable, even though some specific problems still emerged. The detailed UARs have been examined and have been organized according to four main categories:

a) Video and audio: the control and audio quality were judged overall very good. The video instead has been considered not completely satisfactory. The quality seems, in fact, sufficient to allow for general navigation in the environment but not entirely satisfactory in case you need to perform specific visual inspections such as reading a text or recognize the state of some specific objects within the environment. One solution would be to improve the quality of the camera and also to provide it with a zoom feature.

b) Navigation: the navigation in the environment was generally satisfactory. Some difficulties were encountered when the robot had to move in extremely narrow spaces or with obstacles. A suggestion from participants regards the possibility to insert a map and a position indicator of the robot within the environment. This feature could possibly be superfluous in case the secondary user is a son or a person who knows the environment in which the elderly live. On the contrary, it would be particularly useful if the secondary users is a person less familiar with the explored environment (e.g. a formal caregiver or a health professional). In addition some autonomy for helping the remote operator of the robot, when the driving is more critical could ease the navigation.

c) Client Interface: the client interface was satisfactory. The commands for the control of the robot have been judged as clear and easily identifiable. A possible improvement concerns the indicator of the level of charge that could be implemented with a more visible color or through a flashing signal that would attract the attention when the battery is reaching a critical level.

d) Docking: this was the most critical functionality from the point of view of usability. At least half of the participants had difficulties in the docking. This is both because of poor video quality, and the manual docking conducted without visual aids. Possible solutions to this problem are: implementation of an automatic docking functionality or alternatively, providing the base with more visible indicators (e.g. colored) and simultaneously put directional indicators in the interface which can "guide" during manual docking.

As for the SUS usability questionnaire, results show that the GIRAFF application scored 77 of 100 points. Our result

(a)

(b)

Fig. 1. Pictures from the "Thinking Aloud" evaluation session: (a) Reading task; (b) Driving task

[5]The SUS instrument is a reliable tool for measuring the usability of a wide variety of products and services. It is composed of 10 statements that are scored on a 5-point scale of strength of agreement. Final scores for the SUS can range from 0 to 100 where scores above 70 indicate products which are at least passable. Scores in the high 70s to upper 80s guarantee products with a good acceptability. Greatly superior products score better than 90.

[6]The specific choice of this sample was motivated by the fact that the participants were somehow representative of the secondary users we had contacted for the long term test sites. Specifically, the main secondary users were: a man with experience in using PC and technology in general and young boys with skill in both computer usage and video games. Our plan is however to enlarge the sample size also considering other age brackets.

[7]The detailed UARs are not reported for the sake of space. They have been analyzed and grouped into four main categories.

can be interpreted as an index of a good acceptability and ease of use. Therefore, the general usability assessment was quite good, though some aspects could still be improved.

Some common aspects emerged also from the analysis of the content of semi-structured interview. Specifically, referring to the experience of use participants were asked to judge the interaction through the robot relying on a semantic differential with six adjective pairs on 6 point scale. The participants agreed in judging the telepresence experience as *active*, *participatory* and *exciting*. The GIRAFF's height was judged adequate but its base was considered cumbersome.

B. Assessment of primary and secondary users attitude

After the usability assessment results, we started involving possible users of the telepresence system in order to study their opinions on the use of telepresence systems. As stated in [7], before intelligent technologies would be accepted, it is important to understand their perception of the benefits, concern and adoption criteria. In our study, we aim at reproducing as much as possible an "ecological" setting for the experiment. To this purpose we distinguished the role of the users and recruited different participants according to their expected role. Specifically for the secondary users group we recruited users representative of the potential visitors of the elderly users among caregivers, nurses, health workers, etc. For the end user side we interviewed older adults living alone, or possibly receiving some kind of health care assistance. This evaluation was aimed at assessing users reaction toward the possible adoption of the GIRAFF system as a means to visit or provide some kind of service to the elderly users. Aspects investigated were *willingness to adopt the robotic solution*, *possible domains of application*, *advantages* and *disadvantages* and *suggestions for improvements*.

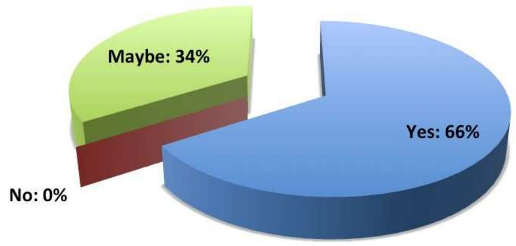
Health workers as secondary users

1) Participants and Procedure: forty-four health workers from different specialist areas were recruited for this study. The sample interviewed so far is composed by 26 women and 18 men with a mean age of 42 years, $SD = 12.2$.

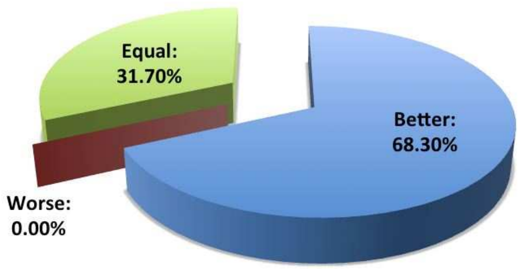
The meeting entailed a tutorial presentation of 20 minutes to describe features and functionalities of the telepresence robot. After this tutorial, a practical session allowed the health workers to operate the system and experience the different functionalities. Following the tutorial a focus group was conducted and a final questionnaire was administrated to assess possible applications of the telepresence robot, the perceived advantages and disadvantages of the system, the patient profile best suited to benefit from the use of telepresence.

2) Results: the results have been grouped according to the following categories:

a) General assessment: a first analysis of the results showed a positive reaction of the participants to the system. In particular 66% of participants would be willing to use GIRAFF as an aid support in his/her profession and no one opposes to the use of robots (see Figure 2a). In addition most of them judge the telepresence robot as a better tool



(a)



(b)

Fig. 2. General assessment of the GIRAFF system: (a) willingness to adopt it; (b) qualitative comparison with traditional teleconference systems like skype

with respect to traditional teleconference system like Skype (see Figure 2b).

b) Profile of potential users: results also identify the categories of people who could benefit from the use of telepresence robots: specifically, the category "self-sufficient or semi-autonomous elderly living alone" has been mentioned by 35% of respondents; 25% of the subjects also indicates "adults and elderly patients in home care and with special needs", such as patients in isolation for infection, dialysis patients or with chronic diseases such as Chronic Obstructive Pulmonary Disease (COPD) or diabetes. A 20% of the responses were grouped into the category "older adults with early or mild dementia". Two other categories were "adults or older adults with physical disabilities" (17%) and "young people and adults with intellectual disabilities" (7%).

c) Application domains: the participants are in favor of the use of robots to train the family caregiver to small nursing tasks and to maintain constant contact with assisted older adult. The possibility of *continuous monitoring* (see Figure 3) of the patient at home is considered the most useful application (59% of participants were in favor of this kind of application). The *support* application follows at 23%, while the *companionship* and *communication* applicative domains seem less suitable. More specifically, 45.5% of the health workers advocate the use of the robot to train a family caregiver to perform small nursing tasks (e.g., treat a bedsore, administer an enema, measuring of vital signs) and to maintain a constant contact with the patient and his family (75% of participants). Finally 60% of participants also says that the robot could alleviate the workload of the family caregiver, but not that of the health workers themselves (50% of people admit to be uncertain about the real possibility of the robot to diminish their daily workload).

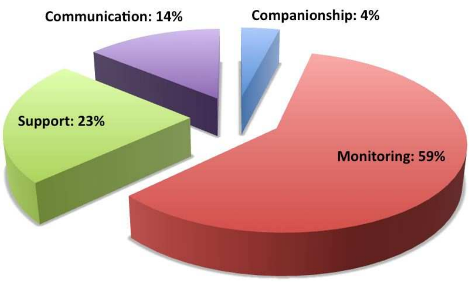


Fig. 3. Favorite GIRAFF's domains of application

d) Advantages and Disadvantages: among the advantages in using the robot, participants listed the following: a) ability to monitor remotely *via* visual communication the physical state of health; b) possibility to follow the management of medication and certain health practices (e.g., control of vital parameters such as level of blood glucose for diabetic patients, supervision of practices related to their care and medication like deep breathing exercises for patients with COPD); c) the possibility for the operator to improve his/her night surveillance activity in hospital and home care cases. Among the disadvantages they reported the poor quality of the video, the bulky size of the base unit, the fact that the robot might not be suitable for all patients, issues related to cost and privacy.

e) Suggested improvements: The focus group conducted at the end of this analysis, highlighted some aspects considered as particularly relevant for using the platform in the healthcare domain for long-term period. These aspects specifically refer to improvements and integration of additional functionalities. Specifically according to participants, the need exists to improve the video quality, especially in relation to night vision; it would be useful to add the zoom functionality to the webcam; the battery duration and recharging modality should be improved (e.g., it would be better if the robot could reach autonomously the docking station); the safe navigation of the robot should be guaranteed. In addition it would be beneficial to enable the call transfer if the client is not connected to the robot via the PC. Finally the transmission of vital parameters to the doctor should be supported. All these suggestions for technical improvements are currently inspiring the future modifications of the GIRAFF system in line with the user centered approach pursued in the EXCITE project.

Older adults as primary users

1) Participants and Procedure: To investigate aspects connected to the end-user interaction with the telepresence system we contacted 10 older adults. Four of them were potential end users who have been asked to participate in the long-term evaluation described in Section V. The remaining participants are involved in a parallel study, also connected to the project that aims to validate the GIRAFF system as a tool for providing remote rehabilitation [9].

The procedure followed in this qualitative research entailed an explanation of the main idea underlying the telepresence system, showing some descriptive materials, a video of the system and, where possible, a practical demonstration of the system itself. The selection of the material and the modality to present the system were decided according to the time availability, and the specific situation presented in each evaluation session. We recorded the interview and we then opted for a qualitative analysis, summarizing the main recurrent cited positive and negative aspects, given the relatively small number of the sample. A more structure study is in our future research plan.

2) Results: A qualitative analysis of the interview have been conducted and the most relevant feedbacks are here reported in terms of positive and negative aspects of the MRP.

a) Positive Aspects: Among the positive aspects most of the subjects reported the following: participants judged the visit through GIRAFF as engaging and "real"; the robot was pleasant to see; the ability of the robot to move in the environment was positively assessed; users felt physically involved during the interaction; participants think that the robot would help someone living alone at home to feel safer; participants judged positively both the audio and the video functionalities; participants think that interaction through the robot was spontaneous.

b) Negative Aspects: Among the most negative aspects we mention: the GIRAFF system is too big and consequently may not be well integrated in a domestic environment due to its size; the battery power may be too short; there may be some problems due to the privacy issue; there were some concerns related to the safe movement of the robot and to its ability of obstacle avoidance; some "intelligent features", like the autonomous recharging of the battery, are missing; the connection to the docking station is "not very intuitive".

Also this effort showed an overall positive reaction to the system, even though some improvements are desired in view of a real usage of the system. It is worth underscoring how the key point here is the fact that qualitative data has been gathered by interviewing "real potential users" like for example a group of caregivers and older adults who can receive visits through the robot.

V. LONG TERM EVALUATION

One of the original features of the EXCITE project consists of realizing long-term experiments involving older adults hosting the robot in their living environment both to communicate with others and to receive assistance services.

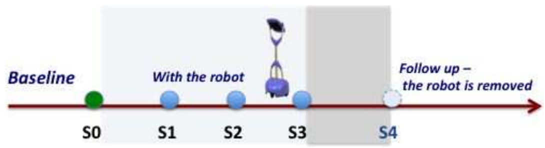


Fig. 4. The Long Term Evaluation timeline

A. Method

Figure 4 gives a general idea of the designed method to evaluate features over time. The evaluation entails a period of N months (with $3 \leq N \leq 12$) during which the end user will have the robot at home and the clients can visit him/her through it. Assessment happens at milestones S_i. Specifically, after an initial assessment (S0 in figure) at the beginning of the experimentation (*baseline*), the variables of interest are measured at regular intervals (S1-3) to observe changes over time. At the last month the GIRAFF will be removed from the end user apartment and the same variables will be assessed again after 2 months from this removal (S4). The general idea is to use a repeated measures method to see changes over time during the long term usage of the robot.

1) Participants and Procedure: Three different cases have been identified to cover different situations in which the robot can be deployed. Specifically, for the secondary user typology we considered (a) a *formal caregiver* belonging to an Health care organization; (b) a *family member (informal caregiver)*; (c) *other relatives or friends* who may visit the elderly person through the robot. The type of material used in the long term evaluation for both the client and the end user depends upon the type of interaction for which the telepresence is used. For this reason, for each of the three mentioned situations we had developed (or selected) a set of questionnaires (almost all validated in the three languages of the involved countries) aimed at monitoring specific variables and to be administrated at specific time both to end users and to clients.

2) Material: For each of the described case we prepared the material to assess the variables under study at the specified intervals. Table I lists in detail the different variables and the related instruments to be used to measure the variables over time.

a) Client side: Specifically on the client side, during the initial step (S0), we use: (a) an informed *consent form* describing the aim and procedure of the study; (b) the *socio demographic data* form to gather some relevant information on the user; (c) we developed on purpose a questionnaire aimed at assessing the client expectation on the GIRAFF's ability to ease the support (*Support Expectation*). It is worth highlighting that we developed two slightly different types of questionnaires for the *formal* and *informal* caregivers, while for the *other relatives and friends* category we designed a questionnaire (*Influence on Relationship Expectation*) on the expectation on GIRAFF as a means to ease and support the remote communication and consequently the social relationship.

During the following step (S1), for all three types of secondary users introduced above we will use: (a) questionnaires based on the SUS inventory [14] to assess the *usability* of the client software; (b) we will ask participant to keep a *diary* to register the "salient" events of the visit through telepresence in terms of encountered problems, good features and so on.

During the subsequent step (S2), in addition to the diary that clients have to keep along the whole experience with the robot, we make a first assessment of ability of GIRAFF to ease the support (or the communication) between the client and the end user through the *Support Assessment* and *Impact on Relationship Assessment* questionnaires. In addition, during this phase we will also use the Temple Presence Inventory [15] that is a tool to measure dimensions of (tele)presence and the Networked Minds Social Presence Inventory ([16]).

At step S3 we use the Positive Affect Negative Affect Scale, PANAS, [17], the Psychosocial Impact of Assistive Devices Scale, PIADS, [18] and a final structured interview to assess the overall experience in terms of the most relevant variables considered in the study.

After two months from the robot removal, S4 will allow assessing the impact of its absence through the *Support Assessment* questionnaire.

b) End user side: For the end user receiving the robot we followed a similar approach, but we focused on some additional variables that is worth dwelling on (see next table). Specifically, we measure: (a) the *perceived loneliness* through the UCLA Loneliness Scale [19], which was developed to assess subjective feelings of loneliness or social isolation; (b) the perceived health status through the Short Form Health Survey (SF12) [20]; (c) the Multidimensional Scale of Perceived Social Support [21]; (d) Geriatric Depression Scale [22]: a modified version of the Health Service Satisfaction Inventory. Finally the Almere [23] model will allow assessing dimensions of technology acceptance.

In the table I, measures highlighted in bold will ensure the repeated measures thus allowing to observe the GIRAFF's influence by changes in response over time. It is worth underscoring how this evaluation plan will allow monitoring the human-robot interaction over time, thus contributing to understand the long term impact of a fully deployed robotic solution.

The actual implementation of this plan in three different European countries will also support a cross-cultural analysis, continuing some work started on this specific topics [24]. The following section briefly reports on the current status of the Italian test sites.

VI. FIRST TEST SITES RUNNING

Two test sites have started in Italy that are representative of the *family-member-elderly* user category.

A. Test site 1

A couple of older adults living in the countryside near Rome are the end users of this test site (see Figure 5). The man has reduced mobility, while the woman has problems with her sight. They are quite independent although their health condition is slowly deteriorating. The secondary users are: their son living in Rome and their grandchild.

Fig. 5. A picture from the first Italian test site

We initially experienced some problems with the technical set-up of this test site. Specifically, the typical layout of the Italian houses has created some problems due to reduced space (particular difficulty emerged in going through doors and due to some narrow passage in the house) to the connection to recharging station and to smoothly move in the house. This highlighted the need to improve

TABLE I

LONG TERM EVALUATION: VARIABLES MEASURED ALONG THE PHASES (S0–S4) AND RELATED MATERIAL

PHASES	S0	S1	S2	S3	S4
CLIENT					
Health Professional	Consent Form, Socio-Demographics Data Form, **Support Expectation**, **Diary**	Usability, **Diary**	**Support assessment**, Temple Presence Inventory, Networked Minds Social Presence Inventory, **Diary**	PANAS, PIADS, Final Interview, **Diary**	**Support Assessment**
Family member	Consent Form, Socio-Demographics Data Form, **Support Expectation** (informal carer), **Diary**	Usability, **Diary**	**Support assessment** (informal carer), Temple Presence Inventory, Networked Minds Social Presence Inventory, **Diary**	PANAS, PIADS, Final Interview, **Diary**,	**Support Assessment** (informal carer)
Relatives friends	Consent Form, Socio-Demographics Data Form, **Influence on Relationship** **Expectation**, Diary	Usability, **Diary**	**Influence on Relationship assessment** (informal carer), Temple Presence Inventory, Networked Minds Social Presence Inventory, **Diary**	PANAS, PIADS, Final Interview, **Diary**	**Influence on Relationship Assessment**
END USER					
Elderly	Consent Form, Socio-Demographics Data Form, **Loneliness (UCLA)**, **Short Form Health Survey (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Almere model**, **Health Service Satisfaction Inventory** (if applies)	**Loneliness (UCLA)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, Attitude_Acceptance, **Health Service Satisfaction Inventory** (if applies)	Temple Presence Inventory, **Almere model**	**Loneliness (UCLA)**, **Short Form Health Survey (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Almere model**, PANAS, PIADS, Final Interview	**Loneliness (UCLA)**, **Short Form Health Survey (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Health Service Satisfaction Inventory** (if applies)

the robot's mobility and to provide an automated recharging functionality. Currently the test site is at step S0 of the evaluation plan. Some robot usability problems are emerging due to the particular fragility of the two older adults who participate in the study. The couple is very interested to the GIRAFF robot, even though its use is currently still limited. Our goal is also to monitor the robot's usage over time to assess the effect of familiarity or habituation.

B. Test site 2

A very active woman living alone in Rome is the end user of our second Italian test sites. Her grandchild and daughter are the main current secondary users. Additionally we are also planning to involve a day care center that will connect to the woman. Also this test site is currently at step S0 of the evaluation plan. However, some preliminary comments can be reported. Both the lady and her grandchild are enthusiastic of the robot. They would also like that the robot do additional things. The lady, as most of the elderly people interviewed, is concerned about possible costs associated to the robots (e.g., the electricity consumption). Overall she really appreciates the possibility to stay in contact with her relatives, also relying on the video capability of the robot. She would also appreciate a sort of service provided by the day care center that would allow her to have a more frequent contact with a doctor or a specialist.

VII. Conclusions

This paper describes the ongoing work that is trying to assess an MRP within the elderly domain. Two important aspects are presented that can be considered as mandatory steps for both a general roadmap in robotics and our specific work.

As a first contribution, we have highlighted the importance of performing *ecological experiments*, i.e., which reproduce as much as possible the actual conditions of use of robotic technology, in terms for instance of real people who use it and real context of use. Although still simple in the results, analysis of the short-term evaluation provides a number of indications "from the field" that are representative of the actual users' expectations, both in relation to the human-robot interaction and to the most urgent technological improvements essential for an effective deployment. In addition to specific suggestions for improving the usability of the systems, we obtained other valuable recommendations that could be used for fielding the system into real world. For example, health workers expressed a number of requests that would be important to fruitfully use the GIRAFF system as a means to support their work. At the same time, the longitudinal tests done in real homes, are highlighting technological barriers that must be necessarily overcome.

The article's second contribution concerns our effort toward a long-term assessment. Other works in the area have highlighted this need but in this article we have proposed a rather elaborated and detailed methodology for the long-term evaluation that is currently being applied to real test sites of elderly people for long periods of time.

In the future we would like to enlarge the sample used in the short term evaluation possibly studying the differences among different groups of people. In addition we hope to continue gathering continuous data from the long term evaluation of the running test sites.

ACKNOWLEDGMENT

Authors are indebted to project partners for the stimulating work environment. Interactions with the colleagues from Örebro University have been fruitful to synthesize the evaluation plan. Authors would like to thank Vittoria Giuliani for comments to preliminary versions of this work.

References

[1] A. Tapus, M. J. Matarić, and B. B. Scassellati, "The Grand Challenges in Socially Assistive Robotics." *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 35 – 42, 2007.

[2] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards Robotic Assistants in Nursing Homes: Challenges and Results," *Robotics and Autonomous Systems*, vol. 42, no. 3–4, pp. 271–281, 2003.

[3] A. Cesta, G. Cortellessa, R. Rasconi, F. Pecora, M. Scopelliti, and L. Tiberio, "Monitoring elderly people with the ROBOCARE Domestic Environment: Interaction synthesis and user evaluation," *Computational Intelligence*, vol. 27, no. 1, pp. 60–82, 2011.

[4] A. Saffiotti, "The Concept of Peis-Ecology: Integrating Robots in Smart Environments," *Acta Futura*, vol. 3, pp. 35–42, 2009.

[5] M. K. Lee and L. Takayama, "Now, I Have a Body: Uses and Social Norms for Mobile Remote Presence in the Workplace," in *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, ser. CHI'11. New York, NY, USA: ACM, 2011, pp. 33–42.

[6] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik, "Exploring Use Cases for Telepresence Robots," in *Proceedings of the 6^{th} Int. Conf. on Human-Robot Interaction*, ser. HRI '11. New York, NY, USA: ACM, 2011, pp. 11–18. [Online]. Available: http://doi.acm.org/10.1145/1957656.1957664

[7] J. B. Beer and L. Takayama, "Mobile Remote Presence for Older Adults: Acceptance, Benefits, and Concerns," in *Proceedings of Human Robot Interaction: HRI 2011*, Lausanne, CH, 2011, pp. 19–26.

[8] A. Kristoffersson, S. Coradeschi, A. Loutfi, and K. Severinson Eklundh, "Towards Evaluation of Social Robotic Telepresence based on Measures of Social and Spatial Presence," in *Proceedings on HRI 2011 Workshop on Social Robotic Telepresence, Lausanne, March*, 2011, pp. 43–49.

[9] L. Tiberio, L. Padua, A. Pellegrino, I. Aprile, G. Cortellessa, and A. Cesta, "Assessing the Tolerance of a Telepresence Robot in Users with Mild Cognitive Impairment – A Protocol for Studying Users' Physiological Response," in *Proceedings on HRI 2011 Workshop on Social Robotic Telepresence, Lausanne, March*, 2011, pp. 23–28.

[10] S. Sabanovic, M. Michalowski, and R. Simmons, "Robots in the Wild: Observing Human-Robot Social Interaction Outside the Lab," in *Proceedings of the International Workshop on Advanced Motion Control*. Istanbul, Turkey: ACM, March 2006.

[11] E. Hutchins, *Cognition in the Wild*. MIT Press, 1995.

[12] T. W. Bickmore and R. W. Picard, "Establishing and Maintaining Long-Term Human-Computer Relationships," *ACM Transactions on Computer Human Interaction*, vol. 12, pp. 293–327, 2005.

[13] J. Nielsen, *Usability engineering*. San Diego, CA: Academic Press, 1993.

[14] J. Sauro and J. Lewis, *Quantifying the User Experience: Practical Statistics for User Research. Software Usability Scale (SUS)*. Chap. 8, pages 198-208. Morgan Kaufmann, 2012.

[15] M. Lombard, T. Ditton, and L. Weinstein, "Measuring Telepresence: The Temple Presence Inventory," in *Proceedings of the Twelfth International Workshop on Presence, Los Angeles, California (USA).*, San Francisco, 2009.

[16] Networked minds social presence inventory (scales only version 1.2). [Online]. Available: http://cogprints.org/6742/

[17] A. Terracciano, R. R. McCrae, and P. T. Costa, "Factorial and Construct Validity of the Italian Positive and Negative Affect Schedule (PANAS)," *European journal of psychological assessment official organ of the European Association of Psychological Assessment*, vol. 19, no. 2, pp. 131–141, 2003.

[18] H. D. J. Jutai, "Psychosocial impact of assistive devices scale (piads)," *Technology and Disability*, vol. 14, no. 3, pp. 107 – 111, 2002.

[19] D. Russell, L. A. Peplau, and C. E. Cutrona, "The Revised UCLA Loneliness Scale: Concurrent and Discriminant Validity Evidence," *Journal of Personality and Social Psychology*, vol. 39, pp. 472–480, 1980.

[20] J. E. J. Ware, M. Kosinski, and S. D. Keller, "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity," *Medical Care*, vol. 34, no. 3, 1996.

[21] G. D. Zimet, N. W. Dahlem, S. G. Zimet, and G. K. Farley, "The Multidimensional Scale of Perceived Social Support," *Journal of Personality Assessment*, vol. 52, no. 1, pp. 30–41, 1988.

[22] J. A. Yesavage, T. L. Brink, T. L. Rose, O. Lum, V. Huang, M. Adey, and V. O. Leirer, "Development and Validation of a Geriatric Depression Screening Scale: a Preliminary Report," *Journal of Psychiatric Research*, vol. 17, no. 1, pp. 37–49, 1983.

[23] M. Heerink, B. J. A. Kröse, V. Evers, and B. J. Wielinga, "Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model," *I. J. Social Robotics*, vol. 2, no. 4, pp. 361–375, 2010.

[24] G. Cortellessa, M. Scopelliti, L. Tiberio, G. Koch Svedberg, A. Loutfi, and F. Pecora, "A Cross-Cultural Evaluation of Domestic Assistive Robots," in *Proceedings of AAAI Fall Symposium on AI in Eldercare: New Solutions to Old Problems*, 2008.

Studying the Influence of Handheld Robotic Media on Social Communications

Takashi Minato[1], Hidenobu Sumioka[1], Shuichi Nishio[1], and Hiroshi Ishiguro[1]

***Abstract*— This paper describes research issues on social robotic telepresence using "Elfoid." It is a portable tele-operated humanoid that is designed to transfer individuals' presence to remote places at anytime, anywhere, to provide a new communication style in which individuals talk with persons in remote locations in such a way that they feel each other's presence. However, it is not known how people adapt to the new communication style and how social communications change by Elfoid. Investigating the influence of Elfoid on social communications is very interesting in the view of social robotic telepresence. This paper introduces Elfoid and shows the position of our studies in social robotic telepresence.**

I. INTRODUCTION

Recent studies have tries to realize immersive telepresence in telecommunication by application of tele-operated robots, which can transmit not only visual and vocal information but also physical information of speakers to remote places (ex. [1], [2]). In particular, the study with Geminoid [2], whose appearance and motion resemble those of an existing person, has revealed that a person operating Geminoid feels its body is his/her own body and another person facing the operated Geminoid also feels it is possessed by the operator, that is, the operator's presence can be transferred. Humanoid robots that resemble humans can be potent media for transferring human's presence. In order to make presence-transmitting medium pervasive in daily lives, a portable Geminoid is desirable: however, it is hard to implement a realistic human-like shape, numbers of actuators to provide human-like motions, and various sensors in a small body. A minimal design of the shape and motions to convey human's presence is necessary to develop a portable presence-transmitting medium.

The tele-operated android "Telenoid" has been developed based on the minimal design of humans [3]. It has a simplified human shape, holdable body, several actuators to express some human-like gestural motions. Based on the same concept, we have developed "Elfoid" (Fig. 1) as a portable tele-operated android [4]. It has also simplified human shape and is designed to transfer a speaker's voice using the cellphone networks and talk to other person in the fashion shown in Fig. 1.

Ogawa et al. [3] have reported that people prefer communication with Telenoid rather than cellphones and aged persons especially like to use Telenoid. From this fact it is

[1]Advanced Telecommunications Research Institute International, Hirosi Ishiguro Lab. and JST CREST, 2-2-2 Hikaridai, Keihanna Science City, Kyoto, Japan minato@atr.jp, sumioka@atr.jp, nishio@ieee.org, ishiguro@sys.es.osaka-u.ac.jp

Fig. 1. The prototype of cellphone-type tele-operated android "Elfoid"

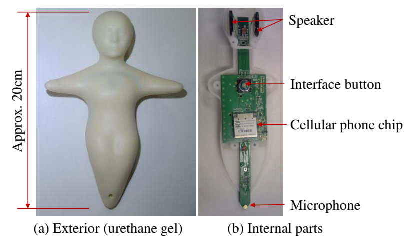


Fig. 2. The details of internal mechanism

expected that Elfoid is also accepted by people and provides a new communication style in which individuals talk with persons in remote locations in such a way that they feel each other's presence. But it is not known how people adapt to this new media. Studying people's adaptation to Elfoid in social communications will contribute to a design methodology of portable presence-transmitting media.

II. DEVELOPMENTAL CONCEPT OF ELFOID

The design concept of Elfoid is that anyone can transfer her presence to a remote place at anytime, anywhere.

a) Anytime, anywhere: A small 3G cellphone unit is embedded (Fig. 2), and people can talk with others in remote places in the same manner as cellphones. The prototype can register one telephone number in the memory, and users can call the number by pushing a button on the chest (a voice dialing system will be embedded in the future). Elfoid also has own telephone number and can receive a phone call.

b) Anyone: The shape is designed as easily recognizable at first glance to be nothing but a human, capable of being interpreted equally as male or female, old or young, that is, neutral human shape. The completely symmetrical face provides neutral gender. Mixing the child-like body proportion with the adult-like face proportion results in age-free appearance. These features aims at enabling it to be an avatar of anyone, differing from Geminoid which closely resembles the original person.

c) Transferring individuals' presence: Ogawa et al. [3] have reported that the holdability of Telenoid contributes conveyance of individuals' presence (that is, physical (tactile) interaction with the communication media is important) through the experiments with Telenoid. Elfoid has soft, pleasant-to-touch exterior (made of urethane gel) and transfers individuals' presence not only by voice and human-like appearance but human-like tactile impression.

The prototype shown in Fig. 2 does not have any sensors and actuators. It can be used as a normal cellular phone. The future works will improve the hardware of Elfoid. We will decide necessary hardware capabilities by studying which sensors and actuators are effective for transmitting human's presence. The appearance of the prototype will be verified whether it can be an avatar of anyone and be accepted by many people. Even the prototype can be easily used by people in their lives. This can enable us to conduct social experiments, not laboratory experiments.

III. STUDYING THE EFFECT OF CELLPHONE-TYPE TELE-OPERATED ANDROID

We expect that people will use Elfoid to have intimate conversations with distant people. In order to study people's adaptation to Elfoid, our research group already has conducted experiments to evaluate people's impression on Telenoid and people's adaptation to Telenoid in shopping malls [3], facilities for the elderly [5], and elementary schools [5].Telenoid provides different interactions compared with Elfoid in that it can move its extremities and users can hug its body. But the form is similar to Elfoid, therefore, it is expected that Telenoid has similar characteristics of presence-transmitting media to Elfoid. The experiments with Telenoid have revealed that people quickly adapt conversation with Telenoid and are impressed by its shape and tactile feeling, and same results are expected on Elfoid.

The important research issue on Elfoid is to study the effect of portable presence-transmitting media in social experiments. Kanda et al. [6], [7] placed a humanoid robot in a classroom of elementary school for two months and observed how the long-term relationship between children and the robot and among children changed. Our study also needs a long-term experiment in the same manner. For example, we distribute a number of Elfoid to a group of people and observe how they use Elfoid and how their relationships change. The focusing points in the social experiments are follows:

- In which situations and to whom do people use Elfoid to talk?
  The user does not always expect the presence of partner in the telecommunication. It is inferred that Elfoid can be easily used among people with close relationships such as lovers and friends but not with e.g., boss-subordinate relationship. Beer and Takayama [8] investigated how their tele-presence robot will be used by aged persons. Their result shows the aged persons want to use the robot to communicate with mainly their friends or families. We will investigate whether people use Elfoid between non-close relationships.
- How does Elfoid establish and change interpersonal relationships?
  Much studies have studied how cellular phones impact on social relationships of people. A survey [9] describes that the cellular phones positively and negatively influence the establishment of young people's relationship. From the results of the experiments with Telenoid, it is expected that Elfoid also increases the users' (especially aged persons') motivation to talk with other persons and facilitates to establish an interpersonal relationship. We will find out whether Elfoid has similar effect. We will also observe how interpersonal relationships among the users are altered owing to the communication style in which they feel each other's presence.
- Which kinds of application do they find?
  People recently find/create a new style to form social communities using the social media such as Facebook and Twitter. We will observe whether the users create a new style to form and change a group or community using the presence-transmitting media.

IV. SUMMARY AND CURRENT PROGRESS

This paper presented a novel communication medium "Elfoid" which provides new communication style in which a user can talk with another person while feeling each other's presence. The developmental concept and its prototype were shown and the research issues were discussed.

As a preparatory step toward social experiments, we are developing Bluetooth version of Elfoid. It works as a Bluetooth speaker-phone by connecting to a normal cellphone. The users can use their cellphones which they already have to use Elfoid. They do not need to change their cellphones, and then it is easy to distribute Elfoid to them.

REFERENCES

[1] S. Tachi, "Two ways of mutual telexistence: Telesar and twister," in *Telecommunication, Teleimmersion and Telexistence*, S. Tachi, Ed. IOS Press, 2003, pp. 3–24.

[2] S. Nishio, H. Ishiguro, and N. Hagita, "Geminoid: Teleoperated android of an existing person," in *Humanoid Robots, New Developments*, A. C. de Pina Filho, Ed. I-Tech Education and Publishing, 2007, pp. 343–352.

[3] K. Ogawa, S. Nishio, K. Koda, G. Balistreri, T. Watanabe, and H. Ishiguro, "Exploring the natural reaction of young and aged person with telenoid in a real world," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 15, no. 5, pp. 592–597, 2011.

[4] T. Minato, S. Nishio, K. Ogawa, and H. Ishiguro, "Development of cellphone-type tele-operated android," in *Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction*, 2012 (to appear).

[5] R. Yamazaki, S. Nishio, K. Ogawa, H. Ishiguro, K. Matsumura, K. Koda, and T. Fujinami, "How does telenoid affect the communication between children in classroom setting?" in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 2012.

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

[7] T. Kanda, S. Nabe, K. Hiraki, H. Ishiguro, and N. Hagita, "Human friendship estimation model for communication robots," *Autonomous Robots*, vol. 24, no. 2, pp. 135–145, 2007.

[8] J. M. Beer and L. Takayama, "Mobile remote presence systems for older adults: acceptance, benefits, and concerns," in *Proceedings of the 6th ACM/IEEE international conference on Human-robot interaction*, 2011, pp. 19–26.

[9] M. Campbell, "The impact of the mobile phone on young people's social life," in *Proceedings of the Social Change in the 21st Century Conference*, 2005.

Robotic Telepresence for 24/07 remote Assistance to Elderly at Home*

D. Lowet, M. Isken, W.P. Lee, F. van Heesch and E.H. Eertink

***Abstract*— Elderly want to live longer independently at home. In the Florence project we investigate to which extent low-cost state of the art robot technology supports elderly people to stay independent. The Florence robot supports robotic telepresence, monitoring and coaching services. In this article we describe the Florence robotic telepresence solution which is targeted at 24/07 remote assistance, not only providing support in emergency situations but also assisting with daily practical issues, social contacts and remote assessment of the level of independence of the elderly. We describe our technical solution and the results from our user tests.**

I. INTRODUCTION

Many elderly, when confronted with a gradual decline in their physical and mental capabilities, still have a strong desire to continue living in their own home despite increasing problems with social connectedness, safety and physical and/or cognitive problems. In the Florence project [1], we are developing a cost-effective robotic platform that supports elderly to stay longer independently in their homes by providing 24/7 remote assistance and coaching. The main objective is to research the acceptability of such a platform for elderly. . The Florence project puts the robot as a central connecting actor between the elderly, AAL services and a smart home environment.

Figure 1: Overview of Florence application scenarios.

*This work is partially funded by the European Commission as part of the FP7 Information and Communication Technologies programme under contract ICT-2009-248730 (Florence).

D. Lowet is with Philips Innovation Group Research, Eindhoven, 5656 AE Netherlands (dietwig.lowet@philips.com).

M. Isken is with OFFIS Institute of Information Technology, Escherweg 2, 26121 Oldenburg, Germany (melvin.isken@offis.de).

W.P. Lee is with Philips Innovation Group Research, Eindhoven, 5656 AE Netherlands (dietwig.lowet@philips.com).

F. van Heesch is with Philips Innovation Group Research, Eindhoven, 5656 AE Netherlands (frank.van.heesch@philips.com).

E.H. Eertink is a principal researcher at Novay, PO Box 589, Enschede, Netherlands (henk.eertink@novay.nl).

The project aims to develop a low-cost solution using currently available state of the art technologies; i.e. the focus is not on developing new hardware, but on combining available technologies into a complete system. The system is modular and extendible, providing a high level API to allow third parties to develop additional AAL services.

The project follows a user-centered, iterative development process. In the first iteration of the Florence platform, the robotic service have been implemented and tested at the home lab environments of Philips and OFFIS with 20 elderly and 10 family members, in the second iteration these services are tested at elderly homes.

Two applications areas are identified for the robotic AAL services: *coaching services*, where the robot autonomously gives advice based on monitored activities, vital signs and the agenda of the elderly, and *remote assistance* services where the robot acts as a telepresence device for services that implement remote assistance for e.g. emergency situations, fitness exercises, and daily activities.

For coaching services, the robot provides advice based on monitored, activities and measured vital signs. In addition, advice and cognitive support can be provided for taking medication, presenting agenda reminders, and giving motivational messages for a healthy lifestyle. The mobile autonomous robot has the benefit of pro-actively giving advice and information at the location of the action.

Robotic telepresence allows continuous remote assistance to the elderly by both family and care providers. Examples are emergency situations such as a fall situations, or providing support for everyday activities like exercising or cooking. The mobility that a robot provides yields a stronger feeling of presence at both sides and improves the quality of remote assistance. Specific aspects that we have addressed for the remote assistance services are safe navigation in a (unknown) home, and naturalness of the interaction.

The 'standalone' Florence service for robotic telepresence is referred to as the 'KEETOU' service (short for: Keep-in-touch). The same telepresence functionality, however, is also instrumental in a number of care and safety related scenarios. A few examples are:

- *Fall Handling*. The Florence system supports fall-detection and smoke detection. When e.g. a fall has been detected by the Florence smart home (external sensor), the robot contacts a telecare center. Their staff is then able to remotely control the robot in the elderly's home in order to review the situation.
- *Assisting with devices*. Many elderly have problems in using modern devices, like microwave ovens, set-top boxes or VCRs. The Florence robot can be called by its users for assistance in operating their devices. Based on the configuration of devices at home, the robot will be

able to select the proper videos, based on the location of the devices or set up a connection with

- *Health Monitoring.* Florence currently supports a WiThings® weighing scale, and a Bluetooth® blood pressure meter. When elderly feel insecure about the measurements, they can autonomously take initiative to discuss their concerns with the telecare center using the KEETOU service.
- *Medication Assistance*. Florence supports medication reminders, both at well-defined times or at relative times (e.g. just before or after having taken lunch). If the elderly are uncertain about which medication to take, the telecare center will support them via a video-call. The telecare center is allowed to read the medication-dispenser log.

This paper will focus on the telepresence functionality of the Florence system.

This paper is organized as follows. We first explain the technical overview of the Florence platform, since this has consequences for the Florence telepresence technical solution. Subsequently, we discuss the Florence telepresence implementation followed by the various user tests that have surrounded the Florence platform development and evaluation. Finally, the telepresence service is benchmarked and conclusions are drawn.

II. The Florence Platform

The Florence project has integrated a low-cost service-robot (~2000 Euro bill of material). The robot is based on currently available state of the art (robot and software) technologies. The Florence robot is a wheel-based, 1.5 meter tall, screen-based mobile device, without arms. The system is based on the Turtlebot platform and is illustrated in Figure 2. The robot senses its environment with a 2D laser scanner, a 3D depth sensor (Kinect) and a wide angle camera, while being connected to a smart home. The Florence project embraces a platform-based approach that supports 3rd party applications [2]. This platform is built on top of the Robotic Operating System (ROS) [3] – a de facto open source standard for robotic software. The Florence platform enables AAL (Ambient Assisted Living) service developers to use high-level APIs for e.g. multi-modal user interaction (including speech and gestures), user activity detection, planning activities and accessing information from the smart home, without requiring detailed knowledge on basic low-level robotic technologies. A number of these AAL services have been developed as a proof of concept.

Its mobility makes also the interaction more natural and social using multiple modalities like touch, speech and gestures. The KEETOU service is just one of the applications that the Florence platform supports and this platform approach has a number of consequences for our robotic telepresence solution:

- To enable an easy workflow for application developers, the Florence robot consists of two computing nodes. One Linux/ROS-based PC –referred to as the *Robot PC*– is part of the mobile base of the robot and runs all robot related software (e.g. autonomous navigation) and one Windows-based-PC with touch screen, referred to as *Touch PC*. The Touch PC, mounted on top of the robot, runs the high level applications and interaction software.
- To enable application developers to easily create new applications, the UI part of the applications can be developed with HTML5 and JavaScript. The communication between UI and the rest of the Florence platform is implemented using HTML5 websockets [3].
- The Florence platform makes use of a smart home environment, containing for example sensors such as PIR sensors, emergency buttons, smoke detectors, a doorbell, etc. and actuators, such as automatic windows, blinds, and doors.

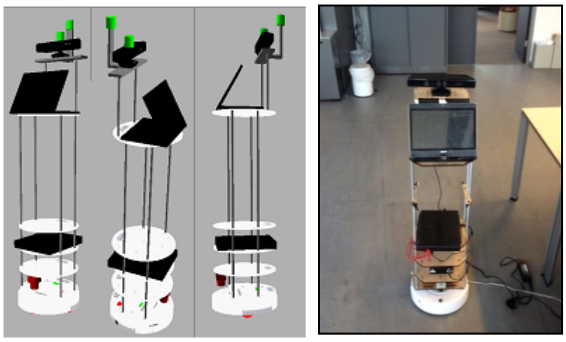


Figure 2. The Florence robot based on the Turtlebot platform.

The robotic services are able to use these home actuators and sensors when needed and available. These home sensors and actuators are used for detecting emergencies and to enable comfort services like easy smart home control. In case of the detection of an emergency situation, sensors can deliver information like the location of the person and the place the emergency has happened. This information can be used by care givers and emergency services to handle the situation in a fast and reliable way.

III. The Florence Telepresence Service

In this section, the telepresence service is described by first listing its requirements, followed by discussing the design and implementation, specific video, audio and teleoperation aspects, and touching upon the privacy and security aspects, respectively.

A. Requirements

The main requirements for the development of the Florence robotic telepresence solutions are:

- *Easy installment of the network access:* this includes the support for NAT and firewall-traversal.
- *Easy and simple to use*. This includes
  - *Automatic optimal video and audio quality:* we believe that the success of robotic telepresence solutions is currently held back by limited audio and video quality which places a high burden on the remote operator. This might be acceptable for dedicated applications but not for elderly care where

non-technical people need to use the system efficiently.
  - *Easy navigation:* this includes obstacle avoidance that prevents the remote user to accidently drive the robot against obstacles.
- *Minimal installation at the remote PC*. The telepresence solution is assumed to be also used by family of the elderly (typically their children). Given the wide variety of PCs and configurations, we aim for a simple setup of the remote PC, making use of existing and proven solutions.

B. *Design and implementation*

The audio/video communication of the KEETOU service implementation is based on Skype. The reasons for selecting Skype are (1) it automatically throttles audio and video quality, creating a fair communication quality even with consumer-grade webcams and typical home network connections and (2) Skype has the ability to deal with firewalls and NAT routers in both homes and companies (e.g. a care organization). This design-choice influences the implementation of our robotic telepresence service. The main drawback of our solution is that Skype is closed; and even though a Skype API [5] is available that allows control of Skype, it is not possible to integrate Skype inside an application (except with a licensing deal using the Skype developer kit). Therefore, it is not possible to completely embed the Skype video stream directly in the application UI (which is HTML based). This adds complexity and is expected to reduce user experience. However, we believe that, despite this drawback, good audio and video quality are more important for a telepresence solution than occasional UI glitches that might occur during the call setup. The Skype software runs on the touch PC, since the Skype output video is shown on this PC.

Skype also provides functionality to send text data over the Skype overlay network via the ap2ap part of the API. Since the Skype API is very low-level and cumbersome to use, we make use of the skype4py library [6] that provides a high-level wrapper for the Skype API and makes it easily accessible for Python applications. Bidirectional communication between the Florence applications (implemented in HTML5/JavaScript), and the python components is based on websockets. The skypeController implements a websocket server and listens to incoming requests from JavaScript code in the browser (e.g. for setting up a call).

At the remote side of the Skype connection a care provider or a family member runs Skype on a PC, referred to as the *Remote PC*, with the addition of an application (remoteControl.py) that enables the remote navigation of the Florence robot. This application also uses the skype4py library to make use of the Skype communication and the Pygame [7] library as a uniform way to handle joysticks and key presses. This application listens to commands from a joystick or keyboard and transmits them to the robot.

Error! Reference source not found.provides a schematic overview of the various KEETOU components and their interactions. Note that the remote control commands are sent from the *Remote PC* via the Skype API to the *Touch PC* and from the *Touch PC* to the *Robot PC*.

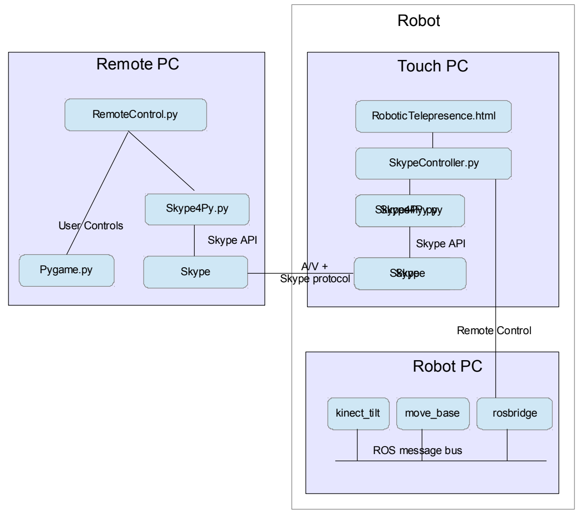


Figure 3. Overview of the Florence KEETOU software components.

0shows two screenshots of the UI of the KEETOU application on the touch PC of the robot. The left screenshot is shown to let the elderly select the person to talk to. The right screenshot depicts the UI that is shown when there is an incoming call.

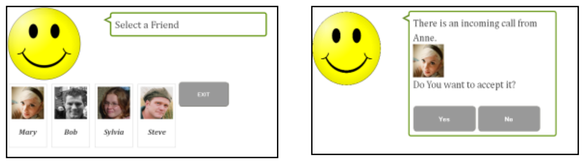


Figure 4. Two screenshots of the KEETOU service. On the left "selection of contact person" to connect to and on the rigth, the screenshot when there is an incoming call from a contact person.

Currently, an open source alternative to Skype, WebRTC [8], is under development by Google, Mozilla and Opera. WebRTC is similar to Skype except that it runs inside the browser. This is important for our solution for two reasons. First, on the robot side, the user interface of application services (fall handling, exercise coach etc) is implemented in HTML5. This allows for tight integration between the application service and the video stream by for example using overlays. Secondly, WebRTC allows anyone with a WebRTC-enabled browser (and the right credentials) to remotely control the robot without the need for installing extra software. We intend to switch the Florence robotic telepresence implementation to WebRTC when a mature version becomes available. The current versions do not yet support HD video and web cameras with hardware encoding.

C. *Video communication aspects*

For natural video communication, ideally the optical and acoustical sensation resembles those at the remote location as closely as possible. The main purpose is achieving media-rich communication between people. This requires careful consideration of both facial expressions and body language. It contributes to form, context or sketches the situation in which the conversation is held. In particular for robotic telepresence, video can provide a means of observation, e.g. when providing support about domestic appliances or during navigation of the robot while avoiding collisions with objects. The first requires a minimum quality of optical resolution while the latter requires a broad overview. As such, the Florence robot uses a 185 degree wide angle lens on a single HD webcam with an embedded encoder.

Its wide angle lens provides sufficient overview of the remote location. This is suits the human visual system very well as it tends to continuously scan the environment by eye and/or head movement. Providing a sharp image over a wide angle from close range, therefore, results in a relatively lifelike experience when scanning this picture. The single camera solution is beneficial to limit system resources, although its high resolution requirement implies a higher computational demand in case an embedded video encoder is absent. In addition, it needs to be stated that the data bus (typically USB) and power usage can be equally important system resources and need to be included in the camera's design choice.

A webcam with embedded encoder and a 1280 by 720 pixel resolution has been chosen to balance acceptable resolution with available system resources. The intelligibility during a conversation and ability to manually navigate the robot is acceptable at a frame-rate of approximately 20Hz. The main advantage of using a webcam is the widespread support for such a device. Besides the system resources, we have to deal with software clients and a complex communication infrastructure. Skype utilizes the embedded encoder and the HD capability of the webcam.

An important downside of using a wide angle lens is the serious distorted sense of distance and proportion it can cause. Seemingly far away objects are in reality much closer. This is especially noticeable when having a conversation with a standing person. One needs a certain level of experience when navigating using this optical configuration to build up confidence and comfort. This is most apparent when navigating the robot through narrow corridors and between door posts.

D. *Audio*

Conventional voice control and speech intelligibility solutions commonly fail when the microphone is not close to the mouth of the speaker, because in this case the microphone not only picks up the desired speech signal but also interfering signals such as background noise, sounds generated by the robot itself, and speech signals that are not intended for controlling the device or part of the conversation. In most cases, this will deteriorate the performance of speech to an unacceptable level.

For robotic telepresence, the user typically is at a distance from the robot and, thus, standard microphone solutions are inadequate. Although solutions for voice control at a distance have been proposed and implemented successfully in professional market segments [9], consumer solutions currently do not perform well enough to allow a broad vocabulary of speech commands. We have used the microphones of the Kinect sensor and limited the set of voice commands to get to an acceptable performance level and rely on the echo cancellation from Skype to improve speech intelligibility during telepresence.

E. *Teleoperation*

As mentioned earlier, we have extended the Skype video call with remote navigation using Sky4Py and Pygame. When navigating the Florence robot for looking around, we achieve panning and 'zooming' by rotating the robot or driving back and forth, respectively, instead of providing panning and zooming functionality to the camera. When navigating the robot for this purpose, however, obstacles are easily overlooked and collisions might occur. To alleviate this, collision avoidance has been added. Collision avoidance is especially important in case of an unreliable network connections or high communication latency. As such, the autonomous navigation capabilities of the Florence robot are used for collision avoidance during teleoperation. Further improvements in teleoperation can be achieved by means of waypoint navigation. This is considered future work.

F. *Privacy and security*

Privacy and security aspects, obviously, need to be considered when enabling a remote-controlled mobile video connection. Currently, KEETOU provides security protection by denying teleoperation access to the people outside the Skype contact list. Hence, it is not possible for the elderly to inadvertently accept friend requests. Furthermore, the elderly must always explicitly accept an incoming call. Further securing the (Skype) connection and outside connectivity, is considered future work.

IV. User Tests and Results

The Florence project is user centered and has adopted a multifold approach to get user feedback during the different development phases; starting with the initial requirements analysis up to the final system implementation. The following section gives an overview of this methodology, followed by the results of the user tests. This section gives a rough overview of the different sessions held and their respective results. More detailed information can be found in the Florence deliverables [1].

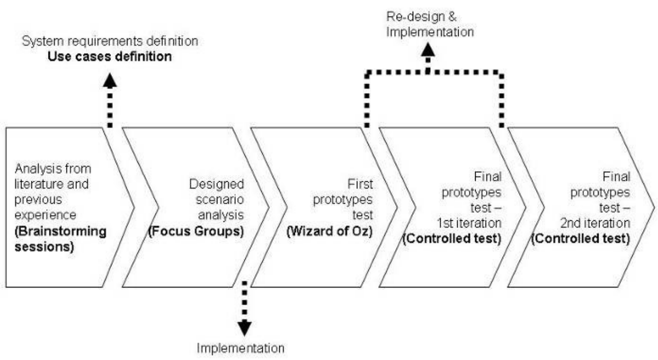


Figure 5. User involvement cycles of Florence

The strategy of the Florence project is to include the target group into the development process. 0shows the separate phases of their involvement in the development process.

Currently, the Florence project is approaching the last phase (Final prototype test). The project has started with brainstorming sessions, refined those results with focus group sessions [10], and implemented a first prototype which was then tested in Wizard of Oz tests [11]. Afterwards, the prototype was refined and tested in controlled real environments (living labs at Philips and OFFIS). The feedback from these tests has been taken into account for the final prototype which is tested in real home environments. In this section, we focus on the user experience results from the Wizard of Oz tests and the tests in the controlled environments. The gathered results from the brainstorming sessions and focus groups are summarized first.

G. *Test setup*

Brainstorming sessions. In order to define the lifestyle and AAL services for the elderly a brainstorming was done by the project team. The participants of this brainstorm session were chosen for their knowledge about elderly needs. From these sessions the following service areas were selected:

- *Coaching*, by giving feedback on specific activities like physical exercises, and advise on activities of daily living.
- *Social inclusion*, by supporting access to the social networks, including web-2.0 and synchronous communication means.
- *Safety*, by using Florence as additional ears and eyes in comfort or safety situations, controlled by service providers or the elderly themselves (crisis or emergency detection, smoke detection, personal alarm.

After the brainstorm sessions also some experts Gerontologist from Ingema helped in defining use cases and scenario definitions.

Focus Groups. Once these use cases were defined, some Focus Groups (FGs) were conducted for defining context and early design with stakeholders. Also, the FGs were used to validate the defined scenarios and use cases with the final users and to identify user requirements and needs. For these sessions, the homogeneity of the participants was considered as important, because people with similar features tend to be more comfortable in the group. The FGs in Spain were also repeated with care professionals (9 persons). All participants in the final FGs were +65 and had some degree of technology acceptance. In total, around 30 participants took part in the FG sessions.

The scenarios that were defined after the brainstorming and FG sessions are listed in the following table:

Acronym	Definition
KEETOU	Keep in touch offers elderly people a telepresence service.
HOMINT	Advanced home interface for remote control of the smart home. controlling lights, doors, windows, etc.
FALHAN	Fall situation *handling*. Once a fall is detected (remote sensor) a telepresence with care provider is automatically set up.
AGEREM	Agenda reminder service.
LIFIMP	Lifestyle improvement. This service helps the user doing e.g. exercises with instructional videos.
COLGAM	Collaborative gaming using videoconference.
LOGSYS	Logging system. For checking some kind of information the user should log in the system.

Wizard of Oz. The Florence robot used in the Wizard of Oz experiment was realized to a level that represents the final system. This includes basic interaction modalities (movement, size, appearance, audio input/output, video input/output, and touch input). The Florence system's autonomous functionality was mostly faked, however, (e.g. using a simulated user interface and remote control) without the test subject noticing this. The following list sums up the software needed for the implementation of the Wizard of Oz test:

- software to play Audio and VOIP software (Videoconferencing. e.g. Skype),
- software to control the movement of the robot and orientation of the camera,
- software to control and see what's displayed on the touch screen (e.g. VNC),
- picture viewer to display mock-ups of the user interface,
- software that allows the wizard to see where the user touched the display,
- the ability to record audio and video streams from the Florence robot.

The user-profile of the Wizard of Oz test participants consisted of people from 59 to 75 years old in an even proportion male-female (Germany 50% Male 50% Female, Spain 40% Male, 60% Female, Netherlands 50% Male 50% Female) with some degree of technology acceptance and some degree on computer use. Their physical state was relatively good, people with serious cognitive impairments or illnesses were excluded.

Additionally, caregivers were invited to participate in the tests to also include their feedback. In total 17 participants took part in the trials. Within the controlled environment, it was assured that all necessary technology to run the robot at full functionality was available. Trained personnel were on-site to help in case of technical problems. These tests were supervised by a regional ethical committee.

Controlled home lab environments. The controlled home environment tests evaluated the first implementation of the Florence platform. Being in a research stage, it was expected that bugs and unforeseen use cases would show up. Therefore, these tests were conducted under constant supervision of an engineer. Also, the controlled environment has the benefit that the robot can be more extensively tuned to its environment, compared to real homes. The results of these tests were used to identify limitations and improve the Florence platform in preparation of the user tests at elderly people's homes.

For the user tests, 24 persons aged between 60 and 85 participated mostly together with a close acquaintance (mostly a son or daughter) for testing the telepresence service with a family member. In addition, professionals of a care institute and nursing students evaluated the user test. In total, 40 persons participated in the tests. The elderly were selected based on the following inclusion criteria:

- living independently at home,

- living alone,
- being healthy, and
- if possible, experiences with accidents like falls (not necessarily personally)

H. General Feedback

Feedback has been collected from the target groups at different stages of development. Feedback from all tested people depended strongly on two main aspects:

- their experience and comfort with technology: surprise vs. comparison with related products,
- their experience with caring for elderly people (e.g. relatives suffering from MCI).

The feedback from the participants was very mixed from "no interest at all" to "very much interest". There was a strong correlation between the interest of the elderly in the robot services and the amount of care/support they currently received or had received in the past. The appreciation of the tested services heavily depends on the experience that people have with similar, relevant technologies. People that were less familiar with video-calls (e.g. using Skype), were more positive towards the entire user experience, as it surprised them more. Others often benchmarked Florence with existing implementations.

I. Feedback regarding telepresence service

The KEETOU service was liked by a large majority of the participants, both by the test subject as well as by the close acquaintance. It was seen as a means to have better contact. Most test subjects mentioned an increased feeling of presence of the remote party compared to their normal way of telecommunicating (mostly phone, some video chat via PC). The main benefits that were mentioned related to:

- showing the remote party around in case help is required, or as a social check.
- the more honest way of communicating ("the remote party can see how I am doing, instead of me just telling them"), and
- the ability to remotely evaluate the environment in case of an emergency.

The subjects would mainly consider the telepresence service for communicating with people they fully trust, like a close relative, care taker, or doctor. Little privacy concerns are seen in this context. During one user test the opposite was mentioned: telepresence with known people was seemed to have little benefit because "one knows what they look like".

Using the KEETOU service for communicating with unknown or less known people was not preferred and the privacy concern of "invading somebody's home" was raised in this context. Concerns about the KEETOU service were sometimes mentioned with respect to:

- being improperly dressed,
- not having the house tidy, and
- ease of use compared to operating a telephone.

J. Feedback regarding fall handling service

The fall handling service was often not distinguished clearly from the KEETOU service because the general functionality of telepresence is the same in both services. The only difference is the trigger of the service. A lot of people mentioned the FALHAN scenario even before it was presented as such. This might indicate that such a service is perceived as important safety functionality. The test subjects that had some experience with falls (themselves or friends that have fallen) had stronger positive opinions about this service.

Most test subjects indicated to find the use of speech to interact with this service as most natural. Having the option to call a family member instead of directly calling an emergency was appreciated as well.

K. Design and behavior

Some test subjects indicated that the Florence robot has authority due to its physical size and presence. With respect to the size and shape of the robot, further comments were made that the Florence robot should not be too colossal (obstacle) and it should appear practical and robust. Also, often it was mentioned that the robot should not look like a toy or puppet or resemble a human being too much. Furthermore, a preference was indicated for the device to fit with the house interior. The example of having the device also function as a lamp was mentioned in this context.

Concerning the behavior of the Florence robot, some of the test subjects commented that the robot should not be too patronizing, yet bring safety and assurance. Concerns were issued that people would be uncomfortable with the robot following them all the time. There were remarks with respect to the ease of use to tell the robot where to go, or not to disturb, and being able to operate it via its screen or by voice.

In general, speech control was preferred over gestures or its touch screen. Most test subjects indicated it should be easier to use the robot than a PC. The friendliness and clear voice of the robot were appreciated. And in the current version where speech input was not yet robust, it was assumed as such, i.e. people spoke back to the Florence robot.

L. Feedback from care professionals and informal care givers

Robotic telepresence was appreciated a lot by almost all of the close acquaintances that participated in the user tests, because it would enable them to often check whether the elderly person is doing well, and quickly assist them in case of problems or emergencies. Similar observations were made by the care providers. They also perceived to have an improved evaluation of the situation at the elderly's place. Professional care providers mentioned to see a huge benefit to better evaluate the validity of an automated alarm, prior to escalating it. Some close acquaintances indicated some skepticism about the ability of the elderly to be comfortable with such technologies.

V. Benchmark with Other Telepresence Robots

Currently, many telepresence robots in research projects and even commercial products are being developed. We can distinguish four categories for these telepresence robots based on their main use case. The first category for telepresence robots is intended for tele-work at remote places where it is too dangerous or cumbersome for humans. These robots are not used for communication. A well known example is the PackBot developed by iRobot [12]. The last few years a number of telepresence robots with communication functionality have been brought to market that are targeted at tele-work for offices and factory environments. Examples here are the Jazz robot of Gostai [13], the VGO robot by VGO communications [14] and the QB robot of Anybots [15]. These telepresence robots are more focused on presence and less on communication. Commonly, these telepresence robots are characterized by having a small screens on the robot. The small screen enables minimal communication without putting too many constraints on the stability of the navigation. Communcation is limited since facial expressions of the remote person cannot easily be made visible. A third category of telepresence robots are targeted at tele-healthcare in hospitals. An example is the the RP-7 from InTouch [16]. In this case, the robot is a US medical device subject to FDA approval. The RP-7 is costly at $100.000. A last category that we distinguish are telepresence robots for elderly care in home environments. In this category, we can distinguish two subtypes. The first is telepresence robots developed only for telepresence, like the giraffe robot [17]. These robots are typically very suited for robotic telepresence with a mobile base and a large screen on the robot. However, they do not provide a platform approach that makes it suitable for other robotic applications. Next to this, there is a considerable number of research projects developing screen-based mobile robots supporting multiple AAL applications of which robotic telepresence is one application service: e.g. Companionable [18], the Kompai developed by Robosoft [19], the Ava robot developed by iRobot [20]. The Florence robot distinguishes itself form these last class by its importance that is put in telepresence. Hardware and software has been developed with high quality telecommunication in mind, instead of considering it as one of the possible applications on an autonomous robot platform. For the Florence robot, this is reflected in not having a (physical) robot head, optimizing the positioning of the screen and wide angle web camera and using a separate compute node for handling robot navigation.

VI. Summary

In the Florence project, robotic telepresence is considered an important robotic application for supporting independent elderly at home. Our user tests indicate that robotic telepresence for elderly is highly appreciated by the elderly, their family and care providers. Not just for communication, as we originally expected, but for helping out with little problems/issues, in emergency situations and for social checks.

In addition, robotic telepresence can play a role in a number of services that many elderly need, like medication management and exercise coach. For these services, the mobility of a robot provides a significant advantage over other implementation of such services. Telepresence robots demand high quality audio (echo cancellation at large distances) and video (wide angle, very high resolution, high dynamic range, stability) when used for communication. In addition, the need for easy remote control and obstacle avoidance already requires a significant level of awareness of the environment and autonomy of the robot. We have addressed a number of these technical challenges using a wide angle camera with high resolution and hardware encoding and by adding automatic obstacle avoidance to the navigation. In the future, we plan to further increase the audio and video quality and develop improved guided remote control with the robot sensing where the remote user intends the robot to go.

VII. Acknowledgment

We would like to thank all partners from the Florence project (funded through the 7th Framework Programme by the EC, grant ICT-2009-248730).

References

[1] Florence project, http://www.florence-project.eu
[2] M.S. Bargh, M. Isken, D. Lowet, N. Snoeck, B. Hebgen, and H. Eertink, "Sensing, Actuation Triggering and Decision Making for Service Robots Deployed in Smart Homes", *in* 2nd International Joint Conference, AmI 2011, Nov. 16-18, 2011, Amsterdam, LNCS 7040
[3] ROS http://www.ros.org/wiki/
[4] Websockets, http://tools.ietf.org/html/rfc6455
[5] Skype API, http://developer.skype.com/public-api-reference-index
[6] Skype4py, http://skype4py.sourceforge.net/doc/html/
[7] Pygame, http://www.pygame.org/
[8] Webrtc, http://www.webrtc.org
[9] B.E. Sarroukh, L.C.A. van Stuivenberg, and C.P. Janse, "Robust Hands-Free Voice Control for Medical applications", in Proc. of the 11th International Workshop on Acoustic Echo and Noise Control, Seattle, Washington USA, 14–17 September 2008
[10] M. Heerink, B. Kröse, V. Evers, B. Wielinga, "The influence of a robot's social abilites on acceptance by elderly users", in Proceedings Ro-man, 2006.
[11] Maguire, M. C.; User-Centred Requirements Handbook. RESPECT Consortium, 1998
[12] Packbot http://www.irobot.com/us/robots/defense/packbot.aspx
[13] Jazz robot http://www.gostai.com/connect/
[14] Vgo robot http://www.vgocom.com/
[15] Qb robot https://www.anybots.com/#front
[16] RP-7 robot http://www.intouchhealth.com/products-and-services/products/rp-7i-robot/
[17] http://www.giraff.org/
[18] Companionable robot http://www.companionable.net/
[19] Kompai robot http://robosoftnews.wordpress.com/2010/03/08/robosoft-introduces-kompai-at-iltci-new-orleans-usa/
[20] Ava robot http://www.irobot.com/ava/

Towards Measurement of Interaction Quality in Social Robotic Telepresence

Annica Kristoffersson[1], Kerstin Severinson Eklundh[2] and Amy Loutfi[1]

Abstract— **This paper presents tools for measuring the quality of interaction in social mobile robotic telepresence. The methodology is in part based on Adam Kendon's theory of F-formations. The theory is based on observations of how bodies naturally orient themselves during interaction between people in real life settings. In addition, two presence questionnaires (Temple Presence Inventory and Networked Minds Social Presence Inventory), designed to measure the users' perceptions of others and the environment when experienced through a communication medium were used. The perceived presence and ease of use are correlated to the spatial formations between the robot and an actor. Use of the tools is validated experimentally on a dataset consisting of interactions between an elder (actor) and 21 different users being trained in piloting a mobile robotic telepresence unit. The evaluation has shown that these tools are suitable for evaluating mobile robotic telepresence and also that correlations between the tools used exist. Further, also from a local user perspective, the spatial formations have affected the perceived comfort in an interaction.**

I. Introduction

Mobile robotic telepresence (MRP) is a combination of teleoperation and telepresence offering a "walking around" capability. A MRP system is a video conferencing system through which the pilot of the system can move around in a remote environment and interact with its inhabitants (local user). The pilot interacts with the local user while at the same time navigating the robot via a computer. The local user experiences the interaction with the pilot via a video conferencing system mounted on a robotic base. Thus, usage of a MRP system includes many sorts of interaction simultaneously. Local users of the system are interacting with another human (HHI) but at the same time, they are interacting with a robot (human-robot interaction, HRI). The pilots of the MRP system are interacting with a computer system (HCI) but at the same time, they are interacting with another human (HHI) while being embodied in a robot which they cannot see themselves. Commercial examples of MRP systems are Giraff (Giraff Technologies [8]), QB (Anybot [23]), Texai (Willow Garage [24]) and VGo (Vgo Communications [25]).

In this paper, we discuss two experiments with a specific MRP system, the Giraff. In the first experiment, 21 alarm operators were trained in steering the Giraff by visiting an elder (actor). Regarding the first experiment, this paper presents an extended analysis of the results presented in [17] by studying the interaction between the pilot and the local user and describing the interaction using measures that relate to the spatial formations that occur between the local user and the pilot[1]. To characterize spatial formations, we take inspiration from Kendon's F-formations which provide a framework to describe the natural positioning/configuration of people when engaged in specific tasks. i.e. how people orient themselves with respect to each other. To characterize the levels of social and spatial presence, a questionnaire based on the Temple Presence Inventory (TPI) [20] and the Networked Minds Social Presence Inventory (Networked Minds) [5] was filled by the pilot user directly after the interaction took place. The TPI and the Networked Minds are two standard questionnaires to assess the perceived presence. The Networked Minds is designed to measure the users' perceptions of others experienced through a communication medium [4] and the TPI is appropriate for use with most media and media content [20]. Both the TPI and Networked Minds have been applied in previous studies in HRI, e.g. [2], [1], [11]. We also present a preliminary analysis of the second experiment in which 10 elderly interacted with a pilot in the Giraff. In this experiment, the pilot intentionally chose one out of two different spatial formations in different steps of a scripted scenario. After the experiment, the elderly were retrospectively interviewed to assess how they perceived the interaction with the person piloting the Giraff and the Giraff itself. The motivation behind both experiments is that we aim to find objective metrics upon which to improve the design of MRP systems. To do this, an understanding of the factors which impact interaction quality is important. Particularly related to social robotic telepresence is the effect of mobility achieved via embodiment (and vice versa), and therefor an aspect to study is how mobility affects the quality of social interaction.

The article is organized as follows, in Section II a description of Kendon's *F-Formation System* and related work is given. Section III details the experimental setups and data collection. Section IV presents our hypotheses. The results are presented and discussed in Section V. Finally, conclusions are drawn in Section VI.

The ExCITE project [7] has been supported by EU under the Ambient Assisted Living Joint Programme (AAL-2009-2-125).

[1]A. Kristoffersson and A. Loutfi are with AASS Center of Applied Autonomous Sensor Systems, rebro University, Sweden `firstname.lastname at oru.se`

[2]K. Severinson Eklundh is with the School of Computer Science and Communication, KTH Royal Institute of Technology, Sweden `kse at csc.kth.se`

[1]This paper is a modified version of an article that will appear in International Journal of Social Robotics. If you use any of the material regarding the first experiment, please cite the original paper [18].

II. Spatial relationships in interaction

Much of the work presented in this article is based on a theory on spatial handling with origins in HHI, Kendon's *F-Formation System*. In this section, we first describe this theory. We then discuss how spatial and embodiment constraints influence the quality of interaction and spatial formations [12], [15], [16], [21]. Communication over a medium is different from HHI as several factors important for the level of intimacy [3] are distorted [10]. Up until this point, there are only a few HRI studies which make use of the F-formations, e.g. [13], [19], [26].

Kendon studies spatial relations that occur whenever two or more people engage in an interaction and his claim is that *"a behaviour of any sort occurs in a three-dimensional world and any activity whatever requires space of some sort", p. 1* [16]. This space must have physical properties, allow the organism to do what it needs to do and be differentiated from other spaces. According to Kendon, it is common in any setting for several individuals to be co-present, but how they orient and space themselves in relation to one another directly reflects how they may be involved with one another. Kendon's F-formation system is a well known theory of spatial relationships. Kendon distinguished three F-formations by observations; (1) vis-a-vis, (2) L-shape and (3) side-by-side are depicted in Fig. 1. They arise when *"two or more people sustain a spatial and orientational relationship, p.209"* [15].

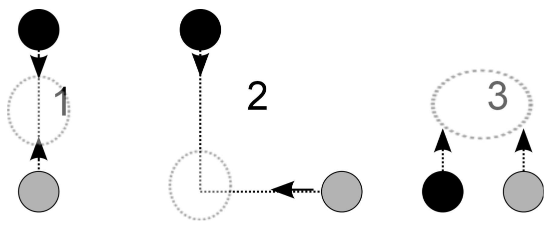


Fig. 1. The different F-formations distinguished by Kendon. 1. Vis-a-vis, 2. L-shape and 3. Side-by-side. The dotted ovals denote the O-space.

The *vis-a-vis* is an arrangement formed by two individuals who are facing each other, the *L-shape* is generated when two individuals are standing perpendicularly to each other facing an object and the *side-by-side* arrangement is when two participants stand close together looking in the same direction. The shared space created by the F-formation is called the *O-space* and is the overlap of transactional segment to which two or more people direct their attention and manipulate objects [15], [21]. The type of F-formation created can be influenced by environmental features such as obstacles and walls. The influence of these features on spatial patterns have yet to be studied in depth [15], [16]. The F-formation spatial arrangement is reconfigured as participants in a group come and go, and change their positions and orientations [19]. However, the orientation of the lower part of the body has a dominant effect on the reconfiguration of the arrangement [15].

According to Marshall et al. [21], physical environments can limit and constrain opportunities for some shared activities, while encouraging others. They concluded that some features of the physical environment may work to discourage the creation of certain F-formations. Hornecker [12] has developed the concept of *embodied constraints*, these restrict what people can do based on the configurations of space and objects and make some behaviors more likely than others. Along these lines, physical embodiment via a MRP system can also constitute an embodied constraint in that the physical properties of the system restrict particular behaviors, e.g. the inability to move backwards in the environment.

The level of intimacy in HHI is an equilibrium *"joint function of eye-contact, physical proximity, intimacy of topic, smiling, etc, p. 293"* [3]. When using MRP systems, this would potentially mean that the experienced intimacy for the local user and pilot engaged in an interaction would be dependent on the above factors. A factor that has been pointed out as being of importance in a MRP system is the ability to adjust the height of the robot to increase the level of eye-contact [6], [14]. It should be noted that an interaction enabled by use of computer-mediated technologies, e.g. a video-conference, can not be equivalent to face-to-face situations according to Heath and Luff [10]. They write that gestures and other forms of body movement including gaze, which are systematically employed in face-to-face communication prove ineffective in a large part in video communication.

In the HRI domain, Kuzuoka et al. [19] examined how reconfiguration of F-formation arrangements occurred as an effect of a guiding robot following a predefined trajectory at a museum when rotating either its head, its upper body or the whole body to guide their listeners when talking about exhibits. The results are corresponding well with Kendon's finding that the orientation of the lower body has a dominant effect on the automatic creations of F-formations. Yamaoka et al. [26] developed a model for how an information-presenting robot should appropriately adjust its position and found that robots that were constrained to an O-space were perceived as more comfortable than robots being close either to the listener or the object the robot was presenting. Hüttenrauch et al. [13] studied spatial formations in real home settings by deploying a service robot and instructing the users to guide the robot around in their own home and to teach the robot rooms, locations and objects. Each trial was video-taped for later analysis and revealed that the participants would lead the robot when passing narrow passages. The formation that occurs in the narrow passage indicates that one can also distinguish another spatial formation in HRI which Hüttenrauch calls *follow* in which the robot follows the user.

In the outlined material on spatial formations in HRI, we are missing studies on how the quality of interaction with others is perceived by both pilots and local users. We are interested in how the perceived quality varies depending on what spatial formations that are created during an interaction.

III. METHOD

In **experiment 1**, 21 alarm operators were invited to a training session during which they made a remote visit to an elder's home via Giraff. The alarm operators respond to alarms coming from elderly who by pushing a button on a necklace get in direct contact with the service. The training session was a preparation for deploying a Giraff in a home with two elderly, one of which was using a wheelchair. This visit was their first use of Giraff and was presented as an opportunity to train on steering and using the Giraff. The training session also served the purpose of collecting data via questionnaires and observations. The training session took place in a smart-home environment. The locality for the training sessions was chosen in order to simulate a close to real experience of use of the robot in a real home setting. This particular setup simulated an elder residence in which one elder(actor) was sitting in a wheelchair. A graphical overview of the remotely visited home can be found in Fig. 2. There are constraints in the physical space. The kitchen is too small to accommodate both a wheelchair and a Giraff at the same time. The distance between the table and television set makes it difficult to fit the Giraff beside the wheelchair. The spatial constraints may discourage creation of F-formations and encourage other spatial formations. The alarm operators who came individually were placed in a room where a laptop equipped with a headset and a mouse was installed. Each training session began with informing the participants about the computer and its connected devices and then instructing them to make a visit to a remote home with the Giraff. They were instructed to interact with the elder as if it was a visit to a real home. Further, they were informed that they would be asked to fill in a questionnaire after completing the training session.

Fig. 2. A model of the visited smart-home. The numbers denote the width of door openings and other spaces, for example 1.51 m in the door opening to the bedroom and the width and depth of the kitchen is 1.8 m. The cylinder 'a' represents the Giraff and the office chair 'b' represents the wheelchair.

The use of an actor was essential to script the visit and to ensure that the interaction was as similar as possible between the visits from different alarm operators. However, two different actors were used during the training sessions.

The procedure outlined below was used for each visit in experiment 1. Here, we use *pilot* to denote the alarm operator, *elder* to denote the actor in the smart home and *researcher* to denote the researcher who sat beside the pilot and provided technical support in case of difficulties (e.g. the pilot cannot find the docking station or does not know which buttons to press on the Giraff Pilot application). Numbers in parentheses, e.g. S1 denote situations in the scripted visit.

1) The researcher instructed the pilot to start the Giraff application, log on to the Giraff server and to connect to the Giraff that was facing the wall.
2) Once connected, the pilot was instructed to undock the Giraff from the docking station by pushing the buttons **Backward** and **Turn**. The pilot was asked to locate the elder. The pilot would find the elder in the bed. No guidance in where to find the elder was given.
3) When the pilot had found the elder, the elder moved over to the wheelchair and asked the pilot to follow to the kitchen (S1).
4) While in the kitchen, the elder started a discussion about a medical issue (S2).
5) The elder then asked for help to find the remote control for the television set. The pilot and the elder moved to the living room (S3) to find it. The pilot would tell the elderly that it was lying on the floor in between the sofa table and the television set (S4).
6) After the pilot had found the remote control, an artificially triggered alarm rang in the bedroom. Depending on the pilot's response, the elder found an appropriate means to conclude the conversation and asked the pilot to return to the docking station.
7) The pilot returned to the docking station and disconnected from the Giraff with help from the researcher if necessary.

For most pilots, the script took in between 200 s to 300 s measuring from when the pilots had undocked and turned the Giraff until the moment when elderly said bye.

In **experiment 2**, 10 elderly were invited to a guiding tour in a showcase environment. The elderly arrived individually and were met with a researcher who presented the Giraff. The elder was asked to sit down in the sofa in the living room and was offered coffee. A second experimenter then connected to the Giraff and began to interact with the elder. In this experiment, a scripted scenario was followed. However, the pilot chose one of two spatial formations in each step.

1) The pilot undocked, navigated the Giraff and stopped in front of the elder and said welcome. The pilot chose either a vis-a-vis or a look-away spatial formation, see Fig. 4.[2]
2) The pilot asked the elder to follow to the kitchen where a number of items on and around a table were shown and explained to the elder. These items are part of the showcase appartment and included a chair for assisting elder and a robot arm. In here the pilot either faced the items on the table (L-shape) or the wall. (look-away)
3) The pilot asked the elder to follow to the bedroom that contained sensors spread around the room. The

[2]The look-away formation is further described in Section IV.

pilot either faced the elder (vis-a-vis) or the bed when describing the content of the room.

4) The pilot drove back to the living room and asked the elder to sit down. Now, a set of paintings, a Roomba and a medicine dosett were shown. Again, the pilot chose either a vis-a-vis or a look-away formation.
5) The pilot said good-bye and returned to the docking station. The first researcher returned.

A. Data Collection

Throughout **experiment 1**, ten permanently installed web cameras positioned at different locations in the ceiling of the smart-home recorded each of the participant's training sessions. The cameras were configured to capture most parts of the apartment from different angles. A snapshot from one of the videos is found in Fig. 3. The figure shows the Giraff and the elder in the wheelchair from several angles allowing analysis of different parameters, e.g. F-formations. Video recordings enable repeated and detailed access to the conduct and interaction of participants, and, more specifically, the interplay of talk, bodily, and material conduct [22]. Recording of video data is an ethical issue. It is argued that when being filmed, people inevitably react to the camera - rendering the data is unreliable [9]. In this experiment however, the participants were unaware of the fact that the Giraff was being recorded on video so they did not react to being filmed. The video recordings were taken without sound and they do not reveal who is currently embodied in the MRP system. However, the choice of camera configuration came with a sacrifice. The video recordings were not able to capture facial expressions from the pilot or sound from the the interaction between the pilot and the elder. The sacrifice limits the possibility to fully understand the interaction and its effects on chosen spatial formations.

Fig. 3. A snap shot from one of the video recordings showing the different angles of video capture.

Upon completing the training, the participants were asked to fill in the questionnaire. Two of the sections in the questionnaire assessed the perceived social and spatial presence. Each dimension of perceived presence in the questionnaire consisted of several questions that were to be answered on a likert scale 1-7 where 1 = not at all and 7 = to a very high degree except for the questions in the TPI dimension *Social richness*. Social richness was assessed by asking the pilots to rate their experience in opposite couples (for example whether they perceived the experience to be Insensitive or Sensitive) on a scale 1-7. The dimensions *Object realism* and *Person realism* originate from the TPI dimension *Perceptual realism*. The original dimension contained more questions about modalities not available in a MRP system. Thus, a subset of questions from the perceptual realism dimension was used, see Appendix A.

In **Experiment 2**, three cameras were used to video capture the experiment. The movies recorded were primarily used in order to perform a voice-recorded retrospective interview with the elderly after having completed the scripted scenario. A retrospective interview technique was used immediately after the elderly were guided. The elderly watched a movie for each step of the guiding tour and were asked to comment to it by responding to a set of questions regarding a number of parameters e.g. the spatial formations.

IV. HYPOTHESES

The participants in **experiment 1** were faced with a multiple of novelties, such as using a video conferencing system to interact with somebody, steering a robot, meeting a new person and a new environment. These factors, and the fact that video conferencing systems cause a distortion in perception made us expect that some pilots would not turn the Giraff in a vis-a-vis formation while interacting with the elder. Two of the situations included in the script enforced a movement of the Giraff together with the elder (S1 and S3) and it was expected that these would lead to situations in which the pilot user would either follow or go ahead of the elder. As already discussed in Section III, the size of the kitchen would potentially limit the possible spatial formations because it could not accommodate a wheelchair and a Giraff at the same time. In combination with the fact that the elder would discuss a medical issue (S2) with the user of the Giraff we expected that the user would form a *vis-a-vis* formation at an appropriate distance to the elder. Another space constraining factor in the apartment was the distance between the table and the television set. We expected that it would be difficult to fit the Giraff and the wheelchair beside each other at this location and therefor expected that a *L-shape* would occur in S4. Thus six patterns of formations were foreseen to occur during the training sessions as shown in Fig. 4. Three of which being F-formations as defined by Kendon and three of which being assumed based on the above presented assumptions. To conclude, the spatial formations were expected to vary throughout the scenario based on the different situations and changes in available space in the scripted scenario.

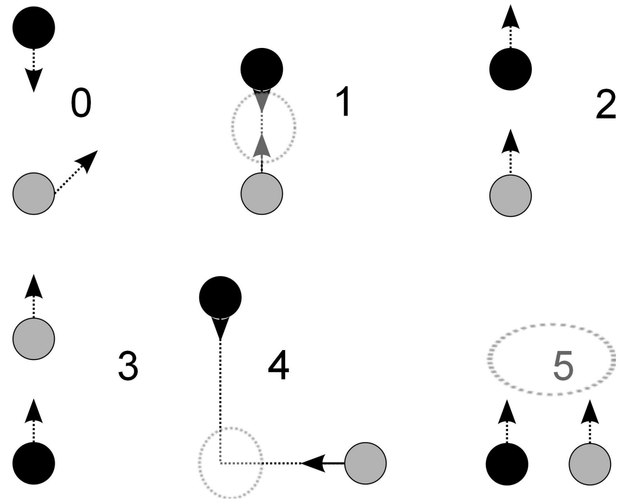


Fig. 4. The formations expected in the experiment, 1, 4 and 5 are F-formations: 0. Look-away, 1. Vis-a-vis, 2. Ahead, 3. Follow, 4. L-shape and 5. Side-by-side. The dotted ovals denote the O-space. The elder is black and the Giraff is grey.

In our previous analysis [17], we analyzed the perceived precence and ease of use during the training session. It was found that the presence questionnaire was suitable for use in this HRI setting. According to Kendon, as discussed in Section II, it is common to be co-present when interacting with others. That is, how people orient and space themselves in relation to one another is a reflection of how they may be involved with one another. Thus it was hypothesized that:

> **[H1]** Relations between chosen formations and perceived presence would exist.

We further expected that the pilots perceiving the MRP system as being easier to use would orient themselves in suitable formations to a higher degree than the ones not orienting themselves in suitable formations..

> **[H2]** Relations between chosen formations and experienced ease of use would exist.

In **experiment 2**, we expected that the elderly would feel more comfortable interacting via the Giraff if the pilot chose the same spatial formations as would naturally occur in real life situations.

V. Results

A. Subjects

The users invited to **experiment 1** were alarm operators. The average age of the 21 users was $\mu_{age} = 42.19$, $\sigma_{age} = 10.34$. Only two of the alarm operators were men, therefore no comparison between genders is done in this study. None of the users had previous experience of Skype or similar systems for communicating with or without video feed. On a likert scale 1-7 where 1 = not at all and 7 = to a very high degree, the experience of using such technologies was $\mu = 1.90$ and $\sigma = 1.67$. Thus there was a dual novelty for the participants in that they lacked experience of both videoconferencing technology and MRP systems.

In **experiment 2**, 3 men and 7 women in the age range 61-82 years old participated. All claimed they had a habit of using computers and three stated they were using video conferencing software.

B. Experiment 1 - Choice of formations based on situations and space

It was expected that the different situations S1-S4 in combination with space constraints, would result in the spatial formations as depicted in Fig. 4. To investigate this, the videos were analyzed in several steps. First, each movie was watched and notes were made on when the different steps in the script occurred and how the interaction between the elder and the pilot took place. A number of fields of information had to be filled for each video. Secondly, all of the 21 video recordings were re-watched and notes were made on how the formations fluctuated during the interactions. These notes were then converted to illustrative graphs showing the fluctuations between different formations.

To exemplify, two graphs showing the occurrences of formations between two different pilots and the elder are presented in Fig. 5 (Pilot 1-1) and Fig. 6 (Pilot 1-10). The x-axis shows the time having elapsed from when the pilot moved the Giraff forward after undocking in Step 2. The y-axis shows the different formations as defined in Fig. 4.

In Fig. 5, Pilot 1-1 found the elder after 55 s. The pilot formed a *vis-a-vis* formation while interacting with the elder in the bedroom. Thereafter, the pilot *followed* the elder to the kitchen (S1) after 78 s. Upon arrival to the kitchen after 100 s, the pilot chose a *vis-a-vis* formation while interacting with the elder in the kitchen (S2). This formation was kept throughout S2 with a short break during which the pilot turned the Giraff around 360 degrees. The pilot *followed* the elder from the kitchen towards the living room (S3) after 183 s. At 195 s, the pilot formed a quick *vis-a-vis* formation after which it continued to lead towards the remote control. The remote control was found (S4) in a *side-by-side* formation after 207 s. This formation lasted for 13 s after which a *vis-a-vis* formation was upheld for 12 s. The elder then went towards the bedroom followed by the pilot, a new *vis-a-vis* occurred at 240 s. This formation lasted for 6 s after which the elder said bye. In this example, the pilot did as expected in S1 and S2.

In Fig. 6, Pilot 1-10 formed a *vis-a-vis* formation while interacting with the elder in the bedroom. Thereafter, the pilot started *following* the elder to the kitchen after 43 s (S1) and arrived there after 62 s. The pilot chose a *vis-a-vis* (S2) formation during the interaction in the kitchen. After 97 s, the pilot chose to go *ahead* to the living room (S3) in order to find the remote control. The remote control was found after 116 s during a *L-shape* formation (S4). This formation lasted for 9 s after which a *Vis-a-vis* was upheld until 172 s when the elder said bye. In this example, the pilot did as expected in S1-S4.

To be able to assess whether the pilots acted according to our expectations and whether the choice of formation correlated with perceived presence and ease of use, the

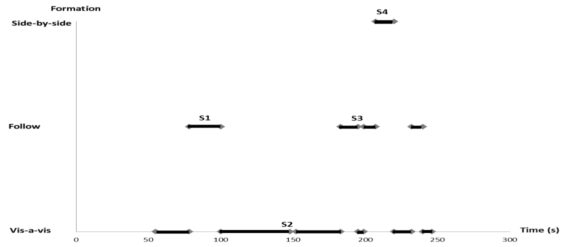


Fig. 5. An illustration of how the formations fluctuated during pilot 1-1's training session. Details on the formations are found in Fig. 4.

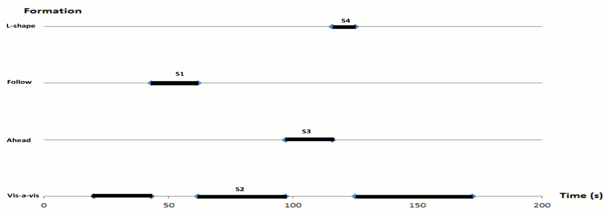


Fig. 6. An illustration of how the formations fluctuated during pilot 1-10's training session. Details on the formations are found in Fig. 4.

actual occurring formations for each participant during the situations S1-S4 were monitored.

In total, 18 out of 21 (86%) pilots chose to *follow* the elder from the bedroom to the kitchen in S1.

Also, 14 out of 21 (67%) pilots chose to form a *vis-a-vis* formation while communicating in the kitchen in S2. The other seven pilots formed a *look-away* formation by looking at the fridge or the wall to the right of the fridge. Two possible reasons are that the Giraff is equipped with a wide-angle lens camera allowing the pilots to see a large field of view, another is the fact that turning the robot around is an effort and requires navigational skills.

It was observed that, 13 out of 21 (62%) pilots chose to go *ahead* of the elder on the way to the living room in S3. When leaving the kitchen, the pilot would have to move out of the way of the elder in order to follow. The fact that 8 pilots did not go ahead of the elder implies that the robot may have been more difficult to navigate than what was expressed in the questionnaire. Another reason may be that the pilots were focusing on the interaction with the elder rather than on navigating the robot.

It was expected that an L-shape would occur when finding the remote control to the television set in the living room (S4). 13 out of 21 (62%) pilots formed a *L-shape* but the other 8 formed a *side-by-side* shape in the space restricted area. It is worth noting that those who formed a *side-by-side* shape had general difficulties navigating the Giraff, as observed by the researcher. The assumption of an L-shape configuration in this situation seems reasonable and corresponds with the fact that comfortable drivers of the Giraff conform to this configuration.

C. Experiment 1 - Relations between formations and perceived presence

One-way Anova tests were done to investigate whether there existed relationships in between the chosen formations in the situations S1-S4 and and the different dimensions of presence outlined in [17], [18]. We found differences in the perceived *Attentional engagement* depending on spatial formation in S1 and *Comprehension* depending on spatial information in S4.

All pilots experienced a relatively high *Attentional engagement* with the elder. However, those who drove *ahead* of the elder experienced a significantly higher level of *Attentional engagement* with respect to those who *followed* the elder in (S1). Specifically F(1,19)=6.36, $\prec 0.05$ ($\bar{x}_{ahead} = 6.11$, $\bar{SD}_{ahead} = 0.23$ and $\bar{x}_{follow} = 5.78$, $\bar{SD}_{follow} = 0.21$).

The pilots who chose the hypothesized *L-shape* in S4 experienced a higher *Comprehension* than the pilots who chose the *side-by-side* formation, Specifically F(1,19)=4.26, $\prec 0.05$ and $\bar{x}_{L-shape} = 5.15$, $\bar{SD}_{L-shape} = 0.96$ among the 13 pilots choosing the hypothesized formation and $\bar{x}_{side-by-side} = 4.33$, $\bar{SD}_{side-by-side} = 0.75$ among the eight pilots who chose the side-by-side formation. See Appendix. A for more details on the dimentsions Attentional Engagement and Comprehension.

Considering that the period of time during which the pilot described the position of the remote control was short (e.g. 13 s for pilot 1-1 in a side-by-side formation and 9 s for pilot 1-10 in a L-shape formation), it is interesting that there exists a significant difference in perceived Comprehension. The numbers indicate that the pilots who formed the hypothesized L-shape were also able to better understand the elder's intentions, thoughts etc. As pointed out in Section V-B, five of the pilot's who chose the side-by-side formation in S4 were observed to steer the robot with more difficulty. It may be the case that they were not able to focus as much on the elder and thus not understand the elder's intentions, thoughts etc. to the same degree.

The majority's choice of formation when describing the position of the remote control and its relation to *Comprehension* is well in line with Kendon. He claims that "In conversations between just two persons, when the topic is disembodied, the arrangement tends to be *L-shaped*...Typically, when two people greet one another and then continue to talk together on some topic, they can be observed to begin with a face-to-face arrangement and then to shift to an L-arrangement as they move from salutation to talk" pp. 8-9. [16]

To summarize, as hypothesized in H1 there are correlations in between what spatial formations the pilots used and how spatially and socially present they felt in the environment and with the elder.

D. Experiment 1 - Relations between formations and perceived ease of use

The pilots were asked to respond to a number of questions such as "How was it to connect to Giraff?" on a likert scale 1-7 where 1 = very difficult and 7 = very easy. The users have responded 5 or higher in average on all questions asked regarding the perceived ease of use. One-way Anova analysis reveals that the perceived ease of use varies depending on the occurring spatial formations in S1 and S2 but not in S3 and S4.

The pilots who did as hypothesized in S1 answered significantly higher on how easy it was to start the Giraff application, F(1,19)=4.58, $\prec 0.05$. The mean value for the pilots choosing to *follow* was $\bar{x}_{follow} = 6.67$, $\bar{SD}_{follow} = 0.59$ while the value for the pilots choosing to go *ahead* was $\bar{x}_{ahead} = 5.67$, $\bar{SD}_{ahead} = 1.53$. It could be the case that pilots who found the interface easier to use were more focused on the task at hand (follow elder to the kitchen).

There is a significant statistical correlation between the pilots' choice of positioning in S2 and how easy the pilots thought it was to leave the docking station (F(1,19)=8.14, $\prec 0.01$) and to make a u-turn (F(1,19)=9.16, $\prec 0.01$). The pilots who chose *vis-a-vis* formation during the interaction with the elder perceived it as easier to both leave the docking station ($\bar{x}_{vis-a-vis} = 6.79$, $\bar{SD}_{vis-a-vis} = 0.43$ compared to $\bar{x}_{look-away} = 5.57$, $\bar{SD}_{look-away} = 1.51$) and to make a u-turn ($\bar{x}_{vis-a-vis} = 6.21$, $\bar{SD}_{vis-a-vis} = 0.58$ compared to $\bar{x}_{look-away} = 4.71$, $\bar{SD}_{look-away} = 1.70$). As previously discussed in Section V-B, one of the possible reasons for not positioning the Giraff in a vis-a-vis formation with the elder in S2 was that an extra effort was needed. The pilots having chosen the vis-a-vis formation also responds that they perceived it as both being easier to leave the docking station and to make a u-turn with the Giraff. For them, steering the Giraff in to a vis-a-vis formation could have been considered as being less of an effort than for the ones having more trouble navigating the Giraff.

To summarize, as hypothesized in H2 there are correlations in between what spatial formations the pilots used and how easy to use they perceived the Giraff system to be.

E. Experiment 2 - Spatial formations from an elderly perspective

While only an initial analysis of video and voice-recording data has been performed it is clear that the importance of eye contact while interacting with a pilot in the Giraff is emphasized by most elderly. This is revealed not only in the retrospective interviews but also in the video data with some of the elderly experiencing the unnatural spatial formations while interacting with the pilot user.

> **[Ex 1]** "[...] it should be turned towards me. The contact is needed."
> **[Ex 2]** "The eye contact was there, I think that part is important."
> **[Ex 3]** "I almost had to move myself so that I could see her.[...] I should see the one I talk to."

A few elderly were concerned about not seeing the pilot during the movement between rooms during the home tour.

> **[Ex 4]** "It felt a bit strange when she had turned towards the table in the kitchen. I was instructed to go there but it felt weird to see her from the back so to say."
> **[Ex 5]** "She turns the back on me when leaving the bedroom. I do not know if she could possibly back out and keep interacting with me on the way out. Technically it should not really be a problem huh so that we can keep the eye contact."

Some elderly chose to move themselves in order to reconfigure the spatial formation to face the pilot (vis-a-vis). Apparent in the bedroom was also that some elderly changed the spatial formations depending on what item that was described in the room. To conclude, spatial formations is of relevance for the local users while interacting with a person piloting the Giraff. Worth noting is also that all elderly would have preferred that the pilot adjusted the height of the Giraff while interacting with the elderly when sitting in the sofa.

VI. Conclusions

In this paper, we have investigated tools for measuring the quality of interaction in social mobile robotic telepresence. Using Kendon's theory of F-formations, that is how bodies naturally orient themselves, the interaction between an elder and mobile robotic telepresence system was monitored during a scripted scenario. In addition a questionnaire that assessed the perceived presence and ease of use was filled by the pilot users. Correlations between the dimensions measured in the questionnaire and the chosen spatial formations emerged. In a second experiment, it was found that spatial formations between the pilot and an elder were also of importance for a comfortable interaction from the perspective of an elder. This work has shown that tools such as F-formations, the Temple Presence Inventory and the Networked Minds Social Presence Inventory are useful for evaluating the quality of interaction in mobile robic telepresence systems. The experiments showed that these tools are suitable for evaluating mobile robotic telepresence and also that the correlations found in experiment 1 can give important guidelines on how to better operate the robotic unit in order to increase the quality of interaction. The intitial analysis of experiment 2 is that using the F-formations are important from the elderly local users point of view. For example, improving the interface in order to allow easier rotation of the robot in order to change spatial formation could lead to a higher perceived comprehension of the thoughts and intentions of the other and a higher quality in the interaction for the pilot as well as the local users. Further experimentation is necessary to better understand the correlations between spatial formations and interaction quality and future work will focus on collection of more experimental data.

Appendix A Supporting definitions on Presence

The *Object Realism* contained two questions:

1) The objects you saw looked like they would have done in reality.
2) The objects you saw sounded like they would have done in reality.

The *Person Realism* contained two questions:

1) The person you met looked like it would have done in reality.
2) The person you met sounded like it would have done in reality.

The level of *co-presence* "is influenced by the degree to which the user and the agent appear to share an environment together, p. 5" [4]. The *co-presence* as used in this study consists of only four questions:

1) I felt that x and I were in the same place.
2) I believe that x felt as if we were in the same place.
3) I was aware of that x was there.
4) x was aware that I was there.

The *attentional engagement* "seek to measure the degree to which the users report attention to the other and the degree to which they perceive the others level of attention towards them, p. 10" [4]. The *Attentional engagement* as used in this study only contains two questions:

1) I payed attention to x.
2) x payed attention to me.

Comprehension is the degree to which the user and the other understand their respective intentions, thoughts etc.

References

[1] S. O. Adalgeirsson, "Mebot: a robotic platform for socially embodied presence," Master's thesis, Massachussetts Institute of Technology, 2009.

[2] S. O. Adalgeirsson and C. Breazeal, "Mebot: a robotic platform for socially embodied presence," in *Proceedings of HRI 2010*, 2010, pp. 15–22.

[3] M. Argyle and J. Dean, "Eye-contact, distance and affiliation," *Sociometry*, vol. 28, no. 3, 1965.

[4] F. Biocca and C. Harms, "Guide to the networked minds social presence inventory," http://cogprints.org/6743/, accessed 6 November 2011.

[5] ——, "Networked minds social presence inventory (scales only version 1.2)," http://cogprints.org/6742/, accessed 6 November 2011.

[6] B. Cohen, J. Lanir, R. Stone, and P. Gurevich, "Requirements and design considerations for a fully immersive robotic telepresence system," in *Proccedings of HRI 2011 Workshop on Social Robotic Telepresence*, 2011, pp. 16–22.

[7] S. Coradeschi, A. Loutfi, A. Kristoffersson, S. von Rump, A. Cesta, G. Cortellessa, and J. Gonzalez, "Towards a methodology for longitudinal evaluation of social robotic telepresence for elderly," in *Proc of the HRI 2011 Workshop Social Robotic Telepresence*, 2011.

[8] Giraff, http://www.giraff.org, accessed 6 Nov 2011.

[9] M. Gottdiener, "Field research and video tape," *Sociological Inquiry*, vol. 49, no. 4, pp. 59–66, 1979.

[10] C. Heath and P. Luff, "Disembodied conduct: Communication through video in a multi-media environment," in *Prodeedings of CHI 1991*, 1991, pp. 99–103.

[11] L. Hoffman and N. C. Krämer, "How should an artificial entity be embodied? comparing the effects of a physically present robot and its virtual representation," in *Proceedings of HRI 2011 Workshop on Social Robotic Telepresence*, 2011.

[12] E. Hornecker, "A design theme for tangible interaction: Embodied facilitation," in *Proceedings of ECSCW 2005*, 2005, pp. 23–43.

[13] H. Hüttenrauch, E. Topp, and K. Severinson-Eklundh, "The art of gate-crashing bringing hri into users' homes," *Interact Stud*, vol. 10, no. 3, pp. 275–298, 2009.

[14] N. Jouppi and S. Thomas, "Telepresence systems with automatic preservation of user head height, local rotation, and remote translation," in *Proceedings of ICRA 2005*, 2005, pp. 62–68.

[15] A. Kendon, *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University Press, 1990.

[16] ——, "Spacing and orientation in co-present interaction," *Lecture Notes in Computer Studies*, vol. 5967, pp. 1–15, 2010.

[17] A. Kristoffersson, S. Coradeschi, A. Loutfi, and K. S. Eklundh, "Towards evaluation of social robotic telepresence based on measures of social and spatial presence," in *Proc of the HRI 2011 Workshop Social Robotic Telepresence*, 2011.

[18] A. Kristoffersson, K. S. Eklundh, and A. Loutfi, "Measuring the quality of interaction in mobile robotic telepresence a pilot's perspective," *Int J of Social Robotics*, 2012, to appear, DOI: 10.1007/s12369-012-0166-7.

[19] H. Kuzuoka, Y. Suzuki, J. Yamashita, and K. Yamazaki, "Reconfiguring spatial formation arrangement by robot body orientation," in *Proceedings of HRI 2010*, 2010, pp. 285–292.

[20] M. Lombard and T. Ditton, "A literature-based presence measurement instrument the temple presence inventory (tpi) (beta)," M.I.N.D. Labs, Temple University, Pennsylvania, USA, Tech. Rep., 2004, accessed 6 November 2011. [Online]. Available: http://astro.temple.edu/~lombard/research/P2scales_1104.doc

[21] P. Marshall, Y. Rogers, and N. Pantidi, "Using f-formations to analyse spatial patterns of interaction in physical environments," in *Proceedings of CSCW 2011*, 2011, pp. 445–454.

[22] R. Meisner, D. von Lehn, C. Heath, A. Burch, B. Gammon, and M. Reisman, "Exibiting performance: Co-participation in science centres and museums," *Int J of Sci Educ*, vol. 29, pp. 1531–1555, 2007.

[23] QB, http://anybots.com, accessed 6 November 2011.

[24] Texai, http://www.willowgarage.com/pages/texai/overview.

[25] VGo, http://www.vgocom.com/, accessed 6 November 2011.

[26] F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita, "A model of proximity control for information-presenting robots," *IEEE Trans on Robot*, vol. 26, no. 1, pp. 187–194, 2010.

Teleoperated android for mediated commmunication: body ownership, personality distortion, and minimal human design

Hidenobu SUMIOKA[1], Shuichi NISHIO[1], and Hiroshi ISHIGURO[1,2]

Abstract— **In this paper we discuss the impact of humanlike appearance on telecommunication and overview our studies with teleoperated androids. Due to humanlike appearance, we show that teleoperated androids affect not only the local people interacting with them but also the teleoperators who control them at other locations. Such androids enhance teleoperator feelings of telepresence by inducing a sense of ownership over particular body parts. We also point out that an appearance mismatch between an android and a teleoperator distorts the latter's personality that is conveyed to the local people. To overcome this problem, we introduce the concept of minimal human likeness design and demonstrate that a new teleoperated android developed with a minimal human likeness design reduces telecommunication distortion. Finally, we discuss some research issues about the sense of ownership over the telerobot's body, the minimal human likeness design, and the interface design.**

I. Introduction

How can we perceive and be perceived as being present at a remote location? Telepresence, which provides such perceptions, is one of the greatest challenges faced by information and communication technology. While such existing personal telecommunication devices as telephones or video chat only transfer user vocal and visual information to a remote location, as telecommunication media, a telerobot conveys other nonverbal information by remotely controlling physical proxies in the location [1], [2], [3], [4]. Such proxies are expected to enhance the quality of life for all people by facilitating human-human daily interaction that requires nonverbal information, including intimate communication, elderly care, and support for sick children [5], [6], [7].

Many telerobots have been invented as physical communication media for teleoperators (*visitors*) in remote sites to interact with partners (*locals*) in local sites. Since they are designed to mediate human-human interaction, they share fundamental communication skills with humans, such as pointing, gesturing and making facial expressions. In contrast, their appearances often take different forms, including humanoids, mobile robots, cute animals or imaginary creatures [1], [2], [3], [4], [8], [9], [10] even though they have to be considered the physical proxies of the visitors by the locals.

This work was supported by JST, CREST and Grant-in-Aid for Scientic Research (S), KAKENHI (20220002).

[1]H. Sumioka, S. Nishio, and H. Ishiguro are with Hiroshi Ishiguro Laboratory, ATR 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan sumioka@atr.jp, nishio@ieee.org

[2]H. Ishiguro is with Graduate School of Eng. Science, Osaka Univ., Machikaneyamacho 1–3, Toyonaka-shi, Osaka, 560–0043 Japan ishiguro@is.sys.es.osaka-u.ac.jp

Obviously, our perception of others strongly depends on their appearance. Research on personality has reported that people judge other's personality based on physical appearance [11]. This suggests that telerobot's appearance affects visitor's personality in human-human interaction mediated by the robot. Studies on human-robot interaction have also reported that humanlike appearance and the natural behavior of autonomous robots cause people to respond to the robots as if they were human [12]. People's expectations of a robot's abilities and its performance are biased by its appearance [13], [14]. These previous works suggest that a telerobot's appearance must be designed carefully for smooth communication. However, insufficient attention has been paid to telerobot appearance for mediated human-human interaction.

Since the creation of Geminoid HI-1, the first teleoperated android (Fig. 2(a)) that closely resembles a living individual, the influence of human likeness on mediated human-human interaction has been investigated [15], [16], [17]. While the geminoids were designed to provide massive nonverbal information including the specific characteristics of its model for locals, the telenoid has been developed to explore minimal human likeness so that it can convey anybody's presence [18].

These studies show the influence of the telerobot's appearance, summarized in Fig. 1, on the visitor's feeling of telepresence and the local's feeling of the visitor's presence. They provide us with valuable insights not only about the enhancement of these feelings but also about their reduction. To understand how to design a telerobot system, we must explore the influence of the telerobot's appearance. However, to the best of our knowledge, no such exploration has been done in telecommunication mediated by telerobots.

In this paper, we describe how a telerobot's appearance positively or negatively affects visitors and locals in mediated human-human interaction and overview studies with teleoperated androids. We first briefly explain geminoid, which is a teleoperated android that resembles a living individual, and its teleoperation system. Then we show that humanlike appearance allows visitors to strongly feel telepresence by inducing a sense of ownership over the geminoid's body. We also show that such an appearance creates in the locals a strong feeling of visitor's presence, although it cannot fully duplicate the visitor due to technological limitation. Next we indicate the negative effects related to the individual characteristics inherent in the geminoid that might prevent visitors from conveying their own personalities to locals. To avoid this problem, a telenoid is introduced whose design

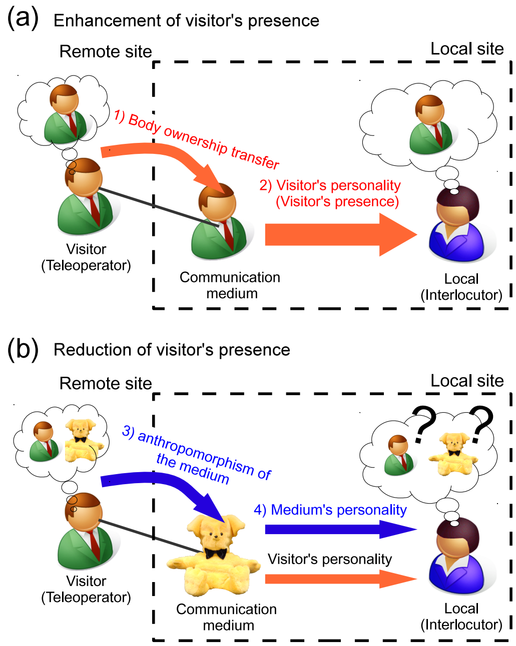


Fig. 1. Overview of appearance's influence on telecommunication: a) enhancement of visitor's presence. When a telerobot's appearance resembles the visitor, 1) the visitor feels a strong telepresence because of body ownership over the telerobot and 2) the local has strong feeling of the visitor's presence. b) Reduction of visitor's presence. When a telerobot's appearance is different from the visitor, the telerobot is anthropomorphized by 3) visitors and 4) locals.

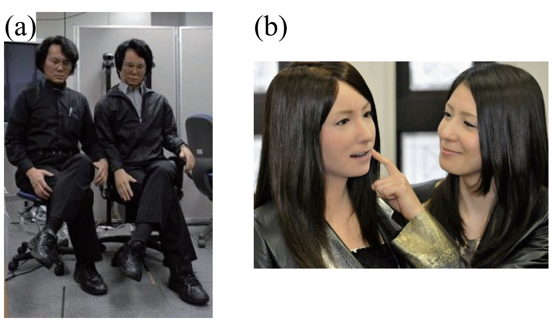
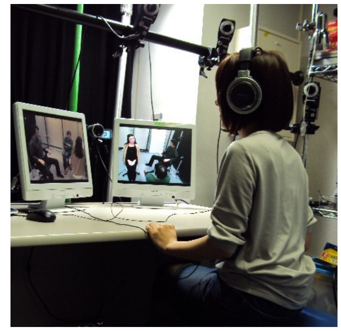


Fig. 2. Geminoids: (a) Geminoid HI-1 (right) and model (left); (b) Geminoid F (left) and model (right)

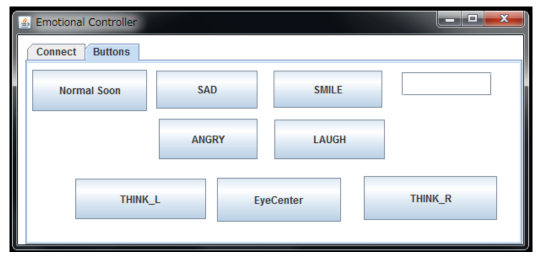
is based on the concept of minimal human likeness. We demonstrate that minimal human likeness design reduces the distortion of the visitor's personality. Finally, we discuss research issues about the sense of ownership over telerobot's body, minimal human likeness design, and interface design.

II. Geminoid and its Teleoperation System

The geminoid, which functions as a duplicate of a living person, has been developed to investigate individual presence (Fig. 2). Geminoid HI-1, which resembles a living male, has 50 degrees of freedom (DoFs) including 13 for facial expressions (Fig. 2 (a)). Geminoid F, which resembles a living female, has 12 DoFs (Fig. 2 (b)), most of which are used for facial expressions.



(a) Teleoperation interface



(b) Example of interface for conscious behavior controller

Fig. 3. Teleoperation system setup and example of conscious behavior controller interface. Setup and interface can be changed based on available devices and required behavior.

Visitors control the geminoid with the teleoperation interface from a remote room (Fig. 3 (a)). Both geminoids have two different controllers: a conscious behavior controller and an unconscious one [15], [16]. The conscious behavior controller has such behaviors predefined by visitors as offering to shake hands and smiling. The visitor can select and run each behavior using a GUI (Fig. 3 (b)). While the conscious behavior controller drives the geminoid's behavior to reflect the visitor's intention, such subtle expressed motions as breathing, blinking, and trembling are added by the unconscious behavior controller to maintain natural behavior. In addition to such semi-autonomy, the system has subsystems to synchronize the geminoid's face directions and lip motions with those of its visitor. There are several possible subsystems to achieve synchronization. A motion capture system is an option to generate precise geminoid movements, although carrying the system is difficult. We can also use a more portable system, such as a face recognition system, which computes the visitor's face direction through camera images, and a speech-driven lip motion generation system, which generates lip motions from the visitor's vocal information of the visitor [19]. Actual motor commands are sent to a control server for the geminoid by TCP/IP.

III. Visitor's Telepresence Enhanced by a Sense of Ownership over Teleoperated Android

Several visitors of geminoid reported that when they looked at another person touching a part of its body, they felt as if they were being touched [16]. A similar phenomenon, a sense of ownership over an artificial body or body part, has been reported in virtual reality and neuroscience. A famous example is the rubber hand illusion [20], where a participant's hand is removed from sight while an artificial rubber hand is placed in view. When an experimenter synchronously strokes both hands, most participants report that in less than a minute they actually feel the touch on the rubber hand. This phenomenon is strongly induced if the same body part as the presented artificial body part is stroked [21]. This indicates that a sense of ownership is evoked over an artificial object if it resembles a human body or a body part.

While in the rubber hand illusion a sense of ownership was induced when visual information matches the tactile one, Nishio *et al.* hypothesized that such a sense was induced when visitors observe the geminoid motion that they control [22], [23]. In their experiment, participants looked at the geminoid's right arm moving through a head-mounted display from a first-person point of view in synchronization with the movements of their real arms. They also experienced two other conditions where the geminoid's arm movements were moved based on those of the participant with a delay and were not moved at all, even though the participant's arm did move. They were told to move their right arms from side to side every three seconds for one minute[1]. After each experience, the participants watched an experimenter gave an injection into the geminoid's arm. Participant reactions were measured as a physiological responses (skin conductance response; SCR) and subjective ratings for a sense of ownership. The results showed that in the synchronous condition, they showed a significant change in SCR and a significantly strong sense of ownership over the geminoid's arm. Although the participants observed the geminoid from a first-person point of view in this study, a follow-up study reported that the sense of ownership occurred even when they observed the geminoid from a third-person point of view [24].

Surprisingly, this phenomenon occurs even when the geminoid is controlled not with the actual body movement of a visitor but with the visitor's brain activity [25]. Alimardani *et al.* explored whether the control of a human-like robot's hand through a noninvasive brain computer interface (BCI) enhanced a sense of ownership over that hand in the absence of visitor's hand motions. Participants wearing an electrode cap made a gripping motion with either the geminoid's left or right hand based on the experimenter's directions by imagining that movement after some practice. After the session, the geminoid's arm was injected and the participant's SCRs and subjective ratings about a sense of ownership were collected. When the geminoid's hands only moved if the motor command from BCI was correct and identical as the cue (match condition), SCR was significantly different from those when the geminoid's hands did not move at all (still condition). In addition, the participants felt a stronger sense of ownership over the geminoid's hands in the match condition than in the still condition.

These results suggest that a teleoperated robot with a humanlike appearance allows visitors to enhance their feelings of being at a local site by evoking a sense of ownership over the robot. This sense seems to occur not only when a visitor observes the robot from a first-person point of view but also from a third-person perspective. Furthermore, a visitor can feel this sense even when the robot is controlled with the visitor's brain activity without the visitor's actual motions. What is interesting is that a sense of ownership can be evoked by the telerobot's humanlike appearance, suggesting the important of the design of the telerobot's appearance.

[1]Due to system delay in controlling the geminoid, the presented movement of the geminoid was created by cutoff animation where the geminoid's arm moves in synchronization with their arm movement.

IV. Locals who feel the presence of visitors

The geminoids also allow locals to feel the presence of visitors. Sakamoto *et al.* compared the presence of a person mediated by different media (geminoid, video chat, and voice chat) in group conversations [15]. In the condition with a geminoid, its head movements (looking at a member or nodding) were remotely controlled by a visitor. In the video chat condition, the visitor's face and voice were presented to other members and only the visitor's voice was presented in the voice chat. The locals were instructed to have three-minute discussions on given topics in all the conditions. After each conversation finished, the locals rated their impressions of the visitor conversations through a medium on a 7-point Likert scale where 1 indicates a negative attitude and 7 stands for a positive one. The results showed that the locals had significantly stronger feeling of the visitor's presence in the geminoid condition than in the other conditions. No significant difference was reported among the geminoid condition and the video chat in the human likeness and naturalness of the presented visitor's behavior.

A feeling of visitor presence seems to be influenced by the human likeness of the telerobot's appearance. Nishio and Ishiguro [26] examined how the appearance of telerobots affected people's attitudes toward them. In their experiment, participants were asked to have a discussion with physical entities with different appearances (the geminoid, the model person, a humanoid robot, and a static object). Through the discussions, the model person tried to persuade the participants by himself or through another physical entity. The impression of the physical entities was rated on 7-point Likert scales about the presence and personality of the visitor and the naturalness of the behavior of the physical entities. Factor analysis results showed that the impression of the visitor's presence and personality changes based on the likeness to humans; the machine vibration of the artificial agents makes them look unnatural. This implies that the human likeness of the telerobot's appearance is an important component to convey a visitor presence to a local site.

Although the geminoids closely resemble living individuals, they cannot fully convey their presence to a local site, as reported by Nishio *et al.* [17], who investigated how children adapted to the geminoid and how much it successfully conveyed its model's presence. In their two-day experiment, two children, the daughter of the model person and a boy who did not know the model person, interacted with a physical entity (the geminoid or the model person). Each child had a conversational task, such as talking about family photographs or video movies, twice with both the geminoid and its model. Two weeks later, both children had the same task with the geminoid and its model again to observe the process of adaptation to the geminoid. The interaction was evaluated by the amount of conversational utterances of the child or the entity for each task, the amount of eye contact, and the amount of the children's body movements. The results suggest that even though both children were scared by the geminoid when they saw it for the first time, they gradually adapted to it. The responses of both children toward the geminoid approached their responses toward the real person. Interviews after the experiments showed that while the model's daughter understood that the model was controlling the geminoid, the boy who was not familiar with the model failed to recognize the geminoid as the model person. He thought that the geminoid was a person wearing a strange mask. Perhaps, it is difficult for locals who do not know a visitor to feel the visitor's presence from the geminoid due to its unnatural movements and appearance.

These studies show that teleoperated androids can provide a strong feeling of visitor presence for locals even though the geminoid cannot fully duplicate its model. The feeling of presence seems to be associated with the human likeness of the telerobot's appearance. However, in locals, the realistic appearance of telerobots might produce the uncanny valley phenomenon [27], which is discussed in the virtual reality community [28], [29] as well as robotics [30], [31], [32]. We will discuss this issue in Section VII-C.

V. Personality distortion derived from telerobot appearance

We demonstrated that a teleoperated android with a very humanlike appearance can enhance both the visitor's feelings of telepresence and the local's feelings of the visitor's presence, even though the geminoid cannot fully duplicate its model. However, does the enhancement occur even when the android's appearance does not resemble its visitor? In this section, we address two kinds of distortions that affect the conveyance of visitor personality. One is caused by the visitors themselves and the other is caused by the locals interacting with telerobots.

A. Distortion by visitors

An ethnographic case study conducted by Straub *et al.* reported that the visitors themselves distorted their personalities due to the geminoid's appearance [33]. They exposed the geminoid to an open public space so that ordinary people could freely experience interaction mediated through the teleoperated android, both on the local site for facing the android and on the remote site for controlling its head motions and voice. The mediated interactions were analyzed for verbal cues about the personality presentation on the side of the visitor and for verbal cues about the personality perception of the geminoid from the local site. Their results revealed that the visitors talked to the locals not only by presenting their own personality but also by presenting a new personality that they created, mixing the robotic and their own personality, or imitating the robot's model person.

Similar distortion to what was reported by Straub *et al.* [33], called the "proteus effect" has been reported in research on virtual avatars [34] and showed that user attitudes and behaviors change depending on the virtual avatar's appearance. The results by Straub *et al.* imply that the proteus effect is also seen with teleoperated androids. This might negatively affect telecommunication because visitors do not convey their own personality to locals.

B. Distortion by locals

In addition to such personality distortion by visitors, the mismatch between visitor's personality and the personality imagined from the telerobot's appearance disturbs the perception of locals. The case study by Straub *et al.* [33] reported that locals anthropomorphized and mentalized the geminoid even when they suspected a visitor in the background. This implies that the telerobots' characteristics inherent in their appearance strongly distort the impression of visitors.

C. Speaker identification under distortions by locals and visitors

As shown above, in mediated human-human interaction locals might be affected by the distortions derived from a visitor and the telerobot's appearance. Is it possible to identify visitors even when personality distortion by a visitor and the anthropomorphization of a telerobot by an local occur?

The results of experiments conducted by Sumioka *et al.* [35] offer insight. They investigated whether locals can identify speakers only from conversational contents, using the "Doppel teleoperation system," which isolates several physical traits from a speaker (Fig. 4). With this system, for each of the communication channels to be transferred, one can choose it in its original form or in the one generated by the system. For example, the speaker's appearance, voice, and body motions can be replaced with the geminoid and its operator while preserving the speech content. In the experiments, participants identified actual speakers among four possible candidates: their acquaintances, the experimenter's assistant, geminoid's operator, and geminoid's model. If they selected the operator or the model, that means that the locals' decision was affected by geminoid's behavior or its appearance since the actual speaker was only selected among their acquaintances and the experimenter's assistant. The results suggest that the locals had difficulty identifying the actual speakers because the impression of the speakers was distorted by geminoid's behavior and its appearance.

This result indicates that it is difficult for locals to identify visitors if their personalities are distorted by the appearance of a telerobot and the visitors themselves. The distortion derived from these sources prevents visitors from conveying their personalities, which might be fundamental information to feel their presence in a remote place.

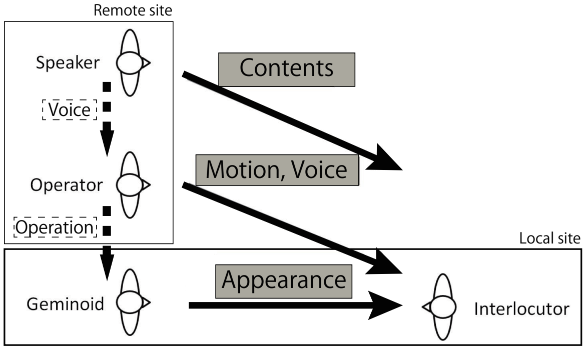
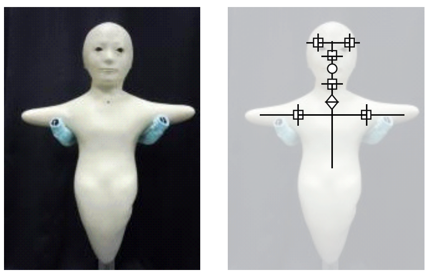


Fig. 4. Communication channels and their sources in Doppel system proposed in [35].



(a) Front view

(b) Layout of joints in telenoid except for one joint for opening the mouth

Fig. 5. Telenoid and its joints

VI. TELENOID: MINIMAL HUMAN LIKENESS DESIGN

As described in Section V, teleoperated androids have both positive and negative effects to achieve visitor feelings of telepresence and local feelings of visitor's presence. One approach to overcome the negative effect is to project an image of the visitor onto displays embedded in the telerobot [1], [2], [3], [4]. However, a problem remains if the robot on which an image of a visitor is projected has a machine-like appearance. Another approach that emphasizes the robot's appearance is to design a telerobot with a minimal human appearance. More precisely, we design a telecommunication robot that has enough humanlike appearance for anybody to feel its telepresence and to be felt by locals in local sites. Such a minimal human likeness enables any person as a visitor to convey his/her personality to locals with less distortion from the robot's appearance and to have a sense of body ownership. The exploration of minimal human likeness is a promising approach to achieve an effective teleoperation system for telecommunication.

To explore such external characteristics, the telenoid[2] has been developed as a testbed with a minimal human likeliness design so that it resembles anybody (Fig. 5(a)). The telenoid, which is 70-cm long and weighs about 4 kg, has nine degrees of freedom (DoFs), most of which are assigned to control its eyes, mouth, and head; the rest are for its right and left hands (Fig. 5(b)). It is covered with soft vinyl chloride[3] to resemble a human.

The teleoperation system for the telenoid is the same as the one for the geminoid except for the predefined behaviors. Due to its simpler body structure, such easy behaviors as waving goodbye or giving a hug were implemented. The system is easy to use and carry because it has a face tracking system based on camera image and a speech-driven lip motion generation system [19]; it only requires a single laptop with a web camera.

A. Design concept

While such existing teleoperated androids as Geminoid HI-1 [15], [16] were designed to resemble a living individual's appearance as closely as possible, the telenoid's objective is to create a minimal human that allows people to feel as if any visitor was actually close. It includes the following key features: 1) an omni-human likeness that enables users to feel any person's presence (e.g., a feeling of being there); 2) holdability that facilitates physical interaction; and 3) mobility that encourages people to use it in a variety of situations [18].

For a minimal human design, the robot's appearance should avoid preconceived ideas. Therefore, we removed as many unnecessary features as possible from the telenoid by: 1) choosing features for communication, e.g., voice, with humans and eliminating non-neutral ones, e.g., beard; 2) reevaluating the chosen features to fit the design requirements by eliminating unnecessary features; and 3) obtaining essential features. As a result, it can be perceived as either male or female, young or old.

B. Reduction of distortion derived from the telerobot's appearance

Kuwamura *et al.* addressed whether the telenoid can reduce the distortion of the visitor's personality and hypothesized that telecommunication robots whose appearances differ from humans distort the visitor's personality more than telerobots that resemble humans [36]. In their experiments, participants had conversations with a visitor who talked through one of three physical entities with different appearances (the telenoid, a stuffed-bear robot that resembled a RobotPhone [8], and a video chat) in three different face-to-face conversations: free talk, hearing a visitor's self-introduction, and being interviewed by a visitor.

The visitor's personality was rated using the Japanese Big Five personality test [37], which is applicable for rating the personalities of others as well as one's own (e.g. [11]). The test, which consists of a 60-item questionnaire answered on a seven-point scale for pairs of antonyms, measures five parameters of personality (extraversion, neuroticism, openness to experience, agreeableness, and conscientiousness), each of which includes twelve different items concerning the parameter. The effect of the distortion by the appearance

[2]The hardware information is about Telenoid R2, the newest telenoid model, which is lighter and smaller than Telenoid R1.

[3]Telenoid R1 was covered with silicon rubber.

of a physical entity was evaluated with the consistency of the answers to the twelve items concerning a parameter in the Big Five test.

The experimental results showed that the stuffed-bear medium had low consistency on extraversion under the interview situation and agreeableness under the self-introduction situation. Such low consistency was not observed for the telenoid and video chat cases. This indicates that the personality transmitted through the stuffed-bear robot was distorted under certain situations, suggesting that a minimal design of a humanlike appearance can reduce the distortion of the visitor's personality.

VII. Discussion and Conclusion

A. Summary

Many telecommunication robots have been developed so far to facilitate telecommunication among humans. However, their appearance was quite different from that of a human, although they were designed to act as physical proxies for a human. To fully convey visitor's information to a local at a local site, we cannot ignore the design issue about the telerobot's appearance. This paper has shown the possible impact of the telerobot's appearance on telecommunication, focusing on both aspects of visitor feelings of telepresence and local feelings of visitor's presence.

B. Investigation on a sense of ownership

An interesting fact is that visitors experience a sense of ownership over teleoperated androids even though no tactile feedback exists. This sense enhances the feeling of telepresence. Although the sense of ownership is only experienced over humanlike body parts [21], it remains unclear how much human likeness is needed to achieve such a sense. Can we have the sense when we use humanoid robots or mobile robots? Another important issue is to investigate what condition is required to evoke a strong sense of ownership. These questions are future work.

One might argue that what we described here is not a sense of body ownership but rather a sense of agency because controlling teleoperated androids provides visitors with the sense that they are causing an action of the androids [38], [39]. Since our aim is facilitating telecommunication, it does not matter what underlying mechanism is driven by humanlike appearance as long as it enables visitors to strongly feel telepresence in a local site. Nevertheless, more deeply understanding the mechanism should be addressed in the future because it improves teleoperation systems.

Note that it also remains unclear whether a visitor's sense of ownership enhances a local's feeling of visitor's presence. Visitors might feel uncomfortable interacting with locals through telerobots unless the locals respond to the telerobots like they responded to the visitors. For example, the locals approach the robots at a distance that makes the visitors feel uncomfortable if they do not consider the robots as the visitors. This problem might not happen when visitors have a strong sense of ownership because they express their discomfort, such as moving the robots away from the locals. Such reactions will help the locals understand the visitors' feelings and treat the robots as the visitors. Therefore, to facilitate telecommunication, we must study the relation between the visitor's sense of ownership and the local's feeling of the visitor's presence.

C. Influence of visitor's presence on locals

A main advantage of using teleoperated androids is that they enable us to provide a strong presence to locals who are at another locations. Group work with video conferencing systems faces a problem because remote people become "out-of-sight and out-of-mind" at local workplaces [40]. This is not true of teleoperated androids due to their strong presence. Therefore, visitors can be actively involved in group work through teleoperated androids. Actually, a field trial about the group work of children using a telenoid suggested that, once the member being mediated by a telenoid was accepted as a group member with the help of the others, all the children worked cooperatively [41]. However, this study only observed group activity for a few days. Lee *et al.* introduced a mobile remote telepresence system into a workplace and observed interaction between it and local workers for 2-18 months to investigate how remote workers are accepted by local workers. They reported that the system supported informal communications and connections among distributed coworkers [42]. Such long-term observation of human-human interaction mediated by a teleoperated android must be conducted as future work to examine how teleoperated androids are accepted by children, workers, and senior citizens and whether teleoperated androids have unique influence compared with mobile robots.

One concern about using teleoperated androids is that locals might experience the uncanny valley, which refers to a sense of unease and discomfort when people look at increasingly realistic virtual/artificial humans [29]. Locals might have weird feelings about them if the android appearances and movements are unnatural. However, anecdotal evidence suggests that locals can overcome the uncanny valley. For example, as described in Section IV, the children's responses toward the geminoid became close to their responses toward the real person, although one mentioned that the geminoid was a person wearing a strange mask [17]. Nishio *et al.* reported that ordinary people who saw the geminoid for the first time had weird and nervous feelings, but shortly after a conversation with it, they interacted with it without any anxious feelings [16]. Such adaptation was also found in short interactions mediated by the telenoid [24]. Almost half of ordinary people interviewed in field trials reported positive impressions about the telenoid. About 36% had negative impressions. Interestingly, however, their attitudes became positive after they hugged it. Positive attitudes toward the telenoid were shared between schoolchildren [41] and the elderly [24], [43] in different countries [44]. These studies imply that brief interaction with teleoperated androids erases or at least reduces anxious impressions about them. We must investigate whether unnaturalness in the teleoperated android's appearance and movements prevents locals from feeling the visitors' presence even after the locals have interacted with the visitors mediated by the androids.

D. Exploration of minimal human likeness design

Although the telenoid's design is based on empirical knowledge derived from research on androids and geminoids, further exploration is needed on the minimal design of human likeness. For example, should legs, hands, and actuators for facial expressions be included in a minimal design of a human appearance? We must consider whether the appearance evokes a visitor's sense of ownership and whether it lacks a strong personality to distort the visitor's personality.

Moreover, note that higher realism in the telerobot's appearance might increase expectations for behavioral realism [45]. Such high expectations negatively affect the impressions of visitors as suggested by a study on group work in virtual reality that reported that the person represented by the more realistic avatar was seen as standoffish and cold because of a lack of expression [46]. Therefore, future work must address the minimal design of a telerobot with a humanlike appearance from both aspects of its appearance and communication skills that are expected by the appearance.

E. Improvement of teleoperation systems

The design of an interface to control a telerobot is another crucial issue to achieve a visitor's feeling of telepresence. A perspective to control the telerobot should be taken into account because the rubber hand illusion does not occur when the rubber arm was not aligned with the orientation of the real arm [47]. A recent study, however, showed that people whose visual ego-center was shifted to behind themselves with a head-mounted display felt a strong illusion of being behind their own physical bodies when their chest was tapped out of sight while they see tapping movement around bottom of their sight in the head-mounted display [48]. As mentioned in Section III, a sense of ownership occurred even when participants observed the geminoid from a third-person point of view [24]. This implies that visitors do not need to have the same point of view as the teleoperated android to feel telepresence. In addition, the input device to control the teleoperated android should also be reconsidered because a sense of ownership occurs even when the geminoid is controlled with the visitor's brain activity [25]. Improving our user friendliness in operating a telerobot will allow ordinary people to accept mediated human-human interaction.

References

[1] E. Paulos and J. Canny. Social tele-embodiment: Understanding presence. *Autonomous Robots*, Vol. 11, No. 1, pp. 87–95, 2001.

[2] H. Kuzuoka, S. Oyama, K. Yamazaki, K. Suzuki, and M. Mitsuishi. Gestureman: a mobile robot that embodies a remote instructor's actions. In *Proc. of the Conference on Computer Supported Cooperative Work*, pp. 155–162. ACM, 2000.

[3] S.O. Adalgeirsson and C. Breazeal. Mebot: a robotic platform for socially embodied presence. In *Proc. of the 5th ACM/IEEE International Conference on Human-robot Interaction*, pp. 15–22, 2010.

[4] S. Tachi, N. Kawakami, H. Nii, K. Watanabe, and K. Minamizawa. Telesarphone: Mutual telexistence master-slave communication system based on retroreflective projection technology. *SICE Journal of Control, Measurement, and System Integration*, Vol. 1, No. 5, pp. 335–344, 2008.

[5] C. DiSalvo, F. Gemperle, J. Forlizzi, and E. Montgomery. The hug: an exploration of robotic form for intimate communication. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 403–408. IEEE, 2003.

[6] S. Coradeschi, A. Loutfi, A. Kristoffersson, S. Von Rump, A. Cesta, G. Cortellessa, and J. Gonzalez. Towards a methodology for longitudinal evaluation of social robotic telepresence for elderly. In *Proc. of Workshop on Social Robotic Telepresence in International Conference on Human-robot Interaction*, 2011.

[7] D.I. Fels and P.L.T. Weiss. Video-mediated communication in the classroom to support sick children: A case study. *International Journal of Industrial Ergonomics*, Vol. 28, No. 5, pp. 251–263, 2001.

[8] D. Sekiguchi, M. Inami, and S. Tachi. Robotphone: Rui for interpersonal communication. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 227–278, 2001.

[9] K. Goris, J. Saldien, I. Vanderniepen, and D. Lefeber. The huggable robot probo, a multi-disciplinary research platform. Research and Education in Robotics—EUROBOT 2008, pp. 29–41, 2009.

[10] W. D. Stiehl, J. K. Lee, C. Breazeal, M. Nalin, A. Morandi, and A. Sanna. The huggable: a platform for research in robotic companions for pediatric care. *Proc. of the 8th International Conference on Interaction Design and Children*, pp. 317–320, 2009.

[11] L.P. Naumann, S. Vazire, P.J. Rentfrow, and S.D. Gosling. Personality judgments based on physical appearance. *Personality and Social Psycholoty Bulletin*, Vol. 35, No. 12, pp. 1661–1671, 2009.

[12] M. Shimada and H. Ishiguro. Motion behavior and its influence on human-likeness in an android robot. In *Proc of the Annual Conference of the Cognitive Science Society*, pp. 2468–2473, 2008.

[13] J. Goetz, S. Kiesler, and A. Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proc. of the 12th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 55–60, 2003.

[14] T. Komatsu and S. Yamada. Adaptation gap hypothesis: How differences between users' expected and perceived agent functions affect their subjective impression. *Journal of Systemics, Cybernetics and Informatics*, Vol. 9, No. 1, pp. 67–74, 2011.

[15] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro, and N. Hagita. Android as a telecommunication medium with a human-like presence. In *Proc. of Interenational Conference on Human-robot Interaction*, pp. 193–200, 2007.

[16] S. Nishio, H. Ishiguro, and N. Hagita. Geminoid: Teleoperated android of an existing person. *Humanoid Robots - New Developments. I-Tech*, 2007.

[17] S. Nishio, H. Ishiguro, and N. Hagita. Can a teleoperated android represent personal presence?—a case study with children. *Psychologia*, Vol. 50, No. 4, pp. 330–342, 2007.

[18] K. Ogawa, S. Nishio, K. Koda, G. Balistreri, T. Watanabe, and H. Ishiguro. Exploring the natural reaction of young and aged person with telenoid in a real world. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 15, No. 5, pp. 592–597, 2011.

[19] C. T. Ishi, L. Chaoran, H. Ishiguro, and N. Hagita. Speech-driven lip motion generation for tele-operated humanoid robots. *International Conference on Auditory-Visual Speech Processing*, pp. 131–135, 2011.

[20] M. Botvinick and J. Cohen. Rubber hands 'feel' touch that eyes see. *Nature*, Vol. 391, No. 6669, pp. 756–756, 1998.

[21] M. Tsakiris, L. Carpenter, D. James, and A. Fotopoulou. Hands only illusion: multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research*, Vol. 204, No. 3, pp. 343–352, 2010.

[22] S. Nishio, T. Watanabe, K. Ogawa, and H. Ishiguro. Body ownership transfer to teleoperated android robot. In *Proc. of International Conference on Social Robotics*, 2012. (To appear).

[23] T. Watanabe, S. Nishio, K. Ogawa, and H. Ishiguro. Body ownership transfer to an android by using tele-operation system. *IEICE Transactions on Information and Systems*, Vol. J94-D, No. 1, pp. 86–93, 2011. In Japanese.

[24] K. Ogawa, K. Taura, S. Nishio, and H. Ishiguro. Effect of perspective change in body ownership transfer to teleoperated android robot. In *Proc. of 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012. (To appear).

[25] M. Alimardani, S. Nishio, and H. Ishiguro. Body ownership transfer to tele-operated android through mind controlling. In *Proc. of Human-agent Interaction Symposium*, pp. I–2A–1, 2011.

[26] S. Nishio and H. Ishiguro. Attitude change induced by different apperances of interaction agents. *International Journal of Machine Consciousness*, Vol. 3, No. 1, pp. 115–126, 2011.
[27] M. Mori, K.F. MacDorman, and N. Kageki. The uncanny valley. *IEEE Robotics & Automation Magazine*, Vol. 19, No. 2, pp. 98–100, 2012.
[28] R.E. Guadagno, J. Blascovich, J.N. Bailenson, and C. Mccall. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, Vol. 10, No. 1, pp. 1–22, 2007.
[29] H. Brenton, M. Gillies, D. Ballin, and D. Chatting. The uncanny valley: does it exist. In *Proc. of Workshop on Human-animated Character Interaction in Conference oo Human Computer Interaction*. Citeseer, 2005.
[30] H. Ishiguro. Android science–toward a new cross-interdisciplinary framework. *Robotics Research*, Vol. 28, pp. 118–127, 2007.
[31] K.F. MacDorman and H. Ishiguro. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, Vol. 7, No. 3, pp. 297–337, 2006.
[32] F.E. Pollick. In search of the uncanny valley. *User Centric Media*, pp. 69–78, 2010.
[33] I. Straub, S. Nishio, and H. Ishiguro. Incorporated identity in interaction with a teleoperated android robot: A case study. In *Proc. of International Symposium on Robot and Human Interactive Communication*, pp. 119–124, 2010.
[34] N. Yee and J. Bailenson. The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, Vol. 33, No. 3, pp. 271–290, 2007.
[35] H. Sumioka, S. Nishio, E. Okamoto, and H. Ishiguro. Isolation of physical traits and conversational content for personality design. In *IEEE International Symposium on Robot and Human Interactive Communication*, 2012. (To appear).
[36] K. Kuwamura, T. Minato, S. Nishio, and H. Ishiguro. Personality distortion in communication through teleoperated robots. In *IEEE International Symposium on Robot and Human Interactive Communication*, 2012. (To appear).
[37] S. Wada. Construction of the Big Five Scales of personality trait terms and concurrent validity with NPI. *Japanese Journal of Psychology*, Vol. 67, No. 1, pp. 61–67, 1996. in Japanese.
[38] E. van den Bos and M. Jeannerod. Sense of body and sense of action both contribute to self-recognition. *Cognition*, Vol. 85, No. 2, pp. 177–187, 2002.
[39] M. Botvinick. Probing the neural basis of body ownership. *Science*, Vol. 305, No. 5685, pp. 782–783, 2004.
[40] G. Venolia, J. Tang, R. Cervantes, S. Bly, G. Robertson, B. Lee, and K. Inkpen. Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams. In *Proc. of Conference on Human Factors in Computing Systems*, pp. 1049–1058. ACM, 2010.
[41] R. Yamazaki, S. Nishio, K. Ogawa, H. Ishiguro, K. Matsumura, K. Koda, and T. Fujinami. How does telenoid affect the communication between children in classroom setting? In *Proc. of the Conference on Human Factors in Computing Systems*, pp. 351–366, 2012.
[42] M.K. Lee and L. Takayama. Now, i have a body: Uses and social norms for mobile remote presence in the workplace. In *Proc. of the Conference on Human Factors in Computing Systems*, pp. 33–42. ACM, 2011.
[43] R. Yamazaki, S. Nishio, H. Ishiguro, T. Minato, M. Norskov, N. Ishiguro, M. Nishikawa, and T. Fujinami. Social inclusion of senior citizens by a teleoperated android: toward inter-generational telecommunity creation. In *Proc. of Workshop on Assistance and Service Robotics in a Human Environment in the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012. (to appear).
[44] R. Yamazaki, S. Nishio, H. Ishiguro, M. Nrskov, N. Ishiguro, and G. Balistreri. Social acceptance of a teleoperated android: Field study on elderly's engagement with an embodied communication medium in denmark. In *Proc. of International Conference on Social Robotics*, 2012. (To appear).
[45] M. Slater and A. Steed. Meeting people virtually: Experiments in shared virtual environments. *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, pp. 146–171, 2002.
[46] J. Tromp, A. Bullock, A. Steed, A. Sadagic, M. Slater, and E. Frécon. Small group behaviour experiments in the coven project. *Computer Graphics and Applications, IEEE*, Vol. 18, No. 6, pp. 53–63, 1998.
[47] F. Pavani, C. Spence, and J. Driver. Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science*, Vol. 11, No. 5, pp. 353–359, 2000.
[48] H.H. Ehrsson. The experimental induction of out-of-body experiences. *Science*, Vol. 317, No. 5841, p. 1048, 2007.