Supporting Referential Gestures in Mobile Remote Presence: A Preliminary Exploration

Victor Kaptelinin

Department of Informatics, Umeå University
90531 Umeå, Sweden
victor.kaptelinin@umu.se

Abstract. The paper discusses recent developments in mobile remote presence (MRP) and argues that providing support for referential gesturing is critically important for exploiting the full potential of MRP systems in creating inclusive smart environments. By bringing in insights from research in interaction design and related fields the paper proposes a set of tentative requirements to referential gesturing support for MRP systems. Some of the key challenges for designing referential gesturing in the context of mobile remote presence are identified.

Keywords: mobile remote presence, embodied telepresence, referential gestures, interaction design, human-computer interaction, computer supported cooperative work.

1 Introduction

Mobile Remote Presence (MRP) systems are remotely controlled devices that serve as physical avatars of their "pilots" and support pilots' embodied social presence in a local setting. MRP systems are increasingly common in various contexts, including healthcare and elderly care environments. The use of MRP systems opens up a possibility to extend the range of physical and social activities available to the elderly and people with special needs, as well as their family, friends, and caretakers, and thus provides more favorable conditions for societal inclusion and empowerment. However, interaction capabilities of existing MRP systems (at least, affordable ones) are limited, which undermines the prospects for using this technology for creating inclusive smart environments.

This paper argues that one of the most significant limitations of many existing MRP systems, which need to be addressed in further research, is a lack of support for referential gesturing. A preliminary exploration, from an interaction design perspective, of challenges and potential solutions associated with addressing this limitation is presented. The paper reports on a work in progress: the analysis in the paper introduces the rationale and general conceptual point of departure for a recently started research project, supported by the Swedish Research Council, which aims to investigate interaction design issues related to mobile remote presence in smart environments.

2 Mobile remote presence, interactional empowerment, and social inclusion

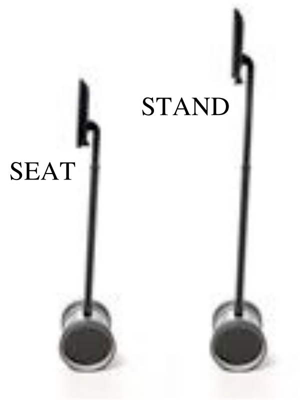

Fig. 1. Telepresence robot Double by Double Robotics (www.doublerobotics.com); SEAT: a sitting height, STAND: a standing height

Mobile Remote Presence (MRP) systems, or telepresence robots, typically comprise the following parts (see Figure 1): (a) a "head", that is, a video-conferencing unit, including a camera, microphone, speakers, and a display; in some models, such as Double by Double Robotics, the unit can simply be a standard tablet computer, (b) a wheeled base, which can be used to move the device in a setting, and (c) an elongated vertical part, such as a pole, which connects the video-conferencing unit and the wheeled base; this part is often designed so that the height of the "head" can be adjusted depending on the context, for instance, on whether the communication partner is standing or sitting. A diversity of MRP systems, which differ in the size of their displays, quality and quantity of cameras, inclusion of additional equipment (such as a laser pointer), and so forth, is currently available on the market (see, e.g., [3], [9], [10]). MRP systems can be controlled by their "pilots" from remote locations via computing devices connected to the Internet.

A key advantage of MRP systems compared to conventional video-conference technologies is that they serve as embodied social proxies of their pilots in a local setting. The pilots can move around in the setting, view people and the environment from various angles, start conversations where and when it is appropriate, join discussions taking place in different locations, and so forth. A number of studies have been conducted into the effect of robotic telepresence on the quality of interaction (e.g., [2], [3], [4], [5], [7], [11]). The effect was found to be generally positive: the perceived social presence of remote workers using the technology was almost at the same level as that of people who were physically present in a setting.

One area, in which the use of MRP systems demonstrated a positive impact on the quality of interaction, is elderly care ([2], [3], [4]). Implementation of MRP systems, such as Giraff by Giraff Technologies, in assisted living settings provided advantages to the elderly, as well as their relatives and caregivers ([3]). Since the visits did not require physical travel, the elderly could enjoy more visits and spend more time interacting with other people.

Arguably, current uses of telepresence robots are just the first steps, and exploiting the full potential of MRP systems in healthcare and assisted living requires further research and development. First, the interaction between local people, e.g., elderly persons, and their visitors, embodied as MRP systems is not always optimal and needs to be improved (e.g., [4]). Second, and more importantly, there are promising novel ways of using the technology, which need to be explored. For instance, one can envision inclusive smart cities of the future that support the elderly, as well as people with special needs, in using MRP technology for visiting, with social presence, remote places that cannot be visited physically.

As argued below, a condition for successfully accomplishing these goals is the development of more advanced solutions for grounding MRP system-mediated social interactions in the specific physical and social contexts of a local setting, and, in particular, providing support for referential gestures.

3 Enabling MRP systems with referential gesturing capabilities: Why is it essential?

While a wide range of telepresence robots is currently commercially available ([3], [9]), most of them are, essentially, just video conferencing systems on wheels. It is true that remote pilots can move around in a local setting, view it from different angles, approach people they want to communicate with, and take an active part in deciding on the place and time of a conversation. However, when a conversation commences, it is usually similar to a conventional videoconference session in the sense that references to the local physical context are not particularly extensive.

Referencing instruments, which are most commonly used in MRP systems, such as QB and MantaroBot ([3], [9]), are laser pointers. Such pointers can be used for instance, for highlighting certain areas of a large screen display, jointly viewed by the pilot and local participants. A disadvantage of laser pointers is that the signals they produce can be easily overlooked if a person is not expecting them, and therefore using such pointers does not necessarily support efficient attention management. In addition, the use of laser pointers can be unsafe. Another solution, implemented in PEBBLES, an MRP system primarily designed for educational settings (e.g., classrooms), is providing robots with a "hand" that can be controlled by the pilot. Empirical evaluation of PEBBLES suggests that the hand may not be an optimal solution: the use of the hand is limited to very simple gestures, such as waving it to draw teacher's attention. In addition, the hand raises some safety concerns ([3]).

Probably, the most advanced functionality for pointing and referential gesturing by a telepresence robot is implemented in a series of experimental systems named GestureMan, created at the University of Tsukuba ([5], [10]). The design of GestureMan robots provides a number of potentially useful insights. It should be noted, however, that the robots are not intended as general-purpose embodied social proxies. Instead, they are designed for specific purposes, such as remote instruction on physical tasks. It is not clear whether the solutions implemented in these technologies are practical or affordable if they are used in other contexts.

Therefore, typical MRP systems support interactions *in* the physical context of a local setting, but not so much *about* the context. Arguably, in many cases this limitation is not critical. For instance, when workers have an impromptu work-related discussion in a hallway, they usually do not discuss the physical context (that is, the hallway) itself. In such cases existing telepresence robots may be generally sufficient to meet workers' communication needs. In a similar vein, when a relative, embodied as an MRP system, visits an elderly person, they do not necessarily talk about objects in the local setting, so supporting referential gestures may not be important.

However, both logical arguments and empirical evidence indicate that users of MRP systems may need something more than just "video conferencing systems on wheels", and extended gestural capabilities is likely to be one of the most needed features. First, supporting referential gesturing is important for a successful collaborative work on a shared object, as opposed to mere communication about the object, so a deeper appropriation of MRP systems for collaboration is likely to require supporting referential gesturing. Second, in some cases communication about the local context appears to be important for successful interaction in existing healthcare and assisted living environments (e.g., in the case of questions like "Did you take *these* pills?"). Third, supporting referential gesturing by telepresence robots is crucial for making it possible for the elderly and people with special needs to employ MRP systems for virtual trips, e.g., in inclusive smart cities. If such scenarios are realized then, apparently, communication with a direct reference to objects in the local context is going to be a key type of communication between the pilot and people in a local setting. Fourth, empirical studies of communication involving MRP systems deployed in real-life settings reveal a number of problems with making appropriate spatial arrangements during such communication (e.g., [4]). Supporting referential gestures, which may help coordinate participants' mental models of their shared spatial context, appears to be a promising way to alleviate these problems.

4 Interaction design of referential gesturing in mobile remote presence: Identifying tentative requirements

Enhancing MRP systems with referential gesturing capabilities can be approached from different perspectives, including software and hardware development, management, and so forth. The analysis in this paper adopts an interaction design perspective. In interaction design and related fields, such as HCI and CSCW, the focus of research is on how people interact with technology and employ it in their everyday activities, rather than on technology per se. Pointing and referential gesturing have been studied from an interaction design perspective for over two decades [8], and some of these studies offer a number of useful insights for exploring possible strategies for supporting referential gesturing in mobile remote presence. The discussion below capitalizes on these studies (e.g., [1], which presents an evaluation methodology for shared workspace groupware) to identify a set of tentative requirements to referential gesturing in mobile remote presence.

Supporting a diversity of communicative purposes. In face-to-face communication gestures serve a diversity of roles. A gesture performed by an actor can refer, for instance, to (a) actor's actual or intended *communication partner*, (b) an *object* in the local setting the actor wants to bring to communication, (c) the communicative *intention* of the actor, or (d) actor's *request for action*, directed to people in the setting. When designing robots' gestural capabilities one should aim for supporting a range of such purposes. It should be noted that it is not uncommon for the same gesture to have several purposes at the same time.

Making the difference between inactive communication mode and active communication mode clearly visible. To avoid possible confusion and address privacy concerns a telepresence robot should provide clear cues signifying its current mode, that is, whether the robot is in an inactive mode (no gesture or other communicative action is performed or planned to be performed) or active mode (a gesture or other communicative action is being performed or can be expected soon).

Coordinating gestures with other communication modalities. The meaning of a referential gesture often depends on how the gesture in question is related to other communication modalities, such as voice, gaze, and posture. For instance, pointing to an object while maintaining eye contact with person A can have a different meaning than pointing to the same object while maintaining eye contact with person B.

Employing unambiguous, widely known, and easily recognizable gestures. Gestural communication with robots is a rather unusual type of activity for many people, and if the gestures performed by telepresence robots are complicated and unclear, people in the local setting may be disoriented.

Providing simple and intuitive gesture control. Typical users serving as pilots of MRP systems and employing widely available technologies, such as conventional computers, may find it challenging to even navigate a system. To avoid potential cognitive overload caused by the need to carry out additional tasks, gesture control features of the pilot's user interface should be as simple and intuitive as possible.

Ensuring safety. Physical gestures performed by telepresence robots in a local setting are associated with an increased risk of causing harm to people and material damage to objects in the environment (including the MRP system itself). A key requirement, therefore, should be taking all necessary steps to minimize the risk. It can be done, for instance, by imposing constraints on robot's movements, introducing emergency controls to be available to local people, and using soft-impact components (e.g., [6]).

Enabling low effort interruption management. MRP systems' pilots are present in two settings simultaneously, which makes them especially prone to interruptions. To avoid sudden disruptions in a local setting (e.g., when an MRP system freezes in the middle of a gesture because the pilot is distracted by what happened in his or her physical setting), pilots should be able to easily deal with such cases, for instance, starting a standard interruption management procedure by simply pressing a button.

This tentative list is intended as a starting point in identifying a more comprehensive set of requirements. It should be noted that meeting the above requirements are associated with a number of challenges. For instance, coordinating gestures with other communication modalities is complicated by problems with using gaze in telerpesence in general, and employing widely known and easily recognizable gestures is complicated by cross-cultural differences in gestural communication.

5 Conclusion

This paper argues that referential gesturing, a centrally important aspect of human communication, should be supported in the design of MRP systems to fully exploit the potential of such systems in the search for novel solutions for assisted living and inclusive smart cities. This paper reports on a work in progress; the analysis presented in the paper has been conducted to inform further design-based exploration of mobile remote presence in smart environments within an ongoing research project.

Acknowledgments. The study was funded by The Swedish Research Council. The author is grateful to Patrik Björnfot and two anonymous reviewers for their helpful comments.

References

1. Baker K., Greenberg S., Gutwin, C. Empirical Development of a Heuristic Evaluation Methodology for Shared Workspace Groupware. In: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, pp. 96—105. ACM Press, NY (2002)

2. Beer, J.M., Takayama, L., Mobile Remote Presence Systems for Older Adults: Acceptance, Benefits, and Concerns. In: Proceedings of the 6th International Conference on Human-Robot Interaction (HRI'11), pp. 19—26. ACM Press, NY (2011)
3. Kristoffersson, A., Coradeschi, S., and Loutfi, A. A Review of Mobile Robotic Telepresence. Advances in Human-Computer Interaction. V. 2013 (2013), article ID 902316, 17 pages.
4. Kristoffersson, A., Coradeschi, S., Severinson Eklundh, K., Loutfi, A. Assessment of Interaction Quality in Mobile Robotic Telepresence: An Elderly Perspective. Interaction Studies 15, 2, 343--357 (2014)
5. Kuzuoka, H., Suzuki, Y., Yamashita, J., Yamazaki, K. Reconfiguring Spatial Formation Arrangement by Robot Body Orientation. In: Proceedings of the 5th International Conference on Human- Robot Interaction (HRI'10), pp. 285–292. ACM Press, NY (2010)
6. Mihaj, D. Pneuduino: A Modular Platform for Controlling Airflow and Pressure. Smashing Robotics, February 2, 2016 (2016). http://smashingrobotics.com/pneuduino-modular-platform-for-controlling-airflow-and-pressure/
7. Rae, I., Mitlu, B., and Takayama, L. Bodies in Motion: Mobility, Presence, and Task Awareness in Telepresence. In: Proceedings of CHI'14, pp. 2153-2162 ACM Press, NY (2014).
8. Schmauks, D., Wille, M. Integration of Communicative Hand Movements into Human-Computer-Interaction. Computers and the Humanities, 25, 2, 129--140 (1991)
9. TelepresenceRobots.com, http://telepresencerobots.com/comparison
10. Veloso, M., Aisen, M., Howard, A., Jenkins, C., Mutlu, B., Scassellati, B. WTEC Panel Report on Human-Robot Interaction: Japan, South Korea, and China. WTEC (2012).
11. Tsui, K.M., Desai, M., Yanco, H.A., Uhlik, C. Exploring Use Cases for Telepresence Robots. In: Proceedings of the 6th International Conference on Human-Robot Interaction (HRI'11), pp. 11–-18. ACM Press, NY (2011)