People with Motor Disabilities Using Gaze to Control Telerobots

Zhang, Guangtao; Hansen, John Paulin

Published in:
Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems

Link to article, DOI:
10.1145/3334480.3382939

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
People with Motor Disabilities Using Gaze to Control Telerobots

Guangtao Zhang
Technical University of Denmark
Kgs. Lyngby, Denmark
guazha@dtu.dk

John Paulin Hansen
Technical University of Denmark
Kgs. Lyngby, Denmark
jpha@dtu.dk

Abstract
Telerobots may give people with motor disabilities access to education, events and places. Eye-gaze interaction with these robots is an option when hands are not functional. Gaze control of telerobots has not yet been evaluated by people from this target group. We conducted a field study with five users in a care-home to investigate their preferences and challenges when driving telerobots via our gaze-controlled robotic telepresence system. We used a Wizard of Oz method to explore gaze and speech interaction, and experience prototyping to consider robot designs and types of displays (e.g., monitors versus head-mounted displays).

Author Keywords
Human-robot interaction, gaze interaction, telepresence, accessibility, motor disabilities, assistive technology, inclusion, robot-mediated communication, navigation, head-mounted displays, experience prototyping, Wizard of Oz.

CCS Concepts
- Human-centered computing → Accessibility; Accessibility systems and tools;

Introduction
Telerobots have several potential usages for people with motor disabilities (MD) [26]. Recent development of telepresence robots [11] may provide people with MD a sense of

Figure 1: One of the participants communicates with his friends via a telepresence robot.
presence within a remote environment. Several manufacturers offer wifi-based telerobots, e.g. Double Robotics and Padbot (see Fig. 2).

Robotic telepresence systems [32] have become increasingly popular for collaboration between geographically distributed teams[12], at academic conferences[23], for children-teacher communication[29], for relationships of long-distance couples [36], for seniors with mobility impairments[27], and for outdoor activities [8]. However, people with MD have been ignored when studying telerobot interaction [34]. Previous studies on assistive technology emphasized the importance of independence for people with MD, who often preferred to retain as much control authority as possible[18]. So while fully autonomous telpresence robots have been suggested in previous work, for instance automatically following a person in a remote environment [4], it is an open question as to level of autonomy people with MD would prefer.

Alternative interaction with telerobots

Keyboard and mouse or joysticks are most commonly used for telerobot systems. Alternative vehicle- and wheelchair-control methods have been suggested for people with MD, for instance based on brain-computer-interaction (BCI) [30, 31, 13, 2, 21], speech [33, 34], gaze interaction [5, 39, 20], and fMRI recording of covert visuospatial attention [1].

Brain-computer interaction [30] may be used by people with severe motor neuron diseases such as Amyotrophic Lateral Sclerosis (ALS), even when they can not move any part of their body - a so-called Locked-In Syndrome. Some of the BCI-methods, however, require substantial training and set-up, since caps and gel-based electrodes can be difficult to apply. Also, information transfer with BCI systems are quite low; for example [9] reported a throughput of 0.05-1.44 bits/s.

Speech interfaces are popular for mobile phones and smart speakers. They have been suggested for telerobot navigation also. Unfortunately, some people with MD have impaired speech. Low-level direction guiding by speech can be cumbersome as a user may have to repeat the same command multiple times (e.g. left, left, left), or rely on precise timing [33].

Gaze interaction has been used by people with MD for more than twenty years, mainly for communication and pointer-input [16]. Throughput values for gaze have been reported to be 2.55 bits/s [17]. Gaze-controlled navigation of robots was first studied by Tall et al. in 2009 [28]. Since then, gaze tracking technology has become inexpensive and accurate. However, it remains a challenge to design interfaces that only respond to intentions, and not just eye movements. Head input has shown to be more precise and less erroneous than gaze [17, 6] but it may also be tiring to use for longer time.

Besides a traditional set-up with monitors, head-mounted displays (HMD) are now being tested for telerobots with a 360 degree camera mounted on the telerobot [5, 39] with gaze- or head- movements changing the field of view [10, 5]. Gaze tracking sensors are built into some recent HMD’s. The increased availability of eye-tracking equipment makes it feasible to utilize this technology for explicit control tasks with robotics [35]. Hence, we have enabled gaze control of telepresence robots with an HMD in previous studies [39, 38]. The key findings were that after training, able-bodied participants could effectively navigate telerobots by gaze. VR-simulations also showed to be a safe and efficient training tool for gaze navigation [37]. However, like other studies, e.g. [20], we did not involve people with MD. So it is important now to engage people with MD in the design of gaze-controlled telepresence robots and explore future use-
cases with them. We initiated this process with five participants in a field test presented in this paper.

**Field study**

Our study applied experience prototyping [3] and the Wizard of Oz technique [24]. The goal was to observe the use and effectiveness of various interfaces and control methods with our target users, rather than to compare a range of complete robotic systems. We therefore used the Wizard of Oz technique for some of our scenarios because of its advantages in early explorations, where complex system intelligence can be delivered by a human assistant instead of being fully implemented. Also, without an open API to the Double robots it would be a demanding task to establish a direct gaze control.

According to principles suggested by [25] when conducting accessibility research, our study-plan and consent form were sent to the caregivers before the study. Observations and conversations during the study were noted by the experimenter, supported by records of the participants’ spontaneous verbalisation, system log files, screen recordings, and room video recordings for post-hoc analysis.

**Participants**

Five people (age 21-55 years; 3 men and 2 women) with different levels of motor impairments participated. They live in a care-home together with 90 people who use wheelchairs for daily mobility. Some of the residents use gaze interaction for communication and some use speech commands for smart-home control. All of our five participants have impaired manual activity, for instance limited gripping, finger or holding ability, due to cerebellum disorders or cerebral palsy. None of them have cognitive impairments or literacy difficulties, but one of them has impaired speech functions. Three of them have experienced VR using an HMD and one has experienced a telerobot.

**Procedure**

Participants conducted the study one by one. After greetings and a brief introduction, we collected demographic information, including experience with gaze control, VR, telerobots and wheelchair control. For the one individual who is non-verbal, this information was provided by a care giver. In particular, participants were informed that if they felt uncomfortable at any time they should stop immediately. Then we showed a demo video of one of our earlier telepresence robot projects [5] with a person controlling a telerobot by his gaze while laying in bed. Two different telerobots (see fig. 2) were standing next to the table we were sitting at.

Initially, we conducted an interview focusing on their expectations and visions of possible usage of telerobots for their daily life activities. Then followed the two experience prototyping sessions as described below. We presented a range of options to them, including types of displays (HMD versus screen), independence levels (independent versus assisted by others), and control methods (gaze, speech or hand).

**Gaze navigation**

In the first experience prototyping session, we used our gaze-controlled robotic telepresence system (see Fig. 5). An HMD (FOVE) was connected with a computer running Unity (see Fig. 3). We first gave the participants the possibility to train navigating of a virtual robot in VR by conducting six trials, similar to the procedure used in [37]. When using the VR simulator for training of gaze control, the simulated “live” video stream was generated from the VR engine. The point-of-view and view height were similar to the real telerobot and so were the virtual robot’s maneuverability. Both the virtual and the real robot (see next section) were navigated in a room of 4 x 5 m.
Then they got to try a real telerobot. A Padbot with a 360-degree RICOH THETA S camera was controlled via a wireless network through a Raspberry Pi (see Fig. 2). We introduced two ways of driving the telerobots, namely one mode termed "invisible control" [38] and the other termed "way-point control" (see Fig. 6). The way-point control was similar to the control mode found in the Double 3 model, where users click on the floor to point at the next location for the robot to drive. In our case, way-points were marked by dwelling with gaze at the floor for a set dwell time.

Participants were asked to use their gaze to navigate the telerobot, wearing a FOVE HMD with build-in eye tracking. They were instructed to drive around a large table standing in the middle of the room and say hello to a person they would meet while driving around it. They did this twice - once using the "invisible control" which was a direct drive-where-you-look steering principle, and once using the way-point method (see Fig. 6). Afterwards, they answered questions about how confident they felt using gaze control and about their preference for one of the two gaze control methods.

Wizard of Oz navigation by speech and gaze
In the second experience prototyping session, we presented a Double robot (see Fig. 2) running the Double web application in a Chrome browser on a laptop (HP elitebook with a 15" monitor) connected to the robot via wireless network.

Two levels of assistance were experienced by the participants. First they could ask a helper to navigate the robot from the corner of the room to their position. Then we gave the participants an open instruction of..."Your next task is to move the robot outside the door. You can just tell the robot what you would like it to do..." This was done to provide a possibility to imagine the level of intelligence they would expect for this type of interaction. In response, they gave different kinds of commands, for instance, simply "move outside"; or "find the door", and then "go outside"; or just simple direction commands like "left, left, stop". Their speech commands were actually executed by the experimenter standing behind them with a joystick. Participants were not informed that their speech commands were manually performed by the experimenter.

Then they were asked to use gaze control. An eye tracker (Tobii Eye trackers 4C) was used for tracking the user's gaze point and shown on top of the live video stream from the Double robot (see Fig. 4). Again, steering of the Double was performed in a Wizard of Oz set-up by an experimenter standing behind the participant manually following the users gaze point with a joystick (Xbox 360 controller). The Double robot was driven from the experiment room through a corridor to the canteen, passing the reception of the care-home - a total distance of around 40 m.

We finally asked participants to try driving the telerobot around by simple speech commands - for those 4 participants who could do so. Four types of commands were possible: Left, right, stop and go. The experimenter executed their commands with his joystick behind their back. Three participants also tried to use hand control in the end.

After the sessions, we talked about their experiences and asked questions like: What would you like to use a telerobot for? How would you like to control the robot? How would you like to communicate with people you meet when driving a telerobot?

Observations
User expectations
All of our participants assumed the robots would be easy to use before they had tried them, and their initial expectations...
were that the telerobots would be fun like a computer game, exciting to drive just with the eyes, a new way of being me. One participant was a bit nervous to use the robots, and asked a helper to hold her hand during the test.

At this early stage, it was difficult for most of them to imagine a use case that they could relate to. Nevertheless, two of them envisioned that they could use telerobots to go shopping by talking with a care giver or a shop assistant via the robot, or just look at the price labels themselves. Once the participants had experienced the telerobots they got a much more clear vision of potential usage, mostly focusing on social interaction and possible inclusion in society. For instance, one participant envisioned that he could get a job in a warehouse driving robot trucks remotely or cleaning floors with a floor cleaning machine. Another participant expressed a wish for the telerobot to have an arm that could do things... Similar applications have been addressed in previous research papers also, for instance, the use of telepresence robots for shopping [36] and for remote work [7].

Control methods
Autonomous telepresence robots [4] might seem like a good solution for our target users. However, fully-autonomous systems may increase workload if users lack trust in the automated system, especially when the underlying mechanism of automation is not clear to them [19]. None of our users preferred to have the telerobot driven for them, either by other people nor by a fully-autonomous “intelligent” system. This is in line with previous observations [14] that our target users prefer to retain control authority. Therefore, semi-autonomous robots seem to be a more viable solution, using intelligent systems to assist in problematic situations only, and sensors to warn when obstacles are detected. However, this is a very complex issue that requires more research addressing differences in user needs, information transfer rates, differences in robot design, and differences in control principles - to name some of the factors that may impact performance and user experience.

When we asked our participants about their feelings of independence using the categories suggested by [22], (in our terms being able to control robots on your own, being able to maintain personal mobility in a remote place and being confident doing so), all of the participants stated that they had the same feeling of independence when using the different control methods (gaze, speech and hand). Since both the gaze and speech were sometimes controlled by another person in the two Wizard of Oz sessions this is not conclusive. However, when asked which input they liked most, all of the participants except one preferred gaze. In particular, the individual who was unable to give speech commands seemed very excited when using gaze interaction - she made a lot of gestures with a smile on her face. One participant was not able to get a sufficiently good gaze calibration with the HMD. The other four participants quickly gained confidence with this input method when driving in the VR simulator.

We had great expectations for the way-point control method. However, most of our participants found it very difficult to use and one of them did not finish driving around the table. Two of them mentioned that setting way-points forced them to look too far down and prevented them from attending to people and looking around in the room, c.f. Fig. 7. Hence, in future designs, we will focus on better ways to mark a way-point on the floor by gaze.

Figure 6: The two types of gaze control UI: invisible control (top), where the user drives in the direction of sight (indicated by a purple circle), and way-point control (bottom), where the user marks the next way-point by a dwell click on the floor (red circle).
them” and that the Double is a robot “... where I could be with my family and they could see me and I could see and talk to them. Interestingly, none of them commented on the option of having the 360 degree view provided by the Padbot, and only rarely did they make use of it. Instead of turning their head when wearing a HMD, they simply turned the robot to change their field of view.

HMDs hold potential advantages for telerobot interaction, for instance improved immersion and natural head control of field of view. However, two participants preferred to use the screen, rather than the HMD. One of them explained that he did not like the HMD because it would prevent him from being completely independent, since he would always need assistance to strap it on and adjust its position. Another participant could not achieve a sufficiently good gaze calibration with the HMD but he had no problems to get calibrated with a screen-based eye-tracker. Another important reason to consider a screen-based display instead of HMD’s is that it allows users to show their face on the remote screen and thus engage in facial communication. This is not possible when the face is covered by a HMD. Some of these problems may be due to current limitations of HMD’s and hopefully better solutions (e.g. smart glasses) will become available in the near future.

Presence and Experience
We frequently asked the drivers where are you now? and the answers to this question were always their remote position (e.g. in the corridor, reception, or canteen) - not their actual physical location (e.g. in the experiment room). This indicates a strong sense of telepresence. Our final - and maybe strongest - impression from this field study was the excitement we observed among all the participants and among their fellow residents and staff in the care-home (see Fig. 1). A staff member even hugged the telerobot when she saw who was driving it. The participant who could not speak, laughed and used her hand gestures when communicating with friend she met on her way. Overall, the participants' assessment of their experiences with telerobots were very positive, with typical statements like “...super fun, easy to use”, and even “proud of myself”.

Conclusion
Based on new insights provided by our target users, we are confident that several additional use-cases may be uncovered, and accessibility of human-telerobot interaction may be improved, if participants get an opportunity to familiarize with this technology and become involved in the design process. Their engagement fully confirmed Liu’s point [15] that providing ways to give and not simply receive help is an important way to support social interaction and integration.

Acknowledgements
Thanks to the people at Jonstrupvang care-home who participated in this study. The research has been supported by the China Scholarship Council and the Bevica Foundation. Thanks to Astrid Kofod Trudslev and Nils David Rasamoel for assisting with the study.

REFERENCES


