Remote Control of Mobile Robots through Human Eye Gaze: The Design and Evaluation of an Interface

Hemin Omer Latif*, SPIE Member, Nasser Sherkat and Ahmad Lotfi
The School of Science and Technology, Nottingham Trent University, United Kingdom

ABSTRACT

Controlling mobile robots remotely requires the operator to monitor the status of the robot through some sort of feedback. Assuming a vision based feedback system is used the operator is required to closely monitor the images while navigating the robot in real time. This will engage the eyes and the hands of the operator. Since the eyes are engaged in the monitoring task anyway, their gaze can be used to navigate the robot in order to free the hands of the operator. However, the challenge here lies in developing an interaction interface that enables an intuitive distinction to be made between monitoring and commanding. This paper presents a novel means of constructing a user interface to meet this challenge. A range of solutions are constructed by augmenting the visual feedback with command regions to investigate the extent to which a user can intuitively control the robot. An experimental platform comprising a mobile robot together with cameras and eye-gaze system is constructed. The design of the system allows control of the robot, control of onboard cameras and control of the interface through eye-gaze. A number of tasks are designed to evaluate the proposed solutions. This paper presents the design considerations and the results of the evaluation. Overall it is found that the proposed solutions provide effective means of successfully navigating the robot for a range of tasks.

Keywords: Human-Robot Interaction, Teleoperation, Mobile Robots Navigation, Eye Gaze Tracking, Man-Machine Interface, Operator Interface, Motion Control.

1. INTRODUCTION

One of the main challenges in developing any robotic platform is the required collaboration between a human being and a robotic agent. Although a number of researchers in the field of robotics aim at fully autonomous agents, many other researchers, in contrast, are more interested in controllable agents [1]. A wide range of these controllable agents require controlling from a remote location, which is mostly known as teleoperation [2]. Therefore, teleoperation remains a widely addressed challenge in a variety of robotic applications up to the present day and a significant amount of work has been done on developing different teleoperation interfaces [3].

Eye tracking, on the other hand, is entering its fourth era [4] and the technology behind it is developing continuously leading to the availability of a number of commercial products in the market at the present time [5-7]. Therefore, the domain of eye tracking applications is expanding rapidly ranging from diagnostic to interactive applications. As part of the development of interactive applications, inputs from the eyes have been used in developing a number of selective user interfaces for human-computer interaction [8]. Hence, few attempts in using eye tracking for human-robot interaction can be seen as well [9].

In many teleoperation applications the operator is required to continuously monitor the status of the robot through a feedback system mechanism. This engages the eyes of the operator in the monitoring task throughout the whole duration of the operation. Meanwhile, the hands of the operator need to be engaged, either partially or fully, in performing the commanding task using a variety of input devices. Therefore, reducing the degrees of engagement in any teleoperation task would give raise to many other opportunities for the operator to deal with.

Since the eyes of the operator are engaged in the monitoring task, inputs from their gaze can be used to aid in the commanding task. This will free the hands of the operator, either partially or fully, from this task. Navigating a mobile...
robot might be the most intuitive task to be implemented through human eyes because people mostly look where they want to go [10]. However, this does not mean that it is free of challenges because interpreting the inputs from the eyes into moving commands is not an easy aim to achieve. Therefore, in this paper a number of the challenges that meet teleoperation through eye gaze are addressed together with the design and the evaluation of an interface.

In this research teleoperation through eye gaze is referred to as TeleGaze. Therefore, this acronym is used throughout this paper. The paper is organized as follows: in section 2 related works are discussed. Section 3 introduces the TeleGaze with a number of design considerations. In section 4 task oriented evaluation of the interface is presented. Section 5 provides sources of limitations followed by pertinent conclusions in section 6.

2. RELATED WORKS

The nature of this work is related to the wide fields of both human-robot interaction and eye gaze tracking. Human-robot interaction has a very multidisciplinary nature and people from different background experience might contribute to the field [11]. Teleoperation therefore, as one of the hot topics in the field of human-robot interaction, gets much attention and contributions. Eye gaze tracking, on the other hand, as it was mentioned earlier has a wide range of applications and many researchers in the world are contributing to the field [4]. Since the focus of this work is remote control of mobile robots using human eyes, the following works are thought to be more related than others.

A controlling interface was developed in [12] in order to control a powered wheelchair through human eye gaze. The interface, which was divided into four command regions and five non-command regions, allowed the operator to drive the wheelchair using inputs from the eyes. Similarly and for the same purpose, earlier in [13] electrooculography was used to issue driving commands using three different interfaces. In both of these works, as the operator was sitting on the powered wheelchair, no feedback information was presented to the user through the interface. Therefore, the interface was totally used as one way communication channel from the user to the robot.

More towards teleoperation in that the control is achieved from a remote location, in [14] a robotic arm was controlled using an experimental eye tracking algorithm. A basic interface provided the operator with feedback through images from a camera mounted in the location. The interface was divided into a feedback region and a commanding region with a number of command buttons. Similarly to experiment an eye tracking algorithm, a robotic toy was controlled in [9]. However, in this work, the location of the eye ball within the eye socket was used. Because the robot did not react based on the point of the gaze on the interface, the interface was used only to provide feedback to the operator.

A more related work to this one is the work in [10] where inputs from the eyes were used to aid in the teleoperation process. A robotic platform was controlled from a remote station with the aid of eye tracking. However, similar to the approach of [9] the robot responded to the position of the eye ball within the eye socket rather than the point of the gaze on the interface.

Different to the works covered above, the focus of this work is the development of the interface rather than an eye tracking mechanism. Furthermore, considering an intuitive interpretation of the inputs from human eyes to navigate a mobile robot from a remote location shapes the aim of this research.

3. REMOTE CONTROL OF MOBILE ROBOTS THROUGH HUMAN EYE GAZE

To design an interface to control the operation of the robot remotely many practical issues need to be considered. As mentioned earlier an appropriate feedback mechanism is also required and this should be integrated into the control interface. Tacking into account all practical issues surrounding the interface design, different approaches are studied for the TeleGaze interface. In this section, the experimental platform and the detailed design considerations are explained.

3.1 The experimental platform:

To test and experiment the interface designed as part of this research, an experimental platform consists of a Wi-Fi enabled mobile robot, an eye tracking system and a teleoperation station is designed. Although one of the objectives of this research is to develop a platform-independent interface, a brief description of the experimental platform used to test the TeleGaze interface is covered here. The TeleGaze experimental platform is shown in Fig. 1.
The eye gaze tracking equipment (Fig. 1a) is a commercial product from the Applied Science Laboratories [5]. The system uses vision based eye gaze tracking algorithms to calculate the point of the gaze by transmitting infrared lights to the eye of the operator and capturing the reflections from the pupil and the cornea of the eye. A calibration process is required prior to obtaining the points of the gaze on the screen.

The mobile robot (Fig. 1c) is a modified wheelchair base equipped with a vision system consisting of two pan/tilt/zoom cameras [15]. The physical setup of the cameras is to provide the user with information from the remote location regarding the status of the robot. The Wi-Fi connectivity is used to establish connection with the teleoperation station.

The teleoperation station is a normal PC running the TeleGaze interface that the user interacts with. Images from the vision system and coordinates of the gaze from the eye gaze tracking system meet in the teleoperation station. Therefore, the TeleGaze interface (Fig. 1b) works as a meeting point between the two ends of the TeleGaze system.

3.2 TeleGaze design considerations:

One of the objectives of the TeleGaze interface developed in this research is to build a two way communication channel. Therefore, the entire display is divided into “action regions” and “no action regions”. The action regions are transparent regions on the display enabling the operator to issue controlling commands whilst monitoring the feedback. The transparency of the action regions helps in not obstructing the actual images from the operator. Hence, commanding as well as monitoring is achieved via the same space. Fig. 2 shows the TeleGaze interface.
Research shows that independently controlling a pan/tilt camera mounted on a mobile robot helps in the task of teleoperation [18]. Therefore, the operator was provided with action regions to control the onboard camera separately from the robot movements. Furthermore, controlling a mobile robot from a remote location requires awareness of any obstacles in the close surroundings of the robot. To gain this awareness, a downward looking camera is mounted on the robot. This helps the operator to get an idea of the distance between the robot and any obstacles in the area. A small view from the downward camera is located in the top-left corner of the interface.

The operator might need to inspect the scene more deeply and frequently. Also, operator’s eyes require rest from time to time. Therefore, the “no action regions” provide the operator with rest areas for the eyes and with the opportunity to inspect parts of the scene. However, to provide the user with a greater opportunity to inspect the scene, a more radical solution is used in the TeleGaze interface.

The operator is provided with the option of using the interface either to interact with the system or to inspect the scene only. The inspection mode of the TeleGaze allows the user to inspect the scene free from action regions. Switching between the interaction mode and the inspection mode is also achieved using inputs from the eyes. In the latter mode, the only action region exists is to switch back to the interaction mode and activate the rest of the action regions. Fig. 3 shows the inspection mode of the TeleGaze interface where only one action region can be seen.

![TeleGaze interface in inspection mode](image)

Fig. 3. The TeleGaze interface in the inspection mode. a) The layout of the interface in the inspection mode b) A snapshot of the interface in the inspection mode. [NB: Only one action region is available to switch back to the interaction mode.]

During the process of developing a TeleGaze interface, it was found that there is not a unique design for the same required level of interaction. Making a final decision, as far as the design is concerned, turned out to be difficult both in terms of personal preference and system performance. Different to previous related attempts in the field [12, 16, 19] this research extended the design possibilities to the actual users of the interface themselves. The design reported here is the results of an observational study carried out earlier and reported in [20] to investigate user preferences in terms of layout and functionality of the interface. Readers are referred to the author’s original publication for detailed results [20].

### 4. TASK ORIENTED EVALUATION AND USER SATISFACTION

To evaluate the TeleGaze interface, a task oriented evaluation was carried out. The purpose of this evaluation was to compare the performance of the TeleGaze with other modalities of interaction and to investigate user satisfactions. This section covers the participants, the task design, the evaluation metrics used in the evaluation and the results of both the evaluation and the user satisfaction questionnaire.

#### 4.1 Participants:

A group of ten participants volunteered to participate in the TeleGaze evaluation study. The age of the participants ranged from first year students to senior lecturers in the school. The participants were both males and females. Among these participants were people with good familiarity with using computers and people with less familiarity. This group...
was chosen to cover a wide range of potential users. All participants were given a brief verbal description of the idea of the evaluation study and then how the interface works.

4.2 The design of a navigational task:

Although a navigational task might differ from one application to another in many terms such as speed, required accuracy, complexity of the task and duration of the task, all these parameters are application dependent and liable to personal judgments. However, there are a number of main functions that are most likely to be required for any navigational tasks. The most likely ones are moving along a straight line, turning to the right and left and finally stopping at a designated point. In order to evaluate all the important functionalities of the system, the track shown in Fig. 4. is designed.

![](image)

Fig. 4. The diagram of the navigational task used in the evaluation.

To evaluate the TeleGaze interface in comparison with some other modes of interaction, the participants were asked to drive the robot along the track using a joystick, a computer mouse and the TeleGaze interface. The TeleGaze interface is designed such as a mouse pointer can be used to interact with the interface instead of inputs from the eye gaze.

One main character of any teleoperation is that the control is from a remote location. This by far makes the task more difficult to perform since the user is not interacting with the robot in real three dimensional spaces. Instead, the interaction will be in two dimensional spaces. In another word, the user is interacting only with feedback and different forms of data rather than real dimensions. Therefore, with all the three modes of interaction, the participants were allowed to look at the robot only through the TeleGaze interface.

A brief explanation of the task was given to each participant with only one minute prior to commencing the actual task to practice. This was to get the participants familiar with each interaction mode and how the robot will respond to the commands. Then the actual time and accuracy of task completion were recorded for each participant and for each interaction mode.

4.3 TeleGaze performance evaluation:

To the present day and due to the diversity of human-robot interaction applications there are no standard metrics to evaluate any newly developed interaction systems. However, there are a number of common metrics in any application domain that are most likely to be used to evaluate the developed system in that domain [21].

In many human-robot interaction applications addressing navigation and teleoperation, the most commonly used metrics are efficiency and effectiveness. These two common metrics are used in the evaluation of the TeleGaze interface too. Although, the definition of these metrics might vary from one application to another, for the purpose of this work efficiency was defined as the time to complete the task and effectiveness was defined as the accuracy to keep on track.

Since the objective of this evaluation is to compare three different interaction modalities, the absolute time of task completion is not a matter of concern as much as the relative time is. The chart shown in Fig. 5. shows the efficiency of all three modes of interaction in the form of the average time of task completion in seconds.
Many commercially available robotic platforms, ranging from toy robots to robots for military applications, are controlled by some sort of joysticks [22]. Based on this information it can be assumed that a joystick is one of the most convenient way to control a robot available so far. Hence, the average time to complete the task using a joystick is the target for the TeleGaze to meet. Based on this assumption and from the average time of task completion, the TeleGaze efficiency is 67% of the joystick efficiency. However, since this efficiency is calculated purely based on the average time then a very significant point should be considered.

The main requirement of the task was to follow the track which the operator could monitor from the view of the downward camera on the top-left corner (Fig. 2). Using the joystick, the operator could look at that view and issue commands for the whole duration of the task. In contrast, with the TeleGaze the user could not issue commands at the times when he/she was looking at the view. In another word, the operator did not issue controlling commands for the total duration of the task which led the overall time of task completion to increase by 43% in comparison with the joystick. This is less likely to happen in a different scenario such as moving from one point to another without checking the view from the downward camera. In this case, the operator would look through the main view and be able to issue commands as well which will increase the efficiency of the TeleGaze. The significant advantage of the TeleGaze should be mentioned here which is completing the navigational task with both hands absolutely free from the task.

Surprisingly, using the mouse instead of the eye gaze to interact with the TeleGaze interface, the efficiency is even higher than that of the joystick as the average time to complete the task is less (Fig. 5). In addition to that, using the mouse only required one hand of the operator to be engaged in the teleoperation while with the joystick both hands were engaged continuously.

To evaluate the effectiveness of the TeleGaze interface, the accuracy was defined as keeping the track between the wheels of the robot. Despite moving out of an absolutely straight line, if no wheels of the robot crossed over the track then the accuracy was calculated as 100%.

Overall, the task was repeated thirty times. Ten times with each interaction mode. Based on the above defined accuracy criterion, eight out of those ten times in each interaction mode the task was completed with full accuracy. Although the accuracy of the other two times was not zero, for a comparison purpose they can be neglected. Therefore, the accuracy of the TeleGaze interface can be calculated as 100% in comparison with that of the joystick and of the mouse.

**4.4 User satisfaction evaluation:**

To find out the level of the user satisfaction for the TeleGaze interface a user satisfaction questionnaire was designed and filled by the participants after the task was completed. All the questions, sixteen in total, were designed in favour of the interface and the participants were asked to show their agreement with each question at a scale of one to five where one was strongly disagree and five is strongly agree. Fig. 6 shows a representation of the average ratings of the questionnaire.
All ratings lie in the area between neutral and strongly agree which means that there is no negative impressions about the TeleGaze interface as an average. The minimum result obtained is for question No. 8 (Fig. 6) that was stating that the operator can perform more complex tasks than the one shown in Fig. 4 with the current TeleGaze interface. The majority of the participants did not agree with this statement but did not disagree either. This shows that the current interface required further development in order to enable the operator to perform more complex tasks with the TeleGaze interface.

The maximum rating average is for two questions (9 and 15) which stated that the user can get more out of the system with more training and experience and that the user would like the system to be developed further as it is worth it. These two findings again emphasize the need for further developments of the TeleGaze interface and user impressions on TeleGaze as a mean for human-robot interaction.

5. SOURCES OF LIMITATIONS

The second part of the questionnaire was to find out the sources of the limits from an actual user prospective or, in another word, why they could not show a better performance. Therefore, with three questions in this part the participants were asked to rate three possible reasons that were thought to limit the performance of the TeleGaze system. Fig. 7 shows the percentage of the likely sources of the limits of the TeleGaze based on the average ratings of the participants. It can be seen from the figure that the majority of the participants agreed with question 3 which stated that the performance of the eye gaze tracking system is the likely source of the limits of the TeleGaze interface. The other two questions were assuming that the design of the interface and the dwell time are the likely sources of the limits.
6. CONCLUSIONS

In this paper, TeleGaze is introduced as a means of controlling mobile robots from a remote location. A TeleGaze interface is developed that provides the operator with control over the navigation of a mobile robot together with onboard cameras as well as the interface itself through inputs from the eyes. A task oriented evaluation was carried out to evaluate the performance of the system and a user satisfaction questionnaire was filled out by the users.

From the work presented in this paper, it can be concluded that TeleGaze is likely to play a significant role in many teleoperation applications. Although TeleGaze did not achieve the same performance level of a joystick, the continuous development in the eye tracking technology will add significantly to the development of TeleGaze. Furthermore, the nature of the task designed to evaluate the interface was more likely to be implemented with a joystick rather than TeleGaze. A different task design might show the advantages of the TeleGaze over conventional input devices. Also improved eye gaze technology is set to result in better usability and carried the advantage of releasing operators’ hands. The authors intend to continue to work on the TeleGaze to meet the performance target of other teleoperation modalities with the advantage of less body engagement.

NB: Video demonstrations of TeleGaze can be accessed at www.hemin.co.uk/telegaze

REFERENCES


