

An Eye-gaze Tracking System for Teleoperation of a Mobile Robot

Carlos Carreto ^{1*}, Daniel Gêgo ¹, Luis Figueiredo ¹

¹ *Research Unit for Inland Development, Polytechnic of Guarda, Guarda, PORTUGAL*

*Corresponding Author: ccarreto@ipg.pt

Citation: Carreto, C., Gêgo, D. and Figueiredo, L. (2018). An Eye-gaze Tracking System for Teleoperation of a Mobile Robot. *Journal of Information Systems Engineering & Management*, 3(2), 16. <https://doi.org/10.20897/jisem.201816>

Published: April 10, 2018

ABSTRACT

Most telerobotic applications rely on a Human-Robot Interface that requires the operator to continuously monitor the state of the robot through visual feedback while uses manual input devices to send commands to control the navigation of the robot. Although this setup is present in many examples of telerobotic applications, it may not be suitable in situations when it is not possible or desirable to have manual input devices, or when the operator has a motor disability that does not allow the use of that type of input devices. Since the operator already uses his/her eyes in the monitoring task, an interface based on the inputs from their gaze could be used to teleoperate the robot. This paper presents a telerobotic platform that uses a user interface based on eye-gaze tracking that enables a user to control the navigation of a teleoperated mobile robot using only his/her eyes as inputs to the system. Details of the operation of the eye-gaze tracking system and the results of a task-oriented evaluation of the developed system are also included.

Keywords: mobile robot, teleoperation interface, human-robot interaction, eye-gaze tracking

INTRODUCTION

In a telerobotic system, a human operator controls the movements of a robot from a remote location. Some of these systems serve only the purpose of teleoperating the robots and others allow the human operators to have a sense of being on the remote location through telepresence. These robotic systems certainly have very interesting applications with enormous benefits to society (Minsky, 1980). Examples of interesting real-world applications are those in the area of Ambient Assisted Living (AAL) where teleoperated robots are starting to be used to remotely enabling the presence of their users and to provide companionship to elderly people (Amedeo et al., 2012). The different platforms combine a robotic mobile base with a remote video conference system for the communication between distributed teams-worker, relatives or health professionals and elderly people at home, or at healthcare facilities (Kyung et al., 2011; Tsui et al., 2011). There are also several mobile telepresence robots commercially available for the general public, such as Double Robotics (Double Robotics, n.d.), Giraff (Giraff, n.d.), QB Avatar (Anybots, n.d.), or R-Bot Synergy Swan (R.Bot, n.d.). These robots are relative cheap and considered as an important tool for inclusion.

Most telerobotic applications rely on a Human-Robot Interface (HRI) that requires the operator to continuously monitor the state of the robot through some sort of visual feedback and to use manual input devices to send commands to control the movements of the robot. This engages the eyes of the operator in the monitoring task and the hands in the controlling task throughout the whole duration of the teleoperation. Although this setup is present in many examples of telerobotic applications, it may not be suitable in various situations. Namely, when it is not possible or desirable to have manual input devices, or when the user has a motor disability that does not allow the use of manual input devices. Also, an effective hands-free teleoperation interface is very interesting to a



Figure 1. The telerobotic platform a) The mobile robot b) The teleoperation station

severely motor-disabled user, enabling him/her to move about with considerable independence, despite the disablement, through the use of a robotic wheelchair (Arai et al., 2011), or through a tele-presence robot (Tonin et al., 2011), for example.

Since the human operator already uses his/her eyes in the monitoring task, inputs from their gaze could be used to aid the controlling task by implementing a hands-free user interface based on eye-gaze tracking.

Eye-gaze tracking has a wide range of applications. There are several commercial available eye-gaze tracking systems such as myGaze (MyGaze, n.d.), SMI Gaze Systems (SMI Vision, n.d.), MagikEye (MagicKey, n.d.) and Eye Track 7 (Applied Science Laboratories, n.d.). Many researchers have contributed to the field by proposing many different applications (Duchowski, 2002), and some of the eye-gaze tracking systems mentioned have been used to develop HRI for telerobotics (Latif, et al. 2008; Jarvis, 2003).

This paper presents the implementation of a telerobotic platform to teleoperate a mobile robot using a novel HRI based on eye-gaze tracking. The aim of this research is to evaluate the feasibility to create a minimalist and low cost user interface based on eye-gaze tracking that enables a user to control the navigation of a teleoperated mobile robot using only his/her eyes as inputs to the system. This is a revisited version of a previous conference paper that was presented in CISTI 2017 (Gêgo et al., 2017).

The rest of the paper is organized as follows: section II presents the telerobotic platform used in the research, section III describes the details of the operation of the eye-gaze interface, section IV discusses the evaluation of the system, and last section VI presents the conclusions.

THE TELEROBOTICS PLATFORM

This section describes the hardware and the software of the experimental telerobotics platform used in the research.

The Hardware of the Platform

Figure 1 shows the experimental platform developed to test and experiment the eye-gaze interface that was developed.

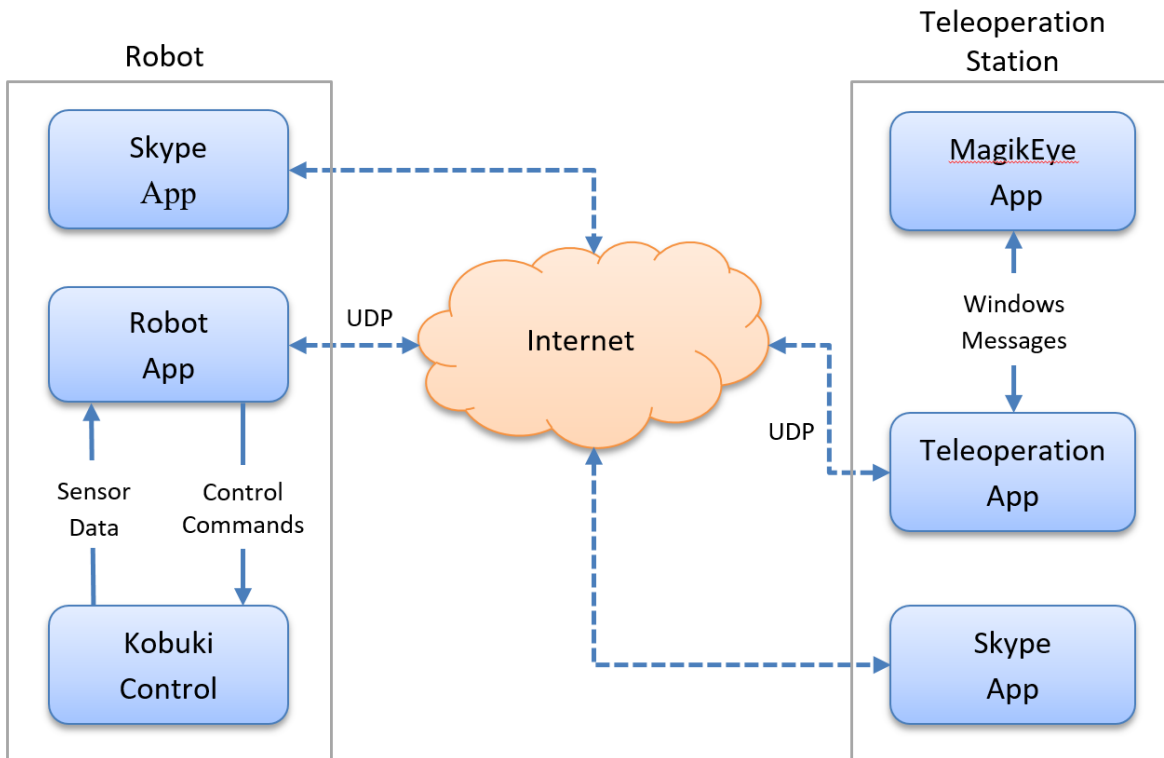


Figure 2. The different software applications

The experimental platform consists of a mobile robot at a remote location and a teleoperation station.

The mobile robot is an adapted version of a Turtlebot II robotic platform (Turtlebot II, n.d.) that comprises a mobile base called Kobuki, a netbook, and a Microsoft Kinect 3D camera sensor. The basic configuration was augmented with a mini-screen and a webcam, both mounted in a tower.

The teleoperation station is a normal laptop equipped with the non-intrusive eye-gaze tracking system MagikEye.

Both the teleoperation station and the mobile robot are Wi-Fi enabled and communicate through wireless Internet.

The Software of the Platform

Figure 2 shows the different software applications that comprise the platform.

The “Skype” application is present in both the robot and the control station and allows video and audio transmission between the two. This application works independently of the others. Although this application provides the means to implement a basic telepresence platform, since the robot also has a screen and a webcam with a microphone, the possibility was not considered in this research and the application was used solely to allow the remote monitoring from the teleoperation station.

The “Robot” application is responsible for controlling the robot navigation. It receives commands from the control station through User Datagram Protocol (UDP) messages that converts into robot specific commands that are sent to the Kobuki base, through a serial connection. The application also receives sensor data from the Kobuki base, through the same serial connection. Although the Kobuki base has several different sensors, in this project the application only used the data regarding the status of the bumpers that uses to suspend a certain control command, if the bumpers indicate the presence of an obstacle.

The “Kobuki Control” refers to the firmware of the robot base responsible for controlling its hardware.

The “MagikEye” application at the control station, implements the eye-gaze tracking system used to develop the user interface for the project.

The “Teleoperation” application implements the user interface to control the remote mobile robot. It receives the eye-gaze tracking data from the MagikEye application through Windows messages, and based on the user interaction, generates commands to control the navigation of the robot that are sent to the remote robot through UDP messages. The application also has the option to control of the robot using the keyboard and the mouse.



Figure 3. The layout of the three action regions

THE USER INTERFACE BASED ON EYE-GAZE TRACKING

This section describes the operation of the user interface and the eye-gaze tracking system.

The User Interface

The user interface of the Teleoperation application described in the last section was implemented with the objective to provide the operator with the capabilities for both controlling and monitoring. Therefore, the user interface must provide access to controlling commands as well as adequate feedback presentation of the remote images captured by the robot. Since the images are presented through the Skype application occupying the entire screen, the user interface was implemented as a transparent layer on top of the entire screen.

The transparency of this layer allows the user to issue commands to control the navigation of the robot whilst monitoring the images of the remote location captured by the robot. Commands are issued when the user looks at certain regions of the transparent layer. [Figure 3](#) shows the layout of the three regions that were defined.

The arrows shown in the figure are just a representation of the commands associated with each region and do not exist in the interface. By looking at one of the three regions, the user can issue a command to make the robot go forward, turn left, or turn right. Each command has associated a speed parameter that the user can also control. The speed value is set proportionally to the position where the user is looking inside a certain region. The area without arrows corresponds to a region not associated with any commands and provides the user with rest for the eyes and with the opportunity to inspect parts of the scene.

The Eye-gaze Tracking System

The eye-gaze tracking system used is called MagikEye and is a commercial product from the MagicKey company (MagicKey, n.d.). It is an alternative point-and-click interface system that allows the user to interact with a computer by computing his/her eye-gaze. The system is composed by a non-intrusive hardware component to capture images of the eyes of the user, and a software component that processes the images and calculates the point of the gaze in the computer screen. The user can move the head that does not interfere with the operation of the application. A calibration process is required prior to obtaining the points of the gaze on the screen. The system uses a very lightweight protocol based on Windows messages that allows the integration with other applications and is available for the Windows platform. The following describes the operation of the system with more detail.

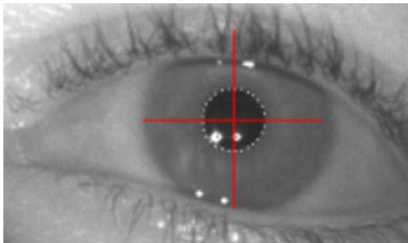
The hardware used by the system can be seen on [Figure 4](#) and is comprised of a high definition camera with a maximum spatial resolution of 1280x1024 pixels, and a color resolution of 8 bits. This camera uses a USB 2.0 interface and has the ability to acquire and transmit 25 frames per second at full resolution and can provide 60 frames per second with a spatial resolution of 1280x400. The camera uses type C lens, with 25mm, which allows the capture of the user's face images with high detail, crucial to the proper functioning of the system. The camera and lens are integrated with two infrared illuminators that emit at a wavelength of 840nm. The emitted infrared meets the EN62471 standard in terms of safety for the user.



Figure 4. The MagikEye hardware



a)



b)

Figure 5. a) Image captured by the camera with a resolution of (1280x400) b) detail of the right eye with the center of the pupil detected.

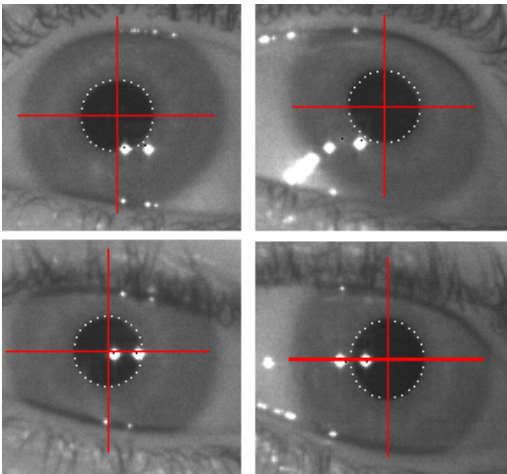


Figure 6. Images of the right eye looking at the four corners of a computer screen

The typical positioning of the camera is near the base of the computer screen facing the user's face, as shown in [Figure 1](#).

The eye-gaze tracking method is based on the detection of the dark pupil obtained by the combination of the angle of the infrared illumination with the lens of the camera, as shown in [Figure 5](#).

The first step of the method is to detect with high precision the center of the pupil. This is accomplished using a modified custom version of the Hough Transform (Duda et al., 1972), optimized for high speed. The algorithm can detect the pupil with high precision even when the pupil has a slightly oval shape and is partially covered by the eyelid.

The second step of the method consists in detecting the position of two white dots closest to the center of the eye that are the reflections of the two infrared illuminators. These reflections (white blobs) are used as a reference to calculate the direction of the user's eye in relation to the camera.

[Figure 6](#) shows a sequence of images that are obtained from the right eye when the user looks at the upper left corner, the upper right corner, the bottom left corner and the lower right corner of a computer screen. As shown in the images, the relative positions of the reflections (white blobs) from the center of the eye are related to the eye-gaze of the user.

The third step of the method maps these relative positions of the reflections into the total resolution of the computer screen, to estimate the point of the gaze.

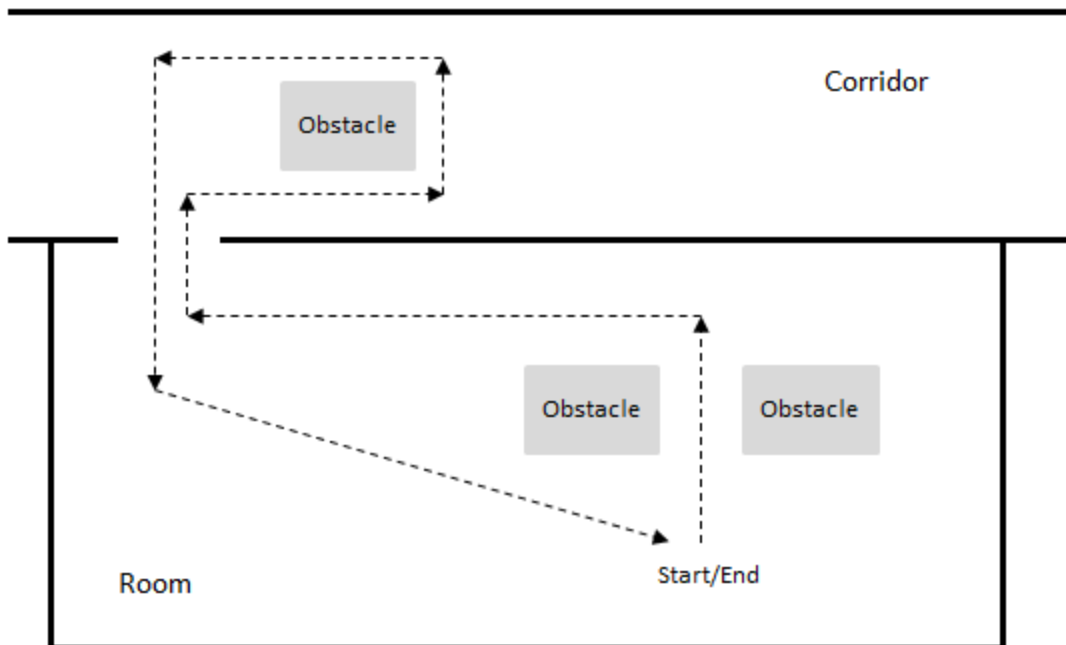


Figure 7. The diagram of the navigation task used in the evaluation

The maximum variation in terms of horizontal or vertical location of the reflections relative to the center of the eye when the user looks at the opposite boundaries of the computer screen does not exceed 40 pixels. This value has to be mapped to represent the full screen resolution. In the case of having a screen with a horizontal resolution of 1680 pixels, this means that the error rate is at least $1680/40 = 42$ pixels. This error rate is increased by the estimation error of the result of the center of the eye and the estimation error of the exact position of the infrared reflections. The following techniques have been implemented to minimize these errors and to increase the accuracy of the system and allow the user to place the cursor of the mouse on any pixel of the computer screen:

A sub-pixel resolution is used to calculate the center of the eye. The center is calculated in decimal terms. To do that, different weights are used to measure the probability of a particular pixel be effectively the center of the eye. Then, the 5 possible centers with highest probabilities are selected and a weighted average of those centers is calculated.

A similar technic is used to determine the center of the white blobs.

The algorithms are optimized to process the largest possible number of images. The system processes 60 frames per second. Since the mouse update is performed at a frequency of 20Hz, the calculations of the final position result from the average of 3 consecutive frames.

The two eyes are processed independently and the results are averaged to calculate the final point of gaze.

A time domain filter is used allowing the stabilization of the final point of gaze in small movements, without affecting large movements, such as when the eyes look rapidly from one side of the screen to the other.

SYSTEM EVALUATION

The performance of the eye-gaze interface was evaluated using a task-oriented evaluation, similar to the one proposed in (Latif et al., 2008). The evaluation had two goals. The first was to compare the performance of the user interface based on eye-gaze tracking with conventional modalities of interaction based on two manual input devices, the keyboard and the mouse. The second was to investigate the user's perception and opinion about the usability of the system.

A navigational task was designed and nine volunteers, aged between 21 and 45 years, performed the same task using all three different modes of interaction. After completing the task each participants filled out a questionnaire on the system usability. The participants were people with good familiarity with using computers but without experience in teleoperating mobile robots or in using user interfaces based on eye-gaze tracking. All participants were given a brief verbal description of the idea of the goals of evaluation study and how the interface works.

The aim of the task was to drive the robot along the track shown in [Figure 7](#).

The track had its beginning and end within a room and included passing through a door. The total length of the track was approximately 22 meters.

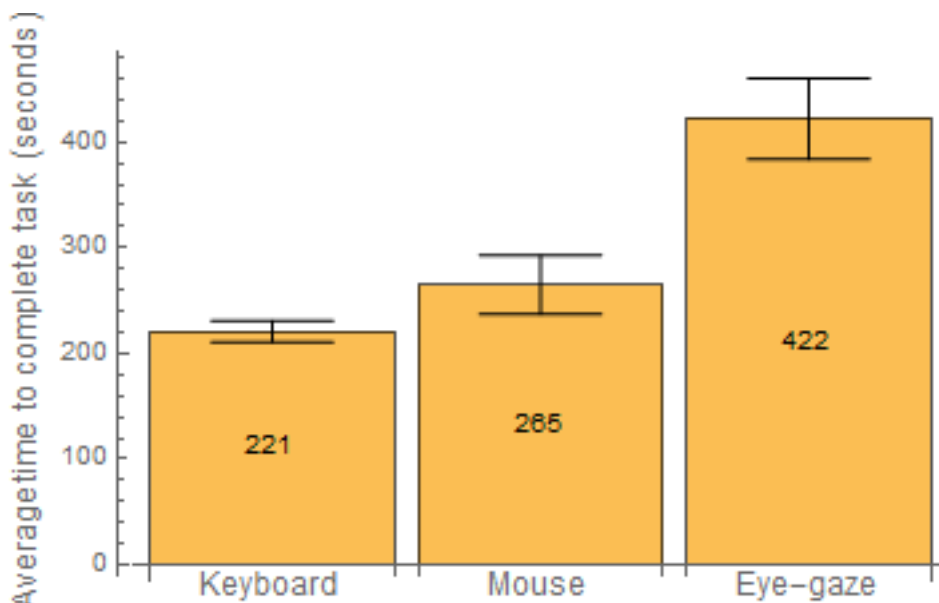


Figure 8. Average time to complete task for the three modes of interaction

Performance Evaluation

One metric commonly used to evaluate the performance of human-robot interaction applications, evolving teleoperation and navigation, is the efficiency. For the purpose of this work, efficiency was defined as the time to complete the navigational task. This time was measured for all three different modes of interaction for each participant starting from the start-point of the track and finishing by coming back to the same point.

A brief explanation of the task was given to each participant before starting the task, but the participants did not undergo any training session, even with the eye-gaze interface.

Before each participant started to execute the task, the MagikEye application was calibrated for that user. Then each participant executed the task three times using the three interaction modes. First the task was completed with the mouse interface, next with the eye-gaze interface, and finally with the keyboard interface.

The efficiency of the three modes of interaction is shown in the chart of [Figure 8](#) in the form of the average time of task completion in seconds. The error bars represent the standard error.

The results showed that the interface based on keyboard input came in the first position in terms of performance. The interface using inputs from the computer mouse came in the second place and finally the eye-gaze interface came in the third place. Nevertheless, the experiment proved the feasibility of the eye-gaze interface as a mean of HRI in teleoperation applications. Despite the relatively low performance of the eye-gaze interface, all participants managed to finish the navigational task using all the modes of interaction.

Usability Evaluation

To evaluate the user's perception and opinion about the usability of the eye-gaze interface, the participants were asked to complete a System Usability Scale (SUS) survey (Brooke, 1996).

The SUS is a simple, widely used 10-statement survey developed as a "quick-and-dirty" subjective measure of system usability. The tool asks users to rate their level of agreement or disagreement to the 10 statements (half worded positively and half negatively) about the system under evaluation. The level of agreement is given using a scale of one to five, where one is strongly disagree and five is strongly agree.

The evaluation of the usability of the system performed by the participants revealed an average SUS score of 70.8 (on a scale of 0 to 100) that is considered to represent a good usability (Brooke, 1996; Nielsen, 1994).

The scores of the individual statements of the SUS survey can be grouped to obtain a set of quality components that characterize the system (Nielsen, 1994).

The chart shown in [Figure 9](#) presents the average scores for the different quality components considered (on a scale of 0 to 4).

It can be seen from the chart that all the quality components obtained a positive score. However, none of the quality components stands in relation to others, and overall, the scores are not high. These results, together with the SUS score, once again prove the feasibility of the system, but emphasize the need for further developments of the system.

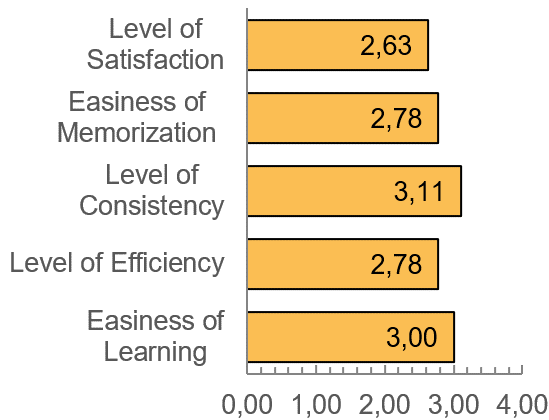


Figure 9. Average scores of the quality components obtained from the SUS survey

CONCLUSIONS

A telerobotic platform that uses a user interface based on eye-gaze tracking was presented. Details of the operation of the eye-gaze tracking system were also included. The system was evaluated using a task-oriented evaluation and the results permit to conclude that the proposed interface is a feasible option as a mean of HRI in teleoperation applications.

The evaluation results and the observations during the experiments identify some possible improvements that can be considered for future work.

The inclusion of a pan-and-tilt mechanism to control the webcam of the robot could improve the overall performance of the system, since the operator will have more flexibility during monitoring and control tasks and could also provide the ability to implement telepresence.

A feature asked by several participants in the evaluation, in relation to the eye-gaze interface, was the possibility to have some kind of visual feedback identifying which region/command is active. This makes sense, because the other modes of manual interaction have an intrinsic feedback given by the tactile and visual senses, but this feedback is lost when using the eye-gaze interface implemented with the transparent layer.

Some functionality could be added to improve the steering control of the robot. For example, would be interesting to be able to drive the robot on a curvature line.

ACKNOWLEDGEMENT

Publication backed by the PEst-OE/EGE/UI4056/2014 UDI/IPG project, funded by the Foundation for Science and Technology.

REFERENCES

- Amedeo, D., Cortellessa, G., Orlandini, A. and Tiberio, L. (2012). Addressing the Long-term Evaluation of a Telepresence Robot for the Elderly. In F. Joaquim and L. N. Ana (eds.) *ICAART* (1), pp. 652-663.
- Anybots (n.d.) QB Avatar. Available at: <https://www.anybots.com>
- Applied Science Laboratories (n.d.) Eye Track 7. Available at: <http://www.asleyetracking.com>
- Arai, K., and Mardiyanto R. (2011). Eyes Based Electric Wheel Chair Control System-I (eye) can control Electric Wheel Chair. *International Journal of Advanced Computer Science and Applications*, 2(12). <https://doi.org/10.14569/IJACSA.2011.021215>
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. *Usability evaluation in industry*, Taylor & Francis, 189–194.
- Double Robotics (n.d.) Double. Available at: <http://www.doublerobotics.com>
- Duchowski, A. (2002). A Breadth-First Survey of Eye Tracking Applications. *Behavior Research Methods, Instruments & Computers*, 34(4), 455-470. <https://doi.org/10.3758/BF03195475>
- Duda, R. O. and Hart, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Communications of the ACM*, 15, 11-15. <https://doi.org/10.1145/361237.361242>

- Gêgo, D., Carreto, C. and Figueiredo, L. (2017). Teleoperation of a mobile robot based on eye-gaze tracking. In *Proceedings of the 12th Iberian Conference on Information Systems and Technologies (CISTI)*, IEEE, 2299-2304. <https://doi.org/10.23919/CISTI.2017.7975673>
- Giraff (n.d.) Giraff. Available at: <http://www.giraff.org>
- Jarvis, R. (2003). A go where you look tele-autonomous rough terrain mobile robot. *Experimental Robotics VIII*. Springer Berlin Heidelberg, 624-633. https://doi.org/10.1007/3-540-36268-1_57
- Kyung, L. M. and Takayama, L. (2011). Now, I have a body: Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 33-42.
- Latif, H. O., Sherkat, N. and Lotfi, A. (2008). Remote Control of Mobile Robots through Human Eye Gaze: The Design and Evaluation of an Interface. In *Proceedings of SPIE Europe Security and Defence 2008*, 71120X-71120X-9.
- MagicKey (n.d.) MagikEye. Available at: <http://www.magickey.ipg.pt/MagicEye.aspx>
- Minsky, M. (1980). Telepresence: A manifesto. *Omni*, 44-52.
- MyGaze (n.d.) myGaze eye tracker. Available at: <http://www.mygaze.com>
- Nielsen, J. (1994). *Usability engineering*. Elsevier.
- R.Bot (n.d.) R.BOT Synergy Swan. Available at: <http://promo.rbot.ru>
- SMI Vision (n.d.) SMI Gaze & Eye Tracking Systems. Available at: <http://www.smivision.com>
- Tonin, L., Carlson, T., Leeb, R. and del Millan, J. R. (2011). Brain-controlled telepresence robot by motor-disabled people. In *Proceedings of the IEEE International Conference of the Engineering in Medicine and Biology Society*, pp. 4227-4230. <https://doi.org/10.1109/IEMBS.2011.6091049>
- Tsui, K., Norton, A., Brooks, D., Yanco, H. and Kontak, D. (2011). Designing telepresence robot systems for use by people with special needs. In *Proceedings of International Symposium on Quality of Life Technologies*, 2.
- Turtlebot II (n.d.) Turtlebot II Documentation. Available at: <http://wiki.ros.org/Robots/TurtleBot>