Distributed Eye Tracking Network for Conveying Gaze of Remote Users in a Robotic Telepresence Scenario

David Rozado
ICT Centre - CSIRO

Belinda Ward
ICT Centre - CSIRO

Fred Pauling
ICT Centre - CSIRO

We propose a distributed network of eye trackers to monitor the gaze behaviour of multiple users of a remote telepresence system. Telepresence systems allow their users to sense a remote location as if they were physically present, and may also project images and audio from tele-present subjects into the physical environment for example via a robot. However, a limitation of such telepresence systems is that for humans around the robot it is difficult to determine where the remote users are focusing their attention. This contrasts with physical presence where body language and other subtle cues provide a hint of gaze points. We use as an illustrative application a robot which facilitates remote museum visits. A panoramic camera on the robot provides remote users with an immersive view within the gallery, and a human museum guide directs the robot around the museum while interacting with the remote visitors. Eye trackers are used to collect information on where remote users are looking. This information is then superimposed into a condensed representation of the panoramic image. The visualization of remote visitors' gaze behaviour provides useful feedback to the museum tour guide about their audiences areas of attention.

Keywords: eye tracking, gaze tracking, telepresence, eye tracking network, gaze aware interfaces, pervasive computing, HCI, gaze responsive interfaces, human computer interaction

Introduction

Telepresence technology allows a person to feel and/or to have an effect as if they were present at a place other than their real location. Telepresence technology requires that the remote user is provided with sensory stimuli conveying information about the distant location in order to re-create the feeling of being in the distant location. This is usually done via visual and audio sensors that capture and transmit the information about an environment to the remote users. It is paramount also to provide people in the robot's location with information about remote visitors' behavior. The option of interacting with the distant location through output effectors can be given to the remote users. This can be achieved by sensing, transmitting and duplicating in the remote location the user's position, movements, actions, or voice. Hence, in realistic telepresence systems, information flows in both directions between the distant locations. The range of possible usages of Telepresence systems scenarios has increased considerably in the last years thanks to advances in robotics and telecommunication technologies. Among its many advantages, telepresence systems enable collaboration independent of geographic location and provide considerable savings in terms of time, travel costs and environmental impact for numerous usage scenarios.

A robotic telepresence system is a particular subtype of telepresence technology that allows a subject to move virtually through a distant location by remotely controlling a wheeled robot which is usually equipped with a camera, a microphone and a loudspeaker. A screen in the mobile robot can optionally display live video of the subject's face allowing the remote user to interact with subjects around the robot.

In this work, we augment an existing telepresence system that allows remote subjects to visit a museum using a semi autonomous mobile robot, an immersive web interface and an educator or museum tour guide near the robot which guides the tour. The system is designed to allow remote visitors to feel as they are on an actual tour in the museum by employing a web browser viewer of the panoramic video streamed by the robot. The innovative part of the system is the usage of a network of eye trackers that monitor the gaze

This work is an Australian Government funded initiative. The work was carried out at CSIRO - ICT Centre. Corresponding author: david.rozado@csiro.au

behavior of the remote museum visitors and transmits their gaze data to the mobile robot. The robot then displays on its on-board screen the gaze behavior of the remote users in order to provide the educator at the museum with a signal that conveys what is attracting the gaze (and hence the attention) of the remote students.

The point of regard (PoR) of a computer user on the screen can be estimated using video oculography gaze tracking. The usage of eye gaze as a pointing or control mechanism is well documented in the literature(Duchowski, ,). Gaze tracking has found a niche application to study how computer users interact with content on a computer screen (Rozado, Rodriguez, & Varona, , ,). The gaze data accumulated during a gaze tracking session can be visualized offline as a heat map data structure. A heat map captures the accumulated degree of attention that a given user places at different points on the screen during a gaze tracking trial. Hence, it is well-established that the direction of gaze reflects the user's focus of attention on the screen (Zhai,), and hence, it provides a hint about intention. In this work, we leverage on these well known facts about eye tracking practice and extend it to a network of eye trackers scenario used in combination to convey the gaze behavior of several remote museum visitors in a remote telepresence systems. Since the gaze coordinates in a remote eye tracker only reflect the PoR on a given plane, we also monitor the field of view within the panoramic video viewer. The combination of the field of view and gaze coordinates maps to a unique point in 3D space that identifies the position within the panoramic sphere video grabbed by the robot that is capturing the user's attention.

The concept of telepresence systems is not new. The work from (Minsky,) coined the term *telepresence* to describe systems that would transform work, manufacturing, energy production and medicine by allowing remote workers to carry out tasks in remote locations through the usage of robots, sensors and telecommunication infrastructure. As a pioneer in the field, Minsky was the first to propose the idea of being able to work in another country or planet through remote telepresence.

One of the first and most popular applications of telepresence is videoconferencing (Duncanson & Williams,). The review from (Egido,) highlighted some of the limitations of the technology at the time it was first proposed, most of which have already been overcome (Lawson, Comber, Gage, & Cullum-Hanshaw,).

Another branch of telepresence that has gathered a lot of attention in the research literature is telesurgery, also known as *surgery at a distance* (Green, Hill, Jensen, & Shah,). The technique allows surgeons to operate robotically on a patient that is located in different geographical coordinates. Among the many benefits of the telesurgery concept, the technology has the potential to revolutionize healthcare delivery by bringing surgical attention to previously inaccessible settings (Holt, Zaidi, Abramson, & Somogyi,).

More recently, telepresence robots that are controlled remotely and allow a human to explore remote geographical locations have been proposed (Schultz, Nakajima, & Nomura,) and have received a considerable amount of research attention (Tsui, Desai, Yanco, & Uhlik,).

An important feature of robotic telepresence systems is their ability to realistically capture their surrounding environment in order to convey a realistic representation of their location to the remote users. Panoramic images can enhance modern robotic telepresence systems. Panoramic cameras provide the functionality of capturing and transmitting a high quality visual representation of the robot's surrounding environment. The work from (Gledhill, Tian, Taylor, & Clarke,) provides a good overview of panoramic imaging technologies summarizing and comparing some of the methods used to carry out the different steps involved in the process: capturing, image processing, image stitching, 3-D reconstruction, rendering and visualization.

Gaze tracking has been used extensively to monitor the gaze behavior of subjects in a variety of circumstances and environments (Duchowski, , ,). However, to our knowledge, there is no literature available on the utility of using gaze tracking to transmit the gaze behavior of remote users to the robot's environment in a telepresence system. The amount of literature on using networks of eye trackers is also very limited due to the high cost of the devices until very recently.

In Summary, in this work, we augment a robotic telepresence system that allows live interaction between remote students and an educator in a museum by providing the museum educator with a convenient visualization of the gaze dynamics of the remote museum visitors. Multiple remote museum visitors connect via a single robot to the museum telepresence tour, but each remote visitor controls their own view within the gallery. Remote students are given a tour of the museum by the educator and can access additional digital content about objects on display through their panoramic viewers augmenting thereby the tour experience. A network of eye tracking software in the remote students' computers monitors and transmit their gaze behavior to the robot. This gaze behavior data is displayed in the robot screen to provide the museum tour guide with a feedback signal about where the remote visitors are paying attention to. This information closes the interaction loop between the educator in the museum and the remote students and offsets some of the nonverbal communication limitations of telepresence systems by signaling to the museum tour guide where the remote museum visitors are paying attention to at any given instant in time.

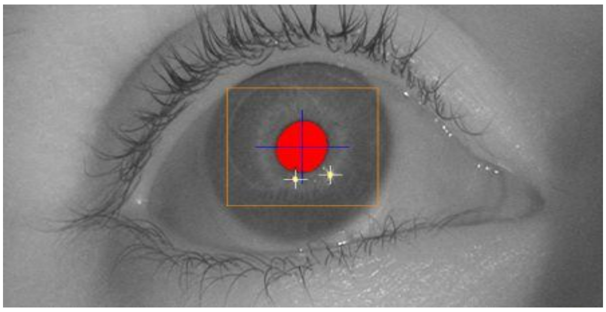

Figure 1. **The Open Source ITU Gaze Tracker Tracking One Eye.** The features being tracked in the image are the pupil center and two corneal reflections. These features are used by the gaze estimation algorithms to determine the PoR of the user on the screen.

Figure 2. **Prototype of the Telepresence Robot.** The figure displays the main components of the robot, a panoramic camera on its top, a screen in the middle and a mobility unit in the bottom.

Method

The entire gaze monitoring and robotic telepresence system comprises several subparts that we describe below.

Eye-tracking system. Eyes are used by humans to obtain information about the surroundings and to communicate information. When something attracts our attention, we position our gaze on it, thus performing a *fixation*. A fixation usually has a duration of at least 150 milliseconds (ms). The fast eye movements that occur between fixations are known as *saccades*, and they are used to reposition the eye so that the object of interest is projected onto the fovea. The direction of gaze thus reflects the focus of *attention* and also provides an indirect hint for *intention*(Velichkovsky & Hansen,).

A video-based gaze tracking system seeks to find where a person is looking, i.e. the Point of Regard (PoR), using images obtained from the eye by one or more cameras. Most systems employ infrared illumination that is invisible to the human eye and hence it is not distracting for the user. Infrared light improves image contrast and produces a reflection on the cornea, known as corneal reflection or glint. Eye features such as the corneal reflections and the center of the pupil/iris can be used to estimate the PoR. Figure 1 shows a screenshot of an eye being tracked by the open-source ITU Gaze Tracker (San Agustin et al., ,). In this case, the center of the pupil and two corneal reflections are the features being tracked.

In most eye tracking studies, the PoR of the user is employed to generate a heat map of attention areas in the 2D scene plane being observed or as a pointing device that substitutes the mouse in gaze interaction paradigms. For this work, eye tracking was used behind the scenes just to capture the gaze behavior of the remote museum visitors. That data was then transmitted to the robot in order for it to display the aggregated gaze behavior of several users as a feedback signal to the museum educator.

A set of 3 Tobii Rex eye trackers and 1 Tobii X1 light eye tracker are used for the experimental part. Both systems track gaze at 30 frames per second (fps) with an average gaze estimation accuracy of 0.5 degrees at 60cm from the screen, (Tobii Technology AB,). A simple exponential smoothing filter, (Kumar, Paepcke, & Winograd,), of the raw gaze tracking data is carried out to prevent jitter in the visualization of the remote visitor's gaze.

Robot

We use a commercially available Research-PatrolBot robot from Adept-MobileRobots to provide mobility and power to the system. The robot base carries a shoulder-height Point Grey Ladybug3 panoramic camera, a touch-screen for controlling the robot and the tour, a web-cam for teleconferencing between the guide and remote visitors, and a computer for running the user interface and video-processing. All data communication is over 802.11n Wi-Fi. See Figure 2 to observe an early prototype of the system.

The robot moves at walking speed within any wheelchair accessible space, and is able to navigate safely and autonomously between pre-defined locations in the museum, as commanded by the guide. The robot navigates to a commanded destination by first planning an efficient route within a pre-made map of the museum, and then following that route whilst reactively avoiding unmapped obstacles (such as people and other objects), that are detected by the on-board laser-scanner. The robot maintains an accurate estimate

Figure 3. **Panoramic Camera.** The figure displays the LadyBug3 panoramic camera on top of the robot that uses several cameras facing different directions to create a panoramic image of the robot's surroundings.

of its location in the museum by tracking its movements using wheel-odometry, and correcting for drift by matching features in the current laser scan with features in the pre-made map.

Museum Tour Guide

The museum tour guide wears a wireless head-set microphone to ensure that he can be heard by the remote students while carrying out a museum tour. The museum tour guide can see who is online and which students have questions via the display on the front of the robot. With the proposed enhancement we are describing in this work, the museum guide is also able to monitor the gaze of the remote visitors on the screen of the robot.

360° Video Stream and Panoramic Viewer

A panorama is a single wide-angle image of the environment around the camera (Gledhill et al.,). The most realistic types capture their environment on the horizontal plane (360°), and 180° in the vertical field of view. There are different ways to capture a panorama: single, rotated about its optical center, single omnidirectional camera (using multiple cameras facing in different directions) or using a stereo panoramic camera from which scene information can be extracted.

The 360° panoramic camera (Figure 3) used in our telepresence system was mounted on top of the robot and captured a high-resolution omnidirectional image of the robot's environment that was further streamed to the browser clients in the remote visitor's computers.

The panoramic camera used in this work uses six 2.0MP cameras to capture a 360° by 140° field of view.

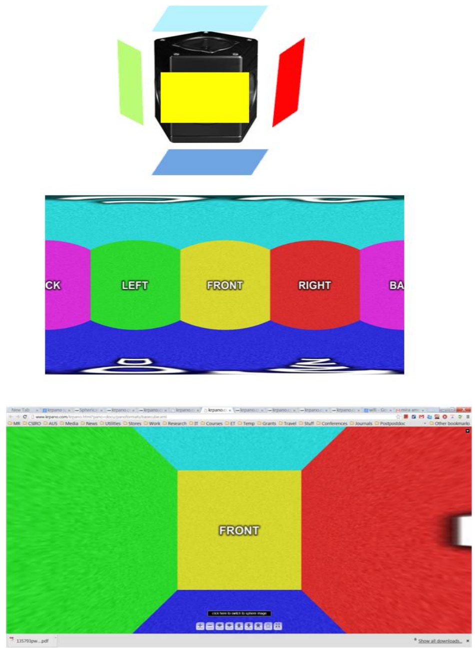

Figure 4. **Panoramic Viewer.** A panoramic camera (Figure top) is used to capture different, but overlapping, views of the robot environment that are combined into a panoramic video stream, and sent to each client, where it is wrapped onto a virtual sphere. The middle of the figure shows a one image 360x180° projection of the spherical panorama being captured[2]. This allows the remote users to use a panoramic viewer such as Krpano that renders the different views into an immersive viewing environment (Figure bottom). Note that in our particular museum robot project, the bottom view displayed in this Figure was not captured since the panoramic camera lacks that capability. The ladybug instead also possess 6 horizontal cameras generating 6 views rather than the 4 illustrated in the Figure. The reason for this discordance is to simplify the conceptual clarification of the system overall functionality.

The six individual images are stitched onto a fixed-radius sphere, and then mapped onto a rectangular image using an equirectangular projection. A continuous sequence of these panoramic images is encoded into a high-definition video stream and sent to each client, where it is decoded and mapped onto a spherical canvas for immersive panoramic viewing using a virtual camera (see Figures 4 and 5).

We use Krpano[3] as the panoramic image viewer used in the remote computers. Krpano is a light and very flexible high-performance viewer for all kind of panoramic images, videos and interactive virtual tours. The viewer is available as a Flash or HTML5 application and is designed for usage inside a Browser on Desktop and Mobile computer devices. For our specific

[3] `http://krpano.com/`

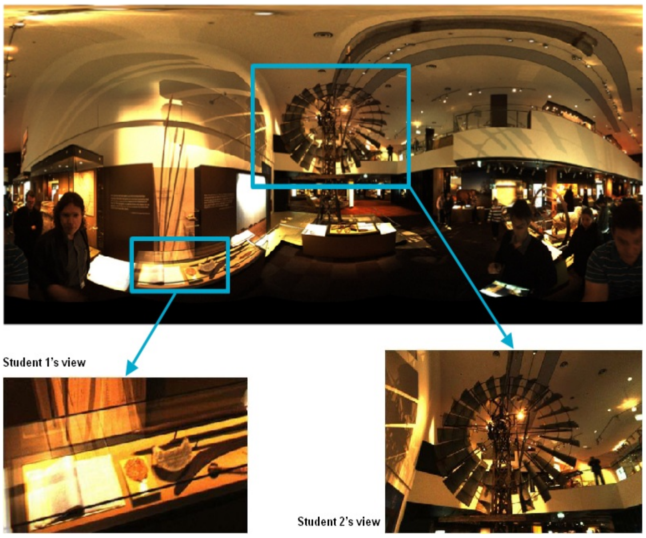

Figure 5. **Creation of an Immersive Telepresence System.** Images from the 360° camera on top of the robot are combined to form a high-resolution 2D omnidirectional image representing the environment around the robot (top image). Individual users can look at different subareas of the panoramic video stream (bottom images).

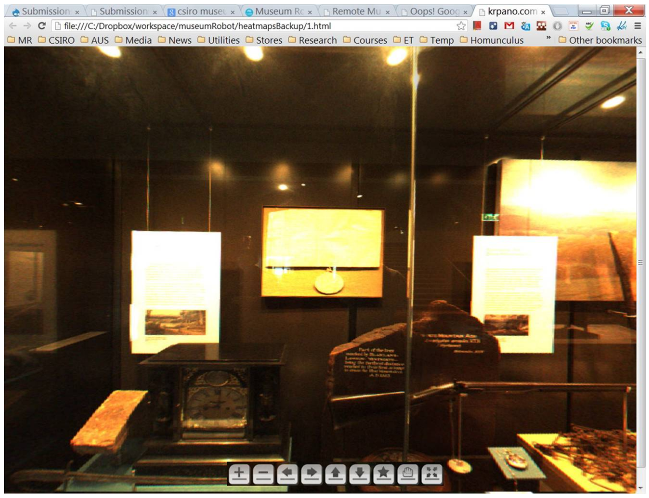

Figure 6. **Panoramic Viewer in the Remote Museum Visitors' Computers.** Remote browser interface client through which student can look around the museum with freedom to look at different planes within the 360° field of view environment.

telepresence museum visit, each remote students can then independently "look around" the gallery using the panoramic viewer within their browser as shown in Figure 6.

Remote Users' Client Computers

Each museum visitor computer is equipped with a low cost eye tracker capturing gaze frames at 30Hz. The eye tracker provides the X, Y coordinates of gaze on the screen plane. Since the remote museum visitor can also control their horizontal and vertical field of view within the panoramic viewer, these parameters are also monitored by a javascript script running behind the scenes in the browser. The horizontal and vertical view points refer to the views that determine the plane being displayed on screen at any given time from the panoramic image. The combine parameters of horizontal and vertical field of view and the X, Y gaze coordinates are transformed in Krpano into spherical coordinates to obtain the gaze point of regard of the user in the spherical space. These four parameters define a unique point in the virtual 3D space where the museum visitor is looking at and those parameters are then send back to the robot, see Figure 7.

Images from the panoramic camera on top of the robot currently being captured are combined to form a condensed 2D omni-directional image. The points of regard of several remote museum visitors are projected into this image to provide feedback to the educator about where the students are paying attention to in the scene, see Figure 7 and 8.

Each remote student computer uses a server-client architecture to gather the data being monitored: gaze coordinates and horizontal and vertical fields of view. A javascript client program embedded in the browser display monitors the panoramic viewer field of view and stream this data through a websocket to a local server that also receives gaze coordinates from a gaze tracker client. This server streams the horizontal and vertical field of view and the gaze coordinates to a further server located in the robot that dispatches all the gaze data streams from different browser clients to a handler that displays this information on top of a condensed 2D representation of the omni-directional video stream currently being captured by the robot. The TCP protocol is used to issue commands to the eye trackers such as start calibration, calculate results from calibration, start tracking, stop tracking. Data transfers between the clients and the robot is carried out using the UDP protocol for its simplicity and due to the fact that missing a gaze frame occasionally is not critically important for the visualization purposes. This is because the system can afford to lose some gaze data packets since this would not impact the visualization of the gaze behavior of remote museum visitors.

Application

The system described here is currently a work in progress in the prototyping stage. We are using a network of 4 eye trackers to monitor the gaze behavior of

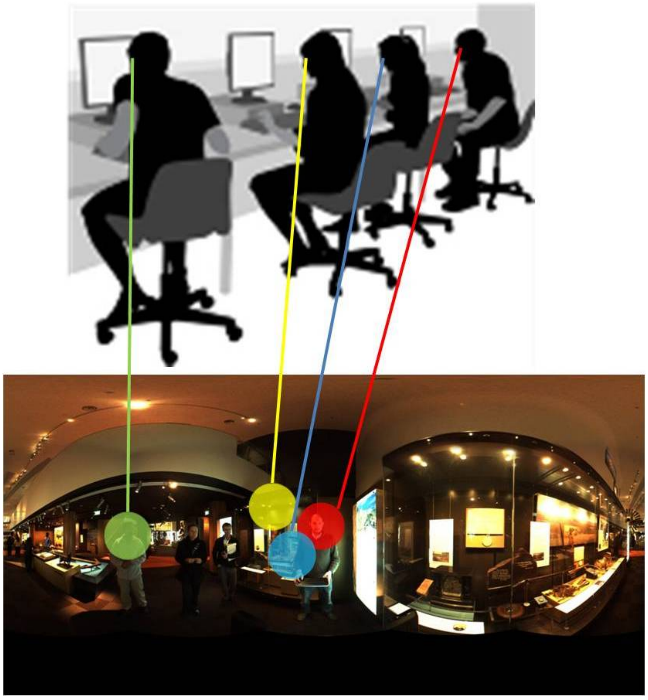

Figure 7. **Eye Tracking Network.** A network of eye trackers on the remote museum visitor's computers captures the gaze behavior of the students. That data is combined with the specific field of view plane within the panoramic video stream of each student and send back to the robot. The robot then combines the gaze dynamics of each remote visitor with a condensed 2D representation of the omni-directional panoramic image.

Figure 8. **Visualization of Remote Museum Visitors' Gaze Dynamics on the Robot.** The aggregated gaze behavior of remote students superimposed into the condensed 2D omni-directional video stream is shown in the robot display so the tour guide can obtain feedback about what areas in the environment around the robot are attracting the visitor's attention.

4 distinct users and stream that data to another computer acting as the robot display where a condensed 2D representation of the omni-directional video stream captured by the panoramic camera is shown.

This work is embedded within a wider project that seeks to integrate the capabilities described here into an actual robotic telepresence system being used by high school students (ages 10-16) to carry out remote museum visits. One of the main features of the system is the ability of the remote students are able to customize the specific view from the panoramic video stream that they want to take at any particular time point during the museum tour.

We have developed scripts in the browser client that stream the horizontal and vertical field of view parameters active on each user's client computer and their gaze coordinates to a server intended to be placed in the robot. The robot's server receives the vector features from each connected client, parses the data and dispatches commands to the display unit to render the gaze coordinates of each user being monitored. This is accomplished by representing each user's gaze as a disc over a condensed 2D representation of the panoramic image. In order to gather a sense about how the remote gaze monitoring system in the robotic telepresence scenario works, the interested reader is encouraged to take a look at the manuscript's associated video at `http://youtu.be/gB89jT2_3oA`. The video provides a good visual overview of a preliminary version of the system at work and how it can be used to monitor the gaze behavior of remote museum visitors. The video also shows how the gaze data of the remote users is superimposed in the condensed 2D representation of the panoramic image being captured by the robot, the key feature of the system described system. The video shows how the museum tour guide actually moves around the museum and how she is closely followed by the semi-autonomous robot while interacting with the students. The main contribution of the system described here to the telepresence scenario is to provide the educator or museum tour guide with a sense of what around her catches the remote students' attention.

An early protototype of the system described here has already been built and tested as shown in the manuscripts accompanying video, reaching the proof of concept status. Future work will strive to carry out user studies with young students from several national high schools in order to study their gaze dynamics during a remote museum visit by following their gaze patterns with a larger network of low cost Tobii Rex eye trackers.

Discussion and Conclusion

Advancements in robotics and communication technologies have propelled the emergence of telepresence technology systems that allow users to perceive and/or interact with a distant region from a remote location. The robotic telepresence system to visit a museum described in this work is one of several examples in the growing field of robotic telepresence. Telepresence system offer many opportunities by for instance facilitating different regions of the world to export specialized skills anywhere at any time. The technology offers numerous additional advantages in terms of time and cost benefits by reducing transportation costs, energy savings and commuting time. Last but not least, telepresence eliminates numerous chemical and physical health hazards of physical presence in dangerous locations.

One particular are where telepresence systems can be particularly advantageous is for education scenarios such as the system presented in this work that allows potential museum visitors such as high-school students to visit institution with relevant education material such as museums from distant geographical locations. National institutions such as museums have a responsibility to reach regional and remote potential visitors who are often unable to visit them physically due to factors such as cost, time and distance. A mobile telepresence system provides an opportunity for regional and remote potential museum visitors to participate in tours on the museums and other educationally relevant institutions.

A key challenge for robotic telepresence systems for visiting distant locations through a robot will be how to leverage the new opportunities provided by telepresence technology to engage users of the system by providing highly realistic user experiences. In the telepresence system presented here, the panoramic viewer in the browser's client is highly immersive and strives to make remote students feel as if they are present in the gallery through the panoramic video stream. The systems is also highly interactive allowing remote students to select their own view within the panoramic image and to engage in discussions with the educator and other students on the tour. The strive to make the tour a realistic experience does not preclude augmenting the tour experience by superimposing in key objects of the museum meta digital content layers for which a given student might wish to know more. The interactive nature of the systems presented here in the context of a museum visit permits interactive learning rather than passive and collaborative learning by allowing students to interact with each other and with the tour guide. This visual aspect greatly enhances communications, allowing for perceptions of facial expressions and other body language.

However, specific telepresence application also need to overcome challenges that make the experience at its current stage unrealistic. An important limitation of telepresence technology that we tried to bridge with our proposed system is the lack of awareness of the humans around the robot or elsewhere about the point of regard in the environment of the remote users. Our system bridges that gap by substituting the body language cues that actual museum visitors would provide the museum tour guide with a visualization of the 3D environment in 2D in the robot display where the gaze patterns of the remote users are superimposed. The importance of high quality sensory feedback is paramount in telepresence scenarios. The network of eye trackers system presented here provides an additional important clue to the museum tour guide about what areas in his/her environment are attracting the student's attention.

The non-invasive nature of the eye trackers being used, by virtue of being video based and not requiring any equipment physically attached to the remote museum visitors, makes them non-invasive. The gaze tracking monitoring being carried out in the background results in a transparent user experience. Often, after the calibration the user is not necessarily aware any more than his gaze is being monitor since the task is being carried out silently in the background.

One potential limitation of the system is the possibility of cognitively overloading the museum tour guide that while carrying out his usual touring tasks has to pay attention to the robot's display in order to monitor where in the environment the students are paying attention to. This potential issue can be limited by the tour guide consciously trying to avoid paying attention continuously to the robot's display but rather concentrating on its traditional tasks of guiding the tour. The educator can then look at the robot's display sparingly at key moments during the tour where he feels the need to obtain feedback about the engagement levels of the tour participants.

The gaze data gathered by the system that we proposed here is simply used to provide a feedback signal to the museum tour guide in the current implementation, but once the capabilities to gather the gaze patterns are in place, all sort of sophisticated studies about learning during museum tour visits or other type or robotic telepresence scenarios can be envisioned. For instance, the gaze behavior of a class of students could be monitored and recorded during the museum visit. After the visit, the students would be required to answer a questionnaire to measure the degree of understanding and acquired knowledge that they achieved through the visit. Using this data, correlations between the gaze behavior of students and different performance levels in the questionnaire should be mined in an attempt to generate insights about the relationships between gaze dynamics and learning performance. This knowledge could be harnessed online to gently push students detected to be become disengaged during the course of a museum visit. Different methods to regain the attention of disengaged students towards the task could maximize the educational benefits of the museum tour.

Any sort of large scale study looking for correlations between learning and gaze dynamics in a robotic telepresence scenario would need a fairly large number of computers equipped with eye trackers to obtain enough statistical power. This could be a limiting factor in case the gaze trackers required by the proposed method would be expensive. However, we reckon that low end eye trackers providing 30 frames per second tracking performance and the standard reported average accuracy of 0.5 degrees at 60 cm from the screen should suffice for numerous analysis and experimental scenarios. The current trend towards lower cost eye trackers from major manufacturers suggests that eye trackers hardware costs will not be a limiting factor for this sort of research.

In summary, the system presented here is first of its kind in capturing the gaze data of several remote museum visitors in a robotic telepresence scenario in order display their gaze activity to a museum tour guide. This important feedback signal helps to bridge the communication gap between the museum tour educator and the distant museum visitors. Further studies should build over the proposed system and use the gaze data generated by the system to find out those students that are not benefiting from the museum tour and provide them with online signals to recapture their attention dynamically in order to improve the educational value of the museum visit experience.

References

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Duncanson, J. P., Williams, A. D. (1973). Video conferencing: Reactions of users. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *15*(5), 471–485.

Egido, C. (1988). Video conferencing as a technology to support group work: a review of its failures. In *Proceedings of the 1988 acm conference on computer-supported cooperative work* (pp. 13–24). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/62266.62268` doi: 10.1145/62266.62268

Gledhill, D., Tian, G. Y., Taylor, D., Clarke, D. (2003). Panoramic imaginga review. *Computers Graphics*, *27*(3), 435 - 445. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0097849303000384` doi: 10.1016/S0097-8493(03)00038-4

Green, P. S., Hill, J. W., Jensen, J. F., Shah, A. (1995). Telepresence surgery. *Engineering in Medicine and Biology Magazine, IEEE*, *14*(3), 324–329.

Hansen, D. W., Ji, Q. (2010). In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478–500. Retrieved from `http://dx.doi.org/10.1109/TPAMI.2009.30` doi: 10.1109/TPAMI.2009.30

Holt, D., Zaidi, A., Abramson, J., Somogyi, R. (2004). Telesurgery: advances and trends. *University of Toronto Medical Journal*, *82*(1), 52–5.

Jacob, R. J., Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, *2*(3), 4.

Kumar, M., Paepcke, A., Winograd, T. (2007, April). EyePoint. In *Proceedings of the sigchi conference on human factors in computing systems - chi '07* (p. 421). New York, New York, USA: ACM Press. Retrieved from `http://dl.acm.org/citation.cfm?id=1240624.1240692` doi: 10.1145/1240624.1240692

Lawson, T., Comber, C., Gage, J., Cullum-Hanshaw, A. (2010). Images of the future for education? videoconferencing: A literature review. *Technology, Pedagogy and Education*, *19*(3), 295–314.

Majaranta, P., Aoki, H., Donegan, M., Hansen, D. W., Hansen, J. P., Hyrskykari, A., Raiha, K. (2012). *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies*. IGI Global. Retrieved from `http://books.google.com/books?id=zSfEQgAACAAJ` doi: 10.4018/978-1-61350-098-9.ch009

Minsky, M. (1980). Telepresence. *Omni*, *2*(9), 45–52.

Rozado, D., Agustin, J. S., Rodriguez, F. B., Varona, P. (2012, January). Gliding and saccadic gaze gesture recognition in real time. *ACM Transactions on Interactive Intelligent Systems*, *1*(2), 1–27. Retrieved from `http://dl.acm.org/citation.cfm?id=2070719.2070723` doi: 10.1145/2070719.2070723

Rozado, D., Rodriguez, F. B., Varona, P. (2011). Gaze Gesture Recognition with Hierarchical Temporal Memory Networks. In J. Cabestany, I. Rojas, G. Joya (Eds.), *Advances in computational intelligence* (Vol. 6691, pp. 1–8). Springer Berlin / Heidelberg.

Rozado, D., Rodriguez, F. B., Varona, P. (2012, August). Low cost remote gaze gesture recognition in real time. *Applied Soft Computing*, *12*(8), 2072–2084. Retrieved from `http://dx.doi.org/10.1016/j.asoc.2012.02.023` doi: 10.1016/j.asoc.2012.02.023

San Agustin, J., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D. W., Hansen, J. P. (2010). Evaluation of a low-cost open-source gaze tracker. In *Etra '10: Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 77–80). New York, NY, USA: ACM. doi:

http://doi.acm.org/10.1145/1743666.1743685

Schultz, R., Nakajima, R., Nomura, J. (1991). Telepresence mobile robot for security applications. In *Industrial electronics, control and instrumentation, 1991. proceedings. iecon'91., 1991 international conference on* (pp. 1063–1066).

Tobii Technology AB. (2012). *Accuracy and precision Test report X1 Light Eye Tracker.* Retrieved 07/09/12, from http://www.tobii.com/Global/Analysis/Training/Metrics/Tobii_X1_Light_Eye_Tracker_Accuracy_and_Precision_Test_Report.pdf

Tsui, K. M., Desai, M., Yanco, H. A., Uhlik, C. (2011). Exploring use cases for telepresence robots. In *Human-robot interaction (hri), 2011 6th acm/ieee international conference on* (pp. 11–18).

Velichkovsky, B., Hansen, J. P. (1996). New Technological Windows into Mind: There is More in Eyes and Brains for Human-Computer Interaction. In *Chi'96* (pp. 496–503).

Zhai, S. (2003, March). What's in the eyes for attentive input. *Communications of the ACM*, *46*(3), 34. Retrieved from http://dl.acm.org/ft_gateway.cfm?id=636795&type=html doi: 10.1145/636772.636795