Web-Based Tele-Robotics Revisited

Elliot Duff, Kane Usher, Ken Taylor and Con Caris

Abstract

This paper describes the development of a open and immersive Web-based augmented virtual environment for the tele-robotic control of mining robots. Many mining tasks must be done in remote locations. The most convenient mechanism for communication for this task is the Internet, for which a browser-based graphical user interface is developed allowing the operator to control the machine by specifying high-level tasks. Rich-user interactivity is provided in the Web-browser interface through the use of X3D and AJAX techniques which provide a means of asynchronous communication between the browser and server.

1 Introduction

Technical advances in the mining industry are dominated by improvements to safety and productivity. Tele-robotics can improve personal safety by removing people from hazardous environments. This has been used for more than half a century with conventional RF technology and is represented in the bottom row of Figure 1. In this figure, video is transmitted to the human operator, who in return, responds with command signals to the robot. Unfortunately the performance of conventional tele-robotic systems has been poor, both in terms of productivity and damage to the machine and there appears to be a complex trade-off between safety and productivity. With advances in machine autonomy, such as dig-to-plan and cooperative behavior, there is an opportunity to redress this imbalance.

During the late 1990s there was a great deal of interest in tele-robotics applications over the web [13]. Although research was conducted to develop a system for the mining industry [4] this technology was never commercially realized. The reasons for this failure is unknown, but at this time, the technology was immature and probably did not provide the appropriate level of immersion and interaction necessary for control (i.e. due to latency and bandwidth issues). With recent advances in network infrastructure, sensor technology and software standards (e.g AJAX and .NET etc) tele-robotics is gaining renewed interest.[1] In the mining industry, there is a demand for visual and haptic user interfaces that are capable of integrating complex geological data with real-time data streams from multiple and remote machine sensors within an open software framework.

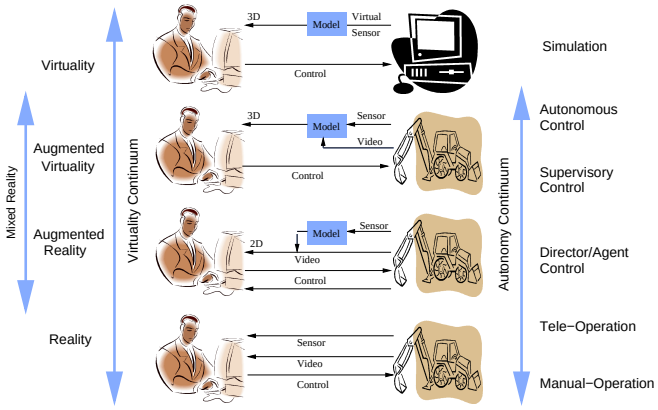

Figure 1: Continuum of virtuality and autonomy.

One of the main criticisms of tele-robotics is that it does not provide sufficient situational awareness [23] to the human operator to sustain the previous manual levels of production. What is required is a higher level of immersion. This may be achievable with mixed-reality user interfaces. The virtuality continuum [17] spans a spectrum from reality to virtuality (see Figure 1). Mixed reality lies between the two extremes and can be divided into Augmented Reality (AR) where data and virtual objects are displayed to the operator on top of a live video stream (third line of Figure 1), or Augmented Virtuality (AV) where real data is placed in a virtual environment [19] (second line of Figure 1). In general, there is a correspondance between the continuum of virtuality and the continuum of autonomy, so that under normal

[1] Ironically, as mining machines have become automated there is an increasing demand for tele-robotics. This is because it provides one of the few ways in which it is possible to operate non-autonomous machines safely in proximity to autonomous machines.

circumstances, AR is suited to director/agent level of autonomy, whilst VR is suited to supervisory levels of autonomy. It is this later interface that is appropriate to the supervision of semi-autonomous mining machines.

For the interface to be immersive and interactive, it is critical that the data be presented to the operator in a natural and intuitive manner. From lessons learnt in the gaming industry, it is clear that the environment should be three dimensional and have a single modality, that is, switching between different virtuality's ("world" representations) is minimized. The gaming industry also reveals that expensive displays are not critical to the immersive experience. To create a virtual world from reality (augmented virtuality) it is necessary to merge disparate and numerous data sets into a coherent whole. Furthermore, to provide timely feedback to the robot, the state of the machine and environment must be measured in real-time.

2 Literature review

One of the first Web-based tele-operation projects[12] involved a mock-up of an archaeological site situated in a radioactive area. Users could join a queue to take control of a 2.5DOF manipulator, searching for 'relics'. This work later evolved into a 'tele-garden' system in which users could log in and remotely tend to a garden. In both cases, latency issues were not explicity dealt with, rather the system had inbuilt checks to ensure that the user could not damage the system [11; 24]. Around the same time in the mid 1990s, researchers at the University of Western Australia also developed a Web-controllable, 6-axis robot with an 'in-hand' camera as feedback to the operator, together with joint angles, gripper pose and Cartesian position [29]. This system has subsequently been converted to a 'tele-laboratory' allowing students to remotely log-in to a manipulator to perform parts of their coursework. Another example is the Kaplan car [15] which allowed a user to control the speed and direction of a remote control car using a video feed as guidance. Again, control is at a low-level and latency is not explicitly dealt with. Xavier [25], a Web-controllable mobile robot, is probably one of the more successful Web-controllable robots. It differs from much of the other work already cited here in allowing high-level task specification — a user can request a position for the robot to go to, and Xavier uses its on-board autonomous navigation system to guide it to the the location. This allows the robot to deal with the low-level control and navigation issues, overcoming many of the problems associated with latency. WebDriver [14] and WITS [2] are examples of Vehicle tele-operation using JAVA over the Web. Fong et al. [10] reviewed much of the tele-operation interface work at CMU ranging from systems with rather low-level control through to systems with much more intelligence at the robot end.

With respect to tele-excavation, the work of Boulanger et al. [4] involved an extensive project for mine virtualization in the coal sands of Canada in the late 1990s. Ballantyne et al. [3; 1] also investigated the use of *Virtual* type displays for excavator tele-operation. Again, real video augmented the feedback to the operator but as with most previous work in this area, the operator control was at a very low-level (i.e specifying points rather than 'tasks'). However, the display enabled the operator to pre-set no-go areas for the excavator and to also mark areas of the excavation site in which the terrain was perhaps to dangerous for the excavator to navigate. Several Japanese groups have also been investigating tele-operation of mining and construction equipment [28; 18; 16].

3 Previous Work

In 1999, CSIRO developed an autonomous underground mining vehicle [22]. The prototype user interface for this system is shown in Figure 2. The left screen displays conventional surveillance video feed from significant locations around the mine (eg. at the dump points). The centre screen displays the tactical state of the selected machine; video stream from the vehicle, laser range data (showing free space around vehicle) and numerous dials (temperature, speed, revs etc). The right screen displays the strategic state of the mine: the position of each vehicle in the mine, the number of passes that it has made in the shift, etc. From the strategic display the operator selects the appropriate load and dump point and the vehicle will automatically drive to the dump point and back. Since the digging has not been automated, the operator is required to directly operate the vehicle though the joystick. Whilst this interface rapidly gained operator acceptance (requiring less than a day to train) the interface in not particularly scalable. In this prototype, the operator could only operate one vehicle at a time. The intention for the commercial system is to simultaneously supervise three or four machines and a more immersive interface is required.

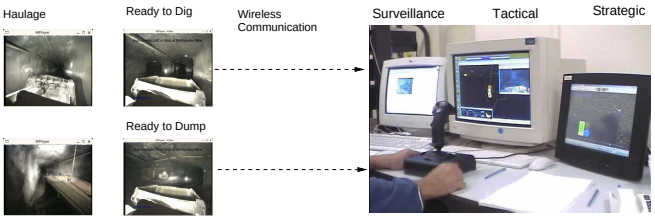

Figure 2: User interface of autonomous underground mining vehicle.

In 2003, CSIRO developed systems to automate the horizon control and face alignment of a longwall shearer [20]. Whilst this removed the need for direct control of the machine, it has not removed the need for people along the longwall. The machine and the environment still need to be inspected. In fact, as machines become increasing automated their is a greater need for preventative maintenance. These people can only be removed if we are able to provide a suffi-

ciently rich immersive environment that will give them the information that they need (i.e. tele-presence). The current visualization system (see Figure 3) implements a synthetic 3D environment that enables the integration of complex geological data with real-time data streams from mining equipment sensors. Sensor data is piped into a Virtual Reality Markup Language (VRML) model that is displayed with an interactive user interface. Artificial colours are used to highlight areas of concern (i.e. sensors that are outside of their normal operational range) and numerical data is overlayed over selected components. The interface can be accessed securely by various users at remote locations via the mines's Intranet. The system is operational in three mines in Australia with an average turn-around latency of 120ms.

Preliminary results indicate that high quality 3D visualization and real-time exception reporting/analysis of automated mining equipment and underground conditions will be essential to give remote users confidence that underground systems are operating optimally and safely. At this stage, survellance video is available via a standard Web browser, however there are plans to install more than 40 cameras along the longwall, so there is a strong desire to integate the video into the 3D data set. If live video is inserted into the appropriate location in the virtual environment, it is possible for the operator to navigate through a *virtual mine* instead of observing a screen full of 2D video cameras. In this way, it is possible to take advantage of the intrinsic spatial awareness of the operator. This technique has been demonstrated in a security application [26].

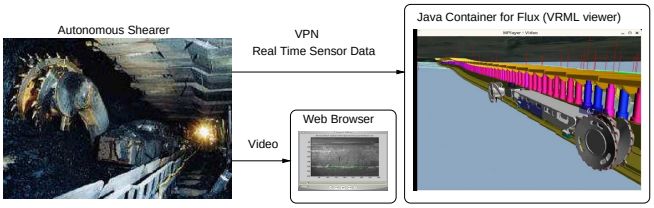

Figure 3: Live 3D visualization of longwall shearer.

4 Design

One of the shortcomings of the previous work has been the Ad-Hoc nature in which the user interfaces have been developed. Whilst components based robotics (such as Player/Stage, ORCA and Microsoft Robotic Studio) and Web-based tele-robotics (such as LabView and Matlab) have moved to the mainstream, the user interfaces ave not provided the level of immersion necessary to provide sufficient situational awareness to control dangerous and expensive equipment in remote and unstructured environments. Some researchers have proposed a framework based upon gaming technology [21], but such technology does not easily allow the *virtual* world to change. Typically the virtual world is built with *prior world knowledge*, that only allows small variations to its structure (i.e. opening a door). This is not suitable for the mining industry where the environment is continually modified. Our proposed framework is intended to be Web browser based and utilise XML standards so it will function on the widest variety of computing equipment possible. It will be open and scale smoothly across both varying bandwidths and varying display and operator equipment capabilities. Ideally it should scale down to mobile phone capabilities and still provide useful information regarding machine operation and performance. The framework should also scale up to take advantage of high bandwidths, where available, and function on large multi-screen operator stations.

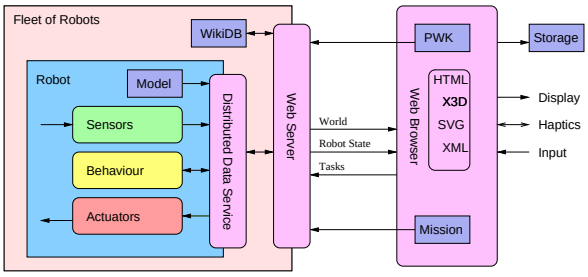

Figure 4: Components of Web-based Augmented Virtual Environment.

A browser-based interface has several advantages over development in other GUI environments. The first is that almost everyone has the pre-requisite software — there is no need to link to specific libraries, run specific operating systems, etc. Of course similar arguments could be made for Java-developed applications, however a browser interface is somewhat lighter than a Java application. A further advantage is that the browser needs no knowledge of the internal workings and network structure of the server and the upstream computing. This is an attractive attribute when providing access to potential clients and customers,.

In the past, user interaction with a Web-page could be quite clumsy with each server request from the browser effectively triggering a HTTP request back to the server from which an entire HTML page was generated and sent back to the browser. New mechanisms for asynchronous communication with the server have since emerged which allow seamless updates of parts of the page, rather than a refresh of the entire page. These methods are rather loosely termed AJAX techniques for Asynchronous Javascript and XML. In effect, AJAX methods allow for much richer user interaction with a Web-page. Good examples of the use of these techniques include some of the emerging Google applications,, such as *Google maps*.

The proposed framework is to use AJAX3D (`www.ajax3d.org`) which merges X3D and AJAX technology. X3D (`www.x3d.org`) is an ISO standard for real-time computer graphics that is presented as an XML document.

It can be viewed with appropriate software (eg. Flux from www.mediamachines.com). It can be inserted into host HTML documents and viewed with plugins to standard Internet browsers. Multimedia streams can be placed onto surfaces in the environment (e.g. video onto billboards). Different browsers can support audio, stereo displays and haptic devices. Being an open standard there are a number of CAD systems that are able to export drawings to X3D (or at least VRML which is backwardly compatible). With AJAX calls to the DOM (Document Object Module) and SAI (Scene Authoring Interface) it is possible to partially load parts of the scene (i.e. update the world) or reload features of the scene (i.e. a joint angle). It also provides for commands from the user to be sent back to the server.

The components of the proposed framework are shown in Figure 4. Here the robots consist of three component classes: **Sensors** that detect the state of the robot and the world (i.e. lasers, encoders); **Behavior** that responds to sensor information (i.e wall following); and **Actuators** that produce a physical response to the behavior (i.e. steering, speed). To facilitate distributed control of the robot (Tele-Robotics) data is registered with a Distributed Data Service[2]. In addition there is a **Model** of the robot in X3D. **PWK** is an XML file that contains the prior world knowledge, such as, a survey map and the approximate location of the robot. This would normally be provided by the remote operator. The **Mission** is a file that contains a list of tasks for the fleet of robots. Once again this would be provided by the remote operator with sufficient authority to do so. **WikiDB** is a database that records configuration data. The nature of the **Storage** has yet to be determined, however it will be required to replay historical events.

The key to the system, is the fact that an association is made in the Web server between a number of sensor states and the X3D model. For example, the measured steering angle is associated with the steering angle of a joint in the X3D model, so that changes in the steering angle are reflected in the X3D Scene displayed on the Web browser. These changes are seen in real-time because the data is provided asynchronously. Likewise, an association is made between the PWK and new world knowledge that is gathered from the on-board sensors. The management of these data sets (in dark blue) and numerous configuration data is handled by a Wiki database. In this way the configuration data can be edited with the web browser itself. The Wiki also has the ability to reformat the wiki pages in response to different display devices, called 'skins'. For example, a low resolution X3D file is sent to a mobile phone, whilst a high resolution X3D file is sent to a high bandwith wide screen display. The Wiki has the added advantage in that it provides a level of security and the ability to keep track of changes. These changes can be "rolled" back if necessary. When combined, these features make the Wiki an ideal collaborative tool. Other than the Distributed Data Server, all the components are open and freely available. The intent of this framework is not to prescribe the back-end (the robot) but to create a interface between the back-end and front-end that can be easily reused and reconfigured.

[2]There are a number of middle-ware products that can do this (CORBA, ICE etc) but here DDX (Dynamic Data eXchange) was used [5] .

5 Implementation

NASA is interested in remote excavation for possible future missions which may involve the construction of lunar habitats [8]. Towards this aim, CSIRO was commissioned to build a system which enabled the demonstration of remote excavation activities with large latencies between the operator and the excavation machinery. The demonstration system consists of a 1:7 scale model dragline, a back-end networking system including a router with a customizable time delay (set to 8 seconds to simulate communications to the moon), and a Web-based interface to control the machine. The graphical interface (see Figure 5) allows for high-level task control of the machine, counteracting the large latencies in the system. The operator is able to command the machine to collect a Digital Terrain Map (DTM) which is then displayed in the browser. Based on this DTM, the operator can then interactively select an excavation region and a dump zone, and then command the dragline to perform the excavation. Feedback to the operator includes live updates of the dragline's state, video feeds, and the digital terrain map in which a model of the dragline can be displayed.

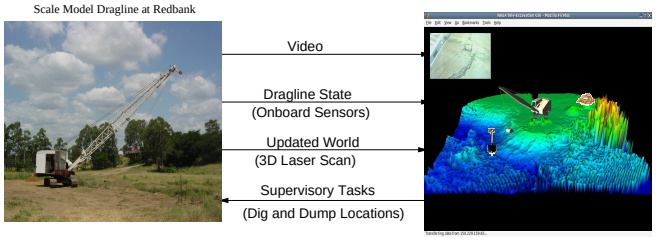

Figure 5: Dragline Web-based Augmented Virtuality.

The principal demonstration for this project involved an operator in the USA controlling the scale-model dragline at Redbank, Queensland, Australia. The scale model dragline has been fitted with suffient sensors and control systems to provide semi-autonous control of the bucket motion [9]. The Internet provides a natural means for communication between the two sites. Connection of the machine to the outside world, via the CSIRO network, is achieved through a complex route involving: a local network on the dragline machine connecting the control computers to the low-level systems which interface to the machine's actuators; a wireless connection to a router located in a nearby workshed; an ADSL link from the Redbank site to the CSIRO QCAT site (via several 'tunnels'

and a microwave link); and a connection to a Web-server located outside of the CSIRO network. Over a six month period, the scale model dragine was controlled from numerous sites around Australia, and two sites in the US; Boston and Washington. Preliminary results show that the machine is capable of taking high level commands from a remote operator and operate autonomously within a changing environment (click-to-dig). A demonstration of 50 consecutive autonomous cycles moved approx. 5.1 m3 of material in 52 minutes (see Figure 6).

6 Summary

This project is certainly not the first to use the Internet as a mechanism for providing connectivity between a user and a machine. In fact, this project is not even the first to perform tele-excavation over the Internet. However, it is one of the few systems developed which have allowed a user to connect to a remote robot using a standard Web-browser, and to allow the subsequent specification of high-level tasks rather than specifying low-level joint or Cartesian space demands. This high-level autonomy provides the key to overcoming the take-up of Web-based user interfaces, the primary drawback of which, even with modern techniques, is somewhat unpredictable latency between the client and server. For short delays of less than two seconds, techniques based on predictive control [27] can be quite effective but for larger latencies the only real solution is to build in some level of autonomy which enables the machine to deal with local control issues, while providing the operator with enough flexibility to achieve the desired task.

This paper describes the development of a augmented virtuality framework, in both project specific and generic terms. Although aspects of the discussion are specific to mining, the paper is intended to provide an argument for the development of future human-robot interfaces to occur in this environment, freeing the end-user of any dependence on operating system or specific applications or libraries. In short, Web-based GUIs should be a serious consideration for the remote supervsion of robotic systems.

At this stage the mission planning is rather primitive. It consists of a simple python script that can be edited via the Wiki. In future, we will be reviewing more sophisticated languages [6] and examining their implementations in Neptus [7] and Maestro (`mars.telascience.org`) for inspiration. Future work includes: improve placement of video onto the 3D surface (eg. registration); investigate different ways of presenting data (e.g. colour, sound, vibration); apply framework to other robotic platforms; test framework with multiple robots and allow other robots to use “virtual” world as a “virtual” sensor; deploy client side storage with the ability to rewind “virtual worlds”; and most importantly, measure the performance of the framework, both in terms of situational awareness and operator acceptance.

Figure 6: Photo of spoil pile after 50 autonomous cycles

7 Acknowledgments

The authors would like to thank the rest of the CSIRO Mining Robotics team for their valued contribution in hardware and software support. Preliminary dragline and rope shovel research has been conducted with the support of the Australian Coal Association Research Program (ACARP) through projects C3007, C5003, C9028,C13040, C12030, and NASA project 62442132.

References

[1] Michael Greenspan amd John Ballantyne and Mike Lipsett. Sticky and slippery collision avoidance for tele-excavation. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 1666–1671, Grenoble, France, September 1997.

[2] Paul Backes, Kam Tso, and Gregory Tharp. Mars pathfinder mission internet-based operations. In *Proceedings AUVSI North America Conference*, 2006.

[3] John Ballantyne and Eldon Wong. A virtual environment display for teleoperated excavation. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 1894–1899, Victoria, Canada, 1998.

[4] P. Boulanger, J.-F. Lapointe, and W Wong. Virtualized reality: an application to open-pit mine monitoring. In *Archives of the 19th Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS)*, volume XXXIII, Part B5/1, pages 92–98, Amsterdam, The Netherlands, July 2000.

[5] Peter Corke, Pavan Sikka, Jonathan Roberts, and Elliot Duff. DDX: A distributed architecture for robot control. In *Australian Conference on Robotics and Automation*, Canberra, Australia, 2004.

[6] Eve Coste-Maniere and Reid Simmons. Architecture, the backbone of robotics systems. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2000.

[7] Paulo S Dias, Rui Gomes, and Jose Pinto. Mission planning and specification in the neptus framework. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2006.

[8] Matthew Dunbabin, Peter Corke, Graeme Winstanley, and Jonathan Roberts. Off-world robotic excavation for large-scale habitat construction and resource extraction. In T. Fong, editor, *To Boldly Go Where No Human-Robot Team Has Gone Before: Papers from the 2006 AAAI Spring Symposium*, Technical Report SS-06-07, pages 95–103. AAAI Press, July 2001.

[9] Matthew Dunbabin, Peter Corke, Graeme Winstanley, and Jonathan Roberts. Off-world robotic excavation for large-scale habitat construction and resource extraction. In *To Boldly Go Where No Human-Robot Team Has Gone Before: Papers from the 2006 Spring Symposium, American Association for Artificial Intelligence*, 2007.

[10] T. Fong and C. Thorpe. Vehicle teleoperation interfaces. *Autonomous Robots*, 11(1):9–18, July 2001.

[11] Soraya Ghiasi. A reusable framework for web-based teleoperation of robotic devices. Master's thesis, University of Colorado, November 1999.

[12] Ken Goldberg, Michael Mascha, Steve Gentner, Nick Rothenberg, C. Sutter, and J. Wiegley. Desktop teleoperation via the World Wide Web. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 654–659, Japan, April 1995.

[13] Ken Goldberg and Roland Siegwart. *Beyond Webcams An Introduction to Online Robots*. MIT Press, 2001.

[14] S. Grange, Terrence W Fong, and C. Baur. Effective vehicle teleoperation on the world wide web. In *International Conference on Robotics and Automation*. IEEE, April 2000.

[15] A. E. Kaplan, Srinivasan Keshav, N. L Schryer, and J. H Venutolo. An Internet accessible telepresence. *Multimedia systems*, 5(2), 1997.

[16] K. Kosuge, K. Takeo, and H Ishida. Teleoperation system of (sic) power shovel for subterranean line work. In *Proceedings of the 23rd International Conference on Industrial Electronics, Control and Instrumentation*, pages 1421–1426, November 1997.

[17] P Milgram, H Takemura, A Utsumi, and F Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. In *Proceedings Telemanipulator and Telepresense Technologies*, volume 2351, page 282, 1994.

[18] M. Minamoto and K Matsunaga. Tele-presence information and remote-controlled task execution. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pages 1102–1106, Victoria, Canada, October 1998.

[19] Peter Rander, P. J. Narayanan, and Takeo Kanade. Virtualized reality: constructing time-varying virtual worlds from real world events. In *VIS '97: Proceedings of the 8th conference on Visualization '97*, pages 277–ff., Los Alamitos, CA, USA, 1997. IEEE Computer Society Press.

[20] David Reid, David Hainsworth, and Jonathan Ralston. Shearer guidance: A major advance in longwall mining. In *Proceedings of the International Conference on Field and Service Robotics*, 2003.

[21] Justin Richer and Jill L. Drury. A video game-based framework for analyzing human-robot interaction: characterizing interface design in real-time interactive multimedia applications. In *HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 266–273, New York, NY, USA, 2006. ACM Press.

[22] Jonathan M. Roberts, Elliot S. Duff, and Peter I. Corke. Reactive navigation and opportunistic localization for autonomous underground mining vehicles. *International Journal of Information Sciences.*, 145(1-2):127–146, 2002.

[23] Brennan P. Sellner, Laura M. Hiatt, Reid Simmons, and Sanjiv Singh. Attaining situational awareness for sliding autonomy. In *HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 80–87, New York, NY, USA, 2006. ACM Press.

[24] D. Seward, F. Margrave, I. Sommerville, and R. Morrey. LUCIE the Robot Excavator - Design for System Safety. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 963–968, Minneapolis, USA, April 1996.

[25] Reid Simmons, J. Fernandez, R. Goodwin, S. Koenig, and Joseph O'Sullivan. Xavier: An autonomous mobile robot on the web. *Robotics and Automation Magazine*, 1999.

[26] Kristian Simsarian and Karl-Peter Akesson. Windows on the world: An example of augmented virtuality.

[27] Andrew C. Smith and Keyvan Hashtrudi-Zaad. Smith predictor type control architectures for time delayed teleoperation. *International Journal of Robotic Research.*, 25(8):797–818, 2006.

[28] H. Taketsugu, Y. Takashi, A. Junichi, I. Masaki, and Y. Hiroaki. Experimental land model of tele-operated underwater backhoe with AR technology. In *Proceedings of the International Symposium on Underwater Technology*, pages 339–344, April 2004.

[29] K. Taylor and J. Trevelyan. Australia's telerobot on the web. In *26th International Symposium on Industrial Robotics*, 1995.