



VIABILITY OF TELEMEDICAL REMOTE ROBOTIC SYSTEMS

A. Smit^{1*} and Dr. A.F. van der Merwe²

¹Department of Industrial Engineering
University of Stellenbosch University, South Africa
14925435@sun.ac.za

²Department of Industrial Engineering
University of Stellenbosch, South Africa
andrevdm@sun.ac.za

ABSTRACT

Worldwide, the demand for telemedicine devices and systems has risen steadily in recent years. The design of remote robotic systems has always been a complex undertaking. Technological limitations affecting bandwidth and access to communications infrastructure have impeded the advancement of a long distance remote robotic telemedicine system training purposes. Assessing the commercial viability and acceptability will be achieved by evaluating responses from expert user groups. This plays a crucial role in the development of future systems as well as providing expert-user data on user perception of such systems and data on the current viability of long distance patient care. An assessment of the information that is required to be presented to the medical professional in both training and diagnostic situations would have to be completed by making use of questionnaires and interviews. The system will be developed using low cost technologies with support structures in place as to facilitate financial feasibility and minimization of complexity along with intuitive control inputs. The results of the study will provide valuable information for the future development of commercially viable robotic telemedicine systems in South Africa.

*Corresponding author

1. INTRODUCTION

Worldwide, the appeal of telesurgery systems has increased rapidly as access and equality to healthcare has become a priority to the World Health Organisation and many likeminded governments. The focus of this paper is the development of a low-cost robotic tele-operation system to assess the viability of robotic tele-medicine systems as applied in training scenarios. This requires the development of a system similar to telesurgical systems.

Telesurgery is a specialized branch of teleoperation. To understand the basis of tele-surgery one first need a basic understanding of teleoperations. Teleoperation is defined by Sheridan [1] as:

“Teleoperations is the extension of a person’s sensing and manipulation capability to the remote location. A teleoperator includes at the minimum artificial sensors, arms and hands, a vehicle for carrying these, and communication channels to and from the human operator.”

Cui et al. [2] defines it as:

“Teleoperation is a means to operate a robot using human intelligence, which requires the availability of adequate human-machine interface. A teleoperation system usually consists of two robot manipulators that are connected in such a way as to allow the human operator control one of the manipulators, which is called the master arm, to generate commands that map to the remote manipulator, which is called the slave arm.”

Cui et al’s definition focuses on a master/slave architecture exclusively whereas Sheridan’s allows for other control structures such as supervisory control etc.

Sheridan also discussed one of the first modern tele-operation systems which were implemented in a project by Raymond Goertz in the 1950s. The project was in response to a need to manipulate hazardous nuclear material in what is known as “hot-cells” remotely [1]. This initial system was little more than a mechanical system with direct linkages; the system later incorporated CCTV displays and servomechanisms [1]. The focus of this paper will be on master-slave systems where the distance at which the operator (master) and robot/manipulator (slave) are separated is arbitrary. Today the most common applications for teleoperations are for hazardous material handling, remote inspection and repair in deep sea oil mining and exploration as is the case with remotely-operated vehicles (ROVs) or manipulation of objects in space exploration [1],[2].

As shown in Satava [3] modern telesurgery has its roots in teleoperations and virtual reality technologies as one of the first fully functioning modern tele-surgery systems was pioneered by NASA’s Ames research center along with the Stanford Research Institute (SRI) for use in hand surgery [3]. As stated by Satava [3] the project incorporated intuitive interfaces and the concept of telepresence in its design. Telepresence design methodology has become an integral part in modern systems such as the Da Vinci surgical system where the surgeon feels as if they are present in the tele-environment. In 1988 another telesurgery application adapted a Puma 560 robot to perform soft-tissue surgery as was discussed in Davies, *A Review of Robotics in surgery* [4]. Some other popular systems that have been used to date are the AESOP endoscopic system, Prodoc, Robodoc, Zeus and the Da Vinci surgical systems [5].

Although transatlantic surgery has been performed successfully (Marescaux et al [6]), large scale implementation of such systems has not yet become viable. Robotic systems have been applied successfully in in-hospital applications, where the doctor and patient have been separated only by a few feet (or in the next room). The Da Vinci surgical system (da Vinci Surgery [7]) is an example of such an application. Although theoretically the distance between input and output can be arbitrary, this functionality is not currently offered in the completed Da Vinci surgical system. These systems have limitations as far as affordability and patient-doctor separation, due to the high data transfer rates needed for the various control inputs as well as dense media feedback from the patient side. This is due mainly because of the time delay introduced by data transmission.

Current systems such as the Zeus and Da Vinci Surgery robotic system focus mainly on minimally invasive procedures. The scaling of input is a major advantage over the traditional laparoscopic procedures when using a robotic system. These systems are expensive to implement and develop. Therefore, to ensure feasibility, the cost has to be passed down to patient level. Unfortunately, the increased costs associated could result that large-scale deployment of such systems become unfeasible.

The use of 3D technologies in robotic assisted surgery offers the added benefit of depth perception. This greatly improves user experience and technical ability of the system as was found in Falk et. al. [8]. System performance was improved while using 3D when compared to 2D views, specifically the speed with which tasks can be completed. Such a 3D vision system has been employed successfully in the Da Vinci surgical system. With a natural posture and intuitive mechanical gripper manipulator design, the system has become a leader in robotic surgery.

The author proposes an innovative approach to training as applied to healthcare with the aid of a remote robotic system. The system aims to be affordable, scalable, and flexible while focusing on intuitive interfaces and ease of use while optimizing the patient experience. The current system under development is aimed at diagnostics and training (e-learning). The proposed system will focus on macro observation and interaction while being scalable to micro manipulation for advanced operations, also while being flexible enough to be able to reconfigure fast, easily and with minimal turnaround time. Development is aided by critical evaluation of supporting technologies. The aim of the evaluation is to find a low cost solution for online tele-operations and create an interaction environment and framework that is flexible, scalable, and intuitive. The system will be developed using common commercially available software and hardware with established support structures.

2. SYSTEM DESIGN AND TECHNOLOGY EVALUATION

Tele robotic (tele-operation) systems consist of three top tier elements, shown in figure 1. They are Primary user interface (Input), Communication and Client-side (Output) respectively. Each of these tiers consists of complex and sometimes interwoven solutions to the function problem. The tiers will be discussed in terms of the various technologies available to system designers as to facilitate system design for a viable solution. Each technology has its own merits and drawbacks along with a degree of inherent ergonomic appeal to the user. The focus will be mainly on the use of alternative technologies; technologies mostly adapted from the entertainment and recreation markets to facilitate shorter learning curves and reduced complexity while maintaining high technical ability.



Figure 1: Tele robotic three top tier elements

It is also of paramount importance that the system is economically feasible for implementation in many sectors of healthcare. These include tele-surgery, dermatology, diagnostics, and training at a distance to name just a few. Some of the more important aspects to consider are the development and running cost of new technologies and hardware as well as the limits and costs associated with bandwidth when utilizing a telecommunication networks.

2.1 Primary User Interface

A user interface which requires minimal adaption time and with intuitive input will enable any healthcare professional to easily perform highly complex tasks. Overly complex input design may, however, negatively affect the user performance and accuracy. The use of intuitive input methods and control allows the user to interface and interact with the system easily and decreases the learning curve. Input modes can include physical interaction, for example moving a manipulator, input from a keyboard, haptic device input or motion capture from inertial sensors, oral interaction (voice commands), and visual interaction (cameras with optical tracking or fiducially markers) to name but a few. Through a combination of these technologies a viable commercial design for tele-operations in medical applications can be found using low-cost alternatives to purpose designed hardware.

When considering the interface environment, one also has to make the distinction between immersive and standard interactive environments. Immersive environments, as the name suggests, immerses the user into a virtual or tele-environment with a good degree of situational awareness. In immersive environments it feels as if the user is present in the virtual/tele-environment i.e. most if not all the sensory feedback/input to the user is from the tele-environment.

2.1.1 Input technologies

2.1.1.1 Tactile input - Haptic Control

The effects of using haptic control/input have been the subject of much research in recent years (Barnes [9]). To understand the immediate benefits of haptic control one first has to define haptic:

“Of or relating to the sense of touch; tactile”.

Haptic control may also be referred to as force-feed back control, where the user has tactile feedback from the tele-environment. This type of input/feedback device is key to an immersive environment. Barnes [9] also showed that user performance increased when using haptic devices, making haptic feedback attractive for applications which require an already high degree of skill to complete successfully although this also showed that tasks completion takes longer for the specific experimental setup.

Haptic feedback is of paramount importance for the user experience to achieve a real-time immersive environment which delivers life-like interaction with any and all objects

and surfaces that can be manipulated. This is even more relevant in the tele-surgery scenario, where highly technical operations need to be completed.

2.1.1.2 Command Based

Command input is the most basic control method for the manipulation of tele-robots and is familiar to most operators of remote robotic systems. The use of a mouse and keyboard to control robots is usually achieved using a mouse and keyboard in conjunction with a virtualized environment. The virtualized environment is well known to most users, where a model of the tele-environment is presented along with environment specifics, such as limits of movement, static obstacles etc. The use of a virtualized environment also has the drawback of unsafe human machine collaboration as the user is unaware of the robot's dynamic surroundings. The virtual environment usually has some form of point of view control to virtually manipulate the user's point of view of the operation field. Using point and click, or serial keyboard commands the robot is instructed to move to the desired position. The most basic is a set of co-ordinates usually X, Y and Z that moves the end effector of the robot to the desired location.

Although intuitive, there is little feedback and thus increases the potential risk of injury or damage to people and property. This type of control is most applicable where there are exact fixed parameters of operation and lends itself to the automation environment.

2.1.2 Verbal Input

Verbal command is one of the more desirable input technologies. This technology has been applied successfully for robotic assisted procedures as is the case with the AESOP robotic system (Mettler et al [16]). Voice command input is desirable in scenarios where either complex two handed control of apparatus are required or physical control input is not possible, for instance in a two handed procedure [10]. Assistive robotic systems are useful when applied to surgery especially in long tedious procedures where consistent and precise assistance can be delivered as needed.

2.1.3 Optical Tracking as input

The field of optical tracking is vast and complex. Using image analysis or Digital Signal Processing (DSP) skeletal and feature tracking can be achieved. Software such as FaceAPI can be used for robust multiple degree of freedom facial tracking. Tracking data such as position and orientation data can be used to manipulate either the virtualized 3D view, similar to virtual reality (VR) applications or to manipulate the physical camera orientation and position using a manipulator.

Gesture Identification has recently been used in an assistive framework to assist surgeons during surgery as a visualization tool, where Magnetic resonance imaging (MRI) scans can be manipulated and displayed without the need to physically interface with the display (Baute [11]). Even with a limited instruction set and optical guidance, these systems can greatly improve the medical professional's experience during surgery.

Alternative optical tracking systems include the low cost technologies included in progressive gaming technologies such as Microsoft Kinect and Nintendo Wii. Optical distance sensors such as the Microsoft Kinect in conjunction with feature recognition software may be used to track features. This data is used to determine orientation while position data may be gathered from the distance sensors. The Nintendo Wii has already been shown to be competent as a control input, where the remote has been used to control an industrial robotic arm (Neto [12]) in conjunction with voice commands.

2.1.4 Inertial motion capture

Inertial motion capture has been used in the film animation industry for years. This technology has also been applied successfully in biomechanic research, for motion capture. A well-known inertial motion capturing system is the Xsens MVN suite, where full body motion capture can be achieved. An example of the use of this technology is presented in Cutti et al [13] where Xsens sensors are used as data capture tools for gait analysis. In general, data is generated as a set of relative accelerations with fixed mounting points known. This acceleration data can be used to calculate the relative motion between nodes. These systems require relatively complex and somewhat creative calibration techniques. Interference also poses a challenge as any ferrous metal will affect the accuracy of some of the sensors. This input technology is suitable for motion tracking, particularly for 6 DOF head tracking. The technology also allows for filtration of large motions or even tremor reduction. The problem with such systems arises from the un-ergonomic input style, where users are often forced to move in unnatural degrees of freedom to get a desired output if care is not taken during the calibration process. Magnetic interference can also affect the accuracy of these systems.

A high level analysis of the input technologies is shown in table 1.

<i>Technology</i>	<i>Cost</i>		<i>Level of Ergonomic Appeal</i>	<i>Ease of Integration</i>
	Implementation	Maintenance		
Haptic Input	High	Low	High	Intermediate
Command Based	Low	Low	Low	high
Oral	Low	Low	Intermediate	low
Optical Tracking	Intermediate/High	Intermediate	High	Intermediate
Inertial motion capture	Intermediate/High	High	Intermediate	Intermediate

Table 1: Technology Comparison

3. COMMUNICATION

The most challenging aspect of tele-operations is the dense video stream which is required for feedback to the user. The system bandwidth is invariably determined mainly by the video quality and still image data that must be transmitted. The burden of media can be lessened by encoding to compress the images. For as near as possible to real-time compression and decompression, hardware encoding is advised as software encoding may add a significant time delay. Delay may lead to poor performance when performing tasks using a remote robotic system to perform small complex tasks.

As the system will be implemented in a diagnostic and training capacity, so the design will be based on a metropolitan scenario with free access to potentially “unlimited” bandwidth and throughput. The system will be designed to achieve a minimum bandwidth requirement. This will be achieved using hardware audio and video compression along with a UDP streaming function with minimal command size for the robot command structure. The final data transfer required will be assessed on a local area network whereas system stability and performance will be assessed at varying throughput speeds,

simulating available data transfer speeds. Wireless data transmissions for emergency care will not be considered as mobility is outside the scope of the current project.

For the initial realization of the system, external data transfer (internet based transfer) will not be tested, thus the security of data streams will not be evaluated. The first phase implementation will only be rolled out over a large TCP/IP network. Upon final implementation, however, data security would have to be evaluated as sensitive data will be transferred in the final realization.

4. CLIENT SIDE: OUTPUT AND FEEDBACK (PATIENT SIDE)

There are two typical cases that need to be considered when developing the patient side interface and equipment. The first is a purely diagnostic situation with only visual and audio feedback while the second includes a remote robotic arm or device with haptic feedback that allows remote operation to achieve the desired output.

Humanoid robots are not always the answer; they are some of the most complex robots in use today. An industrial alternative was identified; the Motoman SDA10 Robot (figure 2) was selected as it approximates a human in such a way as to put the patient at ease while not sacrificing technical performance. Its base humanoid appearance and two arms with 15 DOF in total allows for scalable movement, thus increasing accuracy and control. This system was also chosen because of the ease with which the input can be scaled to a smaller motion output. Industrial robots can easily be adapted for surgery as well as global support and maintenance network availability.



Figure 2: Motoman SDA10 (YASKAWA MOTOMAN ROBOTICS) [17]

Visual feedback is usually given with cameras, while other imaging data can be sent from remote sites to medical professionals. Camera systems are complex and the need for low cost alternatives to high end specialized optics-based diagnostic cameras have been a challenge. As was discussed in Klutke et al [14], the evaluated low cost videoconferencing equipment has also been analyzed, on a qualitative basis, as an alternative and was found to be a viable alternative for meetings and information sharing. This technology is appropriate for scenarios where a full-field view is desired as opposed to point-of-view (POV) view of the region of interest. As the system will require cameras for inspection purposes, the author proposes the use of low cost, action cameras that support HDMI streaming to a capture server which in turn streams via the internet.

With 135 degree standard field of view, the cameras can be used in a wide variety of applications. With these cameras macro inspection and macro view of the operating field

can easily be achieved as they have minimal image distortion. Action cameras have been found to be the most cost effective solution as USB web-cameras did not perform adequately under testing. These cameras also possess 1080p resolution and have easily scalable resolutions and frame rates resulting in a flexible input solution.

The author proposes a natural interaction robotic alternative, reducing the need for human assistance when manipulating the camera. While most action cameras do not possess zoom lens functionality, the manipulation of a robotic arm in conjunction with autofocus functions compensates for the lack of zoom. This is also helpful as the robotic arm is able to mimic the user's height as to help with a more user specific ergonomic perspective.

5. PROPOSED SYSTEM AND EVALUATION

The system will be implemented using intuitive, immersive technologies. Head mounted displays with 6 DOF head tracking will be used to translate user movements to robotic motion output. This will be done in conjunction with haptic feedback manipulation tools, to mimic surgical/diagnostic tools.

One objective of the system developer is to produce a system that utilizes as many open-source technologies as possible. To this end the system will use free streaming servers and clients such as VLC to share media streams. The video system will be tested on a local area network under varying network loads. This will assess the maximum delay that can be imposed on a tele-operation before performance is adversely affected. Further user perception of image quality will be assessed. This will be achieved as follows; the same images and video will be shown repeatedly at varying resolutions and image enhancement levels in a comparison chart. Some images will have different levels of saturation, colour, and intensity to find the combination which requires the least bandwidth while being the most pleasing to the user. This data will also be used to set selection guidelines to which mode the cameras must operate for a given scenario.

Timed exercises using the final system will be carried out, accuracy and repeatability will be assessed using varying image quality, to assess the effects of image quality on system specific performance. This will investigate the effects of image quality on user performance and assist in selecting the optimum resolution for carrying out various tasks while minimizing the total system bandwidth requirement.

Two sets of cameras will be employed. The first is the point of view camera as mounted on the robotic arm and will be manipulated by the medical professional remotely. The other will be a Pan/Tilt/Zoom (PTZ) camera mounted above the robot giving the medical professional a full view of the field of interest. User performance will be compared using the overhead and Point of View (POV) cameras. This will assess the impact of POV cameras on the performance for the tele-operated system.

6. CONCLUSION

This pilot study has shown that it is feasible to assemble a system to view and control a remote operation. However, the accuracy and usability have to be proven in an experimental study. The user interface and response time are the two parameters considered a challenge. The results of the study will prove its feasibility. Based on the results recommendations can also be made as to facilitate future viability and possible cost reductions to assist future feasibility. The modular approach followed would protect the system against redundancy.

7. REFERENCES

- [1] Sheridan, T. B. 1989, Telerobotics , *Automatica* ,25, pp 487-507.
- [2] Cui, J.; Tosunoglu, S.; Roberts, R.; Moore, C. & Repperger, D. 2003, A review of teleoperation system control, *Proceedings of the Florida Conference on Recent Advances in Robotics (FCRAR), FL*.
- [3] Satava R. M. 2002, Surgical Robotics: The early chronicles: a personal historical perspective, *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*, 12, p. 6 - 16.
- [4] Davies, B. 2000, A review of robotics in surgery, *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, Prof Eng Publishing*, 214, pp.129-140.
- [5] Lanfranco, A.; Castellanos, A.; Desai, J. & Meyers, W. 2004, Robotic surgery: a current perspective, *Annals of Surgery, Lippincott, Williams, and Wilkins*, 239, pp 14 - 21.
- [6] Marescaux, J.; Leroy, J.; Rubino, F.; Smith, M.; Vix, M.; Simone, M. & Mutter, D. 2002, Transcontinental robot-assisted remote telesurgery: feasibility and potential applications, *Annals of surgery, Lippincott, Williams, and Wilkins*, 235, pp 487 -492
- [7] da vinci Surgery. (n.d.). *Frequently Asked Questions*. Retrieved August 2011, from da vinci Surgery: <http://www.davincisurgery.com/davinci-surgery/frequently-asked-questions.html>
- [8] Falk, V.; Mintz, D.; Grunenfelder, J.; Fann, J. & Burdon, T. 2001, Influence of three-dimensional vision on surgical telemanipulator performance, *Surgical endoscopy, Springer*, 15, pp. 1282-1288
- [9] Barnes, D. & Counsell, M. (n.d.), Haptic communication for remote mobile manipulator robot operations, *Department of Electronic & Electrical Engineering, Salford, UK*
- [10] Geriatrics & Aging, 2003, First Long-distance Robotic Surgery Successful, But Skeptics Question Safety, *Geriatrics & Aging*, 6, pp. 82
- [11] BAUTE, NICOLE. 2011. *Surgeons use Xbox to keep hands sterile before surgery*. Retrieved June 2011, from Healthzone.ca: <http://www.healthzone.ca/health/newsfeatures/article/960393--surgeons-use-xbox-to-keep-hands-sterile-before-surgery> accessed: 2011
- [12] Neto, P.; Pires, J. & Moreira, A. 2009. High-level programming for industrial robotics: using gestures, speech and force control, *Citeseer*
- [13] Cutti, A.; Ferrari, A.; Garofalo, P.; Raggi, M.; Cappello, A. & Ferrari, A. 2010, 'Outwalk': a protocol for clinical gait analysis based on inertial and magnetic sensors, *Medical and Biological Engineering and Computing, Springer*, 48, pp. 17-25
- [14] Klutke, P.; Gostomzyk, J.; Mattioli, P.; Baruffaldi, F.; Plasencia, A.; Borrell, C.; Pasarin, M.; Crescenzo, E.; Pipitone, E.; Mancini, C. & others 1999. Practical evaluation of standard-based low-cost video conferencing in telemedicine and epidemiological applications, *Informatics for Health and Social Care, Informa Healthcare*, 24, pp135-145
- [16] Mettler, L.; Ibrahim, M. & Jonat, W. 1998, One year of experience working with the aid of a robotic assistant (the voice-controlled optic holder AESOP) in gynaecological endoscopic surgery, *Human Reproduction, ESHRE*, 13, pp. 2748 - 2750
- [17] YASKAWA MOTOMAN ROBOTICS. (n.d.). *Motoman SDA10D Assembly Robot*. Retrieved June 2011, from YASKAWA MOTOMAN ROBOTICS: <http://www.motoman.com/products/robots/models/SDA10D.php>