Proc. 2001 Australian Conference on Robotics and Automation
Sydney, 14-15 November 2001

The Feasibility of Force Control Over the Internet

Laurence Bate and Christopher Cook
School of Electrical, Computer and
Telecommunications Engineering
University of Wollongong
Northfields Avenue
Wollongong, N.S.W, 2522
{lhb01, c.cook}@uow.edu.au

Abstract

This paper describes the challenges requiring to be overcome in order to be able to utilize the Internet as a medium to provide intelligible force feedback for Tele-robotic systems. Presently there are several devices that can now be controlled over the Internet from previous work by others. It is possible that an expansion of this concept to include force feedback could ultimately lead to remote medical operations (Tele-Surgery) utilizing the Internet as the transport. Successful implementation even in its most rudimentary form could potentially provide much more rapid and expert life saving patient care during the first hour after injury. This time is known within the medical circles as the "Golden Hour". This paper discusses initial work being performed to investigate the relevance and practicality of the Internet to support intelligible force control of Telerobotic systems. A review of current Internet Telerobotic equipment followed by new approaches and experimental work is presented.

1 Introduction

Internet control of equipment has been implemented by several people to provide a novel experience for general internet 'surfers'. This has produced a body of knowledge about the behaviour patterns of such users, and developed some interesting insights into the design of interfaces between the Internet and robotic equipment.

The goal of our work is to explore the technologies that are available in the area of force feedback in order to understand, and quantify, the degree of control and monitoring of a surgical operation that can be achieved across the Internet. It is hoped this will facilitate subsequent commercial development of Tele-Surgery force feedback units.

Typical scenarios which would benefit from the introduction of Tele-Surgery include:

- Delivery of some medical services to remote locations; for example to small country hospitals which are unable to cater for the range of specialist medical problems treated in large city hospitals, medical centers or even country homes.
- While a patient is in transit to the medical facility, for example while inside the Ambulance vehicle, ship or aircraft.
- Casualties of war from the front line. One scenario could be the development of a portable back-pack style unit carried by the army medic at the front line. The unit would consist of a two-way satellite link to surgeons located elsewhere in the world.

Tele-Surgery would allow ad hoc collaboration of leading surgeons from several different countries to operate simultaneously on the same patient.

Today one could find the concept of undergoing operation by a surgeon who is located several hundred kilometres away unsettling. However, if the technology were proven to be reliable, and one was located at a remote outpost, with a window of only a few hours to be operated on, the Tele-Surgery option may well be your only potential lifesaver.

Developing such a medical system will present numerous challenges. Tele-Surgery is a proven technology within structured/controlled environments of a hospital operating theatre where communication can be hard-wired and the doctor and patient are in close proximity. Moving outside of such environments will require innovative approaches.

Using the Internet as a communications channel presents many challenges. Firstly, due to the intrinsic nature of the relevant Internet protocols, the transport medium cannot provide a real time protocol. Secondly, the response obtained is dependent on other traffic, on the particular mode of Internet access available, and other unpredictable phenomena. Methods need to be developed to work around this in order to enable real-time control of the surgical operation being remotely undertaken.

Along with this, the issue of security when using the Internet as a communications channel must be addressed. The surgeon and patient need guarantees that the operation cannot be either inadvertently or deliberately sabotaged.

2 Existing Internet Tele-robots

The University of Western Australia Telerobot site went online late 1994, and is an excellent example of the evolution of telerobotic pioneering. This was developed when real time solutions for Internet technology like Java were only just evolving. The site consists of an ABB IRB 1440 robot. The movement requests are entered by the remote Internet user and are submitted and executed using the Common Gateway Interface (CGI) program written in Visual C++.

This is just one of the many telerobotic sites that have been set up to explore the technologies available with the result of producing entertainment for people on the internet.

These online telerobotic implementations use a system where the operator makes a movement request then waits several seconds before receiving the result and then making the next request. Such a system is unacceptable for Tele-Surgery, since it will not provide sufficiently timely or comprehensive information for a Surgeon to safely operate on a patient. None-the-less, these sites contain many attributes that assist in formulating the implementation of ‘Intelligible Force Feedback for Tele-Surgery over the Internet’.

3 Force feedback Telerobotics Over Long Distances

The typical response times of Internet based long haul telerobotic systems to date have been in the order of seconds [Tsumaki, Goshozono, Abe, Uchiyama, 2000]. Because delays in the order of a tenth of a second are sufficient to hinder the performance of a teleoperator as well as compromise the system’s stability [Anderson and Spong, 1988], various methods have been employed to overcome this issue. The basic method is to simply dampen the whole system, usually by reducing the velocity of the manipulator to almost a crawl. This method works, but it requires extreme patience by the operator, as tasks become laborious.

To tackle this issue, a large amount of effort has been put into the viability of predictive systems. Virtual reality based systems produce timely feedback to the operator. The performance improvement comes at the expense of potentially introducing geometrical modeling errors.

4 Existing Tele-Surgery Units

Already there are several tele-operated medical devices developed in the word today, some of which are in active service, fulfilling life saving operations. The essential thrust behind today’s Tele-Surgery is in the area of Minimally Invasive Surgery (MIS). MIS is essentially “Keyhole surgery, performed with instruments and viewing equipment inserted into the body, though a small incision created by the surgeon. The distance separating the surgeon and patient is only a few metres.

The da Vinci^{TM} Surgical System by Intuitive Surgical Inc. is a very recent entry into the world of MIS. The Stanford Research Institute at Stanford University developed the system. Intuitive Surgical brought it to commercial realisation (see Figure. 1). Already 11 complete systems have been deployed, 9 in Europe and 2 in the United States.

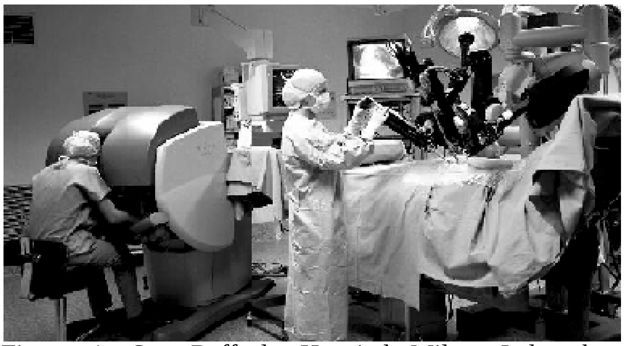

Figure 1. San Raffaele Hospital Milan, Italy photo courtesy of Intuitive Surgical Inc.

The da Vinci™ Surgical System comprises a surgeon's console, a patient-side cart, a 3D endoscope and an interchangeable gripper system. Feedback to the surgeon is primarily vision based and not feeling based (due to the small tissue forces involved). It is interesting that additional force feedback is provided to the surgeon by artificially generating forces at the handles, reproducing and magnifying the relative forces involved in operating the various grippers, scalpels and scissors.

This proven ability to use a computer console to perform operations from a distance (across the room) provides a solid foundation for the development of commercially viable long distance Tele-Surgery systems.

ZEUS^{TM} Robotic Surgical System by Computer Motion is another system (see Figure 2) (that is similar to the da Vinci™ Surgical System three-arm system.

Figure 2. ZEUS^{TM} Robotic Surgical System by Computer Motion

The ZEUS^{TM} system has yet to be deployed commercially. The status as of January 2001 was that Computer Motion had completed an FDA-approved Phase 1 Investigational Device Exemption (IDE) coronary bypass study. From there they had begun a pivotal, randomized multi-center coronary bypass study. In addition, two other multi-center studies had begun for general laproscopic and thoracoscopic clearance of the ZEUS^{TM} System.

5 Ongoing Tele-Surgery Research Activities

The commercial products like the few overviewed in the previous section give a basic indication of the enormous amount of research effort that has been ploughed into the area of Tele-Surgery. These efforts are still continuing at numerous laboratories scattered throughout the world today. A few of the key research efforts are mentioned below.

The Stanford Research Institute was instrumental in developing the technology which is now being utilized in the da Vinci™ Surgical.

NASA has an interest in delivering medical care in space. This is an essential component of successful space exploration. In the case of deep space exploration, the time lag will provide the greatest challenge in real-time interaction.

The delay is in the order of a fraction of a second to the orbiting International Space Station. This delay will increase the further from earth. Even at the speed of light NASA sees deep space real-time interaction as a serious challenge. Recognising this, they have opted for developing a virtual collaborative clinic as the first step. The concept is to have the crewmembers carry out the required emergency operation under the watchful “delayed” eye of medical experts back on earth.

6 Human Reaction Time

Delay in human perception is in the order of 100-200mS. The time required to knowledgably act upon a stimulus can be in the order of 250-600mS [Shumway-Cook and Woollacott, 1995]. This is assuming the teleoperator already had a firm grip on the controls. In fact delays of up to 800-1500mS have been recorded in some circumstances.

Wierwille, Casali, and Repa measured steering reaction times of drivers specifically focusing on the car ahead in a test track experiment when subject to abrupt-onset crosswinds in a moving-base driving simulator [Wierwille, Casali and Repa, 1983]. Reaction times in the order of 300 – 590mS were recorded.

It has been shown that an update rate in the order of 10Hz, is sufficient for humans to sense an accurate representation of the force incurred at the remote station, as well as not to subconsciously upset the tele-operator’s natural reflexes. Local controller updates as low as 5Hz have been shown to be adequate in accurately describing limb motion [Skubic and Volz, 2000].

To satisfy the aims of this project, a refresh rate at 10Hz will be stipulated as an initial starting point and then increased to 1KHz (limit of our present equipment). It is envisaged the capacity of the Internet today will limit the update rate to 250Hz.

7 Types of Communication Delays

7.1 Fixed Delay

Propagation delay is unavoidable in a tele-operated system (due to the limitations of the speed of electricity, the speed of light). Preliminary studies have shown that a transmission delay of 200mS or more leads to problems with surgical accuracy and precision.

The International Telecommunication Union Telecommunications Standardization Sector (ITU-T) noted that with voice calls, most callers notice round trip delays when they exceed 250mS. As a result the ITU-T G.114 recommend the maximum desired one way latency to achieve high quality voice is 150mS. With Round Trip Time (RTT) delays of 500mS or more, a ‘natural’ phone conversation becomes very difficult. These figures for voice traffic will be used as an initial starting guide for our implementation of timely and realistic force feedback.

7.2 Random Time Varying Delay

The Internet is a best effort service that offers no upper bound to response time or bandwidth guarantees. The result is a service that is time varying in a random nature. This fact introduces an extra level of complexity in the teleoperation of a system. A control engineer can deal with the problem of compensating of a constant delay with relative ease. However, a random time varying delay is very difficult to compensate for. Such a situation can often result in destabilising the overall system. The key to timely and stable control of a closed loop system over the Internet is to effectively reduce the variance of the delay.

The problem of controlling a real time tele-system using the Internet as the link has been studied extensively over the past few years. Most researchers have tended to use TCP/IP (with its inherent short comings in ability to deliver data in a timely fashion), seemingly without firstly looking deeply into the IP protocols options available. Generally this past research has concentrated on a variety of complex control methods in order help stabilise a telecontrol system in the presence of TCP/IP delay. Essentially the results have traded off a large amount of system response (delays of 5-6 seconds are not uncommon) in order to achieve stability.

8 TCP/IP

TCP is a connection orientated protocol. Possible congestion due to TCP traffic flows is controlled by the congestion control mechanism that is native to TCP. This congestion control can inflict serious problems on real time applications. In addition to this, TCP has an error correction arrangement in the forms of:

- Ordered delivery
- Duplication detection
- Crash recovery
- Retransmission strategy

By TCP addressing these above issues, TCP offers a guarantee for the reliable transport of packets to destination, thus, shielding the data users from the unreliable nature of the underlying IP network. The downside is the fact that these flow and error control techniques employed by TCP present a major obstacle to achieving time guarantees over the Internet.

For example, the TCP slow start mechanism is used to discover the channel throughput during the initial connection setup and for resumption of a broken connection. This is done by first sending a packet across the channel and waiting for a response. If a response is received, the next packet is sent a bit faster. This procedure is repeated until the speed of the link is discovered. With the half-second delay between responses, throughput is significantly slowed.

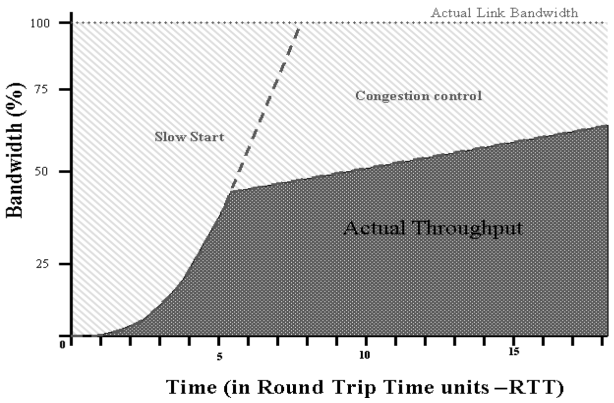

Figure 3. TCP/IP True Throughput

Since this process can take 7-15 Round Trip Times, for a link with a propagation delay of 500ms this can mean that for 3-7 seconds, the link is underutilised. (See figure 3). In situations where the data being sent is small, the entire transaction may be over before the optimum channel throughput has been determined [Allman, Hayes, Kruse and Ostermann, 1997]. This is often the situation with real time applications, where data is often sent as small samples or updates.

One must also consider that in the event of a packet loss, the data flow is halted until successful transmission of the packet concerned is achieved. In addition, when the data flow is resumed, it does so at lower flow rate than it was previously flowing at and then proceeds to slowly increase from this point.

Hence, if slow start could be somehow bypassed, significant gains in TCP/IP performance would be achieved. Pre-tuning the transmission throughput beforehand could allow for the ability to remove slow start. Even so, this method would only be suitable in controlled IP networks. A tremendous amount of research in this area is presently being under-taken by the world's leading satellite companies; it will be interesting to see the outcome.

To further add to overhead, every TCP connection is established by a "3-Way Handshake" between the Receiver and Sender. On links with long propagation delays, this fixed overhead means that even very short data exchanges take at least a few seconds to be completed.

Data links can be noisy, and this has profound effects on the performance of TCP/IP throughput because the slow start congestion control mechanism wrongly detects the noise as network congestion.

Hence, from a real time viewpoint, TCP fails to provide an adequate solution, largely due to the enormous processing overhead it employs in order to provide a reliable path for data.

9 UDP/IP and RTP/UDP/IP

Seemingly, none of the present online tele-operated systems to date have used Real Time Protocol (RTP) running over User Datagram Protocol/ Internet Protocol (UDP/IP). This is probably largely due to the fact UPD/IP is seen an unreliable data medium, whereby data could arrive out of order or not at all.

Even so, RTP/UDP/IP is fast becoming the popular protocol arrangement for streaming data in real time over the Internet. RTP was essentially designed to transport media stream that has strict timing requirements (i.e. voice and video). Every RTP packet contains a timestamp and a mapping between the timestamp and a globally synchronized clock, this mapping allows for the ability to merge multiple streams from different sources.

UDP is a connectionless protocol. This fact gives it very different characteristics to TCP. UDP is an unreliable service due to the fact that delivery and duplication of packets cannot be guaranteed. In addition it is likely that packets will arrive at the destination out of order. Even so, UDP with RTP is a far better option than TCP for real-time applications such as voice or video. Retransmission of a packet 1-2 seconds after it was sent when it contains a 20mS sample (as is the common case for voice) would produce disastrous implications to the real-time voice stream. In addition the cost in time for TCP to detect a packet loss, stop the data stream, request a resend from the point of loss and then finally receive the lost packet can be in the order of several seconds.

As stated, packet loss is unavoidable with UDP/IP, but it can be compensated for in voice streaming by codec loss-concealment schemes. One such codec is G.723.1, which has the ability to interpolate a lost frame by simulating the vocal characteristics of the previous frame and slowly damping the signal. It has been shown that packet loss rates up to the order of 10 percent have little noticeable impact on the audible quality of the speech.

It should be noted also that the connectionless quality of UDP/IP reduces the overhead of the protocol (from TCP/IP 40bytes to UDP/IP 28bytes) and this makes UDP a further preferred choice for constant flow applications such as multimedia and control sessions.

Even though a UDP/IP implementation has a lesser header overhead than that of TCP/IP, the RTP/UDP/IP implementation returns the header overhead back to 40bytes since the RTP component adds an additional 12bytes to the header.

Now 40-45 bytes of overhead would not be an issue if the data packet were in the order of 1500 bytes. The problem is that our implementation only involves packets with a data size in the order of 10-20bytes (due to the sampling rate). Hence a whopping total of 40-45bytes of overhead to transmit a 10-20byte payload.

There are two possible solutions to this problem:

- Increase packet size, at the expense of sample rate and potential delay jitter.
- Use header compression. In the case of voice packets it has been shown that the increased delay incurred from increasing the packet size is unacceptable. For this reason a great amount of research is being undertaken into optimizing header compression.

In summary, utilising UDP/IP in place of TCP/IP will greatly increase network efficacy by:

- Removing the need for having a connection setup before data can start to flow.
- Removing the slow ramping up of data throughput (due to TCP congestion control).

- Low rate packet loss does not halt transmission of the streaming data.

10 Network Response Times

A ping (Packet INternet Groper) routine, through the use of the Internet Control Message Protocol (ICMP) was used to determine the Round Trip Time (RTT) of various destinations on the Internet. This ping routine uses UDP packets. We can clearly see how the RTT is affected by the distance traversed, the size of the data being sent and size of the available bandwidth.

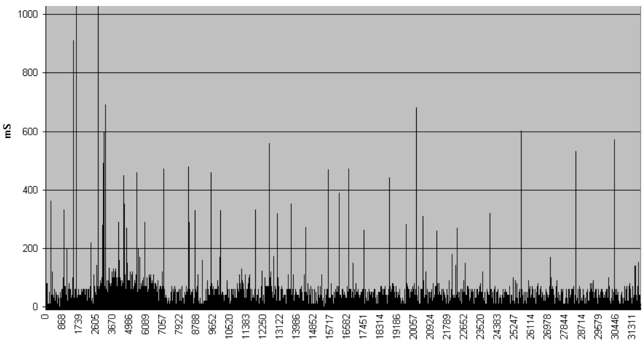

Figure 4. RTT delay between Wollongong (uow.edu.au) to Sydney (www.uts.edu.au) (90Km distance) 23/05/01-140601 Average ping time=13.03mS Packet size = 8bytes

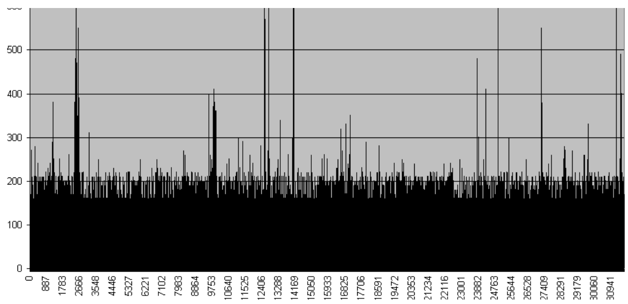

Figure 5. RTT delay between Wollongong, Australia (uow.edu.au) to Berkeley, USA, (www.berkeley.edu) 23/05/01-140601 Ave ping time= 159.97mS Packet size = 8bytes

The results shown in figures 4 & 5 are very encouraging. The measured RTT delay to transverse Wollongong to Sydney return is in the order of 13mS (Figure 4). The real delay could be much lower, due to the fact our measuring device incorrectly records all RTT times of 10mS or below as 10mS. The RTT delay for Wollongong to Berkeley return (figure 5) is in the order of 160mS. The instances where the delay suddenly spikes, will not be an issue as the RTP/UDP/IP protocol stack will override the delay, as opposed to TCP/IP.

It can be clearly seen from inspection of figure 6 that the overall size of the UDP packet affects the RTT delay. This would be a concern when sending large amounts of real time data (e.g. video), but for the case of force feedback, it is envisaged a data packet size of 8 bytes will suffice for carrying all the required information. For this reason the response times in graphs in figures 4 & 5 were measured with a packet size of 8 bytes.

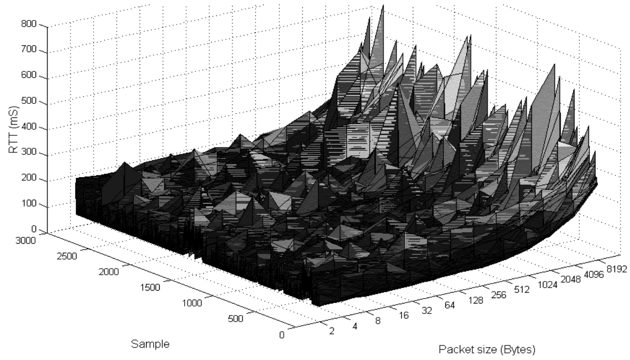

Figure 6. Wollongong, Australia (www.uow.edu.au) to New Zealand (www.Lincoln.ac.nz) 12/05/01

These results fall well inside the ITU-T G.114 recommendation of 150mS as the maximum desired one way latency to achieve high quality voice. As outlined earlier, these figures for voice traffic will be used as an initial starting guide for our implementation of timely and realistic force feedback.

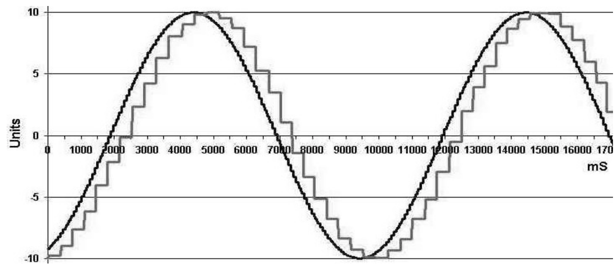

Figure 7. Wollongong, Australia to IRL Auckland, New Zealand 09090 return Average return time = 200mS via 56kps dial up modem from Australia

Figure 7 clearly illustrates UDP as the protocol of choice for real-time control; the received wave closely follows the transmitted sine wave with minimal fluctuations. The same sine wave received over TCP connection, utilizing the Internet as the connection medium is unrecognizable.

11 Experimental Set-up

We have developed a mobile robot whereby remote Internet users can experience collisions within a remote workspace located at the University of Wollongong. This is relatively simply implemented by use of a readily available “off the shelf” force feedback game joystick. While recognizing the limitations of the game joystick, the site creates the opportunity for typical Internet users to experience timely force feedback. Users feel the impact of driving the mobile robot into a wall of the maze (see figure 8 & 9), within 10-100mS of the collision occurring (delay being dependent on geographical location of user). This latency is comparable to human reaction time, so that its effects are insignificant to the operator. This site is not aimed to demonstrate surgical precision, it is to generate comments and useful data from users located all over the world.

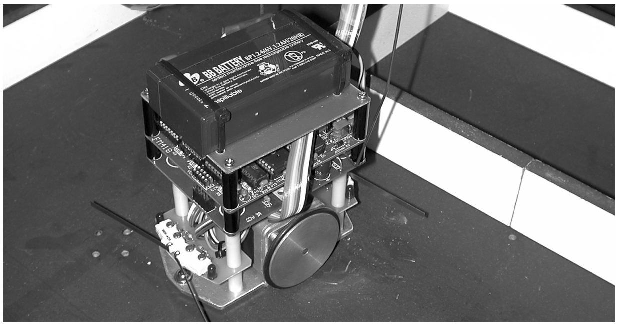

Figure 8. Collision feedback demonstration using a micro controlled mobile robot within a maze at University of Wollongong.

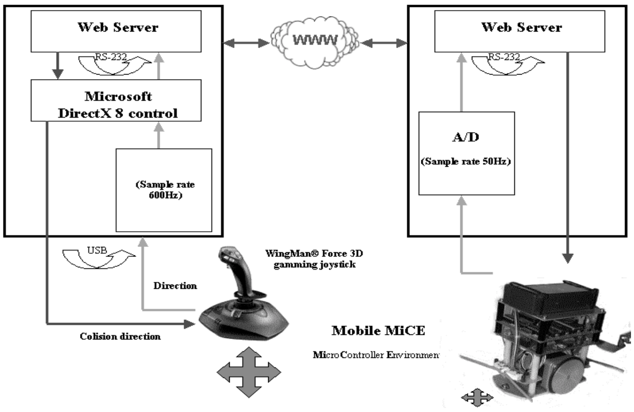

Figure 9. Block diagram of mobile force feedback demonstration.

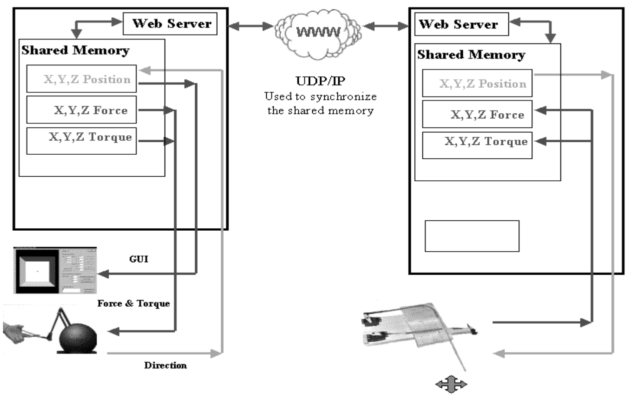

Figure 10 . Block diagram of Precision system, utiising the PHANToM system.

With the support of information obtained above, a precision system with up to 3 degrees of freedom available for the operator, is presently undergoing implementation utilizing a PHANToM haptic device for user input at the master station (see figure 10). Later a second PHANToM will act as actuator at the remote station.

12 The Next Steps in this Project

So far we have proven that it is practical to obtain 'intelligible' and timely control over the Internet. The next steps in the project is to further develop the communication link by more formally implementing the RTP/UDP/IP protocol stack, and then implementing a CODEC on the client side for the handling of lost or timely packets. A buffer to overcome small amounts of packet jitter will be explored, to see if any sizable gains warrant the approximate 50mS overhead required for implementation. Finally the wave variable technique developed by Niemeyer and Slotine will be implemented to guarantee overall system stability [Niemeyer and Slotine, 1998]. Possible extensions on this technique will be investigated, to determine if further improvement of the performance can be achieved.

13. Conclusions

This paper described the Tele-Surgery concept and its possible implementation using the Internet as the medium via which to transport intelligible and timely force feedback for Tele-Surgery. A review of the present online telerobotic sites and present Tele-Surgical units around the world was presented, and their relevance to Tele-Surgery over the Internet discussed. Ways of implementing realistic control, through the use of Internet protocols orientated to real time traffic flow were described.

Experimental work involving evaluation of force control and feedback over the Internet was described.

The next steps in this project will be to ensure system stability regardless of the performance available at the time from the Internet. Expand the number of degrees of freedom available This should eventually enable a convincing demonstration of the Force control over the Internet.

References

[Tsumaki, Goshozono, Abe, Uchiyama, 2000] Y. Tsumaki, T. Goshozono, K. Abe, M. Uchiyama 'Verification of an Advanced Space Teleoperation System Using the Internet', *'IEEE/RSJ International Conference on Intelligent Robots and Systems'*, 2000

[Anderson and Spong, 1988] R. Anderson and M. Spong, 'Bilateral Control of Teleoperators with Time Delay', *'Decision and Control, 1988., Proceedings of the 27th IEEE Conference'*, 7-9 Dec. 1988, Pages 167 - 173

[Shumway-Cook and Woollacott, 1995] A. Shumway-Cook, M. Woollacott, *Motor Control Theory and Practical Applications*, Williams &Wilkins, 1995

[Wierwille, Casali and Repa, 1983] W. Wierwille, J. Casali, B. Repa, 'Driver steering reaction time to abrupt-onset crosswinds, as measured in a moving-base simulator' Human Factors 25:103-116; 1983.

[Skubic and Volz, 2000] M. Skubic, A. Volz, 'Acquiring robust, force-based assembly skills from human demonstration', *'Robotics and Automation, IEEE Transactions'*, Dec. 2000

[Allman, Hayes, Kruse and Ostermann, 1997] M. Allman, C. Hayes, H. Kruse, S. Ostermann, 'TCP Performance Over Satellite Links', *'The Fifth International Conference on Telecommunications Systems, Nashville, TN'*, March, 1997

[Niemeyer and Slotine, 1998] G. Niemeyer, E. Slotine, 'Towards Force-Reflecting Teleoperation over the Internet', *'Robotics and Automation, 1998.*