Building Tele-Presence Framework for Performing Robotic Surgical Procedures

Peter B. Panfilov, Frank M. Cardullo, and Harold W. Lewis, III
Moscow State Institute of Electronics and Mathematics, SUNY at Binghamton, SUNY at Binghamton
{panfilov@miem.edu.ru, cardullo@binghamton.edu, hlewis@binghamton.edu}

Abstract

A research project is being presented aimed at the development of a system of methods that in the future will make it both safe and practical to perform telerobotic surgery. To mitigate the effects of the physically unavoidable delay we propose a telepresence system framework, which is to have the surgeon operate through a simulator running in real-time enhanced with an intelligent controller to provide the safety and efficiency of an operation. Three major research areas must be explored in order to ensure achieving the objectives of our project. They are: simulator as a predictor, image processing, and intelligent control. Each is equally necessary for success of the project. These are diverse, interdisciplinary areas of investigation, thereby requiring a highly coordinated effort by all the members of development team, to ensure an integrated system.

1. Introduction

Robotics is one of key technologies that have a strong potential to change how we live in the 21st century. We have already seen robots exploring surfaces of distant planets and the depths of the ocean, streamlining and speeding up the assembly lines in manufacturing industry. Robotic vacuum cleaners, lawn movers and even pets found their ways to our houses. Among the medical applications of robotics the minimally invasive surgery was the first to demonstrate a real advantages and benefits of introducing robotic devices into operating room over conventional surgical methods. So far, these machines have been used to position an endoscope, perform gallbladder surgery and correct gastroesophogeal reflux and heartburn.

It is interesting to note that in the late 1980s, after its inception the utilization of laparoscopic cholecystectomy grew rapidly. However, minimally invasive surgery (MIS) for other operations has not experienced the same pattern of growth. According to Ballantyne [1], the reason is that in general laparoscopic procedures are hard to learn, perform and master. This is a consequence of the fact that the camera platform is unstable, the instruments have a restrictive number of degrees of freedom and the imagery presented to the surgeon does not offer sufficient depth information. The solution seems to be at hand with the significant growth of robotic surgery. This is surgery where-in the surgeon operates through a robot. In a sense this robot is a telemanipulator under the control of the surgeon. The robotic system provides a stable video platform, added dexterity and in some cases a stereoscopic view of the surgical field.

The surgical robots use technology that allows the human surgeon to get closer to the surgical site than human vision will allow, and work at a smaller scale than conventional surgery permits. The robotic surgery system (e.g., the daVinci robot) consists of two primary components:

- A viewing and control console (operating station)
- A surgical (robotic) arm unit

Sitting at the control console, a few feet from the operating table, the surgeon looks into a viewfinder to examine the 3D images being sent by the camera inside the patient. The images show the surgical site and the two surgical instruments mounted on the tips of the rods. Joystick-like controls, located just underneath the screen, are used by the surgeon to manipulate the surgical instruments. Each time one of the joysticks is moved, a computer sends an electronic signal to one of the instruments, which moves in sync with the movements of the surgeon's hands.

Since proximal robotic surgery seems to be maturing the next logical step in surgical care is to extend to remote applications of robotic surgery. That is to say, the surgeon and the operating console are at one location and the robot and patient at another. The idea of remote robotic surgery, or as some refer to it, telesurgery, has been an objective for some time, especially in the military. This advancement is seen by the military as the means by which the next major improvement in battlefield survivability [2]. In addition to the military application, the technology could be useful if an astronaut were to require emergency surgery while on the space station. Furthermore, perhaps the most ubiquitous application will be in civilian medicine. Patients in medically remote areas would have the option of receiving an operation performed by a renowned surgeon even though the surgeon and patient may be thousands of miles apart.

The long term objective of this research is to develop a framework for remote robotic surgery application which will permit surgery between any two places on earth with a patient in one location and the surgeon in another. The major impediment to remote surgery is the effect of telecommunication delay on the surgeon’s performance. It has been shown in a myriad of studies of human-in-the-loop systems that system delays lead to degraded operator performance and ultimately unstable systems. Since the delay cannot be eliminated, in order to accomplish this objective, the only solution is to mitigate the effects of the delay on the surgeon performing the operation. Relying on many years of experience in human in the loop simulators, including a simulator to train surgeons in laparoscopic cholecystectomy, we propose a method whereby the surgeon operates through a simulator in a virtual environment free from the impediments of telecommunication delay. Since, in robotic surgery, the surgeon is already in a synthetic environment the introduction of a simulator does not significantly alter the physician's perceptual stimuli. The operating station containing the control inceptors and the visual displays is the same as that used to control the surgical robot in the conventional configuration. For robotic surgery systems, one can often use computing power to create novel interfaces such as virtual environment interfaces that enable new paradigms of instrument operation and data visualization.

These systems are inherently distributed systems, often with the robot, an interface generation computer, and special purpose interface devices such as haptic force-feedback devices and head-mounted displays distributed on a local area network in a single room or building. This simple distribution naturally motivates a larger scale distribution of system components across the wide area and/or global networks such as the Internet to enable telepresence experience and teleoperation of the special purpose medical equipment. The result is a distributed virtual operating room consisting of a geographically separated collection of medical instruments, computers, and physicians all interconnected via a computer network.

Most researchers in robotics area now accept the definition of telepresence put forward by Thomas Sheridan [3] as the sense of actually being at a remote or synthetic workplace which users of telerobot or virtual environment systems developed during operation of the system's human interface. Since users of a virtual environment interface are in the same position with respect to simulated effectors in the virtual environment as that of human telerobotics controllers with respect to a remote robot, the two control situations also may be generalized under the term virtual environment (VE). This generalization is reasonable since with respect to the various users' physical viewpoints in either case user experiences *presence in an environment by means of a communication medium.*

To qualify the operators' sense of remote presence during teleoperation or use of virtual environment interfaces as an explanatory scientific concept, Prof. Sheridan identified three determinants of presence [3]: 1) the extent and fidelity of the sensory information that may be displayed to the users, 2) the extent and fidelity of the users' control over the sensory information and 3) the extent and fidelity of the users control of effectors in the environment. These three components suggest measurement by well-established continuous variables associated with physical sensors and effectors, e.g. bandwidth, latency, dynamic range etc.

As emphasized by Stephen Ellis from NASA Ames in [4], presence or equivalently simulation realism is particularly useful as it identifies reasonably independent, easily measurable characteristics of communication channels. In terms of Sheridan's analysis of presence, for example, the fidelity of sensory display could be instantiated as a bandwidth, x_1, or update rate of a visual display. Similarly, the bandwidth or update rate of the control of a virtual effector could provide another potentially continuous variable, x_2, influencing a measure of presence. These signal characteristics are especially useful since they can be considered independently of the users' ability to modify them through control of sensor positions. One may additionally easily imagine that improvements in any of them would lead to improved operator performance, improved realism, and an increased sense of presence within the virtual environment. In fact, with respect to manual control the dynamic requirements to communication channel are classically known to determine the fidelity of control [4].

The proposed surgical telerobotics framework challenges the problem of communication channels' characteristic improvement through a complex combination of systems science, optimal control, and intelligent systems approaches. In our embodiment the simulator acts as a predictor, illustrating for the surgeon what would be the case if there were no delay, and – somewhat simplistically – the image preprocessor at the robot side can be thought of as a corrector. The intelligent controller is designed as an optimizer. Fig. 1 is a simplified view of the proposed framework and the general research areas. The following paragraphs explain the architecture.

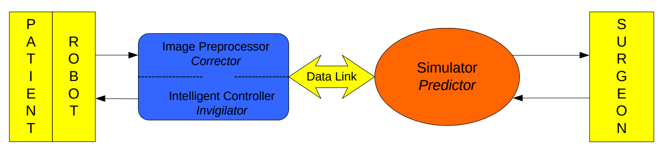

Figure 8 Surgical telerobotics as a tele-presence system

Lombard and his colleagues define presence as "the perceptional illusion of nonmediation" [5], which is the consequence of transparent mediating systems. By "illusion of nonmediation," they refer to a phenomenon in which "… a person fails to perceive or acknowledge the existence of a medium in his or her communication environment and responds as he or she would if the medium were not there" [5]. The three components of the framework working in concert should ultimately constitute a communication medium which is transparent enough to bring the human operator as close as possible to the telepresence "**ideal** of sensing sufficient information about the teleoperator and task environment, and communicating this to the human operator in a sufficiently natural way, that the operator feels physically present at the remote site."[6]

2. The real-time simulator as predictor

Modern simulators tend to be very complex systems in their own right, but one particular aspect of how the simulator will function in this case is where we place the emphasis in calling our simulator a predictor. That is, in order for our time-delay-mitigation scheme to work, the simulator must predict what will happen in the surgical field before it happens. Furthermore, the predictive mechanisms in this case are based on dynamic modeling of a far from trivial sort. Adding more complexity to the task, the system must be designed to allow the models to be updated in real time as the delayed information from the surgical field becomes available. Clearly, dynamics models both for the robot dynamics and organ dynamics are necessary for the simulator to function in this way. Though both are challenging, the organ dynamics modeling is known in medical research circles to be extraordinarily difficult, particularly in the case of soft tissue. For this, we intend to experiment with a variety of approaches, which will include both finite element analysis and continuum analytical models.

The simulator as used here is clearly a perfect example of an anticipatory system. Note that anticipatory behavior is often viewed in the literature [8, 9] as a primary characteristic in intelligent systems.

2.1. The time delay problem in tele-robotic surgery

It is well known that system delays will cause a deterioration of the human-machine system performance. As a matter of fact this is true for any control system, not only a human-in-the-loop control system. Fig. 2 illustrates the time domain effect of delays of 0, 200, 400, and 800 ms where the input is a unit step [7].

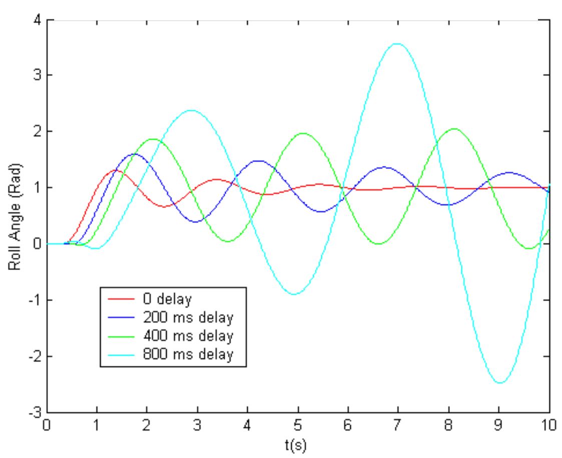

Figure 2 Step responses of a closed loop system with different delays

The graph indicates that as the delay increases, the response lags the input by a greater amount. In addition the 400 ms delay case seems to display limited stability, while the 800 ms delay case clearly exhibits unstable response. The system analyzed includes a fourth order plant and a human operator model, to which the delays are added.

Fig. 3 presents the results of frequency domain analysis of the same system. Here, one observes that the 400 ms delay case yields a phase margin of approximately zero, while the 800 ms case has a negative phase margin. We can then examine human operator performance data in a system with and without delays. There are many such examples in the literature. It is observed that when delays become long, human operators will adopt a move and wait strategy. This allows the operator to observe the results of his/her action before committing to another action. The move and wait strategy may be acceptable for controlling a lunar or Mars rover but it is unacceptable in tightly closed loop applications, robotic surgery being one of them.

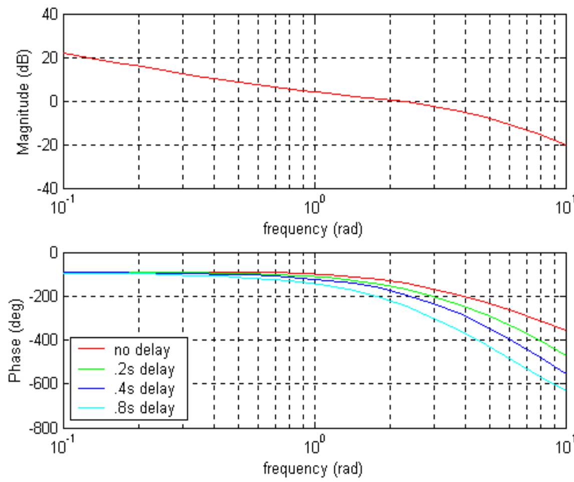

Figure 3 Frequency responses (Bode diagrams) of a closed loop system with different delays

When the delays in an aircraft flight control system become too long the control loop becomes unstable and the aircraft is said to display pilot induced oscillation. This is another case where the move and wait strategy will not work.

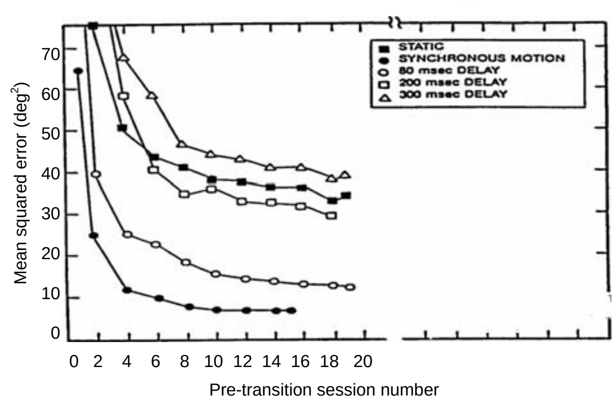

Figure 4 Effect of delay on a system operator

performing a tracking task with/without force feedback

Fig. 4 illustrates the effect of delay on a system operator performing a tracking task with and without force feedback. There were several cases of delay (0, 80, 200 and 300 ms) in the force feedback. In all cases the subjects had a narrow field of view visual presentation. The graph shows that at 200 and 300 ms delay in the force feedback the operator's performance is essentially as bad or worse as with no force feedback [10]. Whereas, at a delay of 80 ms his/her performance is much improved and almost as good as a fully synchronous feedback.

Preliminary studies were conducted using experienced laparoscopists in a suture knot tying task. The task was performed using a laparoscopic training device with delays introduced, in 25 ms intervals, into the video monitor via an analog delay device from Prime Image. The performance metric used in this study was the time it took the subjects to complete the knot. Fig. 5 illustrates the results. For delays up to about 100 ms the execution time remained relatively constant at about 13 seconds. Above the 100 ms point the time increases substantially. Preliminary results seem to indicate that by the time the transport delay approaches 500 ms the time to complete the knot is about 90 seconds. One of the interesting results is that subjects began to experience nausea at delays approaching one second. This result was quite unexpected based on our considerable experience with simulator sickness.

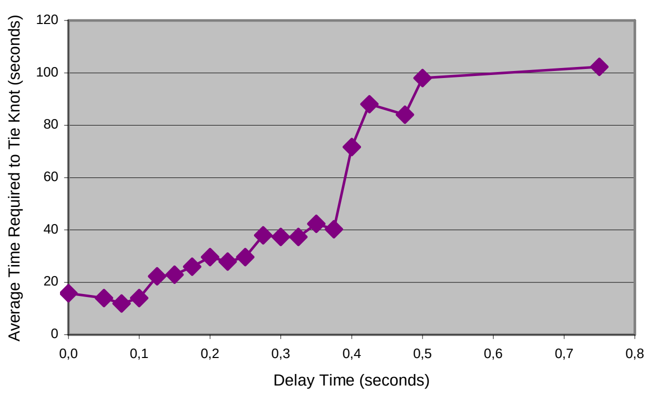

Figure 5 Knot tie time as a function of delay time

2.2. Organ dynamics modeling

Surgery simulation involves biological tissues, it is therefore essential to determine the deformation of the tissue while in contact with a surgical instrument. This is important to transfer a realistic picture of the tissue to the surgeon in real time. The problem is how to determine the interaction between the tool and the tissue. This requires a rigorous model of the material properties, an accurate simulation method to reflect the actual behavior of the tissue, and fast simulation results to enable real-time interactive simulation.

The organ dynamics model must be accurate, timely, and responsive. For the success of this telerobotic system, realistic visual deformation and haptic output of the simulator is essential. The surgeon must see and – it is desirable to feel the organs presented by the simulator as if they were the actual organs of the patient. As an aside it should be noted that the daVinci robot does not provide significant, haptic feel of the tissue. However, as part of this project we will add this capability, since haptic feedback has been demonstrated to mitigate some of the effects of delay by virtue of the fact that the proprioceptive system is faster responding than vision. These simulator outputs depend on the organ dynamics models. The organ dynamics models are based on solid models of the components of the biomechanical system upon which the surgery is being performed, - in this case soft tissues.

The first requirement of the organ dynamics model of the biological system is accuracy: the deformation and force feedback of the simulated organ on which the surgeon is operating must be as close as possible to the actual patient organ. Secondly, the computation that is required to determine the simulated deformation and forces must be performed in a time interval that appears to the surgeon to be real-time. Also, the organ dynamics model must be capable of being quickly updated using the decompressed video feedback that provides corrections without creating visual or haptic discontinuities, "jumps", that could disturb the surgeon. In other words, the organ dynamics model must be responsive to the correcting feedback.

The organ dynamics model employed by the simulator will respond with force and displacement feedback to the surgeon to mimic the actual deformation of the structure, e.g., the organ, at the remote location. The simulation of the deformation allows the surgeon to pursue the surgery on the deformed geometry. This overcomes the delay associated with the transmission from the actual remote operating site. The response of the structure to the actual robotic surgery at the patient end will be used to correct the model, but the corrections must be kept small. A computational model of the dynamic response of the structure to the surgery is needed. The final trade-offs in the design of the model will be between these three constraints: accuracy, timeliness, and responsiveness.

2.2.1. Elements of organ dynamics model. There are three elements of the organ dynamics model: (1) the structure (geometry) of the organ, (2) constitutive material properties of the soft tissues, and (3) the applied forces.

Although a generic base-model of human organs can be a starting point for use in the simulator, the geometry of the organ that is used by the simulator during surgery cannot be a generic model. It must be patient-specific. The geometry of the patient-specific organ is needed to develop an accurate organ model.

There are several techniques that can be used to determine the geometry of an individual organ. Among these are x-ray, magnetic resonance (MR), computed tomography (CT). Chen, et al. [11], have used MR images to model knees. These were analyzed and mapped to a known geometric (cylindrical) kinematic model. Subsequently, a 3D virtual model of the knee was constructed [12]. Gibson, et al. [13], have also taken this approach. In Gibson's research, construction from MRI involves an analyst's visual examination of the segmented images to distinguish tissue intensities. The creation of the geometric model of the organ will require imaging of the patient prior to surgery and could require manual or programmed image analysis.

In addition to the geometry of the structure, constitutive material properties are needed for each type of tissue present: ligament, cartilage, other soft tissues and adjacent bone. To simulate the results from the surgical procedure, the following are needed: (1) the amount of deformation, (2) the force exerted on the surgical instrument, and (3) the resilience of the material. The mechanical properties of biological material are inherently non-linear, inhomogeneous, and anisotropic, for example, soft tissues, such as ligaments, have viscoelastic properties. Bulk, ex vivo material properties are available from the general literature [14], but these can vary significantly from in vivo properties.

Indentation techniques have been used to determine in vivo material properties. Ottensmeyer [15] describes an instrument (TeMPeST 1-D) that was used to determine the viscoelastic properties of porcine liver. The device generates small amplitude oscillations and the force response is recorded. An elastic modulus and viscoelastic stress-strain equation were determined. The indentation technique has been also used by Tay, et al. [16] to determine the force response of pig liver and esophagus tissue. A force vs. displacement plot for in vivo pig esophagus tissue is presented. Material property data can be determined using this force-displacement curve. The apparatus that was used consisted of an indenter with force sensor that was fixed to a programmable force feedback haptic device (Phantom Premium-T 1.0). In addition, the indentation method is employed by Samosky, Grimson and Gray [17] to determine cartilage stiffness.

Vuskovi, et al. [18] used a tissue aspiration method that is a non-contact technique. "The measurement method consists in leaning a tube against the target tissue and gradually reducing the pressure in the tube." The profile of the tissue in the tube is measured. The instrument consists of a tube, mirror and lens with light provided by an optical fiber, and pressure sensor. The acquired data is then fit to a finite element model to determine the material constants. A technique for directly measuring in vivo material properties has been developed by Brouwer, et al. [19]. A device grasps a portion of tissue and stretches and compresses the tissue recording force and displacement. They have successfully applied the technique to determine porcine tissue properties, but this technique does not appear to be transferable to use on humans.

For the simulation, a data acquisition system is needed to determine the forces that are applied to the structure of interest. Forces and deformations have been measured using surgical instruments that are fitted with six-axis force transducers [19]. These transducers have been used to examine tissue by mounting them to a probe. As the probe is moved to selected points on the surface of the tissue (e.g., cartilage), probe tip position and force are recorded [17].

2.2.2. Objectives and scope of organ dynamics modeling research in the current project. Four major goals for the tissue modeling in the project define the scope of the research.

First, the creation of the geometric dynamic model of the organ will require determining material properties for the various types of tissue involved and the loads applied to the tissue due to the surgical instruments. This includes 1) investigation and analysis of the indentation and aspiration techniques, 2) development of a method to determine the required material properties, and 3) conducting experiments to measure forces applied during surgery.

Second, the major objective of the organ model is to determine the change in topology of the organ when it is impacted by a surgical instrument. The topology of the organ is defined by the finite element (FE) model meshes. The location of each node in space is calculated and made available to the image generator. The image generator will then produce the graphical primitives based on the location of the nodes. It is common for collision detection to be carried out in the image generator and it is anticipated that this will be the case here since the image generator is rendering not only the tissue but the surgical implement from the perspective of the video camera.

Third, to determine in detail how to integrate the organ dynamics model most effectively into the overall system, it is necessary to determine and to demonstrate methods for updating the organ dynamics model. As shown on the system block diagram in the figure 6 there are two inputs to the organ dynamics model: surgeon input and image feedback. The direct surgeon input is the motion of the inceptors. This is translated into the motion of the robot effectors through the robot dynamic model resulting in the actions that are applied to the organ dynamics model. The second input, the image feedback, is the decompressed and preprocessed image from the patient end. This image is used to update the state of the organ dynamics model so that it does not deviate from the actual organ.

Fourth, to optimize computational efficiency, several techniques will be employed. A reduced order FE models will be developed. This factor alone will save substantial computation time. We will also create zones around the surgical probe, once it is in contact with tissue. The zone closest to the probe will require the finest mesh for the finite element model and the highest computation rate. The more remote zones will have coarser meshes and slower computation rate because they will be less impacted by probe contact. We anticipate at least three zones. It may also be possible to use lower order models in these outer zones.

2.3. Real-time realistic visual simulation

Visual simulation has become a commonly employed method to provide the environment in which to train surgeons to perform common, yet risky, medical procedures. Typical examples are laparoscopic cholecystectomy and gynecological sterilization. Today the virtual environment comprises a 3D photo-textured polygonal database representing the anatomy of the critical internal organs involved in the procedure.

2.3.1. Dynamic database alteration. The majority of any given surgery is often spent cutting, cauterizing, and removing the fatty and connective tissue that obscures and interferes with the areas of interest. Often the areas of interest can be difficult to differentiate from the surrounding tissue, especially when considering the challenge of spatial cognition. For example, in a cholecystectomy (removal of the gall bladder), the main challenge is in distinguishing the cystic duct from the common bile duct and the hepatic artery. Since the purpose of this research is to ultimately perform live surgery, our objective for free-form volumetric dynamics is to realistically simulate the physical reaction of tissue due to actions such as cutting and tearing as would always be found in a live situation. This includes the simulation of different kinds of tissue ranging from organ tissue to fatty and connective tissue, as well as striations in the tissue. The internal representation used to model tissue must allow for free-form alterations including cutting and tearing, and will provide realistic force-feedback. To simulate tissue with this level of fidelity requires a new approach that goes beyond the limitations of current approaches.

Our initial research is based on the use of free-form iso-spheres or a high-order surface (HOS), with the use of a space-filling Union of Spheres (UoS) as the underlying physical model for tissue. With UoS, objects are modeled using collections of packed spheres. Each sphere can be seen as an independent object with forces acting on them computed by the organ dynamics model.

A major advantage to using UoS over traditional approaches is that the generation of UoS models from medical imagery sources such as MRI or CAT scan data is straightforward. This allows for the possibility of "mission rehearsal" on actual patient data prior to the surgery, and also reduces the challenges and costs of generating realistic anatomical models, and is a perfect fit to include patient-specific data such as MRI or CAT scan information for the actual surgery. Another advantage of a UoS approach is that collision detection becomes a trivial problem. The intersection of an object (such as a surgical instrument) with a sphere is a simple calculation because inside/outside information is readily available. Also, because the spheres have a known spatial location and extent, only a small subset of spheres need be compared to any given object.

The dynamics among neighboring spheres will be governed by the state vector of each sphere calculated by the organ dynamics model.

Associated with each sphere will be numerous properties including a flag indicating if a sphere is an interior or surface sphere. Initially, the minimum number of spheres needed to represent the anatomy will be used. This will reduce both the memory and computational burdens of the force model. At runtime, spheres will have the ability to subdivide and merge, depending on the complexity required to represent a given modification. This will allow the force model to be scalable, capable of automatically taking advantage of future hardware and software improvements. The maximum level of subdivision will be limited by the available computational resources, ensuring the highest possible fidelity at interactive rates.

2.3.2. Spatial database structure. Many of the algorithms that will be used in our approach require data to be efficiently accessible based on spatial location. The algorithms that benefit from exploiting the spatial organization of data include nearest neighbor searching for the propagation of force through the union of spheres and collision detection between spheres and other objects. The visualization system will also require spatial efficiency in order to perform visibility determination to reduce the rendering burden as well as texture manipulation and synthesis for additional realism.

The challenge that arises when using an UoS approach is how to derive the structure of the control points and the texture coordinate values for the generated surface points. One possibility for deriving control point values is that the number and relationships of control points is fixed, but their values can vary. This is similar to how traditional SMD systems represent the underlying physical model of an object. Our goal, however, is to visualize free-form deformation of an isosurface physical model. This requires a more flexible approach to rendering dynamic high order surfaces.

2.3.3. Texture synthesis and mapping. The mapping of textures onto high order surfaces has been a long-standing challenge to computer graphics. Because the surface is procedurally generated, texture coordinates are not easily defined for areas between control points. To address this challenge, a second high order surface can be generated specifically for the computation of texture coordinates. Traditionally, this method incurs substantial overhead. Recent advancements in hardware-accelerated texture techniques make the texturing of high order surfaces feasible for real-time. These advancements include procedural texture coordinate generation and support for 3D textures. 3D textures provide an additional depth component that allows for effects such as volumetric obscuration of the image, impostering, functional lookup, and procedural texture generation and noise. With a 3D texture, the internal volume of an object can have associated texture information. This will allow for the accurate texture mapping of newly introduced surfaces resulting from cutting, tearing, cauterizing, etc.

Texture synthesis will be used to take source imagery from surgical videos and photographs and generate realistic textures for applying to the simulated tissue in the visual scene. Without texture synthesis, objects look artificial because of the noticeable repetition or stretching in the image. Texture synthesis provides a more natural looking result and can be generated at run-time. It will also be used to manipulate textures for representing tissue discoloration due to tearing, cauterizing, pathologies, etc. Texture synthesis will be used to insert the revised texture information from the patient during information transfer. This process will be continued in the real-time process, using the extracted texture from the ongoing patient video stream.

3. Image preprocessor

An image preprocessor module on the patient side of the communications link is essential to the functioning of the overall system. The purpose of this component is to recognize the organs and various other objects in the surgical field from the video imagery coming from the surgical camera. It will feed this information on the location and states of these objects both to the intelligent controller (on the patient side) and to the simulator (on the surgeon side). This image understanding task encompasses the processing, encoding, and recognition tasks that will be accomplished on the patient side, using the video from the surgical camera(s). The problem of image identification has been a focus of research in the community of image understanding for a long time, and the results still seem to indicate that the most effective approach depends on the context of the application.

A new approach to object recognition and categorization in image is under development by members of our team. A camera image is basically a 2D projection of a 3D world in the field of view of the camera. The approach is to reverse this process by determining the 3D world that generated the 2D image. This approach has been successful in extracting 3D objects, including buildings and trees. Although buildings are generally rectilinear, the objects extracted can be of an arbitrary shape. The software to do this in real time has been developed to enhance 2D images, detect specified objects, extract 3D objects, and remove others from the 2D image.

The extraction process is done in two steps. First, the object is detected or recognized. Second, the 3D parameters are extracted. The method being proposed is geometry-based in that it will use a library of 3D models that can describe what is being seen in the camera (organs, veins, etc.). These models are not rigid models, but rather are models that can change shape depending upon specified parameters. These models are likely to be patient specific, having been generated by an offline process prior to the surgery. The method decomposes the 2D image into regions. Regions are contiguous sets of pixels that are determined to belong to a specified set based on any arbitrary criteria. Examples of these criteria are color or texture. All objects in the 2D scene are composed of one or more of these regions. The regions in the image are then analyzed to determine the boundaries of the 2D projected objects, with the result being input to algorithms for determining the depth component, defined as the axis normal to the plane of the camera. This component for any surface is determined using calculations based on color, texture, and diffuse and specular reflection from the surface. The edges of the 2D objects are found with pattern searches and correlated with the library model data to ascertain the actual 3D structure at the time the 2D image was obtained. The accuracy of the resulting 3D model will depend upon a number of factors, including the resolution of the camera, whether the image is in color or not, the quality of illumination, and the availability of accurate models of the objects being displayed.

4. The intelligent controller

This device, located on the robot/patient side of the communication link, performs in two critical roles. In the ultimate system for use on actual patients, the intelligent controller will be necessary to provide both an added measure of safety and an improved level of efficiency in the presence of time delay. Both the safety role and the efficiency-enhancement role require intelligent behavior.

The need for an added element of safety in the presence of time delay is quite obvious. For a variety of reasons, even when the surgeon as well as the various other components of the system are performing perfectly, the existence of time delay prevents the possibility of 100% certainty as to where various tissue will be in relation to the surgical instruments at any given instant in the future. Because the intelligent controller will be proximate to (and linked directly to) the robot, it will interact with the robot without significant delay, and thereby has the potential to control all robot movements instantaneously. Thus, as a last line of defense against the possibility of accidental collisions between surgical instruments and the patient’s vital organs, the intelligent controller will ultimately play a critical role.

The need for improving the level of efficiency over what it would otherwise be in the presence of time delay is also clear. Finishing surgery in a timely manner and preventing unnecessary frustration for the surgeon are always important goals. While it may be true that the delays associated with telerobotic surgery will never allow it to be quite as efficient as proximate robotic surgery, the goal at least must be to make it ultimately as efficient as possible.

Although in the course of the research, we will attempt to apply a variety of advanced approaches to machine intelligence in designing effective intelligent controller prototypes, some basic aspects need not be particularly complex. For the safety aspect, a fairly effective controller could be based on nothing more than a three-dimensional geometric model of the surgical field combined with a simple type of production rule system. A typical rule for the case of gall bladder surgery might look roughly like the following:

IF left end-effector holds sharp instrument AND instrument is within 5 mm of common bile duct AND an override command has not just been submitted by the surgeon
THEN stop movement of left end-effector immediately, send safety alarm signal to surgeon, wait for reset by surgeon

The production rule system for the case of gall bladder surgery may be comprised of perhaps a few dozen such rules. This is a very basic form of the traditional approach to artificial intelligence. Using this as a starting point, we can readily add more sophisticated machine intelligence approaches.

One fairly straightforward addition would be the use of fuzzy logic (FL) in the production rules. FL is simply a calculus for representing mathematically the way humans use somewhat vague concepts, such as “very close” or “rapidly”, when reasoning about complex systems. For example the rule above could be made more sophisticated by changing the antecedent to the following:

IF left end-effector holds sharp instrument AND (instrument is very close to common bile duct OR (instrument is somewhat close to common bile duct AND left end-effector is moving rapidly)) AND an override command has not just been submitted by the surgeon

Naturally, when using fuzzy logic, it will also be necessary to represent in the system those numerical values associated with such terms as “very close”, “somewhat close”, and “rapidly”. It is quite simple to represent such terms in a particular context for relative locations, velocities, accelerations, and any other types of variables we may use.

Fuzzy logic is one component of what is now known widely as “the soft computing approach”. Soft computing (SC) is a term coined by Lotfi Zadeh around 1990 to represent the emerging trend to design complex systems based on hybrids of four component methodologies, each of which had been evolving over the previous three or four decades [20, 21]. These component methodologies of SC are referred to most generically as fuzzy logic (FL), artificial neural networks (ANN), evolutionary computing (EC), and probabilistic reasoning (PR). The key concept of this approach is not just that these four methodologies tend to be powerful in and of themselves, but rather that there tends to be a synergistic effect when two or more of them are combined in appropriate hybrids. The SC approach, also referred to as computational intelligence, may seem to the layman a bit like science fiction, but it is a successful and well established amalgam of methodologies in some fields of engineering and is based ultimately on decades of advanced research.

The success of SC has been demonstrated most graphically in the context of feedback control, particularly in inherently complex control applications. There have been very many citations in the literature of the successful application of SC hybrids, including for example in [22]. Certainly we will experiment with applying them here as well, and we can already state with confidence that they will be useful in the context of the intelligent controller.

We will also experiment with another approach known in general as optimal control techniques, which use a model reference approach. In this case a model of the entire system; patient, robot and surgeon is employed along with a cost function which will be minimized to determine the coefficients of the parameters in the control laws. This need not be viewed as an entirely separate approach in the sense that optimal control concepts are often part of the SC methodology.

5. The integrated tele-robotic surgery system

The total system behaves as the human surgeon would if there were not a performance encumbering delay. Because the simulator through which the surgeon operates is running in real time the surgeon sees reaction to inceptor movements much more quickly than would be the case if he/she were required to wait while the signals made a complete round trip over the long haul network. Fig. 6 is a detailed representation of the proposed system and the general research areas. The following paragraphs briefly explain the architecture.

The signal path from the surgeon’s inceptor movement proceeds to the simulator and simultaneously to the intelligent controller, which commands the robot movement. This simulator is like any other in that it calculates all of the system dynamics in real time and from these computations come changes to the system states, which alter the visual scene observed by the physician. The visual scene is generated by high speed computer graphics engines not unlike those employed by modern flight simulators. However, a unique aspect of the proposed embodiment is that the graphics image is periodically updated by the video image transmitted over the long haul network. This approach ensures that the visual scenes at the patient and at the simulator are never allowed to deviate perceptibly. This update is generated by a complex scheme of image decoding, texture extraction and image format transformation.

The intelligent controller performs the dual role of optimizing robot performance and preventing inadvertent incisions. The research will investigate two general approaches to the design. One approach will use optimal control theory and the other will utilize a hybrid of soft computing techniques (fuzzy control, neural networks and genetic algorithms). Both of these techniques have been used successfully to control autonomous aircraft.

The simulator also calculates appropriate inceptor forces. In the near term, the drive signal math model for the haptic stimuli will be essentially the same as that in the actual robot although it will rely on a sophisticated organ dynamics model to compute the appropriate organ forces interacting with the robot end effectors. Eventually haptic feedback will be applied to enhance the environment for the surgeon. It has been shown in other applications that haptic stimuli, even though artificial, provide information to the operator that improves human performance and dramatically improves the "feeling of presence".

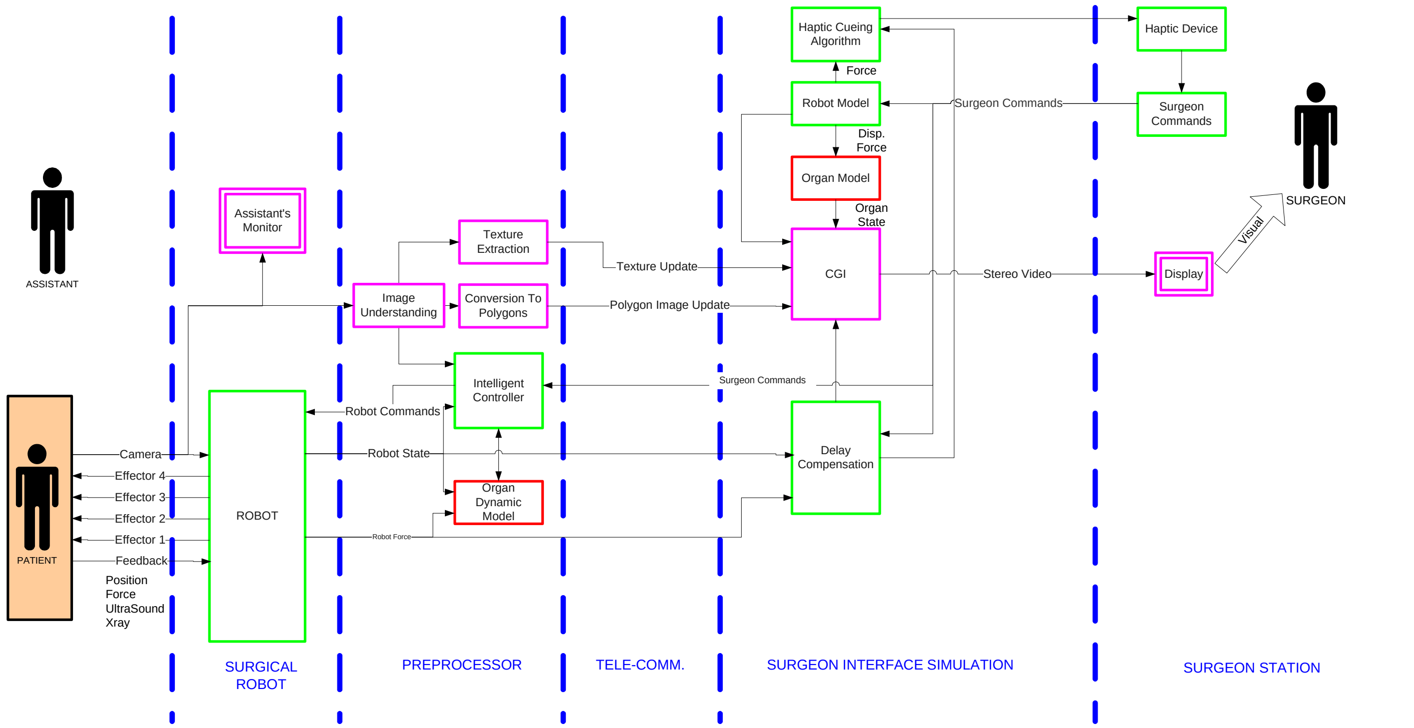

Figure 6 Block diagram of the surgical telerobotics system

Conclusions

The major impediment to remote robotic surgery is the effect of telecommunications delay on the surgeon's performance. An innovative approach to time delay effect mitigation has been presented, which is to have the surgeon operate through a simulator running in real time.

In the proposed telepresence framework the simulator acts as a predictor, providing information to the surgeon consistent with the no delay situation. Clearly, dynamics models both for the robot dynamics and organ dynamics are necessary for the simulator to function in this way. The image preprocessor portion is the essential corrector. The intelligent controller is designed as an invigilator. The total integrated surgical telerobotics system is to behave as the human surgeon would if there were not a performance encumbering delay.

Four major research areas must be explored in order to ensure achieving the objectives of our project. The paper presents a discussion of those key areas as summarized below.

1) Simulator as predictor (Delay analysis and compensation). The delays encountered in remote robotic surgery will be greater than any encountered in human-machine systems analysis, with the possible exception of remote operations in space. Therefore, tests will be conducted to determine the maximum "tolerable" delay. Subsequently novel compensation techniques will be developed. Included will be the development of the real-time simulator, which is at the heart of our approach. The simulator will present real-time, stereoscopic images and artificial haptic stimuli to the surgeon.

2) Organ Modeling. This aspect of the research will be challenging. Our approach is to investigate the relative merits of two different approaches, continuum mechanics and the finite element method. Since the tissue model runs in real time, computational parsimony is imperative.

3) Image Processing. Because of the delay and the possibility of insufficient bandwidth a high level of novel image processing is necessary. This image processing will include several innovative aspects, including; image interpretation, video to graphical conversion, texture extraction, geometric processing, image compression and image generation at the surgeon station.

4) Intelligent Control. Since the approach we propose is in a sense "predictor" based, albeit a very sophisticated predictor, a controller, which not only optimizes end effector

trajectory but also avoids error, is essential. We propose to investigate two different approaches to the controller design. One approach is based on modern control theory; the other one involves soft computing techniques.

References

[1] E. Ballantyne, H. Garth. The Pitfalls of Laparoscopic Surgery: Challenges for Robotics and Telerobotic Surgery. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. Vol.12, No.1, 1-5, 2002.

[2] R.M. Satava. Surgical Robotics: The Early Chronicles. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. Vol.12, No.1, 6-16, 2002.

[3] T.B. Sheridan. Musings on telepresence and virtual presence. *Presence*. 1(1), 120-125, 1992.

[4] S.R. Ellis. Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence", *Presence: Teleoperators and Virtual Environments, MIT Press*, 5(2), 247-259, 1996.

[5] M. Lombard, T. Ditton. Measuring Presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. Paper presented at *Presence 2000: the 3rd International Workshop on Presence* (Delft, Netherlands), 2000.

[6] R. Stone. Haptic feedback: A potted history, from telepresence to virtual reality. In *The First International Workshop on Haptic Human-Computer Interaction* (Glasgow, UK) Springer-Verlag Lecture Notes in Computer Science, 1-7, 2000.

[7] L. Guo, F.M. Cardullo, J.A. Houck, L.C. Kelly, T. Wolters. New Predictive Filters for Compensating the Transport Delay on a Flight Simulator. In *Proc. the AIAA Modelling and Simulation Technologies Conference, Providence (RI), August 17-19* (paper number AIAA 2004-5441), 2004.

[8] G.J. Klir. The Role of Anticipation in Intelligent Systems. Computing Anticipatory Systems, ed. D.M. Dubois, American Institute of Physics, 37-46, 2002.

[9] R. Rosen. Anticipatory Systems in Retrospect and Prospect. *General Systems Yearbook*, Vol. 24, No. 11, 1979.

[10] W.H. Levison, R.E. Lancraft, A.M. Junker. Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task. In *Proc. 15th Annual Conference on Manual Control (AFFDL-TR-79-3134).* Wright Patterson Air Force Base: Air Force Flight Dynamics Laboratory.

[11] J.X. Chen, et al. Knee Surgery Assistance: Patient Model Construction, Motion Simulation, and Biomechanical Visualization. In *IEEE Transactions on Biomedical Engineering.* 48:9. September 2001.

[12] J.X. Chen, S.-M. Kang. Model-order reduction of nonlinear MEMS devices through arclength-based Karhunen-Loeve decomposition. In *Proc. IEEE International Symposium on Circuits and Systems: ISCAS 2001*, Vol. 3, 457-460.

[13] S. Gibson, et al. Simulating Knee Surgery Using Volumetric Object Representations. *CVR Med 1997*, 369-378.

[14] Y.C. Fung. Biomechanics Mechanical Properties of Living Tissue, Second edition. Springer Verlag, New York, 1993.

[15] M.P. Ottensmeyer. In Vivo Measurement of Solid Organ Visco-Elastic Properties. In *Proc. MMVR 2002 - Medicine Meets Virtual Reality 02/10,* 328-333. January 2002.

[16] B.K. Tay, N. Stylopoulus, S. De, D.W. Rattner, M.A. Srinivasan. Measurement of In-vivo Force Response of Intra-abdominal Soft Tissues for Surgical Simulation. In *Proc. MMVR 2002 - Medicine Meets Virtual Reality 02/10,* 514-519. January 2002.

[17] J.T. Samosky, W.E.L Grimson, M.L. Gray. Measurement of Surgical Instrument Force Trajectories and Cartilage Stiffness for Simulation of Arthroscopic Knee Surgery. (http://www.ai.mit.edu/research/abstracts/abstracts2000/pdf/z-samosky.pdf)

[18] V. Vuskovic, M. Kauer, G. Szekely, M. Reidy. Realistic Force Feedback for Virtual Reality Based Diagnostic Surgery Simulators. In *Proc. IEEE ICRA 2000*, 1592-1598.

[19] I. Brouwer, et al. Measuring In Vivo Soft Tissue Properties for Haptic Modeling in Surgical Simulation. In *Proc. MMVR 2001 - Medicine Meets Virtual Reality 01/9,* 69-74. January 2001.

[20] L.A. Zadeh. From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions. In *IEEE Trans. Circuits and Systems –Fundamental Theory and Applications*, vol. 45, No. 1, 105-119. January, 1999.

[21] J.-S.R. Jang, C.-T. Sun, E. Mizutani. Neuro-Fuzzy and Soft Computing, Pearson Education, 1996.

[22] H.W. Lewis, III. The Foundations of Fuzzy Control, Plenum, 1997.