Elucidating, Assessing, and Training Spatial Skills in Minimally Invasive Surgery Using Virtual Environments

Frank Tendick\textsuperscript{1} and Mary Hegarty\textsuperscript{2}

\textsuperscript{1} Department of Surgery, University of California, San Francisco, CA 94143-0475
\textsuperscript{2} Department of Psychology, University of California, Santa Barbara, CA 93106
tendick@eees.berkeley.edu; hegarty@psych.ucsb.edu

Abstract

With the introduction of minimally invasive techniques, surgeons must learn skills and procedures that are radically different from traditional open surgery. Traditional methods of surgical training that were adequate when techniques and instrumentation changed relatively slowly may not be as efficient or effective in training substantially new procedures. Virtual environments are a promising medium for training. Because there are few standardized training methods in surgery, there is little information concerning the essential skills that must be trained and assessed. Consequently, experiments and modeling are needed to develop an understanding of the basis of surgical skill. Although skilled surgeons are often said to have "good hands," in fact, performance in surgery is strongly dependent on spatial skills. In this paper, we describe a collaborative effort to elucidate the role of spatial skills in minimally invasive surgery using virtual environments, and discuss the potential of virtual environments for assessing and training surgical skills.

The Need for Better Surgical Training

Training in surgery is principally based on an apprenticeship model. Residents learn by watching and participating, taking more active roles in the operation as their experience increases. This model endured in part because the techniques of traditional open surgery mostly rely on familiar eye-hand coordination. Consequently, most residents could achieve competence by repeated practice. Although procedures changed, experienced surgeons could learn them relatively quickly because the fundamental techniques were constant. With the introduction of new minimally invasive and image-guided techniques, perceptual-motor relationships are unfamiliar. The introduction and successful adoption of these techniques is often impeded by the inability to effectively train residents and practicing surgeons in their use.

The other major reason for the survival of apprenticeship is the inadequacy of alternatives. Books, videos, and CD-ROMs are poor media for training skills; they are 2-D and the user cannot physically interact with them. Cadavers, animals, and in vitro training models made of synthetic materials can be useful, but they are scarce, expensive, or do not portray the full range of anatomical variations and disease states. Although all of these media are used to some extent, because of their inadequacy surgeons either fail to learn new techniques or traverse much of their learning curve on patients.

We are particularly interested in laparoscopic, or minimally invasive, surgery because it is particularly demanding of spatial skills. This is surgery of the abdomen performed through cannulas, typically 5–10 mm in diameter, inserted through the skin. Long thin instruments are used, with the need to go through the cannula creating a fulcrum and reducing the degrees of freedom of movement (Figure 1). The surgeon watches a video image from an endoscope inserted through one of the cannulas. There are several obvious spatial problems in laparoscopic surgery. The surgeon must adapt to the changing orientations between camera and instruments. The camera must be placed and tissue exposed so that key structures are not obscured and so that instruments will be effective with their limited degrees of freedom. Complex operations are carried out by a team of surgeons, i.e., a camera operator and assistant in addition to the primary surgeon, and the team must be able to communicate spatial plans or have shared mental models.

The Role of Spatial Cognition

Surgical training provides a challenging environment for studying spatial cognition. Although skilled surgeons are often said to have "good hands," in fact, performance in surgery is strongly dependent on spatial skills. The surgeon must develop a mental image of three-dimensional anatomy based on a surface view or cross sections from X-ray, CT, MRI, or ultrasound images. From this model and a goal state based on experience and anatomical knowledge, he or she must plan a strategy to gain exposure of the important anatomy and obtain the desired result. This plan requires complex coordination between a team of assistants using...
Figure 1: Because of the fulcrum at the cannula entry through the abdominal wall, the motion of laparoscopic instruments is constrained to 4 degrees of freedom.

an array of instruments. With the advent of minimally invasive techniques, the surgeon must rely on a video image of the internal anatomy and use instruments constrained by a fulcrum at their passage through the skin. This requires additional mental transformations of the image and careful planning to handle the constraints.

Many aspects of surgery make it an excellent domain in which to study the boundaries of spatial cognition. Anatomical environments can be extremely complex, with intricate 3-D relationships between deformable structures that can vary from patient to patient. The surgeon must visualize relationships beneath the surface that cannot be seen, or construct 3-D mental models from 2-D images. He or she must make transformations between viewpoints—sometimes with millimeter accuracy, as when targeting an unseen tumor with a biopsy needle. Complex spatial and mechanical reasoning is necessary to plan and carry out an action with multiple instruments.

Despite the importance to society of ensuring the competence of surgeons, there has been surprisingly little research on surgical skill and training. Furthermore, this research has relied largely on experienced surgeons’ intuition of the component skills in surgery, and has not been informed by cognitive models (Winckel et al. 1994; Bhoyrul et al. 1994; Hanna et al. 1998; Derossis et al. 1998; Rosser Jr., Rosser, & Savalgi 1998). The perceptual motor consequences of degraded visual information, reduced dexterity, and limited haptic sensation in laparoscopic surgery have been identified (Tendick et al. 1993; Bredveld 1998) and detailed time and motion studies have also been performed (Cao, MacKenzie, & Payande 1996; Sjoerdsmma 1998). Nevertheless, these studies have done little to elucidate the underlying cognitive demands in surgery. Several studies have shown strong correlations between standardized tests of spatial ability and performance ratings on a variety of tasks in open surgery (Gibbons, Gudas, & Gibbons 1983; Gibbons, Baker, & Skinner 1986; Schoneman et al. 1984; Steele, Waldar, & Herbert 1992). It is likely that laparoscopic surgery would be at least as strongly dependent on spatial ability.

The Role of Virtual Environments

Computer-based training in virtual environments has many potential advantages. It is interactive, yet an instructor’s presence is not necessary, so students may practice in their free moments. Any disease state or anatomical variation can be recreated. Simulated positions and forces can be recorded to compare with established performance metrics for assessment and credentialing. Students can also try different techniques and look at anatomy from perspectives that would be impossible during surgery. Although the anatomical environment is quite complex, the essential skills and procedural steps can be taught in simulated environments achievable with current mid-range graphics workstations.

In order to realize the full potential of virtual environments in training, research needs to be conducted to elucidate the specific features of virtual environments that lead to maximum transfer of training from the virtual to the real environment. This transfer might depend on factors such as whether interaction in the environment is learner controlled or passive, whether information in the display is multimodal (e.g., visual and proprioceptive) or uni-modal and the degree of immersion in the environment. It is important to precisely characterize how these differences affect the representations constructed in virtual environments. For example, in a recent study, Richardson, Montello, and Hegarty (1999) found a substantial alignment effect in cognitive maps constructed from a desktop virtual environment, indicating that subjects had learned the route with a preferred orientation corresponding to their initial facing direction in the environment. A likely explanation for this effect, supported by recent research (Chance et al. 1998; Klatzky et al. 1998) is that both vestibular and visual information are necessary to induce egocentric updating and a desktop virtual environment provides no vestibular information.

Despite their novelty, there have already been demonstrations of the usefulness of virtual environments in spatial learning. After moving through a virtual rendition of a building, people have better than random performance at wayfinding in the real building, indicating that people can learn spatial layout from a virtual environments (Regian, Shebilske, & Monk 1992; Baily & Witmer 1994; Wilson, Foreman, & Tlauka 1997). However, to date, spatial knowledge acquired from virtual environments was generally poorer than that acquired in the real environments (Henry & Furness 1993; Richardson, Montello, & Hegarty 1999). We suspect that transfer in simulations of minimally invasive surgery may be substantially better because interaction in the real environment is already reduced by videoscopic imaging and the limited motion and
haptic perception possible with laparoscopic instruments. Consequently, there is less difference between the virtual and real environments.

Virtual environments are interactive and dynamic. These properties will allow us to create situations that would be impossible with figural (pencil and paper) tasks or the static environments explored in common navigation paradigms. We will be able to show physically impossible views, create simulated mechanisms with which the user can interact, graphically portray information otherwise not perceptible (such as internal forces in tissue), and vary the degree of visual and kinesthetic information presented to the user. By changing conditions, we can elucidate the processes underlying performance in complex tasks.

Research Focus

Because of the complexity of surgical tasks, surgeons appear to use significantly different strategies, relying to differing degrees on demanding spatial processes. By studying performance in virtual environments, we can elucidate these strategies and test models of the underlying representations and processes.

Our work has two synergistic thrusts. The first aims to advance fundamental knowledge in cognitive science by examining how spatial cognitive skills are integrated in the performance of complex tasks. Specifically, we plan to study navigation in small-scale three-dimensional environments, mental simulation of mechanical interactions between deformable structures, and how these spatial and mechanical reasoning processes are integrated in solving complex spatial problems. The second thrust is to advance the state of the art in human-computer interaction by designing intelligent systems to train and assist human performance in spatial problem solving. Using the understanding we develop of spatial reasoning strategies, we will develop methods for tracking and identifying users’ interactions with virtual environments, and augmenting these environments to assist training of spatial and motor behaviors. These methods will be based on computational models of the construction of 3-D spatial models from multiple views. To complete the synergism of the two research thrusts, the models will also guide the experiments, performed using the virtual environments we develop to investigate spatial phenomena in ways that otherwise would be impossible.

Our Current Testbed

We have developed a general purpose surgical simulation authoring tool. This modular tool makes it easy to simulate different surgical procedures by changing the anatomic models, physical models of tissue behavior, and visual and haptic interfaces as necessary (Downes et al. 1998; Tendick et al. 1998). Scenes from the environment developed for laparoscopic cholecystectomy, or gallbladder removal, are shown in Figure 2. Input to the simulation is through four-degree-of-freedom (DOF) haptic interfaces with force feedback, which duplicate the kinematics of the motion of laparoscopic instruments through a fulcrum (Figure 3). These are provided for both of the user’s hands, and a third 4 DOF interface without force feedback is used to control simulated laparoscope motion. Visual display for laparoscopic environments is through a single monitor, but can easily be switched to a stereographic or head-mounted display. The simulation has fast algorithms for instrument-tissue contact detection and modeling the deformation of soft tissue. Simulated grasping, electrocautery, stapling, and cutting are implemented. Complex simulations can run on the Silicon Graphics Octane workstation where we have implemented the simulation. The cholecystectomy simulation models include over 12,000 surface triangles, 2,800 of which deform. It runs at an interactive speed of 13 updates per second.

Two simulations are currently implemented in the testbed. The first is to assess and train the use of an angled laparoscope. The second is a simulation of laparoscopic gallbladder removal, or cholecystectomy.

Use of the Angled Laparoscope

In laparoscopic surgery, the fulcrum at the abdominal wall limits the range of motion of the laparoscope.
Figure 3: Laparoscopic haptic interface. The device has a gimbaled fulcrum with a motor driving axial roll. The other 3 DOF are provided by a commercial Phantom interface (Sensable Technologies, Cambridge, MA), for a total of 4 DOF.

Figure 4: Angled laparoscope concept. The laparoscope passes through a cannula, which is constrained by the fulcrum at the abdominal wall. The objective lens is angled with respect to the laparoscope axis.

Consequently, the viewing perspective within the abdomen is also limited. If the objective lens is aligned with the laparoscope axis, it is only possible to view from directions centered at the fulcrum. Some regions may be obscured by neighboring organs, or it may be impossible to view important structures en face. Laparoscopes with the objective lens at an angle with respect to the laparoscope axis are preferred and are often essential for many procedures, as they expand the range of viewing orientations (Figure 4). Although the concept of the angled laparoscope is simple, in practice its use can be difficult. For example, to look into a narrow cavity (shown as a box in Figure 4), the laparoscope objective must point along a line into the cavity. Because of the constrained motion of the laparoscope, there is only one position and orientation of the laparoscope that will place the lens view along this line. (Or, more strictly, there is a narrow range of position and orientation that will suffice, depending on the width of the cavity.) The viewer can only see the location of the cavity relative to the current video image, and consequently must use spatial reasoning to estimate how to achieve the necessary laparoscope location.

To isolate specific spatial skills involved in using an angled laparoscope, we have developed a virtual environment to simulate its use. The environment comprises 6 targets, each a tall box suspended in space at a different position and orientation (Figure 5). Input to the simulation is through the Virtual Laparoscopic Interface (Immersion Corp., Santa Clara CA). The test begins with the laparoscope pointed at the first target. One of the other targets changes color, and the subject must position and orient the laparoscope to view all four corners at the bottom of the target box. When this view is demonstrated, the experimenter hits a key and the process is repeated for the next target in sequence.

The kinematics of the scope and the global position of targets viewed only from the video image cannot be perceived directly, but must be inferred from the user’s kinesthetic sense of the laparoscope orientation as it is held and by a mental model of the location of the scope lens in space and the target’s location relative to the lens. In pilot studies, a few subjects grasped these relationships quickly, while most who were successful at the task used an inefficient strategy that does not rely on a global model, instead making ad hoc reorientations as necessary while attempting to center the target in the image. Some subjects failed to develop any strategy and could not complete the task without substantial coaching.

Laparoscopic Cholecystectomy Simulation

Many experimental and commercial prototype environments for training have tried to simulate entire operations, resulting in low fidelity in each of the component tasks comprising the operation. This is an inefficient and probably ineffective approach. It is relatively easy to learn most steps of a procedure by watching and participating. In every procedure, however, there are a few key steps that are more likely to be performed incorrectly and to result in complications. The significance of these steps might not be obvious, even to an experienced surgeon, until situations arise such as unusual anatomy or uncommon manifestations of disease. The value of a surgical simulator is analogous to the value of a flight simulator. In current practice, pilots are certified to fly by confronting simulated situations, such as wind shear or engine emergencies, that happen only once in a lifetime, if at all. A surgical simulator should train surgeons for the principal pitfalls that underlie the major technical complications. Such training and assessment could be used by medical schools, health administrations, or professional accrediting organizations to enforce standards for granting surgical privileges and for comparing patient outcomes with surgeon skill (Grundfest 1993;
An example of the importance of training critical steps of procedures is the laparoscopic cholecystectomy (gallbladder removal). The bile ducts (Figure 2) carry bile created in the liver to the gallbladder. There it is stored and concentrated until it is released into the intestine. Bile duct injury can be the result of poor technique or misinterpretation of the anatomy. The cystic duct, which leads directly from the gallbladder, must be cut before the gallbladder can be removed. In Figure 2, the cystic duct is easily identified, clipped (i.e., closed with a staple which encircles the duct), and cut. In reality, however, the biliary tree is obscured by connective tissue. The surgeon may confuse the common bile duct (the large duct leading to the intestine) for the cystic duct. If so, the common duct may be inappropriately divided. The repair of this injury is difficult, and since it usually goes unnoticed during the procedure, it requires a second operation.

One prospective study found a high rate (2.2%) of bile duct injuries in procedures performed by inexperienced laparoscopic surgeons (Southern Surgeons Club 1991). Experienced surgeons also caused injuries, although at a lower rate (0.1%). Based on our analysis of 139 bile duct injuries, a few simple rules have been developed to reduce the likelihood of injury (with layperson's explanations in parentheses):

- Use lateral traction (i.e., pull to the side) on the infundibulum (bottom) of the gallbladder during dissection. This draws the cystic duct to full length and maximizes the difference in lie of the cystic and common ducts.
- Dissect any potential space between gallbladder and cystic duct completely. This will help uncover a hidden cystic duct when the gallbladder is adherent to the common duct.
- Clear the triangle of Calot (between the cystic duct, liver, and bile ducts leading from the liver) enough to show the hepatic (liver) side of the infundibulum (bottom) of the gallbladder. This allows the cystic duct to be identified with greater certainty, since it will be found as a continuation of the gallbladder.
- Use an angled scope to gain the optimal (en face) view of the triangle of Calot.
- If the duct about to be clipped will not fit entirely within a 9mm clip (which should close around the duct to seal it), assume it is the common duct (because the common duct has a larger diameter than the cystic duct).
- Any duct that can be traced to disappear behind the duodenum (intestine) has to be the common duct.

The virtual environment shown in Figure 2 has been developed to teach proper techniques that should avoid bile duct injuries. In the current simulation, the user must dissect through a single layer of overlying fat to see the biliary structures. The dissection is achieved by removing small regions of fat with a simulated electrosurgical tool. Although the simulated fat performs the function of hiding the key structures, it is not anatomically accurate. We are developing a version in which the structures are joined by adhesions. The gallbladder must be retracted to expose the cystic duct, which is clipped in two places so that it can be cut between the clips. It is easy to identify the cystic duct in the Visible Human male. In future versions, variations in which greater difficulty is encountered can be created. Anatomic variations of the biliary tree can also be simulated.
Future Directions

Using the tools described above for the creation of virtual environments for surgical training, we will simulate three representative surgical tasks to guide future research.

The first surgical task that we will study involves using the angled laparoscope, described above. Because of the constrained motion of the laparoscope, there is only one position and orientation of the laparoscope that will place the lens view along the optimal line to view the relevant anatomy for a surgical procedure. The surgeon must use spatial reasoning to estimate how to achieve the necessary laparoscope location.

The second task is that of constructing a cognitive map of the relevant anatomy for a specific surgical procedure on a specific patient. The surgeon has to construct this cognitive map from his or her knowledge of the prototypical human anatomy, information provided by MRI scans or ultrasound, and the continually changing visual information from the laparoscope as it is manipulated within the human body. From these three sources of information, the surgeon must plan and maintain a route to the site of the tissue on which the surgery is being performed.

The third surgical task that we will study is that of obtaining exposure to the relevant anatomy for a surgical procedure. This involves achieving an optimal view of the relevant tissues, planning orientation of the tissues and instruments for effective manipulation, and applying traction on tissues to facilitate dissection.

For each of these skills we will conduct preliminary studies in which we will elucidate the range of strategies used by experts and novices to accomplish the skill. From our knowledge of these strategies, we will develop computational models of ideal performance on each of the skills. These models will inform the development of surgical simulators to guide trainees in the development of each skill. We will evaluate the effectiveness of these augmented displays in training each of the skills.

In all of our experiments we will measure individual differences in basic spatial abilities identified in the psychometric literature, i.e., spatial relations (simple rotation), spatial visualization, and spatial orientation abilities (Carroll 1993; Lohman 1979). We will study the relation of each of the spatial abilities to each of the surgical skills identified above. In our training studies we will compare how high- and low-spatial individuals learn from the surgical simulators to examine the extent to which spatial abilities are related to surgical performance.

Throughout the research we will develop and test basic computational models of the strategies subjects use to solve complex problems. These will be based on methods from robotic motion planning and computer vision. We will build on concepts from existing models, but our emphasis will be on using models to test fundamental hypotheses about the nature of knowledge structures and the transformations that must be carried out in solving problems. We will examine the fit of the models to the data that we collect, and where competing modes are possible, we will develop further experiments to distinguish between them. Unlike experiments in mental imagery that must infer internal states from behavior, we have the advantage that we can measure the execution of the strategies in the form of subjects’ actions using behavior recognition techniques. Therefore we have a strong test of our models in that they must be sufficient to characterize the structures and transformations that would be necessary computationally to carry out the strategies. Because we will obtain psychometric data on all our subjects, we may further link our models to known ability factors such as spatial relations and visualization.

In conventional experimental methods, subjects’ strategies are identified from secondary data such as verbal descriptions (Ericsson & Simon 1980) or eye movements (Just & Carpenter 1976) as the subject attempts the task. In the surgical domain, many strategies are embodied in motor action, such as the movement of the laparoscope or ultrasound probe to a goal viewpoint or instrument motion to expose or dissect tissue. This gives direct access to procedural behavior that often cannot be described by the subject. We are developing dynamic behavior recognition methods to identify subjects’ strategies as they perform experimental tasks. The difficulty in identifying the behaviors necessary to elucidate surgical skills is that they are inherently dynamic. Consequently, they cannot be adequately described by a sequence of static positions or forces, but must be defined in terms of directions and velocities. Fortunately, laparoscopic kinematics are 4 DOF, and are further constrained by the tasks we propose. Human motion tends to be simplified by synergies that reduce effective DOF, such as Listing’s law, which describes the reduction of eye and shoulder motion from the 3 DOF possible to 2 DOF in gaze and reaching movements, respectively. Preliminary work shows that we can easily classify basic rotational movements performed by subjects.

In addition to using the haptic interface (Figure 3) to provide force feedback of tissue interaction in the simulations, we propose a novel use of force feedback to provide kinesthetic training. Instead of verbally describing or visually showing a skill, we can guide the user through the desired motion. This could be of value in teaching procedural skills involving perceptual motor relationships that are difficult to verbalize. We have already created a controller capable of guiding the user along a spatial path or a time-dependent trajectory with a simulated stiffness to resist perturbation from the path. This is implemented using a nonlinear sliding mode controller with adaptation to optimize performance as the user muscle stiffness varies.
Discussion

A major issue is realism. Although the models are extracted from Visible Human data and are thus accurate, it is difficult to reproduce the connective tissue and adhesions between organs that cause difficulty in dissection. Although computer power is of course increasing rapidly, a tenfold improvement in detail in each dimension requires a thousandfold increase in computer power for 3-D modeling, which would mandate a high degree of parallelism. Nevertheless, the skills and steps that are critical to train do not depend on detail. For example, bile duct injuries in the cholecystectomy are caused by surgeons “missing the forest for the trees,” and failing to follow basic steps that will allow them to positively identify important anatomical landmarks despite the complex dissection through connective tissue that obscures the landmarks. Consequently, a comparatively simple virtual environment may enhance the user’s learning of fundamental relationships. The environment may also be enhanced to show hidden relationships that can never be observed during real surgery. The differences between reality and simulation will necessitate careful studies of transfer, however.

It is fortuitous that the impoverished interface of minimally invasive surgery, which eliminates cutaneous tactile feedback, reduces kinesthetic force feedback, and limits visual information, also makes it easier to reproduce the actual feedback the surgeon receives. Because the information the user receives is constrained, it is feasible to elucidate how novices and experts use the information in performing complex tasks. As both our understanding and virtual interfaces improve, it may be possible to extend the results to understanding the performance of open surgery and complex skills in other domains in which perception is not so severely constrained.

Acknowledgments

This work was funded in part by the National Science Foundation under grants CDA-9726382 and BCS-9980122, the Office of Naval Research under grants N00014-96-1-1200 and N00014-98-1-0434, and the Army Research Office under grant DAAG55-97-1-0059.

References


