OBJECTIVE EVALUATION OF LAPAROSCOPIC SURGICAL SKILLS USING HIDDEN MARKOV MODELS BASED ON HAPTIC INFORMATION AND TOOL/TISSUE INTERACTIONS

Jacob Rosen[++], Ph.D., Massimiliano Solazzo[+], M.D.,
Blake Hannaford[++], Ph.D., Mika Sinanan[+] , M.D., Ph.D.

[+] *Department of Surgery, Box 356410*
[++] *Department of Electrical Engineering, Box 352500*
University of Washington
Seattle, WA, 98195, USA

ABSTRACT

Laparoscopic surgical skills evaluation of surgery residents is usually a subjective process, carried out in the operating room by senior surgeons. By its nature, this process is performed using fuzzy criteria. The objective of the current study was to develop and assess an objective laparoscopic surgical skill scale using Hidden Markov Models (HMM) based on haptic information, tool/tissue interactions and visual task decomposition. *Methods:* Eight subjects (six residents: first year surgical residents 2xR1, third year surgical residents 2xR3 fifth year surgical residents 2xR5; and two expert laparoscopic surgeons: 2xES) performed laparoscopic cholecystectomy following a specific 7 step protocol on a pig. An instrumented laparoscopic grasper equipped with a three-axis force/torque sensor was used to measure the forces and torques at the hand/tool interface synchronized with a video of the tool operative maneuvers. A synthesis of frame-by-frame video analysis was used to define 14 different types of tool/tissue interactions each one associated with unique force/torque (F/T) signatures. HMMs were developed for each subject representing the surgical skills in terms of haptic information and tool tissue interactions. The statistical distance between the HMMs representing residents at different levels of their training and the HMMs of expert surgeons were calculated in order to evaluate the learning curve of selected steps of laparoscopic cholecystectomy. *Results:* The objective laparoscopic surgical skill learning-curve showed significant differences between all skill levels. The major differences between skill levels were: (i) magnitudes of F/T applied (ii) types of tool/tissue interactions used and with the transition between them and (iii) time intervals spent in each tool/tissue interaction and the overall completion time. The HMM analysis showed that the greatest difference in performance was between R1 and R3 and then decreased as the level of expertise increased. This objective evidence for a learning curve indicates that the surgical residents appear to acquire a major portion of their laparoscopic surgical capabilities between the first and the third years of their residency training.

1. Introduction

The technical performance of surgery in general, and in minimally invasive surgery (MIS) in particular, requires the application of forces and torques (F/T) on and between tissues to achieve specific goals. Parameters that determine the magnitude of the F/T used include the nature and the goal of the tissue being manipulated, the type of the instrument being used, and the skill of the surgeons. One of the more difficult tasks in MIS education is to teach the optimal application of instrument F/T necessary to conduct an operation.

Although the acquisition of laparoscopic technical skill and the assessment of performance are paramount for surgical resident training, the current method of choice for evaluating surgical skill is a subjective evaluation of the performance in the operating room or reviewing a video-tape of surgery. In the last 5 years, use of surgical simulation for training and evaluating surgical skills has become a subject of an ongoing active research.

Simulators for evaluating surgical skills and testing of dexterity in MIS can be roughly divided into three categories: (i) training box including physical objects or latex organ packs (ii) virtual reality simulator including graphical representation of virtual objects or virtual anatomy, and (iii) virtual reality simulator with a force feedback device (haptic display) for simulating forces and torques generated as a result of interaction between the virtual objects or organs and the surgical tools.

One of the most used surgical simulator is the laparoscopic trainer box covered by an opaque membrane through which different trocars are placed at different working angles. The trainee is required to complete several structured laparoscopic tasks which are scored for both precision and speed of performance. Several studies [*1;2;3;4;5*] have found the laparoscopic training box a valuable teaching tool for training and evaluation of basic laparoscopic skills. The performance, evaluated base on using the simulator, improved over the residency training and were correlated with postgraduate year.

A virtual reality simulator for laparoscopic surgery models the movements needed to perform MIS and can generate a score for various aspects of psychomotor skill. A typical example of such a simulator is the MIST-VR system. The MIST-VR uses two laparoscopic instruments mounted on a frame with position sensors which provide instrument movement data that is translated into interactive real time graphics on a PC. Targets appear randomly within the operating volume according to the skill task and can be grasped and manipulated with the instruments. Accuracy and errors during the tasks and completion time are logged**.** Studies preformed using the MIST VR simulator [6,7] concluded that it can objectively assess a number of desirable qualities in laparoscopic surgery, and can distinguish between experienced and novice surgeons.

The use of virtual reality models for teaching complex surgical skills while simulating realistic human/tool and tool/tissue interaction has been a long-term goal of numerous investigators [8,9,10,11,12]. Although the haptic devices providing force feedback to the surgical tool while interacting with the virtual tissue/organ are commercially available (for review see [13]), simulating a realistic force feedback based on biomechanical models of soft tissue is still under active research [14]. The complexity of these biomechanical models is due to the viscoelasticity and non-linear characteristics of soft tissues. Moreover, the F/T data measured in-vivo [15, 16, 17, 18] is crucial for designing and evaluating haptic devices force-feedback telerobotic systems [19, 20, 21, 22] and virtual reality simulators.

The methodology developed in the current study was based on the Hidden Markov Modeling (HMM). HMM were extensively developed in the area of speech recognition [23, 24, 25, 26]. Based on the theory developed for speech recognition HMMs have become useful statistical tools in the fields of human operator modeling in general, and robotics in particular. HMMs

were applied for studying teleoperation [27, 28, 29], human manipulation actions [30], human skills evaluation for the purpose of transferring human skill to robots [31, 32, 33], and manufacturing applications [34, 35]. Gesture recognition with HMMs has also received increasing recent attention from the rehabilitation technology community (see [36] for review). They are also being applied to the recognition of facial expressions from video images [37]. Moreover HMMs may well prove useful in many other emerging applications beyond human computer interfaces, e.g DNA and protein modeling [38], fault diagnosis in nuclear power plants [39], and detection of pulsar signals [40]. These application suggest that the HMMs have high potential to provide better models of the human operator in complex interactive tasks with machines.

The goal of this study was to define the learning curve of MIS based on new quantitative knowledge of the F/T applied by surgeons on their instruments, and the types of tool/tissue interactions used during the course of MIS surgery. This goal was pursued through several steps: (*i*) developing instrumented endoscopic tools which contain embedded sensors capable of measuring and recording F/T information (*ii*) creating a database of F/T signals acquired during actual operating conditions on experimental animals, (*iii*) performing a task decomposition in terms of tool/tissue interactions existed in MIS based on video analysis (*iv*) developing statistical models (HMM) for evaluating an objective laparoscopic skill level.

2. Materials and Methods

2.1 Subjects and Protocol

Eight subjects (six general surgery residents: first year residents - 2xR1, third year residents 2xR3, fifth year residents - 2xR5 and two expert laparoscopic surgeons - 2xES) each completed the experimental protocol. . The protocol consisted of two phases. During the first phase, subjects watched a 45-minute video of the surgical procedure guided by a senior surgeon, to demonstrate the technique of the procedure. This study was not intended to test knowledge of the procedure, but whether the surgeon-subjects could technically perform the procedure. In the second phase, each subject performed a laparoscopic cholecystectomy on a pig using a standardized, 7-step procedure. All surgical procedures and animal care were reviewed and approved by the Animal Care Committee of the University of Washington and the Animal Use Review Division of the U.S. Army Veterinary Corps.

Data from 3 steps of the laparoscopic cholecystectomy (positioning of the gallbladder - LC-1, exposure of the cystic duct - LC-2, and dissection of the gallbladder - LC-3) were recorded. During these steps the instrumented endoscopic tool was used with an atraumatic grasper, a Babcock grasper, and a curved dissector (Fig. 1c).

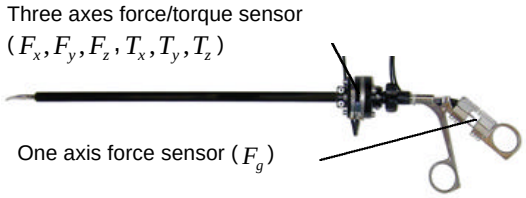
2.2 Experimental System Setup

During the laparoscopic procedures data were acquired from two sources: (*i*) force/torque data measured at the human/tool interface and (*ii*) visual information of the tool tip interacting with the tissues. The two sources of information were synchronized in time and recorded simultaneously for off-line analysis.

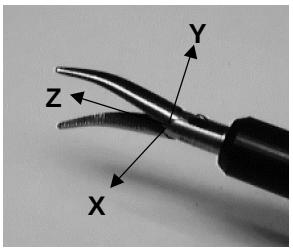
Two sets of sensors measured the F/T at the interface between the surgeons' hand and the endoscopic grasper handle (Fig 1a). The first sensor was a three-axis force/torque sensor (ATI-Mini model) which was mounted into the outer tube (proximal end) of a standard reusable 10-mm endoscopic grasper (Storz). The sensor was capable of simultaneously measuring the three components of force (F_x, F_y, F_z) and three components of torque (T_x, T_y, T_z) in a Cartesian frame (Fig. 1b).

The summation of applied forces and torques at the tool-tip/tissue interface and the tool/trocar interface were transferred through the grasper's structure to the surgeon's hand, as occurs in a normal

instrument, and vice versa. The sensor orientation was such that X and Z axes formed a plane parallel to the tool's internal jaws' contact surfaces and the Y and Z axes defined a plane perpendicular to that surface (Fig. 1b).



(a)

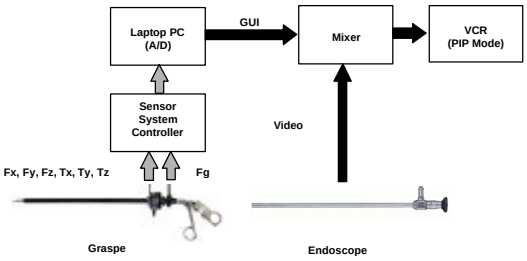


(b)

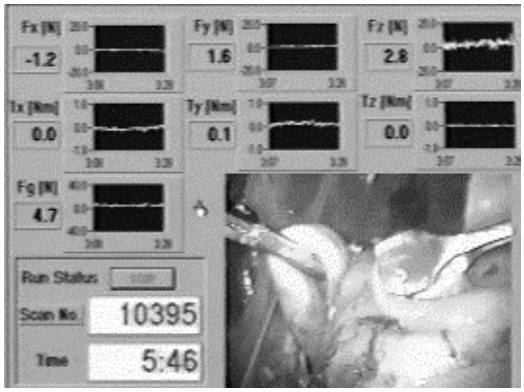
(c)

Figure 1: *The instrumented endoscopic grasper: (a) The grasper with the three axis force/torque sensor implemented on the outer tube and a force sensor located on the instrument handle (b) The tool tip and X,Y,Z frame aligned with the three axis force/torque sensor (c) Tool tips used in the surgical procedure (from left to right): Atraumatic Grasper, Babcock grasper, Curved dissector.*

A second force sensor (Futek - FR1010) was mounted to the endoscopic grasper handle to permit the measurement of grasping force (F_g) applied by the surgeon's fingers on the instrument.



(a)



(b)

Figure 2: *Experimental setup: (a) Block diagram of the experimental setup integrating the force/torque data and the view from the endoscopic camera, (b) Real-Time user interface of force/torque information synchronized with the endoscopic view of the procedure using picture-in-picture mode.*

The grasper had a reticulating feature that enabled the surgeon to change the orientation of the tool tip relative to the grasped tissue without changing the handle orientation. The alignment between tool tip origin relative to the sensor remained unchanged since the outer tube and the tool tip are linked mechanically.

The F/T data were integrated with the laparoscopic camera view of instrument activity (Fig 2). The seven channels F/T data ($F_x, F_y, F_z, T_x, T_y, T_z, F_g$) were sampled at 30 Hz using a laptop computer with a PCMCIA 12 bit A/D card (National Instruments - DAQCard 1200) - Fig 2a. A LabView (National Instruments) application was developed with a graphical user interface for acquiring and visualizing the F/T data in real-time during an actual operation. The video signal from the endoscopic camera

that was monitoring the grasper's tip while interacting with the internal organs/tissues was integrated with the F/T data using a video mixer in a picture-in-picture mode (PIP), allowing to correlation F/T data with instrument activity. The integrated interface was recorded during the operation for off-line frame by frame analysis (Fig. 2 b).

2.3 Data Analysis

Two types of analysis were performed on the raw data: (*i*) Video Analysis, encoding the tool-tip/tissue interaction into states; and (*ii*) Hidden Markov Modeling (HMM), for modeling and comparing the performance of surgeons at different level of their training (R1, R3, R5, ES).

2.3.1 Video Analysis

The video analysis was performed by two expert surgeons encoding the video of each step of the surgical procedure frame by frame (NTSC - 30 frames per second). The encoding process used a code-book of 14 different discrete tool maneuvers in which the endoscopic tool was interacting with the tissue (Table 2). Each identified surgical tool/tissue interaction, had a unique F/T pattern. For example, in the laparoscopic cholecystectomy, isolation of the cystic duct and artery (LC-2) involves performing repeated pushing and spreading (PS-SP - Table 1) maneuvers which in turn requires pushing forces mainly along the Z axis (F_z) and spreading forces (F_g) on the handle.

These 14 states can be grouped into three broader types based on the number of movements performed simultaneously. Fundamental maneuvers were defined as Type I, and included the idle state (moving the tool in space without touching any structures within the insufflated abdomen).

The forces and torques used in idle state represented mainly the interaction of the trocar with the abdominal wall in addition to gravitational and inertial forces.

Type	State Name	State Acronym	Force / Torque						
			Fx	Fy	Fz	Tx	Ty	Tz	Fg
I	Idle	ID	*	*	*	*	*	*	*
	Grasping	GR							+
	Spreading	SP							-
	Pushing	PS			-				
	Sweeping	SW	+/-	+/-		+/-	+/-		
II	Grasping - Pulling	GR-PL			+				+
	Grasping - Pushing	GR-PS			-				+
	Grasping - Sweeping	GR-SW	+/-	+/-		+/-	+/-		+
	Pushing - Spreading	PS-SP			-				-
	Pushing - Sweeping	PS-SW	+/-	+/-	-	+/-	+/-		
	Sweeping - Spreading	SW-SP	+/-	+/-		+/-	+/-		-
III	Grasping - Pulling - Sweeping	GR-PL-SW	+/-	+/-	+	+/-	+/-		+
	Grasping -Pushing - Sweeping	GR-PS-SW	+/-	+/-	-	+/-	+/-		+
	Pushing - Sweeping - Spreading	PS-SW-SP	+/-	+/-	-	+/-	+/-		-

Table 1: *Definition of tool/tissue interactions and the corresponding directions of forces and torques applied in cholecystectomy and Nissen fundoplication during MIS.*

In the grasping and spreading states, compression and tension were applied to tissue by closing/opening the grasper handle. In the pushing state, compression was applied on the tissue by moving the tool along the Z axis. For sweeping, the tool was placed in one position while rotating around the X and Y axes (trocar frame). Type II and type III were states defined as combinations of two or three states defined by type I (Table 2).

2.3.2 Hidden Markov Model (HMM)

During the second step of the data analysis Hidden Markov Models (HMM) and the methodology for evaluating surgical skill in laparoscopic surgery were develop. HMMs were selected for modeling the surgical procedure because their generic architecture fitted very well the nature of laparoscopic surgery task assessment. Moreover, the HMM mathematical formulation provided a very compact form statistically summarized relatively complex tasks such as individual steps of a laparoscopic surgery procedure.

Each laparoscopic surgical step could be decomposed into a series of finite *states* defined by the way the surgeon is interacting with the tissues (Table 1). The surgeon could move from one state to the other or stay in the same state for certain amount of time. Once the surgeon was interacting with the tissue in a specific state a certain F/T signature was applied by the surgeon through the surgical tool to the tissue. These F/T signatures, each defined as an *observation,* was composed of seven components vector of data $(F_x, F_y, F_z, T_x, T_y, T_z, F_g)$. Since the F/T were continues stream of data distributed normally, each state is defined by seven normal distributions functions chartered by a mean and a standard deviation $(N_i(m, s) \quad i = 1...7)$. Combining the 7 element vector into joint multivariable distribution function $f(O)$ was done by using Eq. 1.

$$f(O) = \frac{1}{(\sqrt{2p})^N |\Sigma|^{1/2}} e^{-(O-m)' \Sigma^{-1} (O-m)/2} \qquad (1)$$

where: O is the F/T observation vector; m is the mean vector; Σ is the covariance matrix, and N is the observation vector size.

The state diagram (Fig. 3), describes the decomposed process of a typical laparoscopic surgical procedure step. Circles in this diagram represented states and lines represented transitions between states. The F/T data - observation signals were not included in Fig. 3.

HMM is termed hidden due to the fact that tool/tissue interactions - the states - are hidden and the only observed signals are the F/T data. Although the state could be decomposed manually using a frame-by-frame video analysis, this is time consuming and unnecessary since the data can also be evaluated mathematically by the HMM once its parameters are optimized.

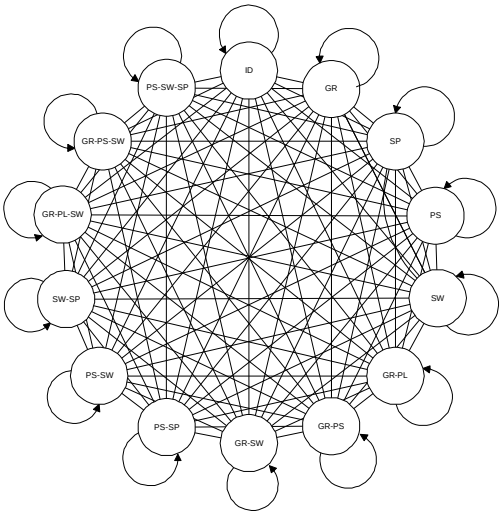


Figure 3: *HMM architecture defined by 14 fully connected state diagram (arrow heads of all the lines connected two states were omitted for simplifying the drawing)*

From the mathematical perspective, four elements should be defined in order to specify a HMM (l) [23]: (*i*) the number of states in the model – *N*, (*ii*) the state

transition probability distribution matrix – *A*, (*iii*) the observation symbol probability distribution matrix – *B*, and (*iv*) the initial state distribution vector– p . The HMM is then defined by the compact notation (7)

$$l = (A, B, p) \qquad (2)$$

Given the HMM architecture there are three basic problems of interest [23]: (*i*) The evaluation problem – Computing the probability (P) of the observation sequence given the model (l) and the observation sequence (O) .

$$\text{Given:} \quad \begin{cases} l = (A, B, p) \\ O = o_1, o_1, ..., o_T \end{cases} \qquad (3)$$

Compute: $\{P(O \mid l)$

(*ii*) Uncover the hidden states – Computing the corresponding hidden state sequence (Q) , given the observation sequence (O) and the model (l) .

$$\text{Given:} \quad \begin{cases} l = (A, B, p) \\ O = o_1, o_2, ..., o_T \end{cases} \qquad (4)$$

Compute: $\{Q = q_1, q_2, ..., q_T$

(*iii*) The training problem – Adjusting the model parameters (A, B, p) to maximize the probability (P) of the observation sequence (O) .

$$\begin{aligned} &\text{Given :} && \{ l = (A, B, p) \\ &\text{Adjust:} && \{A, B, p \\ &\text{Maximize:} && \{P(O \mid l) \end{aligned} \qquad (5)$$

Using the given HMM architecture (Fig. 3), HMMs were trained for each surgeon performing each step of the surgical procedure (8 HMM models, one for each surgeon performing one surgical procedure step). The skill level of each subjects (R1, R3, R5) was evaluated based on the statistical distance between his/her HMMs and the expert surgeons (ES)

Given two HMMs l_1 and l_2 the statistical distances between them $D(l_1, l_2)$ and $D(l_2, l_1)$ were defined by Eq. 6

$$D(l_1, l_2) = \frac{1}{T_{O_2}} [\log P(O_2 \mid l_1) - \log P(O_2 \mid l_2)]$$
$$D(l_2, l_1) = \frac{1}{T_{O_1}} [\log P(O_1 \mid l_1) - \log P(O_1 \mid l_2)] \qquad (6)$$

$D(l_1, l_2)$ is a measure of how well model l_1 matches observations generated by model l_2 relative to how well model l_2 matches observations generated by itself. Since $D(l_1, l_2)$ and $D(l_2, l_1)$ are nonsymmetrical, The natural expression of the symmetrical version is defined by Eq 7.

$$D_S(l_1, l_2) = \frac{D(l_1, l_2) + D(l_2, l_1)}{2} \qquad (7)$$

In order to scale the statistical distance between the various groups (R1, R3, R5) and the expert surgeons (ES), for each surgical procedure the statistical distance between a certain group and the expert group ($D_S(l_{Ri}, l_{ESi})$) was normalized with respect to the distance between the two experts ($D_S(l_{ES1}, l_{ES2})$) - eq. 8.

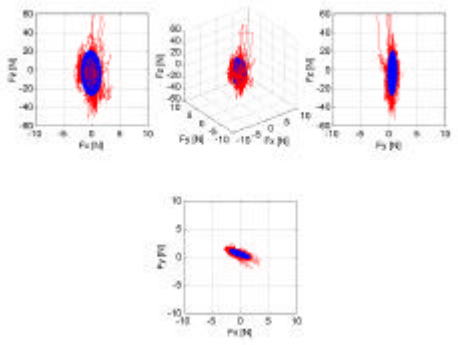
$$\overline{D}_S(l_{Ri}, l_{ESi}) = \frac{D(l_{Ri}, l_{ESi})}{D_S(l_{ES1}, l_{ES2})} \qquad (8)$$

The practical meaning of the normalized statistical distance ($\overline{D}_S(l_{Ri}, l_{ESi})$) is how far each subject from being an expert surgeons.

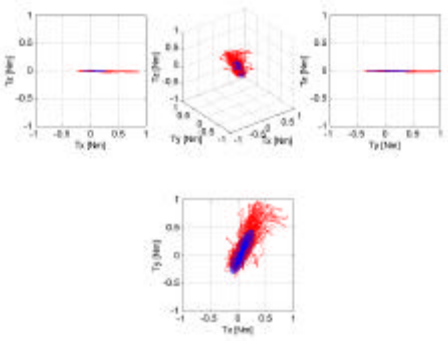
3. Results

Typical raw data of forces and torques were plotted in a 3D space showing the loads developed at the sensor location while dissecting the gallbladder fossae for 425sec by an expert surgeon during laparoscopic cholecystectomy (Fig. 4). The black ellipsoid defines a region including 95% of the F/T samples. The forces along the Z axis (in/out of the trocar) were higher compared to the

forces in the XY plane. On the other hand, torques developed by rotating the tool around the Z axis were extremely low compared to the torques generated while rotating the tool along the X and Y axis while sweeping the tissue or performing lateral retraction. Similar trends in terms of the F/T magnitude ratios between the X, Y, and Z axes were found in the data measured in other steps of the MIS procedures.



(a)



(b)

Figure 4: *Forces and torques measured at the human/tool interface while dissecting the gallbladder fossa during laparoscopic cholecystectomy. For the definitions of the X,Y,Z directions see Fig. 1b. (a) Forces (b) Torques*

Studying the magnitudes of F/T applied by R1 and ES during each step of the MIS procedures for the different tool/tissue interactions (Table 2) using the grand median analysis showed that the F/T magnitudes applied by these groups were significantly different ($p<0.05$) and task dependent (Fig. 5).

Each one of the three steps analyzed in laparoscopic cholecystectomy was represented in Fig. 5 by two pie diagrams.

The pie bar on the left hand side shows the distribution of tool/tissue interactions in which significant differences ($p<0.05$) – stippled sector, and non-significant differences ($p>0.05$) gray sector in terms of the F/T magnitudes were applied by R1 and ES. In tool/tissue interactions where significant differences were identified between R1 and ES (stippled sector), the pie diagram on the right hand side shows the correspondence between high F/T magnitudes and the group of surgeons who applied them (black sector – R1, white sector ES).

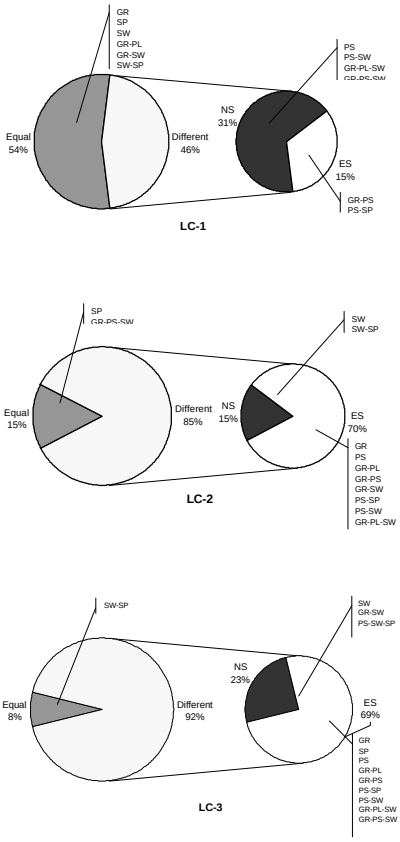


Figure 5: F/T magnitude distributions at different tool/tissue interactions applied by novice surgeons (R1) and expert surgeons (ES) during steps of laparoscopic cholecystectomy *(LC-1 Positioning gallbladder, LC-2: Exposure of cyctic duct, LC-3 Dissection of gallbladder Fossa).*

For example in LC-3 (Fig. 5) out of 92% of the cases where significant difference was observed (left hand pie chart), in 23% of the tool/tissue interactions (e.g. SW, GR-SW, PS-SW-SP) the R1 applied higher F/T magnitudes compare to the ES and in 69% of the cases (e.g. GR, SP, PS, GR-PL etc) the ES applied higher F/T magnitudes compare to the R1 (right hand side pie chart).

By dividing the surgical steps according to the nature of the tool/tissue interactions – e.g., (*i*) tissue manipulation (LC-1) (*ii*) tissue dissection (LC-2, LC-3); the results indicated that higher magnitudes of F/T were applied by the ES compared to the R1 when dissecting the tissues, whereas lower magnitudes of F/T were applied by the ES compared to the R1 during tissues manipulating.

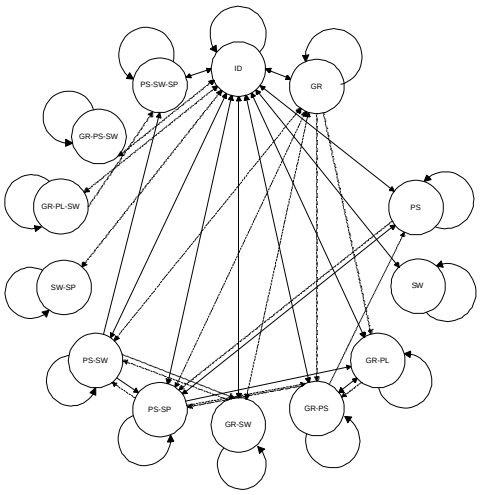


Figure 6: *The state diagram based on a frame-by-frame video analysis of dissecting the gallbladder fossa during laparoscopic cholecystectomy. Each circle represents a different state characterized by a tool/tissue interaction and the arrows represent transitions between states. In some cases, the surgeon stays within the same state; this is depicted by an arrow to the same state. (Dashed line - states and transitions preformed by ES, Dotted line - states and transitions preformed by R1 representing a novice surgeon, Solid line - states and transitions preformed both by R1 and ES)*

Analyzing videotapes of the surgical procedures incorporated the visual view of the tool/tissue interaction and graphs of the F/T at the tool/hand interface frame by frame allowed to define the primary tool/tissue interactions in the MIS procedure, and the direction of forces and torques associated with them (Table 1). Once these tool/tissue interaction archetypes were defined, each step of the surgical procedure could be manually decomposed into a list of tool/tissue interactions. This list was further transformed into a more compact diagram (as shown in Fig. 6) defining a typical tool/tissue transition diagram of a surgical procedure. The tool/tissue transition diagram (state diagram) depicted in Fig. 6 represented the surgical step in which *the gallbladder fossa* was dissected. Although Fig. 6 illustrates a unique/tool tissue transition diagram of a specific Laparoscopic procedure step, the diagram architecture was similar to all the other MIS steps under study.

The Idle state is the only state connected to all the other states in the state transition diagram (Fig. 6). This state, in which no tool/tissue interaction was performed, was mainly used by both expert and novice surgeons to move from one operative state to the other. However, the expert surgeons used the idle state only as a transition state while the novices spent significant amount of time in this state planning the next tool/tissue interaction. Another major difference between expert and novice surgeons was related to the tool/tissue interaction and tool/tissue transitions used by these two groups. Essentially Fig. 6 was constructed from two separate models representing the expert surgeon's model (ES) (solid line and dashed line) and the novice surgeon's model (R1) (solid line and dotted line). For the purpose of evaluating surgical skills using the HMM, the model representing these two groups must share the same architecture. This requirement led to generalized the model architecture as described previously in Figure 3.

This analysis demonstrated several phenomena. First, expert and novice surgeons took different paths to reach the

same goal (Fig. 6). Each group utilized states and transitions not used by the other group. Secondly, studying the median completion time of the novice surgeon group and the expert surgeon group showed a significant difference between these groups (*p<0.05*). The surgical procedure's completion time was longer for the R1 by a factor of 1.5 to 4.8 when compared to the ES. The difference between R1 and ES was more profound in steps requiring higher dexterity and manual skill compared to steps where a specific organ was placed in a specific position (e.g. positioning of the gallbladder). The main factor contributing to the significant difference in the completion times between R1 and ES was the time spent in the idle state. The R1 spent significantly more time in the idle state compare to the ES.

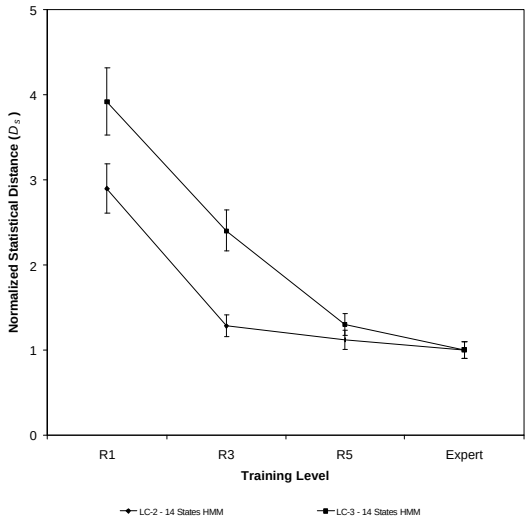


Figure 7: *The learning curve of surgical residence while performing laparoscopic cholecystectomy - Normaized statstical distance between two different HMM architectures (continues 14 state model, and discrete 4 states model) representing the performance of surgical residences (R1, R3, R5) at different year of their training and experts surgeons while performing (LC-2) and (LC-3)*

The final and the most profound analysis included the HMM. HMMs were developed for reach one of the 8 subjects (R1, R3, R5, ES) that performed the three steps of the laparoscopic cholecystectomy (LC-1, LC-2, LC-3). The normalized statistical distances ($\overline{D}_S$ - Eq. 8) between ES and R1 ,R3 ,R5 were plotted in fig 7. The power of this methodology is that it brings together different aspect of the surgical procedure into a single number which practically indicates how far the surgical performance of the subject under study is from an expert performing the same surgical task. In that way $\overline{D}_S$ provides an objective criteria for evaluating surgical performance in MIS

The objective laparoscopic surgical skill learning curve showed significant differences between all skill levels (Fig. 7). The $\overline{D}_S$ value converged to a value of one exponentially as the level of expertise increases. However, the highest gradient was between R1 and R3. This results indicate that the surgical residences acquire a major portion of their laparoscopic surgical capabilities between the first and the third years of their residency training. Calculating the $\overline{D}_S$ values for LC-1 (not plotted in Fig 7) showed no significant difference between the groups. This is correlated with the F/T magnitude differentiation analysis between the R1 and ES as indicated by Fig. 5 . The practical meaning of that result is that LC-1 does not include sufficient haptic information to use this data for differentiating between groups with different skill level. On the other hand LC-2 and LC-3 do provide such information as shown both by Fig. 5 and Fig 7.

4. Discussion

Minimally invasive surgery is a complex task that requires a synthesis between visual and haptic information. Analyzing MIS in terms of these two sources of information is a key step towards developing an objective criteria for training surgeons and evaluating the performance of virtual reality simulators incorporating haptic technology and master/slave robotic systems for teleoperation.

The magnitude of F/T applied by R1 and ES varied based on the task being preformed. High F/T magnitudes were applied by R1 compared to ES while performing tissue manipulation. This might be a result of insufficient dexterity of the R1 that might be a potential for tissue damage. However, low F/T magnitudes were applied by the R1 compared to the ES during tissue dissection, which might also indicate excessive caution to avoid irreversible tissue damage. By doing that more repetition of the dissection movements were required to be preformed by the R1 compare to ES in order to tear the tissue, a process which substantially decrease the efficiency of the MIS procedure. Using the F/T information in real-time during the course of learning as a feedback information to the R1 may improve the learning curve, reducing soft tissue injury and increase the efficiency during endoscopic surgery.

The force/torque signatures and the HMMs are objective criteria for evaluating skills and performance in MIS. The results suggest that HMMs of surgical procedures allow objective quantification of skill based on the statistical distance between HMMs representing surgical residents at different level of their training and HMMs representing expert surgeons. Moreover, this methodology can be used to determine if the performance of a student matches his/her training level.

The power of the methodology using HMM for objective surgical skill assessment is the fact that it is not limited to *in-vivo* condition as demonstrated in the current study. It can be extended to other modalities such as surgical simulators and teleoperation systems.

Increasing the size of the database to include more surgical procedures performed by more surgeons could extend the approach outlined in this study. This information, combined with other feedback data, may form the basis of teaching techniques for optimizing tool usage in MIS. The novice surgeons could practice these skills outside of the operating room on animal models or by using realistic virtual reality simulators until they had achieved the desired level of competence, and compare themselves to norms established by experienced surgeons.

Acknowledgments

This work was supported by the U.S. Army Medical Research and Material Command under DARPA grant- DMAD17-97-1-725, and by a major grant from US Surgical, a division of Tyco, Inc. to the University of Washington, Center for Videoendoscopic Surgery.

References

[1] Rosser J.C., Rosser L.E., Savalgi R.S.: Objective evaluation of a laparoscopic skill program for residents and senior surgeons. *Arch Surg*., 133: 657-661,1998.

[2] Rosser J.C., Rosser L.E., Savalgi R.S.: Skill acquisition and assessment for laparoscopic surgery. *Arch.Surg*., 132: 200-204, 1997.

[3] Derossis A.M., Fried G.M., Abrahamowicz M., Sigman H.H., Barkun J.S., Meakins J.L.: Development of a model for training and evaluation of laparoscopic skills. *Am.J.Surg.*, 175: 482-487, 1998.

[4] Derossis A.M., Antoniuk M., Fried G.M.: Evaluation of laparoscopic skills: a 2-year follow-up during residency training. *C.J.S*., 42; 4: 293-296, 1999.

[5] Melvin W.S., Johnson J.A., Ellison E.C.: Laparoscopic skill enhancement. *Am.J.Surg*., 172: 377-379, 1996.

[6]Gallagher A.G., McClure N., McGuigan J., Crothers I., Browning J.: Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). *Endoscopy*, 31(4): 310-313, 1999.

[7] Chaudhry A., Sutton C., Wood J., Stone R., McCloy R.: Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-

computer interface. *Ann.R.Coll.Surg.Engl.*, 81(4): 281-6, 1999.

[8] Satava R., Virtual Reality Surgical Simulator. Surg Endosc 7: 203-205, 1993
[9] Ota D, Loftin B, Saito T, Lea R, Keller J Virtual Reality in Surgical Education. Comput Biol Med 25: 127-137, 1995

[10] Noar M, Endoscopy Simulation: A Brave New World? Endoscopy 23: 147-149, 1991

[11] Rosen J., Advanced Surgical Technologies for Plastic and Reconstructive Surgery. Otolaryngologic Clinics of North America 31: 357-381, 1998

[12] Bicchi A. et. al., A Sensorized Minimally Invasive Surgery Tool for Detecting Tissue Elastic Properties. Proceedings of the 1996 IEEE International Conference on Robotics and Automation: 884-888, 1996

[13] Chen E. and B. Marcus, Force Feedback for Surgical Simulators, Proceedings of the IEEE, Vol. 86, No., 3, pp. 524-530, March 1998

[14] Berkely et. al, Creating Fast Finite Elemnt Modles from Medical Images, Medicine Meets Virtual Reality , MMVR-2000, Studies in Health Technology and Informatics, Vol. 70 pp. 26-32, IOS Press, 2000.

[15] Morimoto A, Foral R, Kuhlman J, Zucker K, Curet M, Bocklage T, MacFarlane T, Kory L., Force Sensor for Laparoscopic Babcock. In: Proceedings MMVR-97 (Medicine Meets Virtual Reality), K Morgan (ed) IOS Press Amsterdam Netherlands, pp 354-361, 1997.

[16] MacFarlane M., J. Rosen, B. Hannaford, C. Pellegrini, M. Sinanan, Force Feedback Grasper Helps Restore the Sense of Touch in Minimally Invasive Surgery, Journal of Gastrointestinal Surgery, Vol. 3, No. 3, pp. 278-285, May/June 1999.

[17] Rosen J., M. MacFarlane, C. Richards, B. Hannaford, C. Pellegrini, M. Sinanan, Surgeon/Endoscopic Tool Force-Torque Signatures In The Evaluation of Surgical Skills During Minimally Invasive Surgery, Studies in Health Technology and Informatics - Medicine Meets Virtual Reality, Vol. 62, pp. 290-296, January 1999.

[18] Rosen J., C. Richards, B. Hannaford, M. Sinanan, Hidden Markov Models of Minimally Invasive Surgery, Studies in Health Technology and Informatics - Medicine Meets Virtual Reality, Vol. 70 pp. 279-285, January 2000.

[19] B. Hannaford, J. Trujillo, M. Sinanan, M. Moreyra, J. Rosen, J. Brown, R. Lueschke, M. MacFarlane (1998), Computerized Endoscopic Surgical Grasper, In: Proceedings, MMVR-98 (Medicine Meets Virtual Reality), D J Westwood et. al (eds) IOS Press Amsterdam Netherlands, pp 265-271.

[20] Rosen J., B. Hannaford, M. MacFarlane, M. Sinanan, Force Controlled and Teleoperated Endoscopic Grasper for Minimally Invasive Surgery - Experimental Performance Evaluation, IEEE Transactions on Biomedical Engineering, Vol. 46, No. 10, pp. 1212-1221, October 1999.

[21] Buess G. F. et. al., Robotics in allied Technologies, Archive of Surgery Vol. 135 pp. 229-235, February 2000.

[22] Howe R.D. and Y. Matsuoka, Robotics for Surgery, Annu. Rev. Biomed. Eng. Vol. 1, 211-240, 1999

[23] Rabiner L.R., a Tutorial on hidden Markov models and selected application in speech recognition, Proceedings of the IEEE, Vol. 77, No. 2, Feb 1989.

[24] J.K. Baker, “The DRAGON System - An Overview,” IEEE Trans. Acoustics, Speech & Signal Processing, vol. 23, no. 1, pp. 24-29, 1975.

[25] S.E. Levinson, L.R. Rabiner, M.M. Sondhi, “An Introduction to the Application of the Theory of Probabalistic Functions of a Markov Process to Automatic Speech Recognition,” Bell Sys. Tech. Journal, vol. 62, no. 4, pp. 1035-1074, 1983.

[26] R. Reddy, "Foundations and Grand Challenges of Artificial Intelligence: 1988 AAAI Presidential Address," AI Magazine, pp. 9-21, Winter 1988.

[27] B. Hannaford, P. Lee, "Hidden Markov Model Analysis of Force/Torque Information in Telemanipulation," Proceedings 1st International Symposium on Experimental Robotics, Montreal, June 1989.

[28] B. Hannaford, P. Lee, "Hidden Markov Model of Force Torque Information in Telemanipulation," International Journal of Robotics Research, vol. 10, no. 5, pp. 528-539, 1991.

[29] B. Hannaford, P. Lee, "Multi-Dimensional Hidden Markov Model of Telemanipulation Tasks with Varying Outcomes," Proceedings IEEE Intl. Conf. Systems Man and Cybernetics, Los Angeles, CA, Nov. 1990.

[30] P. Pook, D.H. Ballard, "Recognizing Teleoperated Manipulations," Proc. IEEE Robotics and Automation, vol. 2, pp. 578-585, Atlanta, GA, May, 1993.

[31] M.C. Nechyba, Y. Xu, Stochastic similarity for validating human control strategy models, IEEE Transactions on Robotics and Automation. vol.14, no.3. pp. 437-51. June 1998.

[32] Jie-Yang. Yangsheng-Xu. Chen-C-S. Human action learning via hidden Markov model. IEEE Transactions on Systems, Man & Cybernetics, Part A (Systems & Humans). vol.27, no.1. pp. 34-44. Jan. 1997.

[33] K. Itabashi, K. Hirana, T. Suzuki, S. Okuma, F. Fujiwara, "Modeling and Realization of the Peg-in-Hole Task Based on Hidden Markov Model," Proc. IEEE Intl. Conf. on Robotics and Automation, pp. 1142, Leuven, Belgium, May 1998.

[34] B. Hannaford, "Hidden Markov Model Analysis of Manufacturing Process Information," Proc. IROS 1991, Osaka, Japan, Nov. 1991.

[35] B.J. McCarragher, G. Hovland, P. Sikka, P. Aigner, D. Austin, "Hybrid Dynamic Modeling and Control of Constrained Manipulation Systems," IEEE Robotics and Automation Magazine, vol. 4, No. 2, June 1997.

[36] Wachsmuth-I. Frohlich-M., eds., Gesture and Sign Language in Human-Computer Interaction, International Gesture Workshop Proceedings, Springer-Verlag, Berlin, Germany. xi+308 pp. 1998.

[37] Lien-J-J. Kanade-T. Cohn-J-F. Ching-Chung-Li. Automated facial expression recognition based on FACS action units, Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No.98EX107). Nara, Japan. pp. 390-5. 14-16 April 1998.

[38] Baldi P., Brunak S. ,Bioinformatics, MIT Press, 1998.

[39] Kwon K., "Application of HMM to accident Diagnosis in Nuclear Power Plants," Proc. Topical Meeting on Computer Based Human Support Systems, Phil. PA, Amer. Nuclear Soc. June 1995.

[40] P. Freed, "Detecting Pulsars with Hidden Markov Models," 3rd Intl. Symp. On Signal Proc. and Applications, p 179-83, vol. 1, Aug.