Development and Validation of an Assessment Model for Open Surgical Procedures

by

Neil Rittenhouse

A thesis submitted in conformity with the requirements for the degree of Master of Health Science in Clinical Engineering

Institute of Biomaterials and Biomedical Engineering
University of Toronto

© Copyright by Neil Rittenhouse 2010
Development and Validation of a Training and Assessment Model for Open Surgical Procedures

Neil Rittenhouse

Master of Health Science in Clinical Engineering

Institute of Biomaterials and Biomedical Engineering
University of Toronto

2010

Abstract

In surgical education, two of the current primary assessment tools for technical skills are the OSATS checklist and electromagnetic (EM) hand tracking. There are few bench models available for open procedures. The first two chapters of this thesis introduce these concepts further and provide detailed background knowledge. Chapter 3 explores interference that may from the concurrent use of EM tracking and cautery, finding that monopolar cautery interferes, while bipolar cautery does not. Chapter 4 discusses the validation of an assessment tool for open cholecystectomies (OCs) consisting of a porcine bench model and a Wii remote based infrared (IR) hand tracking system. The assessment tool is found to have construct validity and the face validity of the OC model is established. Chapter 5 concludes the thesis and presents several avenues of future research for the improvement of both the OC model and Wii remote based hand tracking system.
Acknowledgments

First and foremost I would like to thank my thesis supervisor Dr. Grantcharov for his constant constructive counsel and guidance throughout the thesis process. I also wish to extend my sincere appreciation to my committee members. To Dr. Sonnadara, thank you the use of your lab space, for your help with statistics, and for answering my frequent questions related to motion analysis. To Dr. Mihailidis, thank you for your assistance with the technical components of the project.

I would like to extend my appreciation to the staff and volunteers at the SSC for their assistance and for putting up with my endless string of requests. This project would not have been possible without their help. In particular I wish to acknowledge Gregory Drozdzal and Ajmal Khan for their help completing the validation experiments. For their assistance developing and troubleshooting model and assessment tool, I wish to acknowledge Glen Kurti, Babak Taati, Bharat Sharma, and Wallace Wee.

Finally, I am grateful to my family, friends, and classmates for their support over the last two years. To the Clinical Engineering class of 2010 – thank you for the friendship and memories. To my family and Karen, thank you for always being there to support me and keeping my spirits up.

To the Natural Sciences and Engineering Research Council of Canada (CGS Master’s Award), and the Network of Excellence in Simulation for Clinical Teaching and Learning, MITACS (Accelerate Scholarship), and the Institute of Biomaterials and Biomedical Engineering at the University of Toronto - thank you for your financial support.
# Table of Contents

Chapter 1 Introduction .................................................................................................................. 1  
1  Introduction.............................................................................................................................. 1  
Chapter 2 Background and Literature Survey ............................................................................ 3  
2  Background.............................................................................................................................. 3  
2.1  Surgical Models .................................................................................................................. 3  
2.1.1  Cholecystectomy Simulation ......................................................................................... 4  
2.1.2  Training in laparoscopic cholecystectomy ................................................................. 4  
2.1.3  Training in open cholecystectomy ............................................................................... 5  
2.2  Objective Assessment of Surgical Skills ........................................................................... 6  
2.2.1  Standardized Rubrics and Checklists ........................................................................... 6  
2.2.2  Motion Analysis ............................................................................................................ 6  
Chapter 3 Electromagnetic Tracking in the Presence of Cautery ............................................ 13  
3  Electromagnetic Tracking in the Presence of Cautery ........................................................... 13  
3.1  Methods............................................................................................................................. 13  
3.1.1  Setup ........................................................................................................................... 13  
3.1.2  Effects of Cautery on Electromagnetic Tracking .......................................................... 14  
3.1.3  Quantification of Error ............................................................................................... 15  
3.2  Results............................................................................................................................... 17  
3.2.1  Effects of Cautery on Electromagnetic Tracking .......................................................... 17  
3.2.2  Quantification of Error ............................................................................................... 19  
3.3  Discussion.......................................................................................................................... 22  
3.4  Summary of Patriot-Cautery Interactions ........................................................................ 23  
Chapter 4 Open Cholecystectomy Model and Assessment Tool ................................................... 25  
4  Open Cholecystectomy Model and Wii Remote Based Hand Tracking ................................. 25  
4.1  Participants......................................................................................................................... 25  
4.2  Materials............................................................................................................................. 25  
4.2.1  Open Cholecystectomy Model Design ..................................................................... 25  
4.2.2  Assessment Tool .......................................................................................................... 26  
4.3  Design............................................................................................................................... 32  
4.4  Procedure........................................................................................................................... 32  
4.4.1  Wii Tracking Validation .............................................................................................. 32  
4.4.2  OC Model and Wii Assessment Validation ................................................................. 33  
4.5  Results............................................................................................................................... 34  
4.5.1  Wii Tracking Validation .............................................................................................. 34  
4.5.2  OC Model and Wii Assessment Validation ................................................................. 37  
4.5.3  Assessment Scores ....................................................................................................... 41  
4.5.4  Survey .......................................................................................................................... 44  
4.6  Discussion........................................................................................................................... 49  
4.6.1  Wii Tracking Validation .............................................................................................. 49  
4.6.2  Model and Assessment ............................................................................................... 51  
4.7  Summary of OC Model with Wii Tracking ...................................................................... 55  
Chapter 5 Conclusions ............................................................................................................... 57  
5  Conclusions............................................................................................................................ 57  
5.1  Future Work ....................................................................................................................... 59  
References ................................................................................................................................... 61
List of Tables

Chapter 2
Table 2-1: Categories for the Classification of Surgical Simulators. ......................................................... 4
Table 2-2: Camera Parameters ....................................................................................................................... 11

Chapter 3
Table 3-1: Monopolar Cautery Trials. ......................................................................................................... 16
Table 3-2: Bipolar Cautery Trials. ................................................................................................................ 16
List of Figures

Chapter 3
Figure 3-1: Grid and Electromagnetic Tracking System Setup .................................................. 14
Figure 3-2: Translation Pattern ........................................................................................................ 15
Figure 3-3: Cautery Lead Wire and Patriot Sensor Setup ................................................................. 16
Figure 3-4: Typical Patriot Translation Result ................................................................................... 17
Figure 3-5: Typical Patriot Translation Result - 2D representation .............................................. 18
Figure 3-6: Typical Patriot Translation Result with Cautery .......................................................... 19
Figure 3-7: Monopolar Cautery Tracking Errors .............................................................................. 20
Figure 3-8: Bipolar Cautery Tracking Errors .................................................................................... 21

Chapter 4
Figure 4-1 OC Model Base ................................................................................................................. 26
Figure 4-2: OC Model and Assessment Tool Equipment Setup ....................................................... 27
Figure 4-3: Origin Square .................................................................................................................. 28
Figure 4-4: Dominant Hand IRM Schematic .................................................................................... 31
Figure 4-5: Typical Manual Assignment View .................................................................................... 31
Figure 4-6: Validation Tracking Perspective Results .......................................................................... 35
Figure 4-7: Motion Analysis Parameters for the Validation Trials .................................................... 36
Figure 4-8: Effects of a Varying Minimum Velocity Threshold on Validation Data ....................... 37
Figure 4-9: OC Motion Analysis Results ............................................................................................ 38
Figure 4-10: Effect of a Varying Minimum Velocity Threshold on OC Data .................................... 39
Figure 4-11: Correlation Coefficients for Assessment Parameters .................................................. 40
Figure 4-12: Mean Movements per Second ....................................................................................... 41
Figure 4-13: Mean OSATS Assessment Scores by Category ............................................................. 42
Figure 4-14: Mean Total OSATS Scores ............................................................................................. 43
Figure 4-15: Correlation between Motion Analysis Parameters and OSATS Scores .................... 44
Figure 4-16: Technical Training Background .................................................................................... 46
Figure 4-17: Didactic Training Background ....................................................................................... 47
Figure 4-18: Operative Experience with Open Procedures ............................................................... 48
Figure 4-19: Open Model Background and Role .............................................................................. 49
List of Appendices

Appendix A: Assessment Sheet ................................................................. 66
Appendix B: Post Questionnaire ............................................................... 67
List of Abbreviations

2D – Two Dimensional
3D – Three Dimensional
AC – Alternating Current
CCD – Charge Coupled Device
DC – Direct Current
EM – Electromagnetic
ICSAD – Imperial College Surgical Assessment Device
IR - Infrared
IRM – Infrared Marker
LC – Laparoscopic Cholecystectomy
OC – Open Cholecystectomy
OR – Operating Room
OSATS – Objective Structured Assessment of Technical Skills
PL – Path Length
ROVIMAS – Robotic Video and Motion Analysis Software
SSC – Surgical Skills Centre
TT – Total Time
VR – Virtual Reality
Chapter 1
Introduction

Public awareness of operative errors and adverse outcomes has resulted in a number of high profile media cases and the Harvard Medical Practice Study\(^1\) and To Err is Human report\(^2\). Such exposure has brought into question the traditional mantra of “see one, do one, teach one” in surgical education and prompted research into curricula to teach technical skills outside of the operating room (OR) and methods to objectively assess surgeons\(^3,4\).

Concurrently, laparoscopy has become increasingly widespread due to a number of advantages for the patient compared with conventional open technique such as reduced operative stress response, less pain, and shortened hospital stay and recovery\(^5\). Currently, laparoscopy is considered the golden standard for cholecystectomy, colorectal, antireflux, and bariatric surgery\(^6\). However, in difficult situations conversion to open technique is often necessary and thus surgeons need to be proficient with the open technique.

The development of laparoscopy has led to reduced education opportunities in open procedures for surgical trainees\(^7\). A number of training models for laparoscopic cholecystectomy have been designed and validated\(^4,8-13\). Yet in current practice there are no validated tools for training and assessment of the skills necessary for open procedures. There has been concern among educators that surgeons educated in the laparoscopic era are not proficient with open cholecystectomy\(^7\).

Traditionally surgical trainees have been evaluated on their technical proficiency on an ongoing basis by their superiors, documented through in-training reports\(^14\). Motivated by patient safety concerns and the advent of novel surgical techniques, surgical educators have sought ways to introduce methods for objectively assessing a surgeon’s technical capability. Checklists\(^15,16\), video scoring\(^17\), and efficiency measures\(^18\) are some of the techniques developed that provide better feedback and assessment of technical competence. These methods are either time consuming to apply, unable to provide real time feedback, or do not provide a comprehensive assessment. Ideally, objective assessment techniques would be similar to an examination – candidates would be able to self-practice, be tested, and provided with a grade. Such an ideal
objective assessment technique could be used to determine resident progression, as well as serve as part of surgical licensing requirements\textsuperscript{19}.

With these concerns in mind, the aim of this study was to develop and an assessment model for open cholecystectomies. All work was conducted at the University of Toronto Surgical Skills Centre (SSC) and in part is based on existing practices/technologies in use at the SSC.
Chapter 2
Background and Literature Survey

Prior to starting the study, a literature survey looking at Surgical Models and Objective Assessment Methods was done. The surgical models section focused on existing surgical models for LCs and open procedures, while the objective assessment section focused on tools for objective assessment, including technologies that could be used as assessment tools.

2.1 Surgical Models

Merriam Webster defines a model as “a device that enables the operator to reproduce or represent under test conditions phenomena likely to occur in actual performance. A surgical model then is any device that reproduces or represents surgical conditions for the purposes of training, teaching, or practicing a specific task or procedure. Surgical models have become more common as educational tools over the past 20 years and are likely to become more so as proficiency based training curriculums are adopted20.

With the increasing popularity of models, classifications were developed to group models into types. Distinguishing features of a model include fidelity, validity, the medium it is constructed from, and what it aims to teach. A summary of these interconnected categories, adapted from McDougall21, is presented in Table 2-1. Training with live animals or human cadaveric specimens is also common but is not included in this summary.
Table 2-1: Categories for the Classification of Surgical Simulators.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fidelity</td>
<td>Refers to the realism of the model. Commonly split into low fidelity and high fidelity.</td>
</tr>
<tr>
<td>Validity</td>
<td>Refers to how valid a model is for teaching. Broken down into face validity (quality of the representation), construction validity (can the model distinguish between experienced and novice surgeons), and predictive validity (ability to predict subsequent intraoperative performance).</td>
</tr>
<tr>
<td>Medium</td>
<td>Refers to the construction of the model. Typically split into virtual reality and training boxes.</td>
</tr>
<tr>
<td>Training Goals</td>
<td>A measure of what the model is trying to teach. Could be as simple as familiarity with laparoscopic instruments or as complicated as a full surgical procedure. Is related to the fidelity of the model.</td>
</tr>
</tbody>
</table>

2.1.1 Cholecystectomy Simulation

Many surgical simulators have focused on cholecystectomies as they are common, straightforward procedures. A cholecystectomy is the removal of the gallbladder and is indicated in the presence of gallstones, gallbladder trauma, and gallbladder cancer\(^{22}\). An open cholecystectomy (OC) is the old standard of care but has been supplanted by laparoscopic cholecystectomy (LC), which is now accepted as the standard of care for the majority of patients. LCs offer patients less pain, shorter hospitalization, and faster recovery\(^7\). An OC is still indicated for gall bladder cancers\(^7\) and in difficult cases due to adhesions after previous procedures, acute inflammation, or abnormal biliary anatomy.

2.1.2 Training in laparoscopic cholecystectomy

Laparoscopic training models for cholecystectomies have tended to be either virtual reality (VR) models (high and low fidelity), or box trainers using synthetic models or animal organs.
VR models are implemented using a standard computer, a display, a torso box, and laparoscopic instrument handles that, in lieu of instrument tips, are connected to sensors. During training, software depictions of the operative area and “instrument tip” locations are displayed on a monitor. Many models have different training tasks, ranging from simple navigation tasks through to full procedures (including LCs)\textsuperscript{23}. LapSim, Lap Mentor, LS500, MIST, and MIST are some of the VR models available\textsuperscript{3,8,24,9,10,25,4}. Construct validity (ability to distinguish between novices and experts) has been established for LapSim\textsuperscript{26}, LapMentor, and others. VR models have been criticized for their high cost and poor reliability that prevent widespread adoption in training curricula\textsuperscript{23} and as well as for presenting no clear advantage over cheaper box trainers\textsuperscript{27}. Nevertheless laparoscopic VR models are recognized as a valuable training tool\textsuperscript{28}.

In addition to VR models, training box models are common. These can be divided into those that use entirely synthetic materials, and those that use some form of animal tissue. Synthetic training box models have tended to focus on low fidelity models for learning basic laparoscopic skills and tasks are often as simple as peg transfer or knot tying\textsuperscript{23,29}. In general, low fidelity training boxes are relatively inexpensive and exhibit similar skill acquisition rates to more costly VR models\textsuperscript{27} though they may be less interesting to trainees\textsuperscript{30}. Mid fidelity\textsuperscript{31} and high fidelity\textsuperscript{32} training box models using synthetic material are available though costs generally scale with increasing realism of the simulated tissue and anatomy. Porcine\textsuperscript{23,33,34} and rabbit\textsuperscript{35} tissue have been used in high fidelity training boxes. Since animal organs can often be obtained from abattoirs at a relatively low cost, training boxes employing animal tissues allow for a high fidelity, low cost model. As with all training tools using animal tissues, training box models do not present accurate anatomy and raise ethical considerations regarding animal rights.

2.1.3 Training in open cholecystectomy

Whereas training models and opportunities for LCs are common, those for OC models are virtually non-existent. Visser et al. highlighted that with only 5-25\% of cholecystectomies being done using open methods, training opportunities for OCs are declining and there is a downward trend in the number of open cases to which a graduating surgical resident is exposed\textsuperscript{7}. Knowledge and skills necessary to perform an OC are still required however, since not all patients are suitable candidates for LCs and even in those that are LC candidates, intraoperative conversion from LC to OC may be required\textsuperscript{7}. Despite this, there has been little work done to
develop an OC model with the exception of a study that focused on evaluating a surgical team’s response to indications for conversion from LC to OC that halted the simulation immediately after conversion.\textsuperscript{19}

### 2.2 Objective Assessment of Surgical Skills

In addition to models for surgical education, there is a growing trend to replace traditional subjective assessment techniques with objective methods.\textsuperscript{36} Research into objective assessment methods has focused on either technical proficiency or OR management.\textsuperscript{37} Methods that have been established for assessing technical proficiency include live or video-based observation using global rating scales and procedure-specific checklists,\textsuperscript{14,38} as well as motion analysis metrics based on hand/instrument movements\textsuperscript{39,11} and gesture recognition.\textsuperscript{40}

#### 2.2.1 Standardized Rubrics and Checklists

Standardized rubrics and procedure specific checklists are well established as assessment tools. In 1996 Martin et al. validated the Objective Structured Assessment of Technical Skill (OSATS), a global rating system for evaluating a surgeon’s performance. Since then OSATS has been shown to be superior to a procedure specific checklist\textsuperscript{15} and modified versions are now routinely used\textsuperscript{41,11,19,21}. An alternative to OSATS is the Global Operative Assessment of Laparoscopic Skills (GOALS) which was developed to provide a global rubric and checklist for laparoscopic procedures.\textsuperscript{16} GOALS has demonstrated validity for dissecting the gall bladder from the liver bed\textsuperscript{16} and full LCs\textsuperscript{42} though evaluator training may be required for video assessment of full procedures.\textsuperscript{38}

#### 2.2.2 Motion Analysis

Motion capture technology “allows \textit{the} study of...complex processes \textit{of human movement} through \textit{the} analysis of kinematic data that represent the relative movements of segments connected with rotating joints”.\textsuperscript{43} As related to the assessment of surgical skills, motion analysis has focused on measures of efficiency and economy of motion. Parameters that may be analyzed include number of movements (NM), number of movements per second (NM/s), total path length (PL), adherence to the optimal path, average velocity, peak velocity, and total time taken (TT).\textsuperscript{44} These parameters are easily tracked with VR models since all information about the instruments’ location in space in known and controlled by the computer.
By contrast, determination of motion analysis parameters with training box models or intraoperatively requires additional tracking technologies. Common technologies used for motion capture include video analysis, accelerometers, electromagnetic (EM) tracking, and optical tracking\(^43\). Within surgical education, use of video analysis has focused on the blinded application of OSATS, while glove mounted accelerometers have been used for gesture recognition\(^40\). EM and optical tracking have been used to collect three dimensional (3D) position information to track hands or instruments\(^43\). From this position data, motion analysis parameters such as NM and PL can be determined\(^39\). Measure of NM, PL, and TT\(^1\) via EM tracking has shown to be strongly correlated with OSATS scores\(^45\) and implicitly surgical ability.

2.2.2.1 EM Tracking Background

EM tracking systems determine 3D position by creating a varying magnetic field to induce electric current in coils according to Faraday’s law of induction. The varying magnetic field can be produced by an alternating current (AC) at a constant frequency, or by a direct current (DC) that is periodically pulsed\(^46,47\). Both AC and pulsed DC systems consist of a magnetic field source and one or more sensors. Within each sensor are three orthogonal coils. During system operation, the magnetic field fluctuates at a known rate. As the sensors move through this fluctuating field, a current is induced in each coil proportional to the strength of the magnetic field at that point. By measuring the strength to the induced current, the 3D position of the sensor can be determined.

Both AC and DC tracking systems are subject to noise from ferromagnetic objects placed near a sensor or the source\(^46,47\). AC tracking systems are also subject to noise from magnetic eddy currents in all nearby metals\(^48\), whereas DC systems are subject to interference from the Earth’s magnetic field\(^49,50\). Hardware and software improvements have resulted in significant advances, and both AC and DC systems are marketed for intraoperative use (Northern Digital’s Aurora and Ascension’s microBIRD respectively). Surgical grade tracking systems are expensive and it is more common to find earlier generation AC systems such as the Polhemus Patriot or Polhemus Isotrak II used in surgical education applications.

---

\(^1\) Strictly speaking TT could be measured with a clock. It is grouped with other motion analysis parameters for simplicity of discussion.
To assess surgical skills, the EM sensors from the tracking system are placed on the dorsum of a surgeon’s hand, enabling hand tracking. Presently two comprehensive collection and analysis systems based on EM tracking have been developed and validated: the Imperial College Surgical Assessment Device (ICSAD), and its successor the Robotic Video and Motion Analysis Software (ROVIMAS)\(^{39}\) which adds simultaneous video recording. Both systems use a Patriot or Isotrak II (Polhemus, Colchester VT) to collect 3D position data for each hand. From the 3D position data determine NM and PL can be determined. For determining NM, a movement is defined as “an instant acceleration from one local static point followed by another static point”\(^{39}\). To ensure that only deliberate movements are included and tremors are not, a minimum velocity threshold of 1.5 cm/s is applied prior to determining NM and PL\(^{39}\).

ROVIMAS has been shown to be able to distinguish between novices and experts during intraoperative LCs\(^{39,51}\). ICSAD has shown significant differences between novices and experts in areas such as knot tying\(^{52,53}\) and labour epidural catheter placement\(^{54}\). It has also shown the immediate benefit of a one day training course for ophthalmology residents\(^{55}\), and has been used to validate a novel suturing technique\(^{56}\).

### 2.2.2.2 Infrared Tracking Background

An alternative to tracking systems based on EM fields are those based on infrared (IR) light. An IR tracking system consists of several cameras sensitive to IR light and an IR marker (IRM) on the object of interest. To provide 3D tracking, the IR cameras must be calibrated to identify the camera properties, as well as the location of the cameras relative to one another. After calibration, any IRM that is visible by at least two cameras can be tracked in three dimensions via stereo triangulation.

The source of IR light varies depending upon if the system is active or passive. In an active system, the IRM is powered, emits IR light, and each IRM can be independently turned on or off, typically at high speeds. As only one IRM is ever powered at any point in time, identification of each IRM is possible. A passive system uses IR light sources that are independent of the IRMs with the IRMs being large reflective markers. To enable unique identification of the IRMs, each object to track must have a minimum of 3 reflective markers arranged in a unique geometry. IRMs can then be identified by geometric reconstruction to achieve 3D tracking. Geometric reconstruction can be used with active IRMs to shrink the IRM size but is less common since
smaller object sizes and a faster update rate can be realized by sequencing the active IRMs\textsuperscript{57}. The benefits of an active systems are a faster update rate, better resolution, smaller marker size, and the ability to track a larger number of IRMs, whereas the primary benefit of a passive system is that each IRM is wireless (e.g. there are no wires on the object or from the object to the control box).

Commercial IR tracking systems using both active and passive technologies are available. Active systems offer excellent resolution and tracking ability, but at a premium price. Passive systems offer good resolution and tracking accuracy at a more affordable price but are limited to larger objects due to the size of the reflective marker. Both systems must have line of sight to the IRMs to be effective. Active and passive IR tracking systems are well established tools for motion analysis in other domains\textsuperscript{58,59}, and passive systems are used in image guided surgery\textsuperscript{60}. To the author’s knowledge, there is only one case of an IR tracking system being applied to the field of surgical education. In this study, a commercial IR tracking system was used to determine the rotation of a laparoscopic instrument during a laparoscopic suturing task\textsuperscript{61}. The scarcity of research combining IR tracking and surgical education is likely due to the high expense of acquiring a system.

\textbf{2.2.2.2.1 Low Cost Wii Remote Based IR Tracking}

Motivated by the high costs associated with commercial IR tracking systems, efforts have been made to develop a lower cost alternative. Hay \textit{et al} developed an IR tracking system using the IR cameras in Wii remotes (Nintendo Inc, Kyoto Japan) and a freely available camera calibration toolbox for Matlab\textsuperscript{62}. Other researchers have used the Wii remotes for novel applications such as injury rehabilitation\textsuperscript{63,64}, robot control\textsuperscript{65-67}, 3D tracking\textsuperscript{68,62}, education\textsuperscript{69,70}, and as a unique interface for conventional tasks such as looking at radiographs\textsuperscript{71-74}.

The Nintendo Wii system and its controller, the Wii remote, were released in 2007. At the time, the Wii remote was a novel input method using accelerometers and an IR camera to complement the traditional buttons. The Wii remote communicates by Bluetooth, enabling connection to a computer. A Microsoft .NET library for communication with Wii remotes was developed by Brian Peek is freely available under the Microsoft Public License\textsuperscript{75}. Technical specifications of the Wii remote are not released by Nintendo, but have been determined by the online community to include a 3 axis accelerometer, infrared camera, vibration motor, speaker, four light emitting
diodes, and assorted buttons contained. The accelerometer is an Analog Devices ADXL335 that has a sensing range of +/- 3g. The camera consists of an IR lens filter with a 30° vertical, 40° horizontal field of view mounted in front of a charged coupled device (CCD) with a resolution of 128 x 96. Images are interpolated to 1024 x 768, and built in hardware processing enables pixel domain tracking of up to 4 sources of IR light. The IR camera with the built in hardware processing to find the IR “blobs” make the Wii remote a relatively simple platform for 3D tracking as the image processing and pattern recognition required to identify objects in each camera are already available. 3D tracking just requires the calibration of the Wii remote cameras as demonstrated by Hay et al and discussed in the following section.

2.2.2.2.1.1 Camera Calibration General

To use the Wii remote cameras as IR tracking tools, it is first necessary to calibrate each camera. Calibration is the process of determining the values of the camera parameters in the camera model that allow the two dimensional (2D) pixel coordinates of an image to be accurately mapped to the 3D world coordinates of the original object. The generalized parameters of interest are summarized in Table 2-2. Savli et al summarized the general steps of camera calibration as:

1. Relate a point from the 3D world coordinate frame to the 3D camera coordinate frame. This is done using a rotation matrix and translation vector.
2. Project the 3D point in the camera coordinate frame onto the image plane to obtain the 2D projection coordinates. This is done by using a projective transformation.
3. Model lens distortion, the disparity between the actual and predicted projection coordinates. This typically uses an iterative approach as the distortion models are non-linear.
4. Transform the camera 2D projection coordinates to the computer 2D pixel coordinates.

Typically a series of points with known geometry, often a grid, are shown to the camera in a variety of orientations. Reconstruction of the known geometry allows determination of the camera parameters. With multiple cameras, each camera is calibrated independently for the same object using a common world coordinate system, often using a technique called bundle adjustment – a process of “refining a visual reconstruction to produce jointly optimal structure and viewing parameter estimates”.
Table 2-2: Camera Parameters.
Extrinsic parameters depend on the arrangement of the cameras in space, whereas intrinsic parameters depend only on a camera’s components. Descriptions adapted from the Caltech Camera Calibration Toolbox

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrinsic</strong></td>
<td></td>
</tr>
<tr>
<td>Focal length</td>
<td>The lens focal length in pixels</td>
</tr>
<tr>
<td>Principal point</td>
<td>The pixel coordinates of the point on the lens with no distortion.</td>
</tr>
<tr>
<td></td>
<td>The camera axis passes through this point. It is often in the centre of the lens.</td>
</tr>
<tr>
<td>Skew coefficient</td>
<td>Defines the angle between the x and y pixel axes. The angle is nominally 90 degrees for most cameras.</td>
</tr>
<tr>
<td>Distortion coefficients</td>
<td>Coefficients for the radial and tangential distortion models.</td>
</tr>
<tr>
<td><strong>Extrinsic</strong></td>
<td></td>
</tr>
<tr>
<td>Rotation matrix</td>
<td>Defines the rotation from the camera coordinate system to a real world coordinate system.</td>
</tr>
<tr>
<td>Translation vector</td>
<td>Defines the translation vector from the camera coordinate system to the real world coordinate system.</td>
</tr>
</tbody>
</table>

In practice, there is a freely available camera calibration toolbox that was developed by Jean-Yves Bouguet out of his PhD dissertation at the California Institute of Technology. This toolbox is capable of determining all of the intrinsic and extrinsic parameters for both cameras in a stereo camera setup by acquiring N frames (N>100) of a checkerboard visible to both cameras at various angles. This toolbox is commonly known as the Caltech Camera Calibration Toolbox and was the basis of the work by Hay et al for their Wii remote based 3D tracking system.

A limitation of the Caltech Camera Calibration Toolbox is that it is only capable of calibrating up to two cameras. For 3 or more cameras, Bouguet refers users to Thomas Svoboda’s Multi-Camera Self-Calibration toolbox. The Svoboda toolbox requires the user to translate a single
point around the desired volume. The potential calibration volume is restricted to those locations that are visible by three or more cameras. For the single point of light, Svoboda et al used a laser in a darkened room. With an IR based camera system, such as the Wii remotes, the process is much simpler as a single IR source can be used.

With a series of points to create a blob for calibration, the Svoboda algorithm roughly follows these steps:

1. Discard misdetected points. In practice this includes reflections, poorly synchronized frames, or any points that lie too far from epipolar lines (lines that appear as a point in one camera but a line in another).
2. Determine the projective structure and optimize it using Bundle Adjustment.
3. Return to 3D space, assess the reprojection error, and continue removing outliers until none remain.
4. Determine the non-linear distortion model by iterating the entire process.
5. Move the real world coordinate system from the centroid of the blob to a known location.

Svoboda’s toolbox requires that there be a minimum of three identical cameras with square pixels (or a minimum of eight cameras if the cameras are not identical). It also can become unstable during adjustment for radial distortion.
Chapter 3
Electromagnetic Tracking in the Presence of Cautery

Previous EM tracking experiments at the SSC have anecdotally shown large signal noise that cannot be explained by known sources of noise (e.g. metallic objects). Upon closer inspection, it was realized that this noise was often observed during the operation of electrosurgery equipment, commonly called electrocautery or shortened to simply cautery. Cautery cuts or coagulates tissue by passing a large current through a narrow point, creating a high current density that in turn leads to heating. The different modes of operation are achieved by using high frequency waveforms with varying duty cycles. In addition to the modes of operation, cautery machines offer the ability to change the base power, scaled from 0-100.

Two types of cautery systems are available: monopolar and bipolar. A monopolar system has a cautery pencil with a fine tip for one electrode and a large grounding pad for the return electrode. Bipolar cautery systems have both electrodes within the instrument and without the use of a grounding pad. During operation, monopolar cautery is operated by buttons on the pencil while bipolar cautery is operated via foot pedals. Typical operation of both systems is in a burst manner, with the cautery only activated when required by the surgeon to control bleeding, cut tissue, etc.

Since cautery relies on a varying, non-static, waveform, there is a variable current present. By Faraday’s law, this varying current will induce a magnetic field, which in turn would be expected to interfere with EM tracking fields, distorting the position accuracy. This chapter discusses the experiments done to confirm the presence of this field distortion and to quantify its magnitude.

3.1 Methods

3.1.1 Setup

To assess if cautery affects EM tracking, a Patriot (Polhemus, Colchester VT) EM tracking system was used. The Patriot EM source was aligned to a grid (30 cm by 27 cm with 3 cm squares), and the grid and source were placed in a volume free of metallic objects (Figure 3-1). To assess both positive and negative position values, the origin was shifted from the default to a central point on the grid.
To complete the setup, a cautery machine (Valleylab ForceFX-C) was wheeled close to Patriot setup, with a bovine tissue specimen placed in a nearby tray. The lead wire of a monopolar cautery pencil was taped directly Patriot sensor in line with the sensor’s X-axis. The Patriot sensor with lead wire was then affixed to a plastic ruler to facilitate movement of the sensor. Monopolar cautery was chosen since the suspicion that cautery affects EM tracking arose from simulated procedures where monopolar cautery was used.

![Image](image.png)

**Figure 3-1: Grid and Electromagnetic Tracking System Setup.**

### 3.1.2 Effects of Cautery on Electromagnetic Tracking

Following setup, the Patriot sensor with lead wire attached was translated stepwise through each row and column on the grid (Figure 3-2). At each 3cm interval, the translation was paused for 2 seconds. Five repetitions of the translation through the grid were made with the cautery machine set to fulgurate at a power of 33. The cautery was operated in the typical manner with the pencil often losing contact with the tissue.

Five trials were conducted with cautery, and five trials without cautery.
Figure 3-2: Translation Pattern. X-Y view is shown with +Z out of the page. The pattern was designed with no motion in the Z-axis.

3.1.3 Quantification of Error

To measure the error introduced by cautery, a Patriot sensor was re-zeroed at a single point in a volume free of metallic objects. The cautery lead wire was taped immediately beside the Patriot sensor and collection was run for 2 seconds (Figure 3-3). Monopolar cautery was run at the settings in Table 3-1 and bipolar according to the settings in Table 3-2. All trials had three repetitions. To reduce the number of trials required, a typical, constant power of 33 was used. To ensure constant current flow, care was taken to ensure that the cautery pencil remained in contact with the tissue at all times.
Figure 3-3: Cautery Lead Wire and Patriot Sensor Setup

Table 3-1: Monopolar Cautery Trials.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Cautery Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Off</td>
</tr>
<tr>
<td>2</td>
<td>Cut - Low</td>
</tr>
<tr>
<td>3</td>
<td>Cut – Pure</td>
</tr>
<tr>
<td>4</td>
<td>Cut – Blend</td>
</tr>
<tr>
<td>5</td>
<td>Coagulate – Desiccate</td>
</tr>
<tr>
<td>6</td>
<td>Coagulate – Fulgurate</td>
</tr>
<tr>
<td>7</td>
<td>Coagulate - Spray</td>
</tr>
</tbody>
</table>

Table 3-2: Bipolar Cautery Trials.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Cautery Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Off</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>Medium</td>
</tr>
<tr>
<td>4</td>
<td>Macro</td>
</tr>
</tbody>
</table>
3.2 Results

3.2.1 Effects of Cautery on Electromagnetic Tracking

Tracking through a grid without cautery shows a steady progression up and down in the Y-axis as expected, as well as periodic progression in the X-axis (Figure 3-4 and Figure 3-5). Of note is the misalignment between the tracking motion and grid in Figure 3-5, indicating that the origin was not precisely aligned with the movement grid. Noise is introduced into the data when cautery is run during grid translation (Figure 3-6).

![Graphs showing tracking results](image)

**Figure 3-4: Typical Patriot Translation Result.**
Motion is stepwise in the Y-axis with periodic shifts in the X-axis. Refer to Figure 3-2 for nominal path.
Figure 3-5: Typical Patriot Translation Result - 2D representation.
Refer to Figure 3-2 for nominal path.
Figure 3-6: Typical Patriot Translation Result with Cautery.
Motion is stepwise in the Y-axis with periodic shifts in the X-axis. Refer to Figure 3-2 for nominal path.

3.2.2 Quantification of Error

Monopolar cautery produces errors in the Patriot 3D position for all modes of operation. For all modes, a large mean positive error in the Y-axis (Figure 3-7a), a large mean negative error in the X-axis (Figure 3-7b) is observed. All three axes for all modes of operation exhibit absolute maximum errors greater than 20 cm (Figure 3-7c). Coagulate generally introduces higher errors than cut. Bipolar cautery shows minimal mean positive error, and minimal mean negative error (Figure 3-8a-b). Maximum errors for bipolar cautery are negligible on low and medium operation modes but noticeable when operating in macro mode (Figure 3-8c).
Figure 3-7: Monopolar Cautery Tracking Errors.
Results are (a) positive mean error, (b) negative mean error, and (c) absolute max error. Error bars represent standard error of the mean.
Figure 3-8: Bipolar Cautery Tracking Errors.
Results are (a) positive mean error, (b) negative mean error, and (c) absolute max error. Error bars represent standard error of the mean.
3.3 Discussion

As was anticipated, monopolar cautery distorts the magnetic fields of EM tracking and makes tracking impossible. The noise observed was in bursts, a fact which can be explained by the manner in which cautery is typically used. Even if the button on the cautery pencil is held, the current flow is impeded if the pencil tip is not in contact with the tissue/grounding pad. This burst nature of the noise is confirmed by the absence of the bursts during the error quantification, where it was ensured that the pencil did not leave tissue during collection.

Quantification of the cautery effects showed a negative bias in the X-axis and a positive bias in the Y-axis for monopolar cautery. These biases are due to the position of the orthogonal coils in the Patriot sensor and would be expected to change if the position of the cautery wire relative to the Patriot sensor was changed. Quantification of the effects of bipolar cautery reveals minimal interference with EM tracking in low and medium operation modes while a small, noticeable absolute error is seen in macro operation. The near absence of error from bipolar cautery can be explained by the fact that there are anti-parallel currents in the two lead wires. These anti-parallel currents induce opposite magnetic fields that effectively cancel out. During macro operation (highest bipolar setting), the imperfect cancelling of these fields becomes noticeable due to the larger currents involved.

The quantification of error occurred for continuous operation of all modes at a constant power of 33 with the lead wire directly beside the EM sensor. It is emphasized that this reflects a worst case scenario for the specified power setting. In practice these circumstances are unlikely to occur as cautery is more likely to be run in a burst mode than continuously. Even if the cautery were run continuously, the field strength will drop proportional to \( \frac{1}{r^2} \), where \( r \) is the shortest distance between the cautery lead wire and the EM sensor. Thus the maximum error case is unlikely to occur.

Regardless, it is clear that monopolar cautery affects EM tracking in a non-linear fashion and care should be taken to not use the two simultaneously. Motion analysis studies that combine the two may produce results that appear correct, but the underlying data will have been distorted by the interference of the cautery induced magnetic field.
It is emphasized that the observed signal distortions are in addition to the well documented effects of metallic objects distorting the tracking accuracy. Furthermore the effects of cautery have only been confirmed for the Polhemus Patriot, an older generation AC based system and not for pulsed DC or newer AC based EM tracking systems. Due to the underlying physics it is unlikely that any EM tracking system will be able to track accurately in the presence of monopolar cautery. Consultation with Northern Digital Inc. confirms that cautery does distort the Aurora signal\textsuperscript{81}. This implies that EM tracking systems used for guided surgery should not be relied upon while monopolar cautery is in use. Despite this Northern Digital Inc. currently has no guidelines for the concurrent use of EM tracking and cautery\textsuperscript{81}. It is possible that the distortions from monopolar cautery may be sufficiently large that surgeons will not rely on EM tracking while using monopolar cautery but this would not extend to bipolar cautery. As even small errors in position can affect the outcome of an operation, bipolar cautery could lead to potential errors. Thus it is recommended that guidelines for the combined intraoperative use of cautery and EM tracking be established. These guidelines should alert surgeons to the potential tracking errors and, with further research, should cover cases when intraoperative cautery and EM tracking can be used (e.g. potentially bipolar cautery in low operating modes).

Though these intraoperative EM tracking systems are still subject to noise from monopolar cautery, they do offer much smaller sensors than the Patriot, as well as rejection of noise from metallic objects. With a proper glove to mount the sensors in, such a system (e.g. Northern Digital Inc.’s Aurora or Ascension’s microBird) could potentially allow a more detailed analysis of hand movements during simulated surgical tasks that do not use monopolar cautery. Were such a system to be considered for assessment via motion analysis, the effects of the metallic instruments would have to be explored.

3.4 Summary of Patriot-Cautery Interactions

As anticipated, monopolar cautery has been shown to have significant distorting effects on EM tracking. By contrast, bipolar cautery did not show significant distorting effects due to the cancelling of the induced magnetic fields by the antiparallel currents. Upper limits on the magnitude of these distorting effects at a power of 33 for all modes of operation using a Polhemus Patriot and Valleylab ForceFX-C cautery machine have been established. These limits will not translate to higher power settings, and, as the cautery waveforms may vary by
manufacturer, are unlikely to be transferable to different cautery machines. It is anticipated that different EM tracking systems will respond in a similar manner, though the collection frequency of each system may influence the results. The best general guide is that EM tracking cannot accurately track objects in the presence of monopolar cautery and that caution should be exercised when using combining EM tracking and bipolar cautery.
Chapter 4
Open Cholecystectomy Model and Assessment Tool

It has previously been mentioned that there are potentially deficiencies in present OC training that residents undergo. To address this concern, an OC model is developed and validated. As monopolar cautery is used in OC and the distorting affects of cautery on EM tracking were documented in Chapter 3, an alternative assessment tool beyond EM tracking is required. Thus a novel assessment tool based on Wii remote IR tracking is developed concurrently with the OC model. This chapter discusses the development and validation of the OC model and Wii tracking system as an assessment tool.

4.1 Participants

To validate the OC model and assessment tool, eleven novices and five experts performed a single OC with the OC model utilizing monopolar cautery as is typical for OCs. Novice surgeons were defined as residents who had performed fewer than 10 OCs and 20 LCs while experts were staff surgeons and fellows who had performed a minimum of one hundred cholecystectomies (LCs or OCs).

4.2 Materials

4.2.1 Open Cholecystectomy Model Design

To simplify development, the OC model was designed based on existing, high fidelity LC models used at the SSC. The OC model consists of an existing plastic torso covered a silicon skin pad with a 6” right subcostal incision. A standard yellow biohazard bag was taped to the base of the torso to enable easy clean up and disposal after each use of the model. To provide a mounting platform for the liver, 3 layers of hand towels folded in thirds were placed in a plastic bag, wrapped an incontinence pad, and taped to the biohazard bag in the base of the torso (Figure 4-1). For each use of the model, a porcine liver was sourced from an abattoir. The liver was prepared by removing the medial and lateral lobes, leaving the central two lobes, the gallbladder and bile ducts. The prepared liver was stapled to the towel mounting platform using a skin stapler and a monopolar cautery grounding pad was affixed to its anterior surface. Note that the grounding pad must be present as only monopolar cautery can be used for OCs.
4.2.2 Assessment Tool

To improve upon the available assessment tools, custom data collection software was developed using LabVIEW (National Instruments, Austin Texas). The developed software provided synchronized data collection from two video perspectives at 30 frames per second each, a Polhemus Patriot at 60 Hz, and up to 5 Nintendo Wii remotes at 60 Hz. Communication with the video devices was via IEEE 1394a (Firewire 400), and communication with the Patriot was over a standard 9 pin, RS232 serial port. Wii remote communication was done with a USB Bluetooth adapter (DLink DBT-120) using the freely available Managed Library for Nintendo’s Wiimote for Microsoft’s .NET programming environment. The IR cameras in the Wii remotes were used as the basis for a calibrated IR based 3D tracking system. To affix IRMs to objects, assorted circuit boards with powered IR emitters were constructed. These IR emitters were wired using standard 3.5 mm audio cables to provide power. With the wired setup, independent control of each IRM was possible and was used for origin identification (refer to 4.2.2.2). The IRMs could easily be made wireless by attaching a battery to provide power, though this would sacrifice the independent control of each IRM. The IR emitters used had a 180° field of view optimized for 950 nm wavelength. To determine the number of Wii remotes to include in the system, cases with three, four, and five remotes were setup, calibrated, assigned an origin, and testing by moving a test IRM to the origin, +X, and +Y locations (refer to 4.2.2.2 for details on these steps). Three and four remotes were found to provide acceptable tracking, while it was not possible to achieve an acceptable calibration with five remotes. As four remotes offered a better field of view than three remotes, four remotes were chosen for the system. Note that it was not possible to try a setup with more than five remotes as a limitation of the equipment used was that a maximum of five Wii remotes could be connected to the computer at any one time.
4.2.2.1 Assessment Tool in the Operative Field

To setup the operative field the OC model was placed in the centre of a raised table. Four Wii remotes were mounted on laparoscopic camera holders (Limbs and Things, Bristol UK). The mounted Wii remotes were then clamped to the table around the OC model. Relative to the model, two were clamped to the inferior and two to the superior (Figure 4-2a). The Wii remotes were numbered in the order they were connected to the computer and this numbering is indicated in Figure 4-2b. The perspectives were chosen so that the surgeon’s left hand would be tracked by pair 2-4 and the right hand by the pair 1-3. Defining the distance the Wii remotes could be from the OC model required finding a balance between perceived IRM size and field of view. This was done by setting up close, medium, and far distances, calibrating the remotes, assigning an origin, and translating a test IRM to the three points of the origin (refer to 4.2.2.2 for details on these steps). If the remotes were too far away, the IRMs on the origin and the test IRM appeared too small to the cameras and inaccuracies were introduced due to the limited resolution of the CCD (128x96). If the remotes were too close, the effective volume was reduced significantly. Through an iterative process, an effective distance was found to be around 1 m from the centre of the model. Angling of the remotes was refined through pilot testing to achieve the acceptable perspectives for tracking the surgeon’s hands.

![Figure 4-2: OC Model and Assessment Tool Equipment Setup.](image)

The (a) operative theatre with Wii remote placement and numbering and (b) the schematic showing the numbering of the remotes are seen.

4.2.2.2 Calibration

Before the Wii remotes could be used tracking, it was necessary to calibrate their cameras to find the intrinsic and extrinsic camera parameters (see Chapter 2). To calibrate the cameras, a single
IR emitter was translated throughout the desired calibration volume. The pixel coordinates of the resulting blob were collected on a per camera basis. These coordinates were then used as the input data for Svoboda’s Multi-Camera Self-Calibration toolbox\(^8\) to find the camera parameters and completing calibration.

Following calibration, the origin of the world coordinate system was located at the centroid of the calibration blob and had unknown scaling. To achieve meaningful 3D tracking, it was necessary to move the origin to a known location and determine the scaling factor. To do so, a square with four IRMs (Figure 4-3) was placed in the centre of the calibrated volume in the desired orientation. To create an origin, three of the IRMs on the square were turned on in sequence. The 3D coordinates of each IRM were determined in the original world coordinate system. Then the rotation matrix and translation vector from each camera to the new coordinate system were determined. The scaling factor was found taking the known square side length (in centimetres) and dividing by the mean of the measured sides.

![Figure 4-3: Origin Square.](image)

Yellow circles are activated IRMs and the dark red circle is an inactive IRM. IRMs are sequenced \(O, +x, +y\). Side lengths (\(l\)) are equal and are used to determine the scaling factor.

During prototyping, it was found that perspectives that are collinear, such as 1-4 and 2-3, were found to be highly inaccurate. Thus useful pairings were identified as 1-2, 1-3, 2-4, 4-4. Furthermore, the system accuracy was found to be sensitive to the origin location. This sensitivity appeared to be exacerbated when the IRMs of the square used for the origin were approximately collinear with the camera axis. The sensitivity to the origin may then be a limitation of the system stemming from the limited resolution of the Wii CCD.
4.2.2.3 IRM Identification and Triangulation

With camera calibration complete and an origin defined, the next step was to define how the pixel coordinates of the IRMs would be mapped to the real world IRMs and then triangulated to obtain 3D coordinates and achieve an IR tracking system. As sequencing the IRMs (like in an active system) was the easiest method, this was tried first. Unfortunately, due to history in the onboard hardware processing for the Wii remotes, when tracking two objects the fastest update rate that could be achieved was 40 ms between samples. This implies 80ms between updates for a single object and is too slow for motion analysis. Alternatively, geometric reconstruction could have been used to identify the mapping from real world IRMs with their pixel values. Unfortunately tracking of two hands in this manner required a minimum of 6 IRMs to be used (3 per hand) and the Wii remotes are only capable of tracking up to 4 IRMs. Thus it was necessary to apply other methods to map the real world IRMs to their proper pixel identities.

For one Wii remote camera, the numbering of the IRM pixel identities at any point in time is determined by the order the real world IRMs become visible to that camera. This implied that in multi-IRM tracking, for a given real world IRM and a single camera, the mapping of that particular IRM to a pixel identity was variable over time, necessitating frequent updates of this mapping. To achieve 3D coordinates for a single IRM, it was necessary that the map for that IRM be established in at least two cameras at all time points, magnifying the complexity.

For the Wii system to be an assessment tool, in practice it was not necessary to be completely accurate in the identification of the map between IRMs and pixel coordinates. As long as the IRMs on one hand were could be identified as belonging to that hand (and not the other), the system would be able to track hand movements. This simplification was justified so long as the IRMs were on a rigid body since then they would move together. If some of the IRM identifications on the rigid body were reversed, small positional errors will be introduced\(^1\) but the general location of the rigid body in space could still be determined. By using this principle, hand tracking can be achieved by mounting a single IRM on one hand and multiple IRMs on the other. In practice, the single IRM was mounted on the non-dominant hand and three IRMs in a right angle triangle orientation (Figure 4-4) were mounted on the dominant hand

---

\(^1\) If applicable, orientation information will be distorted as well.
Based on this, the identification of the hands during the OC trials was done in three steps:

1. Apply a minimum Euclidean distance classifier (MED) to the raw pixel coordinates for each camera. This was to assign cases where the single IRM on the non-dominant hand could be identified.

2. Explicitly identify the dominant hand IRMs by triangulating all possible pairings. Prior to triangulating a pairing, there was a manual check to confirm that the 3 IRMs considered all corresponded to the dominant hand. After triangulation of all possibilities, IRM assignments were made based on minimizing the error given by:

\[
Error = \left| \frac{\alpha}{90} \right| + \left| \frac{l_{12} - L_{12}}{L_{12}} \right| + |l_{123} - 1|
\]

where \( \alpha \) is the actual value of the right angle in degrees, \( l_{12} \) is the ratio between sides \( l_1 \) and \( l_2 \), \( L_{12} \) is the physical ratio between \( l_1 \) and \( l_2 \) (2 for the triangle used), and \( l_{123} \) is the Pythagorean relationship between the sides (Figure 4-4). These parameters were easily identifiable properties of the right angle triangle that could be related to known values or ratios. The weighting of each part was arbitrary but worked in test scenarios.

3. Identification of any remaining IRMs via manual assignment. Manual assignment of the IRMs proceeded by examining graphs of the x and y pixel coordinates in time, with each assignment based on a frame consisting of IRMs that need to be assigned. Based on the four waveforms visible, hand separation was possible since the movement of the IRM on the non-dominant hand was distinct from that of the three IRMs on the dominant hand (Figure 4-4). For expediency, the focus was on separating the hands, not uniquely identifying each IRM on the dominant hand and thus IRM assignment on the dominant hand was done arbitrarily. When the number of IRMs visible to a given Wii remote was less than 3 (making the aforementioned technique impossible), IRM assignment was done by inspecting the video and, where that failed, by probability that a given remote saw a given hand (e.g. remote 4 was more likely to see the left hand and therefore if only one IRM was visible to remote 4, it was assumed that the visible IRM corresponded to the left hand). During the IRM assignment process, jumps were noticed in the data and corrected for.

All triangulation of IRM pairs was done using the stereo triangulation function from the Caltech Camera Calibration toolbox.
Figure 4-4: Dominant Hand IRM Schematic.

Figure 4-5: Typical Manual Assignment View
Colours correspond to the numbering the Wii remote assigned to the IRMs, which in turn corresponds to the order the IRMs became visible to the Wii remotes. Colours are ordered blue = 1, red = 2, green = 3, and purple = 4. The black bars in the middle indicate the frame to be assigned. The single IRM ion the non-dominant hand can be clearly distinguished as the purple and blue trace in the upper portion of the X-axis graph and lower portion of the Y-axis graph. Jumps in the data are indicated and were corrected during the manual IRM assignment process.
4.2.2.4 Determination of Motion Analysis Parameters

To apply the Wii tracking system as an assessment tool, the motion analysis parameters NM, PL and TT were determined from the 3D hand data. These parameters were chosen as they are the parameters most commonly used in the literature, though other parameters such as average velocity and peak velocity could be determined if desired. To simplify comparison between novices and experts, NM and PL comparisons were reported as the total for both hands. ICSAD and ROVIMAS set a minimum velocity threshold value of 1.5 cm/s that must be exceeded for hand motion to be considered deliberate and not simply tremor\textsuperscript{39}. As part of the validation of the Wii system the validity of the 1.5 cm/s threshold value was checked and revised.

4.3 Design

Validation of the Wii tracking system was done by empirically checking the tracking accuracy. Validation of the OC model and of the Wii system as an assessment tool was done by comparing the performance of novices and experts.

4.4 Procedure

With the instrumentation developed, the Wii tracking system was validated for tracking accuracy prior to being validated as an assessment tool concurrently with the OC model. The methods used for both validations are outlined in this section.

4.4.1 Wii Tracking Validation

To validate the Wii tracking in real world coordinates, the Wii and Patriot origins were aligned to a 30 cm by 27 cm grid with 3 cm spacing. A Patriot sensor with an IRM attached was translated stepwise through each row on the grid (Figure 3-1). At each 3cm interval, the translation was paused for 2 seconds. Five repetitions of the translation through the grid were made. Data was separated into five “modalities” – the Patriot, and each of four distinct pairs of Wii remotes.

Previous literature set a minimum velocity threshold to EM velocity data prior to determining NM and PL\textsuperscript{39}. This threshold was to focus on deliberate movements and eliminate hand tremor. The threshold values in the literature were empirically derived specifically for EM hand tracking and were set at 1.2 cm/s for a hand resting on a surface and 1.5 cm/s for a hand in free air\textsuperscript{39}. As
these are specific for hand tracking, it was inappropriate to apply the validation data. Thus NM and PL for every trial were calculated without a minimum velocity threshold. From these results, mean NM and mean PL for each modality were determined and compared with the nominal values of 109 and 327 cm.

To assess the effects of introducing a minimum velocity threshold, mean NM and PL from each modality were found while the threshold value was varied from 0.16 cm/s in 0.01 cm/s steps. This range includes the thresholds of 1.2 cm/s and 1.5 cm/s used in the ROVIMAS system.

### 4.4.2 OC Model and Wii Assessment Validation

During the procedure, subjects’ hand motions were tracked using the Patriot and Wii systems. Each procedure was filmed from two perspectives: an overhead perspective for subsequent assessment, and a table perspective for the removal of extraneous movements from the motion tracking data. All subjects were instructed about the recording methods in place and were not compensated for their time. All procedures were assisted by a graduate student with no prior surgical training. To compensate for the untrained assistant, subjects were instructed in advance that they must use explicit instructions when communicating with the assistant. There was no scrub nurse during the sessions – subjects were asked to select and retrieve their own instruments from a tray placed above the groin area of the “patient”.

To conduct the motion analysis it was first necessary to process the data. First the Wii data was processed according to the details in section 4.2.2.3 to identify the IRMs and separate hands. Following this, extraneous movements were removed from the Wii and Patriot datasets through video analysis. Due to the absence of a scrub nurse and constraints on the Wii tracking volume, all movements to select an instrument were deemed extraneous and removed from the motion analysis for both modalities. To finish the data processing, a minimum velocity threshold was set at 0.5cm/s for the Wii system based on the Wii validation results. The threshold value was set at 1.5cm/s for the Patriot to match established values in the literature.

After the removal extraneous movements, the total NM and PL for both hands for each subject was determined and mean values for NM, PL, and TT for novices and experts were found. Differences in these parameters between novices and experts were checked using a two sided
Mann-Whitney U-test with a significance level of 0.05. Following this, the effect on NM and PL of varying the threshold from 0-2 cm/s in 0.1cm/s intervals was checked descriptively.

To check the relation between NM, PL, and TT, Pearson R correlation coefficients for these three parameters were found. The link between NM and TT was further explored by determining NM/s and checking for significance. As before tests for significance between novices and experts were done using a two sided T-test with a significance level of 0.05.

To assess the validity of the OC model, the overhead video was sent to staff general surgeon for blinded assessment. The staff surgeon was instructed to use an OSATS scoring system, a copy of which is in Appendix A. Test for significance between novices and experts was done using a two sided Mann-Whitney U-test, p<0.05.

Following the their sessions, all novice subjects were emailed a questionnaire asking about their training experiences, OC/LC background, comments on the accuracy of the OC model, and suggestions for improving the OC model. Results of the survey were analyzed in a qualitative manner with the results for select questions reported. A copy of the survey is available in Appendix B.

4.5 Results

4.5.1 Wii Tracking Validation

Mean 3D positions for all five modalities (the Patriot plus the 4 Wii remote pairs) and the nominal 3D position are presented in Figure 4-6. As can be seen from the X-Y view (Figure 4-6a), there is a counter clockwise rotation away from nominal about the Z-axis for the Patriot and a clockwise rotation away from nominal about the Z-axis for the Wii pairs. The 3D plot reveals (Figure 4-6b) that the Wii plane of movement is skewed relative to the nominal. All three perspective views (Figure 4-6a,c,d) reveal a skew in the Wii 1-2 pair that gets worse with increasing X-axis values.
Figure 4-6: Validation Tracking Perspective Results.
Perspectives are (a) X-Y plane overhead (b) 3D (c) X-Z plane front and (d) Y-Z plane side. Position is the mean of five trials at each nominal location (pauses in translation) with the actual movement between pauses excluded. Error ellipsoids represent standard error of the mean in the X, Y and Z directions.

From the grid translation, mean NM and PL are calculated without a minimum velocity threshold. The resulting mean values for NM and PL are above the nominal value for all modalities (Figure 4-7a-b). Wii pair 1-2 has a mean NM and mean PL similar to the Patriot, while the other Wii pairs exhibit mean NM and PL values larger than the Patriot.
Figure 4-7: Motion Analysis Parameters for the Validation Trials.
Results are mean of five trials and all five modalities are shown. Parameters are (a) mean NM and (b) mean PL. Data is without a minimum velocity threshold. Error bars represent standard error of the mean.

The minimum velocity threshold was varied and related to mean NM (Figure 4-8a) and mean PL (Figure 4-8b) from each modality. The threshold range considered is 0-1.6 cm/s in 0.01 cm/s steps, which includes the thresholds of 1.2 cm/s and 1.5 cm/s used in the ROVIMAS system. NM and PL for all modalities exhibit a logarithmic decline in value until a threshold around 0.5 cm/s, after which the values are essentially 0.
Figure 4-8: Effects of a Varying Minimum Velocity Threshold on Validation Data. Results are the mean of five trials for (a) NM and (b) PL. Note how application of the ROVIMAS values would eliminate all of the data. Error bars are excluded for clarity.

4.5.2 OC Model and Wii Assessment Validation

Of the 11 novices who participated in the OC validation, two failed to complete the set task; one worked in a 10” long incision (instead of a standard 6” long incision) due to the failure of the skin pad and the other cut the cystic cut and artery together instead of separating the triangle of Calot. Thus these two subjects are excluded from motion analysis and OSATS assessment. In addition, one subject is excluded from the Wii analysis as the Wii remote field of view did not adequately cover his/her hand movements, and another subject is excluded from the Patriot analysis because the Patriot did not function during the procedure. Both subjects that were selectively excluded are novices, leaving 8 novices for the Wii and Patriot analyses, 9 novices for other analyses, and 5 experts for all analyses. All novices who participated were sent the
survey regardless of their above exclusions. All survey responses received are included in analysis of the survey results.

After these exclusions, total NM, PL, and TT for novices and experts are compared. Novices had a significantly larger NM value compared to experts when measured by both the Patriot (novice median = 3779.5, expert median = 1897) [Mann Whitney U=20, n1=8, n2=5, p=0.016] and the Wii (novice median = 3856.5, expert median = 2044) [Mann Whitney U=20, n1=8, n2=5, p=0.093] (Figure 4-9a). Mean NM for both groups are smaller when measured with the Patriot than with the Wii. There was no significance difference between novices’ and experts’ PL when measured using the Patriot (novice median = 2.945 km, expert median = 1.419 km) or the Wii (novice median = 9843 cm, expert median = 4529.5 cm) (Figure 4-9b) though for both groups PL measured by the Patriot is over an order of magnitude larger than that measured by the Wii. TT reveals that novices (median = 36.18 min.) took significantly longer than experts (median = 19.53 min.) [Mann Whitney U=20, n1=9, n2=5, p=0.016] (Figure 4-9c). During data collection for all subjects, strong noise was observed in the Patriot signal when the cautery was operational.

![Figure 4-9: OC Motion Analysis Results.](image)

Results for (a) NM (N=8) (b) PL (N=8) and (c) TT (N=9) are shown. * indicates a significance of p <0.05 for a Mann-Whitney U-test. Error bars represent standard error of the mean.

As with the validation of the tracking of the Wii system, a varying minimum velocity threshold was applied to the OC data. Varying the threshold shows a consistent difference (though not necessarily a significant difference) between novices and experts in NM and PL regardless of the
exact value used (Figure 4-10a-d). As measured by both the Patriot and Wii systems, NM and PL decline with an increasing minimum velocity threshold value and, with the exception of the Patriot PL values, novices and experts show NM and PL values distinct from one another.

**Figure 4-10: Effect of a Varying Minimum Velocity Threshold on OC Data.** Results are the mean for each group for the parameters (a) Patriot NM, (b) Patriot PL, (c) Wii NM, and (d) Wii PL. Note that there is a clear difference between novices and experts in all plots regardless of the threshold value. Error bars represent standard error of the mean.

The correlations between NM, PL, and TT were assessed. Very weak association is found between NM versus TT measured with the Patriot (Figure 4-11a) or PL versus TT measured with the Patriot (Figure 4-11b). There is strong correlation between NM measured with the Wii and TT (0.90), as well as between PL measured with the Wii and TT (0.84) (Figure 4-11d-e). Correlation also exists between NM and PL for both modalities (Patriot 0.88, Wii 0.79) (Figure 4-11c and Figure 4-11f respectively). Comparison of the mean NM/s between the two groups shows no significant difference in either modality (Figure 4-12).
Figure 4-11: Correlation Coefficients for Assessment Parameters.
There is no correlation for (a) Patriot NM versus TT and (b) Patriot PL versus TT, likely due to the interference from cautery. Correlation is found for (c) Patriot NM versus PL, (d) Wii NM versus TT, (e) Wii PL versus TT, and (f) Wii NM versus PL.
There is no significance in results between the two groups for either modality. Error bars represent standard error of the mean.

### 4.5.3 Assessment Scores

OSATS assessment scores show a significant difference between novices and experts in all categories - Respect for Tissue (novice median = 2, expert median = 4) [Mann-Whitney U=54, $n_1=9$, $n_2=5$, $p=0.034$], Time & Motion (novice median = 2, expert median = 3) [Mann-Whitney U=54.5, $n_1=9$, $n_2=5$, $p=0.021$], Instrument Handling (novice median = 2, expert median = 3) [Mann-Whitney U=53.5, $n_1=9$, $n_2=5$, $p=0.042$], Knowledge of Instruments (novice median = 2, expert median = 4) [Mann-Whitney U=57, $n_1=9$, $n_2=5$, $p=0.010$], Use of Assistants (novice median = 2, expert median = 4) [Mann-Whitney U=57, $n_1=9$, $n_2=5$, $p=0.007$], Flow of Operation & Forward Planning (novice median = 2, expert median = 3) [Mann-Whitney U=57, $n_1=9$, $n_2=5$, $p=0.010$], and Knowledge of Specific Procedure (novice median = 2, expert median = 4) [Mann-Whitney U=59, $n_1=9$, $n_2=5$, $p=0.003$] (Figure 4-13). Mean total assessment scores also show significance between novices (median = 14) and experts (median = 24) [Mann-Whitney U=58, $n_1=9$, $n_2=5$, $p=0.004$] (Figure 4-14). Comparison of the assessment scores with the motion analysis parameters reveals some correlation between the scores and Patriot NM and PL (Figure 4-13).
4-15a-b). There is stronger correlation between the assessment scores and TT (Figure 4-15c), as well as the assessment scores and the Wii NM and PL (Figure 4-15d-e).

**Figure 4-13: Mean OSATS Assessment Scores by Category.**
* indicates a significance of p<0.05 for a Mann-Whitney U-test. Error bars represent standard error of the mean.
Figure 4-14: Mean Total OSATS Scores.
* indicates significance of p<0.05 for a Mann-Whitney U-test. Error bars represent standard error of the mean.
Figure 4-15: Correlation between Motion Analysis Parameters and OSATS Scores. There is some correlation between (a) Patriot NM and (b) Patriot PL compared with OSATS scores. There is a stronger correlation in (c) TT (d) Wii NM and (e) Wii PL and the OSATS scores.

4.5.4 Survey

Of the 11 residents who were emailed the survey, 9 responded (82% response rate). Of the two respondents who did not respond, one was on a research year and had been excluded for failing to dissect the Triangle of Calot and the other was on rotation, completed the proper procedure, and scored a 21/35 on the OSATS. Of the respondents, all received the majority of their technical skills training in open surgery in the OR during the operation. No respondents practiced open surgery on their own time. Generally respondents thought technical training occurred at a high frequency (Figure 4-16a) and was of moderate to good quality (Figure 4-16b).

Only four respondents had previously received formal didactic theoretical training for open surgery, though it was generally of good quality (Figure 4-17a). Didactic training occurred less
than once per week (Figure 4-17b), though respondents generally viewed this frequency as an appropriate (Figure 4-17c).

In general, respondents had observed more procedures than they had assisted with, and had assisted with more procedures than they had been primary surgeon on (Figure 4-18a). Volume of LC cases was larger for all categories, with the number of LC cases respondents had acted as primary surgeon on being comparable to the number of OCs observed. Respondents were split on whether the volume of open cases they performed as the primary surgeon were too low or too high (Figure 4-18b). They were also split on whether the amount of training they received prior to being the primary surgeon on an open procedure was appropriate (Figure 4-18c).

Two respondents had previously seen an open model but only one had previously used an open model (Figure 4-19a). Three quarters of the respondents were interested in training with an open model (Figure 4-19a) and 50% of respondents thought here was a role for open models in a training curriculum, with a further 38% neutral to the idea (Figure 4-19b).

The OC model was generally viewed as being partially accurate with the most common complaints being the liver is too deep (N=3), the liver positioning was not ideal (N=3), and the anatomy was incorrect (N=3). Despite these concerns, 6 respondents thought the model would be useful in a surgical training curriculum.

Improvements that were suggested were to bring the liver/gallbladder closer to the incision, to have more realistic anatomy, and to consider adopting the OC model for other open procedures especially those that cannot be done laparoscopically.
Figure 4-16: Technical Training Background.
Results are total number of respondents to questions concerning the (a) frequency of technical training and (b) quality of the technical training. Only novices were surveyed.
Figure 4-17: Didactic Training Background.
Results are total number of respondents to questions concerning the (a) quality of the didactic teaching sessions, (b) frequency of the didactic teaching sessions and (c) rating of the frequency of the didactic teaching sessions (3=the right frequency). Only novices were surveyed.
Novices were asked to report (a) total number of OC cases observed, assisted, and as primary, as well as LC cases observed, assisted and as primary. Mean results for all respondents are reported, error bars represent standard error of the mean. They were also asked (b) to rate the volume of operative cases performed as the primary surgeon (3 = just right) and (c) if they believed that they had sufficient training in open surgery prior to being the primary surgeon. One respondent had never been the primary surgeon on an open procedure and responded with Not Applicable for the latter two questions.
4.6 Discussion

4.6.1 Wii Tracking Validation

During validation, the Wii tracking system was shown to track a consistent path, albeit not aligned with the reference grid. This misalignment was simply due to the origin not being perfectly aligned to the grid prior to conducting the validation. A further error in the validation was the non-linear skew exhibited by Wii remote pair 1-2. Reasons for this skew are tied into other factors and discussed below. Regardless of the skew, mean NM and PL for all pairs were consistently larger than the nominal.

It is tempting to attribute the increased path length to the Wii origin not being aligned to the reference grid as in Figure 4-6b. Such an explanation would be incorrect however since the tracked motion could be expressed relative to any origin and still have the same net path length (i.e. the rotation matrix does not scale). Rather, this deviation from the nominal is explained primarily by jitter in the path during the manual translation of the IRM (including the Z-axis
jumps as the marker was translated). Some of the increased NM and PL may be due to inaccuracies with the detection and triangulation of IRMs. Any error in the triangulation of the initial origin points will propagate through all subsequent data, since the rotation matrix, and scaling factor will be affected. Errors in the triangulation of tracked IRMs will compound with the origin errors. Both error types could result from a poor calibration result.

In addition to affecting NM and PL, errors in the origin are the most likely explanation for the skew in Wii pair 1-2. A different calibration and origin location (though same Wii remote arrangement) was used for the OC trials. Test movements with the OC calibration and origin indicated that pair 1-2 provided comparable or superior accuracy to the other pairs, emphasizing the sensitivity of the origin location. Anecdotally, it appears that the origin accuracy appears to be the worst when two origin points are approximately collinear with the main axis of the Wii remote camera, though a poor calibration result, operating near the periphery of the Wii remotes’ field of view may also play a role. To achieve higher accuracy, remotes should be restricted to favourable locations and collection should be restricted to the central portion of each lens, with the trade-off being a corresponding reduction in perspectives and working volume available. Beyond these anecdotal fixes, further work should be done to eliminate or reduce the inflexibility of the origin and thus improve the resulting tracking.

Like the Wii system, the Patriot was able to accurately follow the grid though it did exhibit a slight rotation about the Z-axis attributable to its origin not being perfectly aligned to the reference grid. NM and PL for the Patriot were much larger than the nominal values. This can be attributed to jitter during the movements and emphasizing the need for a minimum velocity threshold though some error may have been introduced by the close proximity of the IRM mounted directly above the Patriot sensor.

Varying the minimum velocity threshold before determining NM and PL reveals that the ROVIMAS threshold values are not valid for sensor that is not mounted on a hand. By extension, it is uncertain if the ROVIMAS threshold values can be extended to all people that may be assessed. Variations in sensor placement on the hands, hand size, pulse rate, etc. will affect the rate of hand tremor, which the static threshold value will be unable to respond to. This uncertainty and the sensitivity of NM and PL to the threshold value call into question the validity of the ICSAD and ROVIMAS systems for accurately counting movements. Each system still
has shown validity to distinguish between novices and experts based on NM and PL however. This seeming paradox is discussed along with the OC validation results.

4.6.2 Model and Assessment

The Wii tracking system has been shown to distinguish between novices and experts working on an OC model by measuring NM but not PL. It is not clear why there was not a significant difference between novices’ and experts’ PL. One possible explanation is that since the manual assignment of IRMs focused on separating hands and not uniquely identifying IRMs. The extra PL that is introduced by this process for each subjects’ hands may explain the lack of significance in PL measured by the Wii.

To achieve significance in NM required laborious manual assignment of the hands by looking at raw pixel x/y graphs for each Wii remote. During this process, reflections and inexplicable jumps in IRM pixel identity numbering were observed and corrected for (Figure 4-5). These artefacts in the data emphasize the difficulty of the IRM identification problem. These artefacts, coupled with the limit of viewing 4 IRMs, inability to sequence IRMs, and the time consuming process of assigning the IRMs, add up to many limitations on the Wii system. As a practical assessment tool then, the Wii system is restrictive. One practical use of the Wii system would be to tracking a single object (hand). In this case, rejection of reflections would still be required, but this could be facilitated by an algorithm that would look for and reject large, instantaneous jumps. An alternative method of using the Wii system would be to tracking 4 IRMs that are mounted on a single, rigid body similar to commercial passive IR tracking systems. The IRMs could be identified by geometric reconstruction of the known geometry, which would provide position and orientation of the rigid body. Since four IRMs would be required to track one object, the size of the object would necessarily increase.

One possible solution for tracking multiple objects with the Wii remotes would be to use two different IR wavelengths. To achieve this, a different wavelength IRM would be placed on each hand. Two sets of Wii remotes, each set equipped with lens filters to make the remotes sensitive to one of the chosen wavelengths, would be required. A drawback of this approach is that double the number of Wii remotes would be needed to achieve an equivalent volume. As only 5 remotes can be connected to a computer at once, multiple collection computers would have to be
setup and synchronized. The cost and complexities detract significantly from a system that was designed to provide simple, low cost IR tracking.

Another possible solution is the development of a hybrid system combining EM, IR, and possibly accelerometer tracking. In such a hybrid system, the Wii remotes would operate at a low update rate with sequenced IRMs while another modality filled in the gaps between samples.

A third possible solution to overcome the Wii system limitations is to discard the Wii remotes as the cameras and build a new custom camera system. Such a custom camera system would operate in an analogous fashion to the Wii remote system. It would use lens filters to limit capture to IR wavelengths, and would enable synchronization of the shutters. Provided sufficiently fast cameras were chosen, it would be possible to sequence the IRMs. Furthermore, by using commercially available components, such a system could provide an increased flexibility in camera positioning at a much lower cost compared to commercially available systems.

During the OC trials, the Patriot was able to distinguish between novices and experts based on NM but not PL. This lack of significance in PL with the Patriot is suspicious due to the known affects of cautery on EM tracking. Further inspection of the results reveals that the mean PL of the experts (1701 m) is affected primarily by one subject (3155m) and that the mean expert PL drops to 1338 m if that subject is excluded. There is a corresponding drop of 28% in the standard error of the mean. Upon video analysis, it is revealed that the subject in question was the only one to hold the cautery pencil with the left-hand, causing the lead wire to sit immediately beside the Patriot sensor for large portions of the procedure. Furthermore, all subjects used monopolar cautery and their EM position data was affected to some degree. This is confirmed by looking at the inflated magnitude of the PL results (order of kilometres). Thus the validity of distinguishing between novices and experts based on the Patriot is called into question since the distinctions are based on underlying data that is known to be flawed. Breaking the analysis down by task would yield more insight since portions of the procedure where cautery is not used could be separated out. Such an analysis is beyond the scope of this work but should be undertaken in the future.

As with the validation trials, the effects of the minimum velocity threshold on the OC results were examined. The results of a varying threshold reveal that the precise choice of threshold
value is not important since the relative differences between novices and experts remains roughly constant. However, the threshold value does affect the absolute values recorded for NM and PL.

Application of a varying minimum velocity threshold revealed that the absolute values recorded for NM and PL depend on the threshold value used. An implication is this is that different threshold values explain why the resulting NM for the Wii is consistently higher than those for the Patriot. Were the threshold values adjusted, there is some point at which the results would be equivalent, though it is unlikely this would be a more accurate a representation of reality. This raises a larger concern as the values stated lose their actual meaning. For a given task, X NM and Y PL may allow for identification of novices and experts. However, X will not reflect the precise number of deliberate movements or Y the precise value of the deliberate path length. X and Y will merely be approximations, since tremors above the minimum velocity threshold will be counted, and deliberate movements below the threshold will be discarded. As an alternative, determining and subtracting out a constant tremor value on a per subject basis as is done in functional magnetic resonance imaging (fMRI) could be considered in the future.

Knowing that NM and PL are only approximations that can tell novices and experts apart regardless of the threshold level, it is worth asking if the NM and PL parameters are actually useful in the assessment of surgeons. Novices and experts are also known to be separable based on time taken for the procedure. Intuitively, as a surgeon takes longer to complete a procedure, his/her hands are likely to move more, increasing the NM made and the total distance travelled. This is confirmed by looking at the correlation between NM vs. TT and PL vs. TT, as well as the lack of significance in NM/s. Further exploration of whether NM and PL are indeed useful parameters or are merely alternative methods of measuring TT is needed. Possible alternatives to NM and PL would be gesture recognition algorithms, assessing repeatability and accuracy of movements for a given task, and incorporating information about the quality of the task (e.g. spacing of sutures).

Analysis of the OSASTS assessment results confirmed the validity of the OC model. OSATS scores were correlated with TT, Wii NM, and Wii PL, but not Patriot NM, and PL, adding to the evidence that assessment with EM tracking loses its validity when monopolar cautery is used. Curiously, the model showed validity despite the presence of a minimally trained assistant. This implies that practice sessions with open models could be facilitated by persons hired and trained
specifically to assist with the procedures of interest. For more frequent practice, it is conceivable that a simple robot could take the place of the assistant. Such a robot could be designed to respond to simple directions such as holding instruments in a specified position. It also could enable better practice of most procedures, not just open ones.

The OC model focused on the basic OC procedure, however, in practice, OCs are used for the most difficult, complicated cases. The question is raised then as to whether the OC model should be adapted for more complicated cases. The initial aim was to develop an OC model as an assessment tool and the presented model was validated as a tool for assessing OC skills in the general case. Expanding to specific complicated procedures would require carefully assessing what complications would be introduced and how these would be integrated into a surgical education curriculum. Were the model modified to introduce complications, the modifications would not be easy due to the nature of the porcine liver and modifications would have to be considered on a case by case basis for each potential complication. Alternatively, controlled presentation of unique cases is a hallmark advantage of VR simulators and development of a VR simulator for OCs would facilitate all of these cases to be covered. Such a VR simulator would require the use of expensive gloves and would have a decrease in the accuracy of the tactile feel compared to OC model presented here. Due to the difficulties that would be involved in introducing more complex procedures, it is worth emphasizing that the utility of each complication would have to be carefully assessed before detailed work on a model or simulator for that complication was done.

There was little in common between the two novices who did not respond to the survey, thus the survey respondents appear to be representative of the general surgery population. Analysis of the survey results shows that residents are generally happy with the quality and frequency of and didactic training in open procedures. Self reported values also confirm that residents are exposed to far more LCs than OCs. Residents were split on whether the volume of open cases they were primary surgeon was too small or too large. They were also split on whether they received adequate preparation prior to being primary surgeon in an open case. These split views likely reflects the locations the respondents had done their placements. Within the University of Toronto general surgery community, it is known that Toronto General Hospital (TGH) does far more open procedures than the others. Further exploration as to whether TGH is providing the
bulk of training in open procedures and the implications of this on resident training are worth exploring.

With respect to open surgical models, residents are interested in using them and feel that such models should be part of a surgical curriculum, though few had previously seen an open model. The merits of developing open models and introducing them into a surgical curriculum should be further explored and should take place in conjunction with assessing the actual resident caseload of open procedures within and beyond the University of Toronto environment.

Respondents to the survey highlighted two changes to the OC model, as well as the suggestion to adapt the idea of an open simulator to other procedures. One of the changes, moving the gallbladder closer to the incision, is easily implemented and is recommended for future work. The other suggested change was to make the anatomy more realistic. This could be accomplished using synthetic tissues for the liver and gallbladder, but at the expense of accurate tissue feel and the ability to use cautery. A consequence of the different anatomy in the porcine model is that the experts may find the task easier than novices since the experts are better equipped to handle the different anatomy. Consequently, it could be argued that the differences between novices and experts was primarily due to the different anatomy however the literature has demonstrated animal methods to be valid for “teaching, developing, and refining surgical techniques in both open and laparoscopic approaches”\textsuperscript{82}.

4.7 Summary of OC Model with Wii Tracking

An IR 3D tracking system based on Wii remotes has been developed and shown to provide reasonable accuracy tracking across a grid. The tracking system still requires improvements, in particular the removal of the tracking accuracy being dependent on the origin location.

A novel OC model and assessment tool based on Wii tracking system was developed and validated. The OC model showed face and construct validity but should be improved by placing the liver in a more anatomically correct position. Motion analysis of the OC trials reveals that the Wii system can distinguish between novices and experts based on NM and PL. Further work is required to remove the manual IRM assignment steps from the processing of the Wii data. The Patriot motion tracking was able to distinguish between novices and experts based on NM
but not PL, emphasizing the inappropriateness of EM tracking in the presence of monopolar cautery. TT could also distinguish between novices and experts.

Survey results revealed that residents are receiving open surgical training and are interested in training with open simulators. It is not clear if the level of open surgical training residents are receiving is sufficient and more work is required to investigate this.

NM and PL have been shown to be dependent on the value of minimum velocity threshold, making comparisons of NM and PL values between different modalities difficult. Furthermore, NM and PL have been shown to be correlated with TT, raising questions about their usefulness as assessment parameters. Future work should focus on the exploring the link between NM, PL and TT, as well as exploring other methods of motion analysis such as gesture recognition.
Chapter 5
Conclusions

In Chapter 3, EM tracking was shown to have large position errors in the presence of monopolar cautery. Since the antiparallel wires in bipolar cautery produce magnetic fields that cancel, there are only small tracking errors when bipolar cautery is used. Absolute maximum and mean plus and minus position errors have been established for all modes of operation of monopolar and bipolar cautery. These error limits are valid for the specific test conditions and are expected to change if the test conditions are changed. Regardless of the exact error values, monopolar cautery will affect all EM tracking systems and EM tracking should not be relied upon in situations where monopolar cautery is used. This implies that alternative motion analysis tools are required for assessing surgical tasks involving monopolar cautery. Furthermore, manufacturers’ guidelines should be established related to the intraoperative use of EM tracking and cautery machines. Due to the high sensitivity of guided surgery applications, these guidelines should cover monopolar and bipolar cautery.

In Chapter 4, the development and validation of a novel model and assessment tool for open surgery was discussed. The model was constructed for OCs using a porcine liver and standard plastic torso. The assessment tool focused on an IR tracking system based on calibrated Wii remotes. The Wii system tracking was validated qualitatively by translation through a grid. As an accurate tracking tool, the Wii system requires additional work to eliminate sensitivities to the origin location. To be used for tracking multiple objects, a better method of uniquely identifying IRs is also required.

To validate the OC model and assessment tool, eleven novices and five experts performed one OC each. OSATS scoring of the subjects validated the OC model. Motion analysis using NM, PL, and TT validated the Wii system as an assessment tool. EM tracking could only distinguish between novices and experts on NM, not PL. Furthermore EM tracking PL values were on the order of kilometres. These PL distortions confirm the interfering effects of monopolar cautery on EM tracking.

Responses to the post OC trial survey indicate that face validity of the OC model was partially established through post procedure surveys. Suggestions to improve the face validity included
adjusting the position/angle of the liver and gallbladder relative to the incision, and improving the accuracy of the anatomy. The first suggestion is straightforward to implement and should be done to improve the face validity of the OC model. The second suggestion echoes a familiar refrain of simulators: accurate anatomy can be established with synthetic materials, but with a loss of tissue realism. An OC model could be designed using synthetic materials to have realistic anatomy, but at the expense of tissue feel and the use of cautery.

Additional survey responses revealed that respondents receive most of their technical training in the OR during the procedure and are happy with the quality of the technical and didactic training, as well as the frequency of didactic training. Respondents had typically not previously seen or used an open simulator but expressed interested in training with one and generally saw a role for the OC model in a surgical curriculum. Respondents had far more exposure to LCs, with the number of observed OCs comparable to the number of LCs they had been the primary surgeon on. It is not clear if the number of OCs residents are exposed to is sufficient to ensure proficiency. The broader applicability of these findings is uncertain as there were only nine respondents.

During validation of the Wii tracking and the OC model and assessment tool, exploration of the effects of the minimum velocity threshold was done. The inflated NM and PL values from the Wii tracking validation indicate that some form of threshold is required to remove tremors but it is not clear what the value should be. Inspection of the OC results indicates a consistent gap between novices and experts for NM and PL as the minimum velocity threshold is varied. There is no clear indication as to what the true values of NM and PL are. The original purpose of the threshold was to focus on deliberate movements while eliminating tremor, yet slow deliberate movements are likely to be eliminated as well. Thus it is unknown what the actual number of deliberate movements and deliberate distance travelled is. Results are simply an approximation for a given threshold value and may not translate between different tracking modalities. An alternative approach could be to remove a constant tremor value from the data, as is done in fMRI research.

More detailed analysis of the motion analysis parameters revealed that NM and PL collected with the Wii and TT are all correlated, though the Patriot NM and PL are not. This reinforces that EM tracking is influenced by monopolar cautery. Furthermore, TT, Wii NM and Wii PL are
all correlated with the OSATS scores as well, indicating consistent assessment of the OC model. Novices and experts are indistinguishable on NM/s. The strong correlations of NM and PL with TT and lack of significance in NM/s raises the question of if NM and PL are unique measures or if the significance in NM and PL is a reflection of the correlation with TT. Further investigation of this is recommended.

Given the sensitivity of NM and PL to the minimum velocity threshold, as well as the correlation with TT, it is worth asking if NM and PL are the best parameters to consider for assessment of skills. NM and PL offer the ability to tell novices and experts apart within a controlled setting (e.g. video analysis to ensure the same task is completed, extraneous movements were removed, etc.). However, the raw NM and PL values are not useful feedback to a trainee as they do not help him/her to improve operative performance. Any self directed improvement based solely on NM and PL will result in training to those values, not the desired improved surgical result. Better motion analysis parameters such as gesture recognition or the repeatability and accuracy of specific movements would allow for detailed assessment and potentially automatic feedback, but would be restricted to specific skills and would be unwieldy for an entire procedure. Such parameters could be collected from a hybrid data collection system incorporating video, accelerometers, EM, and IR tracking. Different aspects of such a hybrid system could be used to assess different parameters (e.g. accelerometers for gesture recognition while knowing where the hand is based on IR or EM tracking).

5.1 Future Work

There are several avenues for future work. Interactions between EM tracking and cautery should be explored further. This is important in the field of surgical assessment where EM is often used but is far more important for the intraoperative use of EM tracking. If not already existing, guidelines should be developed to address to avoid reliance on EM tracking while using cautery intraoperatively.

The OC model should be refined to fully establish face validity. The role for open simulators in a surgical curriculum should be explored and additional open models developed as required. Development of a robot assistant could be explored to allow for independent practice on the open models. Such a robot assistant could be integrated with motion tracking technologies to provide automatic feedback. In addition, exploration of alternative assessment parameters and sensing
modalities should be considered along with the usefulness of subtracting a constant tremor signal for each subject in *lieu* of a minimum velocity threshold. The validity of the OC model as a training tool and the utility of adapting the OC model to include surgical complications should also be explored.
References


81. Scholz M. Email Communication: Electrosurgery and the Aurora System. 2010.

# Appendix A: Assessment Sheet

## Open cholecystectomy model validation assessment tool

### Assessor Initials:

(a) *Objective Structured Assessment of Technical Skill (Martin et al, Br J Surg 1998)*

<table>
<thead>
<tr>
<th>General Skill</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect For Tissue</td>
<td>Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments</td>
<td>Careful handling of tissue but occasionally caused inadvertent damage</td>
<td>Consistently handled tissues appropriately with minimal damage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time &amp; Motion</td>
<td>Many unnecessary moves</td>
<td>Efficient time/motion but some unnecessary moves</td>
<td>Economy of movement and maximum efficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrument Handling</td>
<td>Repeatedly makes tentative or awkward moves with instruments</td>
<td>Competent use of instruments although occasionally appeared stiff or awkward</td>
<td>Fluid moves with instruments and no awkwardness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of Instruments</td>
<td>Frequently asked for the wrong instrument or used an inappropriate instrument</td>
<td>Knew the names of most instruments and used appropriate instrument or the task</td>
<td>Obviously familiar with the instruments required and their names</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of Assistants</td>
<td>Consistently placed assistants poorly or failed to use assistants</td>
<td>Good use of assistants most of the time</td>
<td>Strategically used assistant to the best advantage at all times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow of Operation &amp; Forward Planning</td>
<td>Frequently stopped operating or needed to discuss next move</td>
<td>Demonstrated ability for forward planning with steady progression of operative procedure</td>
<td>Obviously planned course of operation with effortless flow from one move to the next</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of Specific Procedure</td>
<td>Deficient knowledge. Needed specific instruction at most operative steps</td>
<td>Knew all important aspects of the operation</td>
<td>Demonstrated familiarity with all aspects of the operation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Post Questionnaire

PRE STUDY EXPERIENCE
The following apply to your previous experience up to the day you participated in the study.

Please provide the following:
PGY level:
Age:
Handedness:
When was the last time you were on a surgical rotation?

Educational Experience
1. Where is the majority of your technical skills training in "open surgery" occurring? Please select one.
   a. in the OR before the operation begins
   b. in the OR during the operation
   c. outside of the OR
   d. a surgical skills lab
   Answer:

2. How frequently do these "open surgery" technical skills teaching sessions occur? Please select one.
   a. 0 x / week
   b. <1 x /week
   c. 1 x / week
   d. 2 x /week
   e. 3 x / week
   f. > 3 x / week
   Answer:

3. If you are receiving technical skills training in "open surgery", how would you rate the quality of the teaching you are receiving?
   a. 1 => lowest quality
   b. 2
   c. 3
   d. 4
   e. 5 => highest quality
   f. N/A
   Answer:

4. Do you practice "open" surgical technique on your own time?
   a. No I do not
   b. Yes on my own (e.g. at the surgical skills lab)
   c. No, most of my practice is in laparoscopic surgery (outside the OR)
   d. Other, please specify
   Answer:

5. During your surgical residency, have you received didactic theoretical training for "open surgery"? (Instruction for proper use of instruments, technique, safety, etc.)?
   a. Yes
   b. No
   Answer:

6. How would you rate the quality of the didactic teaching sessions you are receiving? Please select one.
   a. 1 => lowest quality
   b. 2
   c. 3
   d. 4
e. 5 => highest quality
f. N/A

Answer:

7. How frequently do these teaching sessions occur?
   a. <1 x /week
   b. 1 x / week
   c. 2 x /week
   d. 3 x / week
   e. > 3 x / week

Answer:

9. How would you rate the frequency of these didactic teaching sessions you are receiving? Please select one.
   a. 1 => too little
   b. 2
   c. 3
   d. 4
   e. 5 => too much
   f. N/A

Answer:

10. Have you seen an Open Surgery model? (prior to the one in my study)
    a. Yes
    b. No

Answer:

11. Have you used an open surgery model? (prior to the one in my study)
    a. Yes
    b. No

Answer:

12. Do you think that Open Surgery models have a role in training for "Open surgery"?
    a. 1 => not useful
    b. 2
    c. 3
    d. 4
    e. 5 => very useful
    f. N/A

Answer:

13. Are you interested in training with "Open Surgery" models?
    a. Yes
    b. No

Answer:

Operative Experience

1. How many OPEN cholecystectomy cases have you observed since the beginning of your residency?
   Answer:

2. How many OPEN cholecystectomy cases have you assisted in since the beginning of your residency (you DID NOT dissect the Triangle of Calot)?
   Answer:

3. How many OPEN cholecystectomy cases have you been the primary surgeon for since the beginning of your residency (you dissected the Triangle of Calot)?
   Answer:
4. How many laparoscopic cholecystectomy cases have you observed since the beginning of your residency?
   Answer:

5. How many laparoscopic cholecystectomy cases have you assisted in since the beginning of your residency (you DID NOT dissect the Triangle of Calot)?
   Answer:

6. How many laparoscopic cholecystectomy cases have you been the primary surgeon for since the beginning of your residency (you dissected the Triangle of Calot)?
   Answer:

7. How would you rate the volume of OPEN operative cases you perform as the primary surgeon?
   a. 1 => too little
   b. 2
   c. 3
   d. 4
   e. 5 => too much
   f. N/A
   Answer:

8. Do you believe you have enough training in OPEN surgery prior to performing an OPEN case as the primary surgeon?
   a. 1 => I don't have enough
   b. 2
   c. 3
   d. 4
   e. 5 => I have more than enough
   f. N/A
   Answer:

**POST STUDY**
1. Did the model accurately reproduce an open cholecystectomy procedure? Why or why not?

2. From an operative standpoint, do you think the use of this model in a surgical training curriculum would be beneficial?

3. Please list any suggested changes to the model: