Western Graduate&PostdoctoralStudies

Western University

Scholarship@Western

Electronic Thesis and Dissertation Repository

8-28-2012 12:00 AM

Metrics for Evaluating Surgical Microscope Usage During Myringotomy

Arefin M. Shamsil
The University of Western Ontario

Supervisor
Dr. Hanif M. Ladak
The University of Western Ontario

Graduate Program in Electrical and Computer Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science

METRICS FOR EVALUATING SURGICAL MICROSCOPE USAGE DURING MYRINGOTOMY

(Spine title: Metrics for surgical microscope usage)

(Thesis format: Integrated Article)

by

Arefin Mohammed Shamsil

Graduate Program in Electrical and Computer Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Masters of Engineering Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisors	Examiners
Dr. Hanif M. Ladak	Dr. Vijay Parsa
Dr. Sumit K. Agrawal	Dr. Jagath Samarabandu
Supervisory Committee	Dr. Mahmoud El-Sakka
Dr. Philip C. Doyle	

The thesis by

Arefin Mohammed Shamsil

entitled:

Metrics for Evaluating Surgical Microscope Usage During Myringotomy

is accepted in partial fulfillment of the
requirements for the degree of
Masters of Engineering Science

Date

Chair of the Thesis Examination Board

Abstract

Although teaching and learning surgical microscope manoeuvring is a fundamental step in middle ear surgical training, currently there is no objective method to teach or assess this skill. This thesis presents an experimental study designed to implement and test sets of metrics capable of numerically evaluating microscope manoeuvrability and qualitatively assessing surgical expertise of a subject during a middle ear surgery called myringotomy. The core experiment involved performing a myringotomy surgical procedure on a fixed cadaveric head with intact ear anatomy. As participants, experienced ear-nose-throat (ENT) surgeons and ENT surgical residents were invited. While performing the experimental surgical procedure, their microscope manoeuvring motions were captured as translational and angular coordinates using an optical tracker. These data were analyzed in terms of motion path length, velocity, acceleration, jitter, manoeuvring volume, smoothness, rotation and time. Participants' hand motion, body posture and microscopic view were also video recorded in order to qualitatively assess their surgical expertise via a review panel. The following categories of metrics were identified as discriminatory: time, rotation, volume, smoothness, jitter, proper microscope positioning, proper speculum and tube insertions, no hand jitter, optimum body and arm postures during operation and unobstructed and centered optical field-of-view through the microscope. The future goal is to incorporate these metrics into a myringotomy surgical simulator to train ENT residents.

Keywords: Surgical microscope, performance metrics, myringotomy, training

Co-Authorship

Chapter 2 of this thesis is a modified version of a manuscript that is being prepared for submission to the journal *The Laryngoscope*. This manuscript was co-authored by Arefin Shamsil, Brandon Wickens, Philip Doyle, Hanif Ladak and Sumit Agrawal. Arefin Shamsil implemented all the technical aspects of the study, conducted the experiments, collected all data and analyzed them. Dr. Wickens scheduled participants for the experiments. Dr. Doyle aided with data representation, data analyses and interpretation. Dr. Ladak was the supervisor who set out the goals and milestones. Dr. Agrawal was a joint supervisor and directed the study and provided guidance in the study design.

Acknowledgments

I would like to acknowledge my supervisor Dr. Hanif M. Ladak and my co-supervisor Dr. Sumit K. Agrawal for their tremendous guidance and direction in the course of my research work from conception to completion. I would also like to thank Dr. Philip C. Doyle for providing substantial guidance in data representation and data analyses. I would also like to thank Dr. Brandon Wickens for his contributions in aiding me execute all the experiments. He has put forth substantial effort in contacting and recruiting all the residents in the Western University Otolaryngology residency program. I thank Ehsan Salamati, Reza Mousavi and Alireza Rohani who helped me in validating computational techniques during the technical implementation phases of my research. Finally I thank my parents for their constant support and encouragement throughout my graduate study.

This work was funded by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Research Foundation (ORF) with contributions from Medtronic Canada.

Table of Contents

List of Tables

List of Figures

Chapter 1 Figures

Chapter 2 Figures

List of Appendices

Chapter 1

1 Introduction

1.1 Motivation

In the treatment or diagnosis of ear, nose and throat (ENT) related pathologies, a surgical microscope is often used as the primary device. For example, anatomical structures within the ear are tiny and complexly interconnected (see Figure 1.1a). Therefore, a surgical microscope (Figure 1.1b) is used to see into a targeted ENT orifice (e.g., the ear canal) to optimally locate the anatomical structure of interest. After obtaining an optimal view of the targeted structure, an ENT surgery is performed or a diagnostic procedure is carried out. Since the surgical site (or diagnostic site) is continually visualized through a microscope during an operation, an ENT surgeon must have the surgical skills to operate using a microscope. Human visual perception changes under a microscope due to its optical zoom and focus. Therefore, precise hand-eye coordination through the microscope optics must be mastered to carry out an ENT surgery. In order to develop such skills, one must initially be able to produce and maintain an optimum microscopic view of the anatomy of interest during an operation.

Figure 1.1: (a) An illustration of ear, nose and throat anatomy [1] (b) A Leica M720 OH5 surgical microscope [2].

Efficiently manoeuvring a surgical microscope is a fine skill itself that is inevitably essential to master. Otherwise, a good length of surgical time is wasted to obtain the optimum view of the surgical site. In ENT surgical training, teaching and learning to manoeuvre a surgical microscope is one of the first steps. Yet microscope manoeuvring is often overlooked as a trivial task and there are no quantitative methods to assess microscope manoeuvring skill development. Following the conventional 'see one, do one, teach one' method, trainees are often required to manoeuvre a microscope without any prior practice. By receiving only qualitative feedback on their overall performance on the procedure, trainees do not identify the particular manoeuvring problems they need to attend to. As a result, their skill development time lengthens while teaching microscope manoeuvring becomes a frustrating process for the instructing surgeon. Instructing surgeons agree that the trainees always struggle to manoeuvre a microscope following the most "economic" path. An "economic" path refers to the shortest and smoothest route covered in a short period of time with the least amount of jitter. Trainees repeatedly adjust the final position of the microscope and struggle to obtain the optimum microscopic view. Also, they sometimes have unfocussed vision through the microscope optics and inadequate distance between the eyepiece and the anatomy of interest. These factors collectively limit their microscopic vision and affect their operating efficiency.

ENT is a vast branch of medicine and surgery that broadly makes use of the surgical microscope. To appropriately limit the scope of this research, the focus is specifically on myringotomy – a simple yet delicate surgery in which an incision is made in the eardrum. The procedure is often used to treat middle-ear infections. The motif behind this research project is to track and measure the motion of the surgical microscope during myringotomy in order to compare the microscope manoeuvrability of trainees and experts within a controlled experimental structure. By comparing their manoeuvrability based on a multi-variable metric system (i.e., path length, time, efficiency etc.), it may be possible to determine how well or poorly a trainee performs in comparison to an expert. Furthermore, based on this metric system, a numerical assessment report can be produced to rank the manoeuvring performance of an operator. The end goal (beyond the scope of this thesis) is to incorporate the metric system into a training surgical simulator for myringotomy. This will enable trainees to get automated feedback on their performance and track their improvement over time.

1.2 Background

As noted previously, the focus of this research is on surgery of the ear, particularly myringotomy. As shown in Figure 1.2, the human ear is conceptually subdivided into three sections: the outer ear, the middle ear and the inner ear. Of specific interest is the middle ear as it is the site of myringotomy. The middle ear is composed of very delicate and tiny anatomical structures. Due to their sizes, physiological and anatomical properties, sensitive locations and complex interconnections, middle ear surgeries require precise dexterity and fine hand-eye coordination through a surgical microscope. All ear surgeries are utterly reliant on the surgical microscope without which many ear procedures and ENT procedures in general cannot be mastered. Therefore, mastering microscope manoeuvring and optical focusing are primarily the essential tasks before learning more complex ENT surgical procedures.

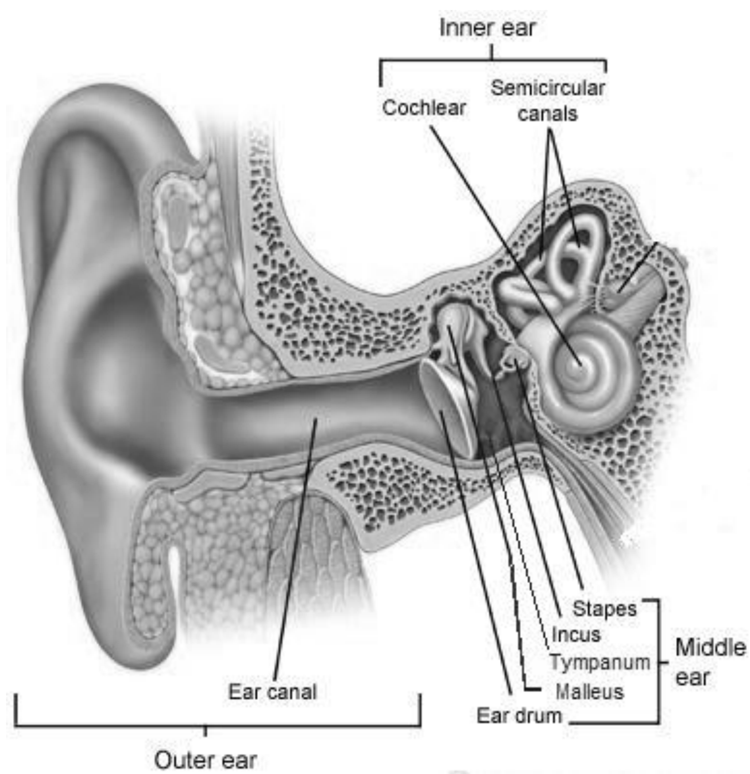

Figure 1.2: Sagittal view of the human ear anatomy showing all subsections and critical structures [3].

Myringotomy is a basic ENT surgical procedure which requires continual use of the surgical microscope. Surgical dexterity and microscope manoeuvrability required in this procedure are the prerequisite skills to be mastered before moving further onto other complex ENT surgeries. In an ENT surgical residency program, myringotomy and other ENT surgeries are taught through the conventional "see one, do one, teach one" apprenticeship approach. Following this approach, during a myringotomy training session, a patient's auditory anatomy is put to risk in the hands of a novice ENT resident. Furthermore, the instructing surgeon becomes liable if any internal anatomical injury occurs during the session. In a myringotomy procedure, as shown in Figure 1.3, the tympanic membrane (also called the eardrum) is accessed through the ear canal using a small surgical speculum and is visualized with a surgical microscope. Guiding a surgical blade down this speculum and through the ear canal, a small incision is then made on the tympanic membrane. A surgical suction tube with controlled reverse pressure is then used to draw out all the accumulated fluid, pus and wax from the middle ear cavity (also called the tympanum) through this incision. Finally a ventilation tube is placed in the tympanic membrane to facilitate continuous drainage of the fluid.

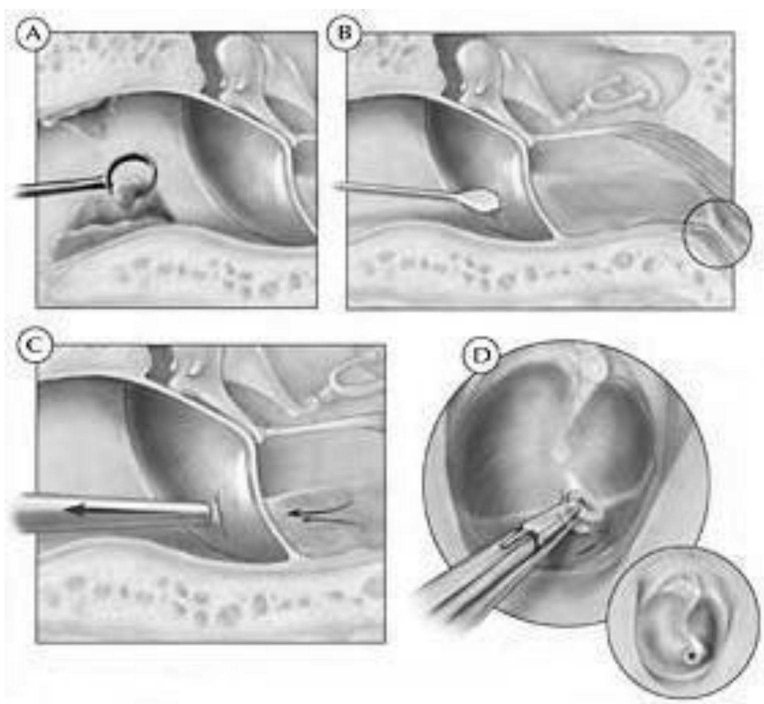

Figure 1.3: Steps of a myringotomy procedure. A: wax removal; B: incision; C: fluid suction; D: tube placement [4].

Although the steps in myringotomy are simple, many common errors can happen as outlined by Montague *et al.* [5]. In fact, many of these errors occur because of the lack of optimum microscopic view and good hand-eye coordination. Some of these errors, as mentioned in [5], include multiple attempts to place the ventilation tube, multiple attempts to complete the myringotomy, inappropriate microscope magnification and improper incision sizes. Seeing an object under a surgical microscope is quite different from seeing it with the naked eye. Due to the object magnification, human visual perception is challenged. Minor movements of the surgical tools on an object appear to be quite large through the microscope, which is why obvious harmful errors are easy to make. A resident must learn to cope with this situation in order to efficiently carry out balanced and calculating moves with his/her surgical tools. One challenge that almost all novice residents face during a microscopic ENT surgery is to obtain a clear microscopic view of the object of interest and maintain that view throughout the surgery. This observation was, in fact, made during a live myringotomy training session at the London Health Sciences Centre – Victoria Hospital. Instructing ENT surgeons have also reported that during surgical microscope navigation, novice residents may have a blind spot through either one or both of the ocular pieces of the microscope, unfocussed vision through either one or both ocular pieces, inadequate distance between the object of interest and the ocular lens and inconsistent stationary position and orientation of the microscope's end-frame throughout the entire surgery. These shortcomings certainly limit the microscopic vision of the resident and consequently disrupt the progression of his/her operating efficiency.

There is no doubt that repeated practice is required to overcome these technical issues. However, the problem is that there are usually not enough actual myringotomy cases for novice residents to practice on. In addition, novice residents practice on live patients under the close observation of an instructing surgeon since patient safety is the prime concern in an apprenticeship approach. However, mistakes can happen even under the watchful eyes of an instructor. These drawbacks lengthen the progression of the resident while putting patients at risk. As an alternative, cadaveric specimens serve as a training resource for residents to practice myringotomy. However, there is often a shortage of cadaveric specimens and once they are practiced on, they cannot be reused.

Apart from cadavers, physical models are also used to simulate myringotomy cases (Section 1.3). The problem with this method is that it is a poor representation of an actual myringotomy and actual middle ear anatomy with proper biomechanical properties. Therefore, residents get the wrong sense of force feedback and interaction between a surgical blade and the model of middle ear structures. Taking all these issues under consideration, it has been deduced that a virtual-reality (VR) based simulator for myringotomy would be a robust training resource compared to all other resources. Over the years, the Auditory Biophysics Laboratory at Western University has developed a prototype of such a VR simulator for myringotomy training. To date, a significant contribution of the research team has been put toward the development of the myringotomy simulation software, construction of a virtual model of the tympanic membrane, implementation of incision algorithms and incorporation of haptics. Brief summaries of these works can be found in Section 1.3. Although the overall system is quite elegant, it still needs further improvement before it can be used as a training resource. The anticipated plan is to model the 3D middle ear anatomy comprised of the ear canal and ear drum with realistic mechanical properties, wax removal and pathological cases. Furthermore, to simulate realism of the surgery, inclusion of a surgical microscope is anticipated along with a manoeuvring performance evaluator integrated into the myringotomy simulator.

1.3 Literature Review

As noted previously, myringotomy is taught through an apprenticeship approach at large. In the literature, however, a few publications are available that present physical model based simulators as training tools for the myringotomy procedure. Among them the Wigan grommet trainer [6], the Bradford grommet trainer [7] and the Artificial Ear trainer [8] are a few examples of physical myringotomy simulators for novice ENT residents to practice the procedure. These models usually contain a synthetic membrane attached at the end of a hollow tube to simulate the eardrum and the ear canal, respectively. Although the use of these models eliminates the risk based practice of myringotomy training on live patients, they provide an unrealistic experience of the procedure as they do not emulate the realistic ear anatomy or the realistic mechanical properties of the eardrum. Lastly, there are no ways to quantify skill development progress of residents with these models. In fact, any assessment of skills based on these models may yet be inaccurate or perhaps be heavily biased. Therefore, to overcome the shortcomings of the apprenticeship approach, which entails risk to the patient, and to overcome the shortcomings of physical models, which entails incorrect ear anatomy, virtual reality (VR) based simulators have been considered. A VR simulator typically adapts the concept of video games in which a computer generated interactive virtual environment is created containing a series of interactive virtual objects. VR simulators are widely used in flight training and are now used in many surgical training procedures (i.e., laparoscopy, endoscopy etc.). The advantage of VR surgical simulator is that precise anatomical models can be generated and be made interactive and responsive in order for a trainee to experience the realism of a surgery without interacting with a live patient. Following this general motif, a VR myringotomy simulator has been developed in the Auditory Biophysics Laboratory at Western University. This myringotomy simulator is the first of its kind and no such simulator exists commercially or in the research literature. In addition, the surgical microscope simulator component needed for this myringotomy simulator is not available yet either. Therefore, no scientific studies have been done to objectively assess microscope manoeuvring skills. As a result, to understand the concept behind objective skill assessment, publications on other studies focussed on other surgical procedures have been reviewed.

1.3.1 Virtual reality based myringotomy simulator

As mentioned in Section 1.2, a VR-based training simulator for myringotomy is being developed in the Auditory Biophysics Laboratory at Western University. The very first prototype of this simulator was developed by Wheeler *et al.* [9] that simulated virtual surgical blade navigation and collision with a virtual ear canal for trauma detection. In this prototype, the surgical microscope is represented by a 3D stereo Visor (eMagin Z800 3D visor, Bellevue, WA) mounted on an adjustable aluminium stand as labelled in Figure 1.4a. The surgical tools to be simulated, labelled as (2) speculum and (3) myringotomy blade in Figure 1.5, were marked with optical markers within the field of view of a stereoscopic optical tracker (Claron MicronTracker 2 S60, Claron

Technology, Toronto, ON) shown in Figure 1.4a. The surgical view of the procedure was virtually simulated in a software graphics and physics engine called OGRE 3D. The simulated virtual scene, shown in Figure 1.4b, included the virtually rendered speculum, ear canal, eardrum and a blade which interactively responded to the optically tracked surgical blade in real space. Although optical tracking was carried out with sub-millimetre accuracy (i.e., 0.25 mm), the low rendering frequency (i.e., 30 frames per second) of the tracking camera caused too much jitter in the corresponding virtual blade. In terms of functional limitations, this prototype simulated the eardrum incision as a drawn line, provided no force feedback during the operation, did not simulate any wax removal or bleeding and did not simulate pathological cases. In addition, a face validity study administered to instructing surgeons in ENT and to ENT surgical residents indicated that tactile feedback on virtual tools during the operation is essential in creating a realistic simulator.

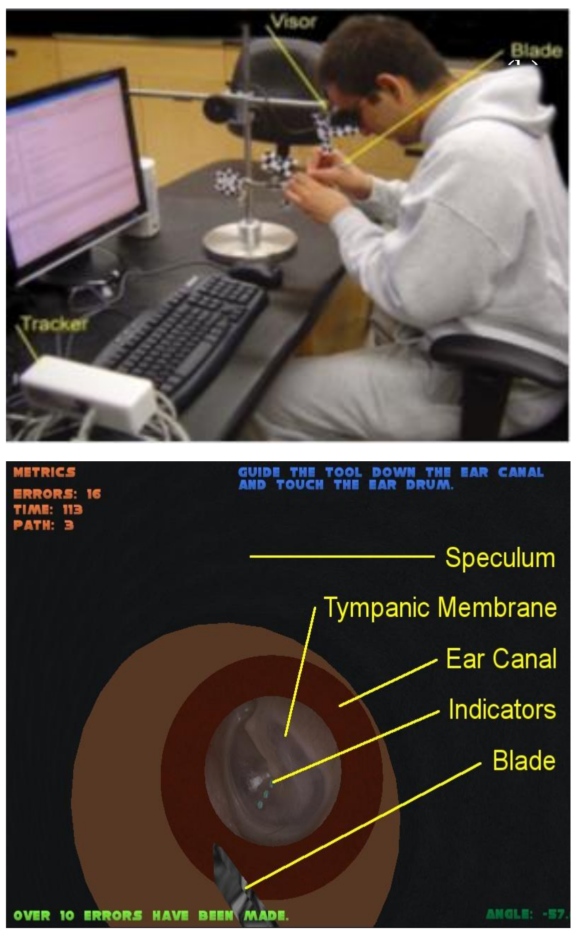

Figure 1.4: (a) Simulator developed by Wheeler *et a.*. (b) Snapshot of screen visible through the 3D visor.

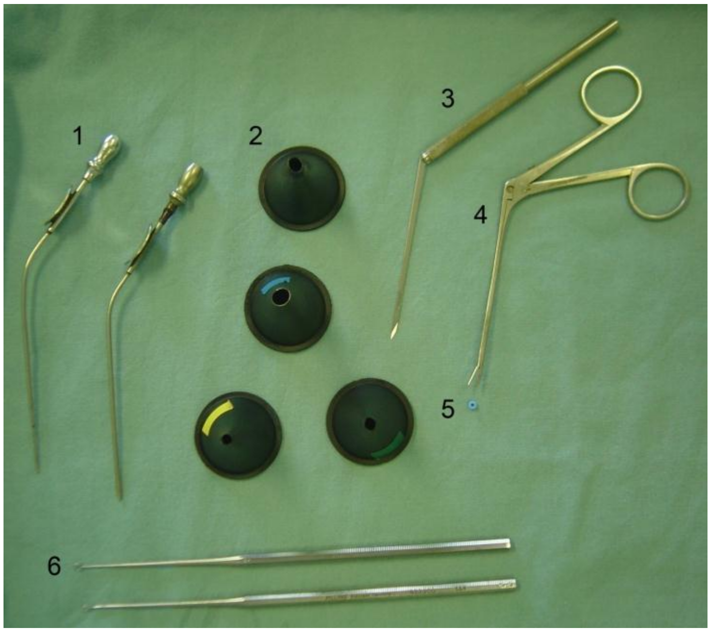

Figure 1.5: Tools used in myringotomy procedure. (1) Suction pipe. (2) Speculum. (3) Blade. (4) Crocodile forceps. (5) Ventilation tube (6) Curettes.

In order to eliminate the jitter issue added by the optical tracker and to add tactile feedback when cutting into simulated tissues using the virtual blade, Rehal included a haptic arm (PHANTOM Omni, Sensable Technologies, Woburn, MA) [10]. To simulate appropriate tactile feedback response with this haptic device, the virtual middle ear anatomy was modelled with corresponding compliant or rigid properties. The new simulator is shown in Figure 1.6. In this case, the ear canal was modelled as a rigid tube which provided a very stiff force feedback when contacted by the virtual blade simulated by the stylus of the haptic device as labelled in Figure 1.6. As confirmed by the practicing ENT surgeons, force feedback from the eardrum upon contact with the surgical blade is smaller than the smallest force rendered by the Omni haptic device. Therefore, the eardrum was modelled as completely compliant. Although inclusion of the haptic device added some realistic sense of interaction with the virtual anatomy, the virtual eardrum did not have realistic topology, nor did it behave realistically as pointed out by the face validity test administered by Sowerby *et al.* [11]. In addition, the eardrum incision was still being simulated with a line drawn on the virtual eardrum surface. In light of these issues, it was concluded that a realistic simulator would include a deformable model of the eardrum that actually cuts open when an incision is made upon it.

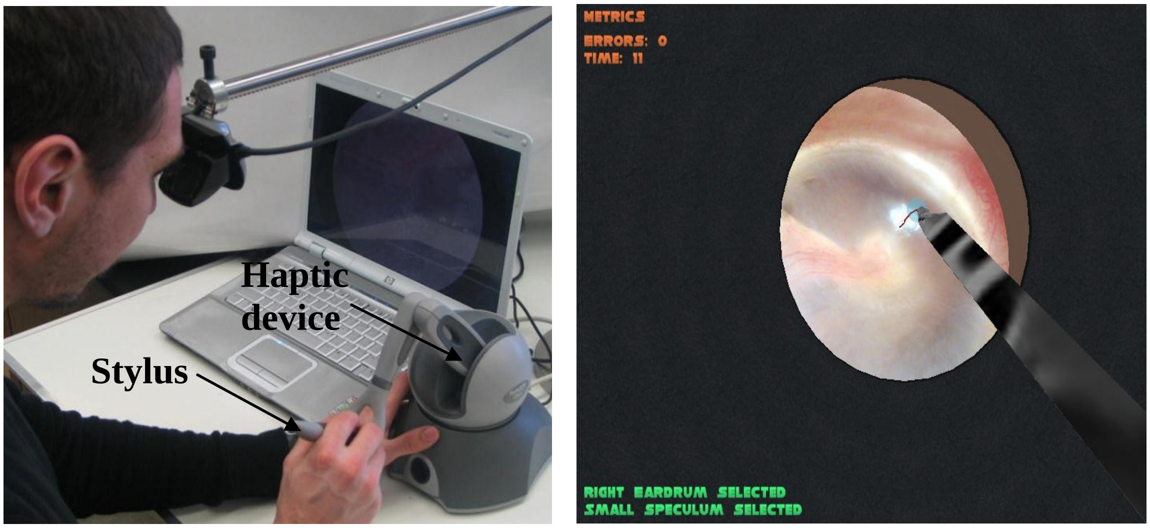

Figure 1.6: Picture on the left shows the setup of the simulator by Rehal which included a haptic device, while the picture on the right shows what is seen through the 3D visor.

The issues associated with the eardrum incision that continued since Wheeler *et al.* were then addressed by Ho *et al.* [12] who developed a topologically correct deformable model of the eardrum and incorporated it into the existing simulator. It was developed using a deformable mass-spring model that could be cut by the virtual surgical tool that existed in the previous simulator(s). A total of three cutting algorithms were implemented and a face validity test on each algorithm was administered to instructing surgeons and surgical residents. Figure 1.7 shows one of the end results of Ho *et al*.

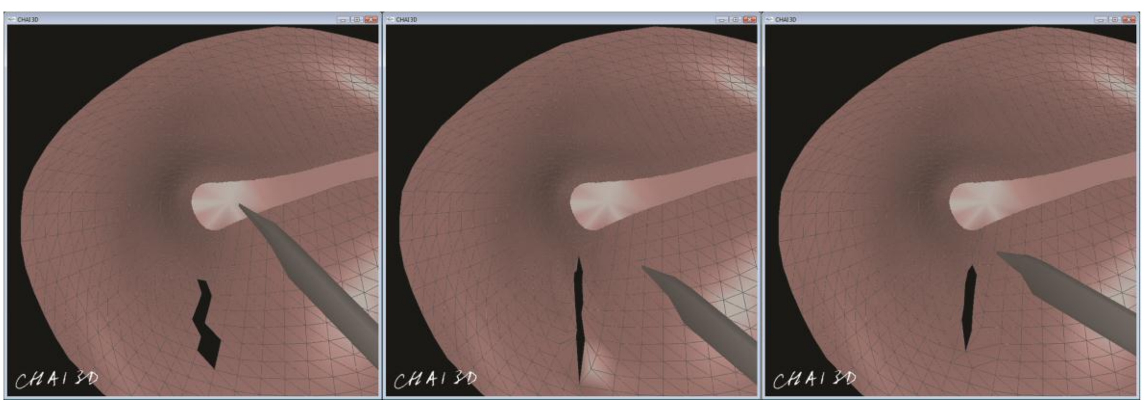

Figure 1.7: A simulation developed by Ho *et al.* which shows the mass-spring model of the virtual eardrum that is cut by the virtual blade using 3 different cutting algorithms that were developed.

All contributions made toward developing the myringotomy simulator thus far always omitted simulation of the surgical microscope. The current representation of the microscope does not contain any articulation or freedom of movement as the real surgical microscope. One's skill development in myringotomy is still incomplete if he/she does not master economic and safe navigation of the surgical microscope. Skilled manoeuvring of the microscope is governed by an optimum set of kinematic quantities such as motion time, motion path length, motion smoothness, motion velocity and motion acceleration. Without knowing what these optimum values are, the level of skill cannot be quantified. In the literature, no particular publications exist that directly address the dynamics behind the microscope's motion during the myringotomy procedure or any ENT procedure for that matter. However, many publications do exist that address the microscope's motion tracking as an integral part of a surgical navigation system (i.e., augmented reality based surgery and image-guided surgery systems). All of these past works are presented briefly in the next two sub-sections.

1.3.2 Application of motion tracking in surgery

Introduction of motion tracking to surgical intervention prompted one of the significant changes in surgical treatment and surgical training over the past decade. Motion tracking made possible the whole domains of image guided surgery, augmented reality based surgery, virtual reality based surgical simulators, surgical tool tracking and hand motion analysis. Though the realization of any of these technologies requires sophisticated integration of medicine and engineering, the end goal is to make surgical treatments easy and safe and make surgical training more reliable, objective and safer. However in the current ENT surgical training approach, as stated in Section 1.2, instructors still use the "see one, do one, teach one" approach. In this approach, there is no way to teach surgical microscope manoeuvrability based on objective measures. In order to do so, one approach is to study the kinematics of surgical microscope manoeuvres during an operation, and this is possible only by acquiring of motion data of the microscope in an accurate and feasible manner.

In searching for a suitable motion tracking technique, application of motion tracking in various domains related to surgery were looked at in literature. From the literature research, it was found that though the application of motion tracking in surgery is problem specific, the format of the acquired raw data of any motion of any object is always the same. That is the raw data are composed of 3D translational coordinates and 3D angular coordinates of discrete points collected over a period of time within a global coordinate system. Since microscope motion tracking is the essential component of this project, the primary interest was to learn from past literature what equipment, software and methods successfully produced accurate raw motion data in various surgical applications that exercised motion tracking. At first, the work of Edwards *et al.* [13] was reviewed. This group integrated augmented reality with a surgical microscope to introduce augmented reality based guided interventions in ENT and neurological surgeries. The main idea behind their work was to project in real-time a patient specific 3D anatomical model generated from pre-operative CT and MR images onto the patient through the microscope's optics. The purpose of this idea was to enable the surgeon to see in real-time all the critical structures at and around the targeted anatomical structure so as to avoid any surgical mistakes and see the surgical changes being done to the anatomy in comparison to the overlaid model. In pursuit of their work, their challenges involved calibration, segmentation, registration, motion tracking and real-time visualization with motion tracking being the integral component as opposed to a separate step. In regards to motion tracking, they had to constantly track the patient's head on the operating bed with the optical tracker mounted on the microscope's end frame. Since the microscope was moved around to position and orient it differently as required during the operation, the microscopic view of the surgical site also changed. Therefore, to always visualize the anatomical model as accurately registered onto the real anatomy, regardless of the line of sight and the distance between the microscope optics and the patient, motion tracking was used. They marked the head with infrared light-emitting diodes and tracked them with the Optotrak 3020 optical tracking system (Northern Digital Inc. (NDI), Waterloo, ON, CA). The stationary global coordinate system of the head was calibrated with the dynamic local coordinate system of the tracker. The tracking system was used to ensure that regardless of the microscope's motion, the 3D position and orientation of the projected model through the microscope optics would align accurately with the actual structure under surgery.

The accuracy of real-time registration in an image-guided surgical navigation system depends on the calibration accuracy of the motion tracker's global/local coordinate system. Coordinate calibration is essential in defining the tracker's field-of-view with an accurate rigid or dynamic frame of reference, point of origin and axes orientations in order to generate accurate motion data. In the computer vision literature, camera calibration dives deep into the fundamentals of coordinate calibration and real-time image processing [14 - 16]. Likewise, tracker calibration is a similar research problem sharing the same fundamentals and is widely applied in surgical navigation systems requiring significant accuracy. Among the many publications on this topic, the work of Xu and Taylor [17] on electromagnetic tracker calibration and the work of García *et al.* [18] on calibration of a surgical microscope with optical trackers are most relevant to this project. Xu and Taylor developed a framework for calibrating an electromagnetic tracker (NDI Aurora system, Waterloo, ON, CA) with an accurate optical tracker (NDI Optotrak system, Waterloo, ON, CA). By simultaneously tracking an object with both the Aurora and Optotrak tracking tools, they first registered the stationary global coordinate systems of both trackers to each other. Later when the object moved, the error field between the dynamic coordinate systems was approximated and minimized using Bernstein polynomial based cost functions. They claimed that by error correction, both position and orientation data can be significantly improved. On the other hand, García *et al.* calibrated the Leica M-500 surgical microscope (Leica Microsystems, Heerbrugg, Switzerland) and the Atracsys easyTrack 200 optical tracker (Atracsys LLC, Le Mont-sur-Lausanne, Switzerland) via three major steps: optical calibration of the microscope's optical field-of-view to that of the tracker, calibration and registration of the tracker's coordinate system with that of the microscope's line of sight based coordinate system by registering both of them to a common reference grid and zoom/focus modelling. They concluded that their calibration method is faster, allows easy re-calibration and is transferrable to other devices (i.e., endoscopes).

A number of research works on specific surgical procedures have been published in the past that utilized motion tracking as an integral component [19 - 21]. Among them, the most recent publication was on Image-Enhanced Surgical Navigation (IESN) for endoscopic sinus surgery (ESS) by Lapeer *et al.* [22]. This group investigated

IESN for ESS mainly considering the rigid anatomy of the nasal cavity and passage. In their IESN system, the real image of the surgical site acquired by the endoscope's optics in real-time was overlaid on its virtual pre-operative CT data. The pre-operative CT data was calibrated to the patient's tracker tool and the real data was calibrated to the endoscope's tracker tool, while both tracker tools were calibrated to the world coordinate system of the optical tracking device. Therefore, registration of real and virtual images is based on the real-time coordinate transformation of the virtual CT data onto the endoscope's local coordinate system that navigates the coordinated map of the real data. Their experimental tests and validation results demonstrated expected accuracy in calibration, registration and motion tracking in accordance with other validated IESN systems. To obtain the most reliable data, the system was tested and validated by expert surgeons as they have the best surgical dexterity and the finest hand-eye coordination in the field. All endoscopic surgeries require long instruments and are performed through small surgical openings, which resemble the general procedure of laparoscopic surgeries. Since the early 2000's, many research works have been published in the literature on objective skill assessment of endoscopic and laparoscopic surgeries using motion tracking technologies. Among them, the recent study on objective assessment of laparoscopic suturing skills using the NDI Aurora electromagnetic tracker conducted by Yamaguchi *et al.* [23] demonstrated an in-house setup of a laparoscopic training system. While tracking the instruments in real-time, the system assessed the laparoscopic suturing and knot tying skills based on *time*, *path length* and *average speed* of the forceps in each hand. Time and path length are the most common skill assessment metrics found in related literature as discussed in Section 1.3.4. However, the computed average speed of a task is a unique metric that demonstrates a significant difference between right hand and left hand as the subjects performed suturing and knot tying tasks. Everyone was right handed; therefore, everyone had a faster speed using their right hands. However, experts had higher right hand speed scores than novices and also performed better overall. Another notable matter in this study was that the experimental setup was done following a standardized protocol to conduct the experiments in a controlled fashion.

Apart from its integral yet direct application in augmented reality, image guidance and surgical skill assessment, motion tracking is also used for testing and validating VR based surgical training simulators. One purpose of these simulators is to replace the constant supervision and guidance of an instructor while mastering certain surgical skill (i.e., drilling, cutting, navigating, suturing, etc.). To date, many VR training simulators with integrated metric based performance evaluators have been developed for research and commercial purposes [24, 25]. Among them, the most relevant simulators to our myringotomy simulator, discussed in Section 1.3.2, are a VR temporal bone drilling simulator [26], a VR mastoidectomy simulator [27], a mastoidectomy simulator [28, 29] and a VR laparoscopy trainer [30]. The publications associated with these simulators discuss in detail the derivation, implementation, evaluation and validation of metrics that constitute the integrated automated performance evaluators of the simulators. These evaluators virtually track the motions of the virtual tools, as they are used to perform a virtual task, and provide an automated metric analysis of the motion. Therefore, in metrics testing and validation studies conducted on the aforementioned VR surgical training simulators, electromagnetic and optical trackers were used to externally track the operator’s hand motion or the motion of the stylus/tooltip of the interfacing haptic device during a practice session. The externally tracked motions were then analyzed based on the same metrics built within the VR trainers. Statistical tests were then conducted for two separate reasons. Firstly, the tests were done to compare the metric results computed by the VR trainers as the experts and residents used them. Secondly, the tests compared the VR trainer metric results with the same metrics computed externally from data collected through motion tracking while the participants used the trainers. Based on the statistical test results, judgments were made on the accuracy, precision, consistency and effectiveness of the VR trainers and the integrated performance evaluators.

1.3.3 Metric based assessment of surgical skills

Various variables factor in when analyzing a dynamic motion path, which is the case for surgical microscope manoeuvres. Whether it is a large gross motion of the microscope or a small fine adjustment of its eyepiece, any change makes a difference in obtaining an optimal view of the object of interest. An expert ENT surgeon with fellowship training may obtain a quick focus on the object of interest through the microscope without thinking much about his/her course of action, while it may yet be the toughest task to accomplish for a resident-in-training. Without knowing what variables collectively define the microscope manoeuvring efficiency, flaws are certainly introduced in the training process. These flaws are mostly ignored so long a satisfactory microscopic view of the object of interest is obtained. However, unfortunately, flaws like longer motion time and path length, lesser path smoothness and more motion jitter combine to slow improvement in microscope manoeuvrability.

In order to define metrics specific to microscope manoeuvring, various applications of metrics were looked at in the literature related to surgical training, motion-based skill assessment, performance evaluation, hand motion analysis and surgical simulators. In regards to surgical training, the most common surgical procedure found in the literature is the laparoscopic procedure. In 2002, Cotin *et al.* [31] followed a scientific approach to define metrics to assess computer-assisted laparoscopic skill training. They used a five degree of freedom motion tracking device and a software platform to track the motion of the laparoscopes in use, process the motion data in real time and provide feedback. By closely observing a series of training sessions in the operating room, they subjectively determined various components of a laparoscopic surgical task that in combination define the skill level of a performer. They listed these components as compact spatial distribution of the tool's tip, tool's motion smoothness, good depth perception, shorter completion time, smaller rotational orientation and ambidexterity. They used simple kinematic metrics such as *time*, *path length*, *smoothness* and *total axial rotation* of the instrument to quantify the identified components. To validate the effectiveness of these metrics statistically, the standardized score (z-score) of the results collected from both novice and experts, were computed and compared. Finally they reported that the higher z-scores correspond to less experience while the lower z-scores correspond to more experience. Based on their statistical analysis, they claimed that their approach and metric definitions can be introduced in VR based training simulators for laparoscopic surgery and perhaps in other surgical simulators as well.

Respecting the conclusion of Cotin *et al.*, Acosta and Temkin introduced the performance metrics into their LapSkills laparoscopic surgical simulator [32] in 2005. In their simulator, they incorporated several fundamental skill training tasks such as laparoscope navigation and exploration, hand-eye coordination, grasping, applying clips and cutting. The advantage of the metrics defined and implemented by Cotin *et al.* is that these metrics are task independent. Hence, they are applicable in any situation. Acosta and Temkin utilized this advantage by combining the appropriate task independent metrics to create their own customized metrics in order to evaluate skill training tasks. For instance, the simulated laparoscopic navigation through a virtual tube with long length and smaller radius would be assessed based on completion *time*, *force* registered on the tube wall, *number of times the tube is touched*, the total *path length* and the *smoothness* of the path travelled by the virtual laparoscope's tip. Acosta and Temkin also conclusively claimed that their simulator can be configured to simulate different surgical procedures and serve as a training tool with metric based real-time assessment feedback.

Following Acosta and Temkin, Oostema *et al.* carried out another study on "time efficient skill assessment using augmented reality simulator" in 2007 [33]. They used the ProMIS hybrid virtual reality trainer [34] to practice on camera navigation, object positioning, sharp dissection and intracorporeal knot tying. They performed these tasks from very easy level to very difficult level. The simulator evaluated the performance based on task *completion time*, *smoothness* and *path length*. Again, data was collected from expert surgeons (i.e., with experience on 1000+ procedures), 3rd year medical students (i.e., with no experience) and residents from year 1 to 5 (i.e., with varying but less experience than experts). The metrics from all three groups were statistically correlated using regression analysis and analysis of variance (ANOVA). From these analyses, they observed statistically significant correlation between experience level and performance in all 3 metrics for all 4 tasks only during the very difficult level. However, smoothness and time scores always showed significant correlation with experience at every difficulty level. Based on their findings, they concluded that although all 3 metrics showed significant results, they are quite general and apply to many procedures. Perhaps if more specific metrics applicable only for a certain laparoscopic procedure are determined, actual skills can be measured in the virtual trainer more deterministically.

In most VR based skill assessment studies, path smoothness is found to be a very important metric. Smoothness is quantified based on the lack of jaggedness in a motion path. In other words, the lesser the jaggedness, the smoother is the motion path. Typically, an expert surgeon would carry out a surgical procedure with surgical tools by optimizing and economizing all of his/her moves. Therefore an expert, even in a complex task like knot tying, would trace a smooth path with the tool, independent of time and path length. This smoothness, however, depends upon one's motion fluidity and dexterity with surgical tools that come with experience. In order to validate the accuracy of smoothness measurement, motion paths containing various levels of path jaggedness need to be considered. This is why smoothness is the most effective metric to do research work related to Parkinson's disease (PD) and vice-versa. Since PD progressively deteriorates one's motor control, the patient is seen to have progressive levels of tremor. The correlation between this tremor increment, or motion jaggedness, and progression of PD is then measurable in terms of smoothness of motion paths created during drawing exercises completed by the PD patients. Several such experimental studies have been done in the past. Among them, the works of Buch and Contreras-Vidal [35], Tresilian *et al.* [36] and Teulings *et al.* [37] were found to demonstrate the effectiveness of smoothness measurement in quantifying the amount of loss of motor control of PD patients in terms of increasing levels of motion path jaggedness. In all of these literary works, the same fundamental kinematic equation was used to quantify the level of jaggedness. It was commonly called Normalized Motion Jerk. This equation essentially quantifies the amount of motion jerk (i.e., vibration or jaggedness) in a motion path. In terms of experimental setups, all of these groups had some default drawings with known motion jerk scores. PD patients from various stages were then asked to redraw those drawings on paper while their hands were tracked and the data was analyzed with the aforementioned equation. Though all the groups followed different statistical analytical techniques, they all arrived at similar conclusions sharing the idea that high motion jerk scores mean unsmooth motion while low scores mean smoother motion. Therefore, by comparing the normalized motion jerk scores of PD patients with the default values and also comparing these scores within PD patients group, the loss of motor control of each patient was quantified. Following this validated approach, the smoothness metric can be similarly applied to surgical tool handling and determine how it defines one's tool handling skill during an operation independent of any other metrics.

Besides smoothness, force also serves as a surgical skill assessment metric. Yamauchi *et al.* presented their developmental and validation work on Endoscopic Sinus Surgery (ESS) training system [38] based on force calculations. Their ESS training system is composed of a physical head dummy model. The interior and exterior of its nasal cavity is an exact replica of a real nasal cavity based on CT scans. The interior platform of the model is equipped with force sensors. In addition, the endoscope's tip is equipped with an optical position sensor to record path length over time. When the endoscope is inserted through the nasal cavity to reach a predetermined target, position is recorded and the force detected on the platform is recorded. For force data collected from novice and expert groups, three indices were determined: maximum, average and integral based on the absolute values obtained from the force sensors. Non-parametric tests were performed on these indices which included Friedman's test and Wilcoxson's test. However, they found statistical significance in the integral force indices only. Since clinically the integral index of the force data reflects both the magnitude and duration of friction cause by the endoscope's tip contacting the tissue during operation, they considered it as the most adequate force index of surgical skill in ESS. Furthermore, their statistical finding validated their consideration. With analysis of variance of integral force indices, under the presumption of normal distribution, they also found significant difference between expert and novice participants. They proposed that since integral force indices demonstrated reliable results for skill assessment, the ESS Training System should be further improved with 6-DOF force sensors and position sensors. Such improvements would make the surgical tool force and motion analysis more sophisticated, more reliable and precise.

Unlike endoscopic or laparoscopic surgeries, many surgeries are performed in close contact with the surgical site. In such cases, performance evaluation can be done via hand motion analysis. Grober *et al.* implemented this concept in 2003 to objectively measure residents' and experts' microsurgical skills and stereoscopic visual acuity [39]. They used the Imperial College Surgical Assessment Device (ICSAD) (Imperial College, London, UK) and electromagnetic trackers mounted on the dorsal surface of the hands. The purpose was to record and analyze the positional data of the operator's hands during a given standardized microsurgical task. The ICSAD collected the data recorded by the trackers and computed the economy of hand motion by computing number of hand movements, hand travel distance, direction and acceleration changes. The values of these metrics were statistically correlated to a previously approved subjective evaluation method using the Spearman ρ and Pearson r correlation coefficients. Paired t-tests were also performed to compare pre- and post-training hand motion analysis scores. Finally they reported that the subjective ratings significantly correlated with the metric based scores in order to claim the effectiveness of their defined metrics in assessing the microsurgical skills of the residents.

Later in 2007, Kinoshita *et al.* further stretched the concept of hand motion analysis to movement of fingers during use of tools [40]. They used the integrated sensor based CyberGlove to capture and record dextrous and cyclical hand motions. They evaluated these hand motions in terms of accuracy, repeatability and efficiency. The formulae of these metrics were derived by integrating multiple statistical functions together. The CyberGlove has 22 bend sensors that measure 3D angular change of and between fingers as they are flexed simultaneously or relative to each other. These angular data are classified using the k-mean method by their pattern, phase-lag and quantity of the movement. In terms of experimentation, they tested the hand motion during chopstick use and during rotating two balls in the hands. They collected these hand motion data from an expert group and a novice group for statistical comparison. They concluded that based on the comparative result they obtained from statistical computation, their evaluation methods gave better understanding of the improvement of accuracy and repeatability of the two particular hand motions. However, the results were not able to fully explain the efficiency of the two hand motions. They also claimed that these methods would be able to evaluate the performance level of particular hand motions within an instructional system. Furthermore, the work of Stefanidis *et al.*[41] validates the usefulness of motion metrics such as path length and smoothness, in addition to time metrics, in assessing improvement of proficiency level while training with simulators.

1.4 Objectives

Economized and safe navigation of a surgical microscope during myringotomy, or during any ENT procedure, is a challenging task to perform for novice ENT surgical residents. Currently there are no guidelines for training in microscope manoeuvring, nor there are any objective measures to assess the microscope manoeuvring performance of an individual. As of now, residents get better at this through self-discovery and repetitive practice. This process takes a longer period of time resulting in a frustrating experience for both residents and instructing surgeons. Therefore, *the central objective of this project is to identify discriminatory numerical metrics to assess surgical microscope manoeuvring performance of an individual during myringotomy*. To achieve this central objective, the project has been broken down into chronological phases. The categorized structure of the project with corresponding objectives is described below.

A. Technical Phase

1. **Motion tracking:** Develop a real-time motion tracking module for the optical tracking system in order to track multiple optical markers, visualize their motion in virtual space and record their discrete motion data separately.

2. **Software testing:** Test and improve the accuracy of real-time data acquisition and the functionality of the tracking module.

B. Analytical Phase I

3. **Metric research:** Derive and implement objective metrics for quantitative analyses of microscope's motion and subjective metrics for qualitative analysis of operating ability and visibility through microscope optics.

4. **Metric validation:** Validate the accuracy and the effectiveness of the derived metrics by testing them on various controlled motion paths that have been previously analysed accurately using a different motion tracking device.

C. Experimental Phase

5. **Experimental design:** Deign the experimental study based on standardized procedures, protocols and settings of all equipment and conduct a few trial runs to

optimize all steps pertaining to data collection and confirm the accuracy of the metric based motion analysis.

6. **Data collection:** Conduct the full experiment resetting all experimental settings for each participant and develop a database containing all raw data categorized by the participant's class: expert or resident.

D. Analytical Phase II

7. **Metric evaluation:** Perform analyses on the numerical and visual data collected during experiments.

8. **Statistical analysis:** Demonstrate all the metric results in appropriate graphic forms and perform parametric/nonparametric tests where appropriate in order to determine all the discriminatory metrics.

1.5 Scope

The focus of this project is on metric derivation and implementation in order to evaluate myringotomy specific surgical microscope manoeuvrability. In general, assessment of manoeuvrability is an analysis of motion dynamics based on parameters such as motion time, motion speed, motion path length, motion smoothness, motion jitter and motion repeatability. Although such analysis is being performed on myringotomy specific motion dynamics, the potential scope of the work is quite broad. Upon successful completion of the metric study of the microscope's motion during myringotomy validated by statistical analyses, the same metric based evaluation technique can be applied to other ENT surgical procedures where microscope manoeuvre is an essential component. However, first to make sure the objectives of this project are feasible within the given timeframe, a set of deliverables are outlined.

1.5.1 Deliverables

1. A motion tracking software module that connects and communicates with the motion tracker unit, to be used in this project, in order to trace, record and virtually visualize the motion(s) of tracker tool(s) within the field-of-view of the tracker.

2. Sets of metrics that evaluate the motion path(s) of the microscope during myringotomy and the operation skills of the operator(s).

3. Representation of the experimental data collected during experimental trails of myringotomy.

4. Analyses and discussion of manoeuvring and operating skills based on parametric and non-parametric statistical tests done on the raw data.

1.6 References

[1] J. Bavosi, “Artwork of ear, nose and throat in a cold sufferer,” File: M1300429-Artwork_of_ear,_nose_throat_in_a_cold_sufferer-SPL.jpg, *Science Photo Library*, [Online]. Available: http://www.sciencephoto.com/media/252538/enlarge# [Accessed: April 12, 2012].

[2] Leica Microsystem, “Leica M720 OH5 Premium Surgical Microscope – A New Dimension of Innovation and Comfort,” www.leica-microsystems.com, [Online]. Available: http://www.leica-microsystems.com/products/surgical-microscopes/ neurosurgery-spine/details/product/leica-m720-oh5-2/ [Accessed: May 19, 2012].

[3] My Child Health, “Middle ear infection in children,” File: middle-ear-infection.jpg, www.mychildhealth.net, [Online]. Available: http://www.mychildhealth.net/wp-content/uploads/2009/04/middle-ear-infection.jpg [Accessed: May 26, 2012].

[4] J. Stowe, “Myringotomy with Bilateral Tube Placement," File: 2008-JS Myringotomy.jpg, *Georgia Health Sciences University*, [Online]. Available: http://www.georgiahealth.edu/medart/MI-Class2008.html [Accessed: May 26, 2012].

[5] M. L. Montague, M. S. W. Lee and S. S. M. Hussain, "Human error identification: an analysis of myringotomy and ventilation tube insertion," *Archives of Otolaryngology-Head and Neck Surgery,* vol. 130, no. 10, pp. 1153, 2004.

[6] T. Walker, S. Duvvi and B. N. Kumar, "The Wigan grommet trainer," *Clinical Otolaryngology*, vol. 31, no. 4, pp. 349-350, 2006.

[7] M. Duijvestein and J. Borgstein, "The Bradford grommet trainer," *Clinical Otolaryngology*, vol. 31, no. 2, pp. 163, 2006.

[8] A. Leong, S. Kundu, P. Martinez-Devesa and C. Aldren, "Artificial ear: a training tool for grommet insertion and manual dexterity," *ORL,* vol. 68, no. 2, pp. 115-117, 2006.

[9] B. Wheeler, P. Doyle, S. Chandarana, S. Agrawal, M. Husein, H. Ladak, “Interactive computer-based simulator for training in blade navigation and targeting in myringotomy,” *Computer Methods and Programs in Biomedicine*, vol. 98, pp. 130-139, 2010.

[10] G. Rehal, "Virtual reality simulation for myringotomy training with haptic feedback," M.E.Sc. thesis, The University of Western Ontario, London, Canada, 2009.

[11] L. J. Sowerby, G. Rehal, M. Husein, P. C. Doyle, S. K. Agrawal and H. M. Ladak, "Development and face validity testing of a three-dimensional myringotomy simulator with haptic feedback," *J. Otolaryngol. Head. Neck. Surg.*, vol. 39, no. 2, pp. 122-129, Apr, 2010.

[12] A.K. Ho, H. Alsaffar, P. C. Doyle, H. M. Ladak, S. K. Agrawal, “Virtual reality myringotomy simulation with real-time deformation: Development and validity testing,” *Laryngoscope*, vol. 122, no. 8, pp. 1844 – 1851, May, 2012.

[13] P. J. Edwards, A. P. King, D. J. Hawkes, O. Fleig, C. R. Maurer, Jr, D. L. Hill, M. R. Fenlon, D. A. de Cunha, R. P. Gaston, S. Chandra, *et al.*, “Stereo augmented reality in the surgical microscope,” *Stud Health Technol Inform.*, vol. 62, pp. 102–108. 1999.

[14] Q. Xu, D. Ye, R. Che, Y. Huang, "Accurate Camera Calibration with New Minimizing Function," Robotics and Biomimetics, 2006. ROBIO '06. IEEE International Conference on, pp.779-784, Dec. 2006.

[15] A. Datta, Kim Jun-Sik, T. Kanade, "Accurate camera calibration using iterative refinement of control points," *Computer Vision Workshops (ICCV Workshops)*, 2009 IEEE 12th International Conference on, vol., no., pp.1201-1208, Sept. 27 2009-Oct. 4 2009.

[16] T.T. Li, H.Y. Zhang, J. Geng, "Geometric calibration of a camera-projector 3D imaging system," *Image and Vision Computing New Zealand (IVCNZ)*, 2010 25th International Conference on, pp.1-8, 8-9 Nov. 2010.

[17] X. Wu; R.H. Taylor,"A framework for calibration of electromagnetic surgical navigation system," *Intelligent Robots and Systems (IROS 2003)* Proceedings. IEEE/RSJ International Conference on, vol.1, pp. 547- 552, 27-31 Oct. 2003.

[18] J. García, R. Thoranaghatte, G. Marti, G. Zheng, M. Caversaccio, M. A. G. Ballester, “Calibration of a surgical microscope with automated zoom lenses using an active optical tracker”, *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, no. 1, pp. 87–93, Mar. 2008.

[19] L. Moody, C. Baber, T. N. Arvanitis, “Objective metrics for the evaluation of simple surgical skills in real and virtual domains,” *Presence*, vol. 12, no. 2, Arp. 2003.

[20] M. Agus, A. Giachetti, E. Gobbetti, G. Zanetti, A. Zorcolo, "Adaptive techniques for real-time haptic and visual simulation of bone dissection," *Virtual Reality, 2003. Proceedings. IEEE* , pp. 102- 109, 22-26 Mar. 2003.

[21] C. E. Reiley, H. C. Lin, D. D. Yuh, G, D. Hager, “Review of methods for objective surgical skill evaluation,” *Surg. Endosc.*, vol. 25, pp. 356–366. 2011.

[22] R. Lapeer, M.S. Chen, G. Gonzalez, A. Linney, G. Alusi, “Image-enhanced surgical navigation for endoscopic sinus surgery: evaluating calibration, registration and tracking,” *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, no. 1, pp. 32–45, Mar. 2008.

[23] S. Yamaguchi, D. Yoshida, H. Kenmotsu, T. Yasunaga, K. Konishi, S. Ieiri, H. Nakashima, K. Tanoue, M. Hashizume, “Objective assessment of laparoscopic suturing skills using a motion-tracking system,” *Surgical Endoscopy*, vol. 25, pp. 771 – 775, 2011.

[24] Simendo, “Simendo Laparoscopy,” [Online]. Available: http://www.simendo.eu/ products/laparoscopy/ [Accessed: Jun 1, 2012].

[25] Voxel-Man, “VOXEL-MAN Tempo,” [Online]. Available: http://www.voxel-man.de/simulator/tempo/[Accessed: Jun 1,2012].

[26] M. Zirkle, D. W. Roberson, R. Leuwer, A. Dubrowski, “Using a Virtual Reality Temporal Bone Simulator to Assess Otolaryngology Trainees,” *Laryngoscope*, vol. 117, issue 2, pp 258 – 263, 2007.

[27] M. Agus, A. Giachetti, E. Gobbetti, G. Zanetti, A. Zorcolo, N.W. John, R.J. Stone, “Mastoidectomy Simulation with Combined Visual and Haptic Feedback,” *Stud Health Technol Inform.*, vol. 85, pp 17-23, 2002.

[28] C. Sewell, D. Morris, N. H. Blevins, S. Agrawal, S. Dutta, F. Barbagli, K. Salisbury, “Validating Metrics for a Mastoidectomy Simulator,” *Studies in Health Technology and Informatics*, vol. 125, pp. 421-426, 2007.

[29] C. Sewell, D. Morris, N.H. Blevins, F. Barbagli, K. Salisbury, “Evaluating drilling and suctioning technique in a mastoidectomy simulator,” *Studies in Health Technology and Inform.*, vol. 125, pp. 427-432, 2007.

[30] E. Hiemstra, E.M. Terveer, M.K. Chmarra, J. Dankelman, F.W. Jansen, “Virtual reality in laparoscopic skills training: is haptic feedback replaceable?,” *Minim Invasive Ther Allied Technol.*, vol. 20, no. 3, pp 179- 184, May 2011.

[31] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, S. Dawson, “Metrics for Laparoscopic Skills Trainers: The Weakest Link!,” Medical Image Computing and Computer Assisted Intervention (MICCAI), *Lecture Notes in Computer Science (LNCS)*, vol. 2488, pp. 34 – 43, 2002.

[32] E. Acosta, B. Temkin, "Haptic laparoscopic skills trainer with practical user evaluation metrics," *Stud Health Technol Inform.*, vol. 111, pp. 8 – 11, 2005.

[33] J. A. Oostema, M. P. Abdel, Jon C. Gould, "Time-efficient laparoscopic skills assessment using an augmented reality simulator," *Surgical Endoscopy*, vol.22, no. 12, pp. 2621- 2624, Dec. 2008.

[34] CAE Healthcare, "ProMIS hybrid virtual reality trainer," [Online]. Available: http://www.cae.com/en/healthcare/promis.simulator.asp [Accessed: May 30, 2012].

[35] E. R. Buch, J. L. Contreras-Vidal, "Effects of Parkinson's disease on visuo-motor adaptation," *Exp Brain Res*, vol. 150, pp. 25–32, 2003.

[36] James R. Tresilian, George E. Stelmach1 and Charles H. Adler, "Stability of reach-to-grasp movement patterns in Parkinson's disease," *Brain*, vol. 120, no. 11, pp. 2093-2111, 1997.

[37] Hans-Leo Teulings, J. L. Contreras-Vidal, G. E. Stelmach, C. H. Adler, "Parkinsonism Reduces Coordination of Fingers, Wrist, and Arm in Fine Motor Control," Experimental Neurology, vol. 146, pp. 159-170, 1997.

[38] Y. Yamauchi, J. Yamashita, O. Morikawa, R. Hashimoto, M. Mochimaru, Y. Fukui, H. Uno, K. Yokoyama, "Surgical Skill Evaluation by Force Data for Endoscopic Sinus Surgery Training System," *LNCS*, vol. 2488, pp. 44–51, 2002.

[39] E. D. Grober, S. J. Hamstra, K. R. Wanzel, E. D. Matsumoto, R. S. Sidhu, K. A. Jarvi, "Validation of novel and objective measures of microsurgical skill: hand motion analysis and stereoscopic visual acuity," *Microsurgery*, vol. 23, pp. 317-322, 2003.

[40] Y. Kinoshita; D. Takeda; A. Sasaki; H. Hashimoto, "Archive and Instruction of the Hand Motion: Analysis and Evaluation of the Hand Motion in Use of Chopsticks," *SICE-ICASE, 2006. International Joint Conference*, pp.2348-2353, 18-21 Oct. 2006.

[41] D. Stefanidis; D. J. Scott; J. R. Korndorffer Jr., "Do Metrics Matter? Time Versus Motion Tracking for Performance Assessment of Proficiency-Based Laparoscopic Skills Training," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 4, no. 2, pp.104-108, 2009.

Chapter 2

2 The clinical study to assess surgical microscope usage during myringotomy

2.1 Introduction

The human auditory system is both anatomically complex and tiny; therefore ear, nose, and throat (ENT) surgeons require a surgical microscope to optimally view anatomical structures of interest during diagnosis and surgical treatment. The operating microscope itself is quite complicated and it presents a variety of challenges to surgical trainees. The first challenge is manoeuvring the microscope into a position that provides the optimal surgical view of the pathological site. The second challenge is having the dexterity and hand-eye coordination to perform microsurgery in a highly magnified field while looking through a microscope.

Learning proper microscope skills is essential for surgical trainees, and it is often overlooked as a prerequisite to learning the actual surgical procedure. First and foremost, it allows them to attain the optimal surgical view within which to perform the operative task. An obstructed view, improperly focussed optics, or inappropriate focal length can lead to surgical errors or the inability to complete the task. Second, trainees are often inefficient in the use of the operating microscope leading to long path lengths, unnecessary microscope repositioning, and wasted operative time.

Unfortunately, very little formal training is given towards surgical microscope use in conventional ENT training. Surgical residents are often required to manoeuvre a microscope without any prior practice or proper techniques to follow. Although they receive informal qualitative feedback on their overall performance, residents often do not know which particular manoeuvring skill they need to improve upon. Despite the use of the operating microscope in a number of surgical specialties and procedures, the individual skills and maneuvers required by trainees have not been objectively studied to date. This study aims to identify these prerequisite skills by objectively comparing the operating microscope usage between surgical experts and residents.

The central hypothesis of this study is that *quantifiable differences exist between experts and residents when manoeuvring a microscope to attain an optimum microscopic view of the eardrum during myringotomy*. Based on this hypothesis, three main objectives were established: (1) conduct a blinded trial to collect data on a group of experts and residents using the operating microscope to perform a procedure; (2) implement software to track the microscope's motion and develop a set of *tracking metrics* to numerically assess the tracked paths; and (3) produce a set of *procedural metrics* to assess one's surgical performance, *positional metrics* to assess the participant's body and arm location, and *optical metrics* to assess the field-of-view produced through microscope.

The surgery performed in this trial was myringotomy with tube insertion. This procedure was chosen as it is one of the first microscopic surgeries performed by surgical trainees, and it is one of the most common ENT procedures performed in North America [1]. In this procedure, trainees must place a speculum into the ear canal, and then position the operating microscope in order to obtain an optimum surgical view of the eardrum. A surgical blade is then guided through the speculum and down the ear canal in order to make a small incision in the eardrum called a myringotomy. Finally, a ventilation tube is carefully placed within the incision in order to provide aeration and allow fluid to drain from the middle ear space.

Although the procedure sounds simple, Montague *et al.* [2] have outlined a number of complications that can occur when trainees perform this surgery. Common errors include inappropriate magnification and view of the surgical site, multiple attempts at tube insertion, and inappropriate incision size. Consequently, the Auditory Biophysics Laboratory (ABL) at Western University, London, Canada is currently developing a virtual reality myringotomy simulator to train ENT residents [3 - 5]. The simulator also includes a 3D visor (eMagin Z800 3D visor, Bellevue, WA) mounted on an adjustable aluminum stand simulating a surgical microscope [6] (Figure 2.1). However, this simulated microscope is quite unrealistic as it does not have any articulations to allow for movement, nor does it have any zoom or focus functionality. Our long-term goal is to improve the realism of the simulated microscope and utilize the results of this study to create automated microscope metrics to provide trainees with feedback on their performance. The focus of this work is solely on the development and evaluation of metrics.

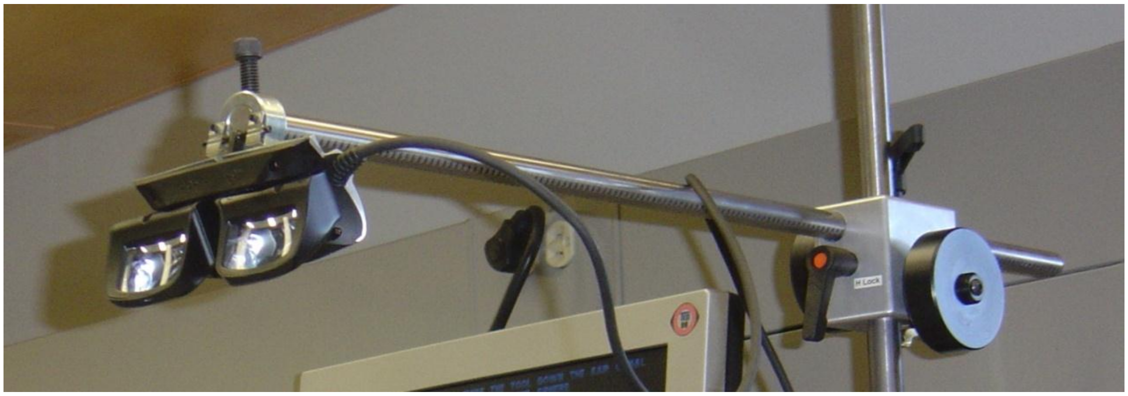

Figure 2.1: The current microscope simulator consisting of a stereo visor mounted on an adjustable aluminum stand [6].

2.2 Materials and methods

2.2.1 Subjects

All subjects were affiliated with the Department of Otolaryngology – Head & Neck Surgery at Western University. The expert group (n=4) consisted of Neurotologists and Pediatric Otolaryngologists performing a high volume of myringotomies in their practices. The Otolaryngology resident/trainee group (n=8) were junior residents in postgraduate years 1 to 3 of a 5-year curriculum. These residents had limited previous exposure to myringotomy and tube placement in the operating room. None of the participants in either group had previously used the particular surgical microscope used in the study (Leica M720 OH5, Leica Microsystems, Wetzlar, Germany), thereby ensuring this was not a confounding variable. Prior to each trial, a 20 minute orientation session was held with each participant reviewing the experimental procedure and thoroughly orienting them to the surgical microscope. The participants were then allowed to practice with the microscope for as long as they needed in order to feel comfortable with its functionality and movement. In addition, each resident was supplied with a baseline questionnaire (Appendix 2) to determine their baseline level of microscopic surgical experience. This was based on 1) time spent on Neurotology and Paediatric rotations and 2) previous exposure to ear examinations, myringotomy, microscope manoeuvring and simulators.

2.2.2 Surgical task

A myringotomy was performed on a fixed cadaveric eardrum by an expert ENT surgeon before making it available for experimentation. During each trial, the subject first had to position himself/herself on a height adjustable chair and relative to the operating table appropriately. Then the subject had to manoeuvre the end-frame of a surgical microscope from a common starting position, place a speculum in the ear canal, and obtain a focussed microscopic view of the pre-existing myringotomy. Finally, the subject had to guide a ventilation tube down the ear canal and appropriately insert it into the myringotomy using otologic instruments.

2.2.3 Experimental Setup and protocols

The experimental setup mimicked the basic operating room setup used during myringotomy and tube insertion cases at London Health Sciences Centre as shown in Figure 2.2. After the participants' orientation time, they were asked to leave the room while the equipment was reset to the same standardized baseline. In particular, the starting position and setting of the equipment was carefully controlled in order to test all subjects in similar conditions.

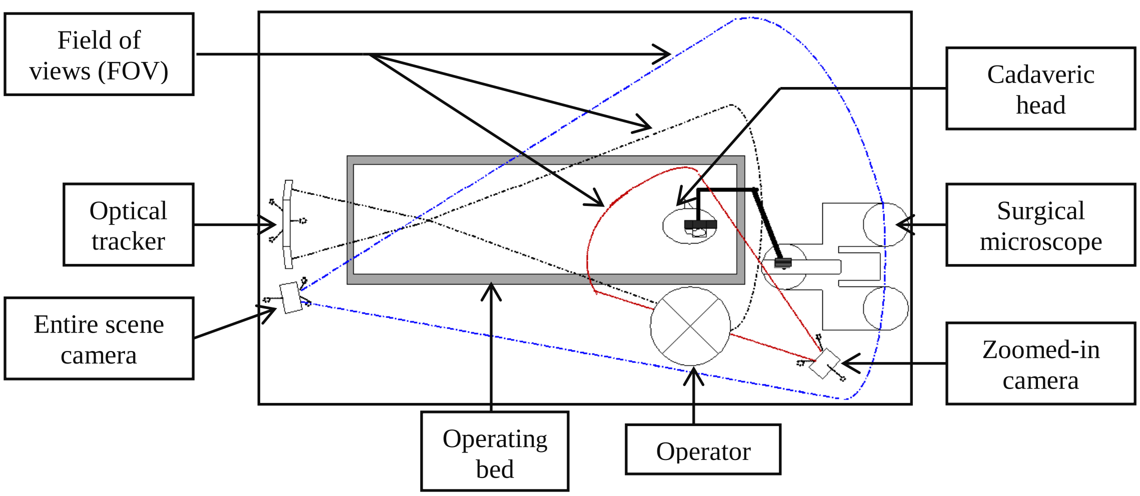

Figure 2.2: Top down schematic view of the experimental setup.

The base of the surgical microscope was parked 30 centimeters from the head of the operating table. The end-frame of the microscope was balanced and placed in the same starting position for each trial. The microscope settings were reset to minimum zoom (M = 2.7), minimum focus (WD = 200), minimum intraocular distance, and neutral tilt (180^o from ocular lens to microscope optics) to ensure that all participants had to perform similar adjustments in order to obtain an optimal view.

The cadaveric head was placed 10 centimeters from the head of the bed in the supine position. It was tilted 45 degrees away from the participant to expose the left ear. Sterile drapes and gloves were used to simulate an intraoperative procedure. The chair and table were placed in the same position and adjusted to the same height at the start of each trial. The instrument tray was placed in the usual position and this consisted of various aural specula, forceps, picks, and a standard Baxter ventilation tube.

The motion of the microscope was objectively measured using an optical tracker (Polaris® Hybrid, Northern Digital Inc., Waterloo, ON). A marker consisting of 3 reflective spheres was placed on the microscope frame, and the tracker with infrared detectors was placed such that the marker always stayed within its field-of-view (FOV) as shown in Figures 2.3, 2.4, and 2.5. The tracker uses reflected infrared light to triangulate the three-dimensional position and rotation of the microscope 60 times per second with an accuracy of $\pm$0.25 mm. Translations were recorded in millimeters and rotations were recorded in degrees. This allows for very smooth and accurate motion paths to be recorded and analyzed.

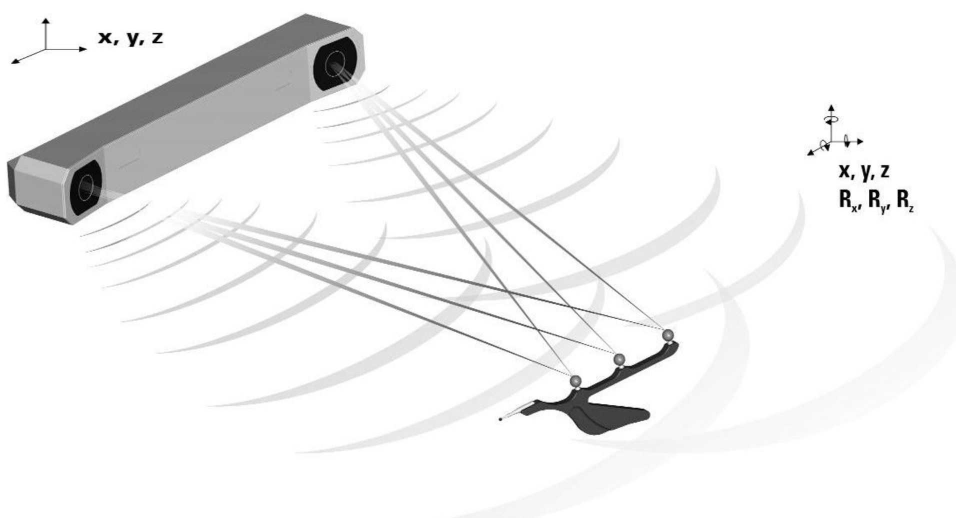

Figure 2.3: Optical tracking with Polaris® Hybrid, Northern Digital Inc., Waterloo, ON [24].

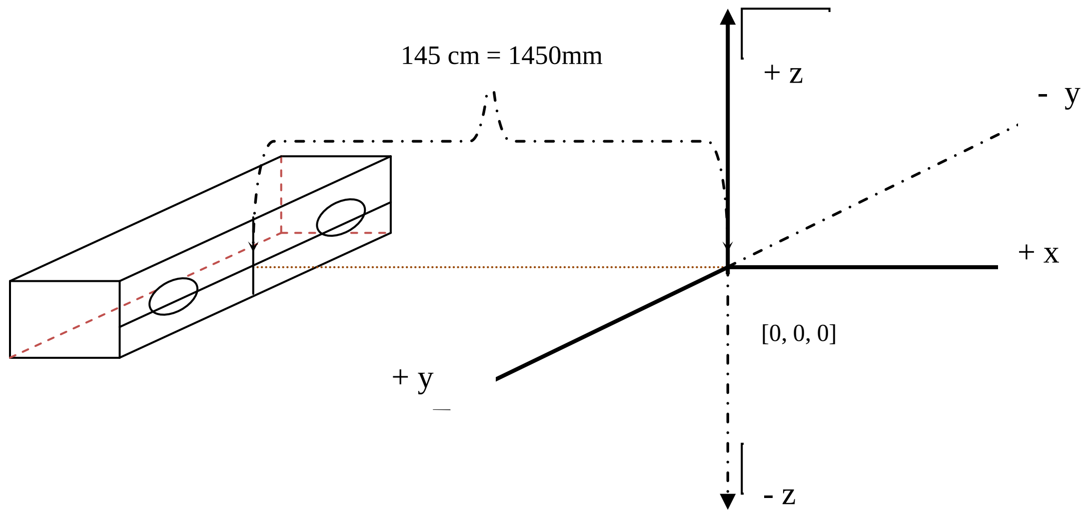

Figure 2.4: The location (145 cm forward in x direction from the front face of the tracker) of the point of origin and the orientation of the global lab coordinate system of the Polaris® tracker [24].

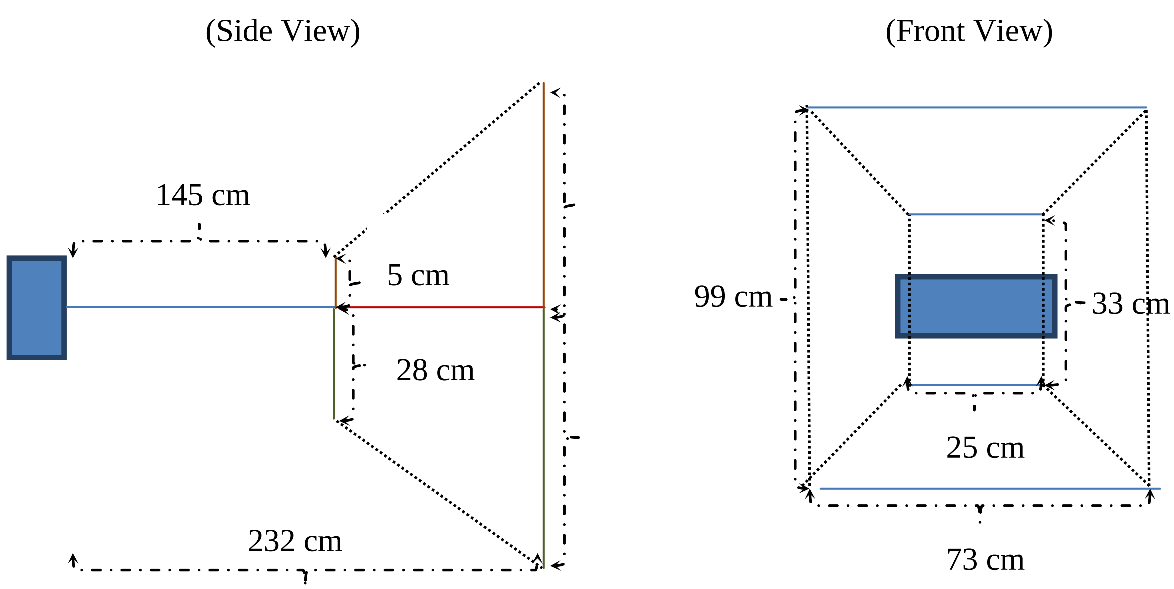

Figure 2.5: Practically measured FOV of Polaris® tracker. Here, the left figure shows the side view of the FOV while the right figure is the front view. These figures show the shape, size, location and orientation of the FOV with respect to the physical tracker.

A module was programmed using MATLAB® (MathWorks®, Natick, MA, USA) and the Image-Guided Surgery Toolkit (IGSTK, Clifton Park, NY, USA) to record the real-time tracking data and display the motion in three dimensions. The setup of the module is shown in Figure 2.6. Two high-definition video cameras were also used to capture the experimental scene. One camera captured the whole scene, whereas the second was zoomed onto the participant's hands to give a clear picture of the hand motion and position. The final experimental scene is shown in Figure 2.7. The internal microscope video camera was also high-definition (720p) and it captured the optical view seen by the participant. This camera was calibrated so that the zoom, FOV, and focus matched the optical view of the participants.

All motion tracking and video capture began when the participant entered the scene, and it ended once the ventilation tube had been placed into the myringotomy and the surgical instruments had been removed from the microscope's optical view. The captured video and animated tracking capture were then time-synchronized and compiled in a single four quadrant split-screen video using Vegas™ Pro 11 (Sony Creative Software, Middleton, WI, USA) video editing software. Figure 2.8 shows a screen shot of this video with the top-right screen showing the entire experimental scene, the top-left screen showing the zoomed-in recording of the hand motion, the bottom-left screen showing the corresponding real-time motion path and orientation of the end-frame, and the bottom-right screen shows the captured optical FOV of the microscope. These videos were then anonymized by blurring the faces of the participants.

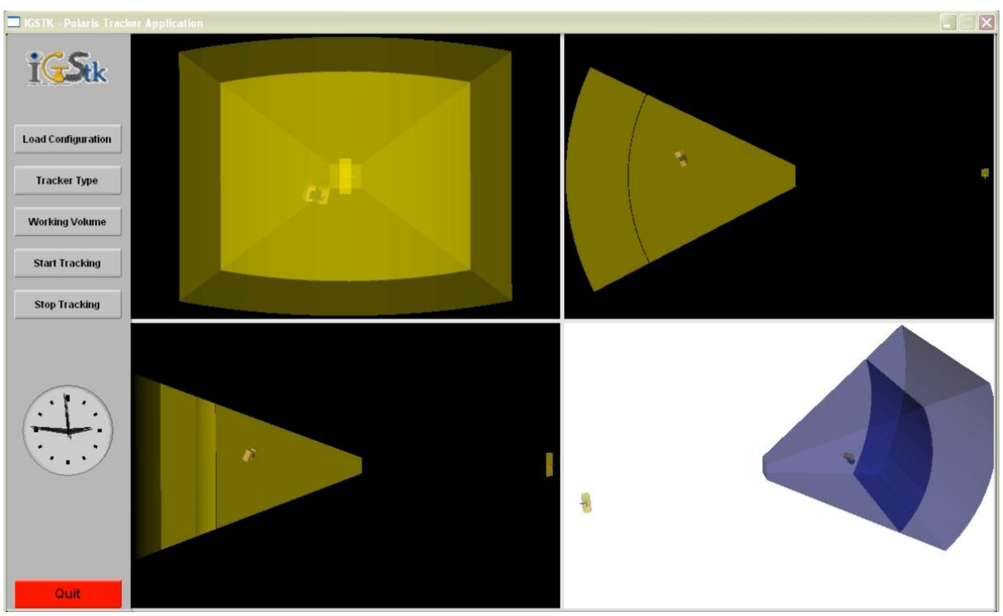

Figure 2.6: Screen shot of the implemented tracking module in IGSTK. The top-left view is the front view of the FOV, bottom-left is its side view, top-right is the top view while the bottom-right is the 3D view. The rectangular object in all of them is the marker that is being tracked in physical space.

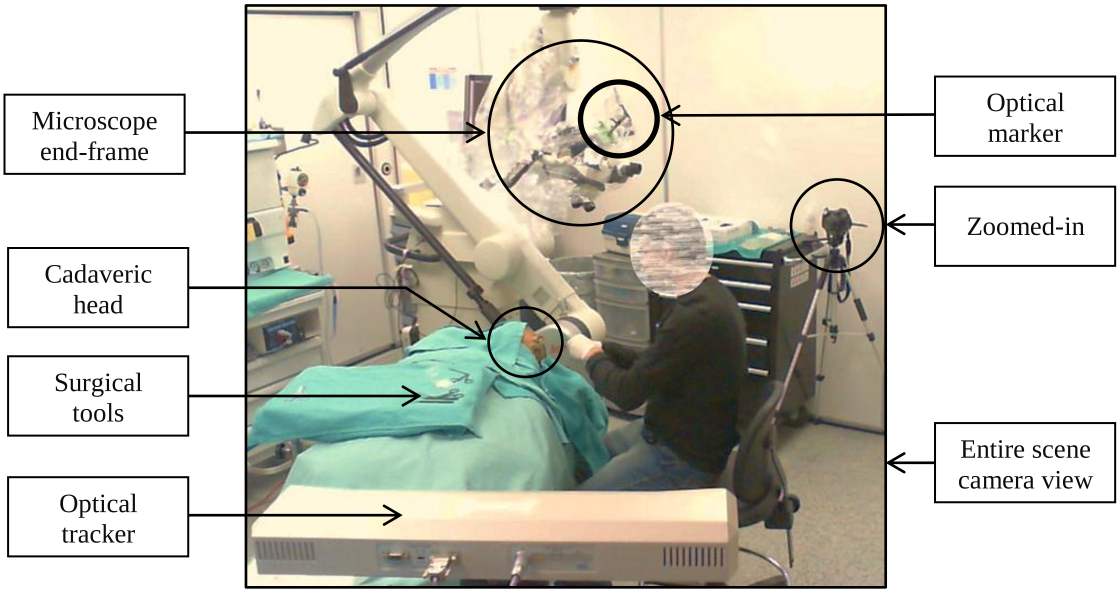

Figure 2.7: The final setup of the experiment scene with all the equipment labeled.

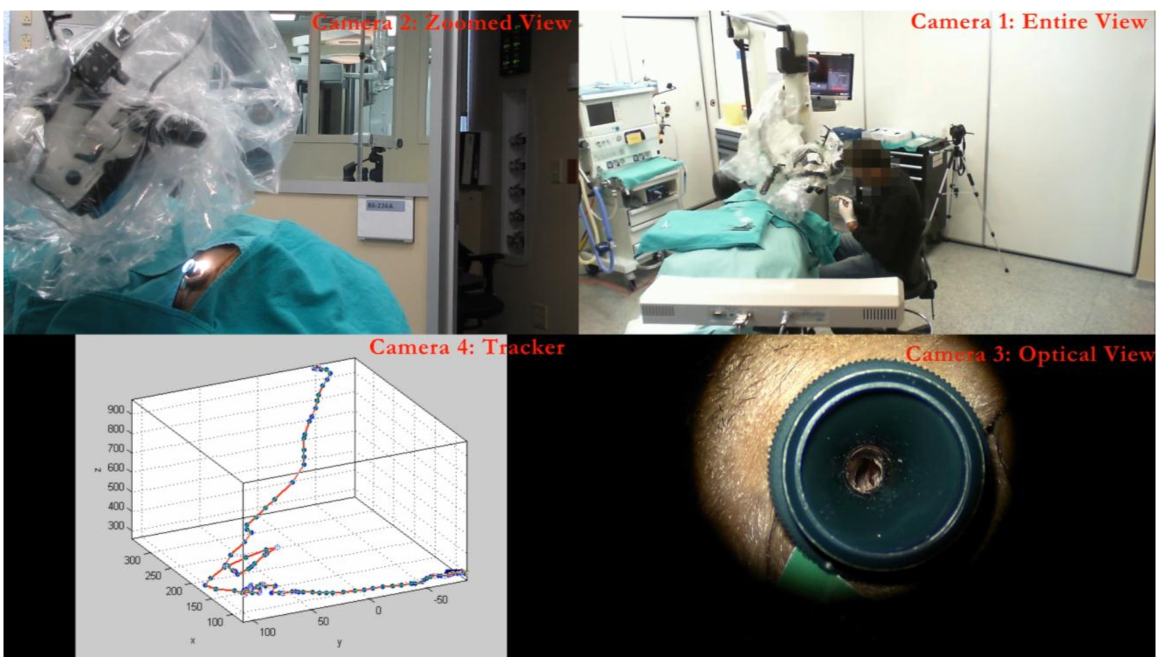

Figure 2.8: Time-synchronized split screen view to simultaneously visualize body and hand motion, tracking motion and surgical tool motion within optical FOV. All the camera views are labeled herein.

2.2.4 Metrics

The entire surgical task was comprehensively analysed using four categories of metrics. The *tracking metrics* were obtained by numerically analyzing the optical tracker's motion data. These tracking metrics were further grouped into motion time, path length, path smoothness, and path jitter.

The other three categories of metrics were tabulated by blinded experts (Neurotology and Pediatric Otolaryngology) from Western University and McGill University (n=3). These expert reviewers were separate from the four surgeons participating in the expert group of the study. The reviewers were each presented the anonymized videos in random order. They analysed the videos for the following categories: 1) *positioning metrics* (assessment of the operator's hand stabilization, arm position, and body position); 2) *optical metrics* (pertaining to optical FOV, focus, zoom, and obstruction); and 3) *procedural metrics* (such as hand jitter, instrument handling, and tube insertion).

2.2.5 Tracking Metrics

In order to analyse the manoeuvred path of the microscope's end-frame, the path was segmented into two parts, gross motion path and fine motion path. The segmented manoeuvred path is demonstrated in Figure 2.9. During each trial, the end-frame was manoeuvred from the common starting position and locked at some point to obtain an initial view of the surgical site. This continuous path, as shown in the figure, from the starting position to the initial stop position is the *gross motion path*. Due to default optical settings during each trial, each subject had to lock the end-frame at the end of the gross motion to adjust the optical parameters. Any manoeuvred path traced after the gross motion, in order to obtain the final view of the surgical site, is defined as the *fine motion path*. However, as shown in Figure 2.10, after initially locking the end-frame followed by a gross motion, some subjects may unlock, manoeuvre, and then lock the end-frame multiple times in order to attain the final view. Appearing as distinguishable standalone curve sequences in the given figure, each of these manoeuvres is considered as a separate fine motion path separated from each other by noticeable pause durations and the lock/unlock of the microscope. Therefore, all the derived tracking metrics were applied separately to the gross motion path, the fine motion path(s) and the total motion path (i.e., from starting position to final position) traced during each trial.

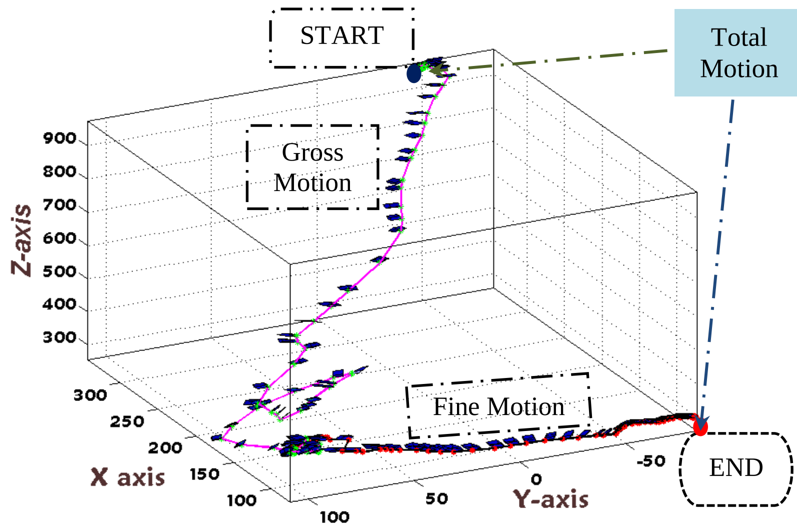

Figure 2.9: 3D motion path and 3D orientation of a virtual box simulating the end-frame. The motion path is segmented into gross motion path and fine motion path with the traced total motion path.

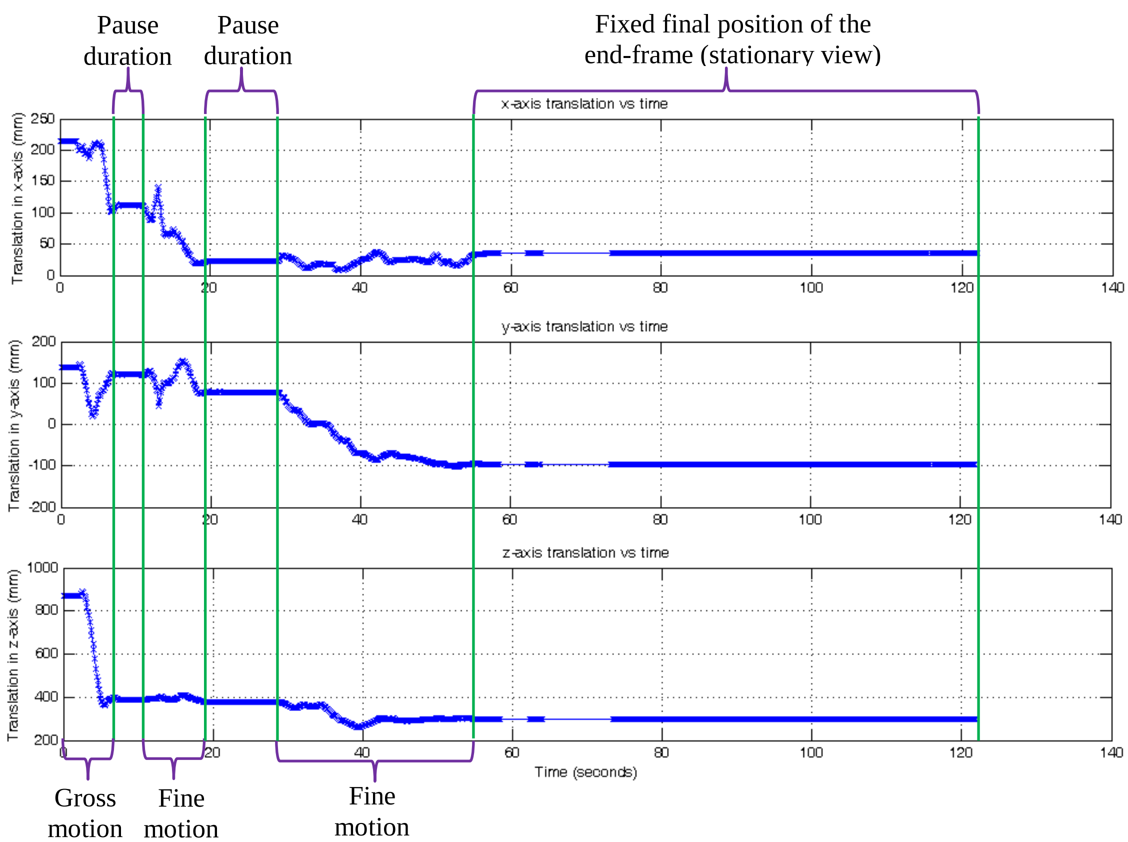

Figure 2.10: Traced motion path shown in segmented portions which are separated by pause durations of the end-frame.

The first derived tracking metric was *time* and it was computed for gross motion, fine motion(s) and total motion separately. For those subjects who traced multiple fine motion paths, the corresponding times were summed together to report the total fine motion time. In addition to motion specific times, the *total completion time* was counted from the instant the microscope's motion started until the instant the subject removed the forceps from the optical FOV after tube insertion. Next, *the total operation time* was counted from the instant any surgical tool was collected until the end of the tube insertion event. Likewise, *tube insertion time* was counted from the instant the tube was brought into the optical FOV of the microscope until it was completely inserted into the myringotomy. Finally, *still time* was the addition of all pause durations needed by the participating subject to prepare in any way to perform the operation after completing the gross motion. Such preparatory events included some or all of the following: optical parameters adjustments, surgical instrument fetching, and loading the ventilation tube onto the forceps. Each of the aforementioned time metrics, excluding the motion specific times, were measured by observing the subject's performance in the synchronized split screen video shown earlier. Generally, subjects with more experience were expected to require the least amount of time to perform any of the aforementioned tasks.

The next tracking metric was *path length* and again it was computed for gross motion, fine motion(s) and total motion separately. Since the motion path was traced in discrete 3D point sequence, the Euclidean Distance equation was used to compute the path length (please refer to Appendix 1 to view all equations). This equation cumulatively computes the actual length between consecutive discrete points for *N* data points. In addition, for all of gross, fine and total motion paths, the ideal path lengths were computed using the same Euclidean Distance equation but considering only the start and the end positions. Next, the efficiency measure of length was computed by dividing the actual path length by the ideal path length. The equation is provided in Appendix 1. This ratio essentially is the measure of length efficiency. If for a particular motion path (i.e., gross motion path), this actual to ideal ratio turns out to be a large value, it is indicative of a very inefficient motion. In other words, the higher the efficiency ratio, proportionally, the greater the actual manoeuvred path than the ideal path. Based on past publication on laparoscopic skill assessment [9], greater motion path traced by a surgical instrument is related to lesser experience. Intuitively, the most economic path, therefore, would be the one with the shortest path length from an initial to a final position complemented by a much smaller actual to ideal path length ratio.

Similarly, to assess the orientation of the end-frame, its *total rotations* about each axis (i.e., about x-axis is called *roll*, about y-axis is called *pitch* and about z-axis is called *yaw*) were separately computed for gross, fine and total motions. The implemented equation again is provided in Appendix 1. This equation cumulatively adds the absolute angular difference between consecutive data points about a single axis. Total rotations about other two axes were computed the same way as well. Again considering only the start and the end positions of the end-frame during gross, fine and total motions, the same equation was used to compute the ideal rotations about each axis. Rotational efficiency about each axis, therefore, was computed as the total rotation to ideal rotation ratio. The orientation of the microscope at the end of the gross and fine motion will vary by participant. Therefore, for any positional or rotational metric, the efficiency measures serve as the data to be compared.

Next to position and rotation, *manoeuvring volume* was derived to quantify the space taken to manoeuvre the end-frame from the initial position to the final position. As discussed in the same laparoscopic skill assessment study [9], compact manoeuvring volume covered by the surgical tools during an operation was considered to be more efficient and indicated greater experience. Therefore, it was computed by individually taking the difference between the maximum and the minimum x, y and z coordinate values recorded in the data series and multiplying these differences together to obtain the actual cubic volume. However, only considering the start position and the end position, the ideal volume was computed the same way in order to compute the normalized volume factor. The manoeuvring volume was also computed for gross motion and all fine motions separately. Figure 2.11 illustrates the quantification of this metric during gross motion as an example.

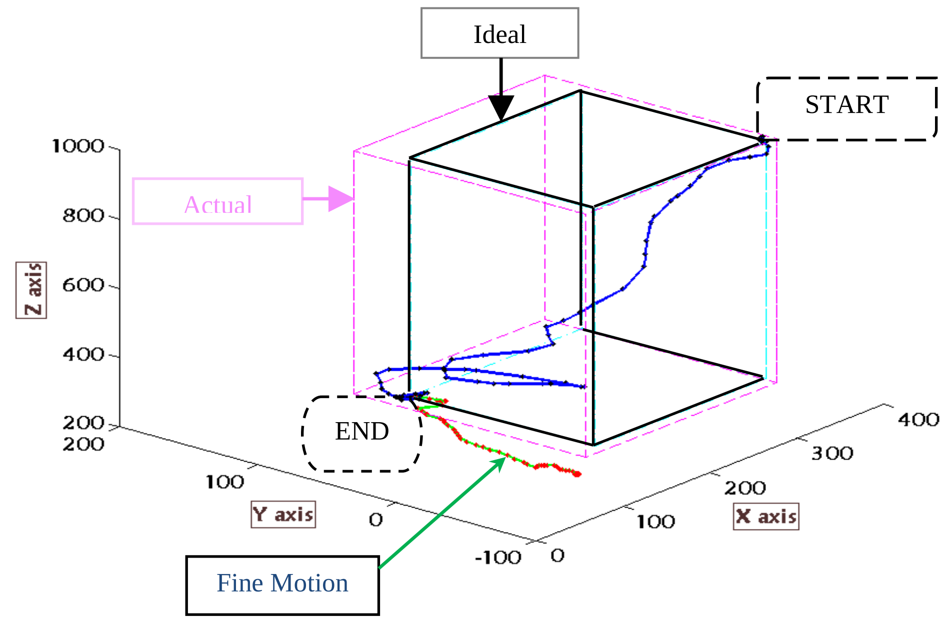

Figure 2.11: Manoeuvring volume computation. Shown here is the volume computation for gross motion

As mentioned earlier, after tracing a gross motion path, some subjects may trace multiple fine motion paths due to multiple sequential unlocking, manoeuvring and locking events of the end-frame. If the time difference between two such successive events is greater than at least two seconds, then each of those events are *re-adjustments* of the end-frame. Typically, subjects with more experience are likely to attain the final view with the least number of re-adjustments, whereas those with less experience are likely to require the most. To implement this metric, a simple algorithm was developed. It essentially scanned the tracked motion data points, recorded during a trial, to find the sets of successive data points with coordinate values that remained the same for more than two seconds. However, due to the RMS tracking error, the coordinate values of the discrete data points, recorded while the end-frame was stationary, fluctuated from each other by at most +0.25 mm to at least –0.25 mm. The algorithm was designed to neglect this fluctuation and mark such successive data points to be the same. The instant any of the coordinate values of a data point is detected to fluctuate from its preceding value beyond the allowable error range, it was counted as a re-adjustment by the algorithm.

The next characteristic considered was *path smoothness* evaluated using the metric normalized motion jerk. It is expected that experienced subjects will trace the smoothest paths while attempting to obtain the final view. To quantify path smoothness, its mathematical equation was adapted from some of the past clinical studies [25 – 28]. In this equation, squared motion jerk (i.e., time derivative of motion acceleration) is integrated over motion time. As the motion smoothness increases, the numerical result of this equation decreases. Since motion path length and time vary from subject to subject, this equation is normalized by a factor specific to a subject's motion path length and time. To complete normalization, the square root of the modified equation is then computed to produce a comparable dimensionless quantity of smoothness. Therefore, a larger value computed via this equation would mean the traced path was proportionally unsmooth, while a smaller value computed would mean a smoother motion path. While implementing the smoothness equation, at first the discretized motion path was differentiated over time to compute the motion velocity and differentiated again to compute the motion acceleration. When acceleration was differentiated again, the resulting unit scale (i.e., mm/s^3) of the signal further decreased while the noise increased demonstrating quantity of acceleration change in a motion path. Physically, change in acceleration means sudden unpredictable jerk when tracing a motion path, hence the name *motion jerk*. If a signal has frequent such jerks, it is indicative of unsmooth motion path and vice versa. (Appendix 1 lists all the aforementioned equations.)

The final metric considered was *motion jitter* which computed the total vibration in the traced motion path as a result of hand jerkiness while manoeuvring the end-frame. Jitter is related to smoothness, such that unsmooth path result from jittery motion. Therefore, to compute jitter, motion jerk had to be considered again. Figure 2.12 shows the plot of jerk amplitude versus time during a fine motion. Essentially, it is the illustration of motion vibration over the motion time. To only extract the valuable jerk signal from this graph, a simple filter was applied to the data. Shown in Appendix 1, the equation of this filter computes a threshold value and filters out everything below it as noise. The threshold value was taken to be 10% of the maximum jerk value computed for a particular motion path as shown in the figure. For the remaining non-zero jerk signal, the area underneath the curve was computed and normalized with respect to the motion path length and time. Again by normalizing the signal, jitter quantification was similarly standardized so that the dimensionless jitter quantity could be comparable from subject to subject. Based on the derivation of the jitter equation, a large value would mean more motion vibration while a smaller value would mean lesser motion vibration. It is arguable that over the same time frame, a jerk signal with significant amount of spikes many still have the same area underneath the curve as a signal with much less spikes. Though it may be true, technically, more spikes mean more vibration in a motion path which cumulatively adds to the motion path length. Therefore, when the equation is normalized by motion specific time and path length, it produces the corresponding jitter quantity. Microscope manoeuvrability was assessed numerically using all the derived tracking metrics that were successfully implemented in MATLAB.

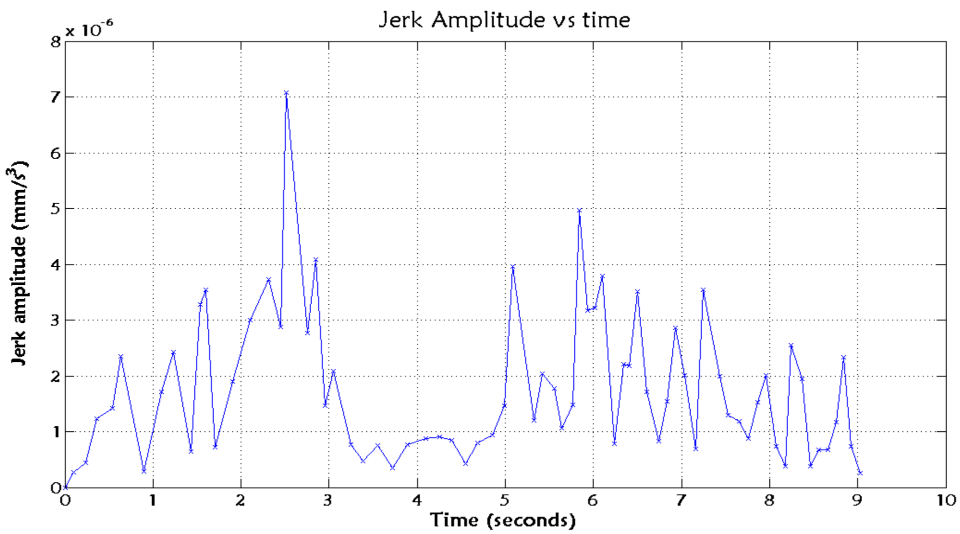

Figure 2.12: Plot of motion Jerk amplitude. It is a vector quantity as it consists of x, y and z components.

2.2.6 Positioning Metrics

A set of positioning metrics were derived to assess a subject's arms and body posture during the operation. A questionnaire was crafted composed of the metric name, evaluation objective, and 2-3 possible choices. In this category, the first metric considered was subject's *arm level* relative to the cadaveric head. Arms maintained fairly stationary and parallel to the cadaveric ear were perceived as being at the optimum level. If the elbows were raised significantly higher or lower than the cadaveric head, then hand jitter and reduced finger articulation were likely to occur due to disproportionate weight distribution on the operating hand and decreased flexibility of the hand muscles. Similarly, the subject's *wrist positioning* was the metric to evaluate stabilization of the wrist against the cadaveric head. The subject's *arm posture*, whether outstretched, flexed too close, or optimum, depended on the subject's *body to bed distance*.

2.2.7 Optical Metrics

Optical metrics were derived in order to assess the subject's quality of vision through the microscope optics. The quality of the optical view is directly dependent on the appropriate combination of the *optical zoom* (i.e., magnification) and the *optical focus*. Since many subjects wore prescription eyewear, they were asked to wear them before attempting to adjust the focus and the zoom. Having restored their vision to approximately 20/20 after wearing their eyewear, each subject was assumed to have the same initial vision through the microscope optics.

FOV obstruction was considered a critical metric as the optical view should not be blocked during the procedure. Complete obstruction of one eye can eliminate depth perception and increase the chances of injury to surrounding structure. Most commonly, this occurred from having the speculum at a poor angle such that the myringotomy could not be seen or from obstruction from the participant's hand during the procedure. Such an event is compared to an optimum optical view in Figure 2.13. Other optical metrics included whether the FOV was centered, intraocular distance adjustment, and intraocular tilt adjustment. Figure 2.14 clarifies the intraocular distance and tilt metrics.

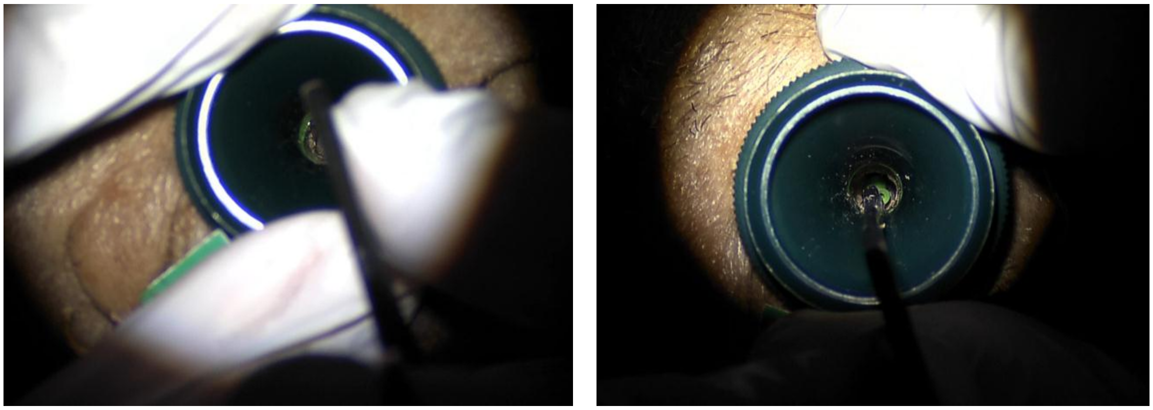

Figure 2.13: Figure on the left shows a bad optical view as the surgical site is completely blocked by the operating hand. Figure on the right shows a really good optical view as it entirely shows the tube being inserted without any obstruction.

Figure 2.14: Figure on the left shows the adjustment range of intraocular distance between the two lenses. Figure on the right shows the range of intraocular tilt adjustment. Courtesy of Leica Microsystems [29].

2.2.8 Procedural Metrics

Procedural metrics focussed on all aspects of the tube insertion task. *Microscope positioning* and *repositioning* considered to determine whether the final microscope position was optimal, too far, or too close to the cadaveric head. *Efficiency of instrument motion, instrument handling, speculum insertion, tube loading, tube insertion,* and *hand jitter* were all separate metrics that were considered and rated on a 2-3 point scale. Evaluation questionnaires for these metrics are provided in Appendix 3.

2.2.9 Statistical Analysis

The experimental project was conducted as a pilot study. Therefore, no existing data are available to compare and validate the present data. The maximum number of experts and residents were recruited from a single ENT residency program in order to better detect and appreciate the noticeable performance difference between the data obtained for these two groups. To compare residents with experts in terms of every tracking metric derived, independent sample t-tests were performed on the metric data produced by the participating subjects in their respected trials. These were the parametric tests performed assuming normal distribution of the data at each metric. In addition, non-parametric tests, specifically Kolmogorov Smirnov tests, were performed assuming non-normal distribution of the metric data at each metric. Later, parametric and non-parametric test results were compared to detect potential statistical trends in a single or multiple tracking metric gathered in this investigation.

Analyses of the optical, positioning, and procedural metrics were performed differently as these metrics were evaluated by a group of expert reviewers. Since the evaluations were done by observing the recorded performance videos, these assessments were personal judgments of the reviewers relative to the performance observed. In order to determine potential differences that existed in the metrics gathered in these categories, patterns of performance between experts and residents were carefully and comprehensively assessed. In this effort to detect performance differences, the appropriate method of preliminary analysis was directed toward assessment of inter-rater agreement. In the evaluation criteria, there were more than 2 raters and the metrics were evaluated using either a binary or a 3 point categorical scale. For this reason, Fleiss' kappa was determined to be the most appropriate statistical measure of inter-rater reliability. Since Fleiss' kappa works for any number of raters who provide categorical ratings for a fixed number of items, this reliability test was conducted independently on each of the aforementioned metrics. The fixed number of items was the total 12 subjects consisting of 4 expert surgeons and 8 residents. Fleiss' kappa (κ) was computed separately for the experts and the residents for each metric. Upon its computation, the outcome was compared to its given significance scale [30] to determine which metrics could relevantly distinguish operational performance of experts from residents.

2.3 Results

2.3.1 Demographics

A total of 12 subjects (8 males, 4 females) participated in the study including both residents (n = 8) and experts (n = 4). The resident group included individuals from ENT postgraduate years (PGY) 1 (n = 3), 2 (n = 3) and 3 (n = 2). The expert group included practicing ENT surgeons, most of whom (n =3) had >6 years of surgical experience in their specialty. However, one expert had just over one year of such experience after completing a post-residency fellowship training program. Regardless of their varying years of practice, all experts had performed at least 500 myringotomy cases. Since all residents were in their learning phase, only they were considered for the assessment. Based on the assessment, PGY 1 residents (n = 3) had no prior ENT surgical experience before participating in the experiment, while the rest combining PGY 2 and PGY 3 residents (n = 5) reported having some experience performing a myrigotomy. This was reflective of the PGY 1 residents having performed no myringotomy case to date, whereas PGY 2 residents indicated they had performed about 10 to 20 cases and PGY 3 residents performed more than 50 cases to date. In terms of time spent in the Neurotology/Paediatric rotation, none of the PGY 1 residents had taken part in it yet; however, PGY 2 residents had spent 3 to 5 months in it, while PGY 3 residents had spent more than 8 months in this rotation. During their time in the residency program, PGY 1 residents reported that they had used a surgical microscope from zero to a maximum of 2 times. On the contrary, PGY 2 residents estimated their use of the microscope to range between 50 to 150 times; PGY 3 residents estimated their use at more than 200 times. In general, a majority of the residents (n = 5) had previous experience working with some sort of microscope compared to the remaining individuals (n = 3) who never used a microscope before. In regards to handedness, almost all residents (n = 7), from a sample size of 8, were right-handed. Furthermore, a majority of them (n = 4) also indicated that they had no preference as to which ear (i.e., left or right) was easier to perform the myringotomy. However, some (n = 3) preferred right ear compared to the one individual who preferred left ear. Finally, information on the residents' extraneous manual skills that may have facilitated their capabilities with the surgical microscope were also collected. This included information on exposure to playing video games and expertise with playing a musical instrument. This information was gathered in order to discover if such exposure helped with development of microscope manoeuvrability and finger control. Based on resident judgments obtained for each of these extraneous activities with a rating of "1" being the least skilled to "10" being the most skilled in such practices, half the residents rated themselves at 6 or greater in video gaming and at 5 or greater specific to playing an instrument.

2.3.2 Assessment of Microscope Manoeuvrability

As outlined earlier, except for the time metrics, all other metric results were normalized in terms of ratio of actual value to ideal value. To graphically appreciate experts' and residents' overall raw data, each group's maximum, minimum, mean, median, standard deviation and variance were computed across all the metrics. Our rationale for calculating these measures of central tendency was based on our desire to avoid misinterpretation of the range of performance that could occur if using only a measure such as the mean. More specifically, because of the small sample sizes studied in this experiment, the potential for an error in accurately representing the collective data of any given group is increased considerably. Mean values obtained from small samples can be greatly influenced by extreme scores. Thus, additional measures of central tendency were generated for comparative purposes.

In calculating and applying these newly computed values, two box plots (i.e., one for residents and one for experts) were produced for each metric. A typical graph is shown in Figure 2.15. In this graph, for each time metric, the right plot demonstrates residents' data and the left plot demonstrates experts' data. Each plot has a maximum and a minimum mark illustrating the actual data range. Typically, a large range would mean high data variability, whereas a small range would mean the opposite. Keep in mind that if in fact a distribution of any dataset is normal, then the mean, median, and mode will be identical; as data become skewed to either the negative or positive side of any given distribution, then these values will differ. The box appearing midway within the data range is reflective of the difference between mean and median. As a result, the size of this difference determines the height of the box. When this box appears in the middle of the data range with low to no height, it reflects normal Gaussian distribution of the data where mean and median approximate each other. In addition, the location of the box in the data range indicates where a majority of the data are concentrated. In the figure, such a scenario can be seen in the plots of *total completion time* for experts, *total still time* for experts and *fine motion time* for residents. If, however, the box appears outside of the midpoint within the depicted range, it demonstrates a non-normal distribution of data. In the figure, the plots for *total completion time* and *total still time* for residents show this pattern. For instance, in the total completion time plot for residents, the box is localized near the maximum mark. This finding indicates that a majority of the residents required fairly long total completion times. It is also apparent in the figure across all the time metrics, that the mean and median values of the experts are comparably much lower than that of the residents. Due to their experience, experts would generally be expected to require less time to perform any surgical task, hence, the lower mean and median values shown. When t-tests were performed on all the metrics, $p \leq 0.05$ was only considered to be significant while a probability level $0.05 < p < 0.1$ was considered as a relative index of metrics with approaching significance [31]. T-tests performed on all the time metrics revealed that only total completion time ($p = 0.009$) and total fine motion time ($p = 0.022$) demonstrated statistically significant difference between experts and residents. However, total still time demonstrated approaching statistical significance ($p = 0.075$) when comparing experts with residents. Similarly, approaching statistical significance was observed in normalized manoeuvring volume during gross motion ($p = 0.056$), in path smoothness during total motion ($p = 0.054$) and fine motion ($p = 0.088$) and lastly in jitter during total motion ($p = 0.079$) and fine motion ($p = 0.075$). Finally, statistical significance was observed in reposition frequency metric ($p = 0.034$). Corresponding box plots of all of these metrics are illustrated in Figure 2.16.

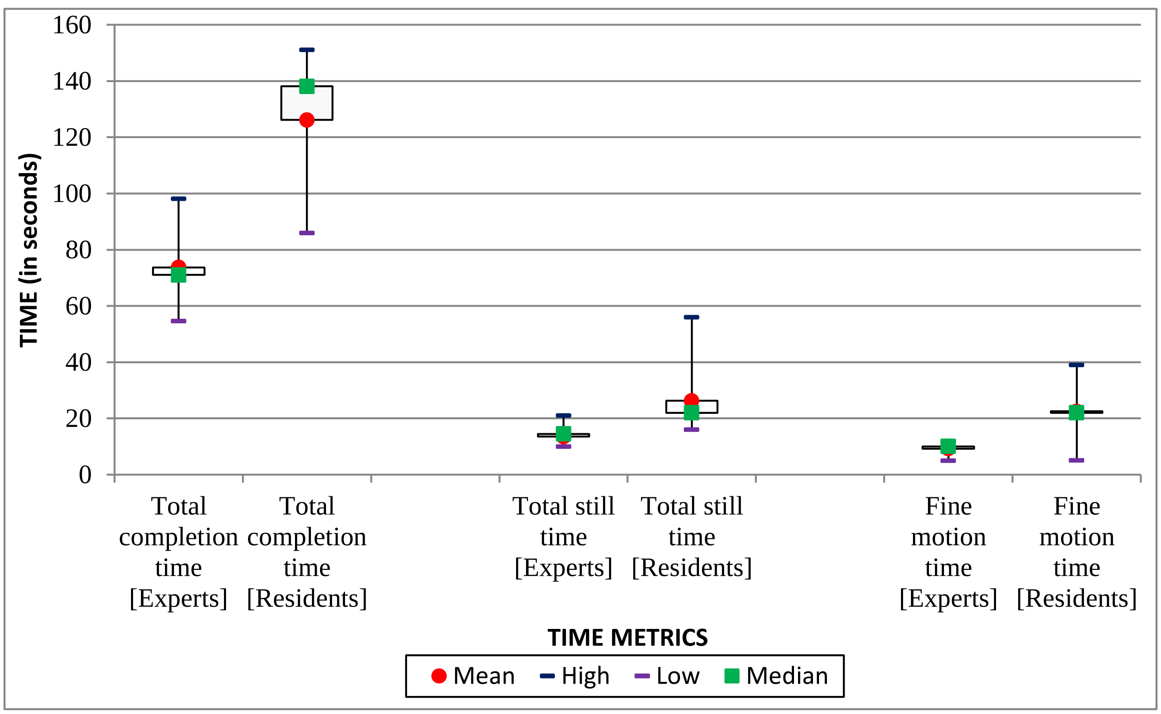

Figure 2.15: Time metrics that came out to be significant and those that showed approaching significance.

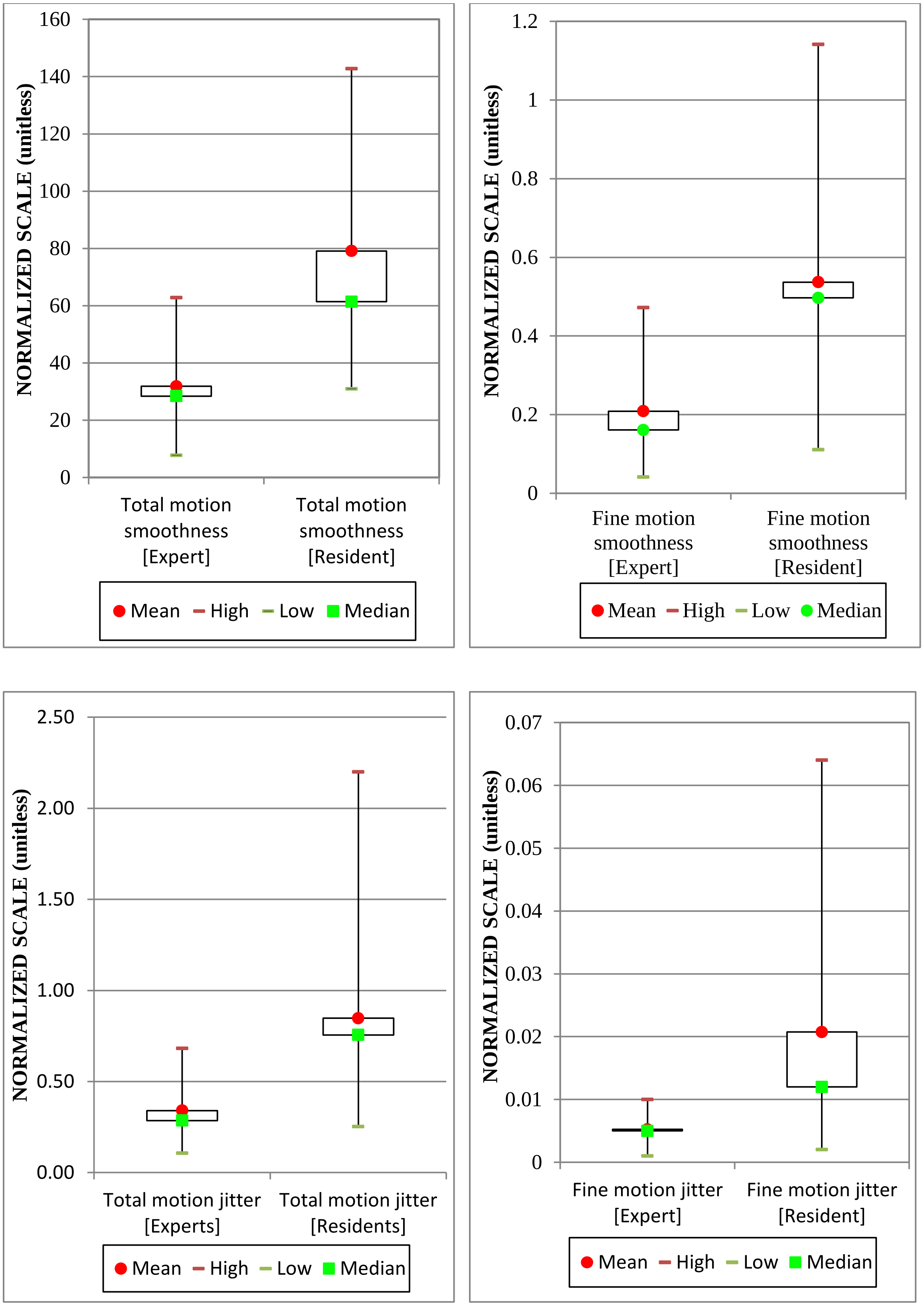

Figure 2.16 (a): All of the above graphs show the performance difference between experts and residents at metrics with approaching significance. Lower mean/median means better performance. (Continued).

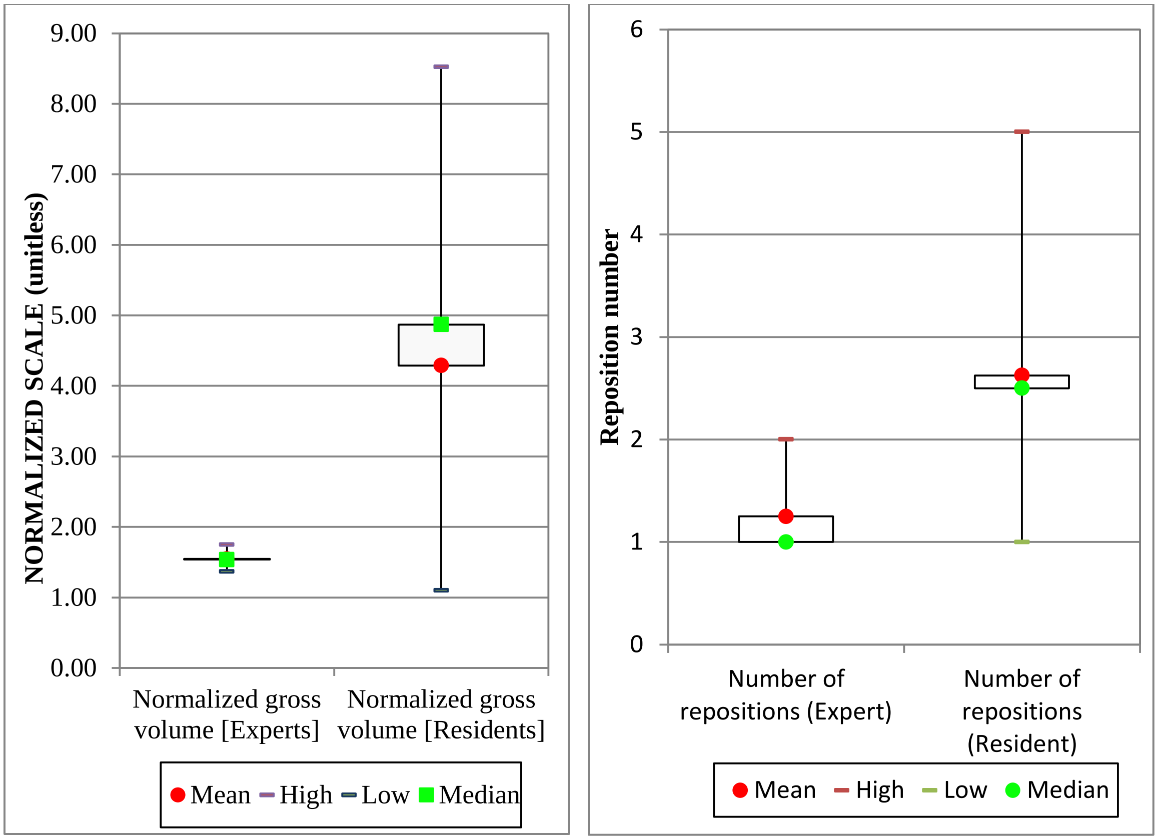

Figure 2.17 (b): All of the above graphs show the performance difference between experts and residents at metrics with approaching significance and significant difference respectively.

In addition to parametric t-tests, non-parametric Kolmogorov Smirnov tests also were performed to determine if any metrics differed across experts and residents. The purpose of these additional tests was to determine the level of consistency between both parametric and non-parametric tests on the same metric(s). If a difference was noted on both tests for a given metric, this finding would add some strength to its value as a potentially discriminating metric(s) capable of differentiating skill levels. Assuming non-normal distribution, these tests revealed four tracking metrics that indicated a significant difference between the groups. These metrics were total completion time ($p = 0.01$), total still time ($p = 0.024$), normalized total rotation in roll direction during fine motion ($p = 0.029$) and normalized total rotation in yaw direction during fine motion ($p = 0.051$). Furthermore, these non-parametric tests were performed whenever a non-normal distribution was observed at a particular metric. All the aforementioned results are summarized in Table 1.

Significant Metrics	Metrics with approaching significance
Total completion time ($p = 0.009$) [Parametric result]	Total still time ($p = 0.075$) [Parametric result]
Total fine motion time ($p = 0.022$) [Parametric result]	Normalized manoeuvring volume during gross motion ($p = 0.056$) [Parametric result]
Reposition frequency metric ($p = 0.034$) [Parametric result]	Path smoothness during total motion ($p = 0.054$) [Parametric result]
Normalized total rotation in roll direction during fine motion ($p = 0.029$) [Non-parametric result]	Path smoothness during fine motion ($p = 0.088$) [Parametric result]
Normalized total rotation in yaw direction during fine motion ($p = 0.051$) [Non-parametric result]	Jitter during total motion ($p = 0.079$) [Parametric result]
	Jitter during fine motion ($p = 0.075$) [Parametric result]

Table 2.1: Summary of significant statistical results obtained by performing statistical tests on tracking metrics data.

2.3.3 Assessment of Operational Metrics

Comprised of optical, positioning, and procedural metrics, Fleiss' kappa based inter-rater agreement results of all the operational metrics are listed in Table 2. To compute the kappa, data from each metric was treated with its formula manually. The formula is provided in Appendix 1. Under optical metrics, intraocular distance metric showed complete agreement for both experts ($\kappa_E = 1.000$) and residents ($\kappa_R = 1.000$). On the contrary, intraocular tilt metric showed complete disagreement for both subjects ($\kappa_E < 0.000$; $\kappa_R < 0.000$). In terms of focus, again there was a complete agreement for experts and substantial agreement for residents ($\kappa_R = 0.644$). For zoom, however, the experts demonstrated lesser agreement ($\kappa_E = 0.172$) compared to that of the residents ($\kappa_R =$

0.281). In regards to optical FOV metrics, there was a moderate agreement on experts (κ_E = 0.494) that they had unobstructed FOV compared to the residents (κ_R = 0.115). Finally, the degree of centered FOV metric demonstrated substantial agreement on the experts (κ_E = 0.625), while it was comparably lower for the residents (κ_R = 0.301). Each of the reviewers independently classified subjects as “expert” or “resident” at the end of their optical metric evaluation. Based on a simple percentage calculation, it was determined that there was an 88.8% inter-rater agreement across the judges when categorizing the subjects collectively based on the optical metrics. When confirmed by the kappa result, it was found that there was a substantial agreement among the reviewers when identifying the experts as ‘Experts’ (κ_E = 0.625). However, when identifying the residents as ‘Experts’ the agreement was much lower (κ_R = 0.234) indicating that most residents were not identified as experts. Careful observation of the raw data revealed that this unexpected agreement was due to misidentification of 2 of the 8 residents as experts by one reviewer. Similarly, the imperfect kappa value for expert identification was due to misidentification of one of the 4 experts as being a resident by all reviewers.

Finally, all positioning metrics demonstrated consistently perfect agreement on the optimum behaviours of the experts (κ_E = 1.000) and consistently low-to-no agreement for that of the residents (see Table 2). However, one metric, that of a subject’s wrist positioning, showed complete disagreement across the reviewers for both experts and residents. Similar to the positioning metrics, a majority of the procedural metrics demonstrated perfect agreement on the optimum behaviours of experts. Those metrics that did not have perfect agreement were tube loading (κ_E = 0.625; κ_R < 0.000), tube insertion (κ_E = 0.400; κ_R < 0.000) and hand jitter (κ_E = 0.172; κ_R = 0.066). Based on the collective evaluations of these metrics, another simple percentage calculation revealed that there was a 94.4% agreement across the reviewers when categorizing subjects as either an expert or a resident. However, agreement on identifying the residents as experts (κ_R = 0.454) was slightly higher as well. This anomaly may have been due to increased sample size and misidentification by 2 reviewers of one of the 8 residents as being an expert. All of the aforementioned findings are summarized in Table 2.

Optical Metrics	Fleiss' kappa (κ_E for experts; κ_R for residents)	
Intraocular distance adjustment	$\kappa_E = 1.000$	$\kappa_R = 1.000$
Intraocular tilt	$\kappa_E < 0.000$	$\kappa_R < 0.000$
Focus	$\kappa_E = 1.000$	$\kappa_R = 0.644$
Zoom	$\kappa_E = 0.172$	$\kappa_R = 0.281$
Unobstructed FOV	$\kappa_E = 0.494$	$\kappa_R = 0.115$
Optimally centred FOV	$\kappa_E = 0.625$	$\kappa_R = 0.301$
Decide if the subject is an expert	$\kappa_E = 0.625$	$\kappa_R = 0.234$
Positioning Metrics		
Subject's optimum arm level	$\kappa_E = 1.000$	$\kappa_R = 0.077$
Subject's optimum wrist position	$\kappa_E < 0.000$	$\kappa_R < 0.000$
Subject's optimum posture	$\kappa_E = 1.000$	$\kappa_R = 0.303$
Subject's optimum body to bed positioning	$\kappa_E = 1.000$	$\kappa_R < 0.000$
Subject's optimum arm posture	$\kappa_E = 1.000$	$\kappa_R < 0.000$
Procedural Metrics		
Microscope's minimum repositioning	$\kappa_E = 1.000$	$\kappa_R = 0.625$
Microscope's proper positioning	$\kappa_E = 1.000$	$\kappa_R = 0.059$
Speculum insertion	$\kappa_E = 1.000$	$\kappa_R = 0.100$
Instrument motion efficiency (unnecessary motion present?)	$\kappa_E = 1.000$	$\kappa_R = 0.251$
Fluid instrument handling	$\kappa_E = 1.000$	$\kappa_R = 0.063$
Accurate tube loading	$\kappa_E = 0.625$	$\kappa_R < 0.000$
Appropriate tube insertion	$\kappa_E = 0.400$	$\kappa_R < 0.000$
No hand jitters	$\kappa_E = 0.172$	$\kappa_R = 0.066$
Decide if the subject is an expert	$\kappa_E = 1.000$	$\kappa_R = 0.454$

Table 2. 2: List of the computed kappa values for subjective inter-rater agreements. Here, $\kappa < 0$ means no agreement; $0 < \kappa < 0.4$ is assumed to be slim to low agreement; $0.4 < \kappa < 0.6$ is assumed to be of moderate agreement; $0.6 < \kappa < 1.0$ is assumed to be substantial agreement; and $\kappa = 1.0$ is known to be at perfect agreement among the raters.

2.4 Discussion

2.4.1 Interpretation of Tracking Metric analyses

Although sample sizes in both participating groups were small and data obtained must be considered relative to their external validity, several findings of value within the context of this project did emerge. First, several statistically significant differences were observed between expert and resident surgical microscope users for some tracking metrics. These differences may potentially be indicative of discriminatory tracking metrics that distinguish experts from novice residents (see Table 1). Because of our concerns related to the small sample size for both groups and the inherent concerns of variability, both parametric and non-parametric tests were used as a preliminary index of these metrics. The *total completion time* metric represents the expected difference between the two groups in that expert surgeons will always be assumed to need less time to complete a given surgical task. It is a perfectly valid judgement based on their years of practice in the real world, and this assumption was confirmed to some extent via statistical analysis. However, the total time metric alone gives little to no information as to what factors contribute to this time difference. Therefore, statistics of other time metrics may help uncover these underlying factors. Based on our data, the time metric reveals that residents spend more time preparing for their performance of a given surgical task. In contrast, experts appear to know the exact sequence of what needs to be done including getting the microscope into position, adjusting the optics, and handling the instruments in order to perform the surgical task. As a result, the time experts require is significantly less than that of residents. Based on the non-parametric test, *still time* is the second significantly different time metric between groups. Its significance is also supported by its corresponding t-test outcome. Therefore, it is not inappropriate to suggest that this time difference is an important factor that quantifies surgical expertise for myringotomy. Still time is the temporal sum of all pauses existing within the operation time, tube insertion time, and fine motion time. However, due to the variability of these fragmented still times, all other time metrics computed for the residents showed large degrees of variability and consequently, resulted in non-significant differences. However, it is important to point out that a non-significant difference does not suggest that other metrics are the same; they just did not meet the critical difference for the statistic used. In this regard, the limitations of a small sample can influence statistical measures in both ways (i.e., inadvertent identification of significance and non-significance). Although the operation time metric (p = 0.15) and tube insertion time metric (p = 0.78) were not found to be significantly different between the two groups, the data gathered clearly outline the better time-wise performance of the experts (see Figure 2.17). Since experts know in what sequence a task must be performed, they need lesser still time to prepare for an operational task; therefore, requiring lesser time during the entire operation and during tube insertion. Finally, significant differences in fine motion time metric outlines one's fine ability to lock in on the final optical view of the surgical site prior to moving to the next sequential step in the procedure. Obtaining an initial view of the surgical site following a gross motion is a basic task, and as such, both groups needed nearly the same time on average to meet this requirement (Figure 2.17). As a result, the non-significant outcome for the gross motion time metric (p = 0.86) is understandable. On the other hand, attaining an unobstructed, focussed, and centred optical view in a short period of time is a certain indicator of a highly skilled performance, hence, our finding of the significant outcome in the fine motion time metric.

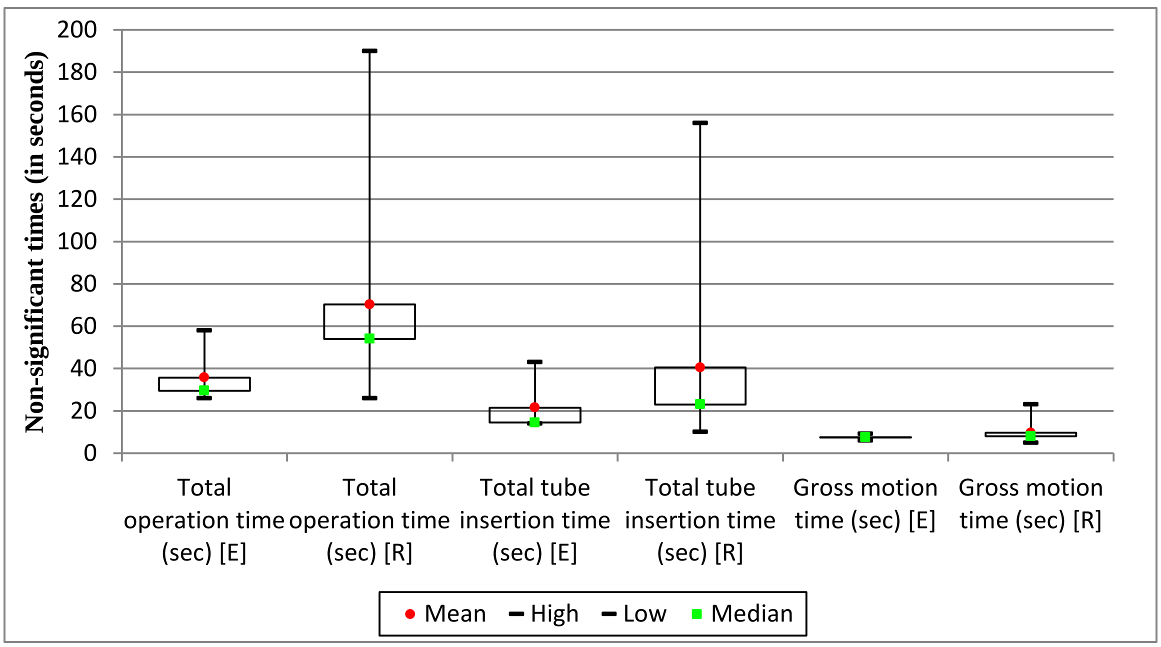

Figure 2.17: All the non-significant time metrics found through t-tests. Group differences are depicted.

Stated earlier in this Section, non-parametric tests revealed significant differences for the normalized *total rotation in roll* and *yaw* directions during fine motion. However, the actual validity of these results remains questionable since corresponding parametric t-tests revealed non-significant differences for the fine motion roll metric ($p = 0.32$) and the fine motion yaw metric ($p = 0.82$). To investigate this anomaly, the corresponding raw data plots (Figure 2.18) were carefully examined. It was found that with exception of the normalized fine motion yaw rotation for experts, all other fine motion rotation metrics had non-normal distributions. Consequently, performing parametric tests and assuming a normal distribution is problematic. But, regardless of this observed disagreement between parametric and non-parametric tests, experts still demonstrated lower yaw and roll averages than that of the residents. Though this observation lacks strong statistical support, it still provides a potentially valid metric for distinguishing groups based on the non-parametric test outcomes. The lower mean and median values of the rotation metrics recorded by the experts indicate that they do not experiment on site to figure out what is the best angular orientation of the end-frame. It is, therefore, highly likely that they intuitively know what would be the end-frame's best orientation given the fixed position and orientation of the cadaveric ear. On the other hand, residents are likely to keep on rotating the end-frame in different angular directions trying to find the best orientation, which cumulatively results in higher rotational average. Collectively, these findings do suggest that several metrics may prove to be of value relative to optimizing surgical training and monitoring the acquisition of such skill for the young surgeon.

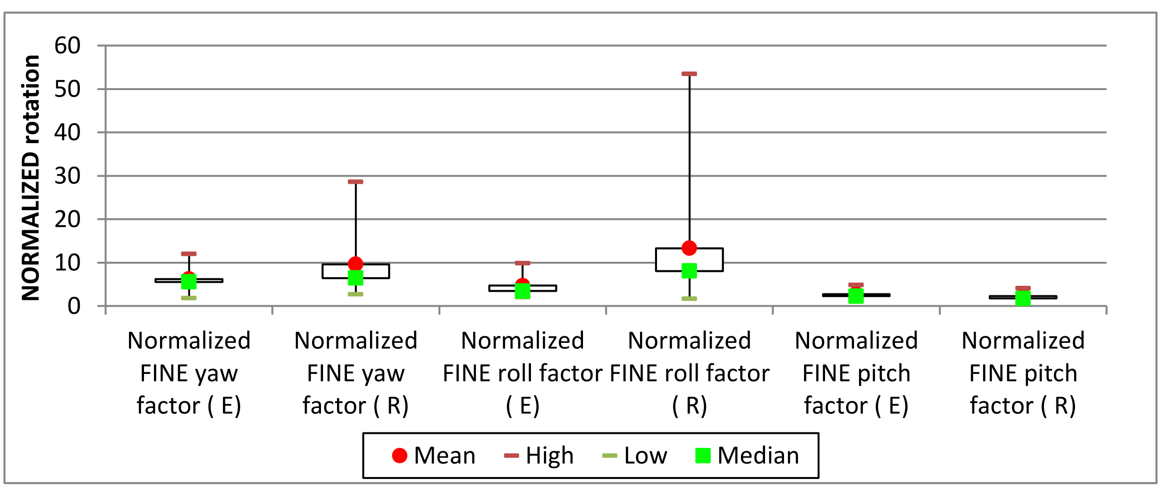

Figure 2.18: Plot of all rotation metric results. Non-parametric tests revealed roll and yaw as significant.

When examining the box plots (Figure 2.16) of the remaining tracking metrics where t-test findings approached statistical significance (i.e., gross motion volume, total and fine motion smoothness, and total and fine motion jitters), it was observed for each of these metrics, both experts' and residents' data appeared to be distributed fairly normally (although it is acknowledged that the sample from which the distribution is inferred is small). However, due to the smaller sample sizes of the participating groups and existence of variability, the outcomes were not exactly below *a priori* probability level of 0.05, but rather between $p < 0.05$ and 0.1. Taking these factors into consideration, interpretation of the plots can be made with reasonable confidence. From visual perspective, all of these metric plots consistently illustrate the predominant nature of the experts' performance through their lower mean and median scores. The low and coinciding mean and median scores of the experts' normalized gross motion volume, with very low variability, indicate that these experts have equally mastered the art of manoeuvring the end-frame to an initial position within a very compact physical space. Comparably, the residents need a much larger space to achieve the same goal, as well as demonstrating greater variability in their scores is simply due to their varying experience with surgical microscopes. Similarly, for total and fine motion smoothness, the low scores of the experts indicate their more refined ability to trace smoother path during the entire manoeuvring event and during locking in on the final optical view. Comparably, the residents performed poorly here as well due to their relative lack of experience with surgical microscope manipulation.

As mentioned earlier, unsmooth paths are traced through jittery motion and jitter is a metric of motion vibration. From a logical point of view, being able to control and minimize this vibration while manoeuvring the microscope is the indication of proficient manoeuvring skills. Therefore, examining the fine motion jitter metric plot, it can be seen that the experts again have very low variability and coinciding low mean and median scores. This observation supports the claim that they are proficient enough to minimize jitter during motion; this ability is fairly consistent across all the experts. On the other hand, residents were seen to have higher variability and comparably higher mean and median scores than that of the experts, a finding that was indicative of their inability to minimize jitter via fine movement. Finally, the plot of reposition frequency metric illustrates the last significant difference between experts and residents. Experts repositioned once or twice at most to land the end-frame in the final position, while the residents repositioned multiple times. This finding once again ties into the same concept of experts intuitively knowing exactly where to position the end-frame as opposed to the residents who ultimately appeared to find that location through trial and error.

Each of the aforementioned metrics shows a unique performance characteristic that may outline distinguishing differences between expert and resident surgeons while performing a myringotomy. These metrics are able to show in consistent manner numerically comparable, distinguishable, and identifiable inherent performance behaviour of experts against residents and vice versa. Therefore, these metrics could potentially serve as rating parameters or an index of performance in a surgical training simulator. However, further validation is required with greater sample sizes for absolute confirmation of these interpretations.

2.4.2 Interpretation of Operational Metrics analyses

The purpose of optical, positioning, and procedural metrics was used in the present study to evaluate each and every operational task performed in the experimental myringotomy in order to identify patterns of consistent performance. As these evaluations were analyzed via Fleiss' kappa (κ), the magnitude of kappa was interpreted in accordance with the significance scale presented by Landis and Koch [30]. Landis and Koch agree that no scale for any kappa coefficient is universal. In fact, kappa magnitude changes when the rating categories increase or decrease for the same numbers of subjects and evaluators. This case is also true when the sample size or the number of evaluators changes, thus keeping the rating categories constant. In this study, there were both binary and ternary categories of rating (i.e., also termed earlier as 2 to 3 point scale), which were considered independently for each metric. In other words, no single metric had both binary and ternary scales. In addition, a fixed number of evaluators ($n = 3$) always evaluated fixed numbers of both experts and residents. As a result, this provided reliability in the kappa values calculated and confidence when interpreting them.

In interpretation of the optical metrics' kappa results, the perfect agreement ($\kappa = 1.0$) on both experts and residents for the intraocular distance metric, indicates that all subjects adjusted this parameter. However, its entirely homogeneous outcome suggests that this metric does not behave as a performance distinguishing feature. Similarly, the complete disagreement ($\kappa < 0$) on intraocular tilt metric indicates that none of the subjects cared much to adjust this parameter at all. Again due to homogeneity of the subjects based on the kappa outcome obtained for experts and residents, this metric can be discarded off as unimportant and incapable of differentiating skilled performance. While focus metric does show perfect agreement on the experts, substantial agreement on the residents ($\kappa_R = 0.644$) was noticed as well. This comparably lower score of the kappa for residents indicates that most residents are able to acquire a focussed vision. As a result, it may not be a suitable skill differentiating metric. As for zoom metric, the kappa results ($\kappa_E = 0.172$, $\kappa_R = 0.281$) do not show any considerable polarity toward experts or toward residents. Since interpretation of an optimum zoom is highly variable from person to person, the low agreement among the reviewers is justifiable. Therefore, zoom may not be a critical metric within the context of the present study. Based on the calculated kappa outcomes, the only optical metrics showing moderate agreement differences were unobstructed FOV ($\kappa_E = 0.494$, $\kappa_R = 0.115$) and centered FOV ($\kappa_E = 0.625$, $\kappa_R = 0.301$). The power of these two metrics can be appreciated when the appropriateness of an optical FOV is questioned. If an optical FOV is partially or completely blocked, then it is certain that the surgeon cannot entirely see the myringotomy (i.e., the surgical site). Similarly, if the optical FOV is not centered, then it is likely that the myringotomy is at or toward the edge of the optical FOV, which will provide very little to constricted visibility during operation. Therefore, the worst ratings of both of these metrics are an indication of potential procedural hazards. Higher agreements upon the experts' abilities to obtain optimally centered and unobstructed FOV, compared to that of the residents, therefore, outlines their high proficiency level.

In contrast to the above information, all positioning metrics dominantly showed significant agreement on better performance by the experts. However, the only exception was wrist positioning. A better performance was signified by optimum arm level, optimum body and arm postures, and optimum body-to-bed distance. Clearly, these three behaviours may represent a composite physical behaviour associated with higher skill levels specific to the present study. Comparably, the agreement on such performance was little or none for the residents. Since there were such substantial differences in agreements between the two participating groups for all the positioning metrics, these features may act as the performance differentiating metrics. Though wrist positioning may be important, the complete disagreements pertaining to it is justifiable based on its raw data. As all subjects were wearing hand gloves while some were wearing long sleeve coats throughout the operation, much of their wrists were covered. That is why wrist positioning could not be determined properly. As for the procedural metrics, significant agreement was observed across all of them for experts, while for residents the agreements were substantially low. This phenomenon again supports the claim that procedural metrics (Table 2), may serve to differentiate an expert's performance. However, among all procedural metrics, minimum repositioning metric showed substantial agreement for the residents (κ_R = 0.625) as well. This high agreement suggests that majority of the residents were able to localize the final position with as few repositions as possible. Therefore, due to the high agreement on both experts and residents at this metric, it may not largely differentiate experts from residents. However, the study should be conducted again with more participants to statistically verify this claim.

Although all or at least a majority of the subjects were identified accurately as experts or residents, one expert was consistently misidentified by all 3 reviewers based on the three categories of operational metrics. To understand this anomaly, the background of the subject was investigated and it was found that the subject was a much younger surgeon with the least amount of practicing experience. Therefore, investigation of this subject's optical metrics data revealed that this subject always had a partially obstructed FOV and it was not centered. These were the most common errors made by the participating residents as well. Even though optimum performance was observed based on all the positioning metrics, this subject did not have instrument handling fluidity, was not able to load the tube onto the crocodile forceps same as the other experts, was not able to perform the tube insertion same as the other experts, and had significant hand jitter when inserting the tube. These errors were also seen across many participating residents while none of the other experts committed any of these errors. Therefore, these shortcomings of this individual could potentially be the toughest skill set that a resident needs to master in order to be fully classified as an expert. Had there been more such subjects, the preceding claim could have been supported or refuted empirically. Hence, recruitment of more participants is necessary for any follow-up study that utilizes the metrics described herein.

2.5 Conclusions

Metric based assessment of microscope manoeuvrability during myringotomy objectively determined how a surgical microscope is used by experts and residents during a myringotomy procedure. It is certain that experts have greater control, have better understanding of an optimum microscopic view, and are more proficient in surgical instrument handling. It is also certain that eventually residents will attain these skills. However, the issue lies within the transformation phase of when these skills are mastered. The duration of this phase is uncertainly variable among the novice residents due to their varying capacity of learning new skills. Without having a structured teaching method comprised of optimum manoeuvring techniques and objective evaluation strategy, residents are left on their own to discover what optical settings, procedural practices and manoeuvring techniques work for them through numerous trial and errors. As with any motor task, skill sets are acquired through direct practice and practice performance can be evaluated. Therefore, having determined and validated sets of potentially discriminatory metrics, which are numerically quantifiable, may help with teaching microscope manoeuvring in a controlled fashion and assess one's manoeuvrability objectively. In addition, upon successful integration of a real enough microscope simulator with the existing myringotomy simulator, the ultimate goal is to incorporate these discriminatory metrics into the integrated simulator to train young surgeons. This will enable ENT residents in training to obtain automated feedback on their microscope manoeuvring performance and track their improvement over time, both short- and long-term. Furthermore, upon successful validation of this study through evaluation of an increased number of subjects recruited by running multi-center trials, the present methodology can potentially be applied in other ENT procedures such as Microlaryngoscopy, Tympanoplasty or Mastoidectomy that require extensive use of surgical microscopes.

2.6 References

[1] W. J. McIsaac, P.C. Coyte, R. Croxford et al., "Otolaryngologists' perceptions of the indications for tympanostomy tube insertion in children," *CMAJ*, vol. 162, pp. 1285-1288. 2000.

[2] M. L. Montague, M. S. W. Lee and S. S. M. Hussain, "Human error identification: an analysis of myringotomy and ventilation tube insertion," *Archives of Otolaryngology-Head and Neck Surgery*, vol. 130, no. 10, pp. 1153, 2004.

[3] A.K. Ho, H. Alsaffar, P. C. Doyle, H. M. Ladak, S. K. Agrawal, "Virtual reality myringotomy simulation with real-time deformation: Development and validity testing," *Laryngoscope*, vol. 122, no. 8, pp. 1844 – 1851, May, 2012.

[4] L. J. Sowerby, G. Rehal, M. Husein, P. C. Doyle, S. K. Agrawal and H. M. Ladak, "Development and face validity testing of a three-dimensional myringotomy simulator with haptic feedback," *J. Otolaryngol. Head. Neck. Surg.*, vol. 39, no. 2, pp. 122-129, Apr, 2010.

[5] B. Wheeler, P. Doyle, S. Chandarana, S. Agrawal, M. Husein, H. Ladak, "Interactive computer-based simulator for training in blade navigation and targeting in myringotomy," *Computer Methods and Programs in Biomedicine*, vol. 98, pp. 130-139, 2010.

[6] B. Wheeler, "Virtual Reality Simulation for Myringotomy Training," *M.E.Sc. thesis*, The University of Western Ontario, London, Canada, 2009.

[7] M. Zirkle, D. W. Roberson, R. Leuwer, A. Dubrowski, "Using a Virtual Reality Temporal Bone Simulator to Assess Otolaryngology Trainees," *Laryngoscope*, vol. 117, issue 2, pp 258 – 263, 2007.

[8] E. Hiemstra, E.M. Terveer, M.K. Chmarra, J. Dankelman, F.W. Jansen, "Virtual reality in laparoscopic skills training: is haptic feedback replaceable?," *Minim Invasive Ther Allied Technol.*, vol. 20, no. 3, pp 179- 184, May 2011.

[9] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, S. Dawson, "Metrics for Laparoscopic Skills Trainers: The Weakest Link!," Medical Image Computing and Computer Assisted Intervention (MICCAI), *Lecture Notes in Computer Science (LNCS)*, vol. 2488, pp. 34 – 43, 2002.

[10] Y. Yamauchi, J. Yamashita, O. Morikawa, R. Hashimoto, M. Mochimaru, Y. Fukui, H. Uno, K. Yokoyama, "Surgical Skill Evaluation by Force Data for Endoscopic Sinus Surgery Training System," *LNCS*, vol. 2488, pp. 44–51, 2002.

[11] Y. Kinoshita; D. Takeda; A. Sasaki; H. Hashimoto, "Archive and Instruction of the Hand Motion: Analysis and Evaluation of the Hand Motion in Use of

Chopsticks," *SICE-ICASE, 2006. International Joint Conference*, pp.2348-2353, 18-21 Oct. 2006.

[12] E. D. Grober, S. J. Hamstra, K. R. Wanzel, E. D. Matsumoto, R. S. Sidhu, K. A. Jarvi, “Validation of novel and objective measures of microsurgical skill: hand motion analysis and stereoscopic visual acuity,” *Microsurgery*, vol. 23, pp. 317-322, 2003.

[13] M. Agus, A. Giachetti, E. Gobbetti, G. Zanetti, A. Zorcolo, N.W. John, R.J. Stone, “Mastoidectomy Simulation with Combined Visual and Haptic Feedback,” *Stud Health Technol Inform.*, vol. 85, pp 17-23, 2002.

[14] C. Sewell, D. Morris, N. H. Blevins, S. Agrawal, S. Dutta, F. Barbagli, K. Salisbury, “Validating Metrics for a Mastoidectomy Simulator,” *Studies in Health Technology and Informatics*, vol. 125, pp. 421-426, 2007.

[15] C. Sewell, D. Morris, N.H. Blevins, F. Barbagli, K. Salisbury, “Evaluating drilling and suctioning technique in a mastoidectomy simulator,” *Studies in Health Technology and Inform.*, vol. 125, pp. 427-432, 2007.

[16] E. Acosta, B. Temkin, “Haptic laparoscopic skills trainer with practical user evaluation metrics,” *Stud Health Technol Inform.*, vol. 111, pp. 8 – 11, 2005.

[17] J. A. Oostema, M. P. Abdel, Jon C. Gould, “Time-efficient laparoscopic skills assessment using an augmented reality simulator,” *Surgical Endoscopy*, vol.22, no. 12, pp. 2621- 2624, Dec. 2008.

[18] L. Moody, C. Baber, T. N. Arvanitis, “Objective metrics for the evaluation of simple surgical skills in real and virtual domains,” *Presence*, vol. 12, no. 2, Arp. 2003.

[19] R. Lapeer, M.S. Chen, G. Gonzalez, A. Linney, G. Alusi, “Image-enhanced surgical navigation for endoscopic sinus surgery: evaluating calibration, registration and tracking,” *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, no. 1, pp. 32–45, Mar. 2008.

[20] P. J. Edwards, A. P. King, D. J. Hawkes, O. Fleig, C. R. Maurer, Jr, D. L. Hill, M. R. Fenlon, D. A. de Cunha, R. P. Gaston, S. Chandra, *et al.*, “Stereo augmented reality in the surgical microscope,” *Stud Health Technol Inform.*, vol. 62, pp. 102–108. 1999.

[21] C. E. Reiley, H. C. Lin, D. D. Yuh, G, D. Hager, “Review of methods for objective surgical skill evaluation,” *Surg. Endosc.*, vol. 25, pp. 356–366. 2011.

[22] D.G. Ezra, R. Aggarwal, M. Michaelides, N. Okhravi, S. Verma, L. Benjamin, P. Bloom, A. Darzi, P. Sullivan, “Skills Acquisition and Assessment after a Microsurgical Skills Course for Ophthalmology Residents,” *Ophthalmology*, vol. 116, no. 2, Feb. 2009.

[23] A. Alvand, T. Khan, S. Al-Ali, W.F. Jackson, A.J. Price, J.L. Rees, “Simple visual parameters for objective assessment of arthroscopic skill,” *J Bone Joint Surg Am*, vol. 94, no. 13, pp. 971-997, Jul. 2012.

[24] Northern Digital Inc. (NDI), “How does the NDI Polaris System Work?,” www.ndigital.com, [Online]. Available: http://www.ndigital.com/medical/documents/polaris/polarishdtw_datasheet.pdf [Accessed: Jun. 2011].

[25] E. R. Buch, J. L. Contreras-Vidal, “Effects of Parkinson’s disease on visuo-motor adaptation,” *Exp Brain Res*, vol. 150, pp. 25–32, 2003.

[26] James R. Tresilian, George E. Stelmach1 and Charles H. Adler, “Stability of reach-to-grasp movement patterns in Parkinson’s disease,” *Brain*, vol. 120, no. 11, pp. 2093-2111, 1997.

[27] Hans-Leo Teulings, J. L. Contreras-Vidal, G. E. Stelmach, C. H. Adler, “Parkinsonism Reduces Coordination of Fingers, Wrist, and Arm in Fine Motor Control,” Experimental Neurology, vol. 146, pp. 159-170, 1997.

[28] A. Machowska-Majchrzak, K. Pierzchała, S. Pietraszek, B. Łabuz-Roszak, W. Bartman, “The usefulness of accelerometric registration with assessment of tremor parameters and their symmetry in differential diagnosis of parkinsonian, essential and cerebellar tremor,” *Neurol Neurochir Pol.*, vol. 46, no. 2, pp. 145-156, Mar. 2012.

[29] Leica Microsystem, “Leica M720 OH5 Premium Surgical Microscope – A New Dimension of Innovation and Comfort,” www.leica-microsystems.com, [Online]. Available: http://www.leica-microsystems.com/products/surgical-microscopes/neurosurgery-spine/details/product/leica-m720-oh5-2/ [Accessed: May 19, 2012].

[30] J. R. Landis and G. G. Koch, “The Measurement of Observer Agreement for Categorical Data,” *Biometric*, vol. 33, no. 1, pp. 159-174, Mar. 1977.

[31] Omni Institute, “*Guide to Interpreting (and making the most of) your incredible years data*,” www.omni.org, [Online]. Available: http://www.omni.org/docs/InvestInKids/GuidetoInterpretingYourOutcomeResults.pdf. [Accessed: Jul 14, 2012].

Chapter 3

3 Future works and conclusions

3.1 Summary of contribution

The novel aspect of this project was the design and evaluation of a first of its kind study to objectively compare skills of experts and residents in surgical microscope usage during myringotomy. The first set of challenges were implementation of application specific tracking software, implementation of a series of *tracking* metrics capable of evaluating motion data produced by the tracking software and derivation of myringotomy-specific *procedural*, *positional*, and *optical* metrics. The next set of challenges included design and implementation of an experiment with appropriate protocols and equipment and then collect microscope motion data and video data from groups of participating subjects in a controlled and unbiased manner. The final set of challenges included evaluation of all the metrics from the collected raw data, implementation of a database containing all the metric evaluations and performing statistical analyses on the database to determine sets of discriminatory metrics differentiating experts' and residents' performances.

Implementation of tracking software proved to be technically challenging. Although Northern Digital Inc. does provide proprietary tracking software for the Polaris® hybrid tracker, it was not suitable for this study due to several shortcomings. By default it was set to collect tracking data at 30 frames per seconds, providing reduced accuracy of the tracked motion path. It did not show the orientation of the optical marker in the form of a virtual object when it was being tracked. Visualization of the virtual object's orientation in real time was needed during the actual experiment in order to detect outliers. The FOV of the tracker, in addition, was set with a default point of origin that could not be modified to match the operating space of the experiment. Lastly and perhaps most importantly, the proprietary software did not have any feature to keep track of time in any way during a tracking event. Since time is the fundamental parameter needed to implement every tracking metric, it was essential to attain corresponding discrete time instants of all discrete motion data composed of position and orientation coordinates. Therefore, to make all these required changes, customized tracking software was needed and as previously noted in Chapter 2 was implemented using IGSTK.

The most challenging and novel contribution of this work involved the selection and/or development of suitable metrics. Specific motion tracking metrics were selected from the literature and were further refined for assessing microscope usage. For example, these metrics were adapted to evaluate gross and fine motion paths in addition to being applied to the whole motion path as is common in the literature. Procedural, positioning and optical metrics were defined through lengthy discussions with instructing surgeons.

Finally, these metrics were evaluated by collecting motion and video data from experts and junior residents as they performed tube insertion into a myringotomy in a cadaveric head. The metrics were statistically analyzed to determine the ones with the greatest potential for discriminating experts from residents.

3.2 Conclusion

AS shown earlier through box plots, experts scored much better than residents in almost all tracking metrics. However, among those metrics, only the ones showing significant levels of statistical differences between these two groups can be considered as discriminatory metrics. By mastering the skill sets to score better in these metrics, a resident may be able to become as efficient as an expert. Identified based on statistical analyses such as parametric t-tests and non-parametric Kolmogorov Smirnov tests, the discriminatory tracking metrics found in this study were *total completion time*, *total still time*, *fine motion time*, *gross motion volume*, *path smoothness* and *path jitter* during total and fine motions, *total roll* (about x-axis) and *yaw* (about z-axis) rotations during fine motion and finally *total repositions* of the end-frame. Experts scored substantially lower in these metrics compared to the residents. When interpreted, these results outlined that the experts needed much less time to complete the entire operation as they minimized their still time. Their fine motion time was also very short as they were able to manoeuvre the end-frame from the initial optical view spot (i.e., at the end of a gross motion) to the optimum optical view spot quite fast. During gross motion, which was the largest single motion of the end-frame, they needed very limited volume of space.

Finally, they obtained the optimum view without having to reposition the end-frame at multiple spots multiple times. When training a novice resident to manoeuvre a microscope efficiently, the instructor or a simulator may compare the resident's performance metrics to a normative database of discriminatory metrics collected from experts to evaluate his/her proficiency.

Procedural, positional and optical metrics were evaluated by a panel of 3 experts by reviewing videos taken during the experimental sessions. The metrics with the most potential for discriminating junior residents from experts were found to be *unobstructed FOV*, *centered FOV*, *optimum arm level* during operation, *optimum body and arm postures*, *optimum body-to-bed distance*, *end-frame's proper positioning*, *speculum insertion*, *motion efficiency of instruments*, *instrument handling fluidity*, *tube loading*, *tube insertion* and finally *hand jitter*. If these metrics are to be incorporated into a simulator such as ours to provide automated feedback during training sessions, significant effort will be needed to implement them in software.

3.3 Future directions

Currently, the sample sizes of the participating subjects are small. Based on the demographics, among the 4 experts, only 2 have been performing middle ear based surgeries for a significant time. While 1 expert mostly performs head and neck based surgeries, and 1 expert is a newly appointed surgeon. Due to their areas of expertise and number of years in practice, some variability in terms of metrics was observed. Similarly, residents from years 1 – 3 were all lumped together to form a single group of residents, hence there was greater performance variability observed within them. In order to obtain consistency within the groups, it would be the best practice to recruit all experts with similar expertise. Similarly, the residents should be sub-grouped according to their corresponding year of residency. Since no single ENT residency program would have sufficient numbers of residents and experts, a multi-centre study would need to be undertaken. It would also be worthwhile to statistically compare residents at various stages of their program (e.g., PGY 1, 2, etc) to quantify progression from one residency year to the next.

Ultimately, the Auditory Biophysics Laboratory will want to incorporate all discriminatory metrics determined from a multi-center study into the current myringotomy simulator once the representation of the microscope is also improved. In the current study, metrics such as unobstructed FOV, centered FOV, speculum insertion, tube insertion and hand jitter are evaluated by a panel of experts who observed video streams acquired using the microscope's internal camera. Digital image processing could be used to compute these metrics automatically. To implement the metric *unobstructed FOV*, the outline of the eardrum must be automatically detected in the video stream and the software must continually check that the trainee's hand does not obstruct the view of the eardrum. Similarly, for the metric *centered FOV*, the incision in the eardrum must be detected in the video, and the software must check that the incision is at the center of the FOV. If the speculum is optimally inserted, then its outline would be perfectly round as opposed to oval when viewed through the microscope. In this case, the outline would need to be detected in the video and the degree of circularity would need to be computed to form a metric representing *speculum insertion*. When the ventilation tube is optimally inserted into the incision, it appears as a circular ring because ventilation tubes are generally right circular cylinders. In this case, the method adopted for computing the *speculum insertion* metric could be adapted for assessing the metric *tube insertion*. Finally, hand jitter causes blade or forceps jitter seen in the video stream. To evaluate *hand jitter*, the tool tip could be automatically detected and tracked in the video sequence and the metric described in Chapter 2 for jitter in microscope motion tracking could be used.

Appendices

Appendix 1: List of all metric equations.

Name	**Equation**
Time	T = $\sum_{i=1}^{N-1} t_{i+1} - t_i$, *t* stands for a particular time instant
Path length	$L_T = \sum_{i=1}^{N-1} \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2 + (z_{i+1}-z_i)^2}$
Total rotation	$R_T = \sum_{i=1}^{N-1} \lvert\theta_{i+1} - \theta_i\rvert$
Average velocity	$V_{Pavg} = \frac{1}{N-1}\sum_{i=1}^{N-1} \sqrt{\left(\frac{x_{i+1}-x_i}{t_{i+1}-t_i}\right)^2 + \left(\frac{y_{i+1}-y_i}{t_{i+1}-t_i}\right)^2 + \left(\frac{z_{i+1}-z_i}{t_{i+1}-t_i}\right)^2}$
Average acceleration	$A_{Pavg} = \frac{1}{N-1}\sum_{i=1}^{N-1} \sqrt{\left(\frac{Vx_{i+1}-Vx_i}{t_{i+1}-t_i}\right)^2 + \left(\frac{Vy_{i+1}-Vy_i}{t_{i+1}-t_i}\right)^2 + \left(\frac{Vz_{i+1}-Vz_i}{t_{i+1}-t_i}\right)^2}$
Normalized motion jerk (Required to quantify path smoothness)	$NJ_P = \sqrt{\frac{t_P^5}{2(N-1)L_P^2}\sum_{i=1}^{N-1}\left\lvert\frac{j_{x_{i+1}}^2 + j_{x_i}^2 + j_{y_{i+1}}^2 + j_{y_i}^2 + j_{z_{i+1}}^2 + j_{z_i}^2}{2}\right\rvert (t_{i+1}-t_i)}$ Here t_P is the total time of the path and L_P is the total path length. The lower the NJ_P, the smoother the path. Here squared jerk or $j^2 = \left(\frac{a_{i+1}-a_i}{i_{i+1}-t_i}\right)^2$ where *a* stands for acceleration and *t* stands for time.
Motion Jitter	$Threshold = 0.1[max\lvert J(t)\rvert]$, where $\lvert J(t)\rvert$ is the jerk amplitude calculated above. Jitter computed for motion jerk above the threshold. $Jitter = \frac{T^2}{L_T}\int_0^T \lvert J(t)\rvert dt$
Manoeuvring volume	$V_T = [max(x_n) - \min(x_n)] \times [max(y_n) - \min(y_n)] \times [max(z_n) - \min(z_n)]$
Efficiency measure	Normalized factor = $\frac{\text{Actual}}{\text{Ideal}}$

Above, *i* represents the i^{th} sample in the discrete data stream, *N* is the total number of data points in a selected motion path.

x_i, y_i, z_i therefore are instantaneous coordinate points at i^{th} sample. Same applies for angular coordinates.

Appendix 2

Baseline Questionnaire

Program: ________________________________

PGY: ① ② ③ ④ ⑤

Please estimate a reasonable answer that applies to *you* for any of the following questions.

1. Do you have any ENT surgical experience? Yes: _____ No: _____
2. How many myringotomies have you performed previously? Please approximate a reasonable number.

   ____
3. If you have trained using a myringotomy simulator (i.e., physical models, software etc.), how many simulated myringotomy have you performed? Please approximate a reasonable number.

   ____
4. How many months have you spent on Neurology and Paediatric rotations?

   ____
5. How many times have you used a surgical microscope for ear examination or debridement so far?

   ____
6. How frequently have you performed a myringotomy (or any ear procedure requiring a microscope) in a day and in a week during your Neurology and Paediatric rotations? (i.e., 3/day; 13/week etc.)

   ____________ ____________
7. Do you have microscope experience at a non-surgical setting (i.e., pathology, histology, biology etc.)? Yes: _____ No: _____
8. Rate your comfort level in manoeuvring surgical microscope. (1 = least; 10 = most).

   1 2 3 4 5 6 7 8 9 10
9. Rate your expertise in video games that require a controller or a keyboard.

   1 2 3 4 5 6 7 8 9 10
10. Rate your musical instrument playing ability.

    1 2 3 4 5 6 7 8 9 10

Appendix 3: Optical, Positional and procedural metric evaluation questionnaire

Optical metrics evaluation table (Please shade/tick your answer)

Evaluator: ______________________ Subject/video number: __________

Metric name	Evaluation standard	Evaluating question	Expert evaluation
Use of Eyes	Intraocular distance adjustment *Default:* minimum	Was it adjusted accordingly?	YES ☐ NO ☐
	Intraocular tilt *Default:* strait 180^o to lens	Was it adjusted accordingly?	YES ☐ NO ☐
Focus	*Default* WD = 200	Focus quality?	Unfocussed ☐ Focussed ☐
Zoom	Magnification *Default* M = 2.7	How is the zoom?	Too much ☐ Too little ☐ ☐ Optimum
Field-of-View (FOV) obstruction	Default: Unobstructed view throughout procedure	Is FOV obstructed during procedure?	Completely ☐ Partially ☐ ☐ Unobstructed
Optical view of the surgical site	Default: myringotomy and surrounding area visible and centred	Is the optimum surgical view obtained during procedure?	YES ☐ NO ☐
Based on these OPTICAL metrics, do you think this subject is an expert or a resident?			EXPERT ☐ RESIDENT ☐

<u>Clinical metrics evaluation table (Please shade/tick your answer)</u>

Evaluator: ______________________ Subject/video number: __________

Name	Evaluation objective	Expert evaluation
Operator's arm level	The vertical level of the operator's arm	☐ Too high ☐ Acceptable ☐ Too low
Operator's chair height	The vertical position of the operator's chair during operation	Too high ☐ Optimum ☐ Too low ☐
Operator's wrist position	Positioning and stabilization (bracing) of the wrist against patient's head	☐ Unacceptable Acceptable ☐
Operator's posture	Position of the operator's body while performing the procedure	Slouched ☐ ☐ Acceptable Stretched ☐
Operator's body to bed distance	Operator's body to bed positioning	Too far ☐ Acceptable ☐ Too close ☐
	Operator's arm posture	Outstretched ☐ Acceptable ☐ Too close ☐
Microscope repositioning	Positions microscope efficiently by minimizing repositioning	☐ Yes No ☐
Speculum insertion	Inserts speculum appropriately	☐ Yes No ☐
Microscope positioning	Positions microscope appropriately	Too close ☐ Acceptable ☐ Too far ☐
Instrument efficiency of time/motion	Does the operator perform unnecessary manoeuvres during the insertions with poor efficiency of motion?	☐ Yes No ☐
Instrument handling	Does the operator have fluid use of instruments with excellent control?	☐ Yes No ☐
Tube loading	Loads tube onto instrument appropriately	☐ Yes No ☐
Tube insertion	Inserts tube appropriately	☐ Yes No ☐
Hand jitters	Hand jitters during tube insertions	Significant Jitter ☐ ☐ Minimal jitter No jitter ☐
Based on these CLINICAL metrics, do you think this subject is an expert of a resident?		☐ RESIDENT EXPERT ☐

Curriculum Vitae

Name:	Arefin Shamsil
Post-secondary Education and Degrees:	McMaster University Hamilton, Ontario, Canada 2006-2010 Bachelor of Engineering
	Western University London, Ontario, Canada 2010-2012 Masters of Engineering Science
Honours and Awards:	Queen Elizabeth II aiming for the top Scholarship (2006)
	McMaster University undergraduate Scholarship (2007, 2008)
	Western Engineering Scholarship (2010, 2012)
Related Work Experience	Junior Engineer (Co-op) EnWin Utilities, Windsor, Ontario, Canada (2006)
	Teaching Assistant Western University, London, Ontario, Canada (2011)
	Graduate Research Assistant Western University, London, Ontario, Canada (2010, 2012)

Conference Proceeding:

Shamsil A, Wickens B, Doyle PC, Agrawal SK, Ladak HM. Implementation of metrics to assess microscope manoeuvring skills during myringotomy. Proc. 35th Canadian Medical & Biological Engineering Conference, Halifax, NS, June 2012.