8-2019

Development and Preliminary Validation of Image-enabled Process Metrics for Assessment of Open Surgery Suturing Skill

Irfan Kil
Clemson University, irfankil@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation
https://tigerprints.clemson.edu/all_dissertations/2462

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
DEVELOPMENT AND PRELIMINARY VALIDATION OF IMAGE-ENABLED PROCESS METRICS FOR ASSESSMENT OF OPEN SURGERY SUTURING SKILL

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Electrical Engineering

by
Irfan Kil
August 2019

Accepted by:
Dr. Richard E. Groff, Committee Chair
Dr. Ravikiran B. Singapogu, Committee Co-Chair
Dr. Adam Hoover
Dr. Ian D. Walker
Abstract

Suturing is a fundamental surgical skill required in a variety of operations, ranging from wound repair to delicate vascular reconstruction. It is essential that surgeons master requisite suturing skills so that he or she can deliver safe and effective care to patients. Due to an increased emphasis on standardized medical training, tools and methods are needed to provide objective assessment and feedback during the learning process. In this thesis, a new surgical simulator for assessment and training of open surgery suturing skill is introduced. The suturing simulator system design, force-based, motion-based, image-based and image-enabled metrics for skill assessment, and a preliminary study of resident and attending surgeons are presented.

The simulator collects synchronized force, motion, video and touch data during radial continuous suturing. The synchronized data is used to extract metrics for suturing skill assessment. The simulator has a camera positioned underneath the suturing membrane, enabling visual tracking of the needle during suturing. Needle tracking data enables extraction of meaningful metrics for both the process and the product of the suturing task. To better simulate surgical conditions, the height of the simulator and the depth of the membrane are both adjustable.

The metrics were motivated by insight provided to us from practicing vascular surgeons. These metrics were based on the physics of needle insertion forces, the maxim to follow the curve of the needle while driving through tissue, and minimizing lateral forces and motions that induce tear.
Experimental data from a study involving subjects with various levels of suturing expertise (attending surgeons and surgery residents) are presented. Analysis shows force-based metrics (absolute maximum force/torque in z-direction), motion-based metrics (yaw, pitch, roll), a physical contact metric, image-based metrics (Stitch Length, Idle Time, Needle Tip Trace Distance, Needle Swept Area, Needle Tip Area and Needle Sway Length) and image-enabled metrics (orthogonal force, tangential force and entry angle) are statistically significant in differentiating suturing skill between attendings and residents.

The results suggest that this simulator and accompanying metrics can be used to assess open surgery suturing skill. Furthermore, analysis shows that 6 of 9 image-based metrics were effective in capturing fine-grain differences in skill level between residents and attendings. Moreover, image-based process metrics may be represented graphically in a manner conducive to training. Image-based metrics especially lend themselves to intuitive visualizations. The combination of fine-grained skill differentiation, the ability to simulate depth of suturing, and the intuitive visualizations of selected image-based metrics makes the suturing simulator and associated suite of metrics well-suited for suturing skills assessment and training.
Dedication

In the name of Allah (God), the Most Gracious, the Most Merciful! This dissertation is dedicated to my beloved parents Omer and Feride, and my lovely sister Fatma Rumeysa for their unconditional support throughout my life. Also, to my grandparents and family, whose prayers bless me every time.
Acknowledgments

First and foremost, I would like to thank Allah (God) for giving me the opportunity, strength, and patient to complete this dissertation. I thank Allah for blessing me with amazing people, opportunities, and great success throughout my education. Without his blessing, this work would not have been completed.

I would like to express my sincere gratitude and deep appreciation to my advisor Dr. Richard E. Groff for his guidance, patience, cooperation, and support during the development of this dissertation. Dr. Groff has provided me with excellent supervision, motivation, enthusiasm, and invaluable comments during the work on this dissertation.

I would like to thank Dr. Ravikiran B. Singapogu for his continued support and motivation throughout this work. My sincerest gratitude deep appreciation also extends to Dr. Adam Hoover and Dr. Ian D. Walker for serving on my committee.

I would also like to thank John F. Eidt, MD for his invaluable insight and feedback to this research. I would also like to thank my project mates Anand Jagannathan and Naren Nagarajan for their contributions to this research.
# Table of Contents

Title Page ................................................................. i
Abstract ................................................................. ii
Dedication ............................................................... iv
Acknowledgments ....................................................... v
List of Tables ............................................................ viii
List of Figures ........................................................... ix

1 Introduction ......................................................... 1
  1.1 Introduction .................................................. 1
  1.2 Overview of Thesis ........................................... 6

2 Literature Review .................................................. 7
  2.1 Surgical Skill Assessment by Human ......................... 7
  2.2 Simulators for Surgical Skill Assessment ................... 12
  2.3 Metrics for Surgical Skill Assessment ....................... 15

3 The Suturing Simulator ............................................. 20
  3.1 Physical Components of the Suturing Simulator ............. 21
    3.1.1 Membrane Housing ..................................... 21
    3.1.2 Suturing Membrane .................................... 22
    3.1.3 Height Adjustable Table ............................... 23
    3.1.4 Sensors ................................................ 23
      3.1.4.1 Physical Touch Sensor ............................. 23
      3.1.4.2 Internal and External Cameras ..................... 25
      3.1.4.3 Force/Torque Sensor .............................. 27
      3.1.4.4 Motion Sensor .................................... 27
  3.2 Data-Collection Stage ........................................ 27
  3.3 Data-Processing Stage ....................................... 29

4 Computer Vision Algorithm ...................................... 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Image Processing Stage</td>
<td>33</td>
</tr>
<tr>
<td>4.2 Computer Vision Algorithm Results</td>
<td>39</td>
</tr>
<tr>
<td>5 Metrics for Suturing Skill Assessment</td>
<td>41</td>
</tr>
<tr>
<td>5.1 Vision-Enabled Partitioning of Suture Cycle</td>
<td>41</td>
</tr>
<tr>
<td>5.2 Extracting Metrics from Time-Series Data</td>
<td>45</td>
</tr>
<tr>
<td>5.3 Force/Torque-based, Motion-based and Physical Contact Metrics</td>
<td>45</td>
</tr>
<tr>
<td>5.3.1 Force/Torque-based Metrics</td>
<td>45</td>
</tr>
<tr>
<td>5.3.2 Motion-based Metrics</td>
<td>46</td>
</tr>
<tr>
<td>5.3.3 Physical Contact Metric</td>
<td>46</td>
</tr>
<tr>
<td>5.4 Image-based Metrics</td>
<td>46</td>
</tr>
<tr>
<td>5.4.1 Needle Tip Path Length</td>
<td>48</td>
</tr>
<tr>
<td>5.4.2 Needle Tip Area</td>
<td>48</td>
</tr>
<tr>
<td>5.4.3 Needle Swept Area</td>
<td>49</td>
</tr>
<tr>
<td>5.4.4 Needle Sway Length</td>
<td>51</td>
</tr>
<tr>
<td>5.4.5 Relationship between Image-based Metrics</td>
<td>51</td>
</tr>
<tr>
<td>5.5 Image-enabled Metrics</td>
<td>53</td>
</tr>
<tr>
<td>5.5.1 Orthogonal and Tangential Forces</td>
<td>53</td>
</tr>
<tr>
<td>5.5.2 Entry Force and Angle</td>
<td>56</td>
</tr>
<tr>
<td>6 System Validation and Experimental Results</td>
<td>61</td>
</tr>
<tr>
<td>6.1 Experimental Setup and Protocol</td>
<td>61</td>
</tr>
<tr>
<td>6.2 Experimental Results and Discussion</td>
<td>62</td>
</tr>
<tr>
<td>6.2.1 Force/Torque-based Metrics</td>
<td>64</td>
</tr>
<tr>
<td>6.2.2 Motion-based Metrics</td>
<td>67</td>
</tr>
<tr>
<td>6.2.3 Physical Contact Metric</td>
<td>68</td>
</tr>
<tr>
<td>6.2.4 Image-based Metrics</td>
<td>69</td>
</tr>
<tr>
<td>6.2.5 Image-enabled Metrics</td>
<td>74</td>
</tr>
<tr>
<td>6.2.5.1 Orthogonal and Tangential Forces</td>
<td>74</td>
</tr>
<tr>
<td>6.2.5.2 Entry Force and Angle</td>
<td>76</td>
</tr>
<tr>
<td>7 Conclusion and Future Work</td>
<td>79</td>
</tr>
<tr>
<td>7.1 Conclusion</td>
<td>79</td>
</tr>
<tr>
<td>7.2 Future Work</td>
<td>80</td>
</tr>
<tr>
<td>Appendices</td>
<td>83</td>
</tr>
<tr>
<td>A Towards Real-Time Surgical Skill Training</td>
<td>84</td>
</tr>
<tr>
<td>A.1 GPU Implementation - CPU to GPU Transition</td>
<td>84</td>
</tr>
<tr>
<td>A.2 Visual Feedback For Training</td>
<td>88</td>
</tr>
<tr>
<td>B Questionnaire</td>
<td>90</td>
</tr>
<tr>
<td>C Publications</td>
<td>92</td>
</tr>
<tr>
<td>Bibliography</td>
<td>93</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Literature Review Metrics used in Suturing Skill Assessment in Minimally Invasive Surgery (MIS) and Open Surgery (OS) .......................... 19

6.1 Statistical Results for Force-based, Motion-based and Physical Contact Metrics .............................................................. 63

6.2 Statistical Results for Image-based and Image-enabled Metrics .......................... 70
List of Figures

1.1 A surgeon during a suturing practice ........................................... 2
1.2 The phases of performing suturing ............................................... 3
1.3 An example of tapered suture needle (Prolene SH, Ethicon US, LLC.) ... 5

2.1 Objective Structured Assessment of Technical Skill (OSATS) [1, 2] .... 8
2.2 Global Operative Assessment of Laparoscopic Skills (GOALS) [3] ..... 9
2.3 Global Evaluative Assessment of Robotic Skills (GEARS) [4] .......... 10
2.4 The FLS tasks: (i) peg transfer, (ii) pattern cutting, (iii) ligating loop,
   (iv) extracorporeal suture, (v) intracorporeal suture [5] ............... 11
2.5 The FVS tasks example: (i) radial, continuous suturing, (ii) a patch an-
   gioplasty, and (iii) end-to-side anastomosis [6] ....................... 12

3.1 Suturing simulator overview ..................................................... 20
3.2 a) membrane housing b) suture membrane .................................. 22
3.3 Adjustable height table demonstration ....................................... 23
3.4 Connection diagram for the physical touch sensor: 100 kΩ resistance
   used. External coating was grounded to eliminate possible external affects. ... 24
3.5 Internal View of Membrane Housing ......................................... 25
3.6 External Camera: A Logitech C920 ............................................. 26
3.7 Force data collected using an ATI MINI 40 force/torque sensor .......... 27
3.8 Motion data collected using InertiaCube4 sensor placed on dorsum of
   subject’s hand. ........................................................................ 28
3.9 System process flow-chart, consisting of two stages. In the Data Collection
   Stage, raw data from multiple sensors were synchronized and logged. In
   the Data Processing Stage, the collected data were used to extract metrics
   for suturing skill. ................................................................. 30

4.1 The algorithm consists of two stages: In the Image Processing Stage,
   the needle and thread are detected and needle entry and exit points are
   identified. In the Metrics Extraction Stage, metrics were computed based
   on information from the Image Processing Stage. .......................... 32
4.2 Camera distortion correction and image quality enhancement stage .... 33
4.3 Marker detection to obtain membrane center and pixel-to-mm conversion
   rate ($k_{cam} = 0.19$ pixels/mm) ................................................. 33
4.4 Steps to obtain endpoints of the needle ...................................... 34
4.5 Illustration of the needle tip movement: Red indicates the needle tip whereas green indicates needle entry point.

4.6 Illustration of the needle swage movement: Green indicates the needle swage whereas red indicates needle exit point.

4.7 Needle Tip Trajectory: Pixel list of the needle tip filtered and weeded to be used in post-processing to compute performance metrics.

4.8 Steps for detection of needle and thread; First, camera distortion was corrected and image visible quality was enhanced. Then, frame was masked to obtain region of interest and conversion from RGB to HSV were achieved. Next, needle and thread were detected via vision algorithm and detections were superimposed to the original image.

4.9 Computer vision frame-by-frame display of a) needle and thread detection along with the entry and exit points; b) needle tip path (in red), and swage path (in yellow); c) needle swept area (in black) with the needle detection (in green).

5.1 Graphical User Interface (GUI) designed to show the synchronized force and motion data, along with videos from external and internal cameras. GUI allows for convenient, interactive investigation of synchronized data.

5.2 Example of synchronized force, torque, motion, and touch data for one active suturing time with suture sub-events labeled. A suture cycle is comprised of active suturing time and idle time. Active suturing time is the total duration of entry, driving, exit and pull out phases, whereas idle time (not shown here) is the time between an end-time of one suture to start time of another. (Note: Blue diamond symbol (○) in touch data indicates the time instance of the physical touch).

5.3 An example of one term in the summation (5.9) to find Needle Tip Area. The shaded area represents the area added to the summation of index i.

5.4 A plot of the sway length $l_{sw}(t)$ for one suture cycle, along with example images illustrating the needle with positive (blue) and negative (green) orientation.

5.5 Frame-by-frame illustration of (a) Needle position (in gray); (b) Cumulative area swept out by the needle body from first frame to current frame (in blue); and (c) Needle tip trajectory from first frame to current frame (in black dashed line). Needle Swept Area metric is the area of the blue in the last frame. Needle Tip Path Length metric is the entire length of the black dashed line in the last frame. Needle Tip Area metric is the total area between the straight-line from entry to exit (red dashed line) and the needle tip path (in green) in the last frame.
5.6 Decomposition of horizontal forces into forces orthogonal and tangential to stitch direction: (a) view of the needle taken from internal camera with force sensor coordinate system overlaid; (b) side view of the needle along with orthogonal and tangential force direction; (c) zoomed in view of the needle, together with the force sensor, orthogonal and tangential forces; (d) X and Y directional forces at suture location in (a), for one active suture time; (e) corresponding orthogonal and tangential forces from (a), for one active suture time. 

5.7 Entry force and angle illustration: (a) Side view of the needle along with decomposition of net force into z-directional and tangential forces at a suture location (b) z-directional needle force for one suture cycle, with start point, needle entry/exit points, and entry phase labeled. 

5.8 Illustration of entry force and angle for needle entry phase. Metrics related to entry force and angle were only calculated for the time duration between the 25th ($t_{25}$) and 75th ($t_{75}$) percentage of the entry phase for each suture. This was chosen to avoid possible signal-to-noise ratio effect on needle angle calculation (shown in red circle) related to small tangential and z-directional force changes. 

5.9 Comparison of the angle calculation between the video frame and image-enabled force-based angle. 

5.10 Example of video frames used to calculate entry angles from the external video for certain needle entry angles: (i) 25° (ii) 45° (iii) 70° (iv) 90° (v) 115° (vi) 145°. Here, a straight needle was used. The calculated angles were then compared to $\alpha_e$, i.e. the image-enabled force-based angle calculation.

6.1 Experimental Results for Force-based Metrics: * indicates statistical significance for $p<0.05$. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) 

6.2 Experimental Results for Torque-based Metrics: * indicates statistical significance for $p<0.05$. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) 

6.3 Experimental Results for Motion-based Metrics: * indicates statistical significance for $p<0.05$. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)
6.4 Experimental Results for Physical Contact Metrics: * indicates statistical significance for \( p < 0.05 \). (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25\(^{th}\) and 75\(^{th}\) percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) .................................................. 68

6.5 Experimental Results for Image-based Metrics: * indicates statistical significance for \( p < 0.05 \). (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25\(^{th}\) and 75\(^{th}\) percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) .................................................. 71

6.6 Needle Swept Area and Needle Tip Trajectory Exemplars for an Attending and a Resident .................................................................................................................. 73

6.7 Experimental Results for Image-enabled Metrics (Orthogonal and Tangential Forces): * indicates statistical significance for \( p < 0.05 \). (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25\(^{th}\) and 75\(^{th}\) percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) .................................................. 75

6.8 A snapshot that illustrates thread pulling during suturing .......................................... 77

6.9 Experimental Results for Image-enabled Metrics (Entry Force and Angle): * indicates statistical significance for \( p < 0.05 \). (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25\(^{th}\) and 75\(^{th}\) percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.) .................................................. 78

1 A flowchart of the system that allows GPU Computation ............................................. 85

2 A flowchart of the final system .......................................................................................... 86

3 Prototype for visual feedback .......................................................................................... 87

4 Visual feedback for subject’s progress ................................................................................. 88
Chapter 1

Introduction

1.1 Introduction

Quantification of a surgeon’s skill has received attention in recent years due to multiple factors including: duty hour restrictions on surgical residents, limited training options, a call for the reduction in medical errors, and a need for structured training [7, 8, 9, 10]. Surgical skill is important due to the direct relationship between surgical performance and post-surgery clinical outcomes such as hospital readmission and complication rates [11]. Inadequate surgical skill has been shown to lead to potential medical errors, currently the third leading cause of death in the USA [10]. Surgical outcomes may be improved through training to improve skill. For this purpose, surgical simulators —capable of simulating an aspect of a surgical procedure and of assessing and/or training the subject’s skill on a given task— have received special attention in recent years. One main advantage of using simulators is the ability to train surgical skills without the use of animals or humans. Another key advantage is the ability to measure skill and its progression over time. Overall, evidence suggests that surgical simulators are an effective tool in the training of surgical skills [12, 13].

When using a surgical simulator to quantify the surgical skill of a subject, it is es-
sential to consider the type of surgery in which the subject is being trained: Open Surgery or Minimally Invasive Surgery (MIS). Though MIS has many advantages, open surgery is still performed for delicate surgeries such as cardiovascular and spine operations. Most currently available simulators focus on MIS including endovascular, laparoscopic, and robotic procedures, while few simulators are available for open surgery. The lack of open surgery simulators may be attributed to the difficulty of measuring a surgeon’s actions without the restrictions on tools and movements imposed by the minimally invasive environment [14]. Despite the associated challenges, quantification of skill in open surgery is especially valuable since it has been shown that there is no established correlation between skill in open surgery and in MIS [15].

The changing landscape of surgical techniques has created demand for more efficient and effective training methods. In conventional surgical skill training, expert surgeons observe and provide feedback to novices during exercises [16]. This type of training may be partly subjective since feedback often depends on the expert surgeon’s preferences and style [17]. Further, training draws expert surgeons away from clinical responsibilities [18]. Medical simulators were developed to address these problems and
Figure 1.2: The phases of performing suturing
to standardize and automate assessment of a surgeon’s skill.

Suturing is a fundamental surgical skill required in a variety of operations, ranging from wound repair in trauma care to delicate vascular reconstruction [19]. Suturing is the name given to the process of stitching a rupture or a tear in the tissue, or stitching a tissue to a tissue or a graft, using a needle and suture. The suture, also called thread, is the term that refers to the material used during the stitching of the wounded tissue.

Suturing has been practiced consistently over the centuries, earning its reputation as a necessary procedure for the proper and fast healing of tissue. Not all sutures are made equal. As such, depending on the nature of the surgical operation, a certain type of suture should be used to perform suturing. The selection of suture type depends on the tissue type, tissue location, tissue thickness and tissue tension. Primarily, sutures are categorized into two main types: absorbable and nonabsorbable [20]. Absorbable sutures are created with material which is capable of being absorbed by the body and are generally used for internal tissue repairs. Nonabsorbable sutures, on the other hand, are usually used for external tissue repairment, or for internal tissue repairs that require long-term healing. Both absorbable and nonabsorbable sutures are made of either synthetic materials or natural fibers. Synthetic sutures are the most common type of suture used in tissue repair since their material properties, such as absorption rate and absorption time, are well-known and predictable. Ethilon, prolene, monocryl, and vicryl are just a few examples of synthetic sutures.

Just as the choice of suture type depends on the surgical operation, so does needle type. The most commonly used are cutting needles and tapered needles, both made of stainless steel. Cutting needles are generally used for closure of small incisions, as in both cosmetic and non-cosmetic plastic surgical operations. Tapered needles, on the other hand, are usually used in procedures involving suturing inside of the body, for the closure or stitching of soft and delicate tissues, as in a vascular surgical procedure [20]. An example of a suture with a tapered surgical needle and its parts can be seen in
Multiple techniques may be used to perform suturing. Primary methods include interrupted and uninterrupted, i.e., continuous suturing [21, 22]. During interrupted suturing, every stitch performed by a surgeon is finished off with a knot and the suture is cut prior to beginning the next stitch. In contrast, during uninterrupted suturing, a surgeon performs multiple stitches one after another in a continuous fashion without cutting or knotting the suture during the process. In order for proper suturing to be performed, it is critical to select the appropriate suture needle size, suture type, and suture technique, with full consideration of the surgical operation to be performed [21, 22].

Suturing is a complex procedure and suturing on an expert level requires the use of certain fundamental techniques. The process of suturing can be divided into the following phases: (i) puncturing a needle into the tissue perpendicularly, (ii) driving the needle through the tissue following the curvature of the needle, (iii) exiting the tissue from an exit point, and (iv) pulling the needle from the tissue completely prior to tightening the suture. Demonstration of the phases can be seen in Fig. 1.2. Learning skilled suturing is essential for novice medical practitioners and has been incorporated into most fundamental skills training curricula, for example, the Fundamentals of Laparoscopic Surgery (FLS) [23, 24, 12] and Fundamentals of Vascular Surgery (FVS) curricula [25]. However, most currently available simulators that provide metrics data have been
developed for minimally invasive surgery; only a handful of attempts have focused on the development of simulators for open surgery [26]. Furthermore, the majority of work on suturing skill focuses on product metrics, i.e., metrics based on analyzing the final results of the task. Process metrics, i.e., metrics that quantify skill by analyzing how the task was performed, provide significantly more insight for skill training and assessment than product metrics, but are also more technically challenging to obtain.

To address the limitations of current surgical simulators, we have developed a suturing simulator which collects synchronized force, motion, touch, and video data as trainees perform suturing. Product and process metrics are extracted from these data and are used to distinguish suturing skill level. A feature of this system is that standard surgical tools (needle holder, needle with surgical thread, etc.) are used on the platform in contrast to simulators which require the use of modified surgical tools (example needle coloring, dots for computer vision tracking, etc.). Also, the system allows suturing at various depth levels, simulating surgery inside a body cavity or on the surface [27].

1.2 Overview of Thesis

This thesis is organized as follows. Chapter 2 summarizes published works on simulators, as well as the metrics for surgical skill assessment that provide the motivation and background for this research. Chapter 3 describes the construction of the suturing simulator and the system processes. Chapter 4 describes a computer vision algorithm used to obtain vital information about needle and thread movement from video data, allowing for the extraction of metrics useful in the assessment of suturing skill. Chapter 5 presents the metrics for skill assessment (force-based, motion-based, physical contact, image-based, and image-enabled metrics). Chapter 6 presents the system validation, experimental setup, and corresponding results of a study. Finally, Chapter 7 presents conclusions and future work.
Chapter 2

Literature Review

Recent research has revealed a correlation between the skill of a surgeon and the quality of clinical outcomes [18, 11]. It is imperative, therefore, that attention is given to measuring surgical skill. For this reason, studies have focused on (i) developing tools, (ii) developing surgical simulators, and (iii) extracting metrics for assessment of surgical skill. In the following sections, published works on assessment tools, simulators, and metrics used to assess surgical skill are presented.

2.1 Surgical Skill Assessment by Human

Traditionally, surgical skill is assessed by expert surgeons who observe and give feedback to a trainee during “on the job” training [16, 17]. To standardize the assessment process, an effort has been placed on developing and validating skill assessment tools for consistent evaluations.

Objective Structured Assessment of Technical Skill (OSATS) is an assessment tool used for grading surgeons’ surgical performance during open surgery [1, 2]. While performing a surgical procedure, surgeons are observed and assessed by expert surgeons in the categories of: (i) respect for tissue, (ii) time and motion, (iii) instrument handling, (iv) knowledge of instrument, (v) flow of operation, (vi) forward planning, (vii) use of
assistants, and (viii) knowledge of specific procedure. An example of OSATS grading chart is shown in Fig 2.1.

Global Operative Assessment of Laparoscopic Skills (GOALS) [3] is an assessment tool for surgeons’ surgical performance in Minimally Invasive Surgery (MIS). This assessment method is an adapted version of OSATS, specific to procedures in MIS. Skills are assessed based on depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy. Fig. 2.2 shows a sample GOALS grade chart with descriptions of each score.

### Figure 2.1: Objective Structured Assessment of Technical Skill (OSATS) [1, 2]

<table>
<thead>
<tr>
<th>Respect for Tissue</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments</td>
<td>Careful handling of tissue but occasionally caused inadvertent damage</td>
<td>Consistently handled tissue appropriately with minimal damage</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time and Motion</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many unnecessary moves</td>
<td>Efficient time/motion but some unnecessary moves</td>
<td>Clear economy of movement and maximum efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instrument Handling</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatedly makes tentative or awkward moves with instruments by inappropriate use of instruments</td>
<td>Competent use of instruments but occasionally appear stiff or awkward</td>
<td>Fluid moves with instruments and no awkwardness</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Knowledge of Instruments</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequently asked for wrong instrument or used inappropriate instrument</td>
<td>Knew names of most instruments and used appropriate instrument</td>
<td>Obviously familiar with the instruments and their names</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flow of Operation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequently stopped operating and seemed unsure of next move</td>
<td>Demonstrated some forward planning with reasonable progression of procedure</td>
<td>Obviously planned course of operation with effortless flow from one move to the next</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use of Assistants</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistently placed assistants poorly or failed to use assistants</td>
<td>Appropriate use of assistants most of time</td>
<td>Strategically used assistants to the best advantage at all time</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Knowledge of Specific Procedure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deficient knowledge. Needed specific instruction at most steps</td>
<td>Knew all important steps of operation</td>
<td>Demonstrated familiarity with all aspects of operation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.2: Global Operative Assessment of Laparoscopic Skills (GOALS) [3]

The Global Evaluative Assessment of Robotic Skills (GEARS) [4] is another assessment tool for surgeons’ surgical performance in robotic surgery, tailored from OSATS and GOALS. Skills are assessed based on depth perception, bimanual dexterity, efficiency, force sensitivity, autonomy, and robotic control, which is very similar to GOALS. Fig. 2.3 shows a sample GEARS grade chart with descriptions of each score.

In all aforementioned tools and curricula, the skill assessments are based on either
direct observation during the task or blinded videotaped assessment after the completion of the task.

The Fundamentals of Laparoscopic Surgery (FLS), a standardized training curriculum, consists of specific tasks used to prepare surgeons for laparoscopic procedures [23, 24, 12]. The tasks included in this curriculum are peg transfer, pattern cutting, ligating loop, extracorporeal suture and intracorporeal suture (see Fig. 2.4). After this training, each trainee is examined for cognitive and technical laparoscopic skills, a necessary graduation requirement for residency programs [28, 23, 29]. The FLS has been validated in earlier studies. It has been shown that FLS results correlate with patient outcomes [30]. However, during assessment of FLS, task time is used as the most im-

<table>
<thead>
<tr>
<th>Depth perception</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constantly overshoots target, wide swings, slow to correct</td>
<td>Some overshooting or missing of target, but quick to correct</td>
<td>Accurately directs instruments in the correct plane to target</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bimanual dexterity</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses only one hand, ignores nondominant hand, poor coordination</td>
<td>Uses both hands, but does not optimize interaction between hands</td>
<td>Expertly uses both hands in a complementary way to provide best exposure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficiency</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inefficient efforts; many uncertain movements; constantly changing focus or persisting without progress</td>
<td>Slow, but planned movements are reasonably organized</td>
<td>Confident, efficient and safe conduct, maintains focus on task, fluid progression</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Force sensitivity</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rough moves, tears tissue, injures nearby structures, poor control, frequent suture breakage</td>
<td>Handles tissues reasonably well, minor trauma to adjacent tissue, rare suture breakage</td>
<td>Applies appropriate tension, negligible injury to adjacent structures, no suture breakage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Autonomy</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unable to complete entire task, even with verbal guidance</td>
<td>Able to complete task safety with moderate guidance</td>
<td>Able to complete task independently without prompting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Robotic control</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistently does not optimize view, hand position, or repeated collisions even with guidance</td>
<td>View is sometimes not optimal. Occasionally needs to relocate arms. Occasional collisions and obstruction of assistant.</td>
<td>Controls camera and hand position optimally and independently. Minimal collisions or obstruction of assistant</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.3: Global Evaluative Assessment of Robotic Skills (GEARS) [4]
Figure 2.4: The FLS tasks: (i) peg transfer, (ii) pattern cutting, (iii) ligating loop, (iv) extracorporeal suture, (v) intracorporeal suture [5]

portant metric and this assessment does not consider penalizing inadequate maneuvers [31].

Similarly, the Fundamentals of Vascular Surgery (FVS) [32, 33], a standardized training curriculum under development, consists of specific tasks used to prepare surgeons for vascular procedures. Trainees practice (i) radial, continuous suturing on a membrane, (ii) a patch angioplasty, and (iii) end-to-end and/or end-to-side anastomosis on a graft tube (see Fig. 2.5). Performance is assessed by an expert surgeon at a later time via visual analysis of the sutured membrane and graft tube.
2.2 Simulators for Surgical Skill Assessment

The adequate skill of a surgeon is key for a successful surgical operation, as well as associated post-surgery clinical outcomes [11]. Thus, to improve the quality of surgical performance, it is necessary that a surgeon is properly trained. In conventional surgical skill training, expert surgeons observe and provide feedback to novices during exercises. This type of training may be partly subjective since feedback often depends on the expert surgeon’s preferences and style. Further, training draws expert surgeons away from clinical responsibilities [18]. Simulators were developed to address these problems and to standardize and automate assessment of a surgeon’s skill. A medical simulator can be defined as a system that is capable of rendering an aspect of a surgical procedure for the purpose of assessing a subject’s skill on performance and/or training a subject’s skill. One of the main advantages of using simulators is the ability to train surgical skills without the use of animals and cadavers. Another key advantage is the ability to measure skill and skill progression. In addition, simulators allow trainees to repeatedly practice a specific skill. Overall, evidence suggests that surgical simulators could be an effective tool in the training of surgical skills [12, 13].
We classify medical simulators into four categories: Bench Models, Instrumented Bench Models, Virtual Reality (VR) simulators, and Augmented Reality (AR) simulators. Bench models are the most widely used simulators, consisting of either synthetic materials or animal tissue for teaching basic surgical skills like suturing and knot tying. Bench Models are inexpensive and portable, making them convenient to use [34, 35]. One limitation of bench models is that they lack built-in performance assessment. In contrast, Instrumented Bench Models combine sensing modalities with the standard Bench Model to provide performance measures. Instrumented Bench Models provide objective skill assessment and have the potential ability to train users. VR simulators combine advanced sensing methodologies with realistic computer graphics to create a virtual surgical environment. High fidelity anatomical features of the surgical environment are produced with state-of-the-art computer graphics. Data from various sensors in the system (motion, force, etc.) are translated to the virtual environment, and the interaction is fed back to the user. Nevertheless, VR simulators have not achieved the level of realism desired by clinical practitioners, are very expensive, and require significant maintenance [36]. Lastly, Augmented Reality simulators possess traits of both Bench Models and VR simulators in that they capture the physical interaction between real tools and materials using sensors and use this information in skill assessment.

Anastakis and his colleagues have conducted a study of 32 residents to investigate whether or not Bench Models can be used as training simulators for acquiring skills necessary for surgical procedure [37]. The study demonstrated that trainees who were trained via Bench Models saw skill improvement similar to those who trained on cadaver models.

LAPSIM [38] is a surgical simulator with the capability of rendering laparoscopic tasks to train subjects via virtual reality. Researchers have studied the validity of using LAPSIM simulators for training purposes [38]. Studies have shown that LAPSIM simulators are effective in training both basic and sophisticated laparoscopic tasks, and that
O’Toole and his associates [26] developed a VR surgical simulator to investigate the simulator’s ability to train and assess suturing skill specific to open surgery. They recruited 8 vascular surgeons from Boston hospital and 12 medical students from Harvard Medical School for the study to perform a suturing task on the simulator. Their experimental results showed that training on their VR surgical simulator could be used to increase suturing performance. The collected data was successful in classifying different expertise levels of suturing.

ProMIS [39] is an example of an AR simulator, which uses both a VR system and a bench model system together. Similar to LAPSIM, ProMIS simulators are also used to train both basic and complex laparoscopic tasks [39, 40, 41]. In one study, Sickle et al. demonstrated the validity of the ProMIS in training laparoscopic suturing tasks. Data collected from the system were used to determine skill level and to distinguish 5 laparoscopic experts from 5 laparoscopic novices [40].

After a thorough review of the literature on simulators, it was found that there are currently very few simulators for the assessment and training of skills for open surgery. The bench models that exist for open surgery do not provide much training information, as they require self-assessment or assessment from an expert surgeon, resulting in a largely subjective assessment of skill. In addition, VR and AR simulators are not particularly suited for open surgery, as they have not achieved the level of realism desired by experts in the field. To address the limitations of current surgical simulators, we have developed an Instrumented Bench Model for open surgery suturing skill. Instrumented Bench Models can provide objective assessment and/or training of skill in open surgery, made possible by the system construction and the use of data obtained from multiple sensors.
2.3 Metrics for Surgical Skill Assessment

Medical simulators were developed to standardize and automate assessment of a surgeon’s skill. A survey of literature confirms that numerous research studies have focused on identifying descriptive, objective metrics using the data collected from the simulators for assessment of surgical skill in order to improve upon traditional assessment practices.

Metrics used for skill assessment can be grouped into two categories: Product and Process metrics. Product metrics are based on measurements that could be obtained from a final product or output, while process metrics are based on measurements that could only be obtained during task performance [42]. In one task from the Fundamentals of Vascular Surgery (FVS) [32], trainees practice suturing on a membrane (e.g. GoreTex®). Performance is assessed by an expert surgeon at a later time via visual analysis of the sutured membrane. Because assessment occurs afterwards, FVS necessarily involves only product measures, e.g. stitch length, stitch consistency, and accuracy. Process metrics, which are used to quantify skill during performance, can provide more insight into skill assessment and can be beneficial for skill training.

These descriptive, objective metrics can be further classified as: force-based metrics [43, 44, 45, 42, 46, 47, 48], motion-based metrics [42, 48, 49, 50, 51] and image-based metrics [51, 52, 53, 54, 55]. The rest of this section is organized as follows. First, literature on previously studied force-based, motion-based and image-based metrics are reviewed. Then, literature results and opportunities are summarized at the end of the section.

Force-based metrics have played a primary role in establishing distinctions between surgical tasks performed by novices versus those performed by experts. Differentiation between skill levels is achieved via parameters extracted from force measurements collected during surgical tasks [44, 45, 42, 46, 47, 48].

15
In one study, Richards and coworkers [44] placed a force/torque sensor on an endoscopic grasper tool to collect force/torque measurements from five expert and five novice surgeons during a laparoscopic cholecystectomy operation on a pig. Collected force/torque data were analyzed using grand median analysis, and experimental results showed that there is a significant difference in skill level between novices and experts. The magnitude of force/torque applied by novices as a whole was found to be higher than that of experts. In a later study [45], the same research group used the endoscopic grasper tool again. Hidden Markow Models, representing surgical skills determined with force/torque measurements, were developed for each subject. These models were then used to classify subjects of different surgical skill levels.

In another study in [42], Dubrowski et al. used force/torque measurements collected from a Gamma F/T transducer to obtain process measures during laparoscopic suturing. A total of six participants, consisting of surgical residents and fellows, were asked to perform a specific suturing task on a synthetic model. Peak forces and total time spent suturing were obtained using the force data and were used in the differentiation of skill level between residents and fellows.

In addition, Horeman et al. [46, 47, 48] used force measurement to extract metrics to classify the participants’ skill level. In one study in [47], 11 experts and 21 novices were asked to exercise suturing and knot-tying on a bench model. Experimental results demonstrated that the metrics of absolute force and peak force can be used to differentiate between different skill levels.

Motion-based metrics, which are extracted using hand and/or surgical tool motion, prove to be helpful in accurately distinguishing between different skill levels [42, 50, 51, 40, 56, 57, 35, 39]. In a study by Sickle et al. [40], motion data collected from the ProMIS augmented reality simulator were used to establish metrics meaningful in the assessment of the laparoscopic suturing skill. Five experts and five novices participated in the study. Subjects each performed three trials of the suturing task on the ProMIS.
The tool path length, as well as the smoothness of the tool motion, were found to be vital metrics in distinguishing experts from novices.

In research studies by Datta et al. [56, 57], motion data collected from the Imperial College Surgical Assessment Device (ICSAD) were used to obtain metrics meaningful in the assessment of skill in open surgical procedures. Here, motion data were obtained using an electromagnetic motion tracking system. Fifty subjects, classified into four various levels of surgical experience, participated in the study. Subjects were asked to perform the tasks of small bowel anastomosis and vein patch insertion. The total number of hand movements made and the path travel by the hand (path length) were investigated. Results showed that the number of movements made during each of the tasks were significantly different for each of the four skill levels. It was concluded that hand movement can be used as a way to measure skill in open surgery.

Bann and coworkers [35] also inspected the motion profiles obtained from suturing and knot-tying tasks in ICSAD to classify surgeons with varying experience. Participants were asked to perform these tasks at surface level on a synthetic skin pad, and at depth level on an Annexe Art jig. The number of movements and total time to complete tasks were both found to be lower for more experienced surgeons.

Pellen and colleagues [39] investigated the motion profiles obtained via optical trackers placed on the instruments used on the ProMIS laparoscopic simulator. One hundred and sixty subjects participated in the study. The study included three main tasks: laparoscope orientation, sharp dissection, and object positioning. Smoothness of the motion, path length, and task performance time were found to be important metrics in distinguishing skill level.

In a study by Dubrowski et al. [42], the motion profiles obtained from subject’s suturing performance on an artificial artery were investigated. Six surgical residents and seven experienced surgeons were recruited for the study, and electromagnetic markers were used to track hand movement. Rotation of the wrist was found to be insightful,
and was used to distinguish expert surgeons from novices.

In one study by Sanches et al. [50], acceleration of hand movement was investigated to extract metrics which were used for assessment of skill. Hand movement was obtained using an iPod Touch, which was placed on the subjects’ wrist. Subject’s hand movement was recorded by the Accelerometer Data Pro application. 8 experts and 5 novices were recruited for the study. Average and maximum hand accelerations were both found to be important parameters, and each was used to distinguish expert surgeons from novices.

The use of computer vision in surgical applications has been explored. For example, Iyer et al. [58] used a computer vision application to establish an automated robot arm for suturing. This study, however, required modification of the needle’s color to ensure needle detection during the procedure, which is not ideal in a real procedure. In another study [59], computer vision was used to track and estimate the pose of a suturing needle in real-time. Further, computer vision has also been used to extract image-based metrics as a means to quantify surgical skill [51, 52, 53, 54, 55].

In one study in [51], Dosis et al. obtained synchronized video and motion data to assess surgical skill during laparoscopic surgery. For this study, 1 experienced and 4 inexperienced surgeons were recruited. Each asked to perform 10 laparoscopic cholecystectomies. Metrics of task completion time, path length, number of movements, velocities, and trajectories were obtained and used to distinguish between skill levels.

Frischknecht et al. [52] utilized an image analysis program on photographs taken post-procedure to assess suturing performance. Metrics that proved most meaningful in ranking the quality of suturing included the number of stitches, stitch length, total bite size, and stitch orientation.

Similarly, Islam et al. [53, 54, 55] attempted to accurately attain surgical skill utilizing computer vision application. Unique to their study was the tracking of hand motion via images during a simulated surgical task on the Fundamental of Laparoscopic
Table 2.1: Literature Review Metrics used in Suturing Skill Assessment in Minimally Invasive Surgery (MIS) and Open Surgery (OS)

<table>
<thead>
<tr>
<th>Surgery Type</th>
<th>Metric Type</th>
<th>Data Source</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimally Invasive Surgery (MIS)</td>
<td>Process</td>
<td>Force</td>
<td>[43, 46, 47, 48]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Motion</td>
<td>[49, 50, 51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Image</td>
<td>[51, 52, 53, 54, 55]</td>
</tr>
<tr>
<td></td>
<td>Product</td>
<td>Image</td>
<td>[52, 53, 54, 55]</td>
</tr>
<tr>
<td>Open Surgery (OS)</td>
<td>Process</td>
<td>Force</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Motion</td>
<td>[35, 42]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Image</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Product</td>
<td>Image</td>
<td>None</td>
</tr>
</tbody>
</table>

Surgery trainer, and the measuring of motion smoothness in order to differentiate between skill levels.

Metrics from the literature are categorized according to surgery type, metric type and data source in Table 2.1. Extraction of process metrics from video is challenging, especially in the case of open surgery, which lacks the constrained environment of MIS. Needle motion contains extensive information about the suturing process, but to our knowledge, the use of computer vision for computing open surgery image-based product and process metrics remains unexplored (see Table 2.1). We have designed a simulator that captures video from below the suture membrane to facilitate extraction of process and product metrics based on needle motion as viewed from below. These metrics were based on the physics of needle insertion forces, following the curve of the needle while driving through tissue, and minimizing lateral forces and motions that induce tear [60, 61, 62]. Motivation for this study arose from gaps found in research on open surgery skill assessment and can be seen in Table 2.1.
The suturing simulator (see Fig. 3.1) was inspired by the need for objective skill assessment in open surgery. The system processes of the suturing simulator are categorized into two main stages: (i) Data Collection, and (ii) Data Processing (Fig. 3.9).
In the Data-Collection stage, the system synchronizes and logs force, motion, video, and touch data during suturing. The Data-Processing stage uses the collected data to extract metrics of suturing skill. Details about the physical components of the suturing simulator, data-collection stage, and data-processing stage are presented in the following sections.

3.1 Physical Components of the Suturing Simulator

3.1.1 Membrane Housing

The cylindrical membrane housing was constructed from clear acrylic and its sides were shielded externally with an aluminum sheet. The aluminum sheet eliminates external lights which may negatively affect video frame rendering during the image processing stage. Eight metal latches along the upper exterior of the membrane housing were used to secure the suturing membrane, (e.g., GoreTex®, artificial leather or other fabric) on which suturing is performed (Fig. 3.2a). The purpose of using latches is to allow for the quick and easy replacement of the suturing material once suturing is completed.

A transparent acrylic cylinder was placed around the membrane housing. An anti-backlash ball-screw was mounted between the cylinder and a stepper motor (Mercury Motor). The level of the cylinder was controlled by this stepper motor, allowing for the height of the cylinder to be set at a desired depth. The stepper motor was operated with a stepper motor driver (Big Easy Driver, Sparkfun Electronics). An Arduino was used to control the system. This setup was constructed so that our simulator could achieve real-life situations where suturing is required at some depth inside a body, with limited hand motion, similar to study in [63].
3.1.2 Suturing Membrane

The suture membrane (see Fig. 3.2b) was designed such that suturing is performed in a radial and uninterrupted fashion, similar to the radial suturing task in the Fundamentals of Vascular Surgery [25, 32].

The suture membrane (see Fig. 3.2b) was designed such that suturing is performed in a radial and uninterrupted fashion, similar to the radial suturing task in the Fundamentals of Vascular Surgery [25, 32]. A circle, representing an incision, was drawn on the membrane. The circle was partitioned by radial lines into equal sections each spanning $30^\circ$. Needle entry points were marked on the radial lines based on the diameter of the needle used. The marks indicated where suturing was to be performed (entry on one side, exit on the other). All membranes were made of artificial leather using a laser cutter.
3.1.3 Height Adjustable Table

The membrane housing was mounted onto an adjustable height table. This allows subjects to set the height of the platform as desired for comfort during suturing [27]. Ergonomic studies of the height of operating tables show that the optimum height of the table lies between 55 cm and 100 cm from the floor up to table surface [64, 65, 66]. The table for the suturing simulator was modified to permit heights between 71 cm and 99 cm (see Fig. 3.3).

3.1.4 Sensors

3.1.4.1 Physical Touch Sensor

Capacitive sensing is a widely applied technology that is utilized to detect and/or measure conductive things, or things that have a different dielectric constant than air [67, 68]. Proximity sensors, human interface devices, cell phones, and tablets are all application examples where capacitive sensing has been used. Using this idea, a capacitive sensing application was employed to detect any physical contact between either
the subjects’ body or the surgical instrument, and the cylinder. To achieve this, the interior and top of the transparent cylinder were lined with flexible conductive films (Indium Tin Oxide coated plastic sheet) and thick aluminum foil, respectively. Connection between the films and the foil were achieved using copper foil adhesive tape. 100 kΩ resistance was chosen in order to make the sensor sensitive enough to detect contacts between a human hand (or surgical instrument) and the conductive materials. All of the connections were made through an Arduino in order to obtain and log the touch data. During testing, it was found that the conductive sensor values were unstable, due to their dependency on the environment and surrounding things. Interference with sensor readings occurred when surrounding materials and the number of people changed near the system. Therefore, the exterior wall of the cylinder was also coated with flexible conductive film and was grounded. Grounding the external coating eliminated any possible external affects, and allowed for stable and accurate sensor readings. Connection diagram for the physical touch sensor is shown in Fig. 3.4.
3.1.4.2 Internal and External Cameras

The membrane housing, shown in Fig. 3.5, was designed in a unique way, with a camera positioned beneath the membrane to record a video of needle and thread movement during suturing. In an earlier stage of the design, a Logitech C920 HD Pro webcam used to record the video. This camera has the capability of recording video with 1080p resolution at 30 fps. However, after conducting multiple experiments with expert surgeons, it was realized that recording the internal video at 30 fps was not fast enough. During suturing, certain surgeons were performing so quickly that the Logitech could not record all needle movement, and thus detection of the needle failed during the image processing stage. Therefore, the Logitech camera was replaced by a Firefly MV USB 2.0 camera (PointGrey Inc.), which has the capability of recording video with 640x480 resolution at 60 fps.

During the earlier stage, the membrane housing dimensions were calculated and designed with consideration of the horizontal field of view (fov), as well as Logitech’s focal length (fc= 3.67mm), so that the entire membrane would be captured in the images [69]. In order to use the same membrane housing design with the Firefly camera, a Fujinon CS-mount lens with a 2.8 mm focal length was integrated to allow for the complete
capture of the membrane on the scene. In addition, a white led strip was placed inside the membrane housing for enhancement and stabilization of lightning (Fig. 3.5). Use of the led strip also aids in accurate needle and thread detection during the process of the video data via computer vision algorithm.

To record membrane and subject’s hand movement during suturing, a Logitech C920 HD USB 2.0 camera was used as an external camera. The external camera was placed on a tripod located in front of the suturing platform. The Logitech has the capability of recording video with 640x480 resolution at 30 fps. Using synchronized external video data and touch data, the locations of each contact made between the subject and the system during suturing were obtained and verified. In addition, the external video recording allows for the suturing exercise to be revisited at any time in the future for additional investigation and analysis.
3.1.4.3 Force/Torque Sensor

A 6-axis force/torque sensor (ATI MINI 40, ATI Industrial Automation Inc.) was placed under the housing to measure forces and torques applied to the membrane during suturing (See Fig. 3.7). The multi-axis force/torque sensor is capable of measuring the forces and torques in both positive and negative directions of x, y and z coordinate system [70]. A M-series National Instruments Data Acquisition Card (NI-DAQ), which was connected to the PCI slot of the PC, was used for the data acquisition between the sensor and the PC.

3.1.4.4 Motion Sensor

To record the wrist motion of subjects during the suturing task, an InertiaCube4 sensor (InterSense Inc., MA) was used. It has the capability of measuring roll, pitch, and yaw with 0°- 360° range, with the maximum frequency of 200 Hz. The yaw direction accuracy is 1°, and pitch and roll direction accuracy is 0.25° [71].

3.2 Data-Collection Stage

Data were collected from the four sensing modes: force/torque, motion, video, and physical contact. Force/torque data were collected using the 6-axis force/torque sensor and logged at 1 kHz during suturing. To obtain force/torque data from the sensor,
software was written using the NI-DAQ Software Development Kit (SDK). Collected force/torque data were filtered with a 10th-order Butterworth lowpass filter with a cutoff frequency of 50 Hz to remove noise and smooth the data. To record hand motion, the InertiaCube4 sensor was placed on the dorsum of the subject’s dominant hand as shown in Fig. 3.8 and logged at 200Hz during suturing. InterSense SDK was used to obtain $\theta_{\text{yaw}}$, $\theta_{\text{pitch}}$, and $\theta_{\text{roll}}$ measurements of the subject’s wrist motion. The internal camera with FlyCapture SDK was used to record needle and suture motion from under the membrane at 60 fps. The external camera was used to record membrane and hand movement at 30 fps. An open source computer vision library (OpenCV 3.0.0) was used to capture and log the external video. For logging touch data, the Arduino capacitive sensing [72] and serial communication libraries [73] were used.

The system presented here extends an earlier version of the suturing platform presented in [74, 75] that featured a single external camera, a force sensor, and a motion sensor for the collection of necessary data. We modified the previous system to include an internal camera to record needle and thread movement. This enables the extraction of vision-based metrics. In the previous system, synchronization of the data stream was achieved in post-processing, whereas our current platform synchronizes data collection on a single PC using a multithreaded implementation and timestamping. The Data
Collection Stage software was written in C++ using Microsoft Visual Studio 2013 [76].

During suturing, all unprocessed (raw) data is synchronized and logged. Logging allows for revisiting the raw data at any time for additional investigation and analysis. The raw data were then used in the Data-Processing stage.

### 3.3 Data-Processing Stage

The main objective of this stage was to use collected raw data to extract metrics that are meaningful in the assessment of suture skill in open surgery. In this stage (see Fig. 3.9), internal video was first processed with a computer vision algorithm to obtain information about needle and thread movement (explained in Chapter 4). This information was then used to identify the individual suture cycles. Next, the data were used to extract metrics for each time the subject is actively suturing (explained in Chapter 5). Finally, each of the metrics was statistically analyzed to investigate their ability to distinguish between varying levels of expertise (explained in Chapter 6). In the following chapters, each of these steps will be explained in detail.
Figure 3.9: System process flow-chart, consisting of two stages. In the Data Collection Stage, raw data from multiple sensors were synchronized and logged. In the Data Processing Stage, the collected data were used to extract metrics for suturing skill.
Chapter 4

Computer Vision Algorithm

The computer vision algorithm involves two successive stages: 1) Image Processing and 2) Metrics Extraction. In the Image Processing stage, the following information is obtained: (i) segmentation of the needle and thread, (ii) tracking of the needle tip and swage, and (iii) detection of needle entry and exit points and times. In the Metrics Extraction stage, the needle information is used to compute metrics for skill assessment. Software for the image processing stage was written in C++ in Microsoft Visual Studio 2013 using the open source computer vision library (OpenCV 3.0.0). Software for the metrics extraction stage was written in MATLAB 2015a software. A flow-chart of the algorithm is provided in Fig. 4.1. In the following section, image processing stage is explained in detail. Metrics extraction stage is presented in Chapter 5.
Figure 4.1: The algorithm consists of two stages: In the Image Processing Stage, the needle and thread are detected and needle entry and exit points are identified. In the Metrics Extraction Stage, metrics were computed based on information from the Image Processing Stage.
4.1 Image Processing Stage

To achieve the detection of both the needle and thread, which is the primary objective of the Image Processing Stage, the video recorded beneath the membrane was processed with an extended version of the computer vision algorithm presented earlier in [69] and [77]. The image processing stage was carried out in the following manner.

Using a pre-computed camera calibration, the algorithm warped each frame to correct for lens distortion. After this correction, contrast and brightness of each frame are adjusted to enhance image quality (see Fig. 4.2). Static marker locations (two red and two green, as seen in Fig.4.3), each with relative known distance (121 mm), are
detected. The markers are used to calibrate the pixel-to-mm conversion rate \( k_{\text{cam}} = 0.19 \) pixels/mm. The markers are also used to identify the membrane center as the intersection point of lines drawn between opposing markers. Next, the image is masked to focus attention on a circular region of interest where all needle and thread movement appear. Each frame is converted from RGB color space to HSV color space. Separate threshold values are applied for (i) detection of the needle and (ii) detection of the needle with thread. Arithmetic and morphological operations are used to segment the thread from the needle and to remove noise. After detection, the needle is enclosed with a green circle, and the thread is marked with blue.

After the needle is detected under the membrane, the end points of the segmented needle are detected as the intersection points of the visible needle and the minimum enclosing circle. Frame-to-frame differences in the needle end points are used to distinguish the needle tip from the entry location and to distinguish the needle swage from the exit location. The needle swage is the point where the thread is attached to the needle. When
the needle first enters the membrane, the needle entry point is recorded and assigned as the needle tip. For each subsequent frame, the distances from the entry point to each of the endpoints of the visible needle are calculated and compared. During the driving phase of the suture process, the needle tip moves away from the needle entry point. Therefore, the furthest endpoint from the needle entry point is identified as the needle tip. Later, after the needle begins to exit, the needle swage moves away from the needle entry point. This coincides with thread appearing in the frame. Hence, when thread is detected, the point closest to the needle entry is assigned to the needle swage. Separate pixel trajectories are recorded for the needle tip and for the needle swage. An example of needle tip and needle swage trajectory detection are illustrated in the Fig.4.9.

Several of the metrics presented in the next section are computed from the needle tip trajectory. Denote the list of pixels on the needle tip trajectory for a given suture location as

\[
\tilde{P} = [\tilde{p}_1 \; \tilde{p}_2 \; \ldots \; \tilde{p}_n]
\]  

(4.1)

where \(\tilde{p}_j = \left[ \tilde{x}_j \; \tilde{y}_j \right]\) are coordinates of the \(j^{th}\) pixel in the list. Note that the length \(\tilde{n}\) of the pixel list in (1) may vary from suture to suture.
Figure 4.7: Needle Tip Trajectory: Pixel list of the needle tip filtered and weeded to be used in post-processing to compute performance metrics.
To reduce the noise in the needle tip trajectory, the pixel values are smoothed as follows: (1) The path is first filtered by a 2nd order Butterworth low-pass filter with a 15 Hz cut-off frequency and (2) The list of filtered pixel values representing the needle tip trajectory was weeded; that is, pixels were removed from the list to guarantee that the Euclidean distance between any two sequential pixels in the resulting list was at least 2.0 pixels, corresponding to 0.38 mm. The filtered and weeded pixel list for the needle tip trajectory is denoted as

\[ P = [p_1 \ p_2 \ \ldots \ p_n], \quad \text{where} \quad p_j = \begin{bmatrix} x_j \\ y_j \end{bmatrix}. \]  

(4.2)

Note that due to weeding, \( n \leq \tilde{n} \). Weeding the pixel list reduces jitter and improves stability of the computed metrics (see Fig.4.7). The filtered, weeded pixel list is used in post-processing to compute performance metrics.

In addition, needle entry and exit times are determined. When the needle enters the membrane, the time is recorded as the needle entry time, \( t_{en} \). Similarly, when the needle completely exits the membrane, the time is recorded as the needle exit time, \( t_{ex} \).
Figure 4.8: Steps for detection of needle and thread; First, camera distortion was corrected and image visible quality was enhanced. Then, frame was masked to obtain region of interest and conversion from RGB to HSV were achieved. Next, needle and thread were detected via vision algorithm and detections were superimposed to the original image.
4.2 Computer Vision Algorithm Results

Steps for detection of the needle and thread are shown in Fig 4.8. Illustration of consecutive frames for detection of needle and thread can be seen in the Fig. 4.9a. First, the needle tip appears and is detected. Then, following frames illustrate the detection of the needle body, which travels from entry to exit point. Next, the thread tailing the needle appears in the frame and is detected. Thread detection is illustrated with the color blue, while needle detection is denoted with a green colored circle surrounding the needle. When the needle exits the membrane completely, corresponding entry and exit points for that suture are obtained and illustrated by yellow and pink colored points on the frame, respectively.

Similarly, successive frames for detection of the needle tip and swage traces are illustrated in the Fig. 4.9b. From entry point to exit point, needle tip points are obtained, and the corresponding tip trace is drawn on the frames, as illustrated with by color red. Then the needle swage (location where the needle and thread are attached) is detected, and the corresponding pixel points are attained. From these points, the swage trace is then drawn on the frames with the color yellow.
Figure 4.9: Computer vision frame-by-frame display of a) needle and thread detection along with the entry and exit points; b) needle tip path (in red), and swage path (in yellow); c) needle swept area (in black) with the needle detection (in green)
Chapter 5

Metrics for Suturing Skill Assessment

5.1 Vision-Enabled Partitioning of Suture Cycle

During continuous suturing, a single suture cycle can be divided into two distinct periods of time: active suturing time and idle time. Active suturing time is the time between needle entry into the tissue and complete needle removal from the tissue. Idle time is the time between the end of one active suturing time to the start of the next. In other words, active suturing is the time taken by subjects to complete one suture, whereas idle time is the time spent preparing for the next suture. Active suturing time may be further decomposed into 4 phases: a) entry phase – puncturing the needle into the tissue; b) driving phase – driving the needle along some path inside the tissue; c) exit phase – exiting the needle tip from the tissue; and d) pull-out phase – pulling the needle completely from the tissue and then tightening the thread.
Figure 5.1: Graphical User Interface (GUI) designed to show the synchronized force and motion data, along with videos from external and internal cameras. GUI allows for convenient, interactive investigation of synchronized data.
Dividing each suture cycle into distinct phases allows for context-specific interpretation of the sensor data. Needle entry and exit times obtained from the computer vision algorithm were used to extract each suture cycle for individual analysis. In addition, a Graphical User Interface (GUI) in MATLAB (Fig. 5.1) was created to display synchronized force, motion, and touch data, as well as video from external and internal cameras. The interface also labels the needle entry, needle exit and thread entry times automatically determined by computer vision. The interface enables convenient, interactive exploration of the synchronized data. An example of synchronized data for one active suturing time with the suture sub-events identified (entry, driving, exit and pull-out phase) is shown in Fig. 5.2.
Figure 5.2: Example of synchronized force, torque, motion, and touch data for one active suturing time with suture sub-events labeled. A suture cycle is comprised of active suturing time and idle time. Active suturing time is the total duration of entry, driving, exit and pull out phases, whereas idle time (not shown here) is the time between an end-time of one suture to start time of another. (Note: Blue diamond symbol (▵) in touch data indicates the time instance of the physical touch)
5.2 Extracting Metrics from Time-Series Data

During suturing, the system collects synchronized time-series data from multiple sensors. Many of the metrics presented in this paper are computed from time series data of a scalar signal \( X(t) \) using one of the following functions:

\[
\text{PEAK}_+(X) = \max_t (X(t)) \tag{5.1}
\]

\[
\text{PEAK}_-(X) = \max_t (-X(t)) = -\min_t (X(t)) \tag{5.2}
\]

\[
\text{PP}(X) = \text{PEAK}_+(X) + \text{PEAK}_-(X) \tag{5.3}
\]

The time interval over which the maximum is taken is specified in the definition of the specific metric. Typically the time interval corresponds to one whole active suture time. Note that \( \text{PEAK}_+(X) \) is the maximum value that signal \( X \) took over the time interval and \( \text{PEAK}_-(X) \) is the negative of the minimum value that signal \( X \) took during the time interval. If signal \( X(t) \) is negative at some point, then \( \text{PEAK}_-(X) \) can be interpreted as the magnitude of peak negative value of \( X(t) \). \( \text{PP}(X) \) is the peak-to-peak amplitude of signal \( X \).

5.3 Force/Torque-based, Motion-based and Physical Contact Metrics

5.3.1 Force/Torque-based Metrics

For each active suturing time, (5.1) - (5.3) were used to compute metrics based on time series for force components \( F_x, F_y, \) and \( F_z \), and torque components \( T_x, T_y, \) and
Based on the coordinate axes (shown in Fig. 5.6), \( \text{peak}_+ (F_z) \) is the maximum force component applied upward on the membrane while \( \text{peak}_- (F_z) \) is the maximum force component applied downward.

### 5.3.2 Motion-based Metrics

Metrics on total range of hand motion were extracted from IMU orientation data using (5.3), specifically \( \text{PP} (\theta_{\text{yaw}}) \), \( \text{PP} (\theta_{\text{pitch}}) \) and \( \text{PP} (\theta_{\text{roll}}) \) for each active suturing time.

### 5.3.3 Physical Contact Metric

The capacitive touch sensor was used to identify and count each instance of physical contact between the subject and the top and/or internal wall of the cylinder around the membrane holder. The total number of touches \( (C_n) \) made during a suture cycle is used as a metric.

### 5.4 Image-based Metrics

Distances from optimal entry point and distance from optimal exit point, called *Entry Distance* \( (d_{oe}) \) and *Exit Distance* \( (d_{ox}) \) from here on, are measurements of performance accuracy and were calculated using Euclidean distance (indicated by \( || \ || \)) between two points as

\[
d_{oe} = k_{\text{cam}} ||p_1 - p_{oe}|| \tag{5.4a}
\]

\[
d_{ox} = k_{\text{cam}} ||p_n - p_{ox}|| \tag{5.4b}
\]

where \( p_{oe} \) and \( p_{ox} \) are the predetermined optimal entry and exit locations as marked on the suture membrane, respectively, and \( p_1 \) and \( p_n \) are the needle entry and exit locations from the filtered, weeded pixel list (4.2). The calibration parameter \( k_{\text{cam}} \) converts the
metric to physical units to make the metric independent of the specific camera and optics used in our suturing simulator.

*Stitch length* \((l_s)\) is the length of the stitch and was calculated from the needle entry and exit points for each suture

\[
l_s = k_{cam} ||p_n - p_1||. \tag{5.5}
\]

*Stitch time* \((t_s)\) is the time to complete a single suture, from needle entry \((t_{en})\) to needle exit \((t_{ex})\). Stitch time was calculated as

\[
t_s = t_{ex} - t_{en}. \tag{5.6}
\]

Similarly, *Idle time* \((t_d)\) is the time between needle exit on one suture \((t_{ex}^i)\) and needle entry on the next suture \((t_{en}^{i+1})\). Idle time after suture \(i\) was calculated as

\[
t_d^i = t_{en}^{i+1} - t_{ex}^i \quad \text{for} \quad 1 \leq i \leq 11 \tag{5.7}
\]

where \(i\) indicates the suture number. Idle time captures a subject’s preparation time for the subsequent suture.

In addition to the previously studied metrics described above, the system computed four new image-based process metrics, described in this paper for the first time. All four of these process metrics are inspired by expert surgeons’ recommended best practice for suturing: “follow the curvature of the needle” [19]. Driving the needle along a path that follows the curvature of the needle minimizes tissue trauma and eases penetration into the tissue [78, 61]. The new metrics *Needle Tip Path Length*, *Needle Tip Area*, *Needle Swept Area* and *Needle Sway Length* are described in detail below.
5.4.1 Needle Tip Path Length

Following the curve of the needle implies that the needle tip, as viewed by the camera under the membrane, should move directly from the entry point to the exit point with minimal lateral motion. *Needle Tip Path Length* metric \((l_{tp})\) is the length of the path followed by the needle tip from entry to exit. Intuitively, shorter path lengths indicate greater skill, longer paths indicate that the needle is straying or wiggling from the ideal path. In practice, path lengths become especially long if the needle holder is repeatedly repositioned on the needle. Needle Tip Path Length is computed as the sum of Euclidean distance between sequential pixels in the filtered, weeded tip trajectory (4.2), specifically,

\[
l_{tp} = k_{cam} \sum_{j=1}^{n-1} \|p_{j+1} - p_j\|. \quad (5.8)
\]

5.4.2 Needle Tip Area

Needle tip area is defined as the absolute area between the needle tip path and the straight line from entry point to exit point. In other words, this metric is a measure of how much the needle tip deviates from the straight line path from the needle entry...
point to exit point, with larger deviations penalized more. This metric is designed to
penalize motion of the needle tip which is orthogonal to the direction of the stitch. The
needle tip area is calculated using the filtered, weeded pixel list (4.2) as
\[ a_t = k_{\text{cam}}^2 \sum_{j=1}^{n-1} |(p_{j+1} - p_j)^T \vec{e}_t| \cdot \frac{1}{2} |(p_{j+1} + p_j)^T \vec{e}_o| \]  
(5.9)
where \( \vec{e}_t \) is the unit vector tangential to stitch direction, i.e. in the \((p_n - p_1)\) direction,
pointing from entry point to exit point, and \( \vec{e}_o \) is the unit vector orthogonal to \( \vec{e}_t \). An
example of the incremental needle tip area due to two sequential pixels of the needle tip
trajectory is shown in Fig. 5.3.

### 5.4.3 Needle Swept Area

Needle swept area is the union of all area covered by the needle body during
suturing. Needle Swept Area will be high if the needle rolls during suturing, even if
the tip does not deviate from the straight line between entry and exit. To compute
Needle Swept Area, all pixels corresponding to the portion of the needle visible below
the membrane from each video frame in the active suturing time are superimposed onto
a single binary image. The total number of “on” pixels in this image is \( \hat{n} \). The Needle
Swept Area in mm² is computed as
\[ a_s = k_{\text{cam}}^2 \hat{n}. \]  
(5.10)
Visualization of the needle swept area can be seen in the last row of Fig. 4.9. Needle
body movement is superimposed for each consecutive frame during the suturing exercise
to obtain the total area swept by the needle. To promote visibility, black represents the
total swept area of the needle while the detection of the needle body is overlaid in green.
Figure 5.4: A plot of the sway length $l_{sw}(t)$ for one suture cycle, along with example images illustrating the needle with positive (blue) and negative (green) orientation.
5.4.4 Needle Sway Length

This metric was designed to measure roll of the needle, i.e. rotation about the axis connecting needle tip and swage, during suturing. Needle Sway Length is computed as follows. After needle detection, the end points of the visible portion of the needle are identified by finding the two points where the needle meets the minimum enclosing circle. The midpoint of the chord connecting the two end points is calculated. The instantaneous needle sway length, \( l_{sw}(t) \), is the signed distance from the midpoint of the chord to the needle body, along a line orthogonal to the chord. The instantaneous needle sway length is stored for each frame (see Fig. 5.4). The Needle Sway Length metric (\( l_{swm} \)) was calculated as

\[
l_{swm} = \max(l_{sw}(t)) - \min(l_{sw}(t)).
\]

This metric captures the maximum deviations in the roll of the needle during a suture. Rolling the needle during a suture violates the maxim to follow the curve of the needle and may result in tissue damage. It was hypothesized that shorter sway length corresponds to less tear to the tissue and a path closer to the optimal needle trajectory through the tissue.

5.4.5 Relationship between Image-based Metrics

The metrics Needle Tip Path Length, Needle Swept Area, Needle Tip Area, and Needle Sway Length are all related to the motion of the needle, but capture distinct aspects of the motion. To illustrate similarities and differences, we will discuss a set of thought experiments, illustrated in Fig. 5.5.
Figure 5.5: Frame-by-frame illustration of (a) Needle position (in gray); (b) Cumulative area swept out by the needle body from first frame to current frame (in blue); and (c) Needle tip trajectory from first frame to current frame (in black dashed line). Needle Swept Area metric is the area of the blue in the last frame. Needle Tip Path Length metric is the entire length of the black dashed line in the last frame. Needle Tip Area metric is the total area between the straight-line from entry to exit (red dashed line) and the needle tip path (in green) in the last frame.
First, consider a suture in which the needle tip frequently makes a small deviation from the straight line path from entry to exit (see Fig. 5.5, images c1 - c7). Here, Needle Tip Area will be insignificant, suggesting high skill level. However, this information alone may be misleading since Needle Tip Path Length could be large here, indicating lower skill level.

Second, consider Fig. 5.5 a7 - a11, in which the needle rolls back and forth about the chord connecting the needle tip to the exit location. Needle Swept Area increases significantly due to this motion, but Needle Tip Path Length does not increase significantly, nor does Needle Tip Area.

Third, consider a case in which both Needle Swept Area and Needle Tip Area are large. In this case, Needle Sway Length will identify the underlying reason, roll or yaw of the needle, most responsible for this large area.

5.5 Image-enabled Metrics

5.5.1 Orthogonal and Tangential Forces

Force applied orthogonal to the stitch direction may increase tissue tearing and should therefore be minimized. The axes of the force sensor are not generally aligned with the directions of the radial stitches, so a change of coordinates is required to determine the force components orthogonal and tangential to the stitch direction. Using the suture entry and exit points detected by computer vision, the suture direction at each suture location can be identified. Then, a change of coordinates can be applied to compute the force tangential to stitch direction and orthogonal to stitch direction (see Fig. 5.6). Calculations of the orthogonal and tangential forces were achieved as follows.

Total force, \( \vec{F} \), can be expressed in the vision coordinate system as:

\[
\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y
\]
where $F_x$ and $F_y$ are the component forces in x and y direction, respectively, as read from the force sensor, and $\vec{e}_x$ and $\vec{e}_y$ are the unit vectors in the vision coordinate frame aligned with the x- and y- axes of the force sensor, respectively. Since the coordinate system of the force sensor is constant, $\vec{e}_x$ and $\vec{e}_y$ were also constant, independent of suture location. The unit vectors $\vec{e}_x$ and $\vec{e}_y$ were precomputed based on a calibration experiment considering force sensor axis.

The same force can also be represented as

$$\vec{F} = F_o \vec{e}_o + F_t \vec{e}_t$$  \hspace{1cm} (5.13)$$

where $F_o$ and $F_t$ are the component forces orthogonal and tangential to the stitch direction in vision coordinate frame, respectively, and $\vec{e}_o$ and $\vec{e}_t$ are the corresponding unit vectors in the vision coordinate frame.

Thus, (5.12) and (5.13) can be rearranged as follows to obtain orthogonal and tangential component forces, $F_o$ and $F_t$:

$$\begin{bmatrix} F_o \\ F_t \end{bmatrix} = \begin{bmatrix} \vec{e}_o & \vec{e}_t \end{bmatrix}^{-1} \begin{bmatrix} \vec{e}_x & \vec{e}_y \end{bmatrix} \begin{bmatrix} F_x \\ F_y \end{bmatrix}. \hspace{1cm} (5.14)$$

Contrary to $\vec{e}_x$ and $\vec{e}_y$, the direction of unit vectors $\vec{e}_o$ and $\vec{e}_t$ depend on suture location. The vectors $\vec{e}_o$ and $\vec{e}_t$ are calculated from the suture entry and exit points, whose values are obtained using the computer vision algorithm in Chapter 4.

Using the aforementioned calculations, orthogonal and tangential forces for each suture location were obtained. For each active suturing time, (5.1) - (5.3) were used to compute metrics based on $F_o$ and $F_t$. 

54
Figure 5.6: Decomposition of horizontal forces into forces orthogonal and tangential to stitch direction: (a) view of the needle taken from internal camera with force sensor coordinate system overlaid; (b) side view of the needle along with orthogonal and tangential force direction; (c) zoomed in view of the needle, together with the force sensor, orthogonal and tangential forces; (d) X and Y directional forces at suture location in (a), for one active suture time; (e) corresponding orthogonal and tangential forces from (a), for one active suture time.
5.5.2 Entry Force and Angle

For the specific skill of surgical suturing, general guidelines suggest that surgeons puncture the tissue to be sutured perpendicular to the surface, i.e., the angle between the needle tip and tissue should be $90^\circ$. The surgeon should then “follow the curvature of the needle” during suturing, which results in minimum tissue trauma [60, 61, 79, 80, 81]. These guidelines can be used in learning as well as assessing suturing skill of medical students, residents, or attendings. To date, few researchers have examined needle entry angle for quantifying surgical skill [82, 83, 26]. The existing work either categorized entry angle qualitatively using human review of video or computed by wrist movement data collected from virtual reality simulators. Wrist movement and needle movement are not perfectly correlated since surgeons may use their fingers together with their wrist to achieve certain needle motions. A more robust method to quantitatively estimate needle entry angle during suturing could be useful for providing objective feedback during training. In this thesis, we present a method to calculate the needle entry angle via integrated force and vision data. To our knowledge, this is the first study to calculate the needle entry angle using vision and force data together.
At the beginning of every suture, as the needle begins to enter the membrane, the z-directional force starts to increase. The computer vision system detects the needle only when the tip has fully penetrated the membrane. Therefore, at the beginning of every suture, there is a time interval between the initial application of z-directional force and detection of needle entry by computer vision. This time interval is defined to be the needle entry phase (see Fig. 5.7b). Since the z-directional force exists before needle penetrates to the membrane, it is necessary to obtain the time instance at which this force begins, i.e., start point. The start point of the entry phase for each of the suturing time was calculated as follows. In each suture cycle, the algorithm works backward from the needle entry time to find a time where the z-directional force falls below 0.01N. This force threshold was chosen based on force sensor noise characteristics. This time
is assigned as the start point. After obtaining the start point, the duration between the start point and needle entry point, i.e., the needle entry phase duration, was calculated (see Fig. 5.7b).

Lateral forces indicate tearing and participants should apply minimal forces in that direction during suturing [61, 84]. Since the force sensor is fixed on the platform housing, x, y and z forces applied to the membrane are measured with respect to the sensor axis; the force directions remain the same regardless of suture location. The orthogonal (\(F_{ort}\)) and tangential (\(F_{tan}\)) force components in the plane of the membrane are computed using a change of coordinates. That is, internal video processed via the computer vision algorithm in Chapter 4 allows us to obtain the suture direction for each suture location. Using this suture direction, a change of coordinates was applied to compute the components of the force tangent to the stitch direction and orthogonal to the stitch direction.

The tangential force and the z-directional force (\(F_z\)) were used to derive entry force (\(F_e\)) and entry angle (\(\alpha_e\)) using the following calculations (see Fig. 5.7a):

\[
F_e(t) = \sqrt{(F_z(t))^2 + (F_{tan}(t))^2}
\]

\[
\alpha_e(t) = \arctan(F_{tan}(t)/|F_z(t)|)
\]

An example of the force and entry angle is shown in Fig. 5.8. The instantaneous entry force and instantaneous entry angle are used to compute the following metrics:

\[
F_e = \frac{1}{N} \sum_{i=1}^{N} F_e(i),
\]

\[
F_{e\text{max}} = \max_{t_{25} \leq t \leq t_{75}} F_e(i),
\]
where $\bar{F}_e$ and $F_{max}$ are the mean and maximum of needle entry force, respectively, and $\bar{\alpha}_e$ is the mean of needle entry angle. $N$ is the total number of data point falling between the 25th percentage ($t_{25}$) and 75th percentage ($t_{75}$) of the entry phase for each suture. Outside of this time range, the signal-to-noise ratio was often too small to produce reliable angle calculations (see Fig. 5.8).

**Validation of Needle Entry Angle**

To verify the needle angle calculation, an experiment was first designed using a straight needle. The external camera was positioned on the side of the system and aligned with the membrane level to record the video. Following this, the straight needle was inserted into the membrane with specific angles. Vision-based analysis was used to calculate the entry angle from video frames and compared to the entry angle from forces. Examples of video frames with needle entry angles are shown in Fig. 5.10 while comparison of these video- and force-based entry angles is shown in Fig. 5.9. The Pearson’s correlation coefficient was 0.99 with a p-value less than 0.01.
Figure 5.10: Example of video frames used to calculate entry angles from the external video for certain needle entry angles: (i) 25° (ii) 45° (iii) 70° (iv) 90° (v) 115° (vi) 145°. Here, a straight needle was used. The calculated angles were then compared to $\alpha_e$, i.e. the image-enabled force-based angle calculation.
Chapter 6

System Validation and Experimental Results

6.1 Experimental Setup and Protocol

Ethics approval for the study was obtained from the applicable Institutional Review Board (Reference # Pro00011886). A total of 15 subjects (6 Attending Surgeons, 8 Surgery Residents and 1 Medical Student) were recruited from a local hospital to participate in the study. Informed consent was obtained from participants prior to participation. Each subject was asked to complete a questionnaire on their background and experiences (see Appendix A.2). The data from 12 subjects\(^1\) (5 Attending Surgeons, 7 Surgery Residents) were used in analysis. The range of surgical suturing experience for attending surgeons was from 7 to 25 years, whereas the range of surgical suturing experience for residents was from 2 to 5 years. The majority of attendings in this study specialized in vascular surgery.

Before suturing, subjects were encouraged to adjust the height of the table.

---

\(^1\)Three subjects did not meet the study criteria and were removed from analysis; 1 attending surgeon (did not meet subject pool definition, not actively practicing), 1 surgery resident (trial interruption), and 1 medical student (did not meet subject pool definition).
(Fig. 3.1) to a comfortable level. Sutures were to be stitched on the radial pattern printed on a synthetic leather membrane (Fig. 3.2b). A red marker in a predetermined location on top of the membrane housing indicated where to perform the first suture. Subjects were instructed to perform continuous, uninterrupted suturing in a counter-clockwise direction on the membrane using a prolene suture needle (SH, 26 mm, 3-0) (Ethicon Inc., Somerville, NJ)). Subjects performed this procedure at two different membrane depths: at “surface” (i.e., 0 in. depth) and at “depth” (i.e., 4 in. depth) (Fig. 3.2a).

The data were analyzed using the Wilcoxon rank sum tests (5% significance level) to identify which metrics showed statistically different performance between attending and resident surgeons. Each stitch was considered as a separate trial. Suturing at the surface and at depth are analyzed separately.

6.2 Experimental Results and Discussion

We analyzed subjects’ performance data for continuous, i.e., uninterrupted, sutures at 12 suture locations on a synthetic leather membrane. First, each subject’s data were partitioned into twelve individual suture cycles using the entry and exit times obtained from the computer vision algorithm. Each stitch was considered as an individual trial. Then, proposed metrics were calculated for each subject. The analysis is twofold: (i) attending versus resident, the suturing performance of attending surgeons and surgical residents were compared for the various metrics at different depth levels, (ii) surface versus depth, how the suturing performance of attending surgeons and surgical residents is affected following the introduction of different depth levels. Since the observed distribution of the metrics was not Gaussian, the Wilcoxon rank sum tests (5% significance level) was used. All the statistical results for force-based, motion-based and physical contact metrics are summarized in Table 6.1. Statistical results for image-based and image-enabled metrics are reported in Table 6.2. Interpretation and discussion of the results presented in the following sections.
Table 6.1: Statistical Results for Force-based, Motion-based and Physical Contact Metrics

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Surface Level</th>
<th>Depth Level</th>
<th>Surfaces vs. Residents</th>
<th>Surface vs. Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Force/Torque-based</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEAK+($F_x$)</td>
<td>3.04x10^{-1}</td>
<td>4.96x10^{-1}</td>
<td>6.34x10^{-1}</td>
<td>4.86x10^{-1}</td>
</tr>
<tr>
<td>PEAK-($F_x$)</td>
<td>1.81x10^{-1}</td>
<td>7.59x10^{-1}</td>
<td>7.78x10^{-1}</td>
<td>2.74x10^{-1}</td>
</tr>
<tr>
<td>PP($F_x$)</td>
<td>6.55x10^{-1}</td>
<td>9.96x10^{-1}</td>
<td>2.92x10^{-1}</td>
<td>6.31x10^{-1}</td>
</tr>
<tr>
<td>PEAK+($F_y$)</td>
<td>9.53x10^{-1}</td>
<td>9.27x10^{-1}</td>
<td>5.97x10^{-1}</td>
<td>6.31x10^{-1}</td>
</tr>
<tr>
<td>PEAK-($F_y$)</td>
<td>1.41x10^{-1}</td>
<td>2.41x10^{-1}</td>
<td>8.15x10^{-1}</td>
<td>7.64x10^{-1}</td>
</tr>
<tr>
<td>PP($F_y$)</td>
<td>3.44x10^{-1}</td>
<td>7.75x10^{-1}</td>
<td>5.41x10^{-1}</td>
<td>8.57x10^{-1}</td>
</tr>
<tr>
<td>PEAK+($F_z$)</td>
<td>9.53x10^{-1}</td>
<td>9.27x10^{-1}</td>
<td>5.97x10^{-1}</td>
<td>6.31x10^{-1}</td>
</tr>
<tr>
<td>PEAK-($F_z$)</td>
<td>1.41x10^{-1}</td>
<td>2.41x10^{-1}</td>
<td>8.15x10^{-1}</td>
<td>7.64x10^{-1}</td>
</tr>
<tr>
<td>PP($F_z$)</td>
<td>3.44x10^{-1}</td>
<td>7.75x10^{-1}</td>
<td>5.41x10^{-1}</td>
<td>8.57x10^{-1}</td>
</tr>
<tr>
<td><strong>Motion based</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP($\theta_{yaw}$)</td>
<td>2.80x10^{-3}</td>
<td>3.19x10^{-3}</td>
<td>6.88x10^{-1}</td>
<td>9.43x10^{-1}</td>
</tr>
<tr>
<td>PP($\theta_{pitch}$)</td>
<td>6.79x10^{-2}</td>
<td>2.01x10^{-2}</td>
<td>1.35x10^{-1}</td>
<td>5.92x10^{-3}</td>
</tr>
<tr>
<td>PP($\theta_{roll}$)</td>
<td>5.20x10^{-7}</td>
<td>2.73x10^{-3}</td>
<td>4.80x10^{-1}</td>
<td>1.21x10^{-1}</td>
</tr>
<tr>
<td>$C_n$</td>
<td>2.52x10^{-3}</td>
<td>7.67x10^{-1}</td>
<td>7.93x10^{-3}</td>
<td>3.58x10^{-1}</td>
</tr>
</tbody>
</table>

Note: Metrics with statistical significance are shown with *
6.2.1 Force/Torque-based Metrics

Table 6.1 shows the p-values for statistical analysis on various force metrics and Fig. 6.1 provides box plots of performance of attending and resident surgeons at surface level and at depth level. A statistical difference in performance \( p < 0.05 \) between attending and resident surgeons was found for metrics \( \text{peak}_{-}(F_z) \) and \( \text{pp}(F_z) \) at both depth and surface as well as for metric \( \text{peak}_{+}(F_z) \) at depth. For \( z \)-directional force metrics, the medians of attendings at both surface and depth level were found to be lower as compared to residents. Similar to an earlier study in laparoscopic suturing [46], our results show that \( z \)-directional force was found to be important for distinguishing between experience levels. In contrast to \( z \)-directional forces, in our study, metrics calculated for \( x \) and \( y \) direction forces at both surface and depth level were found to be non-significant \( p > 0.05 \).

Table 6.1 shows the p-values for statistical analysis on various torque metrics and Fig. 6.2 provides box plots of performance of attending and resident surgeons at surface level and at depth level. Results for torque-based metrics show that only \( z \) directional torques \( (\text{peak}_{+}(T_z), \text{peak}_{-}(T_z) \) and \( \text{pp}(T_z)) \) were significantly different between attendings and residents, at both the depth and surface level \( p < 0.05 \). The \( z \)-axis is vertical, so \( T_z \) is associated with forces orthogonal to the \( z \)-axis applied with a non-zero moment arm. Given the radial suturing pattern, that means \( T_z \) is most closely associated with forces orthogonal to the stitch direction. This motivates direct measurement of the orthogonal force \( F_o \), as explained in Section 5.5.1.
Figure 6.1: Experimental Results for Force-based Metrics: * indicates statistical significance for p<0.05. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)
Figure 6.2: Experimental Results for Torque-based Metrics: * indicates statistical significance for p<0.05. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)
6.2.2 Motion-based Metrics

Previous studies suggest that there is a significant difference in hand movement between expert and novice surgeons during suturing. The rotation of the wrist, indicated by $\theta_{\text{roll}}$, was previously found to be particularly useful in assessment of suturing skill [42, 46]. In the present study, similar to earlier studies, the total range of hand movement for $\text{PP}(\theta_{\text{yaw}})$ and $\text{PP}(\theta_{\text{roll}})$ at both surface and depth, and for $\text{PP}(\theta_{\text{pitch}})$ at depth were found to be statistically significant in differentiating attendings from residents ($p < 0.05$). This suggests that yaw, pitch and roll might be useful for assessment of suturing skill.

Table 6.1 shows the p-values for statistical analysis on various motion metrics and Fig. 6.3 provides box plots of performance of attending and resident surgeons at
Figure 6.4: Experimental Results for Physical Contact Metrics: * indicates statistical significance for p<0.05. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)

surface level and at depth level. Results for yaw, pitch, and roll show that total range of hand movement by attendings are consistently lower than that of residents, regardless of depth. In [42, 46], it was found that experts use greater wrist rotation during suturing. In contrast, our results show that attendings use less wrist rotation. This may be explained by the fact that the majority of attendings in this study were experts in the field of vascular surgery. Due to the intricate nature of this type of surgery, it may be reasonable to assume that significant wrist rotation is not necessary in achieving accurate suturing during the surgical procedure. Also, pitch was found to be statistically significant, but only at depth, possibly because hand motion is more complicated when a subject sutures at depth. Moreover, during the experiments, it was observed that inexperienced suturers tend to reposition the needle holder more often while suturing at depth. The complexity of hand movement during suturing deserves further investigation, specifically for suturing at depth, an essential aspect of vascular suturing.

6.2.3 Physical Contact Metric

We examined the number of times subjects made physical contact with the platform at both surface and depth conditions (see Table 6.1 and Fig. 6.4). Results indicate
that the total number of physical touches \((C_n)\) on surface level for attendings was significantly lower than for residents \((p < 0.05)\), whereas there was no statistical difference between attendings and residents at depth. It should be noted that suturing at depth was introduced to mimic more realistic surgical conditions; however, feedback from attendings after the experiment revealed that requiring a surgeon to suture accurately without touching the top and/or the walls of the cylinder was an overly restrictive constraint. In fact, in certain conditions during surgery, surgeons strategically use boundaries of body cavities, for instance, to augment their forces during suturing.

### 6.2.4 Image-based Metrics

Box plots comparing the performance of attendings and residents in terms of the image-based metrics are presented in Fig. 6.5. Corresponding p-values for the statistical tests for each of the metrics are shown in Table 6.2. The results presented in the figure and table are discussed below.

Results for **Entry Distance** and **Exit Distance** show that entry and exit point accuracy for all subjects was widely distributed on both the surface and depth level and there was no significant difference in performance between attendings and residents. The median **Stitch Length** for attendings was significantly shorter than for residents at both surface and depth \((p < 0.05)\). Similar accuracy for entry and exit locations but differences in stitch length may seem like a contradiction at first glance, but this is resolved by noting that attending surgeons appear to emphasize short stitches rather than accuracy of entry and exit locations.
Table 6.2: Statistical Results for Image-based and Image-enabled Metrics

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Surface Level</th>
<th>Depth Level</th>
<th>Image-based</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry Distance ($d_{oe}$)</td>
<td>1.01x10^{-1}</td>
<td>9.67x10^{-1}</td>
<td>2.42x10^{-1}</td>
<td>6.76x10^{-1}</td>
</tr>
<tr>
<td>Exit Distance ($d_{ox}$)</td>
<td>1.43x10^{-1}</td>
<td>5.22x10^{-1}</td>
<td>3.41x10^{-1}</td>
<td>1.77x10^{-1}</td>
</tr>
<tr>
<td>Stitch length ($l_{s}$)</td>
<td>4.53x10^{-7}*</td>
<td>1.47x10^{-5}*</td>
<td>2.59x10^{-1}</td>
<td>4.92x10^{-1}</td>
</tr>
<tr>
<td>Stitch Time ($t_{s}$)</td>
<td>9.46x10^{-2}</td>
<td>1.36x10^{-1}</td>
<td>1.09x10^{-1}</td>
<td>1.14x10^{-1}</td>
</tr>
<tr>
<td>Idle Time ($t_{d}$)</td>
<td>1.17x10^{-2}*</td>
<td>3.44x10^{-1}</td>
<td>4.59x10^{-2}*</td>
<td>4.25x10^{-1}</td>
</tr>
<tr>
<td>Tip Path Length ($l_{tp}$)</td>
<td>8.84x10^{-4}*</td>
<td>1.19x10^{-7}*</td>
<td>8.37x10^{-2}</td>
<td>5.79x10^{-5}*</td>
</tr>
<tr>
<td>Swept Area ($a_{s}$)</td>
<td>1.45x10^{-5}*</td>
<td>3.27x10^{-11}*</td>
<td>2.16x10^{-1}</td>
<td>4.66x10^{-6}*</td>
</tr>
<tr>
<td>Tip Area ($a_{t}$)</td>
<td>9.51x10^{-4}*</td>
<td>1.76x10^{-10}*</td>
<td>5.58x10^{-4}*</td>
<td>6.71x10^{-11}*</td>
</tr>
<tr>
<td>Sway Length ($l_{sw}$)</td>
<td>5.01x10^{-8}*</td>
<td>3.62x10^{-7}*</td>
<td>8.82x10^{-4}*</td>
<td>4.14x10^{-4}*</td>
</tr>
</tbody>
</table>

| Image-enabled | Peak+($F_{o}$) | 3.32x10^{-3}* | 1.71x10^{-2}*| 8.27x10^{-1}| 5.24x10^{-1}|
|              | Peak-($F_{o}$) | 2.76x10^{-4}* | 1.79x10^{-2}*| 1.39x10^{-3}*| 1.21x10^{-5}*|
|              | PP($F_{o}$)    | 3.01x10^{-6}* | 3.38x10^{-4}*| 2.16x10^{-2}*| 1.95x10^{-3}*|
|              | Peak+($F_{t}$) | 4.59x10^{-2}* | 1.25x10^{-1}| 2.12x10^{-3}| 5.27x10^{-3}*|
|              | Peak-($F_{t}$) | 1.80x10^{-2}* | 7.36x10^{-2}| 3.73x10^{-3}*| 5.75x10^{-3}*|
|              | PP($F_{t}$)    | 1.60x10^{-1}  | 2.26x10^{-1}| 1.08x10^{-2}*| 2.75x10^{-2}*|
|              | $F_{cm_{max}}$ | 6.39x10^{-1}  | 5.30x10^{-1}| 2.22x10^{-1}| 7.28x10^{-2}|
|              | $F_{c}$       | 5.30x10^{-1}  | 4.31x10^{-1}| 3.09x10^{-1}| 1.28x10^{-2}|
|              | $\alpha_{e}$  | 2.02x10^{-1}  | 4.80x10^{-2}*| 7.93x10^{-3}| 9.73x10^{-2}|
|              | $|\alpha_{e} - 90^\circ|$ | 5.05x10^{-3}* | 4.80x10^{-2}*| 7.93x10^{-3}| 9.73x10^{-2}|

Note: Metrics with statistical significance are shown with *
Figure 6.5: Experimental Results for Image-based Metrics: * indicates statistical significance for $p<0.05$. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)
The results show that *Stitch Time* could not be used to distinguish between residents and attendings at either surface or depth levels. Technically, there was a statistically significant difference for *Idle Time* at the surface level, but given the relatively high p-value and the fact that there was no statistically significant difference was observed for *Idle Time* at depth, we suspect that the results for *Idle Time* at surface level was a statistical fluke. In earlier studies [43, 85, 42, 48, 14], metrics *Stitch Time* and *Idle Time* were able to differentiate the skill between experts and novices. In contrast, our results suggest that these temporal metrics are not useful for distinguishing skill level between resident and attending surgeons.

Visualizations of *Needle Tip Path*, *Swage Path*, and the *Needle Swept Area* for a typical attending and a typical resident subject are presented in Fig. 6.6. Note the obvious visual distinctions between the resident and attending examples. Intuitively, better-suturing performance should be associated with the smaller needle swept area and a smoother needle path. Statistical analysis shows that *Needle Swept Area* for attendings is significantly less than for residents as expected at both surface and depth. Similarly, attendings’ needle paths are visually more steady than residents’, suggesting that image-based process metrics are able to quantify the differences in a way that is intuitive and easy to interpret. This visual interpretation may serve as useful feedback for trainees during skill training.

*Needle Tip Path Length* and *Needle Tip Area* results (see Fig. 6.5) at both surface and depth were significantly higher for residents than for attendings. Furthermore, there was also a statistically significant difference between residents’ surface and residents’ depth level performances ($p < 0.05$). During suturing, residents strayed farther from the ideal path and were less accurate in following the curvature of the needle from entry to exit. In contrast, attendings’ performances remained relatively consistent, regardless of membrane level. Similarly, *Needle Swept Area* and *Needle Sway Length* results (see Fig. 6.5) at both surface and depth were significantly higher for residents
than attendings ($p < 0.05$). Large needle sway length indicates that residents had large deviations in the roll of the needle, which caused large swept area. Further, the results agree with expectations based on the maxim to “follow the curvature of the needle.” Specifically, the metrics Needle Swept Area, Needle Tip Path Length, Needle Tip Area, and Needle Sway Length, at both surface and depth levels, are lower for attendings than for residents. Therefore, it may be interpreted that smaller values of the image-based metrics indicate better suturing performance. It is apparent from these results that image-based metrics are better than the temporal and product metrics at distinguishing skill between attendings and residents.

It should be emphasized that the metrics are distinguishing between subpopulations of surgeons, i.e., residents versus attendings. Even the “novice” group, residents, had substantial task-specific experience. In contrast, skill differentiation in other work
[43, 46, 52] was between expert (attendings and/or residents) and novice (medical students and/or no medical background).

As seen from Fig. 6.5, for attendings, there was very little variation in image-based metrics when suturing at surface versus at depth. On the other hand, residents’ performance was worse at depth than at surface. Moreover, statistical tests for *Needle Swept Area* \((p = 4.66\times10^{-6})\) and *Needle Tip Path Length* \((p = 5.79\times10^{-5})\) comparing residents’ performance at surface versus at depth show statistically significant differences, suggesting that suturing at depth is a more challenging task for subjects with less experience. In contrast, similar statistical tests show that attendings’ performance did not significantly vary from surface to depth. These results show that suturing at depth is especially useful for assessing performance that requires advanced skill.

### 6.2.5 Image-enabled Metrics

#### 6.2.5.1 Orthogonal and Tangential Forces

Results show that the metrics obtained from orthogonal force \((F_o)\) were statistically different \((p < 0.05)\) between attendings and residents on both surface and depth levels (See Table 6.2). In addition, tangential force \((F_t)\) metrics were significantly different between attendings and residents at surface \((p < 0.05)\), with the exceptions of PP\((F_t)\). Orthogonal forces applied by attendings were lower than those applied by residents, whereas tangential forces applied by attendings were higher.

In [46], subjects made parallel sutures aligned with the \(y\) axis of the force sensor. It was observed that the maximum absolute forces in \(x\) and \(y\) directions were important for distinguishing between experience levels. Since the stitch direction was unchanged, \(x\) and \(y\) force directions were always orthogonal and tangential to the stitch direction, respectively. The study presented here uses a radial suture membrane with stitches in 12 different directions. This radial membrane is based on the one used in FVS training and is intended to test the subject’s dexterity and preparedness for vascular anastomosis.
Figure 6.7: Experimental Results for Image-enabled Metrics (Orthogonal and Tangential Forces): * indicates statistical significance for p < 0.05. (On each box, the middle line indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points including outliers.)

Since the force sensor was fixed in place, x and y force directions were not generally aligned with stitch direction. Even though x and y directional force metrics were not found to be statistically significant in our study, measurements of forces in x and y directions are required to calculate orthogonal and tangential forces. Reinterpreting the x and y for axes from [46] as orthogonal and tangential to stitch direction, the present study supports that orthogonal forces, and to a lesser extent tangential forces,
are important for distinguishing skilled performance.

6.2.5.2 Entry Force and Angle

Obtaining the angle at which the needle enters the membrane offers insight into suturing skill since general guidelines that surgeons follow emphasize that the needle should enter the tissue perpendicularly. Our system construction allows us to obtain forces applied along the stitch, which makes it possible to calculate the entry force and angle as described in Section 5.5.2.

After angle calculations were verified using the previous method in 5.5.2, we analyzed the performance of residents and attendings for each suture location. Note that subjects used a standard semi-circular needle, often used in vascular suturing, during the experiment. Preliminary analysis revealed that the results for calculation of the angle and forces for each suture location presented confounding results. Subjects tended to pull the thread during continuous suturing to keep the suture away from the needle. The force applied to the thread was confounded with the force due to the needle, distorting the calculation. For the first suture, however, this effect is not present since the suture is not yet in the membrane. Fig. 6.8 illustrates a subject pulling the thread during suturing. While each subject made a total of 12 sutures on each membrane, we only used the subject’s first suture location to calculate the force-based needle angle metrics. In the future studies, we may use a half-circle needle without thread to overcome this problem so that forces applied by thread do not interfere with the needle forces.

Experimental results for the needle entry angle and forces for the first suture location can be seen in Fig. 6.9. Table 6.2 shows the p-values for statistical analysis on various needle entry angle and force metrics. At both surface and depth levels, the median of the mean needle entry angle ($\bar{\alpha}$) was found to be higher for attendings as compared to residents. Moreover, the deviation of attendings entry angle were found to be near 90°. Therefore, the deviation of each group’s entry angle from 90° was measured.
In other words, we analyzed how close the needle entry angle metrics were to 90° for groups at both surface and depth level. The results showed that the deviations of entry angle from 90° were lower for attendings and there was a statistical difference between attendings and residents for both the surface and depth level. In a study by Joice et al. [82], the entry angle was qualitatively assessed by humans based on video data as: (i) less than 80°, (ii) between 80° to 100°, and (iii) above 100°. In that study, it was found that experts use a needle entry angle between 80° to 100°. Our results are qualitatively in agreement with [82]. In particular, angle mean results show that the angles for attendings were spread between 80° to 100° at the surface level. Further, Joice et al.’s study only considers suturing at surface level. In our study, for attending surgeons, we found a statistically significant difference (p value=7.93 × 10^{-3}) for entry angle when suturing at surface versus 4 inches depth. The relationship of depth level and entry angle deserves further study. In addition, our results indicate that the distribution of entry angles for attendings at both surface and depth level are more tightly clustered around the median as compared to entry angles for residents.
As seen from the results, the medians for the mean and maximum forces were lower for attendings in comparison with residents at both surface and depth level, but there were no statistically significant differences between the forces applied by attendings and residents. It should be noted that the statistical power was low for this study since only the first suture location was considered.

During lab experiments with a semi-circular needle, it was observed that an inexperienced suturer can apply forces at an angle different from the tangent to the tip of the needle. The angle of the force and the geometric angle of the needle are distinct concepts which appear to coincide for skilled suturers. Further study is required to investigate this hypothesis.
Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, a custom-designed surgical simulator for the assessment of suturing skill in open surgery was presented. The simulator has the capability of collecting synchronized force, motion, touch, and video data while subjects suture a membrane. The simulator setup allows image information to be continuously captured from beneath the suture membrane. A computer vision algorithm was presented that processes these images to extract product and process metrics useful in the assessment of the surgical skill. Process and product metrics are based on force, motion, physical contact, and video data. Image-enabled metrics that combine video and force data to extract metrics for assessment of surgical skill were also presented. Specifically, needle entry angle and forces orthogonal and tangential to stitch direction were extracted via combining force data with computer vision information. Also, the vision algorithm aided in the identification of suture events and the segmentation of corresponding sensor data. The metrics towards the assessment of suturing skill in open surgery were motivated by insight we received from practicing vascular surgeons. Metrics were based on the physics of needle insertion forces, the common practice of following the curve of the needle while driving
through tissue, and minimizing lateral forces and motions that induce tear.

A study involving attending surgeons and surgical residents was performed to investigate the use of proposed metrics for assessment of open surgery suturing skill. Data from both attendings and residents were presented, and metrics were used to compare the performance of attendings and residents. Analysis shows that force-based metrics (absolute maximum force/torque in z-direction), motion-based metrics (yaw, pitch, roll), physical contact metric, image-based metrics (Distance Exit, Stitch Length, Idle Time, Needle Tip Trace Distance, Needle Swept Area, Needle Tip Area and Needle Sway Length), and image-enabled metrics (orthogonal force, tangential force and entry angle) were found to be statistically significant in differentiating suturing skill between attendings and residents. The image-based metrics were especially effective in capturing fine-grained differences in skill level between residents and attendings. Furthermore, the image-based metrics lend themselves to intuitive visualizations. In addition, the effect of membrane depth level on suturing performance was examined. Statistical tests comparing residents’ performance at surface versus at depth show statistically significant differences, suggesting that suturing at depth is a more challenging task for subjects with less experience. In contrast, similar statistical tests show that attendings’ performance did not significantly vary from surface to depth. These results suggest that suturing at depth is especially useful for assessing performance that requires advanced skill. The combination of fine-grained skill differentiation, the ability to simulate depth of suturing, and the intuitive visualizations of selected image-based metrics makes the suturing simulator and associated suite of metrics well-suited for the assessment and training of suturing skills.

7.2 Future Work

The work presented in this thesis outlined the foundation of designing and testing a simulator capable of assessing and training suturing skill. Currently, this simulator,
along with presented force-based, motion-based, image-based, and image-enabled metrics, is capable of differentiating suturing skill between attending surgeons and surgical residents. Future work should focus on whether or not simulator training outside the operating room is effective in improving skill inside the operating room. The transition of training to operating room is an important step towards making this simulator capable of training surgical skill useful in the operating room.

The following analyses would also be insightful: (i) investigating a possible correlation between presented metrics, (ii) determining the most important metric/metrics for skill assessment and using those metric/metrics to provide feedback to the subject. To effectively examine these questions, one goal would be to conduct a large-scale study of suturing skill assessment using the simulator. A large-scale data set will guarantee higher statistical power for the overall study. This would be particularly true for the entry angle and entry force results, since only the first suture location was considered during the analysis of these metrics. Furthermore, data from both the preliminary study and the large-scale study could be analyzed using various machine learning algorithms.

In addition, future work would focus on extending image-enabled metrics. In this thesis, image-enabled metrics (orthogonal force, tangential force, entry angle and entry force) were extracted combining force and image data. However, our simulator has the capability of collecting motion data as well. Combining image data with motion data will allow for image-enabled motion-based metrics.

Future work should also focus on determining the best method for real-time feedback using this suite of metrics. During thesis work, real time image processing and metrics computation on the suturing simulator were achieved (See Appendix A.1). This allowed for immediate feedback during and after each stitch. A prototype training interface, not yet tested, was also created. In this interface, the subject’s performance on a given stitch was replayed upon completion of the stitch. Information and selected metrics were overlaid on and around the video to provide thorough feedback. Explanation of
this interface can be found in Appendix A.2. This is a significant milestone in preparing
the system for training applications, however, the training interface needs to be further
investigated and developed in order to provide effective metric-driven feedback to users.
Appendices
Appendix A  Towards Real-Time Surgical Skill Training

Fast data-processing is an essential characteristic of any simulator that is capable of producing immediate meaningful feedback information related to data obtained from a participant. The ability of the system to provide immediate feedback information can allow for the system to be used as a potential training simulator. Therefore, we increased the process speed and achieved our suturing simulator to run in real-time and provide real-time feedback, enabling the simulator to be used towards skill training in the future.

In the experimental study, all raw data was logged, and the metrics were computed during post-processing. To prepare for training, where live feedback is desired after each stitch, the rate of image processing had to be optimized to run in real time. We have achieved real time image processing and metrics computation on the suturing simulator, which allows live feedback during and after each stitch.

A.1 GPU Implementation - CPU to GPU Transition

The computer vision algorithm presented in Chapter 4 is capable of processing video data in both real-time and post-processing with CPU computing. However, after testing the performance of the computer vision algorithm during real-time processing, it was found that the time needed to process a single frame took a significantly long time. This was seen in the slowing of the frame rate, which dropped from 60 fps to 5-10 fps in real-time processing. As mentioned previously in Chapter 4, the Logitech camera, initially used in capturing video at 30 fps, was replaced with the Firefly camera, with capability of capturing video at 60 fps. This replacement was necessary to accommodate the speed of expert surgeons. During initial system testing, it was found that experienced surgeons performed suturing faster than the Logitech camera was capable of capturing. In other words, the movement of the needle beneath the membrane was too fast to be captured in the video data. This led to the failure of accurate needle mo-
Figure 1: A flowchart of the system that allows GPU Computation
detection which is vital to the extraction of the presented metrics. For this reason,
post-processing was initially favored over real-time processing in our study. Initially, a
standard PC (Windows 7 Enterprise 64-bit, Intel i7 CPU@2.67GHz, 12GB RAM) was
dedicated to collect synchronized unprocessed/raw force, motion, touch, and video data
during the experiment. The synchronization and collection of the data were achieved in
this PC with a software written in MS Visual Studio 2013 employing OpenCV (v.3.0.0)
and multithreading in C++. After each experiment, data was transferred to a second
computer so that each subject’s data from the experiment can be post-processed and
metrics can be obtained. The second computer was also a standard PC with the spec-
ification of Windows 8.1 Enterprise 64-bit, Intel i7-4790 CPU@3.6GHz, 12GB RAM.
The collected data were then post-processed with a computer vision algorithm also writ-
ten in MS Visual Studio 2013 employing OpenCV (v.3.0.0) library in C++ with CPU
processing.

During post-processing, it was realized that a 5-minute video data could be pro-
cessed approximately in 25 minutes since the vision algorithm involves heavy processing.
This precludes to give a real-time feedback to subjects which is a vital aspect of training with feedback. In an attempt to speed up real-time processing, we improved the efficiency of the functions used in the computer vision algorithm. Using the diagnostic tools provided in Visual Studio, we identified the most computationally expensive functions within our code and then altered them so that processing time is minimized. However, improving the efficiency via altering the code was not sufficient to achieve processing in real-time. To overcome this problem, it was anticipated that the implementation of GPU computing for video frames would significantly decrease the overall processing time. Therefore, we switched from CPU computing to GPU computing. We first adapted our vision algorithm code from CPU processing to GPU processing to accomplish the real-time processing. The PC used to collect synchronized raw data did not have a graphics card capable of GPU computing. Therefore, GPU computing was achieved in a laptop (Windows 10 Home 64-bit, Intel i7-7500 CPU@2.7GHz, 8GB RAM and a graphics card NVIDIA GeForce 940MX. To achieve GPU computing, we have used CUDA Toolkit 8.0, OpenCV with GPU/CUDA module (v.3.0.0) library and MS Visual Studio 2015 development environment. We then extended our both codes on the PC and the Laptop with WinSock API to establish TCP/IP protocol between them. A Cat6 Crossover cable was used to establish a high-speed connection between the PC and the laptop. A flowchart
of the system can be seen in Figure 2. This configuration enables to synchronize and log raw data in the PC and transfer video data from PC to laptop in order to process it in real-time.

This configuration enables to synchronize and log raw data in the PC and transfer video data from PC to laptop in order to process it in real-time. Increased speed of the computer vision algorithm, as a result of these improvements, allows our suturing simulator to run in real time. After achieving the real-time processing using two computers, we have decided to move the whole process to one PC. The software was implemented as two processes, a Capture and Logging Process, which collects all camera, force, and motion data, and a Real Time Computation Process, which processes the video and computes metrics. Both process run on the same Windows 10 Enterprise PC, which has a Xeon CPU at 3.60GHz, 32GB RAM, and an NVIDIA Quadro P5000 GPU. Real time computation was achieved by rewriting the vision processing and metrics computation.
code to use GPU computation using OpenCV with the GPU/CUDA module (v3.0.0) and CUDA Toolkit 8.0. Increased speed of the computer vision algorithm, as a result of these improvements, allows our suturing simulator to run in real time, enabling the simulator to be used towards skill training in the future. This is a significant milestone in preparing the system for training purposes.

A.2 Visual Feedback For Training

Unlike procedural simulators that use high-fidelity graphics to render anatomy-based feedback, our goal here is to create an interface that enables intuitive processing of objective feedback towards skill progression. The challenge with this goal is to present “the right information at right time” in an effective manner. Therefore, a training interface was developed to provide metrics-driven effective feedback to users during training with metrics that might be relevant to improve suturing. In the prototype training interface, a video of the subject’s performance on a given stitch can be replayed immediately after finishing the stitch. Simultaneously, information and metrics can be overlaid on
and around the video, including needle entry and exit locations, needle tip paths and corresponding Needle Tip Distance metric, Needle Swept Area and the corresponding metric, and so forth. A screenshot of the visual feedback can be seen in Figure 3.

In addition to this visual feedback, we can display the subject’s metric results calculated after each suture so that progress can be seen after every suture. An example of this feedback for Needle Swept Area at 12 suture locations can be seen in Figure 4.
Appendix B  Questionnaire

Suturing Study Questionnaire

Age: ___________

Sex (circle one): Male    Female    Prefer not to answer

Height: ___________

Dominant hand (circle one): Left    Right    Either

• With which hand do you draw?    Left    Right    Either
• Which hand would you use to throw a ball to hit a target?    Left    Right    Either
• In which hand would you use an eraser on paper?    Left    Right    Either
• Which hand removes the top card when you are dealing from a deck?    Left    Right    Either
• With which hand do you suture?    Left    Right    Either

Which of the following best describes you (circle one):

Attending Surgeon: Year__________ Specialist (Vascular, General, etc… ) __________

Resident: Year__________

Medical student: Year__________

Intern or Undergraduate student: Year__________ Major__________________

Please list the total hours spent on suturing training in your training (labs, workshops, models used [e.g., pigs feet, synthetic], etc.)

<table>
<thead>
<tr>
<th>Suturing on synthetic models</th>
<th>None</th>
<th>&lt; 5hrs.</th>
<th>5-15 hrs.</th>
<th>15-30 hrs.</th>
<th>30-60 hrs.</th>
<th>&gt;60 hrs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suturing on ex-vivo animal tissue (pigs feet, etc.)</td>
<td>None</td>
<td>&lt; 5hrs.</td>
<td>5-15 hrs.</td>
<td>15-30 hrs.</td>
<td>30-60 hrs.</td>
<td>&gt;60 hrs.</td>
</tr>
<tr>
<td>Suturing in-vivo during a procedure</td>
<td>None</td>
<td>&lt; 5hrs.</td>
<td>5-15 hrs.</td>
<td>15-30 hrs.</td>
<td>30-60 hrs.</td>
<td>&gt;60 hrs.</td>
</tr>
<tr>
<td>Suturing on surgical simulators</td>
<td>None</td>
<td>&lt; 5hrs.</td>
<td>5-15 hrs.</td>
<td>15-30 hrs.</td>
<td>30-60 hrs.</td>
<td>&gt;60 hrs.</td>
</tr>
</tbody>
</table>

Do you currently have any problems with your hands, arms, or neck? Yes    No

If yes, please describe: ____________________________________________________________

Have you ever had any surgery on your hands or arms (including fingers and wrists)? Yes    No

If yes, please describe (including which hand or both): ________________________________________

Do you currently have any vision problems aside from corrected vision? Yes    No

If yes, please describe: ________________________________________________________________

ID: ___________________ 
Date: ___________________
For Attending Surgeon and Resident:

- Approximately how many years have you been practicing surgical procedures involving suturing?
  
  __________ years

- Approximately how many surgical procedures have you performed that involved suturing?
  
  <25  25-100  101-500  501-1,000  1,001-1,500  >1,500  N/A

- Approximately how many Non-Robotic Minimally Invasive Surgeries have you performed that involved suturing?
  
  <25  25-100  101-500  501-1,000  1,001-1,500  >1,500  N/A

- Approximately how many Robotic Minimally Invasive Surgeries have you performed that involved suturing?
  
  <25  25-100  101-500  501-1,000  1,001-1,500  >1,500  N/A

- On a scale of 1-10, how would you rate your surgical suturing skill (1=poor, 10=world-class)?
  
  __________

Post-test feedback

On a scale of 1-10, how would you rate the ease/comfort of using the suturing device? (1=worst; 10=best):

__________________

What are the strengths of the suturing device?

__________________

What are the weaknesses of the suturing device?

__________________

In your opinion, what features would the ideal suturing training device have?

__________________

What would you suggest to improve this system so that it may successfully aid in the suturing learning process?

__________________

Are there any additional comments you would like to add?

__________________
Appendix C  Publications


Bibliography


