Evaluating Smoothness of Force for Surgical Skill Assessment

Simar P. Singh\textsuperscript{1}, Amir Mehdi Shayan\textsuperscript{1}, Jianxin Gao\textsuperscript{2}, Joe Bible\textsuperscript{3}, Richard E Groff\textsuperscript{2}, and Ravikiran Singapogu\textsuperscript{1}

\textsuperscript{1}Department of Bioengineering, Clemson University  
\textsuperscript{2}Department of Electrical and Computer Engineering, Clemson University  
\textsuperscript{3}Department of Mathematical and Statistical Sciences, Clemson University

March 29, 2024

Abstract

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.\textsuperscript{2}

The efficient execution of surgical operations plays a crucial role in optimizing patient outcomes, evidencing the need for effective training methods to improve surgical skills. The gaining traction of medical training simulators for automated skill assessment necessitates instrumented sensors and relevant metrics for targeted feedback on all aspects of a surgical procedure. Traditional metrics that capture a single instance of force, such as peak or range, lack the characterization of the entire force profile and lose subtleties that may limit accurate evaluation of the skilled application of force, a valuable aspect of assessment in surgery. This study introduces novel force metrics inspired by motion smoothness-based measures, analyzed on an extensive dataset of 97 subjects suturing on an open vascular suturing simulator. We validated the effectiveness of these metrics by comparing the metrics’ ability to distinguish between subject skill levels. Our findings highlight the value of these advanced force metrics as robust indicators of suturing performance, demonstrating their valuable potential for more accurate and objective skill assessment in surgical training. Clinical Relevance—The force metrics presented in this study analyze the intricacies of the widespread category of assessment in surgery, “respect for tissue”, greatly benefiting surgical education with an improved evaluation of this aspect of suturing skill.
Evaluating Smoothness of Force for Surgical Skill Assessment

Simar P. Singh¹, Amir Mehdi Shayan¹, Jianxin Gao², Joe Bible³, Richard E. Groff², and Ravikiran Singapogu¹
¹Department of Bioengineering, Clemson University, Clemson, SC, USA
²Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA
³Department of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA

Abstract—The efficient execution of surgical operations plays a crucial role in optimizing patient outcomes, evidencing the need for effective training methods to improve surgical skills. The gaining traction of medical training simulators for automated skill assessment necessitates instrumented sensors and relevant metrics for targeted feedback on all aspects of a surgical procedure. Traditional metrics that capture a single instance of force, such as peak or range, lack the characterization of the entire force profile and lose subtleties that may limit accurate evaluation of the skilled application of force, a valuable aspect of assessment in surgery. This study introduces novel force metrics inspired by motion smoothness-based measures, analyzed on an extensive dataset of 97 subjects suturing on an open vascular suturing simulator. We validated the effectiveness of these metrics by comparing the metrics’ ability to distinguish between subject skill levels. Our findings highlight the value of these advanced force metrics as robust indicators of suturing performance, demonstrating their valuable potential for more accurate and objective skill assessment in surgical training.

Clinical Relevance—The force metrics presented in this study analyze the intricacies of the widespread category of assessment in surgery, “respect for tissue”, greatly benefiting surgical education with an improved evaluation of this aspect of suturing skill.

I. INTRODUCTION

In vascular surgery, successful surgical outcomes depend on the meticulous execution of foundational surgical techniques like suturing. Studies in bariatric and cancer surgeries demonstrating the positive impact of surgeons’ proficiency on favorable patient outcomes highlight the need for efficient surgical skills training [1], [2]. In vascular anastomosis, where delicate vessel connections are crucial, common complications like bleeding and tearing can lead to patient morbidity [3]. Surgeons must apply minimal force on surrounding tissues during the suturing process, avoiding excessive pulling that can tear fragile tissues [4]. Consequently, “respect for tissue” is universally desired in surgery and is a crucial category of surgeons’ skill rating in the Objective Structured Assessment of Technical Skill (OSATS), a well-substantiated rating scale to evaluate medical trainees [5].

Assessing respect for tissue relies on overt visual cues, such as excessive tissue deformation and tearing. However, capturing the quality of force and torque applications remains a challenge due to the limitations of visual observation. Sensing elements offer a more precise solution, capturing subtle dynamics invisible to the eye.

Despite the value of force measurement in surgery, current research in quantifying forces during surgery is limited. This scarcity can be attributed to the technical complexity of instrumenting force sensors for advanced procedures, often involving intricate interfacing within surgical tools. While prior studies have successfully differentiated novices from attending surgeons [6], [7], results become inconsistent when distinguishing resident performance from attending surgeons or across different resident PGY levels [7], [8]. Previous work on the SutureCoach by Kil compared traditional force metrics between attendings and residents, observing significant differences in the z and orthogonal directions [9]. The traditional single-instance force metrics (maximum or range) fall short due to their sensitivity to outliers and failure to represent the entire force profile, as highlighted by Trejos et al. [10]. They proposed using metrics designed to assess skilled motion to quantify the skilled application of force, achieving success by associating these metrics with clinical expertise in laparoscopic surgery. Their work demonstrates the potential of this method to reveal deeper insights into the dynamic forces that comprise suturing technique.

This study presents and evaluates new metrics that measure force during simulated surgical suturing that are inspired by measures of smoothness of motion. These metrics are evaluated on a large dataset encompassing several levels of clinical expertise performing tasks on the SutureCoach platform [11]. The SutureCoach interfaces intricate sensing elements for a holistic evaluation of open vascular suturing skills on a clock face pattern modeled after the Fundamentals of Vascular Surgery (FVS) [12]. This study will focus on cutting-edge force and torque metrics to measure the quality of membrane forces during suturing.

Our central research question focuses on analyzing the entire force profile to discern differences in levels of suturing skill. We hypothesize that analyzing the dynamic flow of force throughout a suture provides a more nuanced understanding of the suturing technique, potentially revealing subtle skill variations unable to be captured by traditional force metrics.
contact with the membrane surface, particularly during needle
side their force profiles, we noticed frequent needle driver
indicative of tissue stress.

relation to the suture direction, allowing analysis of the quality
components [9]. This transformation aligns measured forces in
and y data were transformed into orthogonal and tangential
B. Force/Torque in the Orthogonal and Tangential Directions
barrier. Each trial consists of 12 sutures.
simulating suturing in an anatomical cavity through a raised
trials in two conditions: surface and depth suturing, the latter
Novices: 32 medical students/other) completed up to two
is attached to the bottom of the cylinder.
movement, and a force/torque sensor (ATI Mini40, 1000 Hz)
internal camera (Intel RealSense D435, 60 FPS) tracks needle
Subjects suture on a simulated membrane material affixed
sors synchronized via C++ (Microsoft Visual Studio 2017).
METHODS
II. M
A. SutureCoach Description and Participant Information
The SutureCoach platform records data from multiple sen-
sors synchronized via C++ (Microsoft Visual Studio 2017).
Subjects suture on a simulated membrane material affixed
to the surface of a hollow cylinder, seen in Figure 1a. An
internal camera (Intel RealSense D435, 60 FPS) tracks needle
movement, and a force/torque sensor (ATI Mini40, 1000 Hz)
is attached to the bottom of the cylinder.
97 subjects with varying expertise (Experts: 35 attend-
ing surgeons/fellows, Intermediates: 30 residents (PGY1-5),
Novices: 32 medical students/other) completed up to two
trials in two conditions: surface and depth suturing, the latter
simulating suturing in an anatomical cavity through a raised
barrier. Each trial consists of 12 sutures.
B. Force/Torque in the Orthogonal and Tangential Directions
To gain nuanced insights into force application, raw x
and y data were transformed into orthogonal and tangential
components [9]. This transformation aligns measured forces in
relation to the suture direction, allowing analysis of the quality
of 2D driving forces (tangential) and lateral forces (orthogonal)
indicative of tissue stress.
Further, while observing subjects’ video performance along-
side their force profiles, we noticed frequent needle driver
contact with the membrane surface, particularly during needle
adjustment. This contact generates forces in the z direction,
leading to “unsmooth” signals that do not accurately reflect
skilled needle-membrane interaction. To circumvent this issue,
we calculate torque about the orthogonal axis through a sim-
ilar transformation process. This method effectively captures
angular forces relevant to needle driving for a more accurate
representation of vertical force quality. Both transformation
processes enable a clearer representation of forces during
suturing, allowing for a more precise and robust assessment
of suturing technique. These transformation processes are
illustrated in Figure 1b.

C. The Development of Force and Torque Smoothness Metrics:
Analyzing Yank
Our study investigates the applicability of well-established
motion smoothness metrics [11], [13]–[15] to the domain of
force measurement, specifically within the context of suturing.
Notably, Trejos et al. and Balasubramanian et al. recognized
the parallels between force and kinematic profiles, though
the latter emphasized the potential limitations due to inherent
differences in the nature of the data [10], [16]. The fixed
position of the force sensor can limit smoothness measures
as the metrics evaluate changes in continuous measurements
while accounting for hesitation through intermittency within
the signal. Force sensors record solely during application, thus
creating false intermittency outside of this time frame that
negatively impacts the metric. Our study focuses on individual
sutures, isolating our data from membrane needle contact
to needle exit. This shorter task duration (average surface suture
completion time of 9.93 seconds) and the near-continuous
force/torque application during needle driving could mitigate
the identified limitations.

Building upon Trejos et al.’s pioneering work in adapting
movement metrics for force analysis in surgical skills [10],
we implement established metrics like the interquartile range
of force and the derivatives of force while further refining
smoothness calculations in light of recent advancements in the
field [16], [17]. We begin by applying Log Dimensionless Jerk
(LDLJ), a metric that uses the third derivative of position to
evaluate smoothness [13]. In the context of force, we observed
higher derivatives beyond the first filter low-frequency peaks
as seen in Figure 1c. As such, we propose leveraging the first
derivative of force/torque for smoothness-based metrics.

Recognizing the lack of standardized nomenclature for the
first derivative of force, we adopt the term “yank” proposed
by Lin et al. in a commentary on its successful applications in
sensorimotor systems [18]. As noted in their work, given force
is mass times acceleration, yank inherently aligns with the time
derivatives of motion. Therefore, we propose the novel metric
“Log Dimensionless Yank”, or LDLY, using the first derivative
of force/torque as a force-specific equivalent to LDLJ.

The subsequent challenge arises when adapting Spectral
Arc Length (SPARC), originally designed for assessing mo-
tion smoothness through submovement analysis in velocity,
to force data [14]. Although the yank-jerk analogy might
suggest calculating SPARC on the integral of force (impulse),
this approach would yield some ambiguous measure of submovement behavior more indicative of material deformation than a measure of controlled force. As noted by Melendez-Calderon et al., haphazardly applying SPARC to a random signal would be ill-suited due to the lack of interpretability of the variations in said signal, as SPARC requires a profile that represents the “speed” of a given movement [17]. To align with these requirements, we propose applying SPARC directly to force itself (SPARC_F). This approach is supported by the analogous nature of velocity and force, both characterized by submovements consisting of distinct increases and decreases. SPARC_F can provide meaningful and interpretable measures of its submovements, revealing valuable information about the controlled application of force during suturing tasks. This behavior is evidenced in Figure 1d, which depicts novice and expert force and impulse profiles and their corresponding SPARC values in the bottom right panel. The lack of variability and submovement behavior in impulse results in obscure SPARC values in this domain, validating the use of SPARC_F.

In our work, we incorporate force metrics that quantify the entire profile. F total and Y pks are applied to the magnitude of force. LDLY and SPARC_F are applied to orthogonal torque, lateral orthogonal force, and longitudinal tangential force, denoted with a T_o, o, and t, respectively.

1) **Total Force (F total):** A cumulative sum of the absolute value of point-by-point changes in force, similar to the path length metric.

2) **Number of Yank Peaks (Y pks):** The number of peaks in yank with a minimum prominence of 1.5, chosen as yank profiles displayed frequent peaks of lower magnitudes compared to velocity.

3) **Log Dimensionless Yank (LDLY):** For the dimensionless calculation, maximum force (Fp+), aligning with Balasubramanian et al.’s work, who use maximum velocity for LDLJ [14]. Using the maximum value allows for a measure independent of a movement’s amplitude. LDLY_To uses the corresponding maximum torque for the dimensionless calculation.

\[ LDLY = \ln \left( \frac{T}{F_{p+}^{2}} \int_{t_{entry}}^{t_{exit}} \left( \frac{dF}{dt} \right)^{2} dt \right) \]  

4) **Spectral Arc Length of Force (SPARC_F):** As SPARC is a measure of submovements in a profile, we apply the measure directly to force for a measure of smoothness. \( \omega_c \) is an adaptive cutoff frequency, with further detail found in [16].

\[ \text{SPARC}_F = \int_{0}^{\omega_c} \left[ \left( \frac{1}{\omega_c} \right)^{2} + \left( \frac{d\tilde{F}(\omega)}{d\omega} \right)^{2} \right]^{\frac{1}{2}} d\omega; \]  

\[ \tilde{F}(\omega) = \frac{F(\omega)}{F(0)}, \quad F(\omega) = \mathcal{F}\{F(t)\} \]

D. **Data Processing and Statistical Analysis**

Force data was filtered through a second-order low-pass Butterworth filter with a cutoff frequency of 50 Hz. Derivatives were calculated with a third-order Savitzky-Golay filter with a window span of 101. To validate the efficacy of the new force/torque metrics, we compared differences in metric scores among novices (32 subjects with no experience), intermediates (30 medical residents PGY1-5), and experts (35 medical fellows/attending surgeons). For analysis, we looked at mean metrics by skill level. Mean differences, point estimates and confidence intervals were calculated using Tukey’s HSD correction for multiple comparisons and were analyzed independently for each metric. The results of these analyses are summarized in Figure 2. During data collection, an improper barrier (Figure 1a) setup resulted in biased resting forces for a few trials. These trials were excluded from the analysis to ensure an accurate representation of forces across all trials, resulting in a remaining total of 190 surface trials and 192 depth trials.

III. **Results**

![](Fig. 2. Confidence intervals for pairwise comparisons of metrics across surface (blue) and depth (orange) trials. The smoothness measures do not have a caption as they are dimensionless and not significant. Non-significant differences are indicated by confidence intervals intersecting the zero line and a point estimate indicated with a red x. Significant differences do not intersect zero and are indicated with a black circle point estimate.)

Our analyses indicated significant differences in mean metrics for all eight of the metrics in both surface and depth conditions when comparing experts to novices as well as novices to intermediates. The most consistent results were found with the LDLY metrics, demonstrating differences between experts and intermediates in both trial conditions, as demonstrated in the top row of Figure 2. We observed a small difference between the mean number of Y pks scores for experts and intermediates in the surface condition, but this difference was not seen in
the depth condition. These results provide topical evidence to suggest that LDLY metrics could be instrumental variables in assessing suturing skill level.

IV. DISCUSSION

Proper suturing technique emphasizes the importance of appropriate force application to minimize tissue damage and ensure safe outcomes. Assessing the quality of force application during suturing is crucial for understanding and refining surgical suturing skills. However, traditional metrics like maximum and range of forces offer limited insights and fail to capture the dynamics of how force is applied throughout the suturing process, providing insufficient detail for in-depth skill assessment and feedback. Furthermore, inconsistencies in their effectiveness have been documented in existing research [6]–[8], [10]. To effectively evaluate force in surgical tasks, metrics characterizing the entire force profile are crucial. To gain a deeper understanding of force, we incorporate metrics commonly applied to kinematic motion data to analyze force profiles measured during suturing.

All metrics in this study demonstrated moderate success, with significant differences in mean metric scores between novices and the other groups. $Y_{pks}$ significantly differentiated in all comparisons but expert and intermediate depth scores. The harmonic motions inherent to the fixed membrane during applied needle force are likely enhanced and reflected in $Y_{pks}$, with a larger amount of peaks suggesting a lack of control in the exertion of force. $pk_s$, a preliminary smoothness measure that inspired $Y_{pks}$, has demonstrated improved results over the more rudimentary measures of path length and average velocity [15]. Ultimately, the measure was outclassed by the renowned smoothness metrics of SPARC and LDLJ.

As motion smoothness is increasingly recognized as a robust tool for quantifying skilled surgical motions, we similarly hypothesize that respect for tissue implies a smooth application of force. Inspired by cutting-edge motion smoothness metrics, we developed and applied novel force smoothness metrics: $SPARC_F$ and LDLY. While $SPARC_F$ mean scores did not reveal significant differences between experts and intermediates, its consistent results in separating novice scores from the other groups demonstrate potential for further application of the metric. The most robust results emerged with $LDLY$, with significantly different mean scores across all comparisons and trial conditions and notably larger differences in expert and intermediate mean $LDLY_o$ scores. Given the potential association between smooth orthogonal forces and minimized tissue stress, $LDLY_o$ emerges as a strong candidate for feedback. By adapting and refining established motion smoothness metrics for force analysis, this study strove to provide a more nuanced and accurate assessment of suturing skill.

Overall, the metrics developed in this study, particularly $LDLY$, showcase the immense potential of assessing skillful force application during suturing. These superior findings highlight the need for sophisticated metrics to effectively measure all aspects of a surgical procedure, in this case, force, for proper training. Such results warrant further investigation and validation to confirm their efficacy in assessing surgical skill and providing objective, targeted feedback for skill development.

REFERENCES