Evaluating Smoothness of Force for Surgical Skill Assessment

Simar P. Singh[1], Amir Mehdi Shayan[1], Jianxin Gao[2], Joe Bible[3], Richard E Groff[2], and Ravikiran Singapogu[1]

[1]Department of Bioengineering, Clemson University
[2]Department of Electrical and Computer Engineering, Clemson University
[3]Department of Mathematical and Statistical Sciences, Clemson University

April 27, 2024

Evaluating Smoothness of Force for Surgical Skill Assessment

Simar P. Singh[1], Amir Mehdi Shayan[1], Jianxin Gao[2], Joe Bible[3], Richard E. Groff[2], and Ravikiran Singapogu[1]
[1]Department of Bioengineering, Clemson University, Clemson, SC, USA
[2]Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA
[3]Department 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. Medical training simulators, praised for objective, automated skill assessment, require instrumented sensors and relevant metrics for targeted feedback on important 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 metric scores for subjects with different 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 seek to quantify the widely used category of assessment in surgery, "respect for tissue", potentially benefiting surgical education with an improved evaluation of this aspect of the 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]. In vascular anastomosis, where delicate vessel connections are crucial, common complications like bleeding and tearing can lead to patient morbidity [2]. Surgeons must apply minimal force on surrounding tissues during the suturing process, avoiding excessive pulling that could tear fragile tissues. Consequently, "respect for tissue" is a universal directive in surgery and is included as a skill rating category in the Objective Structured Assessment of Technical Skill (OSATS), a well-substantiated rating scale to evaluate medical trainees [3].

Assessing respect for tissue currently 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. sensors offer a more precise solution, capturing subtle dynamics invisible to the eye. Medical training simulators provide a controlled environment ideal for integrating such objective sensors, paving the way towards automated and targeted feedback to improve trainee skills.

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 [4], results become inconsistent when distinguishing resident performance from attending surgeons or across different resident PGY levels [4], [5]. Previous work on the SutureCoach by Kil and colleagues [6] compared traditional force metrics between attendings and residents, observing significant differences in the z and orthogonal directions. 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. [7]. 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 [8]. The SutureCoach interfaces intricate sensors for a holistic evaluation of open vascular suturing skills on a clock face pattern modeled after the Fundamentals of Vascular Surgery (FVS) [9]. This study will focus on 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.

The research in this publication was supported by an R01 grant (5R01HL146843) from the National Institutes of Health.

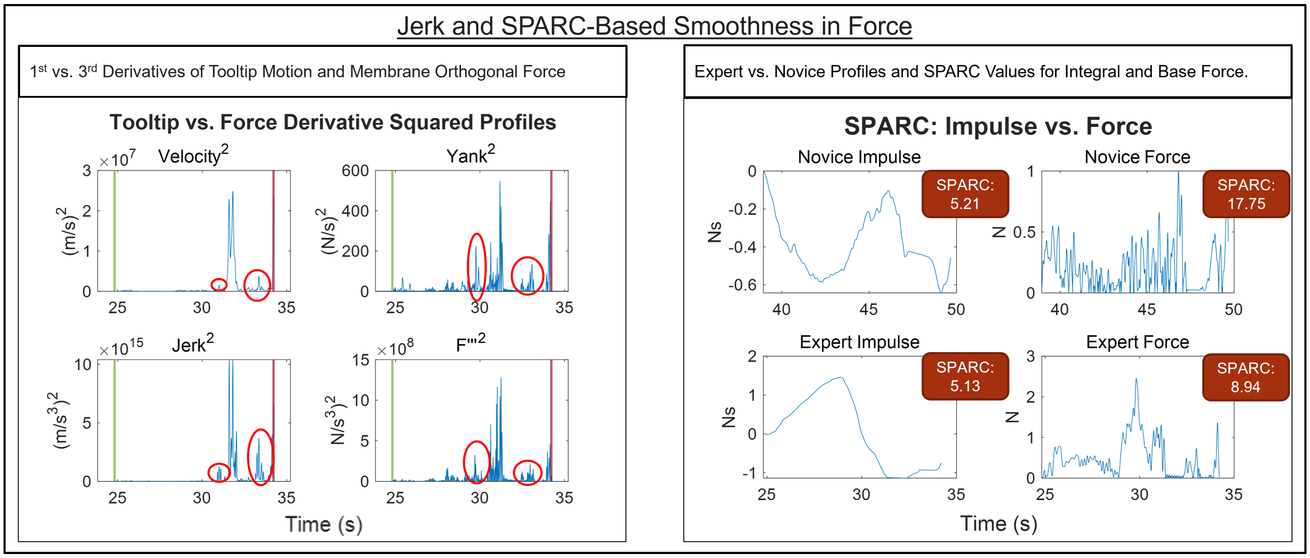

Fig. 1. **(a.)**: A subject suturing on the SutureCoach platform. **(b.)**: Transformation of forces to align with the suture direction. The left image depicts a bottom-up view of a needle detected by the internal camera (custom-filtered for illustrative purposes), where x and y force coordinates (F_x and F_y white vectors) are transformed into orthogonal and tangential forces (F_{ort} and F_{tan} red vectors). The right image depicts a cross-sectional view of needle entry, highlighting the orthogonal torque (T_{ort} curved red arrow) in relation to F_{tan} and z force (F_z). **(c.)**: A comparison of 1st and 3rd derivative tooltip motion vs. force squared profiles. The circles highlight characteristics present in velocity that are emphasized in jerk, whereas circled characteristics in yank are diminished in the third derivative. **(d.)**: Expert vs. novice integral of force (impulse) and force profiles, highlighting the analogous behavior of base force to velocity. This conjecture is further evidenced by the respective SPARC results, demonstrating an ambiguous measure when applied to impulse.

II. Methods

A. *SutureCoach Description and Participant Information*

The SutureCoach platform records data from multiple sensors 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 attending 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 [6]. 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 alongside 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 similar 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 [8], [10], [11] to the domain of force measurement, specifically within the context of suturing. Notably, Trejos et al. [7] and Balasubramanian et al. [11] 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. Motion smoothness metrics evaluate changes in continuous signals while accounting for hesitation through intermittency in the reading, however force sensors measure zero forces outside of contact, creating apparent intermittency that negatively impacts the metric. We address this in our study by focusing 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 limitation.

Building upon Trejos et al.'s [7] pioneering work in adapting movement metrics for force analysis in surgical skills [7], we further refine smoothness calculations in light of recent advancements in the field [11], [12]. We begin by applying Log Dimensionless Jerk (*LDLJ*), a metric that uses the third derivative of position to evaluate smoothness [11]. We observed that the first derivative of force characterizes smooth vs. unsmooth behavior better while higher derivatives of force do not, unlike with motion; see 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. [13] in a commentary on its successful applications in sensorimotor systems. As noted in their work, given $F = ma$, yank inherently aligns with the third derivative of position, i.e., jerk. 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 motion smoothness through submovement analysis in velocity, to force data [11]. Although the yank-jerk analogy might suggest calculating *SPARC* on the integral of force (impulse), this approach would yield an ambiguous measure of submovement behavior more indicative of material deformation than a measure of controlled force. As noted by Melendez-Calderon et al. [12], haphazardly applying *SPARC* to a random signal would be ill-suited due to the lack of interpretability of the variations in the signal, as *SPARC* requires a profile that represents the "speed" of a given movement. To align with these requirements, we propose applying *SPARC* directly to force itself ($SPARC_F$), since the force profile can be conceptually subdivided into movements in a manner analogous to the velocity profile. $SPARC_F$ can provide meaningful and interpretable measures of the submovements, revealing 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 impulse signals display few submovements and result in similar SPARC values, while the force signals display clear submovements with different $SPARC_F$ values.

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 the subscripts $_{To}$, $_{o}$, 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. [11], who use maximum velocity for *LDLJ*. 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.

$$\mathit{LDLY} = \ln\left| \frac{T}{Fp+^2} \int_{t_{entry}}^{t_{exit}} \left(\frac{dF}{dt}\right)^2 dt \right| \tag{1}$$

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. ω_c is an adaptive cutoff frequency, with further detail found in [11].

$$SPARC_F = \int_0^{\omega_c} \left[\left(\frac{1}{\omega_c}\right)^2 + \left(\frac{d\hat{F}(\omega)}{d\omega}\right) \right]^{\frac{1}{2}} d\omega; \tag{2}$$

$$\hat{F}(\omega) = \frac{F(\omega)}{F(0)}, \quad F(\omega) = \mathcal{F}\{F(t)\}$$

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 and 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, inconsistencies in the effectiveness of traditional metrics like maximum and range of forces have been documented in existing research [4], [5], [7]. 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.

Our analyses demonstrated 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

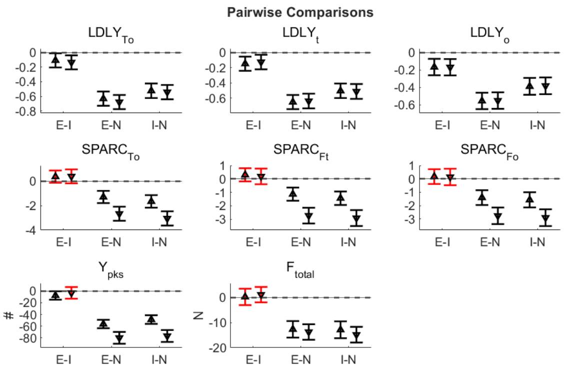

Fig. 2. Confidence intervals for pairwise comparisons of metrics for surface and depth trials (indicated by upwards and downwards triang point estimates). Each interval compares the scores between two of the three groups of clinical expertise: Expert (E), Intermediate (I), and Novice (N). The smoothness measures do not have a caption, as they are dimensionless. Non-significant differences are indicated by confidence intervals intersecting the zero line and colored red. Significant differences are colored black.

novices to intermediates, indicating their lesser skill level and validating base assessment ability with the new metrics. 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. Y_{pks} further significantly differentiated expert and intermediate surface scores. The harmonic motions inherent to the fixed membrane during applied needle force are likely enhanced and reflected in Y_{pks}. Higher values suggest poorer control of forces while suturing.

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 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. Notably larger differences were observed with 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.

IV. Conclusion

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

[1] J. D. Birkmeyer, J. F. Finks, A. O'Reilly, M. Oerline, A. M. Carlin, A. R. Nunn, J. Dimick, M. Banerjee, and N. J. Birkmeyer, "Surgical Skill and Complication Rates after Bariatric Surgery," *New England Journal of Medicine*, vol. 369, no. 15, pp. 1434–1442, Oct. 2013, clinical outcomes in bariatric surgery correlated with surgeon skill.

[2] K. M. Bennett, K. C. Kent, J. Schumacher, C. C. Greenberg, and J. E. Scarborough, "Targeting the Most Important Complications in Vascular Surgery," *Journal of Vascular Surgery*, vol. 65, no. 3, pp. 793–803, Mar. 2017.

[3] J. A. Martin, G. Regehr, R. Reznick, H. Macrae, J. Murnaghan, C. Hutchison, and M. Brown, "Objective Structured Assessment of Technical Skill (OSATS) for Surgical Residents: Objective Structured Assessment of Technical Skill," *British Journal of Surgery*, vol. 84, no. 2, pp. 273–278, Feb. 1997.

[4] S. Montoya-Alvarez, A. Minor-Martínez, R. M. Ordorica-Flores, L. Padilla-Sánchez, J. Tapia-Jurado, and F. Pérez-Escamirosa, "Construct Validity of the Surgforce System for Objective Assessment of Laparoscopic Suturing Skills," *Surgical Endoscopy*, vol. 34, no. 11, pp. 5188–5199, Nov. 2020.

[5] A. Dubrowski, R. Sidhu, J. Park, and H. Carnahan, "Quantification of Motion Characteristics and Forces Applied to Tissues During Suturing," *The American Journal of Surgery*, vol. 190, no. 1, pp. 131–136, Jul. 2005.

[6] I. Kil, J. F. Eidt, R. E. Groff, and R. B. Singapogu, "Assessment of Open Surgery Suturing Skill: Simulator Platform, Force-Based, and Motion-Based Metrics," *Frontiers in Medicine*, vol. 9, p. 897219, Aug. 2022.

[7] A. L. Trejos, R. V. Patel, R. A. Malthaner, and C. M. Schlachta, "Development of Force-Based Metrics for Skills Assessment in Minimally Invasive Surgery," *Surgical Endoscopy*, vol. 28, no. 7, pp. 2106–2119, Jul. 2014.

[8] A. M. Shayan, S. Singh, J. Gao, R. E. Groff, J. Bible, J. F. Eidt, M. Sheahan, S. S. Gandhi, J. V. Blas, and R. Singapogu, "Measuring Hand Movement for Suturing Skill Assessment: A Simulation-Based Study," *Surgery*, p. S0039606023004312, Aug. 2023.

[9] M. Sheahan, A. M. Lalani, J. T. Lee, M. Shames, D. Rigberg, B. Cass, C. Sheahan, and J. Bismuth, "The Fundamentals of Vascular Surgery: When Do Vascular Trainees Achieve Basic Open Surgical Competency?" *Journal of Vascular Surgery*, vol. 72, no. 1, pp. e24–e25, Jul. 2020.

[10] S. Singh, J. Bible, Z. Liu, Z. Zhang, and R. Singapogu, "Motion Smoothness Metrics for Cannulation Skill Assessment: What Factors Matter?" *Frontiers in Robotics and AI*, vol. 8, 2021, publisher: Frontiers.

[11] S. Balasubramanian, A. Melendez-Calderon, A. Roby-Brami, and E. Burdet, "On the Analysis of Movement Smoothness," *Journal of NeuroEngineering and Rehabilitation*, vol. 12, no. 1, p. 112, Dec. 2015.

[12] A. Melendez-Calderon, C. Shirota, and S. Balasubramanian, "Estimating Movement Smoothness From Inertial Measurement Units," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 558771, Jan. 2021.

[13] D. C. Lin, C. P. McGowan, K. P. Blum, and L. H. Ting, "Yank: The Time Derivative of Force Is an Important Biomechanical Variable in Sensorimotor Systems," *Journal of Experimental Biology*, vol. 222, no. 18, p. jeb180414, Sep. 2019.