*Healthcare and Medical Devices,* Vol. 51, 2022, 58–66
https://doi.org/10.54941/ahfe1002099

# Phase-Based Assessment of Arthroscopic Skill Using Motion Smoothness Metrics: A Simulator-Based Proof-of-Concept Study

**Simar Singh, Chelse VanAtter, and Ravikiran Singapogu**

Department of Bioengineering, Clemson University, SC 29634, USA

## ABSTRACT

Minimally invasive surgeries are meticulous procedures that require complex movement within a limited range of motion, requiring intensive training. Medical training simulators often have limited sensing modalities, restricting the quality of metrics to quantify skill. We created a multi-modal 3D printed simulator that is affordable and easily replicable for remote, automated, and phase-based skill assessment in arthroscopy, a type of minimally invasive surgery. In this pilot study, four subjects of non-medical experience levels performed a peg transfer task in the arthroscopy simulator with synchronized motion and video. The task is segmented into phases to determine the relative efficacy of motion smoothness and other calculated metrics. One phase showed more significant differences in metrics than the other phases, demonstrating the potential for phase-based skill evaluation in tasks with more complex maneuvers. Our novel simulator design allowed for metrics computation at a phase-based level, with initial results demonstrating its importance.

**Keywords:** Motion smoothness, Phase segmentation, Arthroscopy, Medical training simulators, Log dimensionless jerk, Spectral arc length

## INTRODUCTION

### Simulators

Minimally invasive surgical techniques have increased in popularity in orthopedic surgery due to faster recovery time and less patient trauma (Garrett et al. 2006). Arthroscopy, a minimally invasive surgery performed on a joint, is used for various surgical operations ranging from partial meniscectomies to torn rotator cuff surgeries. However, arthroscopy has a steep learning curve because of the reduced operating range of motion, increased dexterity due to tools, and the challenge of accurately interpreting 2D images within the 3D anatomical space. Training does seem to play a role in improved clinical outcomes, with experienced surgeons demonstrating shorter operating times (Farnworth et al. 2001). Consequently, there is a need for novice surgeons to have access to more efficient learning methods for becoming proficient in arthroscopy.

 

Medical training simulators have seen an increase in popularity due to their ability to provide efficient practice without a live subject and have been shown to improve basic skills for arthroscopy (Kholinne et al. 2018). As a result, the Fundamentals of Arthroscopy Surgery Training (**FAST**) training simulator was created to learn arthroscopic skills in a simple model. The simulator consists of several modules with tasks like peg transfer to train different aspects of arthroscopic surgery, with the total time taken to complete the task and the number of failures used for skill assessment (Goya et al. 2016). Although these two metrics differentiate between novice and expert surgeons, surgical skill is multi-faceted and nuanced; therefore, the reliance on overall time and failures as primary metrics is rudimentary. There is a need for advanced metrics to quantify skills in arthroscopic simulators that capture the complex nature of motions, forces, and decisions involved in skilled surgery. Due to its relatively recent development, very few groups have researched possibilities for other metrics for the **FAST** trainer thus far.

This study aimed to explore the feasibility of advanced metrics that quantify surgical dexterity through the smoothness of motion in an affordable 3D-printed simulator similar to the **FAST** trainer. Additionally, we also investigated the utility of these metrics in three distinct phases of the peg transfer task. The rationale for this is that there are critical sections where smooth, steady motion is essential during a surgical procedure. On the contrary, there are sections where skilled movement is less important. However, many metrics used in skill assessment are evaluated for the entire task. To facilitate a more robust and interpretable approach to skill evaluation, we explored the value of segmenting the task into multiple phases to isolate critical moments of the task.

## Motion Smoothness

Intuitively, one can readily appreciate that a surgeon should have superior dexterity demonstrated by smooth motion during minimally invasive surgeries. Several previous simulator-based studies have used metrics to capture aspects of dexterity such as path length, the economy of motion, and the number of collisions to determine the level of skill (Ström et al. 2003); however, these metrics are limited in what aspects and to what degree they quantify dexterity in this context. Recent research has revealed the utility of motion smoothness metrics to capture more of the complexity of motion during surgery. Towards this, two motion smoothness metrics have been identified as candidates to evaluate skill in arthroscopy: log dimensionless jerk (*LDLJ*) and spectral arc length (*SPARC*) (Balasubramanian et al. 2015). *LDLJ* relies on the notion that ideal, smooth motion results in minimal discontinuities in acceleration. The measure was initially developed to measure stroke recovery and saw rapid improvements to create a robust measure of motion smoothness (Balasubramanian et al. 2009; Hogan & Sternad 2009; Melendez-Calderon et al. 2020).

In addition to improving jerk's inherent sensitivity to noise, Balasubramanian et al. present a novel metric (called *SPARC*) to calculate motion smoothness by using the arc length of the amplitude of the frequency-normalized

Fourier magnitude spectrum of the velocity profile (Balasubramanian et al. 2012; Hogan & Sternad 2009). Like jerk, this metric relies on the idea that smooth movements will result in lower frequencies, and unsmooth movements will result in larger magnitudes of movements of several different frequencies.

### Previous Work

Although older motion smoothness metrics have successfully differentiated skill in several surgical applications, the more recent formulations have seen fewer uses. A study in 2016 employed dimensionless square jerk (*DLJ*) and *SPARC*, demonstrating correlations of those metrics to scores from a global rating scale for endovascular performance in a Fundamentals of Endovascular Skills (**FEVS**) trainer (Estrada et al. 2016), with newer studies with the same trainer and group continuing to examine the use of *SPARC* for skill assessment (Belroy et al. 2020; O'Malley et al. 2019). In a neurosurgery simulator, *DLJ* differentiated between expert and non-expert performance and among different task constraints in a pegboard placement task (Ghasemloonia et al. 2017), a task similar to a module found in the **FAST** simulator and the focus of this pilot study. In 2018, the first study to our knowledge that quantifies movement skill for shoulder arthroscopy measured hand motions between novice and expert groups in a suture anchoring task (Kholinne et al. 2018), a task similar to another module found in the **FAST** simulator reporting positive results in *DLJ* and time, but not path length.

Given that many formulations of motion smoothness metrics are relatively recent, only a few studies sought to apply these methods for surgical skill assessment. Furthermore, these metrics must be rigorously tested in specific applications that require skilled movement. For example, although initially presented as a superior metric to *LDLJ* due to its insensitivity to noise, *SPARC* demonstrated worse performance in rotational motion smoothness of hand motions (Melendez-Calderon et al. 2020). On the other hand, our recent study found no significant differences in motion smoothness metrics' and other metrics' association with cannulation skill (Singh et al. 2021). Therefore, there exists a need for more testing to understand the efficacy of motion smoothness metrics and their application to surgical skills assessment.

## METHODS

### Simulator and Design

We designed and implemented a custom, affordable, 3D-printed arthroscopic simulator based on the FAST simulator (seen in Figure 1). The simulator comprises a platform with a removable dome with holes to simulate incisions made during arthroscopic surgery. The platform's base has replaceable plates designed for the simulator's various tasks: a ball-in-maze navigation task, a peg transfer task, a shape tracing task, a partial meniscectomy task, and a loose body removal task. This preliminary study was limited to the peg

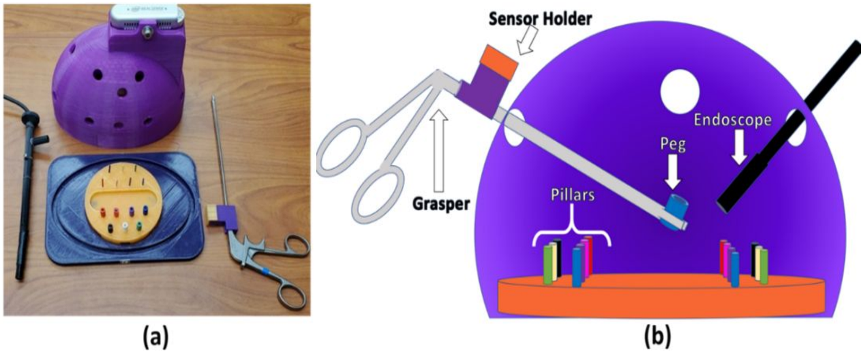


**Figure 1**: (a) Shows an image of the created workstation and (b) shows an illustrated sketch of the peg transfer process.

transfer task, composed of a gap in the center and symmetric, colored pillars on both sides. Subjects were to pick up and transfer colored pegs across the gap to the pillar of the same color. Subjects used an arthroscopic tool (ACUFEX Grasper Handle, Smith & Nephew) to pick up and transfer the pegs. A custom-made endoscope was created from a borescope camera (Depstech Inc.) to view inside the dome. A depth camera (Intel RealSense D435) was mounted on the dome to record video at 60 fps. An electromagnetic (EM) sensor (trakSTAR, Northern Digital Inc.) was mounted on the grasper to measure position and orientation at 60 Hz. The EM sensor position and camera data were recorded, timestamped, and synchronized through custom C++ code. Ferromagnetic metals and electromagnetic distortion sources were removed from the operating environment to reduce metallic interference with the EM sensor. A 3D printed sensor holder was attached to hold the EM sensor at a distance to prevent interference from the grasper. Tooltip calibration was performed to detect motion at the tip of the tool.

Four participants in this pilot study were separated into two groups: two participants with over 20 hours of experience on the simulator (deemed experienced) and two participants with no experience. Participants were asked to transfer the four pegs closest to the gap (considered more accessible pegs to transfer) twice. In total, sixteen pegs were transferred for the experienced participant group and fourteen for the inexperienced group, as two incomplete peg transfers were removed for data analysis.

## Segmentation and Statistical Analysis

Balasubramanian et al. point out that the applications of motion smoothness metrics are task-dependent; i.e., the measurements will not have universal "smooth" and "unsmooth" values and should have constrained start and end times for meaningful interpretation (2015). Therefore, each trial was separated into individual peg transfers (**Attempts**) through manual inspection of the synchronized camera video. By isolating the task into individual **Attempts**, we focused solely on the peg transfer movement and eliminated any movements that were not pertinent to the peg transfer tasks. We further separated

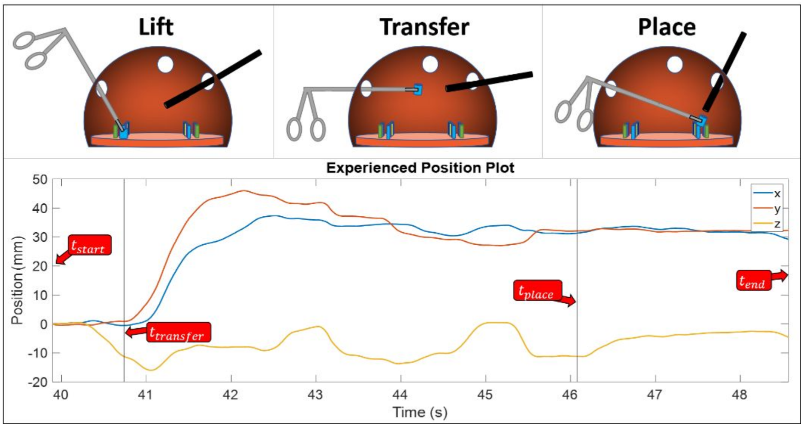


**Figure 2:** The illustration shows each phase of peg transfer along with an experienced user's peg transfer position plot normalized at the start of the trial, with vertical lines signifying phase transitions.

each **Attempt** into three distinct phases (defined below) to analyze metric effectiveness at all levels of the task to explore phase-based task segmentation.

- **Attempt:** $t_{\text{end}} - t_{\text{start}}$
- **Lift:** $t_{\text{transfer}} - t_{\text{start}}$
- **Transfer:** $t_{\text{place}} - t_{\text{transfer}}$
- **Place:** $t_{\text{end}} - t_{\text{palce}}$

$t_{\text{start}}$ is the timestamp the grasper grasps a peg, $t_{\text{transfer}}$ is the timestamp horizontal movement is seen, $t_{\text{place}}$ is the timestamp horizontal movement ceases during a successful **Place**, and $t_{\text{end}}$ is the timestamp the peg stops moving. An illustrated peg transfer process and an example position plot are shown in Figure 2.

The metrics used in this study are defined in Table 1. Data post-processing was performed through MATLAB (v. 2019b). For data processing, a Savitzky-Golay filter of order 3 and a window span of a quarter of a second (15) was used (Singh et al. 2021). Statistical analysis was accomplished through RStudio (v. 1.2.1335). Parametric $T$-tests ($\alpha = 0.05$) were computed following Shapiro-Wilk normality tests. Data analysis revealed two outliers as these trials took around three minutes for completion in contrast with an average completion time of 12.4 seconds (without the two outliers). Consequently, these outliers were removed, resulting in 16 experienced and 12 inexperienced **Attempts**.

## RESULTS

The results of the $t$-tests are shown in Table 2, and the comparisons of mean $\pm$ standard deviation of metric values for both groups separated by phase are shown in Figure 3. All metrics determined a significant difference of means between the two groups in **Attempt**. As hypothesized, most metrics demonstrated significant differences in the **Transfer** phase since much of the

**Table 1.** List of metrics and their descriptions.

| Name | Description | Equation |
|---|---|---|
| Time ($T$) | The total time of the current phase. | $T = t_{phase\_end} - t_{phase\_start}$ |
| Peaks ($Pks$) | The number of local maxima (peaks) in the velocity profile, computed using the built-in MATLAB function. | $Pks = findpeaks\left(\frac{dX}{dt}\right)$; $X = \sqrt{x^2 + y^2 + z^2}$ |
| Path Length ($PL$) | The sum of Euclidean distances between points traversed by the grasper tip. | $PL = \int_{t_{phase\_start}}^{t_{phase\_end}} \frac{dX}{dt}$ |
| Log Dimensionless Jerk ($LDLJ$) | The natural log of jerk integrated and squared. | $-\ln\left\lvert \frac{T^5}{PL^2} \int_{t_{phase\_start}}^{t_{phase\_end}} \left(\frac{d^3X}{dt^3}\right)^2 dt \right\rvert$ |
| Spectral Arc Length ($SPARC$) | As defined in (Balasubramanian, S. et al., 2012), we computed spectral arc length, the arc length of the Fourier transform of the velocity profile, from the provided MATLAB code. | $SPARC = -\int_0^{\omega_c} \left[\left(\frac{1}{\omega_c}\right)^2 + \left(\frac{d\widehat{V}(\omega)}{d\omega}\right)\right]^{\frac{1}{2}} d\omega$; $\widehat{V}(\omega) = \frac{V(\omega)}{V(0)}$ |

**Table 2.** Confidence intervals of t-test significant differences across each phase. Significantly different intervals are bold and noted with a*.

| Confidence Intervals for Difference of Means | | | | |
|---|---|---|---|---|
| | Attempt | Lift | Transfer | Place |
| *LDLJ* | **[1.39 4.59]*** | **[1.05 5.16]*** | **[1.19 5.10]*** | **[0.06 4.84]*** |
| *SPARC* | **[0.62 3.50]*** | [−0.62 1.47] | [0.01 0.53] | [−1.06 2.40] |
| *T* | **[1.64 16.3]*** | [−1.66 0.38] | **[1.07 11.9]*** | [−4.48 0.86] |
| *Pks* | **[10.4 110]*** | [−12.2 2.02] | **[8.00 75.5]*** | [−6.35 33.2] |
| *PL* | **[14.5 151]*** | [−12.0 15.5] | **[8.72 127]*** | [−3.84 30.1] |

variation in motion is seen in this phase. However, *LDLJ* is the sole metric to differentiate between the pilot experienced and novice peg transfers across the phases.

## DISCUSSION

We wish to reiterate that participants in this study had minimal differences in experience on the simulator (twenty hours of experience on the simulator vs. no experience); however, all metrics demonstrated efficacy in distinguishing between the two groups even with this limitation. As shown in Figure 3, the inexperienced group had consistently higher variations indicative of inferior performance across all metrics. Additionally, as demonstrated in Table 2, *LDLJ* determined significant differences between the two groups at all levels, demonstrating its superiority as a metric as it can distinguish across minimal

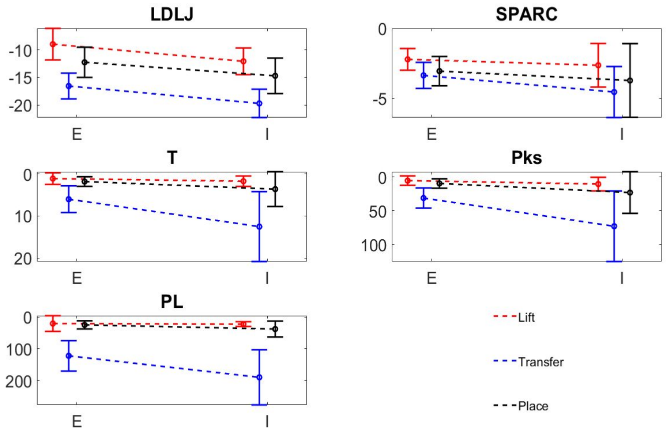


**Figure 3:** A comparison of experienced (E) and inexperienced (I) means ± standard deviations of the different metric values at each phase separated by subplots with a dotted line drawn through the means signifies their differences. A decrease in the y-axis represents a worse performance in that metric.

levels of movements between the two groups in contrast to the other metrics. Therefore, motion smoothness metrics seem to provide a richer characterization of the dexterity required for surgery. However, the metrics' utility and applicability to surgical skill assessment and learning are yet to be fully explored.

The primary goal of this study was to examine skill differences at the subtask level by identifying and isolating critical phases of the peg transfer task. The **Transfer** phase contained worse metric values and higher standard deviations than the other phases for both groups. All metrics but *SPARC* demonstrated significant differences between the two groups at this phase. In previous studies, *SPARC* has demonstrated superior evaluation of surgical skill in comparison to metrics like *T* and *PL*. The lack of *SPARC*'s significant differences can be explained through a small sample size and minimal differences in skill levels between the two groups in the pilot study. We anticipate that future studies examining subjects with more significant skill differences combined with a larger sample size will yield more robust and generalizable results. Additionally, applying phase-based segmentation with tasks with more complex maneuvers will provide greater insight into inter-metric reliability and strength.

In this study, we present the design of a custom novel and affordable 3D printed arthroscopic simulator that incorporates two sensor modalities that could potentially be used for objective, automated skill assessment through sensor metrics. Conventional "box" surgical trainers rely on rudimentary metrics like time and number of errors to evaluate skill. By providing multimodal capture of skill, "smart" simulators allow for structured, meaningful, and individualized training, minimizing the need for expert assessors with time and cost implications.

Additionally, we also present the methodology of isolating a basic and commonly used task in surgical skill assessment into critical phases, which may help to evaluate skill better. This method excludes unimportant phases and computes metrics in known phases, allowing for meaningful metrics interpretation. For example, focusing solely on cutting tissue or suturing by removing the sections between cutting or suturing allows targeted metric computations for motion smoothness. This study indicates that these metrics may help evaluate skills between experts and non-experts with clinical experience. Furthermore, these results also signal the potential of task segmentation to analyze skills at different subtask levels.

## ACKNOWLEDGMENT

Research reported in this publication was supported by the NIGMS of the National Institutes of Health under award number P20 GM121342. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## REFERENCES

Balasubramanian, S., Melendez-Calderon, A., & Burdet, E. (2012). A Robust and Sensitive Metric for Quantifying Movement Smoothness. *IEEE Transactions on Biomedical Engineering*, *59*(8), 2126–2136.

Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A., & Burdet, E. (2015). On the Analysis of Movement Smoothness. *Journal of NeuroEngineering and Rehabilitation*, *12*(1), 112.

Belvroy, V. M., Murali, B., Sheahan, M. G., O'Malley, M. K., & Bismuth, J. (2020). Motion Metrics Reliably Differentiate Competency: Fundamentals of Endovascular and Vascular Surgery. *Journal of Vascular Surgery*, S0741521420304754.

Estrada, S., Duran, C., Schulz, D., Bismuth, J., Byrne, M. D., & O'Malley, M. K. (2016). Smoothness of Surgical Tool Tip Motion Correlates to Skill in Endovascular Tasks. *IEEE Transactions on Human-Machine Systems*, *46*(5), 647–659.

Farnworth, L. R., Lemay, D. E., Wooldridge, T., Mabrey, J. D., Blaschak, M., DeCoster, T. A., Wascher, D. C., & Schenck, R. C. (2001). A Comparison of Operative Times in Arthroscopic ACL Reconstruction Between Orthopaedic Faculty and Residents: The Financial Impact of Orthopaedic Surgical Training in the Operating Room. *The Iowa Orthopaedic Journal*, *21*, 31–35.

Garrett, W., Swiontkowski, M., Weinstein, J., Callaghan, J., Rosier, R., Berry, D., Harrast, J., & Derosa, G. (2006). American Board of Orthopaedic Surgery Practice of the Orthopaedic Surgeon: Part - II, Certification Examination Case Mix. *The Journal of Bone and Joint Surgery-American Volume*, *88*(3), 660–667.

Ghasemloonia, A., Maddahi, Y., Zareinia, K., Lama, S., Dort, J. C., & Sutherland, G. R. (2017). Surgical Skill Assessment Using Motion Quality and Smoothness. *Journal of Surgical Education*, *74*(2), 295–305.

Goyal, S., Radi, M. A., Ramadan, I. K., & Said, H. G. (2016). Arthroscopic Skills Assessment and Use of Box Model for Training in Arthroscopic Surgery Using Sawbones – "FAST" Workstation. *SICOT-J*, *2*.

Hogan, N., & Sternad, D. (2009). Sensitivity of Smoothness Measures to Movement Duration, Amplitude and Arrests. *Journal of Motor Behavior*, *41*(6), 529–534.

Kholinne, E., Gandhi, M. J., Adikrishna, A., Hong, H., Kim, H., Hong, J., & Jeon, I.-H. (2018). The Dimensionless Squared Jerk: An Objective Parameter That Improves Assessment of Hand Motion Analysis during Simulated Shoulder Arthroscopy. *BioMed Research International*, *2018*, 1–8.

Melendez-Calderon, A., Shirota, C., & Balasubramanian, S. (2020). *Estimating Movement Smoothness from Inertial Measurement Units* [Preprint]. Bioengineering.

O'Malley, M. K., Byrne, M. D., Estrada, S., Duran, C., Schulz, D., & Bismuth, J. (2019). Expert Surgeons Can Smoothly Control Robotic Tools With a Discrete Control Interface. *IEEE Transactions on Human-Machine Systems*, *49*(4), 388–394.

Singh, S., Bible, J., Liu, Z., Zhang, Z., & Singapogu, R. (2021). Motion Smoothness Metrics for Cannulation Skill Assessment: What Factors Matter? *Frontiers in Robotics and AI*, *8*, 625003.

Ström, P., Kjellin, A., Hedman, L., Johnson, E., Wredmark, T., & Fellander-Tsai, L. (2003). Validation and Learning in the Procedicus KSA Virtual Reality Surgical Simulator. *Surgical Endoscopy*, *17*(2), 227–231.https://doi.org/10.1007/s00464-002-9078-1