2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob) | 978-1-6654-5849-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/BIOROB52689.2022.9925380

Toward Correcting Anxious Movements Using Haptic Cues on the Da Vinci Surgical Robot

Yi Zheng[1], Marzieh Ershad[2], and Ann Majewicz Fey[1,3]

Abstract— **Surgical movements have an important stylistic quality that individuals without formal surgical training can use to identify expertise. In our prior work, we sought to characterize quantitative metrics associated with surgical style and developed a near-real-time detection framework for stylistic deficiencies using a commercial haptic device. In this paper, we implement bimanual stylistic detection on the da Vinci Research Kit (dVRK) and focus on one stylistic deficiency, "Anxious", which may describe movements under stressful conditions. Our goal is to potentially correct these "Anxious" movements by exploring the effects of three different types of haptic cues (time-variant spring, damper, and spring-damper feedback) on performance during a basic surgical training task using the da Vinci Research Kit (dVRK). Eight subjects were recruited to complete peg transfer tasks using a randomized order of haptic cues and with baseline trials between each task. Overall, all cues lead to a significant improvement over baseline economy of volume and time-variant spring haptic cues lead to significant improvements in reducing the classified "Anxious" movements and corresponded with significantly lower path length and economy of volume for the non-dominant hand. This work is the first step in evaluating our stylistic detection model on a surgical robot and could lay the groundwork for future methods to reduce the negative effect of stress actively and adaptively in the operating room.**

Index Terms— **Human Performance Augmentation, Surgical Robotics: Laparoscopy, Haptics and Haptic Interfaces**

I. INTRODUCTION

There is a direct relationship between surgical outcomes and surgeon skill level [1]–[7]. Thus, providing useful and meaningful feedback to the trainee is critical for patient safety. Traditional surgical skill assessment methods involve an observational approach where a senior surgeon observes a trainee and provides verbal feedback [8]. In recognizing surgical mastery, you often "know it when you see it" - an observation that has led to innovative developments for using crowd-sources to quantify surgical skill assessment [9]–[11]. Other quantitative and data-driven tools are also enabling the next phase of research in surgical skill assessment [12]–[16].

Beyond skill assessment, surgical simulators have long been developed to enhance basic skills outside of the operating room [17]; however, there are concerns that simulators lack meaningful feedback on how trainees should modify their movements to improve performance [18]. The addition of haptic feedback or cue has shown some benefits to simulator-based training [19]–[21], as well as training other human movements [22], [23]; however, this haptic cue is typically task-based (e.g., indicating instrument or tissue collisions) as well as static - meaning that it does not change as a function of trainee learning.

There is an opportunity to design more adaptive and personalized methods of haptic cues. For example, Enayati et al. used haptic guidance to enhance performance in a ring-and-rail following task by providing haptic cues related to optimal orientation of the ring [24]. We aim to adopt a similar strategy for haptic cues; however, our approach is designed to be more global and task-independent. We accomplish this by focusing on the stylistic behaviors that are associated with expert or novice-like movements (e.g., Fluid/Viscous, Crisp/Jittery, Calm/Anxious, among others); a method that was developed through the use of crowd-sourced assessment of surgical styles [11], [25], [26].

In this paper, we focus on a single type of stylistic behavior, namely "Calm/Anxious". This style could potentially be useful to detect, in near-real-time, stressful intra-operative events that could potentially negatively impact surgical performance and thus compromise patient safety [27]–[30].

As a first step in correcting the "Anxious" movements, we implement three stylistic haptic cues (described in Section II) on a open-source telerobotic platform, specifically, the da Vinci Research Kit (dVRK) and associated open-source software [31]–[34] to determine which cue is best at improving performance. This paper builds on our prior work [26] to perform kinematic-based stylistic detection by implementing the method on the dVRK and performing detection on both the dominant and non-dominant hand kinematic data.

II. BACKGROUND AND PRIOR WORK

A. Surgical Skill Assessment Using Stylistic Behavior

In prior work, we presented a novel surgical skill assessment method based on the surgeon's stylistic behavior [11]. We proposed a lexicon of surgical styles, informed by expert surgeons, including stylistic adjectives such as Fluid/Viscous, Smooth/Rough, and Calm/Anxious. We evaluated the ability of stylistic descriptors to differentiate between expertise levels using metrics correlated to crowd-sourced assessments of surgical styles obtained from training videos [25]. We also proposed an automatic, near-real-time method for detecting the quality of performance, based on these behavioral styles and within 0.25 seconds of movement data [26].

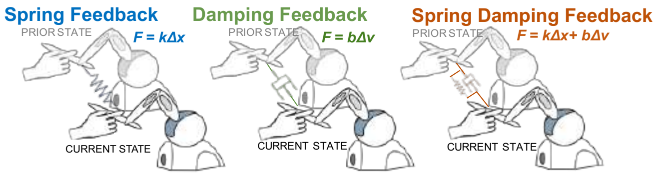
This work was supported in part by NSF grant #1846726 and the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH) under Award UL1TR001105 and 1R01EB030125-01.

[1]Yi Zheng, and Ann Majewicz Fey are with the Department of Mechanical Engineering, the University of Texas at Austin, 204 East Dean Keeton Street, Austin, TX 78712, USA. yi.zheng@austin.utexas.edu

[2]Marzieh Ershad is with Intuitive Surgical, Inc., 1020 Kifer Road Sunnyvale, CA 94086.

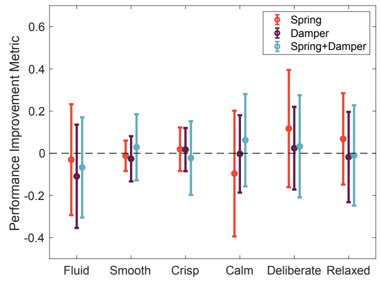
[3]Ann Majewicz Fey is also with the Department of Surgery, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390, USA. ann.majewiczfey@utexas.edu

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.



(a) Stylistic Force Feedback

(b) Experimental Setup



(c) Preliminary Results

Fig. 1: Preliminary work to compare the effects of three different types of haptic cues on six distinct stylistic behaviors detected with our previously developed near-real-time detection algorithm [26], [35]

B. Preliminary Experimental Study of Stylistic Haptic Cues

Most recently, we developed a preliminary adaptive training method consisting of the following elements: (1) a near-real-time detection of stylistic deficiencies in movement based on a crowd sourced assessment of style and a dictionary learning method, optimized for time and computational efficiency using principle component analysis and sparse coding [26] and (2) based on the detection, simple haptic cues (e.g., time-variant spring, damper, and spring-damper haptic cues based on current and prior user positions and velocities) are provided to the human operator to improve style (Fig 1) [35].

We evaluated three different haptic cues: (1) time-variant spring, (2) time-variant damper, and (3) time-variant spring-damper feedback (Fig 1a) across these six behavioral styles in a small human subject study using a commercial haptic device (Geomagic Touch, 3D Systems, SC) and a trajectory following task [36] (Fig 1b).

A generalized metric was developed to compare the performance improvements of the haptic cues over baseline movement trials. Results showed significant differences in changing performance for certain combinations of styles and types of haptic cues, e.g., time-variant spring led to better fluidity, time-variant damper improved crispness, time-variant spring-damper improved "Anxious" movements (Fig 1c). However, the simplistic nature of these uni-manual target reaching tasks could have obscured the potential training benefit of these haptic cues.

Therefore, in this paper, we decide to focus on only one pair of stylistic behavior - "Calm/Anxious" and aim to find the best haptic cue to improve this style during a more challenging and surgically relevant bimanual task.

III. METHODS

We developed a method to recognize the quality of movement through stylistic behavior and apply appropriate near-real-time haptic cues for correcting the "Anxious" movements using a surgical robot platform and a simulated task.

A. Data Acquisition

The da Vinci Research Kit (dVRK) was used in this study for both kinematic movement data acquisition and haptic feedback. Position, linear velocity, and angular velocity from dVRK surgeon-side manipulator (MTMs) were recorded and sent to the detection algorithms using an integrated Robot Operating System (ROS) communication tool [37]. The stylistic detection algorithm was fed a data window of 30 samples. To enable near-real-time detection, the incoming data from the dVRK was read at a frequency of 30Hz through ROS nodes. The detection was performed with the continuously updated data window at the same frequency.

B. Detecting Anxious Movements

The stylistic behavior performance was detected using our previously published method [26]. In order to integrate this method with the dVRK, the proposed classifier was trained using the kinematic data from the MTMs of a da Vinci skill simulator from the JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS) [38]. JIGSAWS contained the kinematic data and the corresponding video of 8 subjects performing 3 different surgical training tasks with 5 repetitions. The ground-truth in training was obtained through crowd-sourcing - a video clip in JIGSAWS was provided to the participant and the participant was required to select an adjective ("Anxious/Calm") to describe the video clip. Our classifier returns 0 if a poor performance ("Anxious") is detected and returns 1 otherwise ("Calm"). These detection algorithms were implemented in MATLAB.

C. Providing Haptic Cues to Correct Anxious Movements

To improve user's performance, haptic cues were provided based on near-real-time style classification. In our study, the *cisst* library and ROS interfaces [31] were used to read kinematic data, as well as publish haptic cues (i.e., the calculated wrench) to the MTMs. When an "Anxious" movement is detected, the haptic cue is turned on. The three types of haptic cues are described below:

- **Time-Variant Spring** (S)**:** This haptic cue was calculated by using the difference between the position of the MTMs at time t (P_t) and the position at time $t-1$ (P_{t-1}).
$$F_s = K_s(P_t - P_{t-1}) \tag{1}$$
The gain K_s was selected as 90 through pilot study where $K_s = 90N/m$ could be noticeably felt and didn't introduce instability nor violate the passivity of the system [24].
- **Time-Variant Damper** (D)**:** This haptic cue was calculated using the difference between the velocity of the

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

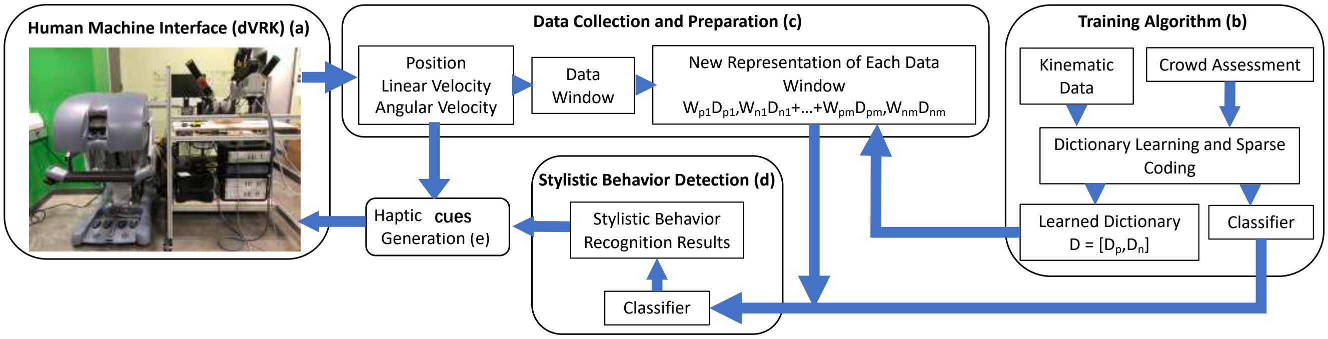


Fig. 2: System Block Diagram: (a) The dVRK and the training environment. (b) The dictionary of stylistic feature codes and a classifier generated to predict stylistic deficiencies [25], [26]. (c) Each window of kinematic measurements is represented as stylistic behaviors by projecting it on the learned dictionary. (d) The quality of the user's style is detected using a classifier. (e) Haptic cues are provided to the user if negative style (i.e., "Anxious" movement) is detected.

Fig. 3: A photo of standard peg transfer task taken by the endoscope on dVRK.

MTMs at time t (V_t) and the position at time $t-1$ (V_{t-1}).

$$F_d = B_d(V_t - V_{t-1}) \tag{2}$$

The gain B_d was chosen to be 15 in this experiment. In pilot study, $B_d = 15N.s/m$ could be perceived by the user and didn't create instabilities nor violate passivity.

- **Time-Variant Spring-Damper** (SD)**:** This haptic cue was calculated using the difference between position and velocity of the MTMs at time t (P_t, V_t) and time $t-1$ (P_{t-1}, V_{t-1})

$$F_{sd} = K_{sd}(P_t - P_{t-1}) + B_{sd}(V_t - V_{t-1}) \tag{3}$$

The gains K_{sd} and B_{sd} were selected as $K_{sd} = 90N/m$, $B_{sd} = 15N.s/m$ to stay consistent with K_s and B_d.

IV. Experimental Setup

A. Subject Recruitment

Eight subjects participated in this study. The subjects were engineering students and staff at the University of Texas at Austin. All subjects were de-identified. Four out of eight subjects had prior experiences in haptics, but none were formally trained in robotic surgery. The study protocol was approved by UT Austin IRB office (#STUDY00000278). Participants had no previously reported muscular-skeletal injuries or diseases, or neurological disorders. Seven subjects were right-hand dominant and one subject was left-hand dominant. The data was analyzed based on "Dominant hand" and "Non-dominant hand" based on self-report.

B. Surgical Training Task

A bimanual peg transfer task was used in this study (Fig 3). The peg transfer task requires subjects to control the Patient Side Manipulators (PSMs) to pick up each block from the left side of the board, complete a mid-air transfer of the block between hands, and place it onto the right side. Then, all blocks need to be moved back to the left side, with a mid-air transfer [39]. This is considered a single "round trip" for the blocks.

C. Experiment Protocol

Each subject was required to finish 6 trials in total. As illustrated in Fig 4, one subject first performed a baseline trial containing an one-round-trip peg transfer task with no haptic cue, followed by a trial containing a two-round-trip peg transfer task with a haptic cue. This procedure was repeated three times for all three types of haptic cues which was presented to each subject in a randomized order. We assigned each type of haptic cues with a number from 1 to 3 (S: 1, D: 2, SD: 3). Then we used $randperm()$ function in MATLAB to generate a uniform pseudo-random order of [1,2,3] for each subject. When the subjects were recruited, the order of the haptic cues was generated and concealed from the subjects - the experimenter did not know the order in advance as it was generated automatically for each subject so there was no possibility of selection bias. The baseline trials were designed to wash out training effects throughout the experiment.

D. User Performance Metric

To quantify performance for each type of haptic cue, we developed a user performance metric, I. For each trial, the performance, R, of the subject was evaluated by the total

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

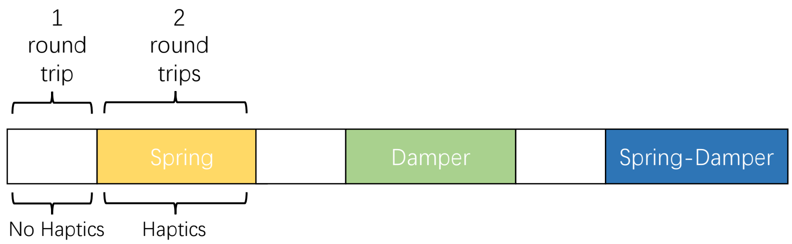


Fig. 4: An example of an experimental protocol for one subject.

sum of '1's' or good performance as being detected (i.e., good performance = Calm style), divided by the total number of detections in that trial.

$$R = \frac{num\_good\_detections}{num\_total\_detections} \tag{4}$$

The performance metric I of one type of haptic cue is then obtained by

$$I = \frac{R_{haptic}}{R_{baseline}} \tag{5}$$

The metric I describes the improvement of performance in the trial with haptic cues over its own baseline trial without haptic cues.

E. Task Performance Metrics

We also included two metrics to quantify task level performance: path length (PL) and economy of volume (EV). Both PL and EV metrics are a single number for each side during each trial.

PL describes the spatial distribution of the MTMs movements in the workspace of the task. A compact "distribution" for path length is a known characteristic of an expert [40]:

$$PL = \sum_{T_{start}}^{T_{end}} \sqrt{(P_{t+1} - P_t)^T (P_{t+1} - P_t)} \tag{6}$$

P is the 3 dimensional position of the movement. EV indicates the efficiency of occupying the space [41], and a larger value of EV indicates better performance:

$$EV = \frac{\sqrt[3]{(x_{max} - x_{min})(y_{max} - y_{min})(z_{max} - z_{min})}}{PL} \tag{7}$$

Moreover, to compare improvement over each baseline trial, we introduced two metrics: path length improvement (PL_{imp}) and economy of volume improvement (EV_{imp}).

$$PL_{imp} = \frac{PL_{haptic}}{PL_{m-baseline}}, EV_{imp} = \frac{EV_{haptic}}{EV_{m-baseline}} \tag{8}$$

As our experimental design required the subject to finish one round trip of a peg transfer task for the baseline trials but two round trips for trials with haptic cues, we needed a modified path length for the baseline trials ($PL_{m-baseline}$). This was calculated as the path length of baseline trial multiplied by two. The $EV_{m-baseline}$ was similarly modified by using the $PL_{m-baseline}$.

Both metrics were calculated by dividing the PL or EV of the trial with haptic cues by the $PL_{m-baseline}$ or $EV_{m-baseline}$ for each individuals baseline trial for a given cue. $PL_{imp} < 1$ and $EV_{imp} > 1$ indicate an improvement over baseline trial. In hindsight, requiring two baseline round trips would have simplified our post-experimental analysis; however, it was important to also the limit the overall duration of the experiment.

F. Analysis Methods

We examined the distribution properties of all the metrics mentioned above. We chose different statistical analysis methods based on the normality test on different metrics. Since we are examining the differences among different types of haptic cues, if the normality test was not rejected for a metric, the ANOVA would be used; If the normality test was rejected, then the Kruskal Wallis test was used to identify the significance (Table. I).

V. RESULTS AND DISCUSSION

To investigate which type of haptic cues can potentially correct user's "Anxious" movements based on stylistic behavior detections, we collected 48 trials (8 subjects, 6 trials each) in total. Data analysis was carried out for all trials. The results include the evaluation of user performance with different types of haptic cues based on "Calm/Anxious" movement detection, as well as a statistical analysis to identify differences between the different haptic cues.

A. Effect of Haptic Cues on Anxious Movements

The effect of each type of haptic cues on "Calm/Anxious" movements is shown in Fig 5. The mean and standard deviation of the quantity associated with good performances (I) for the three types of haptic cues are shown. This is the number of good performances detected in the trials with haptic cues normalized to the number of good performances detected in its baseline trial. The values above the horizontal line crossing at 1 show the improvement of the "Calm/Anxious" style when applying haptic cues with respect to the baseline (no haptic cues).

This plot shows that the S haptic cue has the highest value of I for both hands. This indicates that S haptic cues have a better potential to correcting "Anxious" movements than the other haptic cues tested.

A post-hoc statistical analysis (Table I) was done to determine significant differences in the three types of haptic cues. The normality test to identify a normal distribution in

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

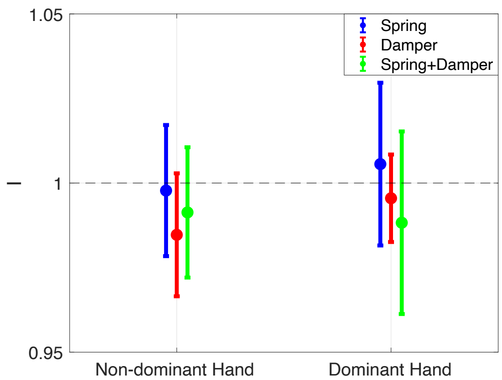


Fig. 5: Comparing the effect of three types of haptic cues on Calm/Anxious style using the user performance metric I.

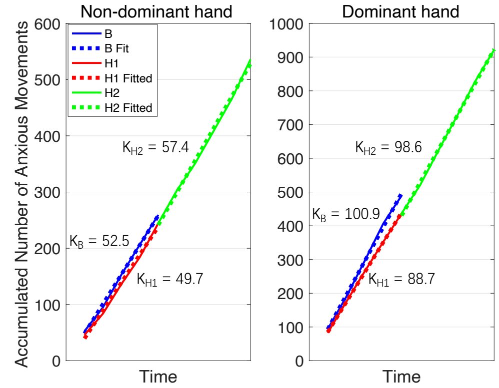


Fig. 6: Accumulated "Anxious" movement detections for all trials. The slopes (K) indicate user performance for each trial (K_B: slope of baseline trials, K_{H1}: slope in first round trip of experiment trials, K_{H2}: slope in second round trip).

the data was rejected and thus, the Kruskal Wallis test was used to identify significantly different groups.

The results from the statistical analysis indicate that for the style "Calm/Anxious", the S haptic cues showed significant difference in improving the user performance compared to the other two types of haptic cues. S haptic cues resulted in a significantly higher I than D in non-dominant hand ($p = 0.0375$) and a significantly higher I than SD in dominant hand ($p = 0.0334$).

However, not every I has a value greater than 1. One reason behind this could be the experimental design: the baseline only contained a single round trip of peg transfer task, while the following trial (with haptic cues) contained two round trips. The fatigue in the extended trial could have had a negative effect on subject's performance, therefore, resulting in a lower value of I. We implemented further analysis on accumulated negative style detections to support this hypothesis. At each time step for each trial, we counted the accumulated number of "Anxious" movements detected, therefore, indicating the growth of negative style movements throughout time. To ensure consistency, the baseline trials were normalized to have the length of 5 and the haptic trials were normalized to have the length of 10 (this is an arbitrary choice of numbers to enable illustration. The only important relationship is to normalize the baseline trials to 1/2 of the experimental trial normalization number). We then took the average across all baseline trials and the average across all trials with haptic cues. Since the growth of "Anxious" movements was found to be linear, we linearly fit the data and recorded the slopes, i.e., the speed of growth. The results in Fig 6 indicate that for the non-dominant hand the first round trip in haptic trials had the lowest slope while the second round trip had the highest slope; similar results were found in dominant hand. This greater slope indicates a faster growth of the number of "Anxious" movements, and thus a worse performance for the second round trip.

Overall, the S haptic cues had a better effect on reducing anxious movements than the other haptic cues, especially for dominant hand.

B. Effect of Haptic Cues on Task Performance

We also analyzed the effect of haptic cues on task performance using two commonly used metrics: path length, and economy of volume, as well as their improvements over baseline. The results are shown in Fig. 7.

Based on the normality test which was not rejected, we used the ANOVA to test PL_{imp} and EV_{imp}. No significant differences were found between the haptic cues (Table I).

We then analyzed PL and EV with all haptic cues. As shown in Fig 8a, S haptic cues show the lowest PL for both dominant and non-dominant hands, indicating the best task performance among all three types of haptic cues.

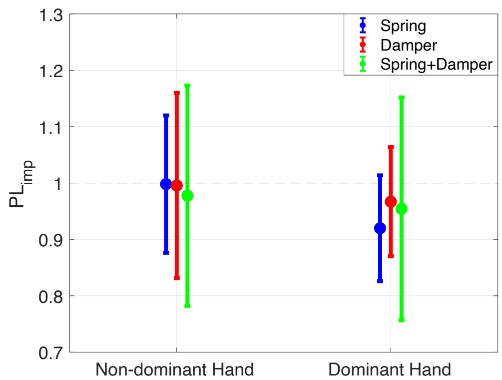
Similar results were found in the analysis of EV where S haptic cues show the highest EV values for both dominant and non-dominant hands, indicating the best task performance among all three types of haptic cues (Fig 8b).

A post-hoc statistical analysis was done. The normality test to identify a normal distribution in PL and EV was rejected, and the Kruskal Wallis test was used. The results indicate that S haptic cues showed significant differences in PL and EV for the non-dominant hand. S haptic cues resulted in a significantly lower PL than D and SD in non-dominant hand ($p = 0.0494$) and a significantly higher EV than SD in the non-dominant hand ($p = 0.0174$). Interestingly, these results seemingly contradict our initial work where SD feedback appeared to lead to best improvements of movements classified as anxious. These two experiments do represent vastly different tasks (i.e., unimanual trajectory following vs. bimanual peg transfer). Further study is needed to better understand the relationships between task difficulty and appropriate and effective feedback strategies for human-in-the-loop systems.

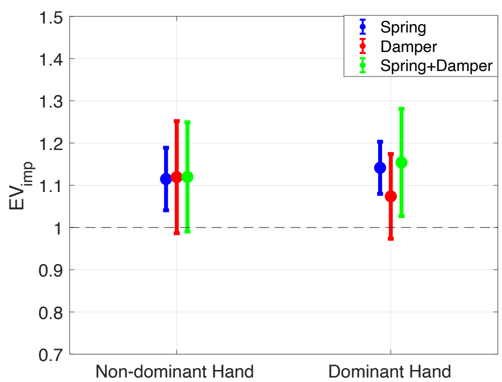
C. Subject Survey

Subjects generally reported a perception of better movements when S haptic cues were applied. They also reported an unpleasant feeling with the D haptic cue. Moreover, four out of eight subjects reported fatigue and arm soreness during

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

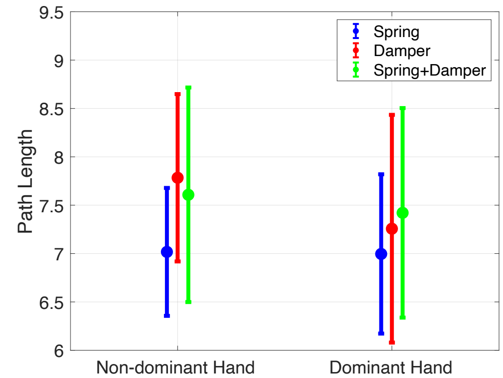


(a) Path Length Improvement over Baseline

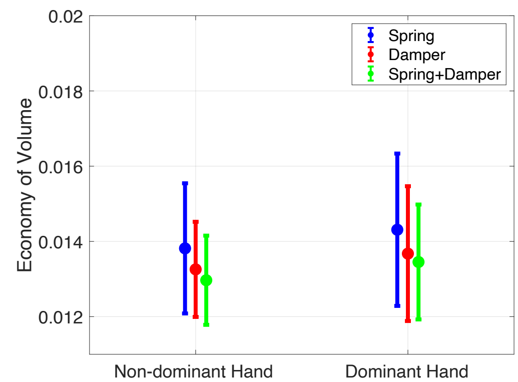


(b) Economy of Volume Improvement over Baseline

Fig. 7: Comparing the effect of three different types of haptic cues on Calm/Anxious style using improvement in task performance metrics.



(a) Path length



(b) Economy of Volume

Fig. 8: Comparing the effect of three haptic cues on Calm/Anxious style using task performance metrics.

two round trips trials and asked for extended break time between trials. This could explain why we saw differences in the growth of accumulated "Anxious" movements in the second round trip (shown in Fig. 6). This could also explain why some of the haptic cues did not result in performance improvements over baseline trials, particularly if the feedback was tiring and distracting.

Based on the analysis of I, PL and EV, the significant differences in these metrics suggested that S haptic cues showed a better performance for correcting "Anxious" movements than the other two types of haptic cues.

VI. CONCLUSIONS

In this study, we proposed a training framework which detects user's "Anxious" movements and applied haptic cues to potentially correct the movements based on our prior work in near-real-time stylistic detection. We focused on the style of "Calm/Anxious" as stressful conditions have known, serious consequences on surgical performance and patient outcomes. Correcting these "Anxious" movements could potentially improve patient safety. We conducted a human-user study to evaluate the effects of different types of haptic cues (Time-variant Spring, Damper, and Spring-Damper) on improving user's performance.

Overall, the time-variant spring haptic cues resulted in better improvement in performance over time-variant damper and time-variant spring-damper, as determined by several quantitative metrics. Four metrics showed significant differences and the common trends in our metrics suggest time-variant spring haptic cues have the potential to correct "Anxious" movement and improve performance in the non-dominant and dominant hand.

One limitation of stylistic behavior detection is the ground truth labeling before training the algorithm. The ground truth, whether a movement is "Calm/Anxious" as well as other adjective pairs, was obtained by video reviewing through crowd sourcing. The subjects in JIGSAWS did not experience external stressors when they were performing the tasks, therefore, making the "Anxious" movements detected in JIGSAWS less representative.

The findings in this paper also have practical application values. Although it is debatable whether doctors should continue the operation in this situation, these findings provide groundwork for providing preventative control strategies to reduce the unwanted effect of stress during surgical training, consequently, improving surgical training outcomes and patient safety.

This study, which is the first of its kind using a dVRK

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Statistical analysis on the effect of haptic cues on correcting Anxious movements. N/A indicates the difference is not significant; ND: Non-dominant hand, D: Dominant hand.

Metric	Handedness	Significance	$pvalue$	Methods
I	ND	S >D	0.0375	Kruskal-Wallis
	D	S >SD	0.0334	Kruskal-Wallis
PL_{imp}	ND	N/A	0.9632	ANOVA
	D	N/A	0.7818	ANOVA
EV_{imp}	ND	N/A	0.9955	ANOVA
	D	N/A	0.2486	ANOVA
PL	ND	S <D, SD	0.0494	Kruskal-Wallis
	D	N/A	0.5307	Kruskal-Wallis
EV	ND	S >SD	0.0174	Kruskal-Wallis
	D	N/A	0.0594	Kruskal-Wallis

platform, paves the way for continued research on adaptively improving user performance using task-independent stylistic behavior detection. In future studies, we will expose subjects to commonly experienced stressors and explore the importance of haptic feedback gains on user-reported experience and performance, as well as start to evaluate different types of haptic cues for other types of surgical movement styles. We will also study more challenging surgical training tasks such as suturing and pattern cutting.

VII. ACKNOWLEDGMENT

The authors would like to thank Anton Deguet and the members in the dVRK community.

REFERENCES

[1] B. A. Boone, M. Zenati, M. E. Hogg, J. Steve, A. J. Moser, D. L. Bartlett, H. J. Zeh, and A. H. Zureikat, "Assessment of quality outcomes for robotic pancreaticoduodenectomy: identification of the learning curve," JAMA surgery, vol. 150, no. 5, pp. 416–422, 2015.

[2] V. Tam, H. J. Zeh, and M. E. Hogg, "Incorporating metrics of surgical proficiency into credentialing and privileging pathways," JAMA surgery, vol. 152, no. 5, pp. 494–495, 2017.

[3] L. M. Knab, A. H. Zureikat, H. J. Zeh, and M. E. Hogg, "Towards standardized robotic surgery in gastrointestinal oncology," Langenbeck's archives of surgery, vol. 402, no. 7, pp. 1003–1014, 2017.

[4] B. Johnson, I. Sorokin, N. Singla, C. Roehrborn, and J. C. Gahan, "Determining the learning curve for robot-assisted simple prostatectomy in surgeons familiar with robotic surgery," Journal of endourology, vol. 32, no. 9, pp. 865–870, 2018.

[5] I. Jacobs, A. Lay, and J. Gahan, "Mp23-17 lack of an experienced bedside assistant may adversely affect outcomes during robotic prostatectomy," The Journal of Urology, 2015.

[6] M. S. Morgan, N. A. Shakir, M. Garcia-Gil, A. Ozayar, J. C. Gahan, J. I. Friedlander, C. G. Roehrborn, and J. A. Cadeddu, "Single-versus dual-console robot-assisted radical prostatectomy: impact on intraoperative and postoperative outcomes in a teaching institution," World journal of urology, vol. 33, no. 6, pp. 781–786, 2015.

[7] R. L. Steinberg, N. Passoni, A. Garbens, B. A. Johnson, and J. C. Gahan, "Initial experience with extraperitoneal robotic-assisted simple prostatectomy using the da vinci sp surgical system," Journal of Robotic Surgery, pp. 1–7, 2019.

[8] C. Reiley, H. Lin, D. Yuh, and G. Hager, "Review of methods for objective surgical skill evaluation," Surgical Endoscopy, vol. 25, pp. 356–366, July 2011.

[9] C. Chen, L. White, T. Kowalewski, R. Aggarwal, C. Lintott, B. Comstock, K. Kuksenok, C. Aragon, D. Holst, and T. Lendvay, "Crowd-sourced assessment of technical skills: A novel method to evaluate surgical performance," Journal of Surgical Research, vol. 187, no. 1, pp. 65–71, mar 2014.

[10] A. Malpani, S. S. Vedula, C. C. G. Chen, and G. D. Hager, "A study of crowdsourced segment-level surgical skill assessment using pairwise rankings," International Journal of Computer Assisted Radiology and Surgery, vol. 10, no. 9, pp. 1435–1447, sep 2015.

[11] M. Ershad, Z. Koesters, R. Rege, and A. Majewicz, "Meaningful assessment of surgical expertise: Semantic labeling with data and crowds," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9900 LNCS. Springer Verlag, oct 2016, pp. 508–515.

[12] H. Rafii-Tari, C. J. Payne, C. Bicknell, K.-W. Kwok, N. J. Cheshire, C. Riga, and G.-Z. Yang, "Objective assessment of endovascular navigation skills with force sensing," Annals of Biomedical Engineering, vol. 45, no. 5, pp. 1315–1327, 2017.

[13] M. J. Fard, S. Ameri, R. Darin Ellis, R. B. Chinnam, A. K. Pandya, and M. D. Klein, "Automated robot-assisted surgical skill evaluation: Predictive analytics approach," The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 14, no. 1, p. e1850, 2018.

[14] N. Ahmidi, L. Tao, S. Sefati, Y. Gao, C. Lea, B. B. Haro, L. Zappella, S. Khudanpur, R. Vidal, and G. D. Hager, "A dataset and benchmarks for segmentation and recognition of gestures in robotic surgery," IEEE Transactions on Biomedical Engineering, vol. 64, no. 9, pp. 2025–2041, 2017.

[15] Y. A. Oquendo, E. W. Riddle, D. Hiller, T. A. Blinman, and K. J. Kuchenbecker, "Automatically rating trainee skill at a pediatric laparoscopic suturing task," Surgical endoscopy, vol. 32, no. 4, pp. 1840–1857, 2018.

[16] Z. Wang and A. M. Fey, "Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery," International journal of computer assisted radiology and surgery, vol. 13, no. 12, pp. 1959–1970, 2018.

[17] S. Kumar, N. Ahmidi, G. Hager, P. Singhal, MD, J. Corso, and V. Krovi, "Surgical performance assessment," Mechanical Engineering, vol. 137, no. 09, pp. S7–S10, Sept. 2015.

[18] C. Sewell, D. Morris, N. H., Blevins, S. Dutta, S. Agrawal, F. Barbagli, and K. Salisbury, "Providing metrics and performance feedback in a surgical simulator," Computer Aided Surgery, vol. 13, no. 2, pp. 63–81, 2008.

[19] L. Panait, E. Akkary, L. Bell, Robert, E. Roberts, Kurt, J. Dudrick, Stanley, and J. Duffy, Andrew, "The role of haptic feedback in laparoscopic simulation training," Journal of Surgical Research, vol. 156, no. 2, pp. 312–316, 2009.

[20] Y. Lin, X. Wang, F. Wu, X. Chen, C. Wang, and G. Shen, "Development and validation of a surgical training simulator with haptic feedback for learning bone-sawing skill," Journal of Biomedical Informatics, vol. 48, pp. 122–129, 2014.

[21] K. Kim, Hyun, W. Rattner, David, and A. Srinivasan, Mandayam, "Virtual-reality-based laparoscopic surgical training: The role of simulation fidelity in haptic feedback," Computer Aided Surgery, vol. 9, no. 5, pp. 227–234, 2004.

[22] D. Feygin, M. Keehner, and F. Tendick, "Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill," in Proceedings - 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS 2002. Institute of Electrical and Electronics Engineers Inc., 2002, pp. 40–47.

[23] W. H. Jantscher, S. Pandey, P. Agarwal, S. H. Richardson, B. R. Lin, M. D. Byrne, and M. K. O'Malley, "Toward improved surgical training: Delivering smoothness feedback using haptic cues," in IEEE Haptics Symposium, HAPTICS, vol. 2018-March. IEEE Computer Society, may 2018, pp. 241–246.

[24] N. Enayati, A. M. Okamura, A. Mariani, E. Pellegrini, M. M. Coad, G. Ferrigno, and E. De Momi, "Robotic assistance-as-needed for enhanced visuomotor learning in surgical robotics training: An experimental study," in 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 6631–6636.

[25] M. Ershad, R. Rege, and M. Fey, A., "Meaningful assessment of robotic surgical style using the wisdom of crowds," International Journal of Computer Assisted Radiology and Surgery, vol. 13, pp. 1037–1048, 2018.

[26] M. Ershad, R. Rege, and A. Majewicz Fey, "Automatic and near real-time stylistic behavior assessment in robotic surgery," International Journal of Computer Assisted Radiology and Surgery, vol. 14, no. 4, pp. 635–643, apr 2019.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.

[27] S. Arora, N. Sevdalis, D. Nestel, T. Tierney, M. Woloshynowych, and R. Kneebone, "Managing intraoperative stress: what do surgeons want from a crisis training program?" American Journal of Surgery, vol. 197, no. 4, pp. 537–543, apr 2009.

[28] K. Moorthy, Y. Munz, A. Dosis, S. Bann, and A. Darzi, "The effect of stress-inducing conditions on the performance of a laparoscopic task," Surgical Endoscopy and Other Interventional Techniques, vol. 17, no. 9, pp. 1481–1484, sep 2003.

[29] K. H. Goodell, C. G. Cao, and S. D. Schwaitzberg, "Effects of cognitive distraction on performance of laparoscopic surgical tasks," Journal of Laparoendoscopic and Advanced Surgical Techniques, vol. 16, no. 2, pp. 94–98, apr 2006.

[30] S. Arora, N. Sevdalis, R. Aggarwal, P. Sirimanna, A. Darzi, and R. Kneebone, "Stress impairs psychomotor performance in novice laparoscopic surgeons," Surgical Endoscopy, vol. 24, no. 10, pp. 2588–2593, 2010.

[31] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da vinci® surgical system," in 2014 IEEE International Conference on Robotics and Automation (ICRA), May 2014, pp. 6434–6439.

[32] A. Kapoor, A. Deguet, and P. Kazanzides, "Software components and frameworks for medical robot control," in Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006. IEEE, 2006, pp. 3813–3818.

[33] A. Deguet, R. Kumar, R. Taylor, and P. Kazanzides, "The cisst libraries for computer assisted intervention systems," Insight, pp. 1–8, 2008.

[34] M. Y. Jung, A. Deguet, and P. Kazanzides, "A component-based architecture for flexible integration of robotic systems," in IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings, 2010, pp. 6107–6112.

[35] M. Ershad, R. Rege, and A. M. Fey, "Adaptive surgical robotic training using real-time stylistic behavior feedback through haptic cues," IEEE Transactions on Medical Robotics and Bionics, vol. 3, no. 4, pp. 959–969, 2021.

[36] Z. Wang and A. Majewicz Fey, "Human-centric predictive model of task difficulty for human-in-the-loop control tasks," PloS one, vol. 13, no. 4, p. e0195053, 2018.

[37] M. QUIGLEY, "Ros: An open-source robot operating system," in International Conference on Robotics and Automation, Open-Source Software Workshop, 2009, 2009.

[38] Y. Gao, S. S. Vedula, C. E. Reiley, N. Ahmidi, B. Varadarajan, H. C. Lin, L. Tao, L. Zappella, B. Béjar, D. D. Yuh, et al., "Jhu-isi gesture and skill assessment working set (jigsaws): A surgical activity dataset for human motion modeling," in MICCAI workshop: M2cai, vol. 3, 2014, p. 3.

[39] E. Ritter, T. Kindelan, C. Michael, E. Pimentel, and M. Bowyer, "Concurrent validity of augmented reality metrics applied to the fundamentals of laparoscopic surgery (fls)," Surgical endoscopy, vol. 21, no. 8, pp. 1441–1445, 2007.

[40] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, and S. Dawson, "Metrics for Laparoscopic Skills Trainers: The Weakest Link!" in Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002, T. Dohi and R. Kikinis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 35–43.

[41] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, Á. Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo, and E. J. Gómez, "EVA: Laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment," Surgical Endoscopy, vol. 27, no. 3, pp. 1029–1039, oct 2013.

Authorized licensed use limited to: University of Texas at Austin. Downloaded on November 05,2022 at 19:59:05 UTC from IEEE Xplore. Restrictions apply.