On the Development of Objective Metrics for Surgical Skills Evaluation Based on Tool Motion

Sean Estrada, Marcia K. O'Malley
Department of Mechanical Engineering
Rice University
Houston, TX, USA
sje1, omalleym @ rice.edu

Cassidy Duran, Daryl Schulz, Jean Bismuth
Debakey Heart & Vascular Center
Department of Cardiovascular Surgery
Methodist Houston
Houston, TX USA
caduran, dgschulz, jbismuth @ houstonmethodist.org

Abstract—**A key competency requirement in endovascular surgery is to optimally visualize and utilize pre-shaped catheters to navigate complex vascular anatomy, yet current performance assessment techniques are limited to grading scales based solely on observation. Since most endovascular procedures involve performing fine motor control tasks that require complex, dexterous movements, this paper explores the potential for a standardized, objective, and quantitative means of measuring technical competence based on analysis of the kinematics of endovascular tool tip motions. To accomplish this goal, we recorded catheter tip movement from twenty subjects performing fundamental endovascular tasks in an inanimate model and in a simulation environment with a virtual representation of the same inanimate model. Several motion-based performance measures that have been shown to reliably assess skill in other domains were computed and tested for correlation with data that were obtained from the global rating scale assessment tool. The metrics that quantified movement quality by assessing movement smoothness produced reliable correlations with the observation-based assessment metrics. These objective and quantitative metrics that capture movement quality could be incorporated into future training protocols to provide detailed feedback on trainee performance.**

Keywords—virtual reality; motion capture; skill assessment

I. INTRODUCTION

Endovascular surgery is a form of minimally invasive surgery (MIS) that was designed to access many regions of the body via major and minor blood vessels. Basic techniques involve the introduction of a catheter percutaneously into a large blood vessel, often the femoral artery, to perform an array of coronary, carotid and cerebral angiographic procedures. The catheter is often used to insert therapeutic devices, such as balloons, stents and coils to assist in treating diseases, such as atherosclerosis, vascular trauma or aneurysms. To help deliver the catheter into the target blood vessel, a sheath and guidewire act as a guide and support. To visualize the movements of these instruments, endovascular surgeons typically use live x-ray (fluoroscopic) images. The imaging system can be repositioned to provide different two-dimensional projection views of the vascular areas being treated during endovascular surgical procedures.

While endovascular procedures offer an advantage over more traditional open surgical treatment techniques by allowing smaller incisions and a reduced chance of infection, there have been numerous reports that the complication rates for some procedures surpass those present after management without surgery [1]. One approach to decreasing complication rates is to develop objective assessment methods, thereby enabling surgical training programs to determine the readiness of a given trainee to perform a particular procedure. The need for objective quantitative assessment tools has indeed been a topic of considerable importance and interest [2][3]. To date, however, surgical skill assessment and clinical evaluation have predominantly remained subjective [4].

Given the complexity of MIS procedures and the limitations of subjective evaluations, researchers have explored the potential to evaluate surgical tool motions as a means to infer skill level. Results from previous studies have demonstrated an ability to assess the skill level of surgeons using quantitative measures derived from motion analysis of surgical tools during laparoscopic procedures on virtual reality simulators [5][6][7], box trainers [8][9] and models [10].

This paper explores the applicability of motion-based measures of performance, such as those previously demonstrated for laparoscopic procedures, to endovascular procedures. As illustrated in Fig. 1, catheter tip motion data is collected during the performance of fundamental endovascular tasks on two different platforms (manual catheterization of a physical inanimate model and the same model on a simulator). Simultaneously, assessments using standardized checklists are collected. Correlations between these measures are computed in order to identify appropriate motion-based measures that can be used to assess skill for endovascular surgeons.
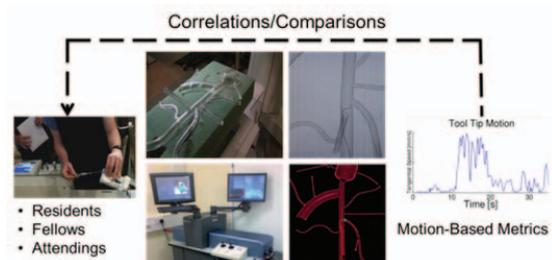
Fig. 1. Approach for identifying motion-based performance metrics of endovascular skills. From left, subject performing endovascular task. Center, experimental platforms (and fluoroscopic images). Right, representative speed trajectory of catheter tip during task execution. Correlations are used to relate the measured motion data to observation-based assessments.

Endovascular surgical training relies heavily on the use of physical models and simulation for training. Often these platforms are used for the rehearsal of specific procedures, since patient-specific data and images can be loaded onto a simulator or used to create detailed anatomical models for the purposes of planning and procedural practice before the surgeon enters the operating room. There is interest in extending the use of inanimate physical models for task specific training rather than rehearsal of full surgical procedures, such as those models that have been developed for laparoscopic training programs. In the laparoscopic domain, the Fundamentals of Laparoscopic Surgery (FLS) program is among the most robust simulation program available for assessing skill level. The program contains both cognitive and skills components, has been correlated with clinical performance, and offers a certification process through validated examinations [8]. Endovascular surgeons are interested in having a comparable program for the training of skills, and to this end, an anatomically inspired inanimate model has been developed to enable task-specific training specifically for endovascular surgery. Further, virtual reality simulation technology enables the replication of this inanimate model in a computer simulation environment.

Given the success of objective assessment of laparoscopic surgical skill using motion-based performance metrics, this paper explores the applicability of such measures for endovascular tasks. The first category of motion-based metrics is those that directly characterize the instrument kinematics (e.g. path length, acceleration, or input frequency). The second category of metrics is those inspired from principles of human motor control. It is hypothesized that these measures, such as smoothness of movement, which are derived from the kinematic data of tool tip motion, will give insight into the *quality* of movement [11]. Movement smoothness is widely regarded as a hallmark of skilled, coordinated movement [12], and has been used as a measure of motor performance in basic motor control tasks [13] and rehabilitation applications [14][15]. However, it has yet to be demonstrated if such metrics are also relevant to more complex tasks such as those required in surgical applications.

II. METHODS

In this study, motion analysis and correlations were used to identify quantitative metrics to differentiate the skill level of residents, fellows, and attending surgeons while they performed a set of fundamental endovascular surgical tasks using manual catheterization with a physical model and in a simulation environment (see Fig. 2).
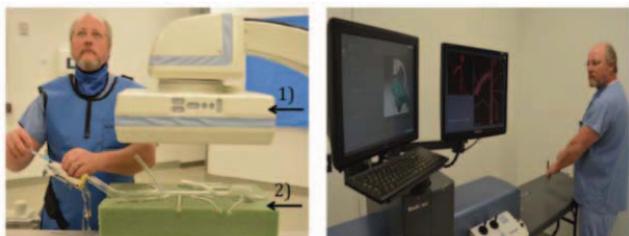
Fig. 2. Experimental set-ups of subject performing catheterization. At left, manual experiment with imaging system to generate fluoroscopic images (1) and inanimate FEVS model (2). At right, Simbionix Angio Mentor simulator.

A. Experimental platforms

The Fundamentals of Endovascular Skills (FEVS) model is non-anatomical (though, anatomically inspired) and was designed to assess basic endovascular skills, as opposed to assessing specific procedures. The model was designed so that fundamental skills of endovascular surgery could be assessed, including basic catheter and guidewire skills, selective catheterization of all vascular beds, and arteriography of all vascular beds [16]. NDI's Window Field Generator with 5-DOF electromagnetic sensors was used to track the position and orientation of the catheter tip when navigating inside of the physical model.

The Simbionix ANGIO Mentor Ultimate is a computer-assisted medical training simulator for endovascular procedures, performed under real-time fluoroscopy on realistic 3D anatomies. The system features realistic 3D anatomies and realistic simulation of tools with real nested instruments and tactile sensation, and can be loaded with custom models, in this case a computer model of the FEVS inanimate model. Video capture software was used to record the performance of all tasks on the simulator and then image processing techniques were used to extract the (x,y) position of the catheter tip, in pixel space.

B. Subjects

Twenty subjects (18 male, 2 female, average age of experts 49, intermediates 39, and novices 32) participated in this study performed at The Methodist Hospital in Houston, TX. Seven subjects were either cardiology or vascular residents, six were cardiovascular fellows, six were attending physicians and one was an expert in a similar domain. The subjects ranged in experience from residents who were less than a year removed from medical school to attending surgeons with more than twenty years of experience in cardiovascular surgery. Novices (10) were those having performed less than fifty cases. Intermediates (6) had performed between 50 and 500 cases. Experts (4) were defined as participants with more than 500 previous cases.

C. Tasks

The subjects performed four fundamental endovascular tasks of comparable difficulty on each platform repeated over three sessions: catheterization into the anterior, first left lateral, posterior and second left lateral branches in the FEVS model (shown in Fig. 3) or in simulation. The first and second left lateral branches were cannulated while the C-arm was at 0º (or, Anterior/Posterior), the anterior branch was cannulated at 75º Left Anterior Oblique (LAO), and the posterior branch was cannulated at A/P until the catheter was sufficiently inside of the first left lateral branch, where the rest of the task was performed at 75º LAO. A sheath, catheter and guidewire were used during each task on each platform to navigate to the branch of interest.

The goal was to either a) successfully cannulate the catheter tip to an identified success point in the FEVS model (between 2-4 cm inside of the branch of interest) within five minutes or b) successfully cannulate the tips of the guidewire, catheter and sheath to color-coded targets inside of the branch

of interest within five minutes using the simulator. While accomplishing all tasks during the experiment, the guidewire was required to be cannulated into the branch of interest prior to the catheter so that the catheter would go into a branch over the guidewire. If the subject failed to do so, the error was deemed "critical" because failure to move a catheter over a guidewire in real-world endovascular procedures could cause severe complications. When such failures occurred, the corresponding assessment data were discarded.
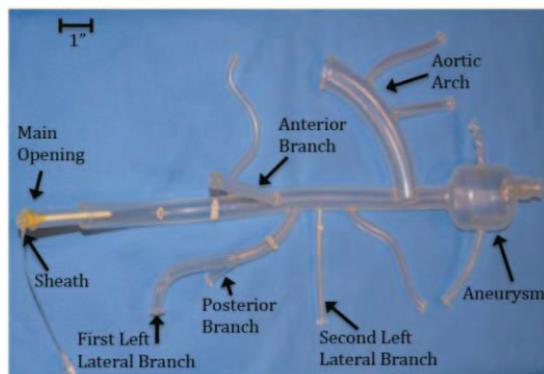
Fig. 3. Inanimate Fundamentals of Endovascular Skills (FEVS) model
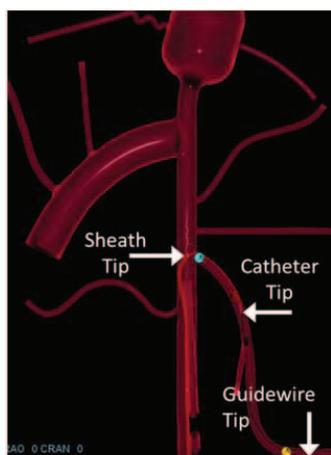
Fig. 4. Simulator visualization of the inanimate model and the sheath, catheter, and guideware tips in view.

D. Performance assessment

Many assessment techniques exist that attempt to evaluate the cognitive capabilities, judgment and decision making, and manual dexterity of surgeons. Structured human grading is an extension of the traditional observational approach where a senior surgeon observes a student and provides verbal feedback. In multiple surgical fields, structured grading techniques often attempt to standardize evaluation through rated checklists that assess performance of fundamental tasks on inanimate bench top models [17]. The Global Rating Scale for Endovascular Performance (GRSEP) is a structured grading tool that is used by a senior clinician to assess each subject based on their performance in endovascular skills. This validated checklist is the most standardized evaluation tool that currently exists for endovascular surgical techniques.

The global rating scale is designed so that higher scores represent greater competence and capability to perform endovascular tasks.

In this study, one GRSEP was completed for each session (all subjects completed three sessions; therefore, three GRESPs were conducted for each subject) as they completed the four endovascular tasks on each platform. The measurable scales of the grading tool were: efficiency, ability to manipulate the surgical tools (catheter and guidewire), use of the device, image quality, image safety, autonomy, and level of difficulty. Each scale was measured by scores 1 to 5, with descriptors associated with each score. Data from the GRSEP were compared to the quantitative data obtained from the motion capture methods described previously.

Score counts from all categories except for Level of Difficulty showed that novices were rated lower compared to more experienced interventionalists. The results helped to both legitimize the global rating scale as a valid assessment tool and verify that the ability groups were chosen appropriately. After comparing these results with the tasks being performed during the experiment, it was decided that efficiency and wire/catheter manipulation scores (see Fig. 5) were the portions of the global rating tool that were most applicable to identify trends with the computed motion metrics. A combined score of efficiency plus wire/catheter manipulation (score from 0 - 10) was computed for tasks performed using manual catheterization on the model and simulator.
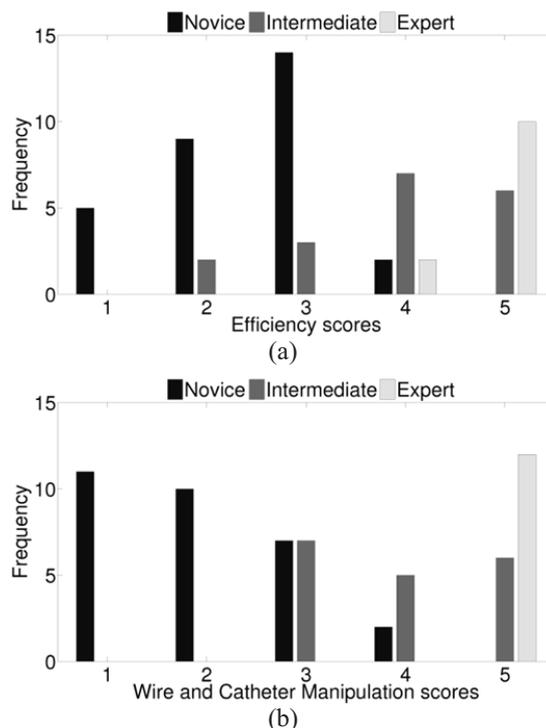
Fig. 5. Global rating scale results for the (a) efficiency and (b) catheter and wire manipulation scales. The trends shows that less experienced subjects scored lower, while more experienced subjects were consistently rated higher.

E. Motion analysis

Observation of fluoroscopy during live endovascular cases suggests that the motion of the catheter tip is a strong indicator of the skill and successful forward progress of the interventionalist. The kinematic movement of the catheter tip was recorded using an electromagnetic (EM) tracking device during manual intervention on the physical model and using video processing techniques for tasks performed on the endovascular simulator.

Three post-processing steps were conducted on data obtained from the electromagnetic sensors and processed images. Filtering techniques were used to remove noise, outliers were removed, and critical failures were defined and excluded from further analysis. Savitzky-Golay filtering was applied to the motion data because of its ability to eliminate most of the noise while preserving the shape qualities of important peaks.

The metrics that were tested for their ability to differentiate skill level among endovascular surgeons were divided into two categories: metrics derived from measured data (path length, average acceleration and input frequency), and motor control inspired metrics (non-dimensional jerk, movement arrest period ratio, total number of speed peaks, submovement extraction measures (number of submovements, duration of a submovement, inter-peak interval, overlap), and spectral arc length). The submovement metrics can be computed with either support bounded lognormal (LGNB) or minimum jerk (MJerk) profile curves. All of these metrics have proven useful for quantifying movement for different motor control tasks performed by both healthy subjects and people recovering from various movement-inhibiting injuries and diseases [7] [9][12][13][15][18][19][20][21][22]. In terms of measuring smoothness of movement, the metrics that typically produce the most reliable results are dimensionless, consistent, sensitive to changes in movement, computationally inexpensive, and use most of the available data. Most of the metrics outlined in this section fit these criteria. All of these metrics were computed for all data that were collected from each of the two experimental platforms and tested for reliable differences among subjects of different levels of experience accomplishing endovascular tasks.

After computing the motion-based metrics, both within-subjects and between-subjects outliers were identified using the 3IQR method, which defines outliers as either less than or greater than three times the factor's interquartile range. Using this method, 30 within-subject outliers were found. All within-subjects outliers were replaced with the individual subject's mean score for that particular metric. No between-subjects outliers were identified. Further, any trials containing critical failures were discarded or replaced with the overall subject mean. If a subject had no more than 33% of the trials as critical failures, the data for tasks with critical failures were replaced with the subject's mean metric score; otherwise, the subject was not included in the analysis. All data from five novice (non-competent) subjects were discarded when computing the correlation coefficients using the motion data from manual catheterization on both the model and simulator because of critical failures. In all, 52 critical failures occurred when cannulating the FEVS model and 37 critical failures occurred when performing catheterization on the simulator.

III. RESULTS

Data were obtained both from the movement of the catheter tip and the global rating scale while subjects executed each of the experimental tasks manually, on the physical model and simulator. After obtaining the combined global rating score per session for each subject, a Pearson's correlation coefficient was computed to assess the relationship between each measured value from each metric to the global rating scores.

Table 1 reports the correlation coefficients for each of these metrics for manual catheterization on the physical model and simulator. The values in the table were obtained by averaging the data for each metric across all tasks and all sessions, since the correlation coefficients across both of those within-subjects factors were not reliably different. An r-value of 0.8 or greater was considered a strong correlation, while a value of r between 0.5 and 0.79 was considered a moderate correlation.

Table 1. Correlation coefficients and p-values for manual catheterization of physical model and simulator. Candidate metrics are indicated in boldface.

Metric	Model (r)	Model (p)	Sim (r)	Sim (p)
Average Acceleration (mm/s^2)	.26	.347	.07	.813
Average Frequency (Hz)	-.09	.744	.29	.289
Completion Time (s)	-.24	.387	-.37	.181
Path Length (mm)	.10	.729	.24	.381
Non-Dimensional Jerk	**.63**	**.012**	**.55**	**.032**
Movement Arrest Period Ratio	-.11	.705	.21	.448
No. Submovements (LGNB)	**.80**	**.001**	**.71**	**.003**
No. Submovements (MJerk)	.55	.032	.48	.069
Avg Submov Duration (LGNB) (s)	.27	.336	**.77**	**.001**
Avg Submov Duration (MJerk) (s)	**.79**	**.001**	**.85**	**.001**
Inter-Peak Interval (LGNB) (s)	.09	.764	-.11	.706
Inter-Peak Interval (Min Jerk) (s)	.10	.736	-.02	.945
Overlap (LGNB) (s)	.35	.196	.19	.488
Overlap (Min Jerk) (s)	.04	.887	-.11	.708
Spectral Arc Length	**.77**	**.001**	**.84**	**.001**
Total Peaks	.37	.170	.37	.174

Motion-based metrics were declared *candidate metrics* if they exhibited more than 33% reliable correlations among the twenty-four measured value combinations (e.g., one of two platforms, one of four tasks, and one of three sessions). From the computed results, five candidate metrics were identified:

- Non-dimensional jerk
- Spectral arc length
- Number of submovements (LGNB profiles)
- Average submovement duration (LGNB profiles)
- Average submovement duration (minimum jerk profiles)

All correlation coefficients for every candidate metric except for the average submovement duration (LGNB curves) metric while manually cannulating the model were reliable. Most metrics showed strong correlation with the combined global rating scale score.

IV. Discussion

The analysis of manual endovascular skills performed on both the physical model and simulator showed that motion-based metrics that quantified smoothness and quality of movement were more strongly correlated to the structured grading assessment than those metrics that simply quantified the catheter tool tip kinematics. Specifically, the candidate metrics of spectral arc length, non-dimensional jerk, number of submovements (using LGNB curves) and average submovement duration (using both LGNB and minimum jerk profile curves) metrics exhibited reliable correlations with the GRSEP scores for manual catheterization.

A limitation of our methodology is that to identify candidate metrics, we promote those quantitative measures that strongly correlate with the GRSEP assessment, despite our prior criticism of such subjective scales. However, instruments like the GSREP which have been validated are shown to increase standardization of clinical skills assessment. Therefore, the GRESP represents the best available benchmark in the endovascular domain.

The non-dimensional jerk metric that was used in this study is completely independent of movement duration and amplitude. Multiple researchers have used jerk-based measures that are independent of movement duration and amplitude to quantify smoothness based on motion data [23][24][25]. Indeed, the non-dimensionality of the jerk metric in this study likely contributed to the strong correlations between this metric and the GRSEP scores.

The spectral arc length metric is unique in that it involves examining the frequency spectrum instead of quantifying task performance from data in the time domain. Since there was not a clear trend of task completion time based on ability group, examining the experimental data in the frequency domain proved to be an effective method to quantify the complexity of the speed trajectory. Additionally, the metric showed multiple strong correlations with the experimental data because the metric is sensitive to alterations in motor behavior, robust to measurement noise, non-dimensional and is ensured to quantify smoothness in a consistent manner [22].

Submovement extraction algorithms are designed to identify the number of discrete units of movement and key characteristics of those movements for some form of motion data. The scattershot algorithm [26] was used to find the number of submovements and associated submovement characteristics (submovement duration, inter-peak interval, overlap) from the experimental data, using both support bounded lognormal (LGNB) and minimum jerk profile curves. After identifying an appropriate error threshold for the stopping criterion, the number of submovements and average submovement duration metrics showed strong correlations with the GRSEP assessment. The mean inter-peak interval and mean overlap scores did not produce strong correlations to the global rating data presumably because, while the algorithm showed differences in number of submovements and average duration from the tangential speed profiles, the submovement trajectories were similar enough in shape to mask inter-peak interval and overlap differences between ability groups.

The metrics that were weakly correlated with the GRSEP scores were primarily those that directly assessed the kinematics of the catheter tip movement, rather than the smoothness-based metrics that sought to capture movement quality. Other weakly correlated metrics included some methods for computing smoothness which required significant portions of data to be discarded or which were not normalized.

Average acceleration did not show strong correlations because of potential issues of dimensionality of the data after differentiating and smoothing the data. Additionally, average frequency did not appear to be an effective metric at discerning between ability groups because most subjects' power spectrums showed that they moved in identical frequency ranges when accomplishing most of the tasks in the experiment. The computation of the path length metric depends on the duration of a movement, and due to the large range of task completion times within each of the ability groups, was not a suitable metric for assessing skill.

Task completion time likely did not show strong correlations because more experienced interventionalists did not consistently complete the tasks faster than the less experienced subjects. These results were contrary to results from studies assessing the performance of laparoscopic surgery [5][10][27], where task completion time was a useful metric for determining skill level. These studies involved subjects performing specific, complex tasks that were designed to take longer for less experienced surgeons.

Counting speed peaks has two intrinsic weaknesses. First, the sum of two or more submovements, each with single-peaked speed profiles, may generate spurious peaks in the composite speed profile [26]. Second, this measure is completely insensitive to the presence or extent of periods of arrest [23]. The Movement Arrest Period Ratio (MAPR) [20] addresses the latter weakness, but is insensitive to fluctuations in speed that do not result in a detectable period of arrest. In addition to these issues, the speed peak counting and MAPR metrics were also not useful as a dependable smoothness metric for the data in this study because the calculation of both metrics essentially discarded much of the available data.

V. Conclusions

Previous studies have shown that caseload and apprenticeship-model-based assessment methods might not be adequate to rate the technical competence for endovascular tasks that require a large amount of hand-eye coordination and manual dexterity. Therefore, the main objective of this paper was to develop objective motion-based measures for assessing skill level among vascular residents, fellows, and attending surgeons performing endovascular tasks. To achieve this goal, catheter tip motion data was collected during the performance of fundamental endovascular tasks on two different platforms (manual catheterization of a physical inanimate model and of the same model on a virtual reality simulator). Electromagnetic sensors and image processing techniques were used to extract the motion of the catheter tip during task execution. Simultaneously, assessments using standardized checklists were conducted. Correlations between these

measures were computed in order to identify appropriate motion-based measures that could be used to differentiate skill level for the endovascular surgeons. From data collected over three sessions for twenty participants, the number of submovements, average submovement duration, spectral arc length and non-dimensional jerk metric results showed significant differences between surgeons of different skill level when performing the four fundamental endovascular tasks. These findings demonstrated that motion-based metrics, particularly those that are based on the principles of motor control and capture movement quality characteristics such as smoothness, were indeed strongly correlated to structured grading assessments of skill. This study has provided insight about skill evaluation based on dexterity as well as motion economy while performing endovascular surgical procedures. Motion analysis could be a key component in future training programs to assessing performance for endovascular tasks.

REFERENCES

[1] M. I. Chimowitz, M. J. Lynn, C. P. Derdeyn, T. N. Turan, D. Fiorella, B. F. Lane, and H. J. Cloft, "Stenting versus aggressive medical therapy for intracranial arterial stenosis," New England Journal of Medicine, vol. 365, no. 11, pp. 993-1003, 2011.

[2] H. C. Lin, I. Shafran, D. Yuh, and G. D. Hager, "Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions," Computer Aided Surgery, vol. 11, no. 5, pp. 220-230, 2006.

[3] G. J. M. Van Hove, P. D. andTuijthof, E. G. G. Verdaasdonk, L. P. S. Stassen, and J. Dankelman, "Objective assessment of technical surgical skills," British Journal of Surgery, vol. 97, no. 7, pp. 972-987, 2010.

[4] T. T. Tsue, J. W. Dugan, and B. Burkey, "Assessment of surgical competency," Otolaryngologic clinics of North America, vol. 40, no. 6, pp. 1237-1259, 2007.

[5] R. V. O'Toole, R. R. Playter, T. M. Krummel, W. C. Blank, N. H. Cornelius, W. R. Roberts, W. J. Bell, and M. Raibert, "Measuring and developing suturing technique with a virtual reality surgical simulator," Journal of the American College of Surgeons, vol.189, no.1, pp.114-127, 1999.

[6] I. Balasundaram, R. Aggarwal, and L. A. Darzi, "Development of a training curriculum for microsurgery," British journal of oral and maxillofacial surgery, vol. 48, no. 8, pp. 598-606, 2010.

[7] 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. Springer, 2002, pp.35-43.

[8] A. O. Castellvi, L. A. Hollett, A. Minhajuddin, D. C. Hogg, S. T. Tesfay, and D. J. Scott, "Maintaining proficiency after fundamentals of laparoscopic surgery training: A 1-year analysis of skill retention for surgery residents," Surgery, vol. 146, no. 2, pp. 387-393, 2009.

[9] E. F. Hofstad, C. Vapenstad, M. K. Chmarra, T. Lang, E. Kuhry, and R. Marvik, "A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance," Surgical endoscopy, vol. 27, no. 3, pp. 854-863, 2013.

[10] S. G. T. Smith, J. Torkington, T. J. Brown, N. J. Tander, and A. Darzi, "Motion analysis: A tool for assessing laparoscopic dexterity in the performance of a laboratory-based laparoscopic cholecystectomy," Surgical Endoscopy, vol.16, no.4, pp. 640-645, 2002.

[11] M. K. Chmara, "Trendo tracking system. motion analysis in minimally invasive surgery," Ph.D. dissertation, Proefschriftmaken.nl, The Netherlands, 2002.

[12] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," The journal of Neuroscience, vol. 5, no. 7, pp. 1688-1703, 1985.

[13] J. C. Huegel, O. Celik, A. Israr, and M. K. O'Malley, "Expertise-based performance measures in a virtual training environment," Presence: Teleoperators & Virtual Environments, vol.18, no.6, pp.449-467, 2009.

[14] O. Celik, M. K. O'Malley, C. Boake, H. S. Levin, N. Yozbatiran, and T. A. Reistetter, "Normalized movement quality measures for therapeutic robots strongly correlate with clinical motor impairment measures," IEEE Trans on Neural Systems and Rehabilitation Engineering, vol. 18, no. 4, pp. 433-444, 2010.

[15] B. Rohrer, S. Fasoli, H. I. Krebs, R. Hughes, B. Volpe, W. R. Frontera, J. Stein, and N. Hogan, "Movement smoothness changes during stroke recovery." The Journal of Neuroscience, vol.22, no.18, pp.8297-8304, 2002.

[16] G. B. Zelenock, T. S. Huber, L. M. Messina, A. B. Lumsden, and G. L. Moneta, Mastery of Vascular and Endovascular Surgery, Section I - Basic Considerations and Peri-operative Care. Lippincott Williams and Wilkins, 2006.

[17] C. Reiley, H. C. Lin, D. D. Yuh, and G. D. Hager, "Review of methods for objective surgical skill evaluation," Surgical Endoscopy, vol. 25, no. 2, pp. 356-366, 2011.

[18] R. Aggarwal, T. Grantcharov, K. Moorthy, T. Milland, P. Papasavas, A. Dosis, and A. Darzi, "An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room," Annals of surgery, vol. 245, no. 6, pp. 992-999, 2007.

[19] H. Nagasaki, "Asymmetric velocity and acceleration profiles of human arm movements," Experimental Brain Research, vol.74, no.2, pp. 319-326, 1989.

[20] H. Beppu, M. Suda, and R. Tanaka, "Analysis of cerebellar motor disorders by visually guided elbow tracking movement," Brain, vol. 107, no. 3, pp. 787-809, 1984.

[21] B. R. Rohrer, "Evolution of movement smoothness and submovement patterns in persons with stroke," Ph.D. dissertation, Massachusetts Institute of Technology, 2002.

[22] S. Balasubramanian, A. Melendez-Calderon, and E. Burdet, "A robust and sensitive metric for quantifying movement smoothness," Biomedical Engineering, IEEE Transactions on, vol. 59, no. 8, pp. 2126-2136, 2012.

[23] N. Hogan and D. Sternad, "Sensitivity of smoothness measures to movement duration, amplitude, and arrests," Journal of motor behavior, vol. 41, no. 6, pp. 529-534, 2009.

[24] K. Takada, K. Yashiro, and M. Takagi, "Reliability and sensitivity of jerk-cost measurement for evaluating irregularity of chewing jaw movements," Physiological measurement, vol.27, no.7, pp.609-622, 2006.

[25] K. Yashiro, T. Nakamura, T. Mizumori, H. Yatani, and K. Takada, "Clinical validity of measuring jerk-cost of jaw movement during speech: Effect of mouthguard design on smoothness of jaw movements," in SICE 2004 Annual Conference. IEEE, 2004, pp. 93-96.

[26] B. Rohrer and N. Hogan, "Avoiding spurious submovement decompositions ii: a scattershot algorithm," Biological cybernetics, vol. 94, no. 5, pp. 409-414, 2006.

[27] I. Oropesa, M. K. Chmarra, P. Sanchez-Gonzlez, P. Lamata, S. P. Rodrigues, S. Enciso, F. M. Sanchez-Margallo, F.-W. Jansen, and E. J. Dankelman, J. Gmez, "Relevance of motion-related assessment metrics in laparoscopic surgery," Surgical Innovation, vol.20, no.3, pp.299-312, 2013.