J Med Syst (2017) 41: 24
DOI 10.1007/s10916-016-0665-4

EDUCATION & TRAINING

Smart Sensor-Based Motion Detection System for Hand Movement Training in Open Surgery

Xinyao Sun[1] · Simon Byrns[2] · Irene Cheng[1] · Bin Zheng[2] · Anup Basu[1]

Received: 20 November 2015 / Accepted: 6 December 2016 / Published online: 20 December 2016

Abstract We introduce a smart sensor-based motion detection technique for objective measurement and assessment of surgical dexterity among users at different experience levels. The goal is to allow trainees to evaluate their performance based on a reference model shared through communication technology, e.g., the Internet, without the physical presence of an evaluating surgeon. While in the current implementation we used a Leap Motion Controller to obtain motion data for analysis, our technique can be applied to motion data captured by other smart sensors, e.g., OptiTrack. To differentiate motions captured from different participants, measurement and assessment in our approach are achieved using two strategies: (1) low level descriptive statistical analysis, and (2) Hidden Markov Model (HMM) classification. Based on our surgical knot tying task experiment, we can conclude that finger motions generated from users with different surgical dexterity, e.g., expert and novice performers, display differences in path length, number of movements and task completion time. In order to validate the discriminatory ability of HMM for classifying different movement patterns, a non-surgical task was included in our analysis. Experimental results demonstrate that our approach had 100 % accuracy in discriminating between expert and novice performances. Our proposed motion analysis technique applied to open surgical procedures is a promising step towards the development of objective computer-assisted assessment and training systems.

Keywords Surgical dexterity · Training performance evaluation · Hidden Markov Model · Smart sensor detection · Leap Motion Controller

This article is part of the Topical Collection on *Education & Training*

✉ Irene Cheng
locheng@ualberta.ca

Xinyao Sun
xinyao1@ualberta.ca

Simon Byrns
sbyms@ualberta.ca

Bin Zheng
bzheng1@ualberta.ca

Anup Basu
basu@ualberta.ca

1 Multimedia Research Center, Department of Computing Science, University of Alberta, Edmonton, AB, Canada

2 Surgical Simulation Research Lab, Department of Surgery, University of Alberta, Edmonton, AB, Canada

Introduction

As a result of technological advances in smart sensing, motion data can be acquired with high precision, which benefits many applications including surgical training. Surgical procedures require a mastery of both technical and judgment skills. The evaluation of this skill set is often carried out towards the end of the training period by certification bodies composed of other surgeons. The current evaluation methods followed by these bodies focus on determining if trainees have obtained the necessary knowledge and judgment for independent practice [1]. However, despite attempts to develop standardized evaluation rubrics, such as the Objective Structured Assessment of Technical Skill (OSATS) examination [2], using other surgeons as human evaluators often results in a significant amount of inherent bias and subjectivity. Furthermore, given the limited number of qualified evaluators and their busy schedules, evaluation cannot be offered on a regular basis to provide helpful feedback to trainees for improvement. We propose using computer-assisted motion analysis, which is able to help measure and decipher patterns in an individual's motor movement, and provide a more objective and reproducible method for evaluating dexterity. We envision a real-time feedback generation system based on comparison of a trainee's motion with a reference database composed of motions from surgeons with varying levels of experience. This database can be accessed through a communication network, e.g., the Internet, thus providing assessment without an evaluator being present in person. This virtual connectivity facilitates information sharing and just-in-time decision making. While computer-assisted motion analysis for training has been discussed for years, its application to open surgical movement analysis is still in its infancy due to the high number of degrees of freedom associated with the motion. In recent years, smart sensors have been increasingly utilized to capture motion data for analysis, e.g., accelerometers are used to analyze motion state [3] and smartphone acceleration sensors are used to classify physical activities [4]. Effective use of computer-assisted systems is a common research topic in healthcare, e.g., Parkinson disease diagnosis [5, 6] and virtual learning [7]. However, these systems have yet to be applied to the evaluation of hand motion performance in open surgery. This paper proposes a smart sensor setup, e.g., the Leap Motion Controller (Leap Motion, San Francisco, CA), for acquiring surgical hand motion without using markers or other devices placed on the hand.

In order to develop an accurate system to measure and evaluate motor skills, it is necessary to understand how human motor skills develop. The most commonly cited contemporary theory is Fitts and Posner's three-stage model that includes the cognitive, integrative and autonomous learning stages [8]. The cognitive stage involves decomposition of a task into discrete steps in order to better understand the mechanics of a movement sequence. In the integrative stage, several discrete steps are concatenated to create movement segments. In the final autonomous stage, the movement sequence is smoothed to generate the complete motion sequence. Note that these stages constitute iterative sub-processes and often involve revisiting earlier stages. Despite the complexity of the human movement, motion analysis can identify patterns in motor movements specific to a particular stage of learning and provides insight into strategies for improving performance. Although this study focuses on surgical tasks, our motion acquisition and analysis technique can be applied to other applications which require comparison of hand motions or gestures.

Motion capture and analysis in surgical procedures

Motion analysis applied to the assessment of surgical dexterity is based on the dynamic system theory of motor skill development [9, 10]. This theory describes how movements made by novice performers become progressively and measurably more efficient with practice. Higher levels of experience consequently result in more timely completion of a task and increased confidence of hand movement. Motion analysis has been used in other fields, particularly in physiotherapy and rehabilitation for gait analysis [11], where kinetic and kinematic data from the experiments can be used to provide guidance for optimizing gait characteristics, improving prosthetic comfort and ambulatory efficiency. Our contribution lies in introducing a robust hand motion detection and assessment system for use in surgical training. Open surgery presents a comparatively more complex environment for motion analysis, as surgeons are able to freely manipulate a variety of tools with both hands, resulting in many degrees of freedom (DOF) for each hand and a plethora of hand-instrument interactions. Hand motion data can be captured by applying a variety of technologies, including force or mechanical, electromagnetic and optical detections. Here we briefly discuss some of the current technology available for acquiring hand motion data.

Mechanically-based motion detection

Gloves with embedded sensors can be worn by a surgeon in order to generate hand and joint position data as well as velocity data. Several commercially available gloves are available for this purpose, including Cyber glove (Cyber glove Systems, San Jose, California), ShapeWrap II (Measured, Fredericton, New Brunswick) and the 5DT Data Glove (Fifth Dimension Technologies, Irvine, California) [11]. These devices measure mechanical deformation of 'bend sensors' that translate hand motion and changes in joint movement into digital signals. Some of these gloves can also be fitted with wireless data systems in order to minimize cables that might obstruct natural movement. However, the gloves themselves are bulky and can impair the user's sense of tactile or haptic feedback.

Electromagnetic-based motion detection

One widely recognized system for obtaining motion data in open surgery is the Imperial College Surgical Assessment Device (ICSAD) [12]. This system is comprised of a Bespoke computer software program and two electromagnetic (EM) 10 mm sensors (Isotrack II, Polhemus Inc, Colchester, Vermont) placed on the dorsum of each hand. Computer software translates the raw movement data obtained from the trackers into three scores of dexterity, including the number of movements of each hand, distance traveled by each hand and the time taken to complete a task. Similar to magnetically-based systems, this traditional form of electromagnetic tracking requires sensors to be placed on hands and fingers, limiting its use in a sterile operating environment.

EM tracking technology has evolved rapidly over the years. For example, trakSTAR^{TM} (Ascension Technology Corp, Shelburne, VT) - a new electromagnetic tracking system which can accommodate multiple sensor units - incorporates tiny sensors that could foreseeably be inserted between two surgical gloves in order to measure the position of each finger in 6 DOF (position and orientation). The magnetic sensor units are tiny with an outer diameter of 1.5 mm. Compared to bulky mechanically based glove systems, sensor embedded gloves using this technology would do little to impede natural movement. The disadvantage is that gloves customized with sensors are needed. Disposal of the gloves after surgery means disposal of the sensors, which can be costly.

Optically-based motion detection

Optically-based surgical motion analysis methods range from analysis of raw video to the use of specially designed motion tracking technologies that employ optical markers. By using a camera or an array of cameras to determine object or hand position, these systems do not impede natural hand movement. However, they have to be properly located to strike a balance between obstruction and occlusion.

In open surgery, Glarner and colleagues conducted a feasibility study in the Department of Surgery at the University of Wisconsin, Madison, where they applied a digital video analysis system to record in the operating room [13]. When the recording is done, an analyst selects a region of interest (ROI) in the video. The template matching tracking algorithm then follows the specified ROI and generates kinematic data including displacement, velocity, and acceleration [14].

Since our goal is to efficiently measure the motion data and evaluate its performance, our technique is designed for any motion data regardless of the type of acquisition system used. Ideally, the acquisition system should not be bulky, or impede natural hand movements. Also, the system should be able to deploy safely in an aseptic environment. Thus, in our case study we used a Leap Motion Controller [15] as the motion detection device. This detector determines hand and finger position in 3D space based on tracking 27 segments of each hand in order to mimic each bone. Similar to the aforementioned optically-based techniques, this markerless and inexpensive sensor does not restrict hand movement. Leap controllers are commonly used in game applications and as a scientific instrument; but to the best of our knowledge, this technology has not been applied to the field of open surgical training.

Objective analysis of hand movements in surgical tasks

Evaluating surgical hand movements is challenging given the operating room environment and multidimensional motion data. Our analysis involves defining a quantitative statistical distance (similarity) between experts and novices. Rosen and colleagues addressed this problem in laparoscopic surgery by using a discrete Markov Model (MM) to reveal the decomposition of a surgical task [16]. They subdivided a complex surgical task in order to select low-level elements that can be associated with quantifiable and measurable parameters. Murphy and colleagues further demonstrated the utility of the MM language model [17]. These results indicate that a stochastic approach might describe the surgical process better than a deterministic approach based on validation via comparison with traditional expert performance [18]. Despite these encouraging findings, a major challenge of motion analysis in open surgery is the high Degrees of Freedom (DOF). Previous studies applied on laparoscopic surgery have had the advantage of a decreased complexity of movement due to more restricted DOF [11]. In contrast, open surgery involves greater DOF and more bimanual (two-handed) maneuvers. As a result, additional parameters such as orientation and position of individual hand components, i.e., finger segments, are necessary to define the current state of a Markov Model. To address this parameter complexity issue in open surgery, we explore the Hidden Markov Model and introduce a cluster based analysis method to detect discrete sets of highly concentrated or clustered information in the sample data for each predefined Markov state, leading to a significant reduction in the data complexity. We achieve data reduction in three steps:

1. Creating a subset of the data associated with each state common to all subjects;
2. Using K-means vector quantization algorithm [19] to identify a number of centers associated with each state;
3. Encoding the raw motion data of the surgical tasks based on these clusters in order to convert the multidimensional data into 1-dimensional vectors with finite symbols.

Processing the data in this fashion effectively generates a discrete Markov Model for each individual. Once the Markov Models are defined to characterize subjects with specific skill levels, it is possible to compute the statistical difference between individual's performance based on their hand motion. By comparing trainees to expert level performance, objective criterion can be generated for evaluating user dexterity.

Our proposed motion detection and evaluation computational model

In order to establish an efficient topology to characterize users with mixed skills, we incorporate the Hidden Markov Model Theory in our computational technique. In a Markov Model (MM), each state has an associated physical meaning, but in Hidden Markov Model (HMM) some of the states are abstract and not related to a specific physical interaction [18]. Although their notations and the architecture are similar, there are several fundamental differences between MM and HMM. In a discrete MM, a unique set of observations that map to each specific state is required. This allows one to match a discrete observation with a vector quantization code-word in order to recognize the corresponding state in subsequent observations. In HMM, even though states can be derived from the same observations, the presence of unobserved states makes establishing relationship between measured parameters and observed states impossible. Nevertheless, HMM supports a more compact model topology that allows the system to model surgical motion in a group of subjects with mixed abilities. By applying an objective evaluation approach making use of the HMM topology, we anticipate eliminating the bias inherent in subjective evaluation. HMM has been used in a similar fashion for activity recognition for personal health applications [21]. Here, we apply HMM to differentiate expert and novice performance in open surgery training. In the following paragraphs we discuss only the basic HMM equations. Readers interested in more detail can refer to Rabiners' review [22].

Given a certain process requiring hand movements and gestures, we represent each task element by one of n distinct hidden states $S_1, S_2, S_3 ... S_n$. However, as "hidden" implies, these states are not directly observable or measurable but the observable sequence of movements are statistically dependent on these hidden states. Fig. 1 depicts the relationship between each observable state O and hidden state S.

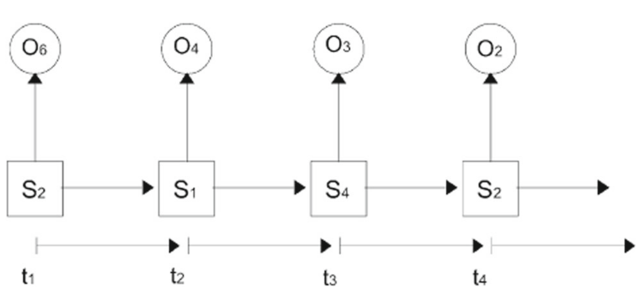


Fig. 1 A Hidden Markov chain representing a sequence of observable states where O_i are the observations, S_j the hidden states that cannot be directly observed, and t_k is the time index [20]

Assume that we have M observable states v_k, $1 \leq k \leq M$. According to the Markov process, the probability of state q at time t only depends on the previous state at time $t - 1$. Thus, we have Eq. 1:

$$P\left[q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_h, q_{t-3} = S_g ...\right] = P\left[q_t = S_j | q_{t-1} = S_i\right] \quad (1)$$

To define a HMM model in our case study, we use the notation $\lambda = (A, B, \pi)$, where A is the state transition probability matrix, which contains the probability of transition from state S_i to S_j. Assuming we have N states, the elements in matrix A are denoted by Eq. 2.

$$a_{ij} = P\left[q_t = S_j | q_{t-1} = S_i\right]$$
$$1 \leq i, j \leq N, a_{ij} \geq 0, \sum_{j=1}^{n} a_{ij} = 1 \quad (2)$$

B is the observation probability matrix, which describes the probability of one state S_j, generates one observation v_k at time t and has elements b_{jk} defined by the following Eq. 3.

$$b_{j(k)} = P\left[v_k \text{ at time } t | q_t = S_j\right]$$
$$1 \leq j \leq N, 1 \leq k \leq M \quad (3)$$

Finally, π is the initial state probability distribution as defined by Eq. 4.

$$\pi_t = P\left[q_1 = S_t\right] 1 \leq t \leq N \quad (4)$$

Hypothesis Based on the above computational model, we hypothesize that:

1. Descriptive statistics applied to hand motion data captured by smart sensor devices would be able to differentiate between novice and expert performance.
2. Furthermore, based on the previous application of HMM to laparoscopy [23], we hypothesize that HMM applied to our tracked data would have at least 80 % discriminatory ability to differentiate between expert and novice performance based on normalized statistical distance to an expert model in open surgery.

Hypothesis validation I - capability to differentiate novice and expert performance

Motion data acquisition

We constructed an open surgery simulator for capturing hand motion data with the Leap Motion Controller. This consisted of an acrylic box with an adjustable system for suspending a monofilament at a consistent position (15 cm) above the Leap sensor. The experimental apparatus is shown in Fig. 2. In this experiment, a nylon monofilament was

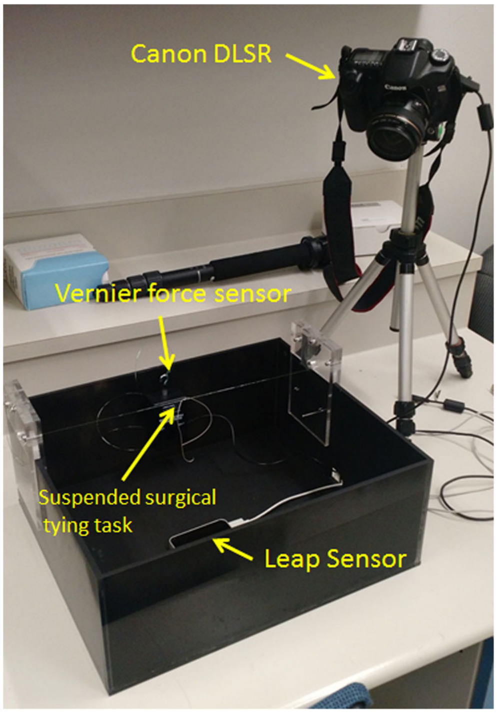


Fig. 2 Experimental setup. Open surgery simulator with Leap Controller and Canon DSLR for real-time motion and video capture respectively

stretched between two adjustable acrylic mounts just above the open top of the box and approximately 15 cm above the Leap sensor. Surgical suture (2-0 Silk or 2-0 polygalactin) was then tied to the nylon filament by each participant. Real time video of the tying task was captured using a Canon 40D Digital Single Lens Reflex (DSLR) Camera in 'Live-view' mode. The Leap sensor was placed cross to the middle bottom, as shown; with the nylon monofilament towards the rear part of the box to ensure that the participant's hands were centered over the Leap device during the suture tying task.

The Leap controller communicated with a Personal Computer (PC) via a Universal Serial Bus (USB). Data was captured using our new custom software developed in C++. Figure 3 depicts our Graphical User Interface (GUI), which illustrates a 3D hand model on the left panel. The right panel displays the corresponding raw video image captured by the camera. An overview of the motion acquisition pipeline is shown in Fig. 4.

In our system, two streams of motion data are captured. The first is recorded at a fixed frequency simultaneously with video from the camera at 60Hz. The second stream contains all the possible motion data captured by the Leap Controller during the recording period (typically greater than 100Hz). Both streams are saved to disk with the first as a space-delimited text file, which contains readable data for the pre-selected components of each hand for use in later analysis (e.g., palm position and palm angle). The second data file contains a series of bytes generated by the Leap Motion API that allows one to replay the tracked data and contains all the motion data that the Leap is able to register. We also implemented a function to generate the

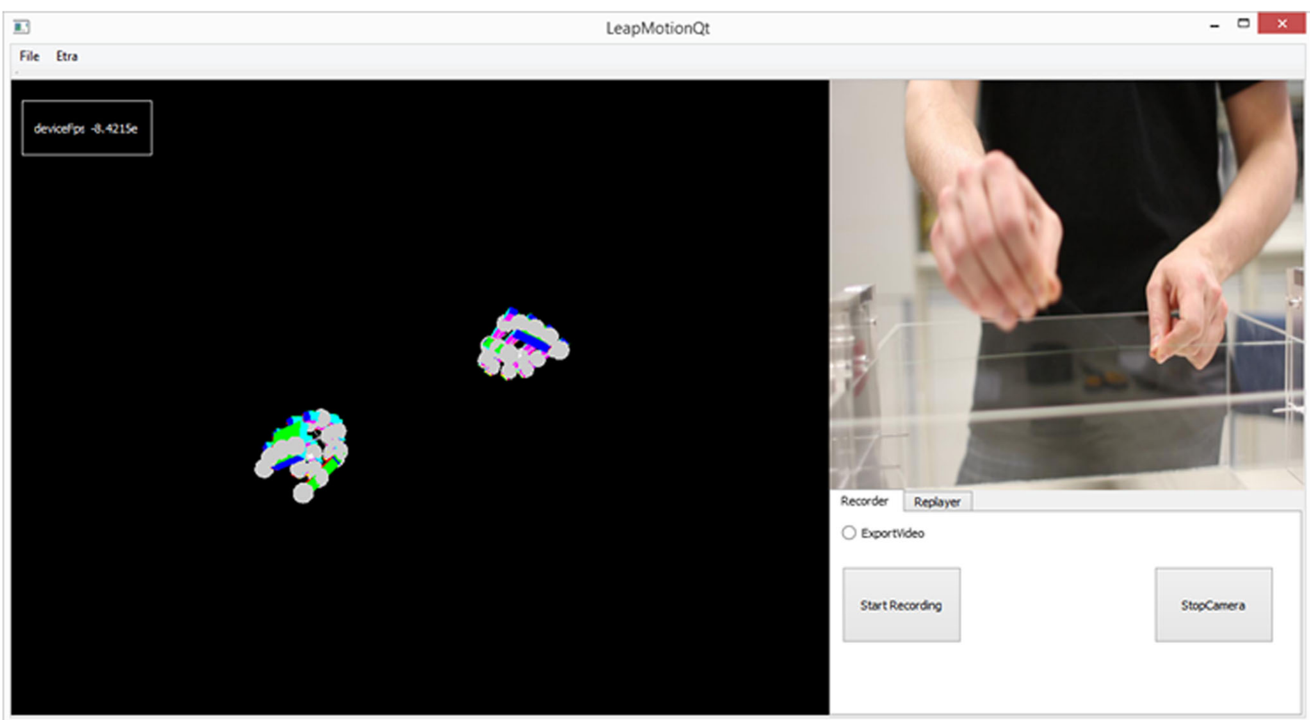


Fig. 3 Screenshot of the GUI for our motion capture application: (*Left*) 3D hand model with Cartesian position data, and (*Right*) video image captured by the digital camera

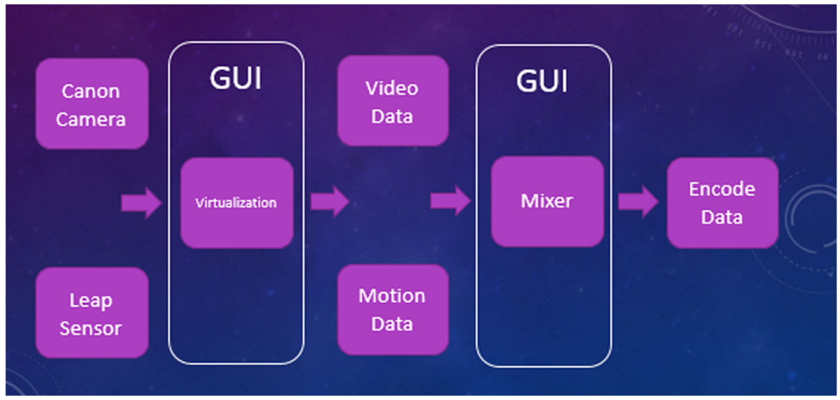


Fig. 4 An overview of our motion data acquisition pipeline

readable text file from a binary Leap recording. This allows us to choose whatever parameters are needed from a single recording in subsequent analysis.

Experimental setting

Early testing of the Leap Controller revealed that the raw motion data generated by the proprietary software driver had an inconsistent frequency. Thus, we implemented an interpolation function to ensure that the captured data was recorded at a constant frequency. After interpolation, we applied a two-pass Butterworth filter algorithm to smooth the data and eliminate high frequency noise in the tracked data.

Two experiments were carried out based on motion data obtained with the Leap Controller. In the first, two participants were asked to perform a sequence of one- and two-handed surgical ties to place a total of five square knots on the nylon monofilament in the training box. Motion data was captured from both procedures and analyzed using traditional metrics: path length, number of movements and total time.

In the second experiment, which was designed to test our objective evaluation algorithm, we chose a simple non-surgical procedure involving an object transfer task using a single hand. This required a participant to lift an object, transfer it from Point A to Point B, and release the object prior to transferring the object back to Point A and again to point B (3 transfers). A total of nine participants were asked to perform this procedure. The first six participants were asked to perform the task as efficiently and smoothly as possible, emulating expert movements. An additional three participants were asked to perform the task with more hesitation, including idling between movements to emulate novice behavior. The performance of the participants was then analyzed.

Motion analysis using hidden Markov models

In order to objectively assess the performance of a participant in a given procedure, we first create a HMM $\lambda = (A, B, \pi)$ that characterizes an expert surgeon's motion pattern. Second, we require an evaluation function to compare a participant's performance with the expert HMM. To achieve this, we need to solve the following HMM steps [22].

1. *Evaluation Step*: In order to compute the probability of the observation sequence that can be generated by the given model λ, we can apply the Viterbi algorithm to find the most likely sequence of states by solving the following:

$$\text{Given:} \begin{cases} \lambda = (A, B, \pi), \\ O = o_1, o_2, o_3 \ldots o_T, \end{cases}$$
$$\text{Compute:} \left\{ P(O|\lambda) \right.$$

2. *Training Step*: To optimize the model parameters (A, B, π) for maximizing the ability of the model to generate the training observation sequence, we can use the Baum-Weltch algorithm.

$$\text{Given:} \left\{ \lambda = (A, B, \pi), \right.$$
$$\text{Adjust:} \left\{ A, B, \pi, \right.$$
$$\text{Compute:} \left\{ P(O|\lambda) \right.$$

Our proposed assessment approach is carried out in four phases and the outcome is described in the Results section. The four phases are:

1. *Data acquisition*: using the system we presented above to capture the required raw data for the predefined procedure.
2. *Data processing*: filtering and smoothing the raw data for use in subsequent analysis.
3. *Data classification and modeling*: vector quantization to transform the multi-dimensional motion data to an one-dimensional observation sequence; optimizing the fitted parameters from the model.
4. *Evaluation*: using statistical functions to compare test performances with the expert model.

Descriptive statistics obtained from hand tying task

In the surgical simulation reported in the ICSAD experiment [12], the authors completed two tasks: a small bowel anastomosis and a vein patch insertion into an artery. Surgical performance was measured based on the number of hand movements made and time taken to complete a task. In a similar fashion, our interpolated and filtered data require tangential velocity analysis in order to calculate the number of movements. The cumulative distance travelled by each hand was calculated to determine the path length, before the total time to complete each task was compared. The number of movements was calculated by using a peak finding algorithm with a threshold set to the mean tangential velocity. These computations were applied to the surgical knot tying tasks, where both a novice and an expert completed five square knots using a one-hand and two-hand typing technique. The results are shown in Table 1. As expected, novices required more movements and time to complete each task. Figure 5 depicts the tangential velocity curves and mean velocity threshold (right hand) of novices and experts in the two-hand tying task. Here, novices are generally slower and have decreased peak velocities compared to experts. This experiment demonstrates the feasibility of our smart sensor-based system as an evaluation metric comparable to the ICSAD system.

Hypothesis validation II - discriminatory ability of hidden Markov model to measure performance

The above differentiation between user performance at different dexterity levels is based on Hidden Markov Model analysis. But does the HMM approach have sufficient discriminatory capability to measure performance?

Gesture definition and data filtering

To address the question above, we conducted a non-surgical transfer task and defined six gestures (states): Idle, Dropping, Grasping, Evaluating, Translating and Releasing, as

Table 1 Surgical knot tying tasks: ICSAD metrics obtained using Leap sensor

Task	Experience Level	Path Length ($mm \times 10^3$)	Number of Movements	Task Completion Time(s)
One-hand	Novice(n = 1)	2.23	43	49
	Expert(n = 1)	1.11	24	22
Two-hand	Novice(n = 1)	3.08	40	65
	Expert(n = 1)	3.05	26	36

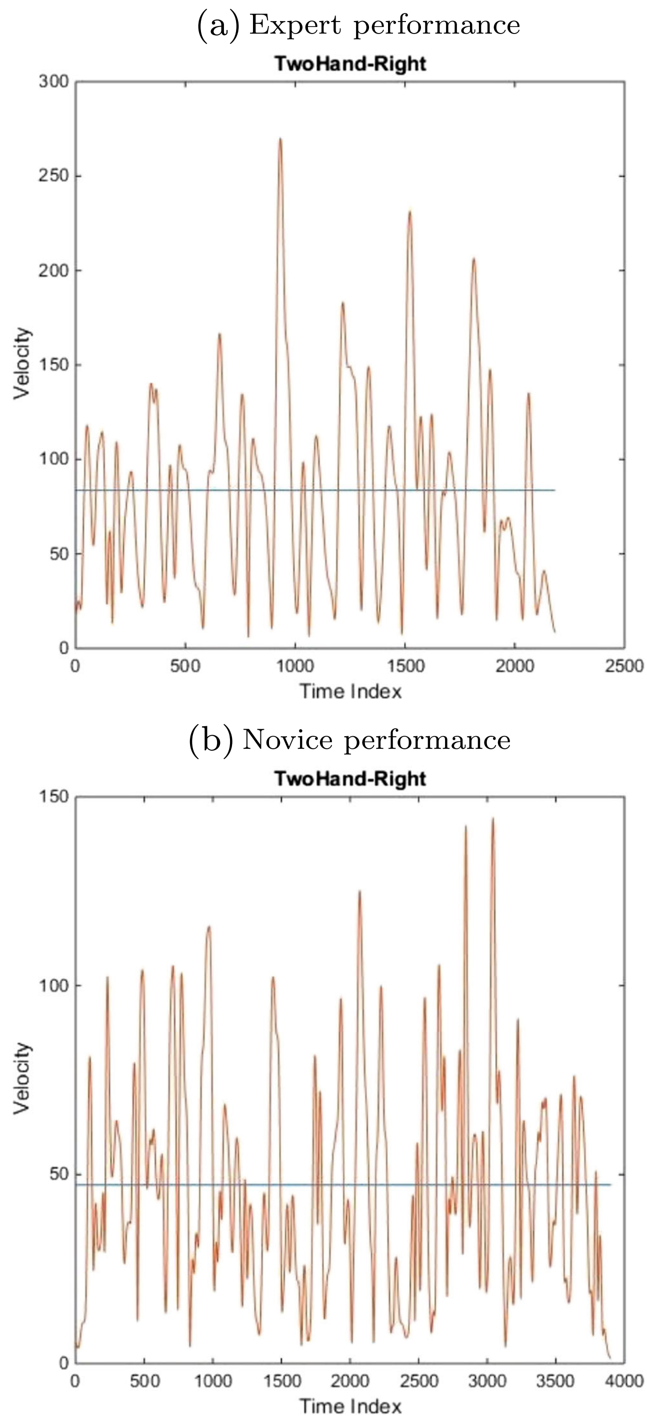


Fig. 5 Tangential velocity curves and mean velocity threshold of the right hand motion when performing two-hand tying task

the key motion segments of a procedure. We then selected the Palm velocity (v_x, v_y) along the x-axis and y-axis, and the velocity (v_s) in the change of distance between the thumb tip and middle finger as depicted in Fig. 6. We selected these points of interest because they feature significant movements in the predefined gestures.

After the hand motion and video capture steps (as illustrated from Fig. 2 to Fig. 4), we applied the interpolation function and Butterworth filter to reduce noise in the raw data (Fig. 7a) and generated a new data set (Fig. 7b) with a frequency of 60 Hz.

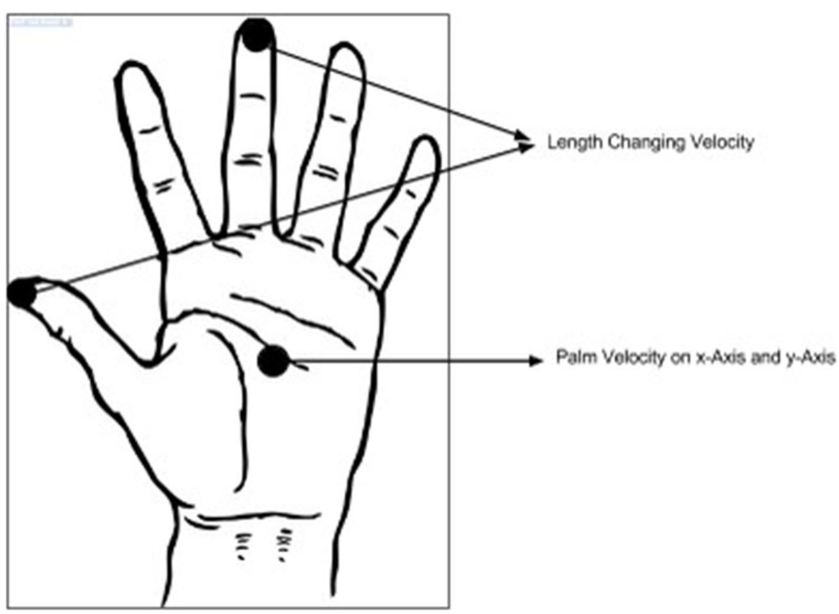
Data classification and modeling

We replayed the recorded synchronized video and motion data, and an expert was asked to identify the Frame IDs corresponding to the start and end of each gesture. This manual segmentation of the motion data served as the ground truth for movement classification allowing us to define the observation distribution matrix B for the data classification and modeling phase, where we built the reference HMM representing expert performance. We used the six predefined gestures as the hidden states. We then used the training function to train the model and initialized all the elements in each row of the state transition probability matrix A with $\frac{1}{N} = \frac{1}{6}$. Initially, the classification was done using frame-by-frame video analysis by an expert; but once initialized, the HMM was able to update matrix A through training algorithms.

After initializing Matrix A, we generated Matrix B which stores the observation probabilities. Each one of the six states was associated with a unique set of velocities (v_x, v_y, v_s). To simplify the theoretical and computational load of the modeling process, we applied an efficient data reduction process to translate the multidimensional data to a one-dimensional observation vector sequence. As part of this process, we applied a K-means vector quantization algorithm [19]. This allowed us to transform the continuous three-dimensional vectors into one-dimensional vectors of 60 observation symbols (10 symbols for each of the 6 states). After applying the K-means algorithms to the motion data representing expert performance, we were able to identify 60 clusters of associated motion data (velocities). The cluster centers were used as the observation symbols for encoding all the multi-dimensional motion data to one-dimensional sequences corresponding to the 60 clusters. To achieve this, we calculated each frame's Euclidean Distance from each of the cluster centers and chose the minimum as the observation symbol for the current frame. We were then able to initialize the observation distribution matrix B corresponding to our expert-encoded observation symbol sequence and the video analysis table using Eq. 5. The initial state probability distribution can be defined based on the assumption that all tasks start in the idle state. Therefore, we have Eq. 6.

$$b_{jk} = \frac{\text{No. of Frames staying at State } S_j \text{ and Using the observation symbol } v_k}{\text{No. of Frames staying at State } S_j}, \quad 1 \leq j \leq N, 1 \leq k \leq M \tag{5}$$

Fig. 6 Selected points of interest on the hand from which velocity data were obtained



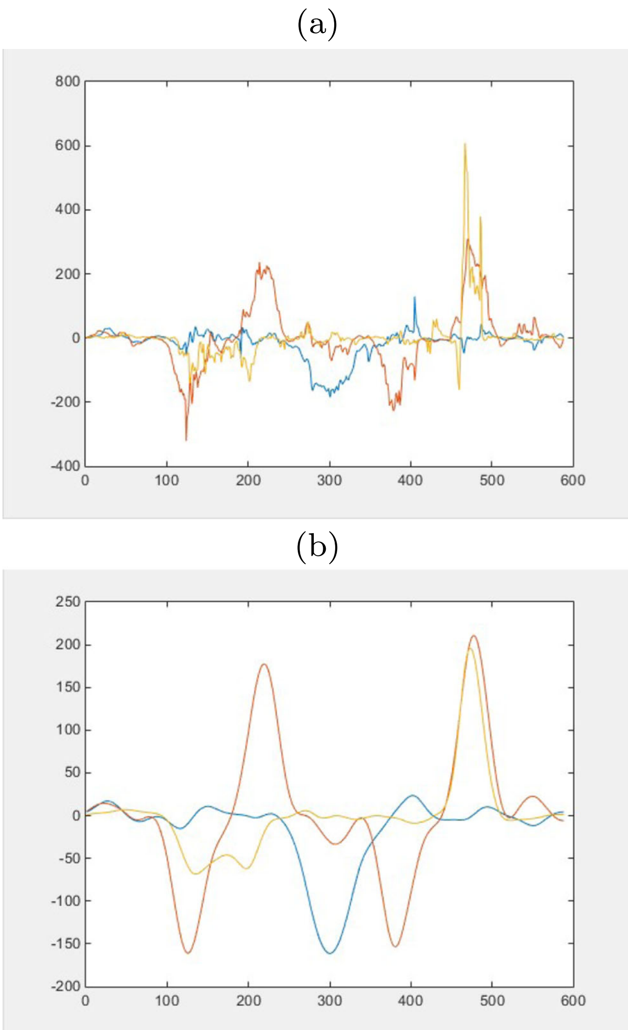


Fig. 7 **a** Raw velocity data over time (v_x-*Blue*, v_y-*Yellow*, v_s-*Orange*) for a portion of the whole procedure performed by an expert, and **b** Velocity over time in (**a**) after interpolation and filtering

$$\pi_{i(idle)} = 1, \pi_{2\ldots N} = 0 \tag{6}$$

The above process describes the initialization of all the parameters necessary to define the HMM, $\lambda = (A, B, \pi)$ for our expert reference model. Next, we used additional sets of motion data captured from the expert(s) to train the model. This step further optimized the parameters in the model, making it more reliable and accurate to describe expert performance. Figures 8 and 9 show the color-coded results of the optimized matrix A and B after training.

Defining the evaluation functions

After obtaining the trained models corresponding to the expert performance λ_E , we applied an evaluation function to compare test behaviors to our expert performance.

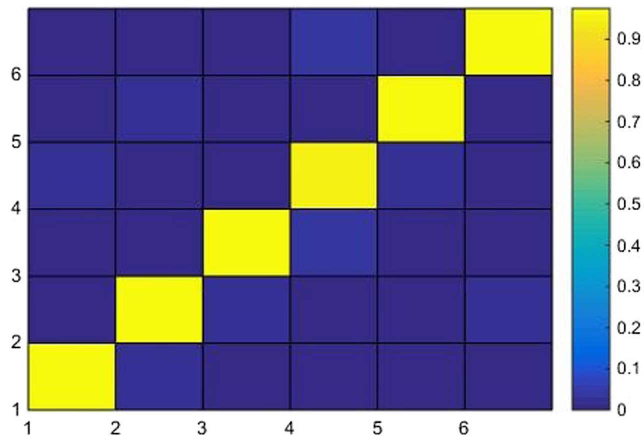


Fig. 8 The probability of transition from each state to any other state

This function provides the probability that a performance described by a given observation symbol sequence is generated by the expert group model λ_E. We normalized the values generated by the evaluation function [20] in order to compare different observations composed of different lengths. We defined NL as the normalized mean probability that a given observation sequence matched the expert model in Eq. 7.

$$NL(O, \lambda_E) = \frac{1}{n}\sum_{i=1}^{n}\left[log P(\hat{O}_i^E|\lambda_E)\right] = 1,$$

$$\hat{O}_i^E \text{ is the observation sequence automatically generated by } \lambda_E \text{ with length } i. \tag{7}$$

$$S(O, \lambda_E) = \frac{|log P(O|\lambda_E) - NL(O|\lambda_E)|}{\frac{1}{n}\sum_{i=1}^{n} |log P(\hat{O}_i^E|\lambda_E) - NL(O|\lambda_E)|} \tag{8}$$

The final evaluation metric for the evaluation score S is then calculated. In Eq. 8, S is the normalized distance of a given observation sequence from the normalized set of observations generated by the expert model. This value can be used to objectively assess performance based on different observation sequences obtained from hand motion data. Lower S scores ($E4$, $E5$ and $E6$) imply that the performance closely approximates the expert model. Novices ($N1$, $N2$ and $N3$) have a larger normalized distance.

Effectiveness of HMM in performance evaluation

In order to evaluate the effectiveness of HMM for hand motion data classification, we collected motion data from the non-surgical object transfer procedure described in "Experimental setting". From a total of nine participants, six experts were in Group $E1$: (E_1, E_2, E_3) and Group $E2$: (E_4, E_5, E_6). Three novices were included in Group N:

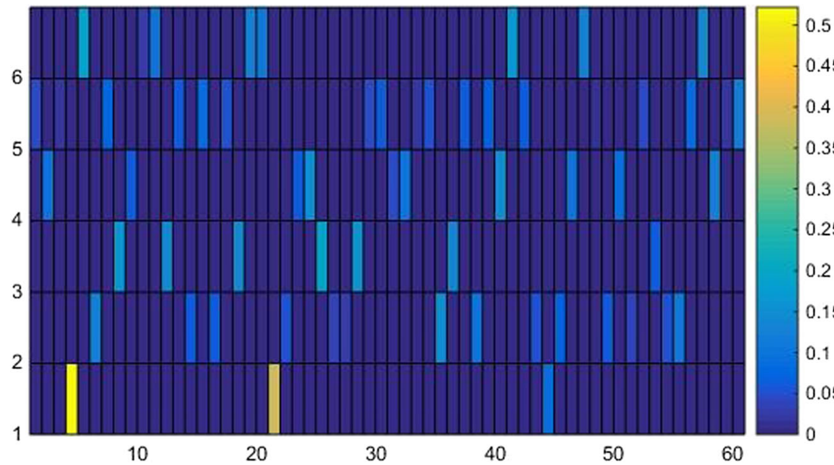


Fig. 9 The probability of each state generating an observation symbol

(N_1, N_2, N_3). Training of the expert reference model utilized motion data from Group $E1$. The evaluation function was then applied to the remaining subjects in Group $E2$ and Group N. Figure 10 shows the original plots of the filtered motion data with (v_x, v_y, v_s) parameters for each replicate.

Visual inspection of the velocity profiles already suggested a difference in the motion data between experts and novices. The novice velocities were more discreet when compared to the amount of overlap in the expert trials. Next, we applied the HMM decoding function to identify the hidden states and compared the decoded gestures

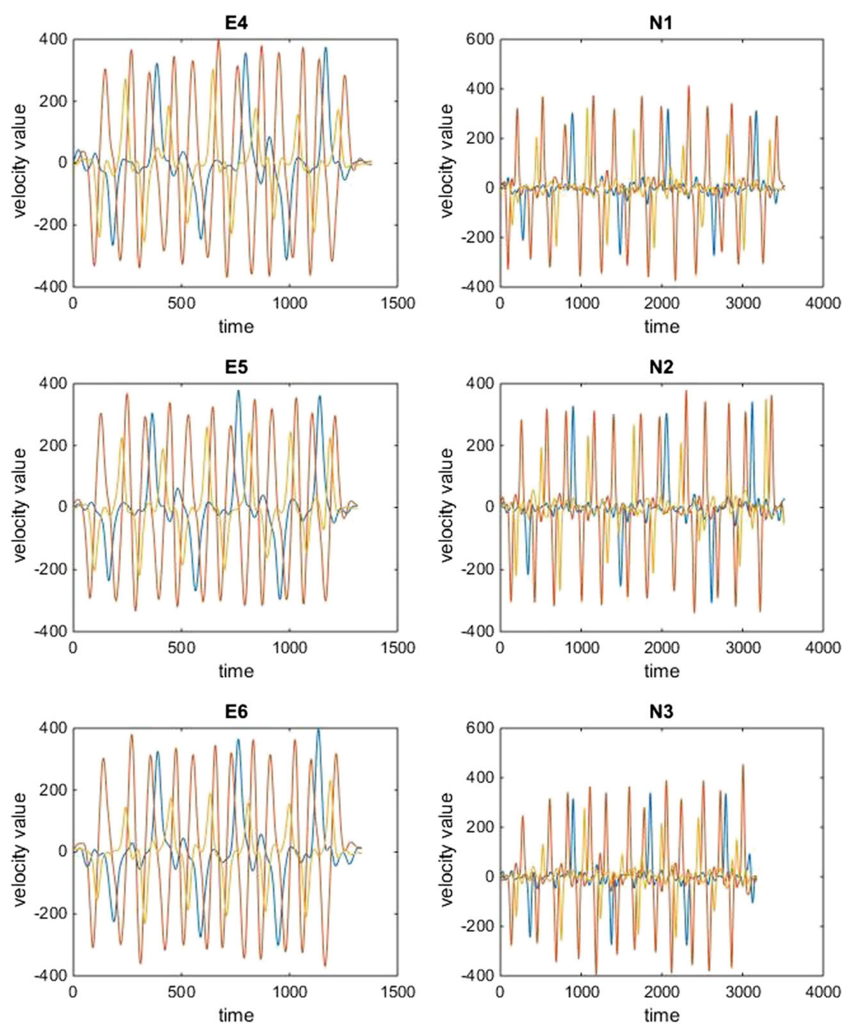


Fig. 10 Velocity over time obtained from 3 experts and 3 novices during the non-surgical transfer task

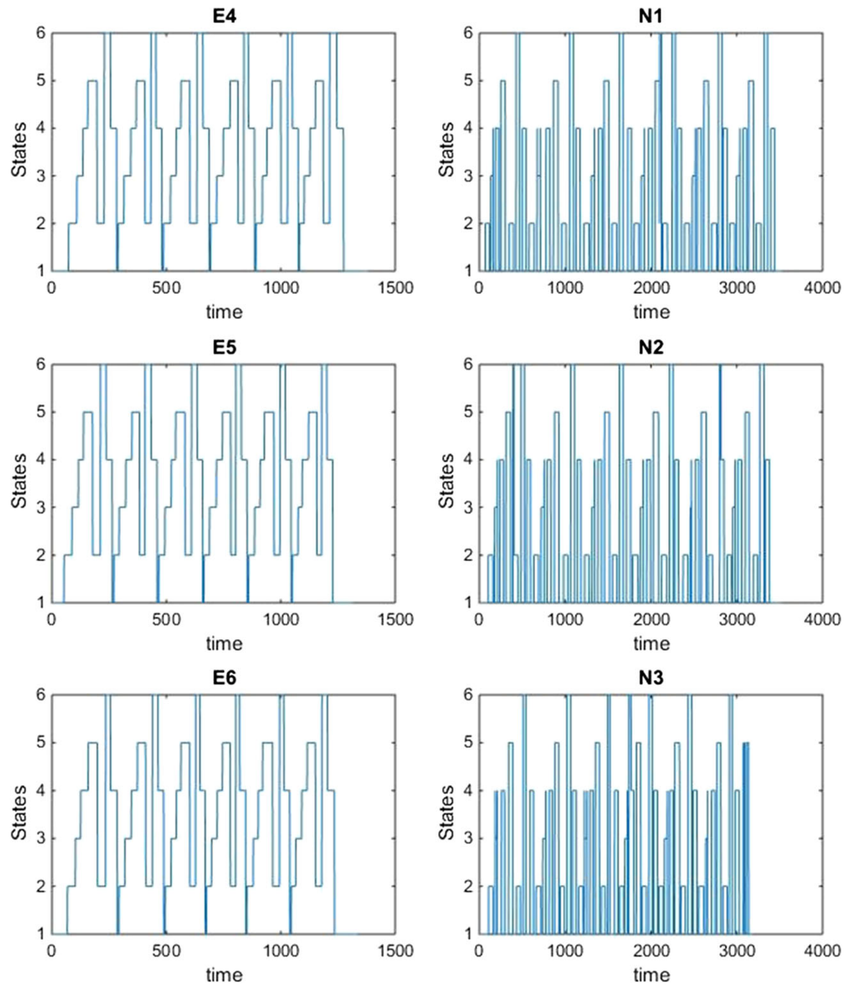


Fig. 11 Decoded state transitions obtained from 3 experts and 3 novices during the non-surgical transfer task

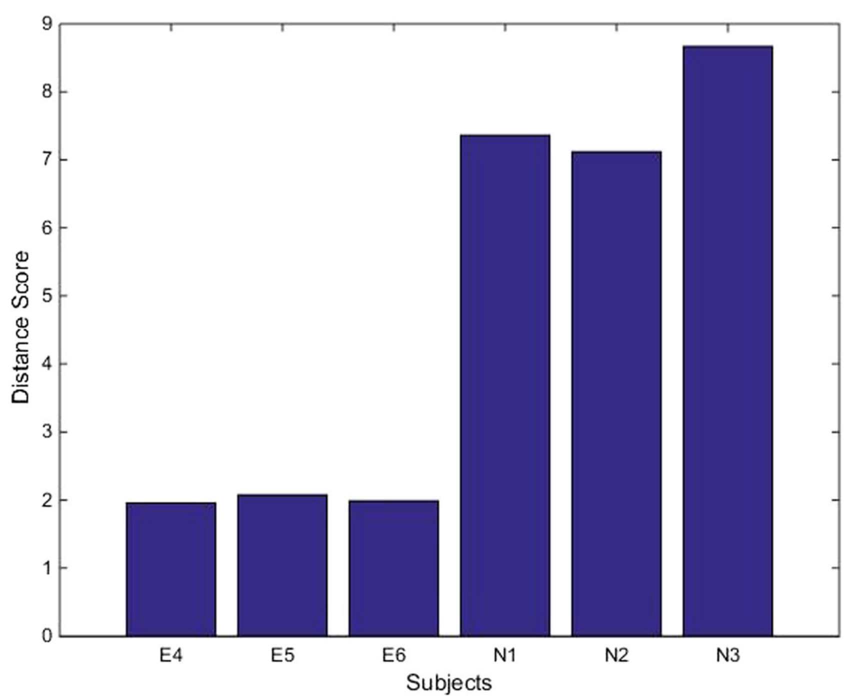


Fig. 12 Normalized distances obtained from 3 experts and 3 novices during the non-surgical transfer task

Table 2 Normalized distances obtained from 3 experts and 3 novices in a non-surgical transfer task

Experience	Statistical	t-test
Level	Distance	P-value
Novice	7.1±0.80	0.007
Expert	2.01±0.06	

(Fig. 11). We then calculated the normalized distance (S_{E4}, S_{E5}, S_{E6}, S_{N1}, S_{N2}, S_{N3}) for each novice and expert performance. The comparison of normalized distance is shown in Fig. 12.

We performed a statistical comparison between the three replicates using Students t-test and found that the distance S between the expert and novice groups was statistically significant as supported by the P-vlaue of 0.007 in Table 2. Differing from the ICSAD simulated tasks performed at depth with lower DOF [24], we investigate open surgery tasks and are able to discriminate between skill levels of subjects.

Discussion & future work

Surgical procedure movements can be complex and the navigation path length may vary depending on specific operations and patient conditions. Only with a sufficiently large library of gestures and predefined hidden markov states, will the system be able to classify with confidence all the maneuvers in a particular operation. Advanced studies on the rapidly developing smart sensor technology will improve the accuracy and reliability of motion data acquisition. Nevertheless, the encouraging results of our study demonstrate that hand motion data obtained from smart sensors has the capability to evaluate individual performance at different dexterity levels.

Traditional metrics make use of a combination of motion analysis techniques such as path length, movement volume and number of movements. These metrics can be useful in providing a baseline assessment of the movement efficiency. The goal of our approach is to complement these traditional motion analysis techniques, by exploiting hidden markov modeling that is necessary as a computational framework in order to compare performance between different participants at a subtask level, which is missing in traditional motion analysis metrics. While video based analysis methods have been deployed in the past, our smart sensor technique provides much richer information including 3D hand position and much higher degrees of freedom. In addition, our motion analysis algorithm is sensor-platform independent and can be applied to motion data captured from different sensors. Our computer-assisted system can measure performance differences between experts and novices using both low level descriptive statistics and hidden markov modeling. To the best of our knowledge, this is the first application of a markerless tracking system for objectively measuring surgical dexterity.

We believe that in order to improve the objectivity and effectiveness of hand motion assessment, analysis needs to be performed at the subtask or surgeme [25] level. Since establishing a robust hidden markov model is a fundamental step towards surgeme analysis, we focus on building such a model in this paper. In future work, we will explore the use of a multi-Leap system, in conjunction with a potential future release of head-mounted Leap sensors, for hand motion capture. By fusing multi-view data acquired from different sensors, we will investigate whether motion classification accuracy can be improved, and how such a step will affect the performance of real-time hand motion assessment.

References

1. Reznick, R., Regehr, G., MacRae, H., Martin, J., and McCulloch, W., Testing technical skill via an innovative "bench station" examination. *Am. J. Surg.* 173(3):226–230, 1997.
2. Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., and Brown, M., Objective structured assessment of technical skill (osats) for surgical residents. *Br. J. Surg.* 84(2):273–278, 1997.
3. Jin, G. H., Lee, S. B., and Lee, T. S., Context awareness of human motion states using accelerometer. *J. Med. Syst.* 32:93–100, 2008.
4. Arif, M., Bilal, M., Kattan, A., and Ahamed, S. I., Better physical activity classification using smartphone acceleration sensor. *J. Med. Syst.* 38:1–10, 2014.
5. Sakar, C. O., and Kursun, O., Telediagnosis of parkinson's disease using measurements of dysphonia. *J. Med. Syst.* 34(4):591–599, 2010.
6. Ozcift, A., Svm feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of parkinson disease. *J. Med. Syst.* 36(4):2142–2147, 2012.
7. Choi, K.-S., Chan, S.-H., and Pang, W.-M., Virtual suturing simulation based on commodity physics engine for medical learning. *J. Med. Syst.* 36:1781–1793, 2012.
8. Fitts, P. M., and Posner, M. I. *Human performance*. Oxford England: Brooks/Cole, 1967. http://psycnet.apa.org/psycinfo/1967-35040-000.
9. Rosenbaum, D. A. *Human motor control*: Academic Press, 2009.
10. Stergiou, N., Harbourne, R. T., and Cavanaugh, J. T., Optimal movement variability: a new theoretical perspective for neurologic physical therapy. *J. Neurol. Phys. Ther.* 30(3):120–129, 2006.
11. Reiley, C. E., Lin, H. C., Yuh, D. D., and Hager, G. D., Review of methods for objective surgical skill evaluation. *Surg. Endosc.* 25(2):356–366, 2011.
12. Datta, V., Mackay, S., Mandalia, M., and Darzi, A., The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *J. Am. Coll. Surg.* 193(5):479–485, 2001.
13. Glarner, C. E., Hu, Y.-Y., Chen, C.-H., Radwin, R. G., Zhao, Q., Craven, M. W., Wiegmann, D. A., Pugh, C. M., Carty, M. J., and Greenberg, C. C., Quantifying technical skills during open operations using video-based motion analysis. *Surgery* 156(3):729–734, 2014.
14. Radwin, R. G., Automated video exposure assessment of repetitive motion. In: Proceedings of the Human Factors and

Ergonomics Society Annual Meeting, Vol. 55, pp. 995–996: SAGE Publications, 2011.
15. Guna, J., Jakus, G., Pogačnik, M., Tomažič, S., and Sodnik, J., An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors* 14(2):3702–3720, 2014.
16. Rosen, J., Brown, J. D., Chang, L., Sinanan, M. N., and Hannaford, B., Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete markov model. *IEEE Trans. Biomed. Eng.* 53(3):399–413, 2006.
17. Murphy, T. E., Vignes, C. M., Yuh, D. D., and Okamura, A. M., Automatic motion recognition and skill evaluation for dynamic tasks. *Proc. Euro. Haptics*, 363, 2003.
18. Leong, J. J. H., Nicolaou, M., Atallah, L., Mylonas, G. P., Darzi, A. W., and Yang, G.-Z., Hmm assessment of quality of movement trajectory in laparoscopic surgery. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006, pp. 752–759: Springer, 2006.
19. Gorman, P., Krummel, T., Webster, R., Smith, M., and Hutchens, D., A prototype haptic lumbar puncture simulator. *Stud. Health Technol. Inform.* 106–109, 2000.
20. Megali, G., Sinigaglia, S., Tonet, O., and Dario, P., Modelling and evaluation of surgical performance using hidden Markov models. *IEEE Trans. Biomed. Eng.* 53(10):1911–1919, 2006.
21. Li, Z., Wei, Z., Yue, Y., Wang, H., Jia, W., Burke, L. E., Baranowski, T., and Sun, M., An adaptive hidden markov model for activity recognition based on a wearable multi-sensor device. *J. Med. Syst.* 39(5):57, 2015.
22. Rabiner, L. R., A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE* 77(2):257–286, 1989.
23. Loukas, C., and Georgiou, E., Multivariate autoregressive modeling of hand kinematics for laparoscopic skills assessment of surgical trainees. *IEEE Trans. Biomed. Eng.* 58(11):3289–3297, 2011.
24. Royal College of Surgeons of England. Basic surgical skills instructors' handbook. http://www.journalacs.org/article/S1072-7515(01)01041-9/references.
25. Reiley, C. E., and Hager, G. D., Task versus subtask surgical skill evaluation of robotic minimally invasive surgery. In: Medical image computing and computer-assisted intervention–MICCAI 2009, pp. 435–442: Springer, 2009.