sEMG-based Motion Recognition for Robotic Surgery Training - A Preliminary Study

Chenji Li, Chao Liu, Arnaud Huaulmé, Nabil Zemiti, Pierre Jannin, Philippe Poignet

sEMG-based Motion Recognition for Robotic Surgery Training - A Preliminary Study

Chenji Li[1], Chao Liu[1], Arnaud Huaulmé[2], Nabil Zemiti[1], Pierre Jannin[2], Philippe Poignet[1]

*Abstract***— Robotic surgery represents a major breakthrough in the evolution of medical technology. Accordingly, efficient skill training and assessment methods should be developed to meet the surgeon's need of acquiring such robotic skills over a relatively short learning curve in a safe manner. Different from conventional training and assessment methods, we aim to explore the surface electromyography (sEMG) signal during the training process in order to obtain semantic and interpretable information to help the trainee better understand and improve his/her training performance. As a preliminary study, motion primitive recognition based on sEMG signal is studied in this work. Using machine learning (ML) technique, it is shown that the sEMG-based motion recognition method is feasible and promising for hand motions along 3 Cartesian axes in the virtual reality (VR) environment of a commercial robotic surgery training platform, which will hence serve as the basis for new robotic surgical skill assessment criterion and training guidance based on muscle activity information. Considering certain motion patterns were less accurately recognized than others, more data collection and deep learning-based analysis will be carried out to further improve the recognition accuracy in future research.**

I. Introduction

With the advancement of robot technology, robotic surgery has been more and more adopted worldwide in clinical practice, with da Vinci surgical robot system as the most well-known and successful example. Surgical robots show great advantages in minimally invasive surgery, especially in laparoscopic surgery [1,2], such as improved ergonomics and visualization, increased operation dexterity, proper hand-eye coordination, eliminating the fulcrum effect and making instrument manipulation more intuitive etc [3].

Surgical training has been long proven to be important to increase operation safety and clinical outcome [4]. Although considered an evolution of conventional laparoscopic surgery, the skills required for robotic surgeons are console-based and maneuvers without haptic feedback. Traditional training method like "see one, do one, teach one" is expensive, human resource heavy and prone to patient safety concerns [5]. The need for formal assessment of competency to ensure safe and sustained growth has led various groups to propose training programs in robotic surgery to improve cognitive and procedural skills before reaching the operating room. Both basic console skills (such as camera, pedal, finger control) and advanced console skills (such as excision, suturing and use of diathermy) can be developed in a mentored simulation environment, either undertaken in a VR simulator, dry lab or a wet lab [3]. These training systems possess numerous advantages over traditional training method such as less involvement of the trainer is requested, the training materials are much less costly, reusable, risk-free and thus enable "do as many as possible". Nevertheless, it is noticed that all existing surgical training platforms, especially VR-based simulators, mainly rely on assessment metrics using directly measurable kinematic and temporal data of the operation or the simulation (completion time, path length, instrument collisions, instrument velocity, acceleration and motion smoothness, workspace overlapping, etc.) [6] to evaluate the trainee's skills and generate training scores. Although such feedback information is useful for objective evaluation, few clues can be obtained informing how to further improve the trainee's skills.

In order to address this problem and better overcome the learning curve, we aim to explore the possibility of investigating the trainee's muscle activities by resorting to machine learning and artificial intelligence (AI) to obtain semantic and interpretable information (e.g., dominant active muscle group sequence during the simulated operation and the according activity status, etc.) that hides behind the consequential kinematic training results. It is expected that such semantic information is easier to perceive and understand to the trainee than abstract and numerical information used in existing assessment metrics. Towards this direction, we carried out the preliminary study on motion recognition using sEMG signals with a commercial VR-based robotic surgery training platform. By breaking down the simulated procedure into segments, the kinematic motion primitives are obtained which are usually slow in the context of surgical operation and involve relatively lower level of muscle activity and sparse information compared to those in fast and complex motion trajectories. To address these challenges, we used a variant-length sliding windows proportional to the overall sEMG signal length to retain regional and global muscle activity property. Based on that, a

* This work was supported in part by the French National Agency for Research (Agence Nationale pour la Recherche, ANR) within the Investissements d'Avenir Program (Labex CAMI, ANR-11-LABX0004, Labex NUMEV, ANR-10-LABX-20), and by the ROBOTEX 2.0 (Grants ROBOTEX ANR-10-EQPX-44-01 and TIRREX ANR-21-ESRE-0015)

[1]C. Li, C. Liu, N. Zemiti and P. Poignet are with LIRMM, UMR 5506, Université de Montpellier-CNRS, 34095 Montpellier, France (cli; liu; zemiti; poignet@lirmm.fr)

[2]A. Huaulmé and P. Jannin are with LTSI-INSERM UMR 1099, Université de Rennes 1, 35500 Rennes, France (arnaud.huaulme; pierre.jannin@univ-rennes1.fr)

feedforward neural network (FFNN) is employed for the motion recognition task. The obtained results are promising and show the feasibility of recognizing simple motion primitives with sEMG signals in a VR surgical training environment, which builds as the first basis towards more complex motion recognition with sEMG signals in surgical training.

II. Experiment

A. Material and set up

The robotic surgery simulation platform used in the experiment is BBZ console (BBZ s.r.l., Vigasio, Italy). The device simulates the da Vinci master console and integrates the Xron simulation software which creates the VR training environment. Two joysticks (Right, Left) control the tool motion in the VR.

The sEMG data for this work was acquired by using the Trigno Research+ system (Delsys, USA). Six electrodes with a sampling rate of 2000 Hz were placed on the right arm's extensor carpi ulnaris (ECU), bicep brachii, tricep brachii, anterior deltoid, intermediate deltoid and posterior deltoid [7].

Three lines of different colors were set in the virtual scene to guide the motion (Fig. 1). The red, green and blue guide-lines correspond to the motions along the X, Y and Z axes of the virtual reality Cartesian space respectively. On each guide-line, there are 2 white markers indicating the motion starting and ending points in the VR. Motions along these 3 axes were analyzed in this preliminary study as they represent 3 typical motion primitives and can constitute more complex motions that mimic real motions during training based on this simulation platform.

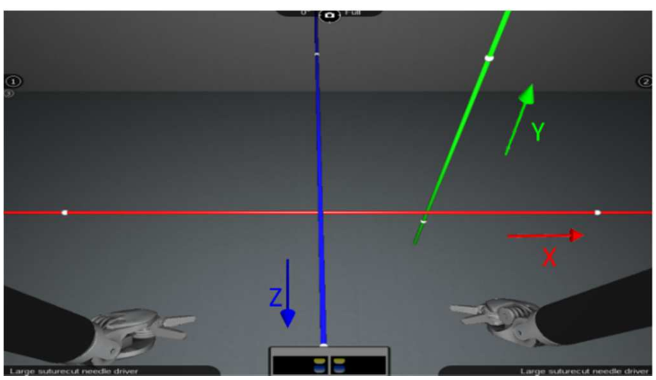

Fig. 1. Three guide-lines in the virtual scene of BBZ simulator

To standardize the starting position of each motion of different subjects, 3D-printed frames were designed to provide reference points for motion along different axis (Fig. 2). The BBZ simulator dimensions and the joystick working space during the training were taken into consideration in order to simulate the natural state of the trainer's motion during the real training.

B. Experiment protocol

The experimental procedures involving human subjects described in this paper were approved by the Institutional Ethics Committee. Seven individuals (ranging from 25 to 28 years of age) with no surgical operation experience and no surgical training experience volunteered to participate in the experiment, and all subjects are right-handed (and in this preliminary study we only right upper limb motion).

Fig. 2. 3D-printed frame in experiment (for Y-axis motion). The monitor shows the position of the Right instrument in the virtual reality.

Before the real experiment, each subject was given 10 minutes to try one simple integrated training course to gain basic understanding on how to maneuver in the VR environment of the BBZ simulator. During the experiment, the subjects were asked to manipulate the Right joystick of BBZ simulator to move the Right instrument in virtual reality along one guide-line back and forth twice. The movement of the instrument in virtual reality needs to be smooth, continuous and follow the guide-line with the least deviation as possible.

As there were two directions of motion along each guide-line in the VR, there were six different motions in the experiment in total. In one test along one guide-line, two round-trips between two white markers are performed. Event trigger was designed to indicate the start/end of each motion by closing and opening the joystick gripper to facilitate the sEMG data segmentation. Each subject repeated the test along each guide-line four times, so there would be twelve tests in total. Between each test, subjects were given a pause time of about two minutes to avoid muscle fatigue.

Through the experiment, all data collected from the seven subjects would be used for data processing and motion classification by using machine learning technique.

III. Data Analysis

A. Data preparation

The sEMG data were firstly bandpass filtered between 10 and 450Hz. The notch filter of 50 Hz was also used. The sEMG data of one test were segmented based on the instrument's kinematic data in virtual reality and event triggers (Fig. 3). For each segment, the first 0.5s and the last 0.5 seconds of data were discarded to eliminate the effect of gripper opening and closing motion on the sEMG data.

B. Sliding windows and feature extraction

In reality, the surgical operation and training motions are relatively slower than most daily life activities, the corresponding muscle activity level is also lower and less dynamic. In addition, the inconsistent speed of the subjects' motion resulted in different sEMG signal lengths. A fixed length sliding window for all sEMG signals might not be suitable, it would be necessary to analyze sEMG data from a larger scale and more global perspective. Our strategy is to set the sliding window length being proportional to sEMG signal and to set the sliding step being proportional to sliding window length. In this work, a sliding window of one-third the length of the sEMG signal is used with a sliding step of one-third the length of sliding windows.

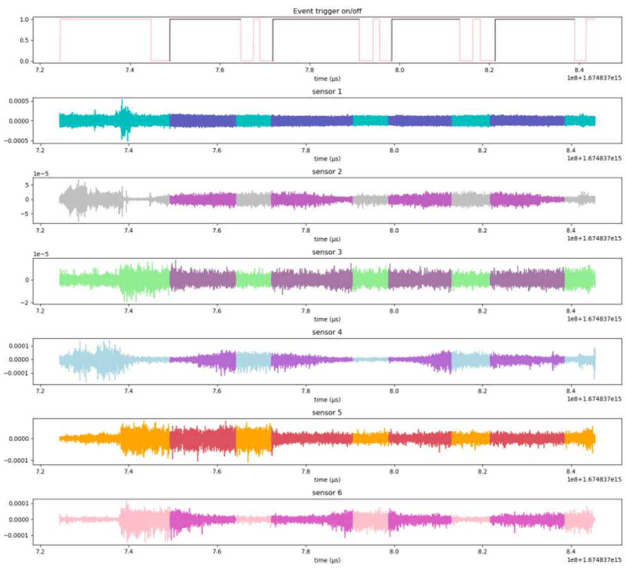

Fig. 3. sEMG data segmented by event trigger

For each sliding window, a total of 16 features [6] are extracted in time domain (Integrated EMG (IEMG), Root Mean Square (RMS), Mean Absolute Value (MAV), Waveform Length (WL), Modified Mean Absolute Value 1 (MMAV1), Zero Crossing (ZC), Modified Mean Absolute Value 2 (MMAV2), Slope Sign Change (SSC), Simple Square Integral (SSI), Willison Amplitude (WAMP), Variance of EMG (VAR)) and in frequency domain (Autoregressive Coefficients (AR), Modified Mean Frequency (MMNF), Mean Frequency (MNF), Modified Median Frequency (MMDF), Median Frequency (MDF)).

After feature extraction for a sliding window, a feature vector of dimension (N,1) is obtained. Then features vectors of all the sliding windows in one sEMG signal sample will be integrated together into a new feature vector of dimension (M*N,1) where M is the number of sliding windows (Fig. 4).

C. Classification Model

Many types of classification methods have been used in literature to recognize sEMG patterns, such as Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), Feedforward neural network (FFNN), Convolutional neural network (CNN), Recurrent neural network (RNN) and Hybrid neural network, etc. [9, 10].

In this work, we use a FFNN model for our preliminary study to classify the surgical training motion. The dataset size in work is not large and hence is not suitable for the data-demanding deep learning techniques such as CNN and RNN. Among the aforementioned machine learning methods, FFNN has proven to have good classification accuracy and is easy to build in terms of parameter tuning. Moreover, FFNN is robust to the individual difference among the subjects, which fits perfectly the case of sEMG measurement in our experiment. Therefore, FFNN represents an ideal analysis tool in this work with good balance between design simplicity and classification accuracy.

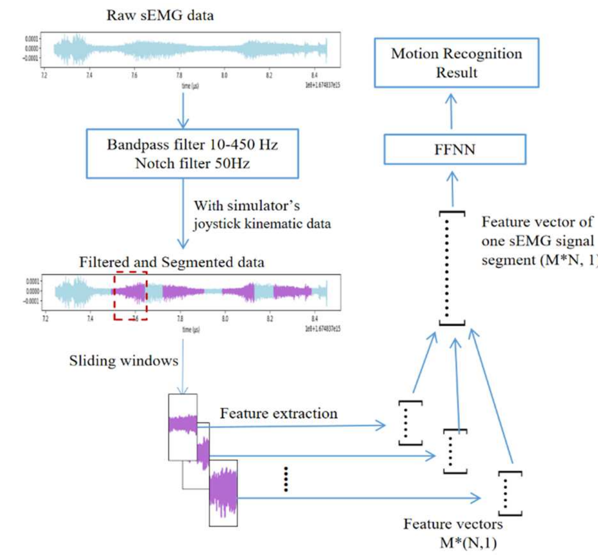

Fig. 4. Diagram of the data analysis workflow

IV. Results and Discussions

For all data collected from the seven subjects, the data from five subjects were used as the training dataset and the data from the remaining two subjects were used as the test dataset.

Eventually, six motion patterns are classified through the FFNN model and are labeled as 0 to 5 (Label 0: left to right along X axis (→red), Label 1: right to left along X axis (←red), Label 2: down to up along Y axis (↑ green), Label 3: up to down along Y axis (↓ green), Label 4: near to far along Z axis (✕blue), Label 5: far to near along Z axis (• blue)).

The result of prediction and the performance of the FFNN model can be presented by a confusion matrix (Fig. 5) which shows the number of samples with their real labels versus their predicted labels. A metrics for performance evaluation is also given in Tab. I.

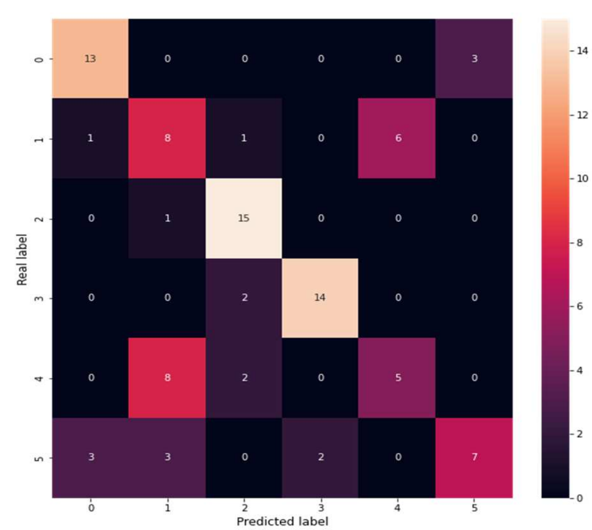

Fig. 5 Confusion Matrix of prediction for 6 motions

TABLE I. REPORT SHOWING THE MAIN CLASSIFICATION METRICS

Label	Precision	Recall	F1-score	Support
0	0.76	0.81	0.78	16
1	0.4	0.5	0.44	16
2	0.75	0.9375	0.83	16
3	0.875	0.875	0.875	16
4	0.45	0.33	0.38	15
5	0.7	0.46	0.56	15
* The missing 1 support for Label 4&5 is due to 1 missed trigger event.				
Accuracy			0.66	

From the F1-score, the motion of Label 3 (up to down along Y axis (green)) has the best score about 0.875 and both of its precision and recall are about 87.5%. The second-best recognized motion is Label 2 (down to up along Y axis). Its recall is about 94% but the precision is about 75%. From the confusion matrix, it shows that the model recognized almost all motion Label 2 except 1 sample mislabeled. However, the model mislabeled two motion samples of Label 4 and two motion samples of Label 5 as motion Label 3 which leads to a low precision. The motion Label 2 and the motion Label 3 are both motions along Y axis. According to feedback from the subjects, the motions along Y axis are the most difficult among all, which makes them exert more effort and muscle control to complete the experiment tasks. From a kinematic point of view, the biceps have relatively more muscular activity when moving along the Y direction, while the deltoids are also involved in the movement, which also made the muscle information rich. In the cases where there are more mislabels in the confusion matrix, it is found that the mislabeled motions are similar to the real motions in terms of the muscle groups used and thus causes difficulties to the motion recognition.

A cross-validation has been carried out by randomly changing the seven subjects in the training and test datasets. The obtained result reached an accuracy of 64.89% with the confusion matrix as shown in Fig. 6.

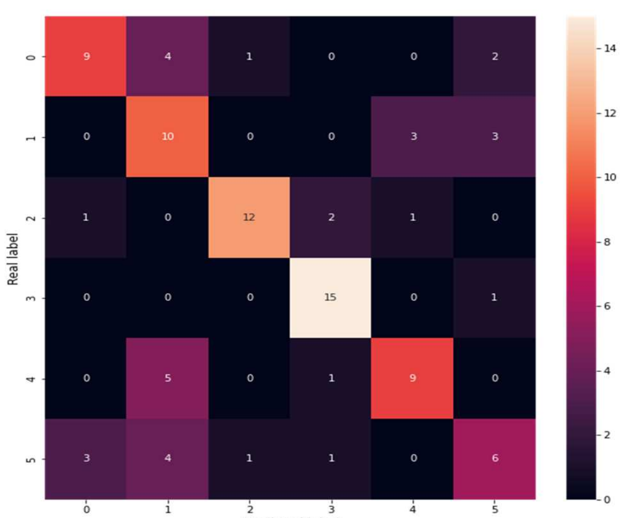

Fig. 6 Confusion Matrix of the cross-validation

In this preliminary study, we used a sliding window that varies its length as the sEMG signal length changes because the length of the sEMG signals varies among different subjects and in different experiment tests. Future research will firstly investigate the Dynamic Time Warping algorithm [11], which is widely used to align two time series of different data length and thus can align the sEMG signal to eliminate the distortion in the time axis. Also, in this work the FFNN model with feature extraction has been used since the collected data size is not very large. Therefore, another important future work is to collect more data with more subjects. With the increased size of collected data in our future experiments, deep learning methods like CNN, RNN and hybrid neural network which are widely applied in motion recognition will be explored to see if the motion recognition accuracy can be further improved. Moreover, through deep learning methods, more spatial and temporal features can be mined to help improving the recognition accuracy. Upon achieving single hand (Right) motion recognition with high accuracy, the Left joystick can be added in for more complex and thus more challenging two-hand motion recognition tasks.

V. CONCLUSION

In this paper, a preliminary study to recognize slow motion primitives through sEMG signal using a commercial robotic surgical training simulator was carried out as an initial step towards developing new surgical skill training and assessment method based on sEMG signal. The obtained results are encouraging and verified the feasibility of motion detection with sEMG signals on VR-based surgical simulation platform. Future works are planned to further improve the motion recognition accuracy as explained in the discussion.

REFERENCES

[1] Lanfranco AR, Castellanos AE, Desai JP, Meyers WC. Robotic surgery: a current perspective. Ann Surg. 239(1):14-21. Jan 2004.

[2] Ploussard G. Robotic surgery in urology: facts and reality. What are the real advantages of robotic approaches for prostate cancer patients? Curr Opin Urol. 28(2):153-158, 2018 Mar.

[3] Sridhar, A.N., Briggs, T.P., Kelly, J.D. et al. Training in Robotic Surgery—an Overview. Curr Urol Rep. 18, 58 (2017).

[4] Fritz T, Stachel N, Braun BJ. Evidence in surgical training - a review. Innov Surg Sci. 4(1):7-13, Apr. 2019.

[5] Rohrich RJ. "See one, do one, teach one": an old adage with a new twist. Plast Reconstr Surg.118:257–258. 2006.

[6] Hong, M., Rozenblit, J.W. & Hamilton, A.J. Simulation-based surgical training systems in laparoscopic surgery: a current review. Virtual Reality 25, 491–510 (2021).

[7] Phinyomark A., Limsakul C., Phukpattaranont P. A Novel Feature Extraction for Robust EMG Pattern Recognition. arXiv:0912.3973 (https://doi.org/10.48550/arXiv.0912.3973)

[8] Soangra R, Sivakumar R, Anirudh ER, Reddy Y SV, John EB. Evaluation of surgical skill using machine learning with optimal wearable sensor locations. PLoS One. 17(6):e0267936, Jun 2022.

[9] Zhou Y, Chen C, Ni J, Ni G, Li M, Xu G, et al. EMG Signal Processing for Hand Motion Pattern Recognition Using Machine Learning Algorithms. Arch Orthop. 2020; 1(1): 17-26.

[10] Xiong D., Zhang D., Zhao X. and Zhao Y., Deep Learning for EMG-based Human-Machine Interaction: A Review. in IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 3, pp. 512-533, March 2021

[11] Keogh E. J. and Pazzani M. J. Scaling up dynamic time warping for datamining applications. In Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '00), 285–289. New York, NY, USA, 2000.