Physics-based task classification of da Vinci robot surgical procedures

M. Selvaggio[1], L. Villani[1], B. Siciliano[1] and F. Ficuciello[1]
[1]*The authors are with with PRISMA Lab, Department of Electrical Engineering and Information Technology, University of Naples Federico II, via Claudio 21, 80125, Naples, Italy, email: fanny.ficuciello@unina.it*

***Abstract*—In this paper, a machine learning algorithm for automatic segmentation of surgical subtasks in teleoperated robot-assisted minimally-invasive surgery (RAMIS) is presented. To improve previously developed methods, we propose to adopt the interaction force measurement in the learning process. This allows us to perform physical considerations of the surgical states and to appropriately select parameters for the learning model. The classification of the surgical states is performed using a Gaussian Mixture Models (GMM) combined with Hidden Markov Models (HMM) to infer the more explicable hidden states of suturing procedures. The experiments are performed using data retrieved from the da Vinci Research Kit (dVRK) and a Force/Torque sensor integrated in a phantom used for suturing. The results demonstrate good accuracy of surgical gestures recognition.**

***Keywords*—Minimally invasive robotic surgery, surgical gestures classification.**

I. Introduction

Combination of teleoperated tasks execution with adaptive assistance strategies is a very promising approach in robotic surgery. In order to employ adaptive and time-varying shared control methods, such as virtual fixtures, task classification constitutes an essential step. It allows to assess surgeon skills and intentions in both training and real interventions. The objective of this work is to develop a reliable method for the automatic classification of surgical tasks. This procedure is often challenging if it relies only on kinematic information. Vision-based techniques might be employed but they bring difficulties related to automatic video interpretation [1]. Hence, we propose to combine kinematic data with the interaction force measurements in the learning process. This is motivated by the purpose of giving physical interpretation to clustered states.

In this work we use Gaussian Mixture Models (GMM) for state classification and combine them with Hidden Markov Models (HMM) for task encoding. The convenience of using a HMM for modelling surgical movements is that it provides analogies with human behavior, which can be thought as a doubly stochastic process, involving a hidden, immeasurable human mental state and a measurable, observable human action. The strength of HMMs is that they do not require any priori definition of what surgical expertise is.

Several works can be found in literature that developed and applied similar approaches. Zappella et al. propose several methods for automatic surgical gesture classification using video and kinematic data [1]. Berthet-Rayne et al. develop a human collaborative framework for bimanual surgical tasks based on learned models [2]: they combine active constraints and learning from demonstration (LfD). Despinoy et al. take into account the operating gesture workflow to provide more intuitive training as well as more accurate gesture and procedural knowledge assessment solutions [3]. Recently, Perez-Del-Pulgar et al. develop a LfD approach based on the use of force information for peg-in-hole tasks [4]. The authors address the problem of learning from demonstration trajectories that depend on contact forces instead of depending on time.

Our aim is to adapt the proposed approaches to minimally invasive surgical procedures, e.g. suturing, using the dVRK. This work is an extension of the work presented in [5].

II. Experimental setup

Our experimental setup is composed by a dVRK full surgical system with the open controllers (https://github.com/jhu-dvrk) commanded in teleoperation mode. The user performs several suturing procedures on a sponge phantom that is intended to act as dummy tissue. The interaction forces between the tool tip and the environment are measured by an ATI nano 17 Force/Torque sensor integrated in the phantom. During our experiments, the Patient Side Manipulators (PSMs) Cartesian state (positions and velocities) and the provided forces are collected at a sampling rate of 200 hz. In addition an external Kinect2 RGB camera is used to collect videos of the training and test procedures which are essential for the verification of the clustering procedure. ROS is used to collect all the data in a synchronized way. Data are downsampled at 60 Hz which are sufficient to describe human gesture.

In this work, GMM are used to cluster states of the suturing procedure. We fix the value of the number of clusters to $K = 4$. This value is established in an heuristic way by observing several clustering processes using different K. The aim is to identify four states of a surgical re-constructive procedure that are easy to be physically explained by the user; they are: *idle*, *interaction*, *free motion* and *thread traction*.

The underlaying physical properties of the encoded states are summarized in the table I. This criteria are used to perform the GMM parameters initialization (means and variances of the gaussians). Thus, we use as data to represent our states the vector $\boldsymbol{x} = (\boldsymbol{p}, \boldsymbol{v}, \boldsymbol{f})$ with $\boldsymbol{p} \in \mathbb{R}^6$ being the Cartesian pose of the manipulator, $\boldsymbol{v} \in \mathbb{R}^6$ its linear and angular velocity and $\boldsymbol{f} \in \mathbb{R}^6$ the force and torques exerted on its tool tip. These recorded data are then clustered using GMM.

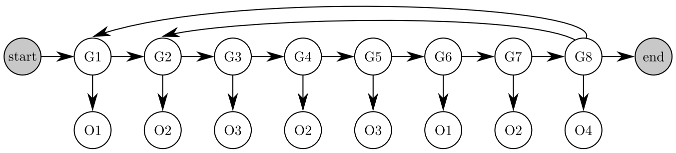

Fig. 1: HMM of suturing procedure. Gi= i-th gesture (hidden state), Oi = i-th cluster (observation).

To validate the learning procedure, once the GMM has been trained, we perform the evaluation on a test set acquired using the same setup. We use 30 suturing procedures as training set and evaluated the unsupervised classification procedure on 2 sequences.

The clustered sequences are then used as observations to train an HMM. More in details, we model the suturing procedure as a graph of hidden states having the following grammar:

- (G1) = idle
- (G2) = move to needle
- (G3) = grasp needle
- (G4) = move to suturing place
- (G5) = suturing
- (G6) = move other arm
- (G7) = thread scrolling
- (G8) = thread traction

Figure 1 represents the relationship, order, and flow of gestures during the execution of the suturing task. We choose to model the suture as a Bakis HMM, i.e. a left-to-right model suitable to model temporal processes. We add two outer loops to account for cyclical operations. Each hidden state has an unique output observation that is constituted by our previously clustered data (see table I):

- (O1) = idle
- (O2) = interaction
- (O3) = free motion
- (O4) = thread traction

As it can be seen in the result section II-A this allows the recognition of different gestures (hidden states) having same observations (clusters) but sequentially related. Both the training and the test set data are processed using the Statistical and Machine Learning Toolbox in MATLAB.

A. Results

An example of position trajectory recorded during a robotic surgical re-constructive procedure is shown in Fig. 2. By

TABLE I: Physical interpretation of the clustered states.

state	velocity	force
idle	low	low
interaction	low	high
free motion	high	low
thread traction	high	high

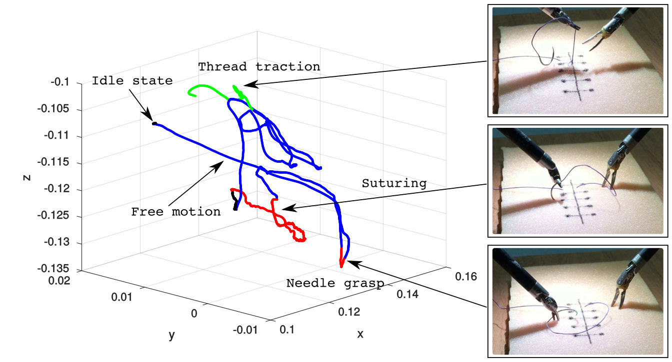

Fig. 2: Example of trajectory recorded during a robotic surgical re-constructive procedure. The graph on the left depicts tool tip Cartesian positions while the picture on the right the corresponding suturing states. Different colors correspond to different clusters.

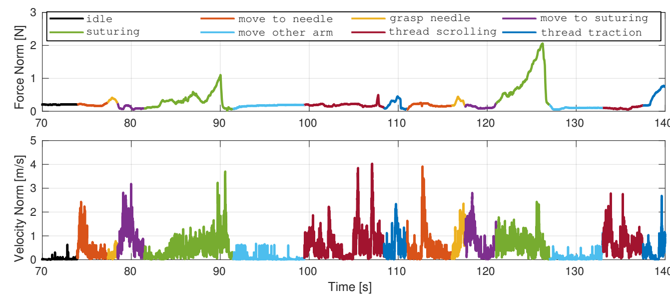

Fig. 3: HMM decoding of the robotic surgical re-constructive procedure shown on force norm signal. Different states are shown in different colors.

modeling our suturing procedure as composed of 8 hidden states and 4 observable quantities, i.e., the clusters (see Fig. 1), and training the model using their sequence, we are able to discriminate different gestures. The results are interpretable looking at the Fig. 3 in which two sequential stitches are considered. In our simple case, the HMM decoding done using the Viterbi algorithm had accuracy 96%. The approach revealed to accurately identify the hidden states and so to discern between different interaction purposes.

III. Conclusions and Discussion

This work demonstrates the use of GMM and HMM for task classification of a robotic surgical re-constructive procedure using kinematic and interaction force information. A GMM has been trained on a training set of expert demonstrations and its output has been used as observable states of an HMM. The GMM parameters have been chosen according to physical interpretation of suturing states. The method has demonstrated good accuracy in the task recognition of surgical re-constructive procedures. These results are encouraging in sight of the development of adaptive assistance methods for robotic aided surgical interventions. Further investigation are needed on the use of solely differential quantities (*i.e.* velocities) to improve spatial generalization. Future works will focus on the comparison of the proposed method with existing ones.

REFERENCES

[1] L. Zappella, B. Bjar, G. Hager, and R. Vidal, "Surgical gesture classification from video and kinematic data," *Medical Image Analysis*, vol. 17, no. 7, pp. 732 – 745, 2013.

[2] P. Berthet-Rayne, M. Power, H. King, and G. Z. Yang, "Hubot: A three state human-robot collaborative framework for bimanual surgical tasks based on learned models," in *IEEE International Conference on Robotics and Automation*, pp. 715–722, May 2016.

[3] F. Despinoy, D. Bouget, G. Forestier, C. Penet, N. Zemiti, P. Poignet, and P. Jannin, "Unsupervised Trajectory Segmentation for Surgical Gesture Recognition in Robotic Training," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 6, pp. 1280–1291, 2016.

[4] C. J. P. del Pulgar, J. Smisek, V. F. Muoz, and A. Schiele, "Using learning from demonstration to generate real-time guidance for haptic shared control," in *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3205–3210, Oct 2016.

[5] M. Selvaggio, G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Task classification of robotic surgical reconstructive procedures using force measurements," in *7th Joint workshop on new technologies for computer/robot assisted surgery*, Sep 2017.