Task Classification of Robotic Surgical Reconstructive Procedures using Force Measurements

M. Selvaggio, G.A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano

Department of Information Technology and Electrical Engineering, Università degli Studi di Napoli Federico II

INTRODUCTION

The development of surgical tasks and skills level classification methods and its combination 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 [1], task classification constitutes an essential step. It allows to assess surgeon skills and intentions in both training and real interventions. Our objective 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 sensors might be employed but they usually require fine parameters tuning and a huge programming effort. Hence, we propose to adopt the interaction force measurement in the learning process. In this paper, a force-based unsupervised segmentation approach of reconstructive surgical gestures is presented. In the past, similar approaches have been investigated: Zappella et al. have proposed several methods for automatic surgical gesture classification using video and kinematic data [2]; Pierre et al. have developed a human collaborative framework for bimanual surgical tasks based on learned model where they combine active constraints and learning from demonstration [3]; in the work of Despinoy et al. the operating gesture workflow has been taken into account, in order to provide more intuitive training as well as more accurate solutions for procedural knowledge assessment [4]; in the work of Perez-Del-Pulgar et al. the authors address the problem of learning from demonstration trajectories that depend on contact forces instead of depending solely on time [5]. In our work, we use force and kinematic data to train a Gaussian Mixture Model (GMM) in order to cluster subtasks during a robotic surgical reconstructive procedure. Comparing our approach with the manual annotations of the surgical gestures, an average matching score of 88.32% is observed for the fully automated gesture recognition process.

MATERIALS AND METHODS

Our experimental setup is composed by the da Vinci Research Kit commanded in teleoperation mode via open controllers[1]. The user teleoperates the Patient Side Manipulators (PSMs) using the Master Tool Manipulators (MTMs) by observing the scene through the endoscopic stereo camera. The complete robot dynamical model has been previously identified in order to estimate external

[1]`https://github.com/jhu-dvrk`

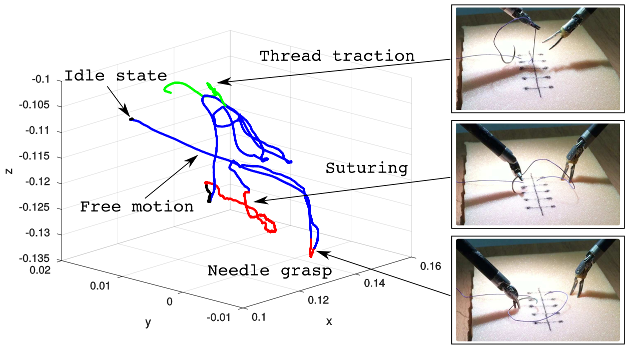

Fig. 1. Example of task classification during a robotic surgical reconstructive procedure. The graph on the left depicts the Patient Side Manipulator tool tip Cartesian trajectory segmented using our approach. The pictures on the right represent some relevant suturing states.

forces from motor current measurements [6]. This method has been used in combination with a recently developed force sensor integrated into the robot trocar [7] that is able to measure the interaction forces (components orthogonal to the instrument's axis) between the tool tip and the environment. The adopted force sensor guarantees a resolution of 0.01 N and a range of measurement of $[-10, 10]$ N that are suitable for measuring interaction forces between the surgical instrument and soft tissues throughout most of robot-aided surgical interventions. This is also the case of in-vivo suturing procedures where the forces can be of very low intensity ($||f|| \leq 1$ N) during the interaction with the tissues but also $||f|| \geq 5$ N, in some cases, when the thread traction is executed. During our experiments the PSMs Cartesian state (positions and velocities) and the measured 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 lately exploited to verify the accuracy of our method. ROS is employed to collect all the data in a synchronized way. During the demonstration phase we obtain a sequence of n elements of sensory information $\hat{x}_n = (x_1, x_2, \ldots, x_n)$. At each time step we encode a tuple $x_i = (p_i, v_i, f_i)$ with p_i being the Cartesian position of the manipulator, v_i its velocity and f_i the force exerted at the tool tip. These data are classified in an unsupervised way using GMM and Expectation-Maximization (EM) approach (see Fig. 1). A GMM is parametrized by two types of values: the mixture component weights and the component means and variances/covariances. For a multivariate GMM with K components, the k^{th} component has a mean μ_k and covariance matrix Σ_k. Given a tuple x the probability that this belongs to an encoded GMM is:

$$P(x) = \sum_{k=1}^{K} \phi_k \mathcal{N}(x \mid \mu_k, \Sigma_k) \tag{1}$$

where $\mathcal{N}$ denotes the classical multivariate normal distribution and the sum of ϕ_k is unitary:

$$\mathcal{N}(x|\mu_k, \Sigma_k) = \frac{\exp\left(-\frac{1}{2}(x-\mu_k)^{\mathrm{T}} \Sigma_k^{-1} (x-\mu_k)\right)}{\sqrt{(2\pi)^K \mid \Sigma_k \mid}} \tag{2}$$

$$\sum_{k=1}^{K} \phi_k = 1 \tag{3}$$

The log-likelihood function of a GMM can be written as follows:

$$L(\hat{x}_n) = \sum_{i=1}^{n} \log(P(x_i)) \tag{4}$$

The learning objective is to find a set of GMM parameters that maximizes Eq. (4). To this end, the EM algorithm iteratively maximizes the likelihood of a statistical model given the training sequence $\hat{x}_n$ and a predefined number of Gaussians. In this work, we have chosen this value to be K=4 in order to identify four states of a surgical reconstructive procedure. This value has been established in a heuristic way by observing several surgical reconstructive interventions performed by expert surgeons. We have used 30 suturing procedures as training set and evaluated the unsupervised classification procedure using 2 sequences. Both the training and the test data have been *offline* processed using the Statistical and Machine Learning Toolbox in MATLAB.

RESULTS

In order to validate our clustering procedure we trained the GMM and performed the evaluation using test set data. Demonstration and test phases consisted of suturing procedures conducted on a sponge phantom intended to act as dummy tissue. The result of a classification test is shown in Fig. 2. Here, only the time history of the measured force norm is reported, since it represents the most significant quantity for this evaluation. For the sake of clarity, only PSM1 data and states are shown but same results hold for PSM2. The four states we aimed at identifying were: *idle*, *interaction*, *free motion* and *thread traction*. To give an insightful explanation of the graph, the teleoperated robot is, at the beginning, in the *idle* configuration, then starts to move to perform its first action, *i.e.* needle grasping. The *interaction* state identified between 75 and 80 s is due to the contact occurred with the tissue while grasping the needle. Then, the operator moves and the next contact is detected while the needle is passing the phantom between 82 and 92 s. Successively, a new *idle* state is identified while the needle is regrasped by the PSM2. A *free motion* state is identified during the process of thread scrolling performed in alternation with the PSM2. Finally, the *thread traction* state concludes the suturing procedure. Our method allows to classify

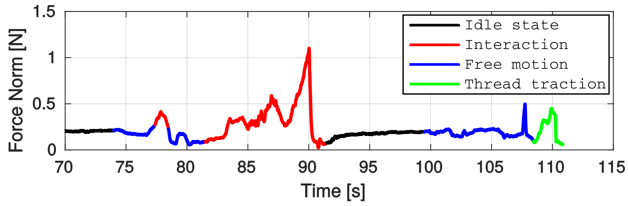

Fig. 2. Force norm variation during a reconstructive procedure used as test set. Different states are classified and shown using different colors.

the correct sequence of states with 88.32% of accuracy during the fully automated gesture recognition process. This result has been calculated by comparing the obtained results to manually annotated data. The annotation phase have exploited the above mentioned recorded videos.

CONCLUSIONS AND DISCUSSION

This paper demonstrates that is feasible to use interaction force information to reliably classify the states of a robotic surgical reconstructive procedure. A GMM has been trained using a set of demonstrations performed by expert surgeons and has been used to cluster test sets. The presented method allows the fully automated gesture recognition with an accuracy of 88.32% with respect to manual annotations of the surgical gestures. These results are encouraging in sight of the development of adaptive assistance strategies for robotic aided surgical interventions. However, some limitations of the method have been identified. Indeed, the use of position information does not allow spatial generalization. Moreover, with the adoption of GMM and force measurements, it is difficult to distinguish between different contact states, *e.g.* needle grasping and suturing. As future works, further investigations on the use of solely differential quantities (velocities), and on the exploitation of tasks sequence information are needed.

REFERENCES

[1] M. Selvaggio, G. Notomista, F. Chen, B. Gao, F. Trapani, and D. Caldwell, "Enhancing bilateral teleoperation using camera-based online virtual fixtures generation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2016, pp. 1483–1488.

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

[3] 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*, May 2016, pp. 715–722.

[4] 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.

[5] C. J. Perez-Del-Pulgar, J. Smisek, V. F. Munoz, 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*, Oct 2016, pp. 3205–3210.

[6] G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Modelling and identification of the da Vinci research kit robotic arms," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[7] G. A. Fontanelli, L. Buonocore, F. Ficuciello, L. Villani, and B. Siciliano, "A novel force sensing integrated into the trocar for minimally invasive robotic surgery," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*.