2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Macau, China, November 4-8, 2019

Vision-based Virtual Fixtures Generation for Robotic-Assisted Polyp Dissection Procedures

Rocco Moccia, Mario Selvaggio, Luigi Villani, Bruno Siciliano, Fanny Ficuciello[1]

Abstract— **Polyp dissection requires very accurate detection of the region of interest and high-precision cutting with adequate safety margins. Robot-assisted polyp dissection is a solution to accomplish high-quality intervention. This paper proposes a method to constrain the robot to follow an accurate dissection path based on Virtual Fixtures (VF). The VFs are created via specific control points obtained directly from images of the surgical scene and are updated by the vision algorithm. The VF constraints can autonomously adapt themselves to environment changing during the surgical intervention. The entire pipeline is validated through experiments on the da Vinci Research Kit (dVRK) robot.**

I. INTRODUCTION

Nowadays, colorectal cancer (CRC) is one of the major health problems. The majority of CRCs arises from adenomas or "polyps" growths on the inner surface of the colon. Endoscopic detection and removal of colorectal polyps significantly reduces the incidence and mortality of CRC, justifying the development of efficient polyp dissection procedures, which require precise movements, high dexterity and enhanced surgeon's skills for region identification and accurate path definition [1]. Particularly, sessile polyps are flat serrated growths adhered to colon surface, difficult to be detected and removed.

Since its introduction, Minimally Invasive Robotic Surgery (MIRS) completely changes surgical procedures, improving surgeon's technical skills, especially in tasks executed at limits of human capabilities. The da Vinci robotic system (Intuitive Surgical Inc., Sunnyvale, CA) is the most widely used robotic system for robot-assisted laparoscopic procedures. This robotic system provides the surgeon with improved dexterity, introducing tremor filtering, motion scaling, and stereoscopic vision. The benefits and safety of MIRS in gastric surgery have been evaluated, showing the enhancement of the quality in surgical procedures [2].

In the case of robotic polypectomy, the surgical operation is performed following precise steps:

- colonscopy for polyp detection;
- safety margins definition around the polyp;
- path planning for cutting execution.

Traditionally, the polyp detection is performed by the surgeon based only on his/her experience in the identification of specific surgical features (colours and textures), allowing the definition of the region of intervention. Moreover, the safety margins' definition is executed by the surgeon that marks selected points in telemanipulation modality by producing cautery spots around the polyp. Finally, the surgeon performs a first cutting operation considering the defined margins, then he/she lifts the surface of the polyp to execute another cutting operation on the underlying tissue while keeping the focus on predefined margins. Furthermore, change in polyp's shape and rigid displacements, due to patient's movements during the task, may compromise the correct execution of the procedure, modifying the predetermined safety margins.

In this scenario, images are the main feedback the surgeon can use to correctly operate in the surgical site. Therefore, in view of the increasingly autonomy in surgical robotics, vision-based techniques play an important role and can arise by extending computer vision algorithms to surgical scenarios [3]. Moreover, a large number of surgical tasks could benefits from the application of advanced shared-control techniques [4]. Particularly, Virtual Fixtures are commonly recognized as a powerful method to improve surgeon's performances, increasing accuracy and precision [5].

A. Related Works

Related works can be divided in two main topics: (*i*) vision algorithms for polyp detection and (*ii*) Virtual Fixtures techniques applied to surgical robotics.

1) Vision algorithms: Automatic polyp detection is a hot research topic. Bernal et al. [6] compared the performances of different polyp detection methods. Some detection methods exploit classic image processing techniques to obtain polyp boundaries. Hwang [7] used speeded up robust features (SURF) and quantized them with K-means clustering to represent the images as a histogram of visual words. The features are then, classified with SVM classifier. Zhu et al. [8] developed a computer-aided detection of colonic polyps based on polyp curvature estimation. Other authors exploit texture and color information as region descriptors in the image. Karkanis et al. [9] proposed an approach based on wavelet decomposition, while Hwang and Celebi [10] used watershed segmentation with initial markers selected using Gabor texture features and K-means clustering. Recent developments in deep learning, e.g., the use of convolutional neural networks (CNNs) have made significant advances in this field. In most of all, the differences among the methods are based on the selection of the specific network architecture or on the data-set for training. Particularly, Riberio et al. [11]

[1]The authors are with the Department of Electrical Engineering and Information Technology, Università degli Studi di Napoli Federico II, 80125 Napoli, Italy. Corresponding author's email: rocco.moccia@unina.it. This research has been partially funded by the EC Seventh Framework Programme (FP7) within RoDyMan project 320992, partially within the Italian National Project BARTOLO (POR FESR 2014-2020), partially by the IEEE RAS Techincal Committee on Haptics under the "Innovation in haptics" research programme.

used CNNs for automated classification of colonic mucosa for colon polyp staging.

Remarkably, none of these methods are currently used in clinic procedures. There are several reasons behind this, for example: some of them are not suitable for real-time constraints; some are built on theoretical model of the polyp; CNNs performance is strongly related to the quality and amount of realistic images used to train the network; finally, the clinical environment, with the presence of smoke, blood and specular highlights, makes the detection process a complicated task. For these reasons, more feasible and reliable methods for polyp detection needs to be developed and implemented to ensure the correct execution of polyp resection pipeline.

2) Virtual Fixtures: VFs can be classified into two categories: forbidden-region virtual fixtures, suitable for simulating barriers around forbidden regions, and guidance virtual fixtures, showing attractive behavior pulling the robot end-effector towards a desired path.

The first author to have introduced VFs was Rosenberg in his work [12]. Since its introduction, these shared control techniques have had great success in surgical applications [5]. One of the major obstacles in vision-based virtual fixtures generation in surgical procedures is the ability to adapt the VF geometry to a changing environment. More recently, VF generation is supported by the introduction of vision-based techniques. Bettini et al. [13] proposed a VF application in vitreoretinal eye surgery, using computer vision for providing a reference trajectory to the virtual fixtures control algorithm. Rydén et al. [14] showed a method for creating forbidden region virtual fixtures protecting an object from unwanted contact using point cloud streamed by an RGB-D camera. Yamamoto et al.[15] developed an interface for teleoperated RMIS providing vision-based forbidden-region virtual fixtures (FRVF) and augmented visual feedback. Very few papers make a significant consideration of adaptive VFs, where the constraint geometry autonomously moves as a result of environmental changes.

B. Contribution

This paper proposes a vision-based method for robot-aided polyp dissection using the da Vinci Research Kit (dVRK) robot, which is a research version of the da Vinci robotic system [16]. The method exploits basic computer vision concepts, simplified setup with colored object and impedance control to enforce guidance VF constraints. The VF path is adapted to the change in polyp's shape and environment displacement, that may occur during the dissection procedure, by updating new control points from the vision algorithm. The goal of the work is to define a functioning pipeline to assist the surgeon in the dissection task, by enhancing the quality of intervention.

An experimental setup recreates patient's anatomy using phantoms and the proposed approach includes vision algorithms for detection and segmentation of the polyp and for path planning of the cutting task. Then, a VF, i.e. a constraint that restricts the motion of the robot's tip along the path through a haptic guidance force rendered to the surgeon, is generated.

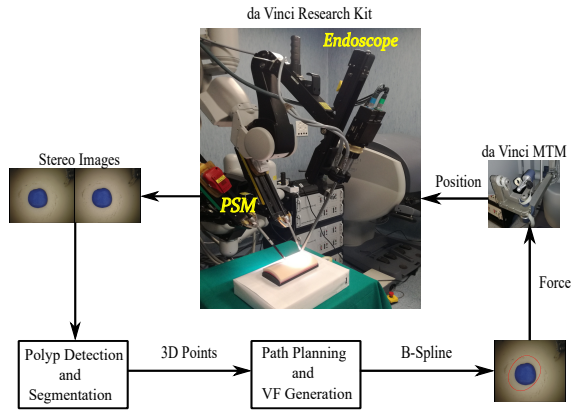


Fig. 1. Overview of the system. The approach starts from stereo endoscopic images processing, including detection and segmentation of the region of interest, and leads to the definition of accurate points needed in path planning and VF generation, executed through haptic guidance forces rendered to the user.

The work is a natural continuation of [17] and [18], towards fully autonomous surgical interventions. The proposed pipeline is articulated as follows:

1) Pre-operative calibration;
2) Stereo endoscopic images acquisition, with the dVRK endoscope in fixed position;
3) Detection and segmentation of the polyp, allowing the computation of control points, defining the polyp contours adjusted with a safety margin;
4) Point 3D reconstruction and path planning;
5) VF generation on the pre-planned path, by rendering a force to the user when the the robot exceeds the path.

II. SYSTEM DESCRIPTION

Figure 1 shows the overview of the system, composed by the dVRK robot and an experimental setup intended to replicate a surgical scene, adopting a plastic phantom to reproduce patient's skin and a blue object representing the polyp.

A. dVRK Robot

The dVRK robot is used in teleoperation mode, with the two Patient Side Manipulators (PSMs) commanded by two Master Tool Manipulators (MTMs), using the open controller developed by [16]. One of the MTMs is controlled through an impedance controller, which requires the measurement of external forces to the user. Considering an n-degree-of-freedom manipulator and a task space vector $x \in \mathbb{R}^r$ with $r \leq n$, the impedance dynamics is achieved through control:

$$M\ddot{\tilde{x}} + D\dot{\tilde{x}} = f_h + f_{VF}(\cdot), \tag{1}$$

where $\tilde{x} = x_d - x$, with x_d as the desired value of the robot task space variable, $M \in \mathbb{R}^{r\times r}$ and $D \in \mathbb{R}^{r\times r}$ are the inertia and damping matrices, designed to be fixed, diagonal and positive definite, $f_h \in \mathbb{R}^r$ is the external forces applied by

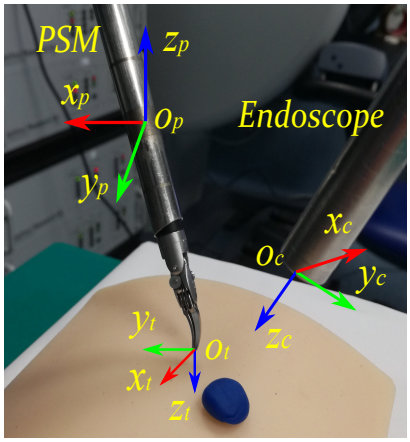


Fig. 2. Reference frames definition. $\mathcal{F}_c : (O_c - x_c y_c z_c) =$ endoscope reference frame; $\mathcal{F}_p : (O_p - x_p y_p z_p) =$ inertial reference frame; $\mathcal{F}_t : (O_t - x_t y_t z_t) =$ tool reference frame.

the user and $f_{VF}(\cdot)$ is the additional force due to the VF. This dynamics is obtained by setting the torque control input $\tau \in \mathbb{R}^n$ of the master robot as

$$\tau = B(q)v + N(q,\dot{q}) + J^T(q) f_h, \tag{2}$$

$$v = J_A^{-1}(q) M^{-1} (M\ddot{x}_d + D\dot{\tilde{x}} - M\dot{J}_A(q,\dot{q})\dot{q} - f_{h,A}), \tag{3}$$

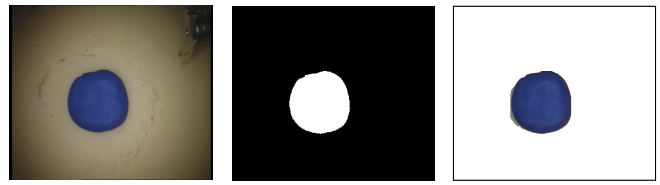
where $B(q) \in \mathbb{R}^{r \times n}$ is the joint space inertia matrix, $J(q)$, $J_A(q) \in \mathbb{R}^{r \times n}$ are the geometric and the analytic Jacobians respectively, and

$$N(q,\dot{q}) = C(q,\dot{q})\dot{q} + g(q) + h(q,\dot{q}) \tag{4}$$

are terms for Coriolis and centrifugal contributions $C(q,\dot{q})\dot{q}$, gravity $g(q)$, friction and disturbance torques $h(q,\dot{q})$ [19] and [17]. The force estimation is performed by resorting to a nonlinear observer [17], [20], and [21]. The system includes an endoscopic camera manipulator (ECM), consisting in a stereo camera with 5 mm baseline. The dVRK dynamic model was computed and identified in [22].

B. Calibration and Reference Frame Definition

The proposed method requires pre-operative calibration. Figure 2 shows the reference frames used in the work. Each PSM is a 7-Degrees-of-Freedom (DoFs) actuated arm moving around a fixed Remote Center of Motion (RCM). An inertial reference frame $\mathcal{F}_p : (O_p - x_p y_b z_p)$ is considered, with the origin positioned in the PSM's RCM. The generated VF is expressed in this inertial reference frame. The origin of the reference frame $\mathcal{F}_t : (O_t - x_t y_t z_t)$ is positioned in the PSM tool tip. The current position of the tool tip in Cartesian space, computed through the direct kinematics of the dVRK, provides the coordinates of point O_t in the frame $\mathcal{F}_p$. A Zhang's stereo camera calibration is performed [23] to estimate the transformation between the two endoscopic cameras and to define the camera reference frame $\mathcal{F}_c : (O_c - x_c y_c z_c)$. Then, by positioning the tool tip in ten different points, the transformation T_c^b between $\mathcal{F}_p$ and $\mathcal{F}_c$ is computed adopting an absolute orientation formulation [24]. It is important to underline that this kind of calibration still remains an open problem in robotics, and this solution ensures an error lower that 1 cm.



(a) (b) (c)

Fig. 3. Segmentation method: (a) Original frame; (b) Binary mask; (c) Segmented Image.

C. Vision Algorithm: Segmentation and 3D Reconstruction

The system takes stereo endoscopic images as input and a pre-process step is performed. A watershed transformation is performed on the left gray-scale image. The transformation operates on the image like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges. This method allows defining the object region in the image, that is later used as seed point defining a bounding box around it. Relying on the defined region, the widespread *GrabCut* segmentation method [25] is applied. This technique is based on graph cuts, addressing the visual segmentation task as a energy minimization problem, based on foreground (polyp) and background models. Giving the input image I, $\alpha = (\alpha)_{i=1}^{P}$ is the set of unknown binary label of the pixels ($\alpha_i = 0$ for the background pixels, $\alpha_i = 1$ for foreground), with P as the number of pixel. The algorithm estimates the values α minimizing the energy function:

$$E(\alpha) = E_{data}(\alpha) + \gamma E_{smooth}(\alpha) \tag{5}$$

with $E_{data}(\alpha) = \sum_i U_i(\alpha_i)$ and $U_i(\alpha_i)$ defines the probability for a pixel to belong to the foreground or background [26]. As in [26], a modifications of this algorithm is adopted, solving the minimization problem by a graph cuts minimization algorithm and defining the statistical models for the data energy function as a Gaussian Mixture Models based on color distribution. The background and foreground layers are defined by a bounding box around the seed point created by the watershed transformation around the object, allowing isolating the object inside image. Figure 3 shows the results of segmentation method.

At this point, the homographic transformation H between the original left and right images is computed, using Scale-Invariant Feature Transform (SIFT) for features detection [27] and Fast Library for Approximate Nearest Neighbors (FLANN) for matching [28]. The left segmented image allows detecting contour's points of the object on the left image plane, computing the Hull convex approximation. The corresponding contour's points on right image plane are computed applying the previously computed H transformation. Finally, the 3D position of the contour's points is

reconstructed using triangulation method with direct linear transform, expressed in the camera frame $\mathcal{F}_c$.

D. Path Planning and VF Generation

The contour's points are mapped in the inertial reference frame $\mathcal{F}_p$ through the transformation T_c^b. In particular, the points' coordinates are adjusted with a safety margin that allows performing the cutting in safe conditions. Once these 3D points are determined, they are used to build the VF geometry. As in [17], the path for cutting is formulated through a parametric curve. In this case, a closed B-Spline curve is adopted, defined in its 1-dimensional form by:

$$\Gamma(s) = \sum_{i=0}^{n} N_{i,k}(s)\, p_i, \tag{6}$$

where $\Gamma(s)$ denotes the curve, k its order, $s \in [0,1]$ is the normalized curve parameter and $N_{i,k}$ are its basis function. The 3D points identified by our vision algorithm are used as controls points of the curve (p_i). Then, the VF is defined as the surface created by sweeping the polyp contours, determined by the B-Spline curve, along the axis that is perpendicular to the tissue's plane, which is set as coincident to the axis perpendicular to the camera. A simple constraint enforcement method is selected, consisting in the application of a spring-damper like force imposed on the path:

$$f_{VF} = K_p(x_d - x) - K_d\dot{x}, \tag{7}$$

where K_p and K_d are properly designed diagonal and positive definite matrices and x_d is the set point belonging to the constraint geometry having minimum distance from the current position x.

For the defined B-Spline, the Newton-Raphson (NR) method is used to find the nearest point on the curve x_d starting from the current robot position. This represents a general method for finding the extrema of a given function in an iterative manner. As explained in [17], the desired point on the curve is the minimum distance point $x_d = \Gamma(\bar{s})$, obtained by finding the correspondent spline parameter $\bar{s}$. The customary NR update law is:

$$s_{k+1} = s_k + \frac{\delta(x, s_k)}{\delta'(x, s_k)}, \tag{8}$$

where $\delta(x,s) : \mathbb{R}^r \times \mathbb{R} \to \mathbb{R}$ is the distance function the point x and the spline $\Gamma(s)$, defined by:

$$\delta(x,s) = \sqrt{(x - \Gamma(s))^T (x - \Gamma(s))} \tag{9}$$

$\delta'(x, s_k)$ is the derivative at s_k of $\delta(x,s)$ with respect to the curve parameter s.

Finally, the VF is mapped to the MTM robot, considering the z-axis value defined by the tissue's normal and displaying the attractive force f through impedance control of the MTM, realized thanks to the dVRK dynamic model with the parameters identified in [22].

(a) (b) (c)

Fig. 4. (a) VF path (red line); (b) The polyp's shape and position during the task changes due to dVRK tool interaction; (C) VF path (red line) adapted after the change in polyp position.

III. EXPERIMENTAL VALIDATION

The proposed vision-based assistive control is evaluated by executing multiple dissection tasks. In all the sessions, a blue object with 1 cm diameter is considered to reproduce the polyp attached to a silicon rubber phantom commonly used by the surgeons for training. The safety margin is set on 1 cm from the polyp's edge, as usual in surgical dissection operations.

During the task, the user performs the dissection procedure by commanding the PSM of the dVRK through the MTM manipulator and following the cutting path defined by the vision algorithm. The dissection task is repeated twice, in the first session the VF is activated while in the second session it is deactivated. In the first session, whenever the user exceeds the defined path, the attractive force pulls him/her towards the planned path. During the second session, without the VF generation, the cutting path is simply defined by a circle, with 1.5 cm diameter, centred on the polyp. The user can perform the dissection task, trying to follow the imposed circle, projected on the endoscopic image, without any haptic constraint.

To evaluate the accuracy and precision in the dissection task, the mean absolute error is computed between the current position of the PSM and the desired position defined as the minimum distance point on the B-Spline (as defined in Section II-D). To evaluate the performance of the 3D reconstruction, the current position of the PSM tool's tip is recorded through the direct kinematics of the dVRK and the vision algorithm proposed, positioning the tip in 50 different workspace positions. Then, a comparison is made between the means of the z-axis values through a statistical unpaired t-test with a significance level $\alpha = 0.05$.

The B-Splines projected on the left camera image are shown in Fig. 4. It is possible to notice that the vision algorithm adapts the B-Spline, at frequency of 25 Hz, after a change in polyp position and shape, that occurred after the interaction with the dVRK tool during the task. This allows keeping the VF constraint on the specific path during the entire dissection procedure, autonomously moving as a results of environmental changes.

A. Results

Figure 5 shows the PSM position and B-Spline during a dissection task in the xy plane of the PSM reference frame; Figure 6 shows the PSM position and the circle during a

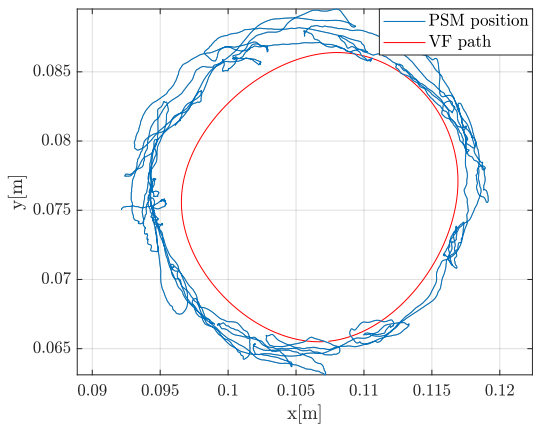


Fig. 5. VF path (red line) and PSM position during the dissection task.

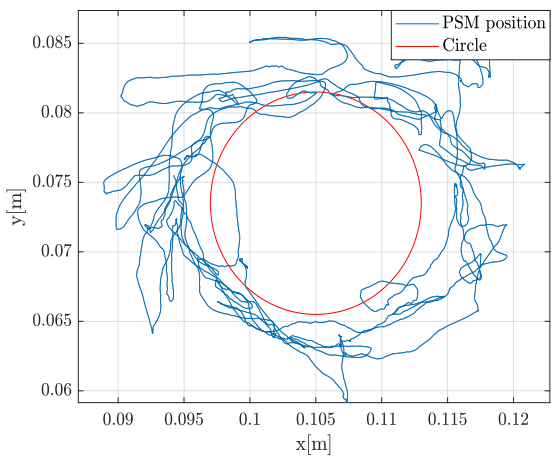


Fig. 6. PSM position and the guided circle (red line).

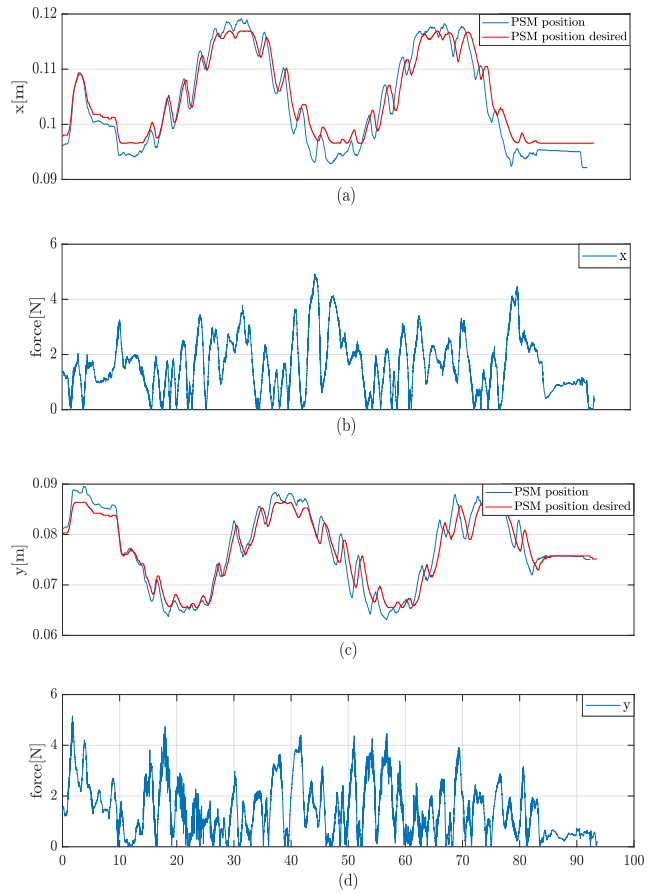


Fig. 7. Dissection experiment. Duration: 100 seconds. Time histories of: (a) PSM position and the desired position (minimum distance point on the B-Spline) along the x-axis; (b) The related estimated force norm along x-axis; (c) PSM position and the desired position (minimum distance point on the B-Spline) along the y-axis; (d) The related estimated force norm along y-axis.

dissection task in free motion. The dissection tasks with VF generation presents a mean absolute error equal to 2.1 mm along x-axis and 1.7 mm along y-axis. Conversely, the dissection tasks without VF generation has a mean absolute error equal to 13 mm along x-axis and 31.6 mm along y-axis. The results of the t-test for the 3D reconstruction evaluation proved to be statistically significant with a probability level of $p = 0.9831$ ($\alpha = 0.05$).

Figure 7(a) shows the PSM position and the desired position (minimum distance point on the B-Spline) along x-axis. The related estimated force norm along x-axis is represented in Figure 7(b). Moreover, Figure 7(c) reports the corresponding PSM position and desired position along y-axis, and, finally, Figure 7(d) shows the estimated force norm along y-axis. Figure 8 contains the estimated haptic guidance forces rendered to the user through the master side (MTM) during the tasks. Figure 7 and 8 are referred to the same dissection task.

B. Discussion

As it is possible to notice in Fig. 5, the user follows the determined path during the procedure, while in Fig. 6 a more irregular path is performed by the user in free motion without the VF assistance. The mean absolute error values present a significant reduction ($\sim$ 10 mm along x-axis and $\sim$ 30 mm along y-axis) in the VF-based dissection task. This suggests that the introduction of the VF improves the accuracy of the procedure, helping the user performing a more precise cutting path. It may be noted that in Fig. 8 the maximum reached force is about 4 N, allowing the surgeon to perform the task accurately, without experiencing excessive force. The results of 3D reconstruction t-test show the absence of statistically significant differences between z-axis values computed through the vision algorithm and the robot's kinematics. In Fig. 7, the peaks in force norms occur when the PSM position distances the desired position. Thus, the user continuously experiences guidance forces when the robot position is not on the VF path.

IV. CONCLUSION AND FUTURE WORKS

This paper introduces an experimentally validated vision-based method for VF generation in minimally invasive robotic surgery polyp dissection tasks. The vision algorithm allows creating a constraint path for cutting, through VF

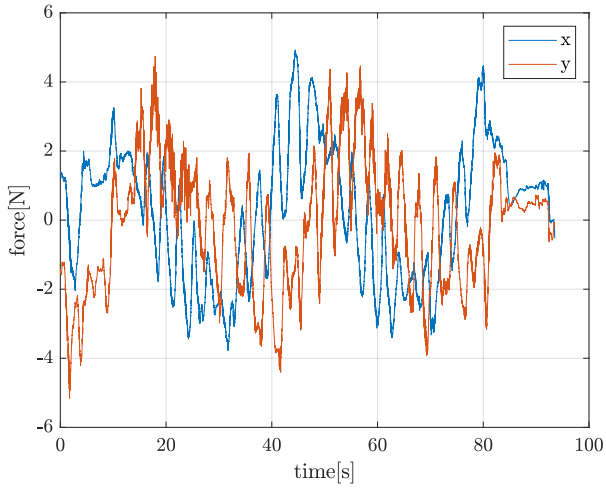


Fig. 8. Estimated Haptic guidance forces displayed to the user through the MTM.

generation. The VF path is updated by the vision algorithm, allowing considering environment displacement during the dissection task. The proposed strategies are evaluated through multiple dissection experiments on dVRK, showing good results in improving accuracy and precision of intervention and thus suggesting the feasibility of the proposed pipeline.

The goal of future works is to consider more advanced computer vision techniques, allowing to extend the proposed method to realistic surgical scenarios with regards to polyp detection, to integrate tissue's deformation and to enforce the method also in presence of occlusions. Also, an accurate study on medical procedures will be considered for a correct definition of safety margins for cutting. Finally, a validation involving surgeons could be conducted to proof the effectiveness of the pipeline.

REFERENCES

[1] M. Ferlitsch, A. Moss, and A. Arezzo et al, "Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESG) Clinical Guideline," *Endoscopy*, vol. 49, no. 3, pp. 270–297, 2017.

[2] A. Patriti, G. Ceccarelli, R. Bellochi, A. Bartoli, A. Spaziani, L. Di Zitti, and L. Casciola, "Robot-assisted laparoscopic total and partial gastric resection with D2 lymph node dissetion for adenocarcinoma," *Surgical Endoscopy*, vol. 22, no. 12, pp. 2753–2760, 2008.

[3] D. Stoyanov, "Surgical Vision," *Annals of Biomedical Engineering*, vol. 40, no. 2, pp. 332–345, 2012.

[4] S. A. Bowyer, B. L. Davies, and F. R. Baena, "Active constraints/Virtual fixtures: A survay," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 138–157, 2014.

[5] N. Enayati, E. De Momi, and G. Ferrigno, "Haptics in robot-assisted surgery: Challenges and benefits," *IEEE Reviews in Biomedical Engineering*, vol. 9, pp. 49–65, 2016.

[6] J. Bernal and D. Stoyanov et al., "Comparative validation of polyp detection methods in video colonoscopy: Results from the MICCAI 2015 endoscopic vision challenge," *IEEE Transactions On Medical Imaging*, vol. 36, no. 6, pp. 1231–1249, 2017.

[7] S. Hwang, "Bag-of-visual-words approach to abnormal image detection in wireless capsule endoscopy videos," in *Advances in Visual Computing*. Germany: Springer: Berlin/Heidelberg, 2011, pp. 320–327.

[8] H. Zhu, Y. Fan, and Z. Liang, "Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps," *Proc. International Workshop on Computational Challenges and Clinical Opportunities in Virtual Colonoscopy and Abdominal Imaging*, pp. 9–14, 2010.

[9] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, and M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 3, pp. 141–152, 2003.

[10] S. Hwang and M. E. Celebi, "Polyp detection in wireless capsule endoscopy videos based on image segmentation and geometric feature," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 678–681, 2010.

[11] E. Riberio, A. Uhl, and M. Hafner, "Colonic polyp classification with convolutional neural networks," *Proc. IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 253–258, 2016.

[12] L. B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation," *Proc. IEEE Virtual Reality Annual International Symposium*, pp. 76–82, 1993.

[13] A. Bettini, P. Marayong, S. Lang, A. M. Okamura, and G. D. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Transactions on Robotics*, vol. 20, no. 6, pp. 953–966, 2004.

[14] F. Rydén and H. J. Chizeck, "Forbidden-region virtual fixtures from streaming point clouds: Remotely touching and protecting a beating heart," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 3308–3313.

[15] T. Yamamoto, N. Abolhassani, S. Jung, A. M. Okamura, and T. N. Judkins, "Augmented reality and haptic interfaces for robot-assisted surgery," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 8, no. 1, pp. 45–56, 2012.

[16] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da Vinci®Surgical System," *Proc. IEEE International Conference on Robotics and Automation*, pp. 6434–6439, 2014.

[17] M. Selvaggio, G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Passive Virtual Fixtures adaptation in minimally invasive robotic surgery," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3129–3136, 2018.

[18] R. Moccia, M. Selvaggio, B. Siciliano, A. Arezzo, and F. Ficuciello, "Vision-based Virtual Fixtures generation for MIRS dissection tasks," *9th Joint Workshop on New Technologies for Computer/Robot Assisted Surgery*, 2019.

[19] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*. New York, NY, USA: Springer-Verlag, 2009.

[20] A. De Luca and R. Mattone, "Actuator failure detection and isolation using generalized momenta," *Proc. IEEE International Conference on Robotics and Automation*, pp. 634–639, 2003.

[21] ——, "Sensorless robot collision detection and hybrid force/motion control," *Proc. IEEE International Conference on Robotics and Automation*, pp. 999–1004, 2005.

[22] G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Modelling and identification of the da Vinci Research Kit robotic arms," *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1464–1469, 2017.

[23] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," *Proc. 7th IEEE International Conference on Computer Vision*, vol. 1, pp. 666–673, 1999.

[24] B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, vol. 4, no. 4, pp. 629–642, 1987.

[25] C. Rother, V. Kolmogorov, and A. Blake, "Grubcut: Interactive foreground extraction using graph cuts," *ACM Transactions on Graphics*, vol. 4, pp. 309–314, 2004.

[26] A. Petit, V. Lippiello, and B. Siciliano, "Real-time tracking of 3d elastic objects with RGB-D sensor," *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3914–3921, 2015.

[27] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[28] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," *Proc. International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 2, pp. 331–340, 2009.