

Vision-Based Dynamic Virtual Fixtures for Tools Collision Avoidance in Robotic Surgery

Rocco Moccia[✉], *Student Member, IEEE*, Cristina Iacono[✉], Bruno Siciliano[✉], *Fellow, IEEE*,
and Fanny Ficuciello[✉], *Senior Member, IEEE*

Abstract—In robot-aided surgery, during the execution of typical bimanual procedures such as dissection, surgical tools can collide and create serious damage to the robot or tissues. The da Vinci robot is one of the most advanced and certainly the most widespread robotic system dedicated to minimally invasive surgery. Although the procedures performed by da Vinci-like surgical robots are teleoperated, potential collisions between surgical tools are a very sensitive issue declared by surgeons. Shared control techniques based on Virtual Fixtures (VF) can be an effective way to help the surgeon prevent tools collision. This letter presents a surgical tools collision avoidance method that uses Forbidden Region Virtual Fixtures. Tool clashing is avoided by rendering a repulsive force to the surgeon. To ensure the correct definition of the VF, a marker-less tool tracking method, using deep neural network architecture for tool segmentation, is adopted. The use of direct kinematics for tools collision avoidance is affected by tools position error introduced by robot component elasticity during tools interaction with the environment. On the other hand, kinematics information can help in case of occlusions of the camera. Therefore, this work proposes an Extended Kalman Filter (EKF) for pose estimation which ensures a more robust application of VF on the tool, coupling vision and kinematics information. The entire pipeline is tested in different tasks using the da Vinci Research Kit system.

Index Terms—Surgical robotics, collision avoidance, virtual fixtures, surgical tool tracking, haptic feedback.

I. INTRODUCTION

MINIMALLY Invasive Robotic Surgery (MIRS) has completely changed surgical procedures. Enhanced dexterity, ergonomics, motion scaling, and tremor filtering, are well-known advantages introduced with respect to classical laparoscopy. With the da Vinci robotic system (Intuitive Surgical Inc., Sunnyvale, CA) the surgeon performs tasks in teleoperation mode using only visual information of the surgical scene provided by a 3D stereo viewer. During the execution of a

surgical procedure, two or more tools can come dangerously close to each other. The surgeon has a very limited vision on the surgical site which reduces dexterity and increases the cognitive workload, making the task most difficult to be performed. The view the surgeon has could be insufficient to avoid collision, thus this issue can cause tools or tissues damage. Experienced surgeons develop strong capabilities to compensate for the lack of haptic information, recreating the perception of haptic feedback from visual cues of the surgical scene [1]. However, recent studies demonstrate different performances in MIRS procedures between experienced and novice surgeons, suggesting that haptic feedback effects performances differently based on the operator's level of experience with the robot [2]. Actually, haptic feedback could significantly affect the performances of novice surgeons, reducing training duration and improving the effectiveness of the procedures.

A large number of surgical tasks can benefit from the introduction of collision avoidance techniques. During robotic polypectomy, one surgical tool has to cut around the polyp while another tool keeps raised the surface of the polyp [3]. In this procedure, the surgeon performs a first cutting operation, then lifts the surface of the polyp and executes another cutting task. An automatic robotic assistance to avoid collision between the surgical tools can alleviate the surgeon workload during the execution of this task, and can allow the surgeon focusing on following the polyp margins. In procedures requiring tissues removal with the use of electrocautery, the direct coupling that occurs with a conductor, such as another tool, could burn non-targeted tissue [4]. Collision between surgical tools in MIRS can be avoided with the application of advanced shared control techniques. In particular, Virtual Fixtures can impose collision avoidance by rendering haptic cues to the surgeon. Forbidden Region Virtual Fixtures (FRVF) restrict the motion of the robot's tool tip through a repulsive force rendered to the surgeon. The da Vinci Research Kit (dVRK) is an open-source mechatronic system, constituted by the first-generation of the da Vinci robotic system equipped with electronics, firmware, and software developed on purpose to create an open control architecture. The dVRK allows testing new control methods and it is already used to test VF-based methods [3] [5]. Since dVRK robot joints are driven through cables that introduce elasticity, backlash and non-linear friction [6], tools position information obtained through direct kinematics is affected by errors and thus requires correction. Therefore, to ensure a correct application of the VF, a method for surgical tool tracking is strictly needed.

Manuscript received September 10, 2019; accepted January 7, 2020. Date of publication January 28, 2020; date of current version February 6, 2020. This letter was recommended for publication by Associate Editor E. De Momi and Editor P. Valdastrì upon evaluation of the reviewers' comments. This work was supported by EC Seventh Framework Programme (FP7) within the Italian National Projects BARTOLO under Grant POR FESR 2014-2020 and PROSCAN under Grant PON 2014-2020. (Rocco Moccia and Cristina Iacono are co-first authors.) (Corresponding author: Rocco Moccia.)

The authors are with the Department of Electrical Engineering and Information Technology, Università degli Studi di Napoli Federico II, 80125 Napoli, Italy (e-mail: rocco.moccia@unina.it; iaconocristina94@gmail.com; siciliano@ieee.org; fanny.ficuciello@unina.it).

This letter has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the authors.

Digital Object Identifier 10.1109/LRA.2020.2969941

A. Related Works

Rosenberg was the first author to introduce VF [7]. Since its introduction, shared control techniques have had great success in surgical applications and for collision and obstacle avoidance. Li *et al.* presented an on-line collision avoidance method for real-time interactive control of a surgical robot in geometrically complex environments, such as the sinus cavities [8]. Ren *et al.* [9] proposed dynamic active constraints using medical images. The system builds potential fields to reduce the contact force between the tool tips. Xia *et al.* [10] reduced the proportional gain in an admittance control law according to the distance between the tool tip and the nearest obstacle. This allowed the system to smoothly avoid collisions. Rydén *et al.* showed a method to create forbidden region virtual fixtures to protect an object from undesired contacts, using point cloud streamed by an RGB-D camera [11]. This method uses depth information and it is generally applicable only for collision avoidance on the tool tips. In [12], the authors proposed a vector-field-inequalities method to provide dynamic active-constraints for collision avoidance of any number of robots and moving objects sharing the same workspace. Banach *et al.* proposed a Forbidden Region Active Constraints strategy to avoid surgical tool clashing and, at the same time, the collision with patient anatomy using elastoplastic frictional force control model [4]. In this work, the current poses of the tools are tracked in real-time from the robot's kinematics of the dVRK. Thus, a non-compensated position error could compromise the effectiveness of the method.

In the literature, most of the tracking methods used to correct the surgical tool position error are realized using the sensors of the robotic system or using external sensors integration, but still obtaining limited accuracy. A significant improvement is introduced by image-based approaches, detecting the tool's position and orientation in the camera reference frame. In [13], the authors presented a survey about vision-based and marker-less surgical tool tracking. The works can be classified based on the segmentation and tracking methods [14], most of them exploiting Random Forest (RF) or convolutional neural networks (CNN) techniques. In [15], the authors combine a region-based segmentation technique with point-based pose estimation, using prior knowledge of the instrument shape through classification with a Random Forest (RF), besides temporal motion is incorporated with a Kalman filter. Du *et al.* [16] proposed a 2D tracker based on a Generalized Hough Transform using SIFT features which can both handle complex environmental changes and recover from tracking failure. They extended the work in [17], presenting a 2D pose estimation framework for articulated endoscopic surgical instruments, which involves a fully convolutional detection-regression network (FCN) and a multi-instrument parsing component.

B. Contribution

This letter proposes a surgical tool collision-avoidance method in MIRS. The goal is to improve safety in surgical procedures, enhancing especially novice surgeons abilities. The method is tested on the da Vinci Research Kit (dVRK). Forbidden Region Virtual Fixtures are used to avoid surgical tool clashing, by rendering a repulsive force to the surgeon which is

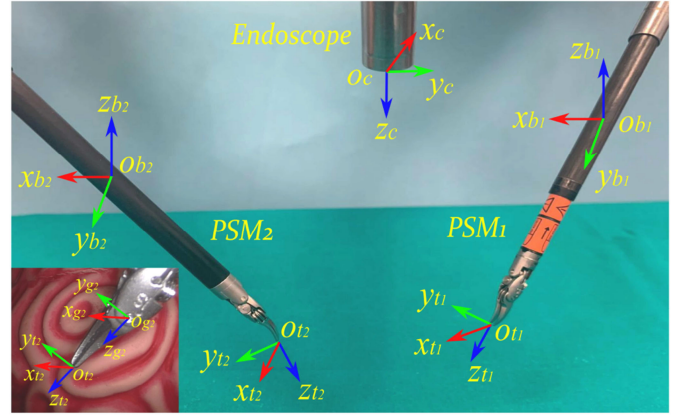


Fig. 1. Experimental setup and frames definition. Gripper frame: $\mathcal{F}_{g_2} : (O_{g_2} - x_{g_2} y_{g_2} z_{g_2})$ and Tool tip frame $\mathcal{F}_{t_2} : (O_{t_2} - x_{t_2} y_{t_2} z_{t_2})$.

inversely proportional to the distance between tools. The method includes a marker-less surgical tool tracking technique using an Extended Kalman Filter (EKF) that couple vision and kinematics information to enhance the robustness of VF application. Visual information allows overcoming the large position error, that occurs on the dVRK kinematics, especially when the surgical tools interact with the environment. While kinematics data reinforce the method in the presence of visual occlusions. To validate the method, an extensively study involving human subjects is conducted on two groups of surgeons, namely experts and novice surgeons, each group is constituted by 6 subjects. The goal is to demonstrate significant improvement in performances caused by the introduction of force cues. The pipeline of the method is articulated as follows:

- 1) Pre-operative calibration and stereo endoscopic images acquisition;
- 2) tool segmentation and tool tip pose estimation from vision algorithm;
- 3) kinematic and vision data fusion with EKF;
- 4) VF generation and force rendering.

II. SYSTEM DESCRIPTION

A. dVRK Robot

The dVRK robot is composed of two Patient Side Manipulators (PSMs) and an Endoscope Camera Manipulator (ECM) commanded by two Master Tool Manipulators (MTMs). Full control of the dVRK robotic arms is possible thanks to an open controller developed by [18]. To generate force cues, as in [3] the MTMs are controlled through an impedance controller. The surgical scene can be seen by the surgeon thanks to an endoscope, including a stereo camera with 5 mm baseline. In Fig. 1, the reference frames definition is shown. The base frame, $\mathcal{F}_{b_1} : (O_{b_1} - x_{b_1} y_{b_1} z_{b_1})$, is positioned at the PSM1 Remote Center of Motion (RCM). Likewise, $\mathcal{F}_{b_2} : (O_{b_2} - x_{b_2} y_{b_2} z_{b_2})$ is the base frame centred in the PSM2 RCM. All the measurements in this work will be expressed referring to the base frame $\mathcal{F}_{b_2}$ of the PSM2. The frames $\mathcal{F}_{g_1} : (O_{g_1} - x_{g_1} y_{g_1} z_{g_1})$ and $\mathcal{F}_{g_2} : (O_{g_2} - x_{g_2} y_{g_2} z_{g_2})$ are the grippers frames. The direct kinematics of the dVRK allows computing the current position

of the tools in the Cartesian space, providing the coordinates of O_{g_1} and O_{g_2} in $\mathcal{F}_{b_1}$ and $\mathcal{F}_{b_2}$ respectively. The reference frames $\mathcal{F}_{t_1} : (O_{t_1} - x_{t_1}y_{t_1}z_{t_1})$ and $\mathcal{F}_{t_2} : (O_{t_2} - x_{t_2}y_{t_2}z_{t_2})$ have their origins in the PSM1 and PSM2 tool tips, respectively. For each PSM, the frames $\mathcal{F}_g$ and $\mathcal{F}_t$ have the same orientation and the origin of $\mathcal{F}_t$ is translated of 1 cm along the z-axis of $\mathcal{F}_g$. As in [3], Zhang stereo camera calibration is performed to define the camera reference frame $\mathcal{F}_c : (O_c - x_cy_cz_c)$ and a hand-eye calibration is performed to find the transformation $T_c^{b_2}$ between $\mathcal{F}_{b_2}$ and $\mathcal{F}_c$. During the calibration process, the tool is placed in ten fixed positions and the transformation is computed adopting an absolute orientation formulation [19]. A hand-eye calibration is performed to find the transformation $T_{b_1}^{b_2}$ between the fixed frames of each robotic arm.

B. Tool Segmentation and 3D Reconstruction

The method directly uses laparoscopic images to track the surgical instrument. A deep learning solution for instrument semantic segmentation is employed. It is based on U-Net architecture, which is a fully convolutional neural network, composed by a contracting path to capture context and by an expanding path that enables precise localization [20]. The system adopts the U-Net modification proposed in [21], called TerausNet that uses pre-trained VGG16 networks as an encoder. The network is trained using the dataset provided for MICCAI 2017 Endoscopic Vision Sub-Challenge: Robotic Instrument Segmentation¹ consisting of 8×225 -frame sequences of high resolution stereo camera images acquired from a da Vinci Xi surgical system during several different porcine procedures, with 2 Hz frame rate. The output of the model is an image in which each pixel is the probability value of belonging to the instrument or background area. Then the binary segmentation is obtained, in which all the instrument pixel values are set as 255 and all the background pixel values are set as 0. The homographic transformation H between the original left and right images is computed, using Scale-Invariant Feature Transform (SIFT) for features detection and Fast Library for Approximate Nearest Neighbors (FLANN) for matching, as in [3]. To detect the tool tip on the image plane, the search area range is reduced by re-projecting the tip kinematic position on the image plane and by constructing a rectangle centered on the projected point. Then, the 3D position of the PSM2 tip, expressed in the camera frame $\mathcal{F}_c$, is reconstructed by using a triangulation method with direct linear transform. The tool orientation is computed solving PnP problem, which allows computing the orientation of the object from a set of n correspondences between 3D points and their 2D projections [22]. In this case, the line of symmetry of tool is computed, allowing the identification of four specific points on the line in the image plane and their correspondent 3D coordinates thanks to the knowledge of the tool's geometry. Finally, using transformation $T_c^{b_2}$, the tool tip position and orientation of PSM2 is found, expressed in the base frame $\mathcal{F}_{b_2}$. Fig. 2 shows the results of the segmentation method.

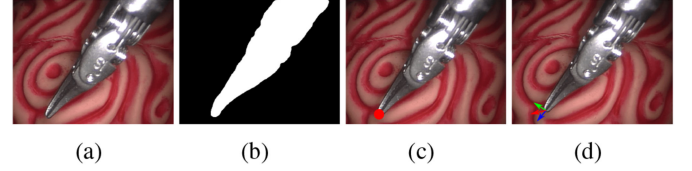


Fig. 2. Segmentation method: (a) Original frame; (b) Binary mask; (c) Point identification in the image plane; (d) Reference frame definition.

C. Surgical Tool Tracking

For the estimation and tracking of the instrument pose, the Extended Kalman Filter (EKF) is used. Kalman filtering allows combining visual information from the endoscope with the robot kinematics [23]. The entire formulation is referred to PSM2 and the subscript 2 is omitted in this subsection.

The filter provides an estimate of the tool tip pose $\zeta = [p_t, q_t]^T$, being p_t the true tool position, and $q_t = [\eta_t, \epsilon_t]^T$ its quaternion-based true orientation in the base frame $\mathcal{F}_b$. The prediction step provides a preliminary estimation of the instrument pose through the linear and angular velocities of the gripper provided by the manipulator kinematics. Then, the vision-based estimated pose is used in the filter correction step. The process dynamics for the state vector ζ is given by:

$$\begin{cases} \dot{p}_t = v_g + S(\omega_g) r_{gt} + n_p \\ \dot{q}_t = \frac{1}{2} \Omega(\omega_g) q_t + n_q \end{cases} \quad (1)$$

where $[v_g, \omega_g]^T$ are the linear and angular velocity of the gripper in $\mathcal{F}_b$, $S(\cdot)$ is the skew-symmetric operator, r_{gt} is the position vector of the tool tip respect to the gripper, $n = [n_p, n_q]^T \sim \mathcal{N}(0, N)$ is the process noise and

$$\Omega(\omega) = \begin{bmatrix} 0 & -\omega^T \\ \omega & S(\omega) \end{bmatrix}. \quad (2)$$

The error state vector is defined as $\tilde{\zeta} = [\tilde{p}, \delta\tilde{\theta}]^T$. The orientation error $\delta\tilde{\theta}$ is the 3×1 small-angle approximation vector of the quaternion orientation error. The vision algorithm computes the 3D pose of the tool tip, so the measurement model is given by:

$$y = \zeta + m \quad (3)$$

where $m \sim \mathcal{N}(0, M)$ is the measurement noise. Then:

$$F = \begin{bmatrix} S(\omega_g) & O_3 \\ O_3 & S(\omega_g) \end{bmatrix}; \quad H = \begin{bmatrix} I_3 & O_3 \\ O_3 & I_3 \end{bmatrix} \quad (4)$$

where F and H are respectively the control and measurement matrix used in the EKF implementation. The output of the EKF consists in the current pose of frame $\mathcal{F}_t$ of PSM2 with respect to the base frame $\mathcal{F}_b$.

D. Virtual Fixtures Generation

The collision avoidance between the two tools is ensured by the application of a FRVF. To this purpose, the VF is defined as the swept surface along the tool axis, the forbidden region is around the PSM2. The VF has a cylindrical shape with a radius which is double the tool radius.

¹ <http://endovissub2017-roboticinstrumentssegmentation.grand-challenge.org/>

Assuming that the last two joints are kept still, the cylinder axis direction corresponds to z_{t_2} axis of $\mathcal{F}_{t_2}$ tracked by the EKF. The current pose of PSM1 is tracked in $\mathcal{F}_{b_1}$ using the dVRK kinematics and then mapped in $\mathcal{F}_{b_2}$ through the transformation matrix $T_{b_1}^{b_2}$. The minimum distance between the PSM1 tool tip position $\mathbf{x}$ and the cylindrical FRVF corresponds to the length of the line segment which joins perpendicularly the point to the axis minus the radius of the cylinder. A constraint enforcement method is defined, consisting in the application of a spring-damper like force:

$$\mathbf{f}_{vf}(\tilde{\mathbf{x}}, \dot{\tilde{\mathbf{x}}}) = -\mathbf{K}_{vf}\tilde{\mathbf{x}} - \mathbf{D}_{vf}\dot{\tilde{\mathbf{x}}} \quad (5)$$

where $\tilde{\mathbf{x}} = \mathbf{x}_d - \mathbf{x}$ is the displacement between the point $\mathbf{x}_d$, belonging to the constraint geometry having minimum distance from $\mathbf{x}$. The matrices $\mathbf{K}_{vf}$ and $\mathbf{D}_{vf}$ are properly designed diagonal and positive definite. The external force is not directly measurable, it is estimated by resorting to a non-linear dynamic observer [5], [24], and [25]. Finally, the force imposed by the Virtual Fixture is mapped on the MTM, so that it exhibits a repulsive behaviour and it pulls the robot end-effector away from the forbidden region. The MTM impedance controller exhibits a closed-loop behaviour that can be described by

$$\mathbf{M}\ddot{\tilde{\mathbf{x}}} + \hat{\mathbf{D}}\dot{\tilde{\mathbf{x}}} + \mathbf{K}_{vf}\tilde{\mathbf{x}} = \mathbf{f}_h \quad (6)$$

where $\hat{\mathbf{D}} = \mathbf{D} + \mathbf{D}_{vf}$ contains the damping assigned both by the impedance control and the constraint enforcement method.

III. EXPERIMENTAL EVALUATION

The experimental validation is performed on dVRK robot, which is controlled at the MTM by an impedance controller as described in Section II-A, with $m_{ii} = 1.5$ and $d_{ii} = 0$, being the (i, i) entries of the matrices $\mathbf{M}$ and $\mathbf{D}$, respectively. The $\mathbf{D}_{vf}$ has been adapted according to the stiffness variation such that $d_{vf,ii} = 2\sqrt{m_{ii}k_{vf,ii}}$ where $d_{vf,ii}$ and $k_{vf,ii}$ are the diagonal values of the matrices $\mathbf{D}_{vf}$ and $\mathbf{K}_{vf}$, respectively and $k_{vf} = 8$ N/m, as in [5]. The dVRK dynamic parameters are identified in [6]. During the experiments, two Endowristda Vinci tools are used: curved scissors and Prograsp forceps. The kinematics data from the dVRK are acquired at 200 Hz, while the vision-based system estimated the tool position at the camera frame rate of 25 Hz. The EKF approach allows overcoming this limitation, providing tool pose at 200 Hz. The tool segmentation is performed using GPU implementation on an NVIDIA GTX 1080 Ti to speed up computation.

A. Tracking Method Evaluation

The proposed tracking method is preliminarily evaluated on a simple task, executed with dVRK robot. The task is planned ad hoc to reduce the variability introduced by the telemanipulation, and, thus, to obtain a reference target to measure the error. Two specific points, placed on a phantom tissue at a distance of 15 mm, are recorded offline from kinematic data, by holding the tool steady in the given positions. In this condition, the position error introduced by the kinematics, in the two selected points, is minimized since the tool is fixed and the interaction force with the phantom goes towards zero. After

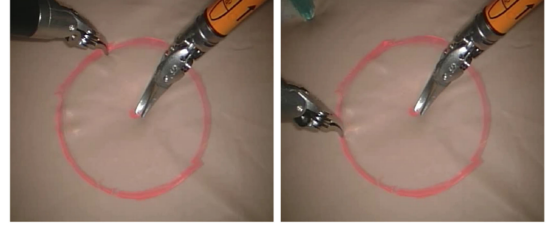


Fig. 3. Two different frames of the second evaluation experiment; PSM1 tool holds the center of the circle; PSM2 moves following the circle.

that, a linear path is defined analytically between the two points to serve as ground truth for the evaluation. The experiment, conducted to evaluate our tracking method, consists in moving the tool, in teleoperation mode, along the defined linear path drawn on the phantom. The task is performed slowly, with a duration of 12 seconds, to minimize the error along the linear path introduced by telemanipulation. During the task execution, the surgical tool is tracked using the EKF method. Then, the estimated pose is compared to the target linear path, obtaining mean absolute errors of 0.126 ± 0.08 mm along the x-axis, and 0.02 ± 0.01 mm along the y-axis. The results demonstrate the goodness of our tracking method. Furthermore, we computed the error obtained just using kinematics information, obtaining mean absolute errors of 0.135 ± 0.08 mm along the x-axis, and 0.02 ± 0.01 mm along the y-axis. As expected, the pose error is similar to the one obtained with our tracking method because of the absence of interaction with the environment.

B. Collision Avoidance Evaluation

The collision avoidance strategy is evaluated in two different tasks. During the first evaluation test, the PSM1 tool is fixed and the PSM2 is moved by the user in teleoperation mode towards PSM1. The collision avoidance strategy is applied during the entire duration of the test. Fig. 4 shows the distance between the two surgical tools, computed considering the proposed tracking method, and the related haptic force norm rendered to the user through the master side (MTM) during the task. The maximum reached force is 3.2 N.

The second evaluation test consists in a human subject study to show significant differences in performance caused by the introduction of force feedback. The study involves 12 subjects divided into two groups, 6 experienced and 6 novice surgeons, based on self-evaluation about their experience in the use of da Vinci Robotic system for minimally invasive surgical procedures. The study is articulated in two experiments using the dVRK robot in teleoperation mode. Taking inspiration from [4], the test simulates burning tissue with an electrocautery device. During each test, the subject keeps the PSM1 centered in the middle of a circle with a diameter of 20 mm. Meanwhile, the PSM2 has to follow the circular path for 270° from a definite starting point, as shown in Fig. 3. In the first experiment, the subjects perform the test 5 times moving the surgical tool in free motion. In the second experiment referred as VF constraint tasks, they perform the same task 5 times with the proposed

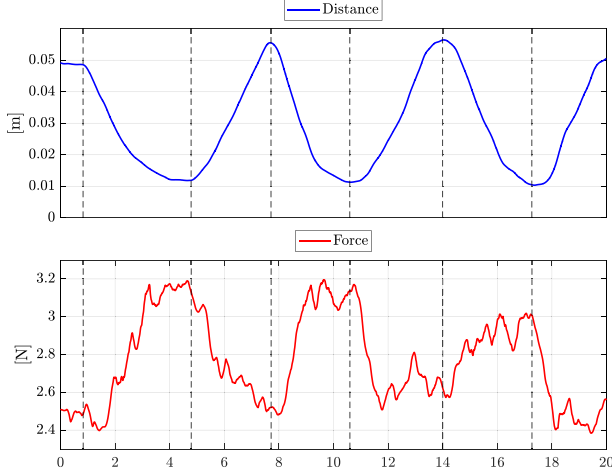


Fig. 4. First evaluation experiment. Duration: 20 seconds. Time histories of: (Blue) Distance between surgical tools; (Red) Related estimated force norm.

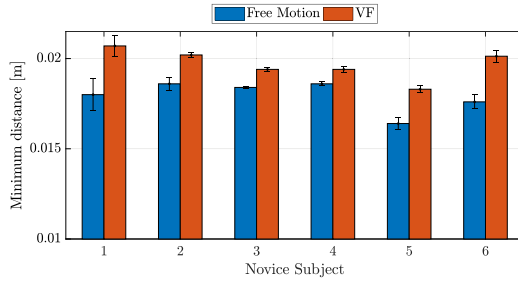


Fig. 5. Novice subjects: Mean values of minimum distance between tools with standard error bars.

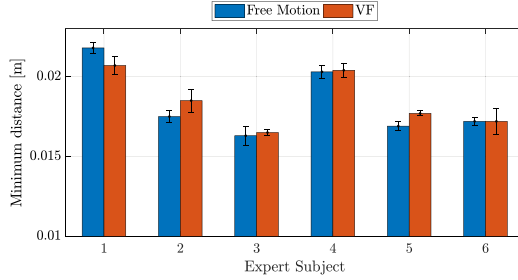


Fig. 6. Expert subjects: Mean values of minimum distance between tools with standard error bars.

collision avoidance constraint applied. Each task has an average execution time of 10 seconds. Each subject is asked to try the test in advance, to become familiar with the task itself and with the dVRK platform. The minimum distance between the tools is considered as performance parameter and it is computed using the proposed tracking method, in the VF constraint test the maximum force felt during the task is also computed.

Fig. 5 and Fig. 6 shows the mean values of the minimum distance between tools for novice and expert subjects during free motion and VF constraint tasks. The error bars represent the standard error of the means. To demonstrate the statistical relevance of the results, a comparison is made between the mean values of minimum distance, through a statistical unpaired

TABLE I
MAXIMUM FORCE AND T-TEST RESULTS ON MINIMUM
DISTANCE FOR NOVICE AND EXPERT USERS

Novice	test	p	F_M [N]	Expert	test	p	F_M [N]
1	1	0.0044	2.4416	1	0	0.1352	3.4527
2	1	0.0127	3.0749	2	0	0.0856	2.8175
3	1	0.0030	3.3411	3	0	0.8286	3.5239
4	1	0.0219	2.8188	4	0	0.8757	2.6180
5	1	0.0206	3.9998	5	0	0.1140	3.0035
6	1	0.0012	3.4170	6	0	1	2.8800

t-test, with a significance level $\alpha = 0.05$. As presented in Table I, the test shows statistically significant differences between the means for all subject in the novice group. Moreover, it presents an increase in the minimum distance values of $\sim 10\%$ in collision tests with respect to free-hand tests. The estimated force norm, rendered to the novice users through the master side (MTM), during the collision avoidance tasks, has a mean value of 3.1822 ± 0.5368 N. The expert group presents a mean force norm of 3.0493 ± 0.3629 N.

C. Discussion

A real comparison between the EKF-based tracking method and the kinematic measurements cannot be reliable without the use of an external sensor providing the ground truth of the tool pose. Indeed, we point out that the method used in Section III-A to evaluate the tracking algorithm is affected by the error due to the low resolution of the camera, and by the variability given by telemanipulation, even if it is minimized thanks to the ad hoc designed experiment. The experiments minimize the interaction with the environment introducing low position error in kinematic data, but ensuring the correct execution of the tool's movement following the defined path. In future works, we aim to significantly improve the validation method by means of a more advanced calibration technique, and we aim to compare the tracking method with the kinematics measurements using an external camera to measure the true tool pose that can be considered as more appropriate ground truth for tracking evaluation.

Fig. 4 shows the repulsive force felt at the MTM when the distance between PSMs decreases. The method to generate the force is designed to have small forces values such as to be slightly perceived by the surgeon. This is because the purpose is not to interfere with surgeon's actions, but to serve as an alarm to remind the presence of the other instrument in the proximity. Indeed, during the experiments, the maximum value of the force norm is 3.2 N.

The human subject study has shown a statistically significant difference regarding the mean of the minimum distance between the tools for the novice subjects. The VF test outperformed the free-hand test and this result suggests that feeling a haptic force during the task allows maintaining a safe distance between the surgical tools. On the contrary, in the free-hand test, the subject has no force feedback during the task and could dangerously reduce the distance between the tools. The maximum reached force is lower than 4 N and it does not create variation in the task performance.

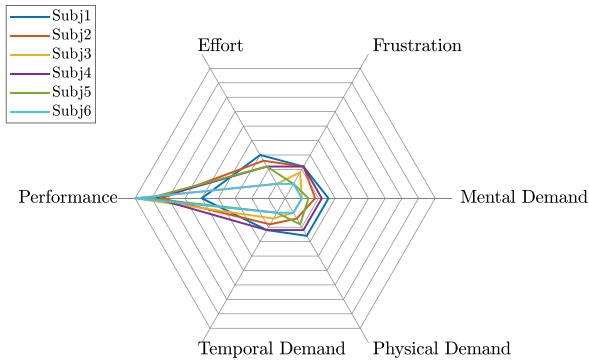


Fig. 7. Radar graph of the TLX results on expert users.

As concern the expert subjects, the test does not show a statistically significant difference in the VF constraint task with respect to the free motion task. Nevertheless, they were asked to compile the NASA-TLX questionnaire [26] to assess the perceived workload. The results of the questionnaires shown in Fig. 7 assess that the VF constrained task is not mentally, physically and temporally demanding and the force feedback does not affect negatively on the performances. On the contrary, it represents a comfortable reminder of the collision risk, that diminish the user mental workload. Similar results were obtained for novice surgeons.

IV. CONCLUSION

This letter introduces a method based on haptic guidance and virtual fixtures that allows avoiding surgical tools collision in MIRS. A marker-less algorithm allows estimating the PSM position and orientation, using kinematic and visual information. The PSM estimated pose is used to generate a FRVF, that aims to avoid collision between the two instruments through a repulsive force felt at the MTM during the surgical task execution. The proposed strategies are evaluated through multiple experiments on dVRK, showing good results in improving novice surgeon's performance. Furthermore, the use of VF allows also expert surgeon to better focus on the task, as far as the haptic force are small enough to suggest that the tools are dangerously close without affecting the performance. Therefore, the method can be considered effective both in a training stage of novice surgeon, as well as when the level of expertise increases. The goal for future works is to improve the accuracy of the tool pose estimation. For this purpose, more advanced method for hand-eye calibration and 3D reconstruction will be considered.

REFERENCES

- [1] M. E. Hagen, J. J. Meehan, I. Inan, and P. Morel, "Visual clues act as a substitute for haptic feedback in robotic surgery," *Surgical Endoscopy*, vol. 22, no. 6, pp. 1505–1508, 2008.
- [2] J. K. Koehn and K. J. Kuchenbecker, "Surgeons and non-surgeons prefer haptic feedback of instrument vibrations during robotic surgery," *Surgical Endoscopy*, vol. 29, no. 10, pp. 2970–2983, 2015.
- [3] R. Moccia, M. Selvaggio, L. Villani, B. Siciliano, and F. Ficuciello, "Vision-based virtual fixtures generation for robotic-assisted polyp dissection procedures," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2019, pp. 7928–7933.
- [4] A. Banach, K. Leibrandt, M. Grammatikopoulou, and G. Yang, "Active constraints for tool-shaft collision avoidance in minimally invasive surgery," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2019, pp. 1556–1562.
- [5] M. Selvaggio, G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Passive virtual fixtures adaptation in minimally invasive robotic surgery," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3129–3136, Oct. 2018.
- [6] G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, "Modelling and identification of the da Vinci Research Kit robotic arms," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1464–1469.
- [7] L. B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation," in *Proc. IEEE Virtual Reality Annu. Int. Symp.*, 1993, pp. 76–82.
- [8] M. Li, M. Ishii, and R. H. Taylor, "Spatial motion constraints using virtual fixtures generated by anatomy," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 4–19, Feb. 2007.
- [9] J. Ren, R. V. Patel, K. A. McIsaac, G. Guiraudon, and T. M. Peters, "Dynamic 3-D virtual fixtures for minimally invasive beating heart procedures," *IEEE Trans. Med. Imag.*, vol. 27, no. 8, pp. 1061–1070, Aug. 2008.
- [10] T. Xia *et al.*, "An integrated system for planning, navigation and robotic assistance for skull base surgery," *Int. J. Med. Robot.*, vol. 4, no. 4, pp. 321–330, 2008.
- [11] 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 Int. Conf. Intell. Robots Syst.*, 2012, pp. 3308–3313.
- [12] M. M. Marinho, B. V. Adorno, K. Harada, and M. Mitsuishi, "Dynamic active constraints for surgical robots using vector-field inequalities," *IEEE Trans. Robot.*, vol. 35, no. 5, pp. 1166–1185, Oct. 2019.
- [13] D. Bouget, M. Allan, D. Stoyanov, and P. Jannin, "Vision-based and marker-less surgical tool detection and tracking: a review of the literature," *Med. Image Anal.*, vol. 35, pp. 633–654, 2017.
- [14] S. Bodenstedt *et al.*, "Comparative evaluation of instrument segmentation and tracking methods in minimally invasive surgery," *Med. Image Anal.*, 2018, *arXiv:1805.02475*.
- [15] S. Allan *et al.*, "2D-3D pose tracking of rigid instruments in minimally invasive surgery," *Inf. Process. Comput.-Assisted Interventions*, vol. 8498, pp. 1–10, 2014.
- [16] X. Du *et al.*, "Combined 2D and 3D tracking of surgical instruments for minimally invasive and robotic-assisted surgery," *Int. J. Comput. Assisted Radiol. Surgery*, vol. 11, no. 6, pp. 1109–1119, 2016.
- [17] X. Du *et al.*, "Articulated multi-instrument 2-d pose estimation using fully convolutional networks," *IEEE Trans. Med. Imag.*, vol. 37, no. 5, pp. 1276–1287, May 2018.
- [18] 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," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2014, pp. 6434–6439.
- [19] B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer. A*, vol. 4, no. 4, pp. 629–642, 1987.
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention MICCAI*, 2015, pp. 234–241.
- [21] A. A. Shvets, A. Rakhlin, A. A. Kalinin, and V. I. Iglovikov, "Automatic instrument segmentation in robot-assisted surgery using deep learning," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl.*, 2018, pp. 624–628.
- [22] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *Int. J. Comput. Vision*, vol. 81, no. 2, pp. 155–166, Feb. 2009.
- [23] M. Ferro, G. A. Fontanelli, F. Ficuciello, B. Siciliano, M. Vendittelli, "Vision-based suturing needle tracking with extended kalman filter," in *Proc. Comput./Robot Assisted Surgery Workshop*, 2017, pp. 36–37.
- [24] A. D. Luca and R. Mattone, "Actuator failure detection and isolation using generalized momenta," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2003, pp. 634–639.
- [25] A. D. Luca and R. Mattone, "Sensorless robot collision detection and hybrid force/motion control," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 999–1004.
- [26] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.