2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

SCAN: System for Camera Autonomous Navigation in Robotic-Assisted Surgery

Tommaso Da Col[1], Andrea Mariani[2], Anton Deguet[3], Arianna Menciassi[2], Peter Kazanzides[3] and Elena De Momi[1]

Abstract— **Robot-Assisted systems for Minimally Invasive Surgery enhance the surgeon capability, however, direct control over both the surgical tools and the endoscope results in an increased workload that leads to longer operation times. This work investigates the introduction of SCAN (System for Camera Autonomous Navigation) to overcome this limitation. An experimental study involving 12 participants was carried out with the da Vinci Research Kit. Each user tested two novel camera control modalities, autonomous and semi-autonomous, as well as the current manual control of the camera, while carrying out a dry-lab task. Among the camera control modalities, the autonomous navigation achieved better objective performances and the highest user confidence. Moreover, the autonomous control (along with the semi-autonomous one) was able to optimize some metrics related to the robotic surgery workflow.**

Index Terms— Medical Robots and Systems, Human-Centered Automation, Telerobotics and Teleoperation

I. INTRODUCTION

A. Research Field

Robotic-Assisted Minimally Invasive Surgery (RAMIS) is becoming more and more relevant in the surgical scenario [1]. In general, a surgical robot is a device that provides the surgeons with a new set of versatile features that extend their capability to treat patients [2]. Medical robots can be classified in many different categories. For example, the da Vinci Surgical System (dVSS, Intuitive Surgical Inc., Sunnyvale, California) is one of the most widely used RAMIS platforms and it belongs to the surgeon extender category. This type of system is generally composed of leader robotic arms manipulated by the surgeon with the purpose to teleoperate the follower arms mounting the surgical tools and supporting the endoscope on the patient side. The reason for RAMIS success lies in the advantages that it introduces for both the patient and the surgeon [3]. Smaller scars, less blood loss and lower pain make the hospitalization and the recovery times shorter. On the other hand, motion scaling and tremor filtering increase the dexterity of the surgeon, that is moreover provided with an immersive tridimensional vision. This last advantage is of particular interest. The visualization modalities, in fact, changed enormously across the evolution of surgery [4]. While in open surgery the surgical site is directly viewed, in Laparoscopic MIS (LMIS) the surgeon views the images coming from the inside of the patient body on a monitor. This approach lacks 3D perception and it provides only a restricted Field of View (FoV). The advent of RAMIS has been able to improve vision while maintaining all the other advantages of MIS. Another important aspect is that the control modalities of the endoscope changed as well [5]. In LMIS, an assistant is in charge to sustain and position the endoscope based on surgeon instructions. This aspect requires a perfect collaboration between the two that is often difficult to obtain. Robotic-Assisted Endoscopic Manipulators (RAEMs) have been developed to give back to the surgeons the direct control over their FoV. Moreover in RAMIS, the surgeon directly teleoperates both the surgical tools and the endoscope. This approach brings out two main problems. First, switching continuously between the teleoperation of the instruments and the camera reduces the smoothness of RAMIS procedures, thus causing longer operating times. Second, dealing with the new control dynamics of the endoscope makes camera motion complicated and it increases the cognitive workload for the surgeon, leading to errors in the execution of surgical operations [6]. These limitations motivated the attempts to exploit the use of different types of Human-Machine Interfaces (HMIs) to assist the surgeon in positioning the endoscope and to speed up the procedure. Different kinds of control such as hands manipulation, head movement, voice commands and foot interfaces have been developed and tested. One alternative way to reduce the cognitive workload for surgeons is to move towards the autonomy of the Operating Room (OR). Generally, the term autonomy in the field of surgery means that a defined surgical task is performed autonomously by a not-human system, for example, a robot. However, the step from simple task execution to autonomous decision making is difficult to implement. Autonomy is more than a repetition of predefined movements; it involves perception of the environment and corresponding adaption of behavior if needed. Robotic surgical systems can help by taking control of certain segments of the surgical workflow to accomplish some repetitive tasks such as, for example, positioning the camera.

B. Related Works

In order to reduce the time involved in tasks other than the actual teleoperation of the surgical tools, such as the endoscope positioning, many different types of HMIs have been developed and tested. The Automated Endoscopic System for Optimal Positioning (AESOP, Computer Motion,

[1] are with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133, Italy, (tommaso.dacol/elena.demomi@mail.polimi.it).
[2] are with The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, 56121, Italy, (an.mariani/arianna.menciassi@santannapisa.it).
[3] are with the Department of Computer Science, Johns Hopkins University, Baltimore, 21218, Maryland, (anton.deguet/pkaz@jhu.edu).

Inc., Santa Barbara, California, and later Intuitive Surgical Inc., Sunnyvale, California) [7] was the first RAEM that obtained FDA approval. Different HMIs have been evaluated to control it. The system was integrated with a foot interface that in [8] was found to be more effective and accurate than a human assistant in manipulating the laparoscope, without implying a longer learning curve [9]. Voice control was found to be more accurate with respect to the previous approach even if it implied slower learning and less intuitive manipulation [10]. LAPMAN (Medsys, Gembloux, Belgium) [11] and SOLOASSIST (AKTORmed GmbH, Barbing, Germany) [12] are examples of RAEMs that are integrated with joysticks that allow hand control. EndoAssist (Armstrong Healthcare, High Wycombe, United Kingdom) [13] and Free-Hand (Freehand 2010 Ltd., Guildford, United Kingdom) [14] can be guided by head movement. ViKY® (EndoControl Medical, La Tronche, France) [15] or LARS (IBM, Armonk, New York) [16] can be controlled by using voice commands. Other approaches to the endoscope control of the dVSS have also been investigated. These studies have been conducted on the da Vinci Research Kit [31], the research platform for the dVSS. In [17] the iNMI (improved Novel Master Interfaces), an interface composed by capacitive touch sensor arrays, allowed to position the endoscope by using the index fingers. The system's stereo-viewers were integrated in [18] with sensors to record information related to the head movements and to use it to control the camera. It is complex to determine which one among these approaches offers the best performances, but for sure none of them eliminates the necessity for the surgeon to directly control the endoscope. An alternative to overcome this issue is the development of autonomous camera navigation systems able to center the optimal Region of Interest (RoI). Some researches have been conducted to provide the already depicted RAEMs with autonomy. The AESOP was integrated in [19] with an IR eye tracking system to center the view on the user gaze point. The AutoLap had a follow-me mode of operation in which the system used image-guidance to track and follow a designated tool [20]. The ViKY® exploited the use of a novel tip tracking control method to follow the surgical instruments during the execution of a procedure in [21]. In fact, the operators manipulate the instruments in order to achieve the desired task's goals, and this indicates that their intent is closely related to the way in which they moves the tools. A work carried out using the dVRK, reports the implementation of a mid-point tracking algorithm. The scene was centered on the centroid of the tools' Center of Mass (CoM) while the endoscope insertion was adjusted on the basis of the tools relative distance, in order to maintain them inside the FoV [22]. In [23], [24] and [25], a novel flexible endoscope designed to be mounted on the dVRK follower arms was presented. It was integrated with an image-based algorithm for autonomous camera navigation. Other possible strategies are the eye-tracking technologies. The user's point of gaze, in fact, is often used to make inferences about the user's attention [26], [27]. In [28], this approach was used to automatically center the laparoscopic camera viewpoint at a user point of gaze or to indicate the respective intents of two users through simultaneous displays. However, there exist safety and implementation issues that limit the application of automation intra-operatively [29].

C. Research Hypothesis

This project introduces the SCAN (System for Camera Autonomous Navigation), that tries to solve the mentioned problems and to overcome the limitations related both to the current camera control and to the novel systems previously listed in the following ways. Using the same setup as a current system (i.e., the dVSS), it features autonomous and semi-autonomous camera control modalities that may be able to shorten the total time of surgical operations by reducing the time needed for positioning the endoscope. Moreover this research is centered on the investigation of shared autonomy. This means leaving the surgeons in the loop, always providing them with a clear clue about the autonomous movement of the endoscope; additionally, the surgeons can overrule the camera movement at any time, with the possibility to adapt it for specific surgical steps that require special viewpoints.

II. METHODS

This section presents implementation details of the SCAN system and materials and methods of an experimental study to evaluate it during the execution of a task in a dry-lab.

A. System Setup

Thanks to the possibility to act on all the control software levels and its similarity with the dVSS, the dVRK was the chosen platform to carry out this research. The dVRK is an open-source mechatronic platform [31], [32]. It is composed by the hardware of the first generation dVSS, integrated with customized control electronics, firmware and software. The structure is divided into two main parts: the surgeon console and the patient side. The patient side features four follower arms. Three of them are called Patient Side Manipulators (PSMs) and they mount interchangeable surgical instruments that enter the patient and are used to perform different procedures. The fourth arm, referred to as the Endoscopic Camera Manipulator (ECM), is in charge of supporting and positioning the endoscope. The surgeon side consists of two leader arms called Master Tool Manipulators (MTMs) to tele-manipulate the follower arms, a 3D imaging system (High Resolution Stereo-Viewer, HRSV) and a foot pedal tray with different buttons and associated functions. Our experimental setup was composed of two PSMs mounting Large Needle Driver tools and the ECM equipped with a straight 12mm endoscope.

B. SCAN: System for Camera Autonomous Navigation

The autonomous camera navigation algorithm that we tested in a virtual reality framework in [30] was developed to control the ECM of the dVRK and improved by adding the possibility to center the scene on three different points of the image, to decide the best zooming factor and also to enable or disable the autonomous tracking of the tools. The next paragraphs will first present the different types of camera control: manual, autonomous and semi-autonomous. Later, the functioning of the algorithm in different scene center modalities will be described: left/right tool, tools midpoint and wide-view.

The use of three modalities to control the ECM was investigated.

- Manual camera Control (MC) modality: the endoscope positioning workflow was exactly the same as the real dVSS. In order to switch from the teleoperation of the surgical tools to teleoperation of the endoscope, the users had to continuously press a foot pedal at the surgeon console. Then, by moving the MTMs, they were able to change their FoV;
- Autonomous camera Control (AC) modality: the system tracked the tools and adjusted the position of the endoscope accordingly, resulting in a continuous movement of the scene. The view was dependent on the scene center modality (which details are presented in the next paragraph) selected by the user with a foot pedal;
- Semi-autonomous Camera (SAC) control modality: the user could decide between triggering the autonomous endoscope movement by pressing a foot pedal or directly controlling the camera in the same way as in the manual camera control modality.

As already stated, the users could decide to center the image on the left/right tools or their CoMs midpoint (all the variables were obtained through kinematics tracking), and also to modify the magnitude of the zoom. In particular, they could choose to adjust the zoom in relation to the distance between the tools, or maintain it fixed. All the different cases are described in the following paragraphs. In the *left/right tool* scene center modality, the endoscope position was computed using the following formula:

$$\mathbf{EP}^t = \mathbf{SC}^t - z\frac{\mathbf{SC}^t - \mathbf{RCM}}{\|\mathbf{SC}^t - \mathbf{RCM}\|} \quad (1)$$

where the vectors are in bold and they are all expressed with respect to the same reference frame. EP is the Endoscope end effector Position, the Scene Center (SC) is the position of the left or right tool end effector and the RCM is the position of the Remote Center of Motion. In MIS, the RCM corresponds to a small incision, through which the surgical tools or the endoscope enter the patient's skin. z was the zooming factor and was chosen $z = 0.1m$ in this case. Note that the word zoom refers here to the relative distance between the SC and the EP that in this case is maintained fixed, not the optical zoom of the cameras. All the main elements composing the camera can be seen in Fig. 1.

In the *tools midpoint* scene center modality, the SC was computed as:

$$\mathbf{FP}^t = \frac{\mathbf{CoM}_l^t + \mathbf{CoM}_r^t}{2} \quad (2)$$

where CoM_l and CoM_r are the Centers of Mass of the left and right tools, respectively. Then the EP was obtained as:

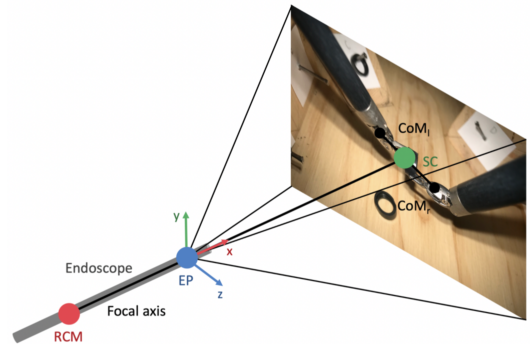

Fig. 1: The main components of the endoscope model: Remote Center of Motion (RCM), Endoscope end effector Position (EP) and Scene Center (SC). In this image, the scene center corresponds to the midpoint of the Centers of Mass of the surgical tools.

$$\mathbf{CC}^t = \mathbf{RCM} + z_{MAX}\left(1 - \frac{\|\mathbf{CoM}_l^t - \mathbf{CoM}_r^t\|}{d_{MAX}}\right) f\mathbf{x} \quad (3)$$

where $z_{MAX} = 0.85$ and $d_{MAX} = 0.2m$ are respectively the maximum zoom allowed and the maximum distance between the tools (these values are set to match the task workspace), while $f = 0.1m$ is the camera nominal focal length and x is the camera focal axis. In this case, the zoom is variable and dependent on the relative distance between the left and right tools. Lastly, in the *wide view* scene center modality, the SC corresponds to the tools midpoint and is computed according to (2), while (1) is modified in the following way to compute the EP:

$$\mathbf{EP}^t = \mathbf{RCM} + z\frac{\mathbf{SC}^t - \mathbf{RCM}}{\|\mathbf{SC}^t - \mathbf{RCM}\|} \quad (4)$$

with $z = 0.1m$ resulting in a fixed zoom once again.

C. Task

The Laparoscopic Skills Training and Testing (LASTT) method was used in this study [33]. The LASST is currently used in the practical skill assessment of a structured diploma for surgeons [34]. It is a validated practical test to measure the competence level of an individual in basic LMIS psychomotor skills: camera handling, hand-eye coordination and bi-manual coordination. The task we chose was optimized for RAMIS basic skills evaluation. The platform was composed by eight rings and the same number of destination pegs, each one identified by a letter. The user had to pick one ring at a time from the center of the platform and place it on a destination peg following the alphabetical order. The user was provided with a graphic overlay on the images shown in the HRSV, which showed the next letter to locate. The position of the letters changed across each repetition. To achieve a valid placement, the distance between the endoscope tip and the destination peg needed to be under a distance threshold d = 0.1m; this choice aimed to force the user to move the

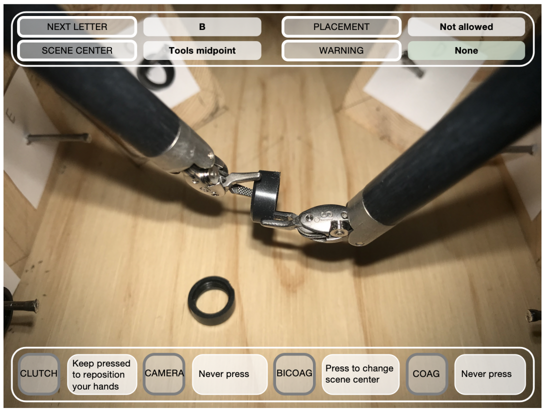

Fig. 2: The Graphic Overlays viewed by the user during the execution of the task, in the top window are shown: 1. Next placement letter; 2. Validity of the placement, related to the distance between the endoscope tip and destination peg. In this case, the placement was not yet allowed because the endoscope was too far away from the scene resulting in a sub-optimal viewpoint. 3. Scene center modality, in this case *Tools midpoint*, which meant that the SC was the midpoint of the COM of the tools. 4. Warning, in this case empty, provided the user with a warning if one of the two tools went out of the FoV. In the bottom window, information related to the functions of the different pedals is shown.

endoscope and evaluate the camera manipulation skills. The information related to the validity of the placement was shown by means of another visual feedback. The graphic overlays contained additional information such as the current scene center (left/right tool, tools midpoint or wide view) chosen by the user and provided feedback related to the optimal use of the endoscope, showing a warning if one of the two tools went outside the endoscope FoV. Other graphic overlays included suggestions about the pedal functioning. They all can be seen in Fig. 2.

D. Performance Metrics

In order to assess the user's performance, we considered both objective and subjective metrics.

Regarding the objective metrics, the LASTT method assessed the user's skill through a validated metric depending on the task execution time. This overall performance was computed as: $P = N/t$ [33]. N was the number of ring placements achieved in a task repetition (8 maximum) and t the duration of the repetition expressed in s (8 minutes maximum). In order to evaluate objectively the endoscope manipulation skill, we considered the total time in which at least one of the two surgical tools was outside the FoV: $tFoV$ in s [35]. Other objective metrics were the percentage time of actual teleoperation of the surgical tools: $teleop$ (without considering the switching to control the endoscope or the clutching to reposition them) and the total time of clutch: $tClutch$ in s. The subjective metrics were extracted from a post-experiment survey, whose questions included to express a preference between the different types of camera control and to evaluate different aspects related to these modalities.

E. Acquisition Protocol

A user study was carried out on a population of 12 non-medical subjects (21 to 58 years old, 11 males and 1 female, all right-handed) to evaluate the SCAN system (HIRB00000701). After declaring verbal consent, the subjects were introduced to the dVRK system. Then, the users were shown videos related to the task execution with all the different endoscope control modalities. After this introductory phase, they were given the possibility to familiarize themselves with the system, trying all the different types of endoscope control with a 3 minutes session for each modality. All the subjects performed 2 repetitions of the task with each endoscope control modality: manual, autonomous and semi-autonomous. The order in which the user carried out the task under the different types of control was randomized with permuted-block randomization (with the additional constraint of having one repetition for each camera control modality in the first 3 repetitions and the same order in the final 3 repetitions). A repetition ended if the user achieved the placement of all the 8 rings or if the repetition time reached 8 minutes.

F. Statistical Analysis

To compare the objective metrics between the different camera control modalities, non-parametric statistical significance tests were exploited, taking into account the small sample size. Considering the metrics as dependent variables and different camera control modalities as independent factors, the Wilcoxon rank sum test was used. Note that the group composed by the first 3 repetitions and the second one that contained the remaining 3 were analyzed separately. Statistically significant differences were assessed at $p < 0.05$. We used MATLAB for the statistical analysis, in particular the command `ranksum()`.

III. Results

The experimental protocol was completed by all 12 subjects. The main research hypothesis was to investigate the effects of the introduction of the SCAN in the field of RAMIS. In order to evaluate its effectiveness and usability, in this section we first report the objective metrics related to the task execution, and we later present the subjective opinions the users expressed with respect to the camera control modalities.

A. Objective Metrics

The left box of Fig. 3 on the following page shows the boxplots related to the overall performances P of the users in the different camera control modalities: Manual Control (MC), Autonomous Control (AC) and Semi-Autonomous Control (SAC). These performances are divided into 2 groups composed by the first and the last 3 repetitions, respectively. The results of the Wilcoxon rank sum test highlighted statistically significant higher performances using the AC control modality with respect to both the MC and SAC across

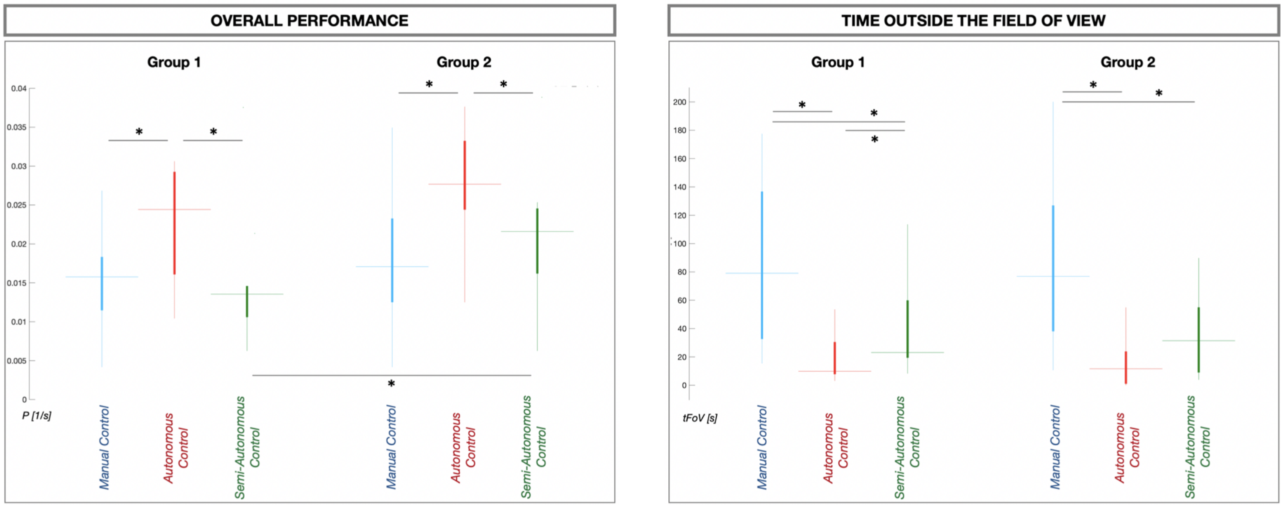

Fig. 3: Objective Metrics. The box on the left is related to the overall performance, P, the one on the right to the time in which at least one of the two tools was outside the FoV, $tFoV$. In the boxplots, the median is identified by a horizontal line, the first and the third quartiles are depicted as bold line edges, while the thin line edges represent the whiskers. The stars denote a statistically significant difference between the control modality medians.

both the groups ($p^1_{ACvsMC} = 0.0194$, $p^2_{ACvsMC} = 0.0043$; $p^1_{ACvsSAC} = 0.0073$, $p^2_{ACvsSAC} = 0.0226$). This difference is related to the additional task of manually controlling the camera, accentuated in the MC but still present with the SAC; the difference may be also determined by the use of many different pedals, especially for the SAC modality. A significant difference related to the performance of group 2 with respect to group 1 ($p^{1vs2}_{SAC} = 0.0304$) in SAC is also present. This could suggest that this modality requires a learning curve with respect to the other modalities that did not feature a significant difference between the first and second repetitions.

In the same figure, the boxplots related to the metric $tFoV$ are shown on the right. For both the AC and SAC, the times in which at least one of the two surgical tools were outside of the field of view were significantly shorter with respect to the MC ($p^1_{ACvsMC} < 0.001$, $p^2_{ACvsMC} = 0.0014$; $p^1_{SACvsMC} = 0.0351$, $p^2_{SACvsMC} = 0.0166$). The main reasons behind these results are that some of the scene center modalities, (*tools midpoint* and *wide view*), optimized the camera behaviour by automatically centering both the tools inside the field of view. Additionally, the $tFoV$ metric was significantly lower for the AC with respect to the SAC in the first group ($p^1_{ACvsSAC} = 0.0304$), while this did not apply in the second group (maybe due to a learning effect). Moreover in the AC and SAC, the use of visual feedback was investigated and the users were provided with a graphic overlay warning to remind them to pay attention to this important aspect.

To evaluate the ability of the SCAN to reduce the time in which the leader arms' movement was not related to the actual teleoperation of the tools (such as endoscope teleoperation or leader arms re-positioning by using the clutch foot pedal), two additional metrics were analyzed: the percentage of time in which the users teleoperated the tools over the entire duration of the repetition ($teleop$) and the total time of clutching ($tClutch$). Regarding the former, the statistical analysis showed significant difference between the repetition in AC and MC in both the groups ($p^{1,2}_{ACvsMC} < 0.001$); the same applied to the comparison of SAC vs. MC ($p^{1,2}_{SACvsMC} < 0.001$), and AC vs. SAC ($p^{1,2}_{ACvsSAC} < 0.001$). This is simply due to the fact that in the AC the teleoperation of the tools was always enabled except during the clutching; in SAC, when the user decided to trigger the autonomous movement of the endoscope instead of controlling it manually, it was possible to more quickly re-engage teleoperation of the tools. Moreover, in this modality the autonomous re-positioning of the leader arms during

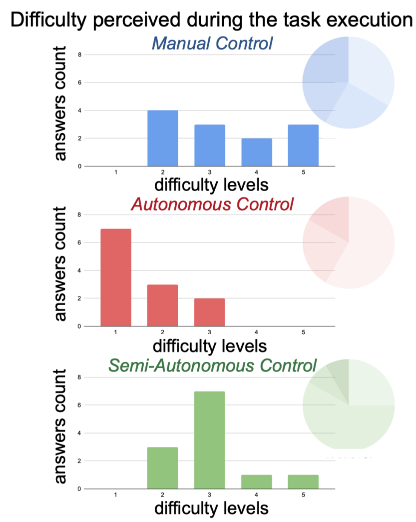

Fig. 4: Subjective Metrics. Column charts related to the difficulty level (the lower, the easier) perceived by the user during the execution of the task in the different camera modalities.

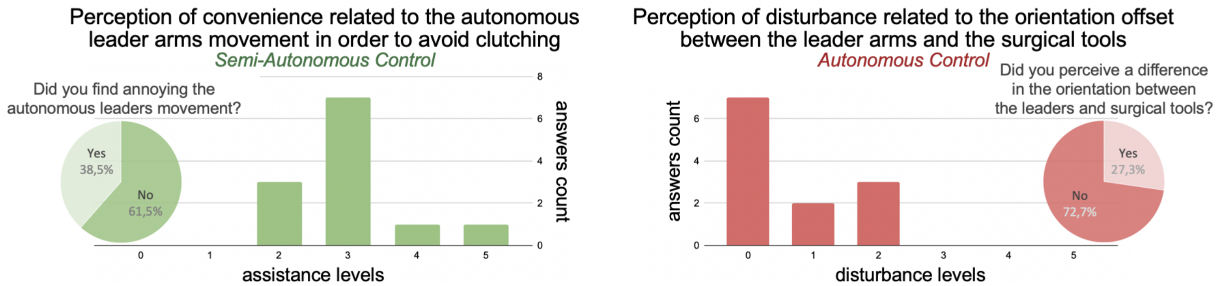

Fig. 5: Subjective Metrics. Column charts and pie charts related to the disturbance perceived by the user regarding the orientation offset and the autonomous leader arms movement with autonomous and semi-autonomous control, respectively.

the camera movement reduced the need for clutching. This second aspect is underlined even in the analysis of the total amount of time spent by the users to re-position the leader arms. Statistical difference between the SAC and the MC was present for both the groups ($p^1_{SACvsMC} = 0.0404$, $p^2_{SACvsMC} = 0.0120$).

B. Subjective Metrics

The answers to the post-experiment survey were analyzed in order to understand the subjective feedback and usability related to the different camera control modalities. Fig. 4 on the previous page shows column charts related to the count of the difficulty level perception in the execution of the task for each control modality. The users were also asked to express a preference among the different types of endoscope control. The predominance of AC, 91.7% of the users, assumes particular importance if we consider that 77% of them were familiar or experienced with the dVRK, the dVSS or their VR simulators. This means that they already know how to control the camera with the manual method, whereas it was the first time for them to exploit the use of the autonomous system.

Two other important pieces of information can be extracted from the data in Fig. 5. On the right, the perception of the orientation offset between the leader arms and the surgical tools is shown. In the current practice, the system reflects the orientation of the surgical instruments on the leader arms before enabling their teleoperation (for example, when recovering from the endoscope control). This provides the users with the feeling that the orientation of their hands are exactly the same as the tools that they are controlling. The continuous teleoperation workflow achieved with the AC does not allow this feature. However, the users' feedback seems to show that this aspect was not important at least for the execution of this particular task. On the left, the answers related to the perception of disturbance generated by the leader arms' autonomous re-positioning during the autonomous camera movement in SAC are shown. In this case, even if the users perceived the importance of this feature to reduce the need of clutching (as shown also by the objective data), this kind of movement was not perceived as natural by all the users.

IV. Conclusion

This work investigates the introduction of the SCAN (System for Camera Autonomous Navigation) in the field of robotic surgery, both in terms of performance enhancement and usability. The main outcomes of this study are:

- The autonomous camera control allowed the users to perform significantly better with respect to the manual camera control. This was probably due to both the time gain and the reduction of the workload of the user who did not need to manually control the endoscope. In fact, a shorter time for performing tasks that are not related to the direct teleoperation of the surgical tools characterized the autonomous with respect to the manual control. On the other side, the answers to the survey seem to suggest that the users subjectively perceived the task to be easier while using autonomous camera control.
- The semi-autonomous control modality was not capable to improve the users' performances, showing outcomes statistically equal to the manual one. This could be due to the need for learning how to deal with this less intuitive mode.
- Both the autonomous camera control and the semi-autonomous camera control allowed to optimize metrics related to the endoscope manipulation, such as the time in which the tools were left outside the field of view. This was achieved thanks to camera control logic and to the warning-feedback provided to the user.

Some secondary outcomes are:

- The orientation offset between the leader arms and the followers in the autonomous modality was not perceived as a relevant disturbance, at least not for the range of orientation offsets encountered during the experiments.
- The introduction of partial autonomy related to the clutching allowed the optimization of additional portions of the teleoperation workflow.

This work constitutes a pilot study but the results justify a deeper research related to cooperative-autonomy for the endoscope control in RAMIS. Future works should apply the SCAN usability investigation in a task that is more related to the operating theatre (e.g., performing a surgical procedure by means of an anatomical phantom or ex-vivo) and they should involve surgical residents and surgeons.

Moreover, the semi-autonomous camera control has to be further investigated in terms of effectiveness and learning curve. This can constitute a step forward to the selection of the right trade-off between autonomy and user control: from the one side, the advantages of the autonomous system in terms of smoothness of the operational workflow, and from the other side, the manual control in terms of safety and versatility.

REFERENCES

[1] S. Maeso, M. Reza, J. A. Mayol, J. A. Blasco, M. Guerra, E. Andradas and M. N. Plana, “Efficacy of the Da Vinci Surgical System in Abdominal Surgery Compared with that of Laparoscopy: a Systematic Review and Meta-Analysis”, Ann Surg, vol. 252, no. 2, pp. 254-262, Aug. 2010.

[2] R. H. Taylor, and D. Stoianovici, “Medical Robotics in Computer-Integrated Surgery”, IEEE Trans Robot Autom, vol. 19, no. 5, pp. 756-781, Oct. 2003.

[3] K. Moorthy, Y. Munz, A. Dosis, J. Hernandez, S. Martin, F. Bello, and A. Darzi, “Dexterity Enhancement with Robotic Surgery”, Surg Endosc, vol. 18, no. 5, pp. 790-795, May 2004.

[4] A. Pandya, L. A. Reisner, B. King, N. Lucas, A. Composto, M. Klein and R. D. Ellis, “A Review of Camera Viewpoint Automation in Robotic and Laparoscopic Surgery”, Surg Endosc, vol. 28, no. 7, pp. 2145-2158, Jul. 2014.

[5] I. Nisky, A. M. Okamura and M. H. Hsieh, “Effects of Robotic Manipulators on Movements of Novices and Surgeons”, Robotics, vol. 3, no. 3, pp. 310-329, Aug. 2014.

[6] B. Zheng, M. A. Cassera, D. V. Martinec, G. O. Spaun and L. L. Swanstrom, “Measuring Mental Workload During the Performance of Advanced Laparoscopic Tasks”, Surg Endosc, vol. 24, no. 1, pp. 45-50, Jan. 2010.

[7] S. W. Unger, H. M. Unger and R. T. Bass, “AESOP Robotic Arm”, Surg Endosc, vol. 8, no. 9, pp. 1131, Aug. 1994.

[8] A. W. Partin, J. B. Adams, R. G. Moore and L. R. Kavoussi, “Complete Robot-Assisted Laparoscopic Urologic Surgery: a Preliminary Report”, J Am Coll Surg, vol. 181, no. 6, pp. 552-557, Nov. 1995.

[9] L. K. Jacobs, V. Shayani, and J. M. Sackier, “Determination of the Learning Curve of the AESOP Robot”, Surg Endosc, vol. 11, no. 1, pp. 54-55, Jan. 1997.

[10] M. E. Allaf, S. V. Jackman, P. G. Schulam, J. A. Cadeddu, B. R. Lee, R. G. Moore, L. R. Kavoussi “Laparoscopic Visual Weld: Voice vs Foot Pedal Interfaces for Control of the AESOP Robot”, Surg Endosc, vol. 12, no. 12, pp. 1415-1418, Dec. 1998.

[11] R. Polet, and J. Donnez “Using a Laparoscope Manipulator (LAPMAN) in Laparoscopic Gynecological Surgery”, Surg Tech Int, vol. 17, pp. 187-191, 2008.

[12] S, Gillen, B. Pletzer, A. Heiligensetzer, P. WolfJorg Kleeff and H. Feussnert, “Solo-Surgical Laparoscopic Cholecystectomy with a Joystick-Guided Camera Device: a Case-Control Study”, Surg Tech Int, vol. 28, no. 1, pp. 164-170, Jan. 2014.

[13] J. M. Gilbert, “The EndoAssist Robotic Camera Holder as an Aid to the Introduction of Laparoscopic Colorectal Surgery”, Surg End, vol. 91, no. 5, pp. 389-393, Jul. 2009.

[14] M. H. Abedin Nasab, “Handbook of Robotics and Image Guided Surgery”, 2nd ed. vol. 1 pp. 57-78, Sep. 2019.

[15] M. Takahashi, M. Takahashi, N. Nishinari, H. Matsuya, T. Tosha, Y. Minagawa, O. Shimooki and T. Abe, “Clinical evaluation of complete solo surgery with the ViKY® robotic laparoscope manipulator”, IEEE Surg Endosc, vol. 31, pp. 981-986, Jul 2016.

[16] R.H. Taylor. J. Funda, B. Eldridge, S. Gomory, K. Gruben, D. LaRose, M. Talamini, L. Kavoussi and J. Anderson, “A Telerobotic Assistant for Laparoscopic Surgery”, IEEE Eng Med Biol, vol. 10, no. 1, pp. 1-13, May 2005.

[17] M. Kim, C. Lee, N. Hong, Y. J. Kim and S.Kim, “Development of Stereo Endoscope System with its Innovative Master Interface for Continuous Surgical Operation”, Biomed Eng, pp. 16-81, Dec. 2017.

[18] N. Hong, M. Kim, C. Lee and S.Kim, “Head-Mounted Interface for Intuitive Vision Control And Continuous Surgical Operation in a Surgical Robot System”, Med Biol Eng Comput, vol. 57, no. 3, pp. 601-614, Mar. 2019.

[19] S. M. Ali, L. A. Reisner, B. King, A. Cao, G. Auner , M. Klein, and A. K. Pandya, “Eye Gaze Tracking for Endoscopic Camera Positioning: an Application of a Hardware/Software Interface Developed to Automate Aesop”, Stud Health Technol Inform, vol. 132, pp. 4-7, Dec. 2007.

[20] P. J. M. Wijsman, I. A. M. J. Broeders, H. J. Brenkman, A. Szold, A. Forgione, H. W. R. Schreuder, E. C. J. Consten, W. A. Draaisma, P. M. Verheijen, J. P. Ruurda and Y. Kaufman, “First Experience with the AutoLapTM System: an Image-Based Robotic Camera Steering Device”, Surg Endosc, vol. 32, no. 5, pp. 2560-2566, May 2018.

[21] A. Agustinos, R. Wolf, J. A. Long, P. Cinquin and S. Voros, “Visual Servoing of a Robotic Endoscope Holder Based on Surgical Instrument Tracking”, presented at IEEE RAS EMBS Int Conf Biomed Robot Biomechatron, Sao Paulo, Brazil, Aug. 2014.

[22] S. Eslamian, L. A. Reisner and A. K. Pandya, “Development and Evaluation of an Autonomous Camera Control Algorithm on the Da Vinci Surgical System”, Int J Med Robotics Comput Assist Surg, vol. 2036, pp. 1-16, Sep 2019.

[23] X. Ma, C. Song, P. W. Chiu and Z. Li, “Autonomous Flexible Endoscope for Minimally Invasive Surgery with Enhanced Safety”, Robot Autom Letter, vol. 4, no. 3, pp. 2607 - 2613, Jan 2019.

[24] X. Ma, C. Song, X. Xia, P. W. Chiu and Z. Li, “Visual Servo of a 6-DOF Robotic Stereo Flexible Endoscope Based on da Vincix Research Kit (dVRK) System”, IEEE Robot Autom Lett, vol. 5, no. 2, pp. 820 - 827, Apr 2020.

[25] C. Song , X. Ma, P. W. Chiu, C. C. Chong and Z. Li,“A robotic flexible endoscope with shared autonomy: a study of mockup cholecystectomy”, Surg Endosc, pp. 1-12, Oct 2019.

[26] J. H. Goldberg and J. C. Schryver., “Eye-Gaze-Contingent Control of the Computer Interface: Methodology and Example for Zoom Detection”, Behav Res Methods Instrum Comput, vol. 27, no. 3, pp. 338–350, Sep. 1995.

[27] D. W. Hansen, H. H. T. Skovsgaard, J. P. Hansen, and E. Mollenbach, “Noise Tolerant Selection by Gaze-Controlled Pan and Zoom in 3D”, Eye Track Res app, pp. 205–212, Mar. 2008.

[28] K. W. Kwok, L. W. Sun, G. P. Mylonas, D. R. C. James, F. Orihuela Espina and G. Z. Yang, “Collaborative Gaze Channelling for Improved Cooperation During Robotic Assisted Surgery”, Ann Biomed Eng, vol. 40, no. 10, pp. 2156–2167, Oct. 2012.

[29] M. Kranzfelder, C. Staub, A. Fiolka, A. Schneider, S. Gillen, D. Wilhelm, H. Friess, A. Knoll, H. Feussner, “Toward Increased Autonomy In The Surgical OR: Needs, Requests, and Expectations”, Surg Endosc, vol. 27, no. 5, pp. 1681-1688, May 2013.

[30] A. Mariani, G. Colaci, T. Da Col, N. Sanna, E. Vendrame, A. Menciassi and E. De Momi, “An Experimental Comparison Towards Autonomous Camera Navigation to Optimize Training in Robot Assisted Surgery”, Robot Autom Letter, vol. 5, no. 2, pp. 1461–1467, Jan. 2020.

[31] 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 Int Conf Robot Autom, pp. 6434–6439, Jun. 2014.

[32] Z. Chen, A. Deguet, R. H. Taylor, and P. Kazanzides, “Software Architecture of the da Vinci Research Kit”, IEEE Int Conf Robot Comp, pp. 6434–6439, Apr. 2017.

[33] R. Campo, C. Reising, Y. V. Belle, J. Nassif, P. O’Donovan and C. R. Molina, “A Valid Model for Testing and Training Laparoscopic Psychomotor Skills”, Surg Endosc, vol. 7, pp. 133–141, Mar. 2010.

[34] R. Campo, Wattiez, A., Tanos V., A. Di Spiezio Sardo and Grimbis G. “Gynaecological Endoscopic Surgical Education and Assessment. A Diploma Programme in Gynaecological Endoscopic Surgery”, Eur J Obstet Gynecol Reprod Biol, vol. 199, pp. 183–186, Feb. 2016.

[35] A. M. Jarc and M.J. Curet, “Viewpoint Matters: Objective Performance Metrics for Surgeon Endoscope Control during Robot-Assisted Surgery”, Surg Endosc, vol. 31, no. 3, pp. 1192-1202, Mar. 2017.