HAL
open science

Visual servoing of a robotic endoscope holder based on surgical instrument tracking

Anthony Agustinos, Jean-Alexandre Long, Philippe Cinquin, Rémi Wolf, Sandrine Voros

To cite this version:

Anthony Agustinos, Jean-Alexandre Long, Philippe Cinquin, Rémi Wolf, Sandrine Voros. Visual servoing of a robotic endoscope holder based on surgical instrument tracking. Biomedical Robotics and Biomechatronics, Aug 2014, Sao Paulo, Brazil. pp.13 - 18, 10.1109/BIOROB.2014.6913744 . hal-01164895

HAL Id: hal-01164895

https://hal.science/hal-01164895v1

Submitted on 9 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Visual Servoing of a robotic endoscope holder based on surgical instrument tracking*

A. Agustinos[1], R. Wolf[1], J. A. Long[2], P. Cinquin[1,3], S. Voros[4]

***Abstract*— We propose an image-based control for a robotic endoscope holder during laparoscopic surgery. Our aim is to provide more comfort to the practitioner during surgery by automatically positioning the endoscope at his request. To do so, we propose to maintain one or more instruments roughly at the center of the laparoscopic image through different command modes. The originality of this method relies on the direct use of the endoscopic image and the absence of artificial markers added to the instruments. The application is validated on a test bench with a commercial robotic endoscope holder.**

I. INTRODUCTION

Laparoscopic surgery is a minimally invasive technique which accurately reproduces the principles of conventional surgery with minimal physical trauma. Surgeons can perform an operation on the abdomen through small incisions in which trocars are positioned, allowing for the insertion of surgical instruments and the endoscope. Compared to conventional surgery, laparoscopy offers many advantages for the patient: the significant reduction of the size of incisions allows a decrease in intra-operative bleeding, risk of infection, post-operative pain, and duration of hospitalization. In return, laproscopic surgery is much more complex to perform than open surgery: the constraints are mostly ergonomic [1], with a reduction of instrument mobility due to fixed insertion points on the abdominal cavity, a loss of tactile sense, a limited field of view and a need for a good coordination of the surgeon and the assistant manipulating the instrument. In this paper, we focus on the last issue: the surgeon no longer has direct control of his vision, which can disrupt the hand-eye coordination and requires perfect teamwork with the assistant. Several studies have shown that precision handling of the endoscope by the assistant during long operations can be degraded over time with tremors and contacts with the surgical site [2], [3].

To overcome the challenges related to the manual manipulation of the endoscope and potentially eliminate the need for an assistant during laparoscopic surgery, this task can be entrusted to a **robotic endoscope holder** (**REH**). The first REH AESOP® [4] and its successors have improved precision and stability of the endoscopic image. However, these robots are often bulky and have basic commands (left, right, up, down, zoom in and zoom out). Several works have been conducted to automate the control of REH by tracking surgical instruments. The aim is to avoid decomposition of the robot's displacements by a series of orders and to have fluid and rectilinear trajectories in the endoscopic images. The majority of these works uses color-based approaches (e.g [5], [6]) where surgical instruments are detected using the color information and often artificial markers in the images. Several works have also been carried out on the semi-automatic manipulation of instruments (e.g [7]). However, unlike automated movement of REH, the automatic navigation of surgical instruments by a **robot instrument holder** (**RIH**) into the abdominal cavity remains dangerous. In [7], a laser is projected onto organs to determine the distance and orientation of the instrument. This allows to position automatically the instrument in safety. Tele-manipulated systems such as the DaVinci® [8] also allow to move instrument with a master/slave control. In this paper, we also demonstrate the feasibility of controlling a RIH that is not rigidly linked to a REH. We propose a semi-autonomous control of a lightweight REH able to follow standard surgical instruments and keep them centered in the endoscopic images. Our method for the localization of surgical instruments [9] is briefly presented Section III.A. A vision sensor based control was established to control the REH (Section III.B) and track instruments with several control modes in the image (Section III.C).

II. MATERIAL AND SYSTEM CALIBRATION

A. Material

The REH used in this study is the lightweight, body mounted ViKY® [10]. More than 300 surgeries have been performed worldwide with the system since its commercialization. It has three degrees of freedom (DoF) (shown in Fig.1(a)) with three motors, each corresponding to one DoF. Two control interfaces are available for its clinical use: a voice command and a control footpad. The camera used is an OLYMPUS OTV600 CCD (Charge Coupled Device) designed for laparoscopy (resolution of 470,000 pixels). The endoscope used is an OLYMPUS WA53000A. We digitize the laparoscopic images using an external acquisition card from IC Imaging Source, which acquires images at a resolution of 720x480 pixels and a frequency of 25 Hz.

*This work is partially funded by the French governement's
- "Agence Nationale de la Recheche, TecSan" program, through the ANR TESCAN 2009 DEPORRA (Dispositif et systèmes ciblés pour la PrOstatectomie Radicale Robotisée Augmentée) project.
- "Fonds Unique Interministèriel (FUI)" program, through the Fluoromis II project.

[1]A. Agustinos and R. Wolf are with the UJF-Grenoble 1 / CNRS, TIMC-IMAG UMR 5525, Grenoble, F-38041, France.
[2]J. A. Long is with the Department of Urology, University Hospital, Grenoble, F-38041, France.
[3]P. Cinquin is with the Centre d'Investigation Clinique - Innovation Technologique, INSERM, CHU de Grenoble, UJF-Grenoble 1, CIT803, Grenoble, F-38041, France.
[4]S. Voros is with the UJF-Grenoble 1 / CNRS / INSERM, TIMC-IMAG UMR 5525, Grenoble, F-38041, France.

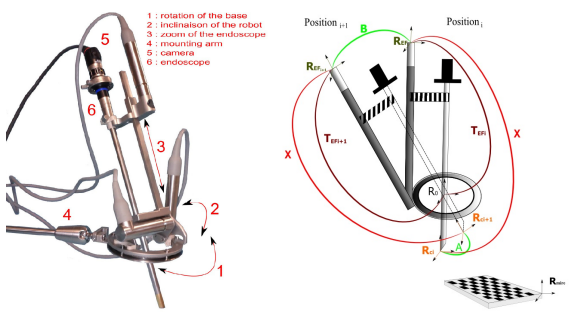
B. System calibration

To combine measures in the image to the reference frame of the robot, several calibration steps are necessary.

1) Endoscopic camera calibration: Our method for the 3D localization and orientation of an instrument in the camera frame (see Section III.A) requires an intrinsic calibration. It is performed using Zhang's procedure [11], which has been extensively validated by the community. This intrinsic calibration allows computing the 3D view line in the camera frame which corresponds to a 2D point in the image.

2) Modeling of the robot ViKY®: For the modeling of the robot, the classical approach of Denavit-Hartenberg [12] has been used. This permits determining the direct geometric model ${}_0\mathbf{T}^{\mathrm{EF}}$ of the robot, i.e the rigid transformation between the robot's reference frame R_0 and the end-effector frame R_{EF} (see Fig.1(a)).

3) Hand-Eye calibration: In order to link image points to the reference frame of the robot, we need to solve a hand-eye calibration problem. This involves estimating the rigid transformation (rotation and translation) $\mathbf{X}$ between the camera frame R_c and the end-effector frame R_{EF} of the robot (see Fig.1(b)).



(a) ViKY® (b) HE calibration

Fig. 1: a) The ViKY® robot endoscope holder, b) Geometric relationship $\mathbf{X}$ between the camera frame R_c and the end-effector frame R_{EF} of the robot.

The HE calibration is performed by measuring displacements of the system {end-effector/camera}, as shown in Fig.1(b), where $\mathbf{A}$ is the rigid transformation resulting from the camera motion, $\mathbf{B}$ is the rigid transformation resulting from the motion of the end-effector of the robot and $\mathbf{X}$ is the unknown rigid transformation between the camera frame and the end-effector frame. Finding $\mathbf{X}$ corresponds to solving the system $\mathbf{AX} = \mathbf{XB}$. This system is solved using Tsai's approach [13], which consists in taking multiple shots of a calibration chessboard for a series of movements $\mathbf{N}$ of the robot (where $N > 3$). We have a sterilizable calibration grid at our disposal that can be used in surgical conditions.

III. METHOD

In this section, we present a few command modes to control the ViKY® based on the visual servoing of surgical instruments. In III.A. we briefly present the instrument's localization method and detection of the instrument tip. In III.B. we present our visual servoing control, that exploits the calibration step presented in the previous section. In III.C. we present the different commands that we implemented.

A. 3D localization of surgical instruments

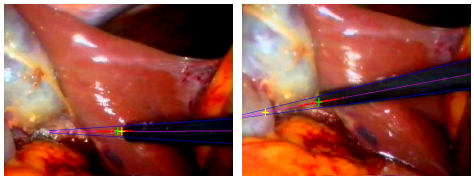
For clarity, the proposed method for the 3D localization of surgical instruments based on the analysis of 2D laparoscopic images is briefly described here. The reader may refer to [9] for a more detailed description of each step of the method.

- An **initialization step** consists in locating the 3D insertion point of each instrument in the patient's abdominal cavity, providing strong constraints for the localization of the instruments in the laparoscopic images.
- A **3D geometrical instrument model** (a cylinder of known diameter and length) represents the instrument in 3D. All the possible orientations of the instrument inserted through an insertion point I are represented by a geode centered on I. The geode is decomposed in cells, on which particles are sampled corresponding to candidate locations of the instrument.
- Based on this model, and the camera calibration, the **3D axis of the instrument** in the camera frame is found using the CondenSation algorithm [14]: particles are sampled on the geode surface randomly and converge to the geode cell that corresponds to the instrument's orientation, based on image measurements. As we will show in the results section, the choice of the number of particles is important.
- Finally, once the 3D orientation of the instrument in the camera frame is known, the camera calibration is used to project the 3D axis in the laparoscopic image to obtain the 2D axis of the instrument. The **position of the instrument's tip** is searched along this 2D axis. In this step the CondenSation algorithm is also used.

It should be noted that what we call the "instrument tip" detection, is in fact a detection of the end of the tool shaft. However, since we detect with our method the instrument's axis in 3D and since the laparoscopic tools have normalized sizes, it is very easy to compute the actual tool tip by adding an offset length to the detected shaft/tip interface, along the 3D axis of the instrument. One advantage of our method is that if the tip is obscured by surgical stage, the target will still match the visible end of the tool. This means that the instrument is not "lost", which is not the case when artificial markers are used and are hidden by overlapping structures (see Fig.2(b)).

B. 2D visual servoing control of the robot endoscope holder

For our application, we want to use information provided by a vision sensor to control the movement of the REH. Our aim is to control the movement of the camera to keep the tip of surgical instruments at the center of the laparoscopic image. We accomplish this by minimizing the error e, between a desired state of the target s^* (the image center), and its current state s (the tool tip position in the image). To do so, we chose to use a state-of-the-art 2D visual servoing control approach (image-based control) [15].



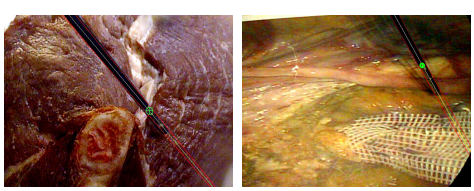
(a) Visible tip (b) Occluded tip

Fig. 2: (a) Typical example of a tool detection: projection of the 3D tool's axis (pink line), borders in the image (blue lines) and tool tip (green dot).(b) When the tool tip is invisible, the green tip corresponds to the tip dectected by our method the yellow tip correponds to the probable position of the real tool tip).

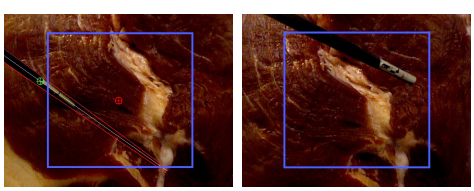
C. Proposed command modes

In this section, we present four simple possible tracking modes that all use as input the instrument's localization method and robot control presented in the previous section. It should be noted that several studies have been conducted on the recognition of a surgical step [16], [17] in laparoscopic surgery. We thus consider that, in term, it will be possible to automatically choose between several tracking modes according to the surgical step.

1) Tracking of a single instrument: When only one instrument is present in the image, we follow its tip as determined in Section III.A. We consider two tracking modes, namely direct and "degraded" tracking. For the direct tracking mode, the tip of the instrument is tracked continuously (see Fig.3(a)). For the "degraded" tracking mode, the tip of the instrument is only tracked if it exits in a zone centered in the image (see Fig.3(b)). The direct mode tracking could be interesting for the exploration of the abdominal cavity and the degraded mode for a suturing gesture where the surgeon might need a stable image except if the instrument leaves the field of view.



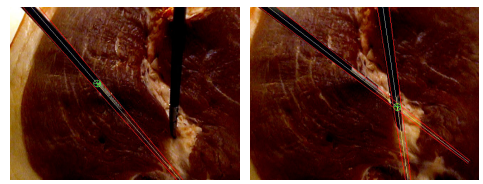
(a) Direct mode tracking



(b) Degraded mode tracking

Fig. 3: Detection of a single instrument in the field of view. The tip of the instrument is represented by a green point. The red lines and the yellow axis correspond to the borders of the instrument and its axis. In (b), the instrument is not tracked inside the blue rectangle while in (a) tracking is continuous.

2) Tracking of several instruments: When two instruments are present in the image, the instrument to track can be selected by identifying its insertion point (see Fig.4(a)). The single instrument tracking could allow the surgeon to choose during the surgery which instrument is more suitable to guide the camera. Let us note that the tracking of the intersection of the instruments is also a possibility (see Fig.4(b)), but we must still study its interest and feasibility in clinical practice.



(a) Single instrument (b) Intersection mode

Fig. 4: Detection of two instruments in the field of view. The chose primitive is represented by green point.

D. 3D insertion point update

The localization method presented in [9], makes the assumption that the camera is fixed. However, during the visual servoing, the camera moves and the coordinates of the insertion point for a given position of the camera are no longer correct and must be updated. The insertion point, determined in the initialization step (Section III.A.) for a reference position of the camera, can easily be updated using the Hand-Eye matrix $\mathbf{X}$ and the geometric model ${}_{\mathbf{0}}\mathbf{T}^{\mathbf{EF}}$ of the robot (Fig.5):

$$\mathbf{P_{i+1}} = \mathbf{X}^{-1}{}_{\mathbf{0}}\mathbf{T}^{\mathbf{EF_i}^{-1}}\ {}_{\mathbf{0}}\mathbf{T}^{\mathbf{EF_{i+1}}}\mathbf{X}\mathbf{P_i} \tag{1}$$

Here, $\mathbf{P_i}$ and $\mathbf{P_{i+1}}$ are the insertion points for the positions i, $i+1$ respectively, in the mobile camera frame. To compensate for the errors in the computation of the new insertion point that can be due to an imperfect geometrical model of the robot, an imperfect Hand-Eye calibration and small movements of the insertion point, we add Gaussian noise to the position of the insertion point. This noise is a 3x1 vector in which each component follows a normal distribution $\mathbf{N}(0,5)$, the standard deviation of the displacements being selected from previous work [18]. This allows us to vary randomly the position of the insertion point in 3D space at each iteration of the instrument's localization algorithm. Hence, the algorithm can converge to the couple {insertion point, instrument's orientation} corresponding to the "best" detection.

E. Towards the control of a robot instrument holder

An extension of the work presented above is to position a RIH towards a target at a desired orientation and depth. To do so we use our instruments localization method to find the geometric transformation between the REH and the RIH. Then the robot's geometric models are sufficient to localize the tip of an instrument manipulated by the RIH. Compared to the REH control, we have a deported camera and we must control three DoFs (the RIH's orientation and depth). To

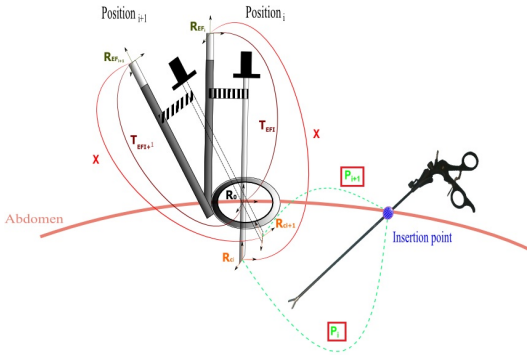


Fig. 5: Determination of the insertion point for two successive positions of the robot.

demonstrate the feasibility of this approach, we use a second ViKY® mounted as IH. However, the automatic navigation of an instrument in the abdominal cavity can be dangerous so ultimately, the best solution would be to use this approach with a co-manipulated RIH.

The control of the RIH, requires us to know the geometric transformation **T** (rotation and translation) between the camera frame R_c of the RIH and the reference frame R_0 of the RIH (Fig.6). The relationship between a 3D point $\mathbf{P_0}$ in the reference frame of the robot and the same 3D point in the camera's referential $\mathbf{P_c}$ can be expressed as :

$$\begin{pmatrix} \mathbf{P_c} \\ 1 \end{pmatrix} = \begin{pmatrix} {}_{\mathbf{c}}\mathbf{R}^{\mathbf{o}}_{\mathbf{3x3}} & {}_{\mathbf{c}}\mathbf{t}^{\mathbf{o}}_{\mathbf{1x3}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_0 \\ 1 \end{pmatrix} \tag{2}$$

where ${}_{\mathbf{c}}\mathbf{R}^{\mathbf{o}}$ is the rotation and ${}_{\mathbf{c}}\mathbf{t}^{\mathbf{o}}$ the translation matrix of the transformation **T**. The translation ${}_{\mathbf{c}}\mathbf{t}^{\mathbf{o}}$ is determined

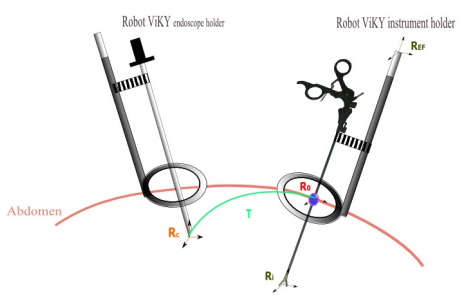


Fig. 6: Geometric relationship **T** between the camera frame R_c of the REH and the reference frame R_0 of the RIH

by finding the insertion point of the RIH using the same method as the one used to find the insertion point of an instrument (initialization step of Section III.A.). The rotation ${}_{\mathbf{c}}\mathbf{R}^{\mathbf{o}}$ between the camera and the reference frame of the RIH is then determined by measuring 3D coordinates of the instrument's tip in the frames R_0 and R_c for **N** displacements of the robot. Due to our localization method, we can compute the 3D position of the instrument in the camera frame $\mathbf{P_c}$. Thanks to the geometric model of the RIH, we can determine the 3D position of the instrument in the RIH's frame P_0:

$$\text{for } 1 < i < N,\ \mathbf{P_{ci}} - {}_{\mathbf{c}}\,\mathbf{t}^{\mathbf{o}} = {}_{\mathbf{c}}\,\mathbf{R}^{\mathbf{0}}\mathbf{P_{0i}},\ \text{where } N > 3 \tag{3}$$

We can determine ${}_{\mathbf{c}}\mathbf{R}^{\mathbf{o}}$ by solving the linear system (5) using a SVD decomposition coupled by a RANSAC [19] to eliminate the outliers.

IV. RESULTS

We performed several experiments on a test bench consisting of a surgery trainer box on which the robot ViKY® is directly positioned, and a piece of meat as background. For the computations, we used an Intel Xeon PC 2.67 GHz, 3.48 GB RAM. To calibrate the camera, we used a 4x7 planar chessboard with 7 mm square size. The calibration procedure involved taking 20 images of the chessboard pattern for different orientations and depths covering the entire work area. For the Hand-Eye calibration, a series of 12 robot displacements for which the calibration chessboard was visible was automatically performed. As described in Section II.B, we solved the system $\mathbf{AX} = \mathbf{XB}$ using the measured robot displacements and computed calibration chessboard displacements. To validate the whole calibration process, we computed the average reprojection error for the image set of the Hand-Eye calibration. An example of five camera calibrations and Hand-Eye calibrations are shown in Table I.

TABLE I: Result of camera and Hand-Eye calibrations

RMS error intrinsic calibration (pixel)	RMS error hand-eye calibration (pixel)
0.398	9.3
0.34	7.8
0.359	10.1
0.295	9.8
0.354	8.2

We deem the camera calibration is accurate in the sense that all calibrations exhibit sub-pixel errors. The Hand-Eye calibrations have a maximum reprojection error of 10.1 pixels. We consider that an error detection of the instrument tip of 10 pixels, is sufficient for our application. Indeed, the automatic positioning of the REH does not require sub-pixel precision. Moreover, a 2D visual servoing is robust to calibration errors [15]. Finally, as we will see in Section IV.C. our localization method allows us to compensate for calibration errors thanks to the random noise that we added to the insertion point.

A. Localization method: precision and computation time

The compromise between computation time and precision plays a significant role in our application. This compromise is essentially determined by the number of particles used. In Table II, we have listed the computation time and associated error in relation to the number of particles used to detect the instrument's 3D orientation.

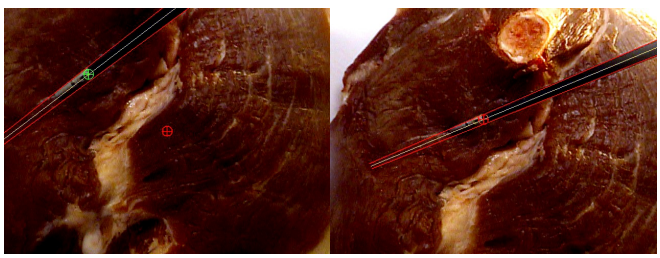
From our experiments, we deem that the optimal trade-off between speed and accuracy is reached when using 1000 particles.

TABLE II: Precision and computation time of the localization method

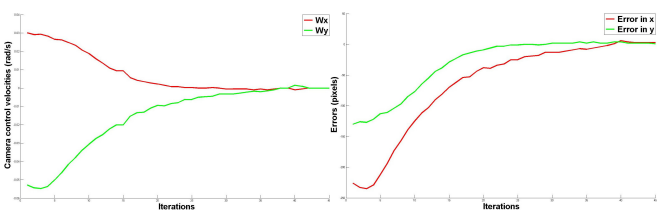
Number of particles	Frequency (Hz)	2D Angular error
500	16.1	1.22
1000	12.8	0.74
2000	10.6	0.72
5000	7.4	0.56

B. Single instrument tracking

Fig.7 shows sample images acquired during the execution of the 2D visual servoing to track an instrument tip in the image. In Fig.7, the red circle represents the center of the image and the desired position of the tip and the green circle is the current position of the tip. Fig.7(a), (b), show the evolution of the scene during positioning task. (c) shows the 3D camera velocity (rad/s) and (d) the error $s - s^*$ (pixel).



(a) Initial image (b) Final image



(c) Camera control velocities (d) Errors in the image

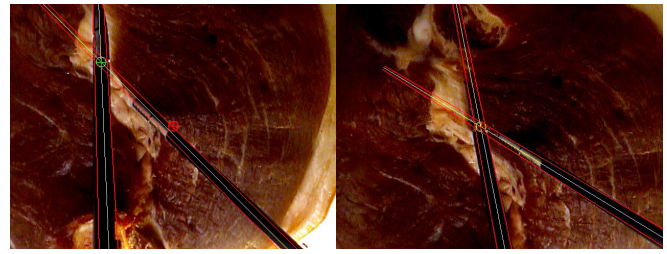
Fig. 7: Results the tracking of single instrument.

For this experiment, λ was empirically set to 0.25 which permits the choice of the convergence velocity. After about 45 iterations, about 4 seconds (at a frequency of 12 Hz), the error on each coordinate was less than 0.5 pixel.

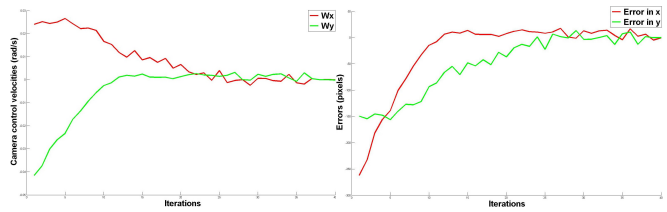
C. Tracking of two instruments

Fig.8 shows the images acquired during the execution of the 2D visual servoing to track the intersection of two instruments.

For this experiment, λ was also set to 0.25. After approximately 40 iterations about 6 sec (at a frequency of 7 Hz), the error on each coordinate is inferior to 1 pixel. It should be noted that small variations in the detection of two instruments in the image involve an oscillation of the point corresponding to their intersection. This results in oscillations in the errors curves in the image and control camera.



(a) Initial image (b) Final image



(c) Camera control velocities (d) Errors in the image

Fig. 8: Results the tracking of single instrument.

D. Accuracy of the insertion point update

We first performed an experiment to evaluate the errors induced by the computation of an insertion point when the robot is fixed at the reference position: we computed 10 times the insertion point for the reference position of the REH using the initialization step of the localization's method. We obtained a mean error of 6 mm relative to the gravity center of the 10 insertion points.

In a second step, we estimated the precision of the update of the insertion point using the HE calibration (Section III.C.), with and without the addition of Gaussian noise.

Our procedure consisted of :

(a) computing the position of the insertion point for a reference robot position, as given by the insertion point initialization method of Section III.A. (ground truth)
(b) moving the REH to 10 different postitions
(c) computing the insertion point using the inialization method for each "new" robot position
(d) moving back the REH to the reference position
(e) estimating the position of the 10 insertion points using the HE calibration and robot model, as given by equation (10)
(f) estimating the positions of the 10 insertion points of e) with addition of Gaussian noise in the tools localization method

These computations are illustrated in Fig.9: the red dot corresponds to the insertion point of (a). The blue dots correspond to the 10 insertion points computed using (e) and the green dots correspond to the 10 insertion points computed in (f) where Gaussian noise was included.

Fig.9 shows that the addition of Gaussian noise increases the accuracy of the insertion point estimation compared to the imperfect Hand-Eye matrix and robot model alone. Indeed, the mean error for the insertion point measurements with Gaussian noise compared to the ground truth was evaluated to 5.25 mm, compared to 38.98 mm for the insertion point measurements without Gaussian noise.

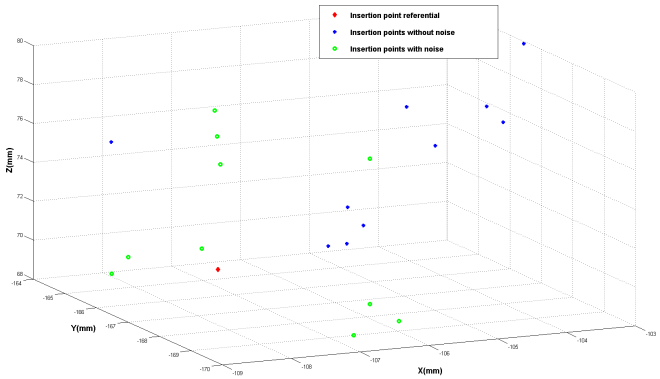


Fig. 9: Comparison of the new insertion points computations, with and without noise.

E. Accuracy of the geometric transformation between the camera and the RIH

To evaluate the accuracy of the calibration between RIH and the REH, we have developed a simple command of the RIH. This command consists of controlling the movement of the instrument so as to position its tip at a given position in the image with a fixed insertion depth. This servoing is only based on the calibration and the REH and RIH geometric models. A printout of a surgical image has been used for the background. The evaluation consists of:

(a) computing the 3D instrument's tip ($\mathbf{P_0}$) in the RIH reference frame
(b) estimating this point in the camera's frame ($\mathbf{P_c}$) using the rigid transformation $\mathbf{T}$ defined in Section III.E
(c) projecting ($\mathbf{P_c}$) in the image thanks to the camera's model
(d) comparing it to the manual identification of the tool's tip in the image.

The calibration between the camera and the RIH is relatively accurate with a mean error of 13 pixels on 30 images between the estimated and the real position's tip.

V. CONCLUSION /FUTURE WORKS

In this paper, we have presented several control modes of a robotic endoscope holder, using an image-analysis based method for the instruments localization that does not require artificial markers. We have shown that it is possible to minimize the errors of calibration and localization of the instruments using Gaussian noise around the insertion point. We have also demonstrated that it is possible to use the instrument's localization method to estimate the transformation between a robot endoscope holder and a robot instrument holder and to control the positioning of an instrument.

In future works, these methods should be evaluated in conditions closer to the clinical reality (cadaver experiments), in order to evaluate the feasibility of the whole process in more realistic conditions. This will also allow us to work on the surgeon/robot interface using, for example, footpad or voice command to start or stop control modes of robots. Using GPU programmaing could be an interesting way to improve the computation time of our method and thus gain precision. In the case where a REH and a RIH collaborate, the real-time image based localization of instruments would be unnecessary, thanks to the calibration and robots modelling. However, it could be useful as a background task for online recalibration.

This opens the possibility to have a non-rigidly lightweight robotic environment allowing a cooperation between a robot endoscope holder and a robot instrument holder.

REFERENCES

[1] R. Agha, BSc and G. Muir, "Does laparoscopic surgery spell the end of the open surgeon?" in Journal of the Royal Society of Medecine, vol. 96, Nov. 2003, pp. 544-546.

[2] M. Winkler, S. Erbse, K. Radermacher, G. Rau and W. Rath, " An automatic camera-holding system for gynecologic laparoscopy" in Journal of the American Association of Gynecologic Laparoscopists, vol. 8, May 2001, pp. 303-306.

[3] K. T. den Boer, J. Dankelman, D. J. Gouma and H. G. Stassen, "Peroperative analysis of the surgical procedure" in Journal of the Surgical Endoscopy and Other Interventional Techniques, vol. 16, March 2002, pp. 492-499.

[4] J. Sackier and Y. Wang, "Robotically assisted laparoscopic surgery" in Journal of the Surgical Endoscopy, vol. 8, Janv. 1994, pp. 63-66.

[5] A. Casals, J. Amat and E. Laporte, "Automatic guidance of an assistant robot in laparoscopic surgery" in Proc. IEEE int. conf. Robot. Autom., Apr. 1996, pp. 22-28.

[6] M. Groeger, K. Arbter and G. Hizinger, "Motion tracking for minimally invasive robotic surgery" in Medical Robotics, I-Tech Education and Publishing, Jan. 2008, pp. 117-148.

[7] K. Krupa, J. Gangloff, C. Doignon, M.F. de Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux, "Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing" in IEEE trans. Robotics and Automation, vol.19, Oct. 2003, pp. 842-853.

[8] I. A. M. J. Broeders and J. Ruurda, "Robotics revolutionizing surgery: the Intuitive Surgical "Da Vinci" system" in Industrial Robot: An International Journal, vol. 28, 2001, pp. 387-391.

[9] R. Wolf, J. Duchateau, P. Cinquin and S. Voros, "3D tracking of laparoscopic instruments using statistical and geometric modeling" in MICCAI, vol. 6891, 2011, pp. 203-210.

[10] J. A. Long, P. Cinquin, J. Troccaz, S. Voros, P. Berkelman, J. L. Descotes, C. Letoublon and J. J. Rambeaud, "Development of Miniaturized Light Endoscope-Holder Robot for Laparoscopic Surgery", in Journal of Endourology, vol. 21, Aug. 2007, pp. 911-914.

[11] Z. Zhang, "A flexible new technique for camera calibration", IEEE Trans. on Pattern Analysis Machine Intelligence, vol. 22, 2000, pp. 1330-1334.

[12] J. Denavit and R. S. Hartenberg, "A kinematic notation for lower-pair mechanisms based on matrices" in Journal of Appl. Mechanics, vol.22, June 1955, pp. 215-221.

[13] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration" in IEEE trans. on Robotics and Automation, vol.5, 1989, pp. 345-358.

[14] M. Isard and A. Blake, "Condensation-conditional density propagation for visual tracking" in int. Journal of computer vision, vol.29, 1998, pp. 5-28.

[15] F. Chaumette, and S. Hutchinson, "Visual servo control. I. Basic approaches" in IEEE Mag. Robotics & Automation, vol.13, Dec. 2006, pp. 82-90.

[16] N. Padoy, T. Blum, S. A. Ahmadi, H. Feussner, M. O. Berger and N. Navab, "Statistical modeling and recognition of surgical workflow" in Medical Image Analysis, vol.16, 2012, pp. 632-641.

[17] K. Seong-Young, J. Kim, K. Dong-Soo and L. Woo-Jung, "Intelligent interaction between surgeon and laparoscopic assistant robot system" in IEEE int. Workshop on Robot and Human Interactive Communication, Aug. 2005, pp. 60-65.

[18] S. Voros, J. A. Long and P. Cinquin, "Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder" in MICCAI, vol. 4190, Jan. 2006, pp. 535-542.

[19] M. A. Fischler and R. C. Bolles, "Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography" in Proc. Image Understanding Workshop, Apr. 1980, pp. 71-88.