

AUTOMATIC POSITIONING OF SURGICAL INSTRUMENTS DURING LAPAROSCOPIC SURGERY WITH ROBOTS USING AUTOMATIC VISUAL FEEDBACK ^{*,**}

A. KRUPA¹, M. DE MATHELIN¹, C. DOIGNON¹, J. GANGLOFF¹, G. MOREL², L. SOLER³,
J. LEROY³ AND J. MARESCAUX³

Abstract. This paper presents a robotic vision system that automatically retrieves and positions surgical instruments in the robotized laparoscopic surgical environment. The goal of the automated task is to bring the instrument from an unknown or hidden position to a desired location specified by the surgeon on the endoscopic image. To achieve this task, a special instrument-holder is designed that projects laser dot patterns onto the organ surface which are seen on the endoscopic images. Then, the surgical instrument positioning is done using an automatic visual servoing from the endoscopic image. Our approach is successfully validated in a real surgical environment by performing experiments on living pigs in the surgical training room of IRCAD.

INTRODUCTION

Robotic laparoscopic systems have recently appeared, *e.g.*, ZEUS (Computer Motion, Inc.) or DaVinci (Intuitive Surgical, Inc.). With these systems, robot arms are used to manipulate the surgical instruments as well as the endoscope. The surgeon tele-operates the robot through master arms using the visual feedback from the laparoscope. This reduces the surgeon tiredness, and potentially increases motion accuracy by the use of a high master/slave motion ratio. Furthermore, teleoperation allows to perform long distance surgical procedures (see *e.g.* [1]). In a typical surgical procedure, the surgeon first drives the laparoscope into a region of interest (for example by voice, with the AESOP system of Computer Motion Inc.). Then, he or she drives the surgical instruments at the operating position. A practical difficulty lies in the fact that the instruments are generally not in the field of view at the start of the procedure. Therefore, the surgeon must either blindly move the instruments into the field of view of the endoscope, or zoom out with the endoscope as to get a larger view. Similarly, when the surgeon zoom in or moves the endoscope during surgery, the instruments may leave the endoscope's field of view. Consequently, instruments may have to be moved blindly with a risk of an undesirable contact between instruments and internal organs. Therefore, in order to assist the surgeon, we propose a visual

* The financial support of the french ministry of research ("ACI jeunes chercheurs" program) is gratefully acknowledged.

** The experimental part of this work has been made possible thanks to the collaboration of Computer Motion Inc. that has graciously provided the medical robots. In particular, we would like to thank Dr. Moji Ghodoussi for his technical support in the use of the medical robots.

¹ Strasbourg I University - LSIIT (UMR CNRS 7005) - ENSPS, Bd. S. Brant, 67400 Illkirch, France.
e-mail: alexandre.krupa@ensps.u-strasbg.fr

² Now with the LRP6 (CNRS FRE 2705), Paris 6 University, France.

³ Institut de Recherche sur le Cancer de l'Appareil Digestif), Strasbourg, France.

servoing system that automatically brings the instrument at the center of the endoscopic image (see, [2] for earlier works in that direction). Furthermore, when the instrument is in the field of view, the system provides a measurement of the distance between the organ and the tip of the instrument. Our system also allows to move automatically the instruments at a 3-D location chosen by the surgeon. The surgeon specifies the position in the image (with, *e.g.*, a touch-screen or a mouse-type device) and sets the distance value between the instrument and the organ at his or her convenience. Note that prior research was conducted on visual servoing techniques in laparoscopic surgery in, *e.g.*, [3], [4] and [5]. However, these systems were guiding the endoscope toward the region of interest instead of the instruments.

Our system includes a special device designed to hold the surgical instruments with tiny laser pointers. This laser pointing instrument-holder is used to project laser spots in the laparoscopic image even if the surgical instrument is not in the field of view. The image of the projected laser spots is used to automatically guide the instrument in the field of view. A major difficulty in designing this automatic instrument positioning system lies in the unknown relative position between the endoscope and the robot holding the instrument, and in the use of monocular vision that lacks depth information. In [6], the authors added an extra laser-pointing endoscope with an optical galvano scanner and a 955 *fps* high-speed camera to estimate the 3-D surface of the scanned organ. However, their method requires to know the relative position between the laser-pointing endoscope and the camera which is estimated by adding an external camera watching the whole system (OPTOTRAK system). In our approach, instead, we robustly estimate in real time the distance between instrument and organ by using images of optical markers mounted on the tip of the instrument and images of the laser spots projected by the same instrument. Then, visual feedback allows to automatically position the instrument at a desired location. The paper is organized as follows. Section 1 describes the system configuration and the endoscopic laser pointing instrument prototype. In section 2, robust image processings leading to laser patterns and optical markers detection are briefly explained. The control scheme used to position the instrument by automatic visual feedback is described in section 3. In the final section, we present experimental results in real surgical conditions at the operating room of IRCAD.

1. SYSTEM DESCRIPTION

1.1. System configuration

The system configuration used to perform the autonomous positioning of the surgical instrument is shown in Figure 1.a. The system includes a laparoscopic surgical robot and an endoscopic laser pointing instrument-holder. The robot allows to move the instrument across a trocar placed at a first incision point. A second trocar is put in order to insert an endoscopic optical lens which provides the feedback view and whose relative position to the robot base frame is generally unknown.

1.2. Endoscopic laser pointing instrument-holder

We have designed the prototype of an endoscopic laser pointing instrument-holder shown in figure 1.b. It is a 30 *cm* long metallic pipe with 10 *mm* external diameter to be inserted through a 12 *mm* standard trocar. Its internal diameter is 5 *mm* so that a standard laparoscopic surgical instrument can fit inside. The head of the instrument-holder contains miniature laser collimators connected to optical fibers which are linked to external controlled laser sources. This device allows to remote laser sources which can not be integrated in the head of the instrument due to their size. The laser pointing instrument-holder mounted on a cleaning-suction instrument is shown in figure 1.c. The instrument-holder, with the surgical instrument inside, is held by the end-effector of the robot. It allows to project optical markers on the surgical scene. Three tiny light emitting diodes are fixed on the tip of the instrument to provide the optical markers needed to compute the depth information.

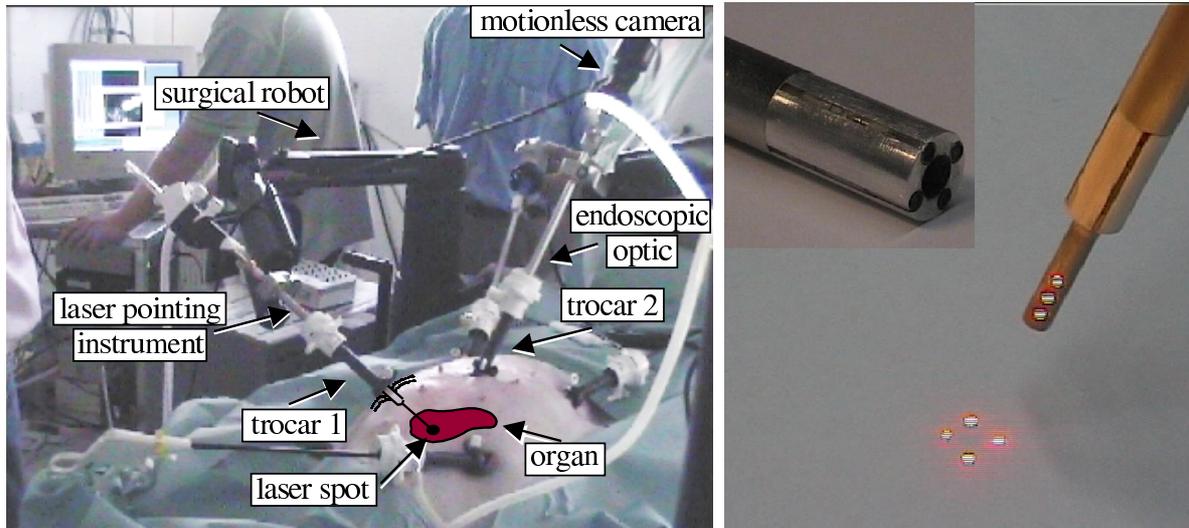


FIGURE 1. (a) System configuration - (b)(c) Endoscopic laser pointing instrument-holder

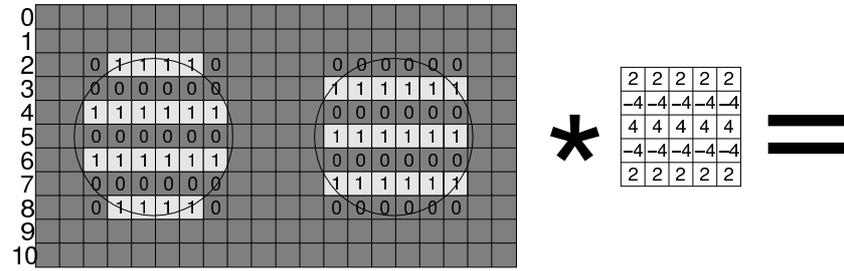
2. ROBUST DETECTION OF LASERS SPOTS AND OPTICAL MARKERS

Robust detection of markers from endoscopic image is quite a challenging issue. In our experiments, we encountered three types of problem that make this task very difficult :

- *i) Lighting conditions* : the light source is on the tip of the endoscope. In this configuration, the reflection is maximal in the center of the image yielding highly saturated areas of pixels.
- *ii) Viscosity of the organs* : this accentuates the reflections of the endoscopic light producing speckles in the image. Furthermore, projected laser spots are diffused yielding large spots of light with unclear contours.
- *iii) Breathing motion* : due to the high magnification factor of the endoscope, the motion in the endoscopic image due to breathing can be very fast. This can make tracking algorithms fail.

We propose a new method for real-time robust detection of markers in a highly noisy scene like an endoscopic view. This technique is based on luminous markers that are blinking at the same frequency than the image acquisition. By switching the marker on in one field of an interlaced image and turning it off in the other field, it is possible to obtain very robust features in the image. Fig.1.a explains how the feature detection works. In this example, we have two blinking dot-shaped markers. The left marker is switched on while the even field of the image is acquired whereas the right marker is switched on in the odd field. To simplify the explanations, only two levels for the pixels (0 for dark and 1 for bright) are used in Fig.1.a.

line N°



line N°

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	14	20	26	26	20	14	8	2	-4	-8	-12	-16	-20	-20	-16	-12	0	0
3	0	0	-20	-28	-36	-36	-28	-20	-12	-4	6	12	18	24	30	30	24	18	0	0
4	0	0	22	30	38	38	30	22	14	6	-8	-16	-24	-32	-40	-40	-32	-24	0	0
5	0	0	-24	-32	-40	-40	-32	-24	-16	-8	8	16	24	32	40	40	32	24	0	0
6	0	0	22	30	38	38	30	22	14	6	-8	-16	-24	-32	-40	-40	-32	-24	0	0
7	0	0	-20	-28	-36	-36	-28	-20	-12	-4	6	12	18	24	30	30	24	18	0	0
8	0	0	14	20	26	26	20	14	8	2	-4	-8	-12	-16	-20	-20	-16	-12	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



FIGURE 2.(a) Robust detection of optical markers. - (b) Image processing.

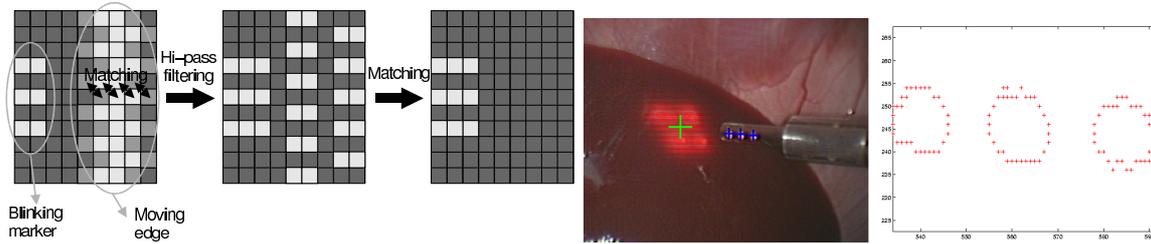


FIGURE 3.(a) Suppression of artifacts due to motion. - (b) Detection of optical markers

The result of the convolution of the image with a 5 by 5 vertical high-pass filter mask shows that, the two markers can be easily detected with a simple thresholding. Furthermore, it's easy to separate the two markers by thresholding separately the even and the odd field in the image. The filtering of the whole image can be performed in real-time due to the symmetry of the convolution mask (for a 768x572 image, the filtering takes about 5 ms on a Pentium IV 1.7 GHz). This detection is very robust, even in a noisy image. Indeed, blinking markers yield patterns in the image whose vertical frequency is the spatial Nyquist frequency of the visual sensor. Usually, in order to avoid aliasing, the lens is designed so that the higher frequencies in the image are cut. So, objects in the scene cannot produce the same image than the blinking markers (one line bright, the next dark, and so on ...). The only other source of vertical high frequency in the image is motion as shown in Fig.3. In this example, the left pattern in the initial image is produced by a blinking marker and the right pattern is produced by the image of an edge moving from left to right in the image. After high-pass filtering and thresholding, the blinking marker is detected as expected but also the moving edge. The artifacts due to the moving edge are removed by a matching algorithm. The horizontal pattern around the detected pixel is compared with the horizontal patterns in the lines that are next to this pixel. If they match, then the pixel is removed. This matching algorithm is very fast since it is limited to the detected pixels. Our setup uses two kind of optical markers that are blinking alternatively: lasers that are projected on the organs, and CMS (Component Mounted on Surface) LEDs (Light Emitting Diodes) that are attached on the tip of the tool. A robust detection of the geometric center of the projected laser spots in the image plane is needed in our system. Due to the complexity of the organ surface, laser spots may be occluded. Therefore a high redundancy factor is achieved by using four laser pointers. We have found in our experiments *in vivo* with four laser sources that the computation of the geometric center is always possible with a small bias even if three spots are occluded. Fig. 2.b shows images resulting from different steps of the image processing applied to the laser spots (from left to right and top to bottom : original image - high-pass filtering - matching - detected geometric center). CMS LEDs markers are turned on during the odd field and turned off during the even field. Edge detection is applied on the result of high-pass filtering and matching in the odd field. Edges detector always yields contours with many pixels of thickness. Thinning operations are performed on the extracted set of pixels based on the comparison of gradient magnitude and direction of each pixel with their neighbors (non maxima suppression) producing a 1-pixel wide edge: this thinning is required to apply hysteresis thresholding and an edge tracking algorithm. Then, contours are merged by using a method called *mutual favorite pairing* [7] that merges neighboring contour chains into a single chain. Finally, the contours are fitted by ellipses (see Fig.3.b - (left) detection of optical markers and laser spots (+) - (right) contours detection of optical markers (odd frame)).

3. INSTRUMENT POSITIONING WITH VISUAL SERVOING

In laparoscopic surgery, instrument displacements are constrained to 4 DOF through an incision point. Only three rotations and one translation are allowed. Let R_L be a frame of reference at the tip of the instrument and R_C the camera frame (see figure 3.a). We define the instrument's velocity vector $\dot{W} = (\omega_x \ \omega_y \ \omega_z \ v_z)^T$ where ω_x , ω_y and ω_z are rotational velocities of frame R_L with respect to camera frame R_C expressed in R_L and v_z is the translational velocity of the instrument along the z_L axis.

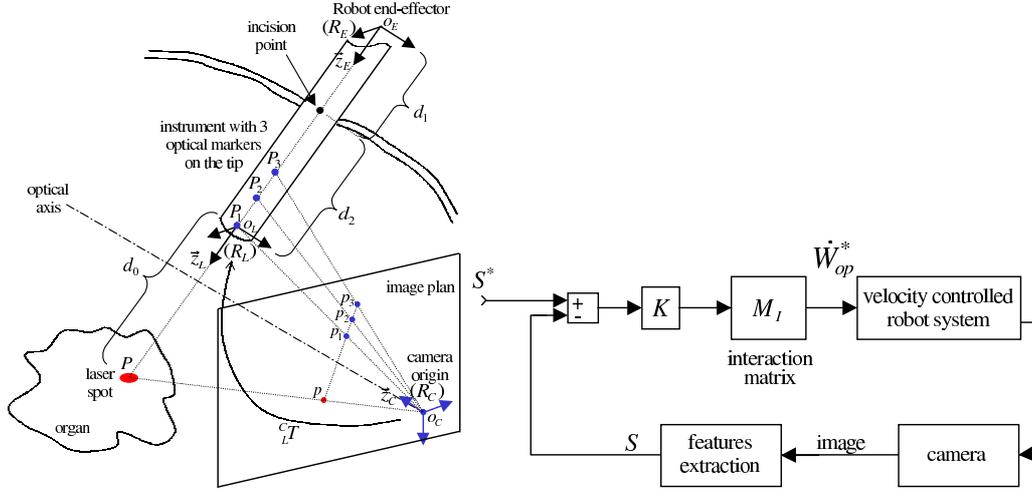


FIGURE 4. (a) System geometry. - (b) Control scheme.

In the case of a symmetrical instrument like the cleaning-suction instrument, it is not necessary to turn the instrument around its own axis to do the desired task. So only 3 DOF will be actuated for this kind of instrument and the operational space velocity vector is reduced to $\dot{W}_{op} = (\omega_x \ \omega_y \ v_z)^T$. For a non symmetrical instrument like a clamping instrument, the rotational speed ω_z can be kept in vector \dot{W}_{op} . To perform a positioning of the instrument, the surgeon specifies the target point on the screen. The required visual features are image coordinates $S_p = (u_p \ v_p)^T$ of the center of mass of the projected laser spots. This center of mass is seen as the projection of the instrument axis onto the organ. Adding optical markers on the tip of the instrument allows to compute the distance d_0 between the organ and the instrument (see section 3.1). Figure 3.b shows the control scheme used for this positioning. The feature vector to control is $S = (u_p \ v_p \ d_0)^T$. If the classical image-based control strategy is used for visual servoing (cf. [8]), the reference velocity vector \dot{W}_{op}^* is chosen proportional to the error on the features: $\dot{W}_{op}^* = M_I K (S^* - S)$, where K is a positive diagonal gain matrix, M_I is the interaction matrix, $M_I = J_S^{-1}$ in classical visual servoing where J_S is a Jacobian matrix relating the instrument velocity vector \dot{W}_{op} to the image feature velocity vector \dot{S} , ($\dot{S} = J_S \dot{W}_{op}$). Furthermore, $S^* = (S_p^* \ d_0^*)^T$, where $S_p^* = (u_p^* \ v_p^*)^T$ is the desired instrument position in the image and d_0^* is the desired distance between the organ and the tip of the instrument. The distance between organ and instrument can be measured only if the laser spots and the instrument markers are visible. When the tip of the instrument is not in the field of view, only S_p can be controlled. Therefore, the instrument retrieval and positioning algorithm must have three stages:

- : **Stage 1:** Positioning of the laser spot at the center of the image by visual servoing of S_p only; control of $(\omega_x \ \omega_y)^T$ only, with $\dot{W}_{op}^* = (\omega_x^* \ \omega_y^* \ 0)^T$.
- : **Stage 2:** Bringing down the instrument along its axis until the optical markers are in the field of view; open-loop motion at constant speed v_z^* with $\dot{W}_{op}^* = (0 \ 0 \ v_z^*)^T$.
- : **Stage 3:** Full visual servoing.

3.1. Depth measurement

Adding optical markers on the tip of the instrument allows to measure the distance d_0 . We put 3 markers P_1, P_2, P_3 along the instrument axis in such a way that we can assume that P_1, P_2, P_3 are collinear with the center of mass of the laser spots P (see Figure 3). Then, the projective cross-ratio τ is computed from points

P, P_1, P_2, P_3 and their respective projection in order to estimate the depth d_0 :

$$\tau = \frac{\left(\frac{\overline{PP_2}}{\overline{P_1P_2}}\right)}{\left(\frac{\overline{PP_3}}{\overline{P_1P_3}}\right)} = \frac{\left(\frac{\overline{FP_2}}{\overline{P_1P_2}}\right)}{\left(\frac{\overline{FP_3}}{\overline{P_1P_3}}\right)} \quad d_0 = \overline{PP_1} = (\tau - 1) \frac{\overline{P_1P_2} \cdot \overline{P_1P_3}}{\overline{P_1P_3} - \tau \cdot \overline{P_1P_2}} \quad (1)$$

3.2. Interaction matrix

To control the position of the instrument in the image plane, one needs to compute the Jacobian matrix J_S which relates the visual features velocity \dot{S} and the instrument velocity \dot{W}_{op} . This Jacobian matrix can be computed analytically if the camera is motionless and if the relative position of the instrument with respect to the camera can be measured as well as the depth variation (if the instrument is in the field of view). For safety reasons, it is not suitable that the laser spots centering error signals, $(S_p^* - S_p)$, induces depth motions ($v_z^* \neq 0$). Furthermore, when the instrument is going up or down ($v_z^* \neq 0$), this motion does not induce an error on the laser spot centering. Therefore, it is recommended to choose the interaction matrix M_I as follows:

$$M_I = \begin{pmatrix} A_{[2 \times 2]} & 0_{[2 \times 1]} \\ 0_{[1 \times 2]} & -1 \end{pmatrix} \quad \hat{A} = \begin{pmatrix} \frac{\Delta u_p}{\omega_x^* \Delta T} & \frac{\Delta u_p}{\omega_y^* \Delta T} \\ \frac{\Delta v_p}{\omega_x^* \Delta T} & \frac{\Delta v_p}{\omega_y^* \Delta T} \end{pmatrix}^{-1} \quad (2)$$

The matrix A in (2) can be analytically computed from J_S if the instrument is visible in the image. When it is not the case, as is stage 1 of the algorithm, this matrix is identified in an initial procedure. This initial identification procedure consists in applying a constant velocity reference signal $\dot{W}_{op}^* = (\omega_x^* \ \omega_y^* \ 0)^T$ during a short time interval ΔT . Small variations of laser spot image coordinates are measured and A is estimated as in (2). The stability of the visual feedback loop is always guaranteed as long as $A\hat{A}^{-1}$ remains positive definite [8]. Should the laser spot centering performance degrades, it is always possible to start again the identification procedure.

4. EXPERIMENTS

Experiments in real surgical conditions on living pigs were conducted in the operating room of IRCAD (see Figure 1.a).

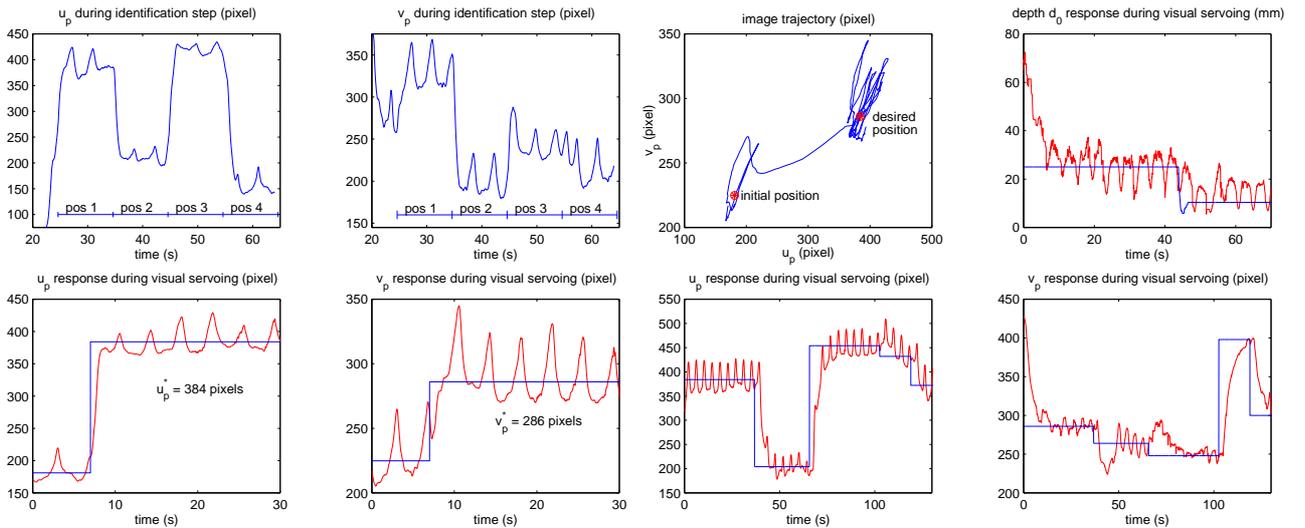


FIGURE 5.(a)(b)-(c)(d) — (e)(f)-(g)(h) Experimental measurements.

The experimental surgical robotic task was the autonomous retrieval of a instrument not seen in the image and then its positioning at a desired 3-D position. We use a bi-processor PC computer (1.7 GHz) for image processing and for controlling, via a serial link, the surgical robot from ComputerMotion. A standard 50fps PAL endoscopic camera hold by a second robot (at standstill) is linked to a PCI image capture board that grabs grey-level images of the observed scene. For each image the center of mass of the laser spots is detected in about 20 ms. The successive steps in the autonomous retrieval and positioning procedure are as following:

- : **Step 1:** changing the orientation of the instrument by applying rotational velocity trajectories (on ω_x^* and ω_y^*) in open loop to scan the organ surface with the laser spot until they appear in the endoscopic view.
- : **Step 2:** automatic identification of the interaction matrix A (cf.(2)).
- : **Step 3:** centering of the laser spots in the image by a 2-D visual servoing.
- : **Step 4:** descent of the instrument by applying a velocity reference signal v_z^* in open loop until it appears in the image, keeping orientation servoing.
- : **Step 5:** real time estimation of the distance d_0 and depth servoing to reach the desired distance, keeping orientation servoing.
- : **Step 6:** new positioning of the instrument to desired 3-D locations by automatic visual servoing at the command of the surgeon. The surgeon indicates on the screen the new laser point image coordinates, $S_p^* = (u_p^* \ v_p^*)^T$, and specifies the new desired distance d_0^* to be reached. Then the visual servoing algorithm performs the positioning.

Figure 4.a-d shows experimental measurements of the laser image coordinates u_p and v_p during the identification stage of A and during the centering stage by visual servoing. For the identification procedure, four positions have been considered to relate variations of the laser image position and angular variations. We can notice a significant perturbation due to the breathing during the visual servoing. For robust identification purpose, we average several measurements of small displacements. This allows to reduce the effect of the breathing which acts like a disturbance. Figure 4.e shows the 2D trajectory obtained (in the image) during the centering step by visual servoing. The oscillating motion around the initial and desired position are also due to the effect of breathing that acts as a periodic perturbation. Figure 4.f shows the measured distance d_0 during the depth servoing at step 5. Responses 4.g and 4.h display the laser image coordinates when the surgeon specifies new positions to be reached in the image.

5. CONCLUSION

The robot vision system presented in this paper automatically retrieves and positions laparoscopic surgical instrument using visual servoing by means of laser pointers and optical markers. To add structured lights on the scene, we designed a laser pointing instrument-holder which can be mounted with any standard instrument in laparoscopic surgery. Successful experiments have been held with a surgical robot on living pigs in a surgical room. In these experiments, the surgeon was able to automatically retrieve a surgical instrument that was out of the field of view and then position it at a desired 3-D location. Our algorithm is essentially based on visual servoing techniques and does not require the measurement or the knowledge of the initial respective position of the endoscope and the surgical instrument. Currently we are working on a new control scheme that better rejects disturbance due to breathing by using repetitive predictive control.

REFERENCES

- [1] J. Marescaux, J. Leroy, M. Gagner, and al. Transatlantic robot-assisted telesurgery. *Nature*, 413:379–380, 2001.
- [2] A. Krupa, J. Gangloff, M. de Mathelin, C. Doignon, G. Morel, L. Soler, J. Leroy, and J. Marescaux. Autonomous retrieval and positioning of surgical instruments in robotized laparoscopic surgery using visual servoing and laser pointers. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 3769–3774, Washington, D.C., May 2002.
- [3] G.-Q. Wei, K. Arbter, and G. Hirzinger. Real-time visual servoing for laparoscopic surgery. *IEEE Engineering in Medicine and Biology*, 16(1):40–45, 1997.

- [4] A. Casals, J. Amat, D. Prats, and E. Laporte. Vision guided robotic system for laparoscopic surgery. In *Proc. of the IFAC International Congress on Advanced Robotics*, pages 33–36, Barcelona, Spain, 1995.
- [5] Y.F. Wang, D.R. Uecker, and Y. Wang. A new framework for vision-enabled and robotically assisted minimally. *Computerized Medical Imaging and Graphics*, 22:429–437, 1998.
- [6] M. Hayashibe and Y. Nakamura. Laser-pointing endoscope system for intra-operative 3Dgeometric registration. In *Proc. of the 2001 IEEE International Conference on Robotics and Automation*, Seoul, Korea, May 2001.
- [7] D.P. Huttenlocher and S. Ullman. Recognizing solid objects by alignment with an image. *The International Journal of Computer Vision*, 5(2):195–212, 1990.
- [8] F. Chaumette, P. Rives, and B. Espiau. Classification and realization of the different vision-based tasks. In *World Scientific Series in Robotics and Automated Systems, Vol. 7*, pages 199–228. Visual Servoing, K. Hashimoto (ed.), Singapore, 1993.