Autonomous 3-D Positioning of Surgical Instruments in Robotized Laparoscopic Surgery Using Visual Servoing

Alexandre Krupa, *Associate Member, IEEE*, Jacques Gangloff, *Member, IEEE*, Christophe Doignon, *Member, IEEE*, Michel F. de Mathelin, *Member, IEEE*, Guillaume Morel, *Member, IEEE*, Joël Leroy, Luc Soler, and Jacques Marescaux

Abstract—This paper presents a robotic vision system that automatically retrieves and positions surgical instruments during robotized laparoscopic surgical operations. The instrument is mounted on the end-effector of a surgical robot which is controlled by visual servoing. The goal of the automated task is to safely bring the instrument at a desired three-dimensional location from an unknown or hidden position. Light-emitting diodes are attached on the tip of the instrument, and a specific instrument holder fitted with optical fibers is used to project laser dots on the surface of the organs. These optical markers are detected in the endoscopic image and allow localizing the instrument with respect to the scene. The instrument is recovered and centered in the image plane by means of a visual servoing algorithm using feature errors in the image. With this system, the surgeon can specify a desired relative position between the instrument and the pointed organ. The relationship between the velocity screw of the surgical instrument and the velocity of the markers in the image is estimated online and, for safety reasons, a multistages servoing scheme is proposed. Our approach has been successfully validated in a real surgical environment by performing experiments on living tissues in the surgical training room of the Institut de Recherche sur les Cancers de l'Appareil Digestif (IRCAD), Strasbourg, France.

Index Terms—Medical robotics, minimally invasive surgery, visual servoing.

I. INTRODUCTION

IN LAPAROSCOPIC surgery, small incisions are made in the human abdomen. Various surgical instruments and an endoscopic optical lens are inserted through trocars at each incision point. Looking at the monitor device, the surgeon moves the instruments in order to perform the desired surgical task. One drawback of this surgical technique is due to the posture of the surgeon, which can be very tiring. Teleoperated robotic laparoscopic systems have recently appeared. There exist several commercial systems, e.g., ZEUS (Computer Motion, Inc., Santa Barbara, CA) or Da Vinci (Intuitive Surgical, Inc., Mountain View, CA). With these systems, robot arms are used to manipulate surgical instruments as well as the endoscope. The surgeon teleoperates the robot through master arms using the visual feedback from the laparoscopic image. This reduces the surgeon's tiredness, and potentially increases motion accuracy by the use of a high master–slave motion ratio. We focus our research in this field on expanding the potentialities of such systems by providing "automatic modes" using visual servoing (see [7] and [8] for earlier works in that direction). For this purpose, the robot controller uses visual information from the laparoscopic images to move instruments, through a visual servo loop, toward their desired location.

Note that prior research was conducted on visual servoing techniques in laparoscopic surgery to automatically guide the camera toward the region of interest (see, e.g., [1], [15], and [16]). However, in a typical surgical procedure, it is usually the other way around: 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 usually not in the field of view at the start of the procedure. Therefore, the surgeon must either blindly move the instruments or zoom out with the endoscope in order to get a larger field of view. Similarly, when the surgeon zooms 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 undesirable contact between instruments and organs.

Therefore, in order to assist the surgeon, we propose a visual servoing system that automatically brings the instruments at the center of the endoscopic image in a safe manner. This system can be used also to move the instruments at a position specified by the surgeon in the image (with, e.g., a touch screen or a mouse-type device). This system allows doing away with the practice of moving the endoscope in order to vizualize the instrument at any time it is introduced to the patient. It 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 guide the instrument. Visibility of laser spots in the image is sufficient to guarantee that the instrument is not blocked by unseen tissue. Because of the poor structuration of the scene and the difficult lighting conditions, several laser pointers are used to guarantee the robustness of the instrument recovery system. A difficulty in designing this automatic instrument recovery system lies in the unknown relative position between the camera and the robot arm holding the instrument, and in the monocular vision that induces a lack of depth information. This problem is also tackled in [3], where an intraoperative three-dimensional (3-D) geometric registration system is presented. The authors add a second endoscope with an optical galvano-scanner. Then, a 955 frames per second (fps) high-speed camera is used with the first endoscopic lens to estimate the 3-D surface of the scanned organ. Furthermore, external cameras watching the whole surgical scene (the Optotrak system) are added to measure the relative position between the laser-pointing endoscope and the camera.

In our approach, only one monocular endoscopic vision system is needed for the surgeon and the autonomous 3-D positioning. The camera has two functions: to give the surgeon a visual feedback, and to provide measurements of the position of optical markers. The relative position from the instrument to the organ is estimated by using images of blinking optical markers mounted on the tip of the instrument and images of blinking laser spots projected by the same instrument.

Note that many commercially available tracking systems make also use of passive or active blinking optical markers synchonized with image acquisition [17]. The most famous among these systems is the Optotrak from Nothern Digital, Inc., Waterloo, ON, Canada, which uses synchronized infrared light-emitting diode (LED) markers tracked by three infrared (IR)-sensitive cameras. However, in the case of all these systems, the imaging system is dedicated to the markers detection task, since they are the only features seen by the camera(s). This greatly simplifies the image processing: there is no need to segment the whole image to extract the markers' locations.

In our system, only one standard, commercially available, endoscopic camera is used for both 3-D measurement and surgeon visual feedback. To do so, we propose a novel method to extract efficiently, in real time, with a high signal-to-noise (S/N) ratio, markers in a scene as complex as an inner human body environment. Furthermore, with our method, it is easy to remove, by software, images of the markers from the endoscopic image and give to the surgeon a quasi-unmodified visual feedback.

The paper is organized as follows. Section II describes the system configuration with the endoscopic laser-pointing instrument holder. Robust image processing for laser spots and LEDs detection is explained in Section III. The control scheme used to position the instrument by automatic visual feedback is described in Section IV. The method for estimating the distance from the instrument to the organ is also presented. In Section V, we show experimental results in real surgical conditions at the operating room of the Institut de Recherche sur les Cancers de l'Appareil Digestif (IRCAD), Strasbourg, France.
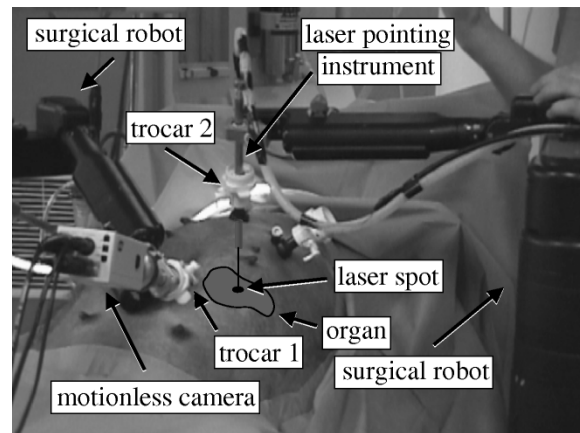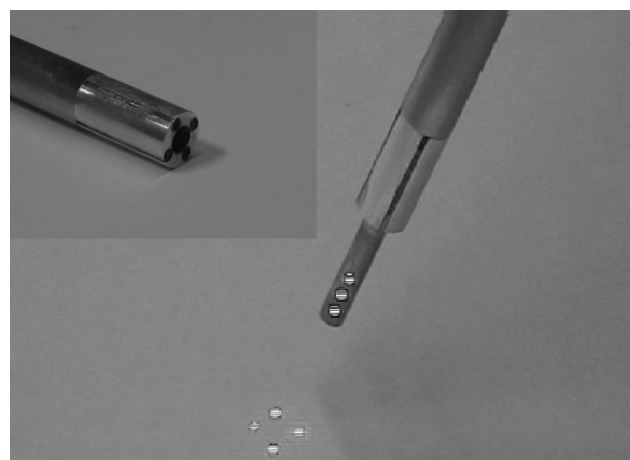

Fig. 1. System configuration.

Fig. 2. Endoscopic laser-pointing instrument holder.

II. SYSTEM DESCRIPTION

A. System Configuration

The system configuration used to perform the autonomous positioning of the surgical instrument is shown in Fig. 1. The system includes a laparoscopic surgical robot, an endoscopic optical lens, and an endoscopic laser-pointing instrument holder. The robotic arm allows moving the instrument across a trocar placed at a first incision point. The surgical instrument is mounted into the laser-pointing instrument holder. This instrument holder projects laser patterns on the organ surface in order to provide information about the relative orientation of the instrument with respect to the organ, even if the surgical instrument is not in the camera's field of view. Another incision point is made in order to insert an endoscopic optical lens which provides the visual feedback and whose location, relative to the robot base frame, is generally unknown.

B. Endoscopic Laser-Pointing Instrument Holder

The prototype of an endoscopic laser-pointing instrument holder is shown in Fig. 2. This instrument holder, with the surgical instrument inside, is held by the end-effector of the robot. It is a 30-cm-long metallic pipe, with a 10 mm external diameter to be inserted through a 12 mm standard trocar. Its internal diameter is 5 mm, so that a standard laparoscopic

line N°

```
      0 1 1 1 0        0 0 0 0 0 0
      0 0 0 0 0 0      1 1 1 1 1 1
      1 1 1 1 1 1      0 0 0 0 0 0
      0 0 0 0 0 0      1 1 1 1 1 1
      1 1 1 1 1 1      0 0 0 0 0 0
      0 0 0 0 0 0      1 1 1 1 1 1
      0 1 1 1 0        0 0 0 0 0 0
```

convolution mask:
```
 2  2  2  2
-4 -4 -4 -4 -4
 4  4  4  4  4
-4 -4 -4 -4 -4
 2  2  2  2
```

line N°

```
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
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
```
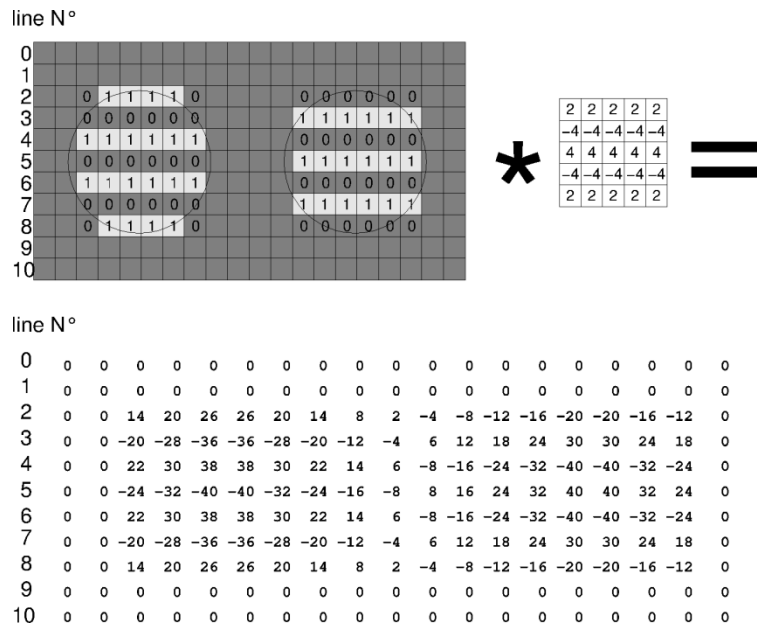
Fig. 3.   Robust detection of optical markers.

surgical instrument can fit inside. The head of the instrument holder contains miniature laser collimators connected to optical fibers which are linked to externally controlled laser sources. This device allows using remote laser sources which can not be integrated in the head of the instrument due to their size. Optical markers are also added on the tip of the surgical instrument. These markers (made up with three LEDs) are directly seen in the image. They are used in conjunction with the image of the projected laser pattern in order to measure the distance between the pointed organ and the instrument.

III. Robust Detection of Laser Spots and Optical Markers

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

1) *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.
2) *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 fuzzy contours.
3) *Breathing motion*: Due to the high magnification factor of the endoscope, the motion in the endoscopic image due to breathing is of high magnitude. This may lead to a failure of the tracking algorithm.

To cope with these difficulties, we have developed 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 as the image acquisition. By switching the marker on when acquiring one field of an interlaced image and turning it off when acquiring the other field, it is possible to obtain very robust features in the image. Fig. 3 explains how the feature detection works. In this example, we use two blinking disk-shaped markers. The left marker is switched on during the even field acquisition, whereas the right marker is switched on during the odd field. To simplify the explanations, only two levels for the pixels (0 for dark and 1 for bright) are used in Fig. 3.

The result of the convolution of the image with a 5×5 vertical high-pass filter mask shows that the two markers can be easily detected with a simple thresholding procedure. Furthermore, it is 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 768×572 image, it takes 5 ms with a Pentium IV 1.7 GHz).

This detection is very robust to image noise. 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 as the blinking markers (one line bright, the next dark, and so on…). The only other source of vertical high-frequency components in the image is motion, as shown in Fig. 4.

In this example, the left pattern in the original 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 is very fast since it is limited to the detected pixels.

Our setup uses two kinds of optical markers that are blinking alternatively: lasers that are projected on the organs, and component mounted on surface (CMS) LEDs that are attached on the
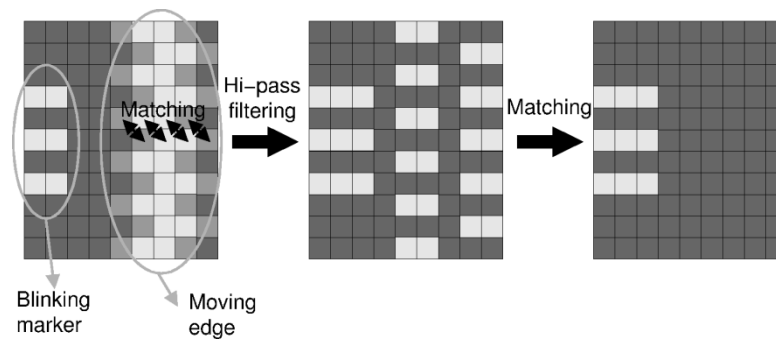
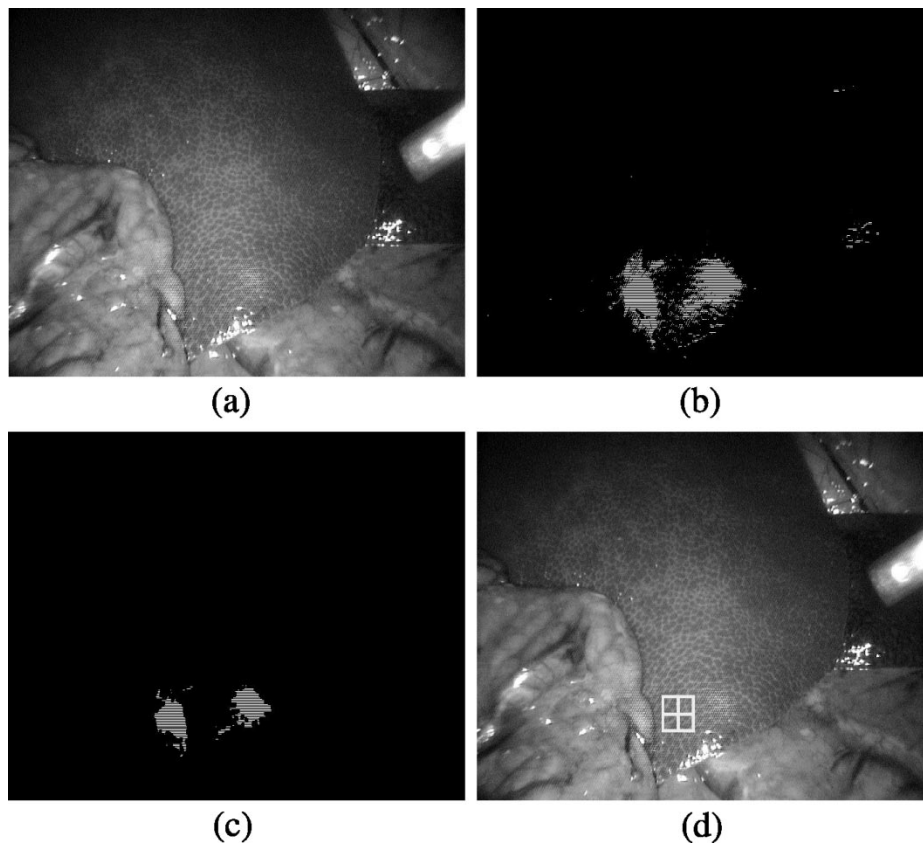Fig. 4. Suppression of artifacts due to motion.

Fig. 5. Detection of laser spots. (a) Original interlaced image. (b) High-pass filtering and thresholding on even frame. (c) Matching. (d) Localization of center of mass (square).

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 limited bias, even if three spots are occluded. Fig. 5 shows images resulting from different steps of the image processing applied to the laser spots.

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 (nonmaxima 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* [6] that merges neighboring contour chains into a single chain. Finally, the contours are fitted by ellipses (see Fig. 6).

For safety reasons, we have added the simple following test: the S/N ratio is monitored by setting a threshold on the minimum number of pixels for each detected marker. If the test fails, the visual servoing is immediately stopped.

Furthermore, to reduce the effect of noise, a low-pass filter is applied on the time-varying image feature coordinates. Areas of interest around detected markers are also used in order to reduce the processing time.
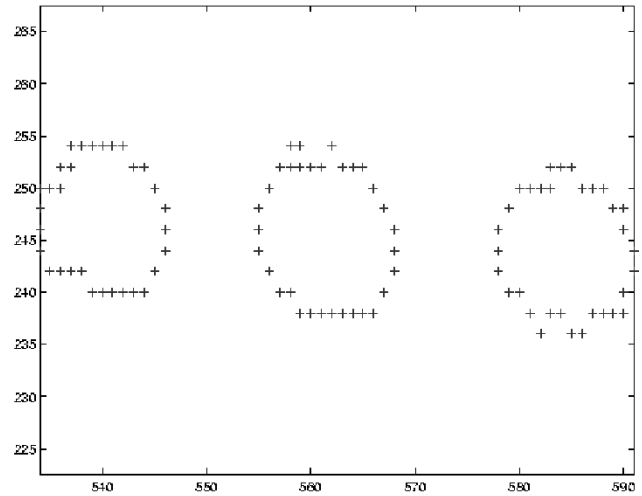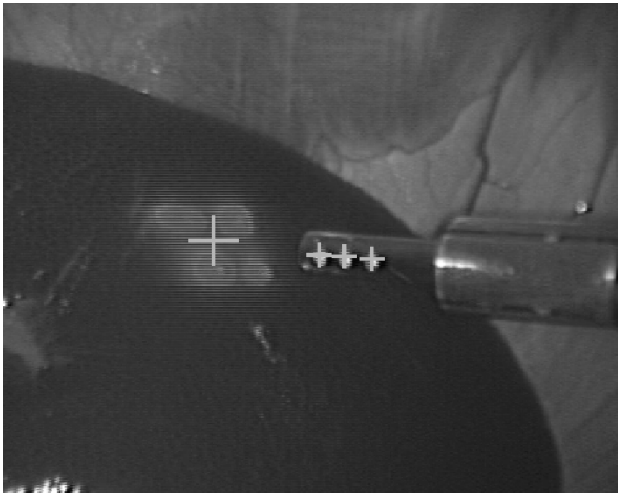
Fig. 6. (left) Detection of optical markers and laser spots ($+$). (right) Contours detection of optical markers (odd frame).

Since the markers appear only on one out of two image lines, and since the areas of the laser and LED markers do not overlap, it is possible to remove these markers from the image by software. For each marker, each detected pixel can be replaced by the nearest pixel that is unaffected by the light of the marker. Therefore, the surgeon does not see the blinking markers in the image, which is more comfortable. This method was validated with two standard endoscopic imaging systems: the Stryker 888 and the Stryker 988.

IV. INSTRUMENT POSITIONING WITH VISUAL SERVOING

The objective of the proposed visual servoing is to guide and to position the instrument mounted on the end-effector of the medical robot. In laparoscopic surgery, displacement are reduced to four degrees of freedom (DOFs), since translational displacements perpendicular to the incision point axis are not allowed by the trocar (see Fig. 7). In the case of a symmetrical instrument like, e.g., the cleaning-suction instrument, it is not necessary to turn the instrument around its own axis to do the desired task. For practical convenience, rotation around the instrument axis is constrained in a way to keep optical markers visible. In our system, a slow visual servoing is performed, based on the ellipses minor/major semiaxes ratio fitting the image projections of optical markers. Since this motion does not contribute to position the tip of the instrument, it is not further considered.

A. Depth Estimation

To perform the 3-D positioning, we need to estimate the distance between the organ and the instrument (depth d_0, in Fig. 7). Three optical markers, P_1, P_2, and P_3, are placed along the tool axis and are assumed to be collinear with the center of mass, P, of the laser spots (see Fig. 7). Under this assumption, a cross ratio, τ, can be computed using these four geometric points [12]. This cross ratio can also be computed in the image using their respective projections p_1, p_2, p_3, and p, assuming the optical markers are in the camera field of view (see Fig. 7). Since a one-dimensional (1-D) projective basis can be defined either with $\{P_1, P_2, P_3\}$ or their respective images $\{p_1, p_2, p_3\}$, the
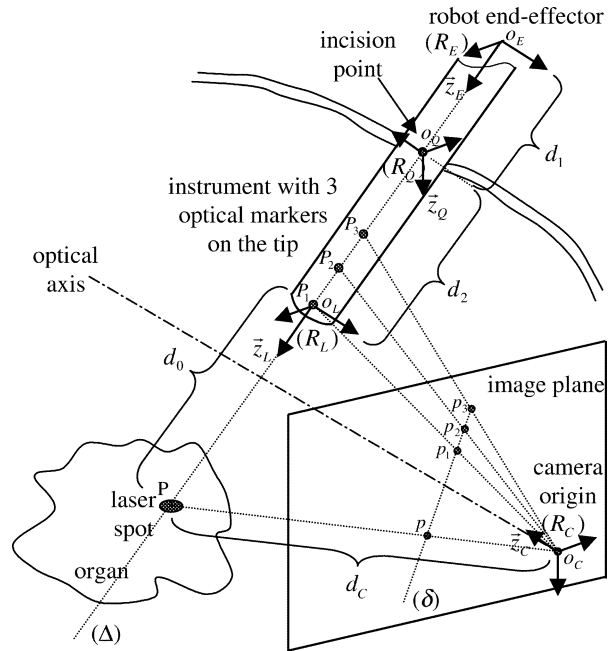

Fig. 7. Basic geometry involved.

selected cross ratio built with the fourth point (P or p) is a projective invariant that can be used to estimate the depth d_0. Indeed, a 1-D homography H exists between these two projective bases, so that the straight line Δ corresponding to the instrument axis is transformed, in the image, into a line $\delta = H(\Delta)$

$$\tau = \frac{\left(\dfrac{\overline{pp_2}}{\overline{p_1p_2}}\right)}{\left(\dfrac{\overline{pp_3}}{\overline{p_1p_3}}\right)} = \frac{\left(\dfrac{\overline{PP_2}}{\overline{P_1P_2}}\right)}{\left(\dfrac{\overline{PP_3}}{\overline{P_1P_3}}\right)} \tag{1}$$

$$d_0 = \overline{PP_1} = (1-\tau)\frac{\overline{P_1P_3}}{\tau - \dfrac{\overline{P_1P_3}}{\overline{P_1P_2}}} = \alpha\,\frac{1-\tau}{\tau - \beta} \tag{2}$$

where α and β depend only on the known relative position of P_1, P_2, and P_3. Similar computations lead to the same relationship
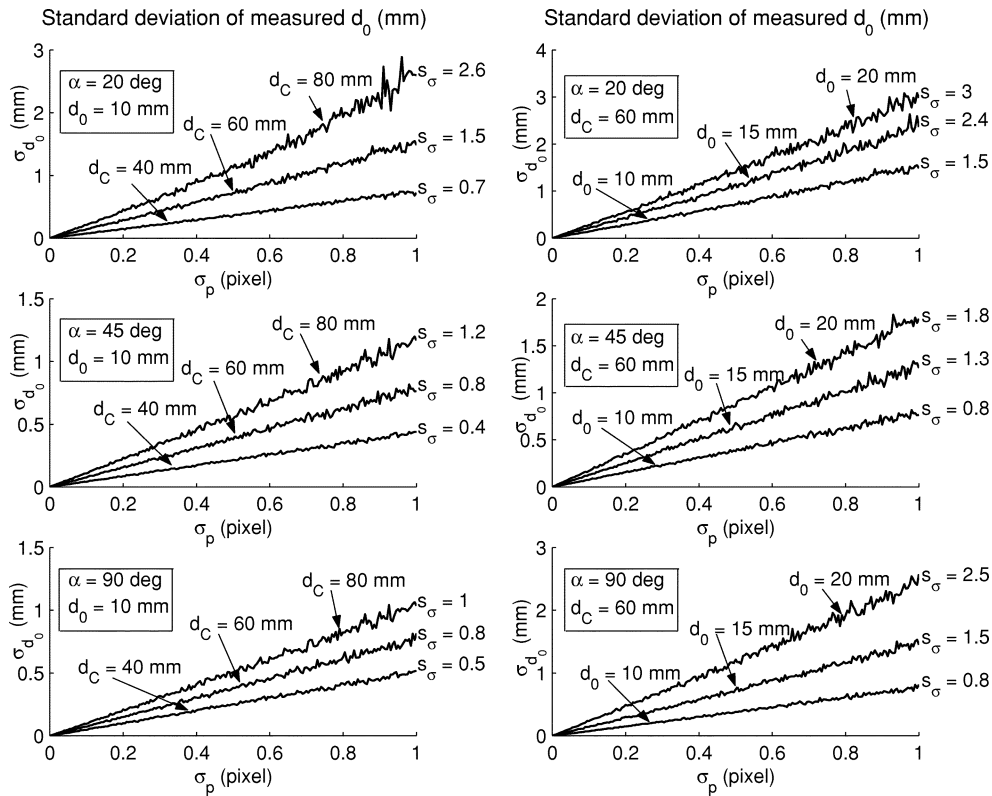
Fig. 8. Standard deviation σ_{d_0} of the estimated depth d_0 (in millimeters) as a function of the standard deviation σ_p of the markers image coordinates (in pixels) for several geometrical configurations.

between d_2 and another cross ratio μ defined with the points P_1, P_2, P_3, O_Q and their respective projections, provided that o_q, the perspective projection of the incision point O_Q, can be recovered. Since O_Q is generally not in the camera field of view, this can be achieved by considering a displacement of the surgical instrument between two configurations, yielding straight lines δ and δ' in the image. Then o_q is the intersection of these lines, since O_Q is motionless. Finally

$$\mu = \frac{\left(\frac{\overline{p_1 p_3}}{\overline{p_2 p_3}}\right)}{\left(\frac{\overline{p_1 o_q}}{\overline{p_2 o_q}}\right)} = \frac{\left(\frac{\overline{P_1 P_3}}{\overline{P_2 P_3}}\right)}{\left(\frac{\overline{P_1 O_Q}}{\overline{P_2 O_Q}}\right)} \quad (3)$$

$$d_2 = \overline{P_1 O_Q} = \frac{\frac{\alpha}{1-\beta}}{\mu + \frac{\beta}{1-\beta}}. \quad (4)$$

Fig. 8 shows the results of a sensitivity analysis on the depth estimation. The standard deviation σ_{d_0} of the estimated depth is plotted as a function of the standard deviation σ_p of the markers coordinates (pixel) in the image plane for several geometrical configurations of the camera and surgical instrument. These configurations are defined by the angle α between the camera's optical axis and the instrument axis, the depth d_0 and the depth d_C between the camera and the laser spot. It can be seen that for standard configurations, the sensitivity of the depth measurement with respect to noise (that is, $s_\sigma = d(\sigma_{d_0})/d\sigma_p$ in Fig. 8) is proportional to the distance d_c and d_0. The sensitivity, s_σ, varies in the interval $[0.4, 3]$ corresponding to $[0.4, 3]$ mm for

σ_{d_0} if $\sigma_p = 1$ pixel. Experimentally, σ_p is typically 0.5 pixel, resulting in $\sigma_{d_0} = 1$ mm. In practice, this noise does not affect the precision of the positioning, due to the low-pass filter effect of the visual servoing.

B. Visual Servoing

In our approach, we combine image feature coordinates and depth information to position the instrument with respect to the pointed organ. There exist previous works about this type of combination (see, e.g., [10] and [11]), however the depth, d_0, of concern here is independent of the position of the camera and it can be estimated with an uncalibrated camera. A feature vector **S** is built with image coordinates of the perspective projection of the laser spots center $S_p = (u_p, v_p)^T$, and the depth d_0 between the pointed organ and the instrument $S = [S_p^T \quad d_0]^T$. In our visual servoing scheme, the robot arm is velocity controlled. Therefore, the key issue is to express the interaction matrix relating the derivative of **S** and the velocity screw of the surgical instrument reduced to three DOFs $\dot{W}_{op} = (\omega_x, \omega_y, v_z = \dot{d}_2)^T$ (see the Appendix for more details)

$$\begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \underbrace{\begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix}}_{\mathbf{J}_S} \begin{bmatrix} \omega_x \\ \omega_y \\ \dot{d}_0 + v_z \end{bmatrix}. \quad (5)$$

Even though all components of $\mathbf{J}_S$ could be recovered from images of optical markers and camera parameters, $\mathbf{J}_S$ is not invertible. Therefore, the velocity screw applied to the robot, $\dot{W}_{op}^* = (\omega_x^*, \omega_y^*, v_z^*)^T$, cannot be directly computed without
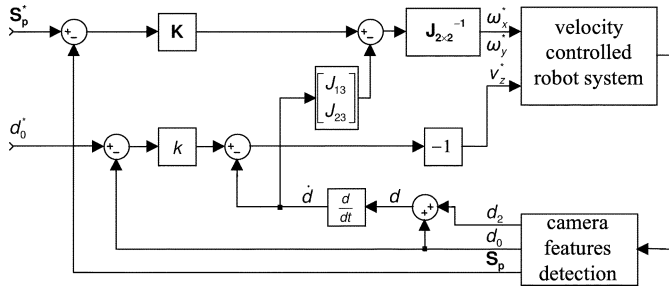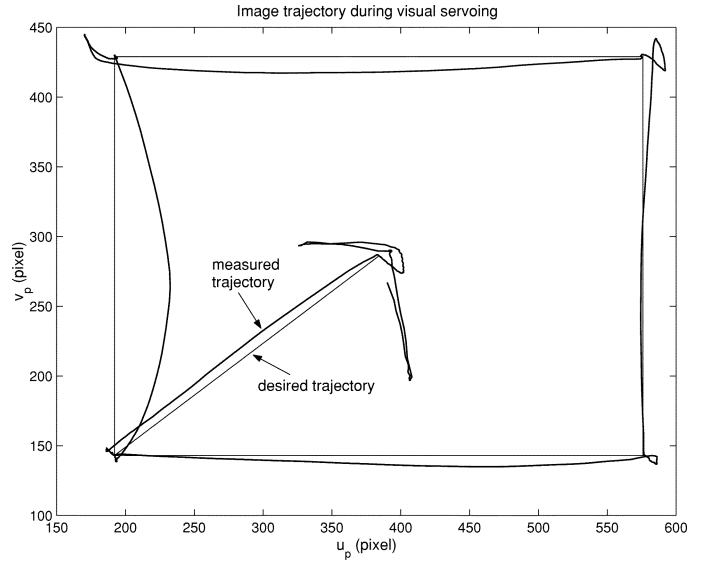
Fig. 9.   Full visual servoing scheme.

Fig. 10.   Initial online identification of the interaction matrix (displacements around the image center) and image-based visual servoing along a square using this identification. (1 mm ≈ 25 pixels.)

some additional assumptions (like, e.g., the surface of the organ in the neighborhood of the pointed direction is planar). Furthermore, when the instrument is not in the camera field of view, d_0 cannot be measured. Therefore, we propose to decompose the visual servoing in two control loops that partly decouple the control of the pointed direction given by $(u_p, v_p)^T$ and the control of the depth d_0. The instrument recovery algorithm is split into three stages.

Instrument Recovery and Positioning Procedure:

• **Stage 1:** Positioning of the laser spot projection, p, at the center of the image by visual servoing of $S_p = (u_p, v_p)^T$ only.

It means that only ω_x and ω_y are controlled. For safety reasons, $v_z^* = 0$ during this stage. Thus, from (5), we have

$$\begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} - \begin{bmatrix} J_{13} \\ J_{23} \end{bmatrix} \dot{d} = \underbrace{\begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}}_{\mathbf{J}_{2\times 2}} \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix}. \qquad (6)$$

Assuming a classical proportional visual feedback [5], the control signal applied to the robot is $\dot{W}_{op}^* = (\omega_x^*, \omega_y^*, 0)^T$, with

$$\begin{bmatrix} \omega_x^* \\ \omega_y^* \end{bmatrix} = \mathbf{J}_{2\times 2}^{-1} \left(\mathbf{K} \begin{bmatrix} u_p^* - u_p \\ v_p^* - v_p \end{bmatrix} - \begin{bmatrix} J_{13} \\ J_{23} \end{bmatrix} \dot{d} \right) \qquad (7)$$

where $\mathbf{K}$ is a positive constant gain matrix.

• **Stage 2:** Bringing down the instrument along its axis until the optical markers are in the field of view.

This is done by an open-loop motion at constant speed v_z^* with $\dot{W}_{op}^* = (0, 0, v_z^*)^T$.

• **Stage 3:** Full visual servoing.

Since strong deformations may be induced by breathing, an entire decoupling (i.e., $\omega_x^* = 0$, $\omega_y^* = 0$) is not suitable. The first stage control, as in (7), must go on in order to reject disturbances. Since $v_z = \dot{d} - \dot{d}_0$, a proportional visual feedback law based on the measurement of d_0 with the cross ratio is given by

$$v_z^* = \dot{d} - k \left(d_0^* - d_0(\tau, \alpha, \beta) \right) \qquad (8)$$

where k is a positive scalar and d_0 is a function of the cross ratio τ, α, and β. The full servoing scheme is shown in Fig. 9.

C. Implementation Issues

The signal $\dot{d}$ in (7) and (8) can be obtained by derivating (2) and (4). However, since d is slowly varying at stage 1, and since

S_p is generally constant at stages 2 and 3, the approximation $\dot{d} \approx 0$ is made during practical experiments resulting in an approximately decoupled behavior.

For practical convenience, the upper (2×2) submatrix $\mathbf{J}_{2\times 2}$ of $\mathbf{J}_S$ must be computed even if the optical markers are not visible. When the instrument is out of the field of view, this submatrix is identified in an initial procedure. This identification consists in applying a constant rotational velocity reference $\dot{W}_{op}^* = (\omega_x^*, \omega_y^*, 0)^T$ during a short time interval ΔT (see Fig. 10). Small variations of laser spot image coordinates are measured and the estimated $\widehat{\mathbf{J}}_\omega$ of the interaction matrix is given by

$$\widehat{\mathbf{J}}_\omega = \begin{bmatrix} \dfrac{\Delta u_p}{\omega_x^* \Delta T} & \dfrac{\Delta u_p}{\omega_y^* \Delta T} \\ \dfrac{\Delta v_p}{\omega_x^* \Delta T} & \dfrac{\Delta v_p}{\omega_y^* \Delta T} \end{bmatrix}. \qquad (9)$$

It is not suitable to try to compensate induced depth motions during the centering stage, since the instrument is not usually in the field of view at that stage. Furthermore, when the instrument is going up or down ($v_z^* \neq 0$), no bias appears on the laser spot centering. Therefore, it is recommended in practice to choose the interaction matrix $\mathbf{M}_I$, mapping $\dot{W}_{op}^*$ into $\dot{S}^*$ with the following structure:

$$\mathbf{M}_I = \begin{bmatrix} \widehat{\mathbf{J}}_\omega & 0_{[2\times 1]} \\ 0_{[1\times 2]} & -1 \end{bmatrix}. \qquad (10)$$

This leads to the experimental control scheme shown in Fig. 11, with $\mathbf{K} > 0$. The bandwith of the visual control loop is directly proportional to $\mathbf{K}$.

For the stability analysis, we consider an ellipsoid as a geometric model for the abdominal cavity, so that $\dot{d}$ is related to ω_x and ω_y. In this case, the interactions matrix, in (5), is reduced to a 2×2 matrix $\mathbf{J}_\omega$

$$\begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \mathbf{J}_\omega \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix} \qquad (11)$$
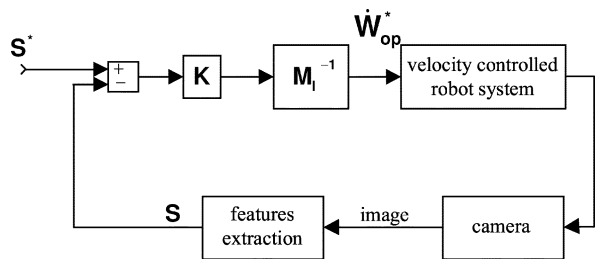
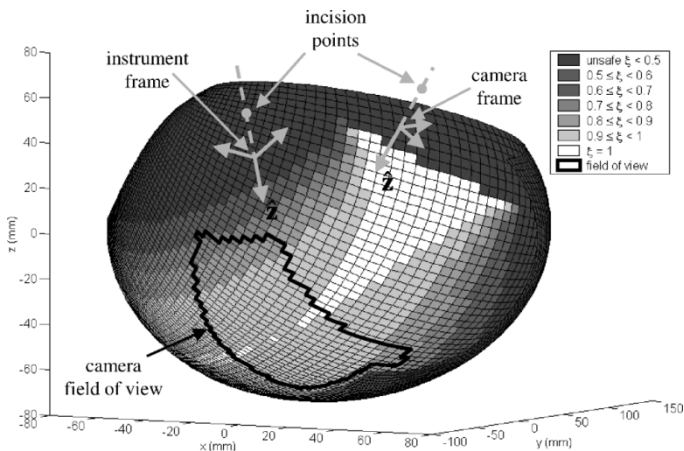Fig. 11. Visual servoing scheme used for experiments.

Fig. 12. Stability analysis. Laser spot surface area delimiting the stability of the control vision loop for a constant identified Jacobian.

and the stability of the visual feedback loop is guaranteed as long as $\mathbf{J}_\omega[\widehat{\mathbf{J}}_\omega]^{-1}$ remains positive definite [2]. In our application, if the camera and the incision point are motionless, the stability is ensured in a workspace much larger than the region covered during experiments. To quantify the stability properties, we have modeled the organ as an ellipsoid. The estimated Jacobian $\widehat{\mathbf{J}}_\omega$ is constant and correspond to a nominal configuration. We have then computed $\mathbf{J}_\omega$ when the laser spot is moved across the organ surface, and computed the eigenvalues of $\mathbf{J}_\omega[\widehat{\mathbf{J}}_\omega]^{-1}$ in the different configurations. Unsafe configurations, corresponding to a damping factor $\zeta < 0.5$, are clearly out of the camera field of view, which is represented by the black contours (see Fig. 12). This leads to a good robustness over the whole image that was experimentally verified.

Furthermore, an accidental motion of the endoscope could also affect $\mathbf{J}_\omega$, and thus the stability. However, in practice, experiments have demonstrated that a rotation of the endoscope as big as $60°$ still preserves the stability of the system. Should the convergence properties be degraded in case of an exceptional change of $\mathbf{J}_\omega$, the tracking performances can be easily monitored and a reidentification of $\widehat{\mathbf{J}}_\omega$ can be programmed ([13], [14]). The validity of the Jacobian matrix can be measured by two ways: the monitoring of the rotational motions of the endoscope or the monitoring of the trajectory error signal in the image (the optimal trajectory should be a straight line for an image-based servoing).

V. EXPERIMENTS

Experiments in real surgical conditions were conducted on living tissues in the operating room of IRCAD (see Fig. 1). The experimental surgical robotic task was the autonomous recovery
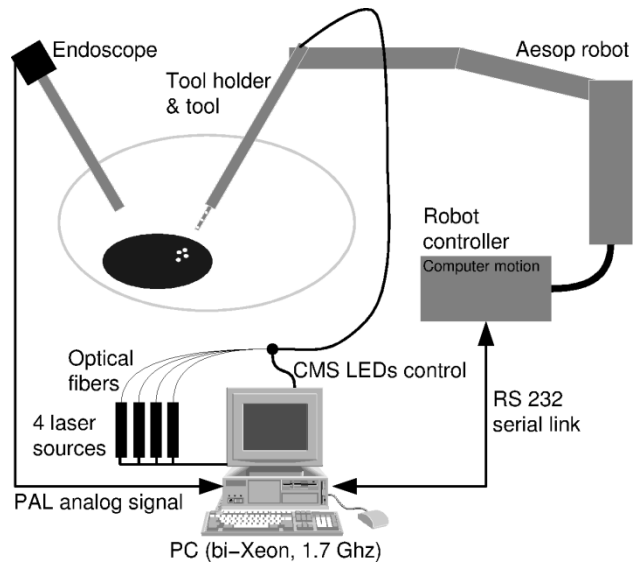

Fig. 13. Experimental setup.

of a instrument not seen in the initial image and then its positioning at a desired 3-D position.

A. Experimental Setup

We use a bi-processor PC computer (1.7 GHz) running Linux for image processing and for controlling, via a serial link, the Computer Motion surgical robot. A standard 50 fps PAL endoscopic camera held by a second robot (at standstill) is linked to a PCI image capture board that grabs images of the observed scene (see Fig. 13). We have modified the driver of the acquisition board in order to use the vertical blank interrupt as a mean to synchronize the blinking markers. The TTL synchronization signals that control the state of the lasers and the LEDs are provided by the PC's parallel port. For each image, the center of mass of the laser spots and centers of the three LEDs are detected in about 20 ms.

B. Experimental Task

Successive steps in the autonomous recovery and positioning are as follows.

Step 1) Changing the orientation of the instrument by applying rotational velocity trajectories (ω_x^* and ω_y^*) in open loop in order to scan the organ surface with the laser spots until they appear in the endoscopic view.

Step 2) Automatic identification of the components of the interaction matrix $\mathbf{J}_\omega$ [cf. (9)].

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, while the orientation servoing is running with a fixed desired set point.

Step 5) Real-time estimation of the distance d_0 and depth servoing to reach the desired distance, while orientation servoing is running with a fixed desired set point.

Step 6) New positioning of the instrument toward a desired 3-D location by automatic visual servoing under
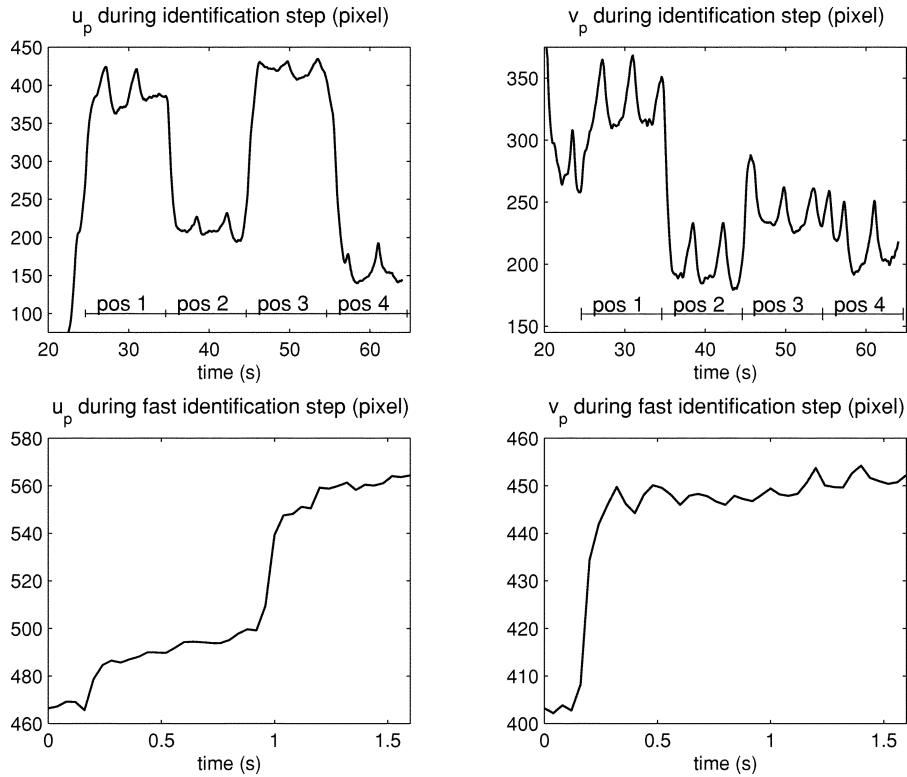
Fig. 14.   Experimental measurements for the identification procedure of the interaction matrix. (top) Slow identification procedure (averaging the effect of breathing) and (bottom) fast identification procedure (short time interval between two regular breaths). (1 mm $\approx$ 25 pixels.)
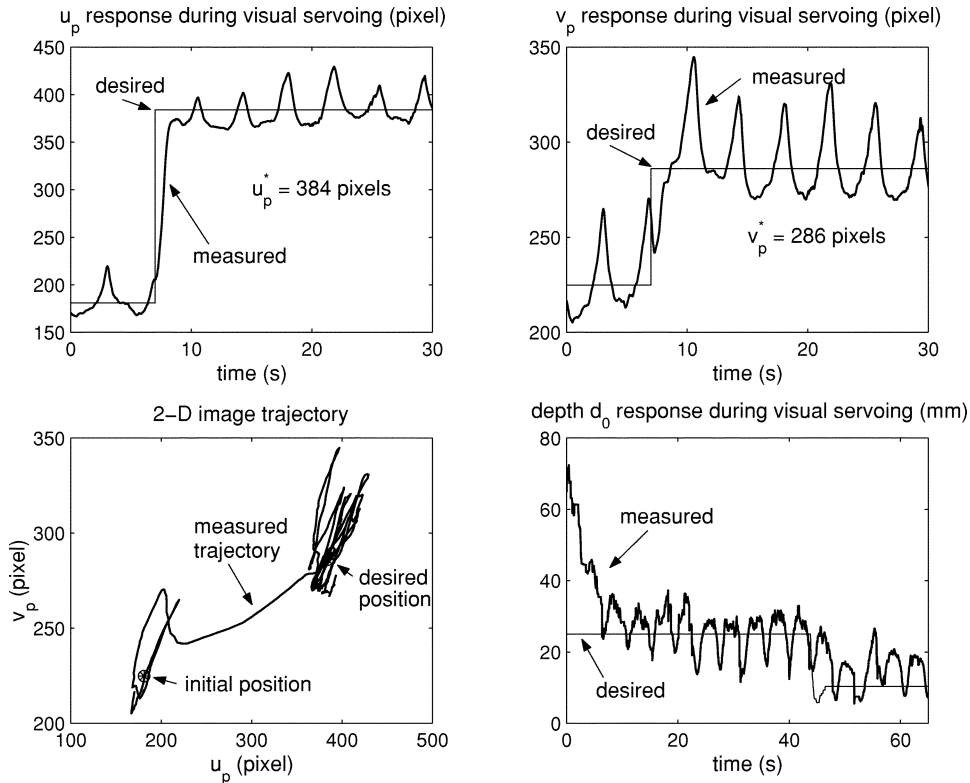

Fig. 15.   Experimental results of the 3-D positioning. (top) Image centering. (bottom left) 2-D trajectory. (bottom right) Depth d_0 control by visual servoing. (1 mm $\approx$ 25 pixels.)

the surgeon's control. The surgeon indicates on the screen the new laser point image coordinates, $S_p^* = (u_p^* \quad v_p^*)^T$, and specifies the new desired distance d_0^* to be reached. Then, the visual servoing algorithm performs the 3-D positioning.

C. *Experimental Measurements*

Fig. 14 shows experimental measurements of the laser image coordinates u_p and v_p during the identification stage of $\mathbf{J}_\omega$. For the identification procedure, four positions have been consid-
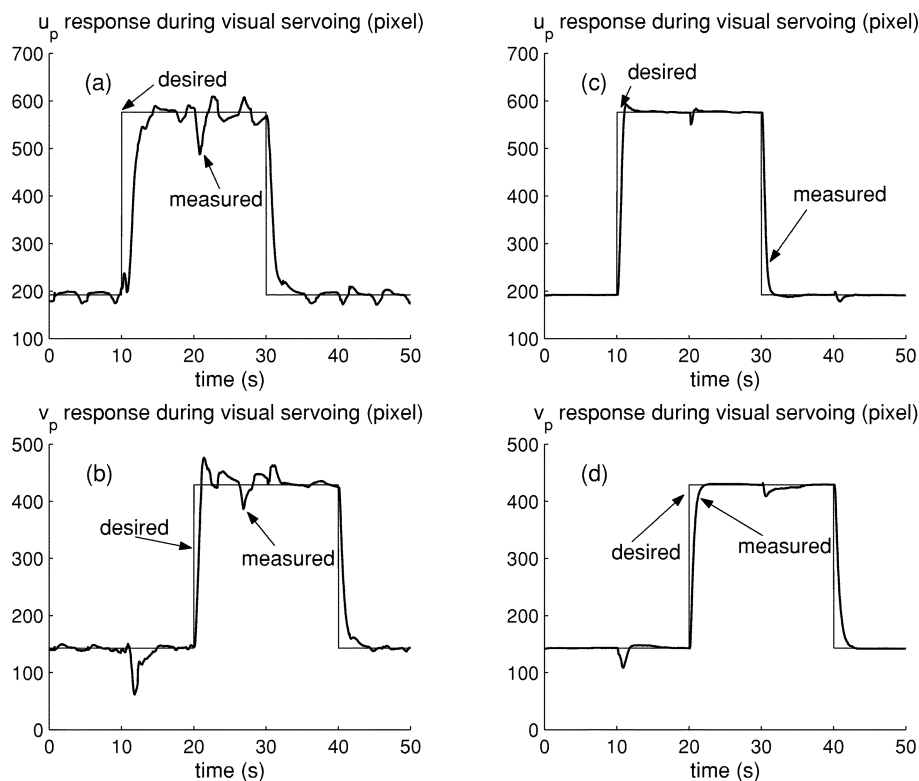
Fig. 16. New desired respective positions of the surgical instrument with respect to the pointed organ specified in the image by the surgeon (step 6). (left) Responses obtained on living tissue. (right) Responses obtained by the use of an endo-trainer with no disturbance due to breathing. (1 mm ≈ 25 pixels.)

TABLE I
TIME PERFORMANCES OF THE RECOVERING AND POSITIONING TASKS FOR A
SET OF 10 EXPERIMENTS

Autonomous recovering task			
step	mean (s)	min (s)	max (s)
1	4.22	0.8	11.96
2	1.64	1.64	1.64
3	1.30	0.48	1.84
4	2.96	1.92	3.84
all	10.12	4.84	19.28
Autonomous 3D positioning task			
step	mean (s)	min (s)	max (s)
3	1.31	1	2.28
4	0.65	0	1.76
5	1.34	0.68	3.28
all	3.30	1.68	7.32

ered to relate variations of the laser image positions and angular variations (see also Fig. 10). One can notice a significant perturbation due to the breathing during visual servoing. For robust identification purposes, we average several measurements of small displacements. This allows reducing the effect of breathing, which acts as a disturbance.

Fig. 15, top and bottom left, shows the 2-D 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 periodical perturbation. Fig. 15, bottom right, shows the measured distance d_0 during the depth servoing at step 5.

Fig. 16, left, displays the laser spot image coordinates when the surgeon specifies new positions to be reached in the image,

at step 6. These results (on living tissues) should be compared with those obtained by the use of an endo-trainer on Fig. 16, right. Note that the fact that the instrument seems to go shortly in the wrong direction at times $t = 10$ s and $t = 20$ s is due to a nonperfect decoupling between u_p and v_p by the identified Jacobian matrix. With our experimental setup, the maximum achieved bandwith is about 1 rad/s. Table 1 shows the time performances of the system. A set of 10 experiments was performed on the instrument recovering task. It takes typically 10 s to bring the instrument in the image center (5 s is the best and 20 s is the worst). For the autonomous 3-D positioning, the time is typically 4 s (2 s is the best and 8 s is the worst). This should be compared with a teleoperated system to the time it takes for a surgeon to command vocally an AESOP system, holding the endoscope, and to bring the instrument and the camera back to the operation field.

VI. CONCLUSION

The robot vision system presented in this paper automatically positions a laparoscopic surgical instrument 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. To position the surgical instrument, we propose a visual servoing algorithm that combines pixel coordinates of the laser spots and the estimated distance between organ and instrument. 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 method is essentially based on visual servoing techniques and online identification of the interaction matrix. It does not require the knowledge of the initial respective position of the endoscope and the surgical instrument.

APPENDIX
DERIVATION OF THE INTERACTION MATRIX

We derive here the relationship between the velocity screw $\dot{W}$ of the instrument and the time derivative of the feature vector $\mathbf{S}$. Let R_L be a reference frame at the tip of the instrument, R_Q the incision point reference frame, and R_C the camera reference frame (see Fig. 7).

Here we derive this interaction matrix in the case where the incision point frame R_Q and the camera frame R_C are motionless. The DOFs that can be controlled are the insertion velocity $\dot{d}_2$ and the rotational velocity of the instrument, $(\mathbf{\Omega}_{L,Q})_{/R_L} = (\omega_x \quad \omega_y \quad \omega_z)^T$, with respect to the incision point frame R_Q.

Let $(\mathbf{V}_{R_Q}^{O_L})_{/R_L}$ be the velocity of the tip of the instrument O_L with respect to the frame R_Q expressed in R_L. We have

$$(\mathbf{V}_{R_Q}^{O_L})_{/R_L} = [\, d_2\omega_y \quad -d_2\omega_x \quad \dot{d}_2 \,]^T. \tag{12}$$

On the other hand, the velocity $(\mathbf{V}_{R_C}^P)_{/R_L}$ of the laser spots center, $P = (x_p, y_p, z_p)^T_{/R_C}$, with respect to the camera frame R_C can be expressed in the instrument frame R_L as follows:

$$
\begin{aligned}
(\mathbf{V}_{R_C}^P)_{/R_L} &= (\mathbf{V}_{R_C}^{O_L})_{/R_L} + (\mathbf{V}_{R_L}^P)_{/R_L} \\
&\quad + (\mathbf{\Omega}_{L,C})_{/R_L} \times \mathbf{O}_L\mathbf{P} \\
&= \mathbf{R}_{CL}^T(\dot{x}_p, \dot{y}_p, \dot{z}_p)^T_{/R_C}
\end{aligned}
\tag{13}
$$

where $\mathbf{R}_{CL} = (r_1, r_2, r_3)^T$ is the rotation matrix between the camera frame and the instrument frame. In the previous equation, $(\mathbf{V}_{R_L}^P)_{/R_L} = \dot{d}_0\mathbf{z}^L$. Since there is no relative motion between R_Q and R_C, $(\mathbf{V}_{R_C}^{O_L})_{/R_L} = (\mathbf{V}_{R_Q}^{O_L})_{/R_L}$ and $(\mathbf{\Omega}_{L,C})_{/R_L} = (\mathbf{\Omega}_{L,Q})_{/R_L}$. From (12) and (13), we have

$$[\, \dot{x}_p \quad \dot{y}_p \quad \dot{z}_p \,]^T = \mathbf{R}_{CL}\,[\, d\,\omega_y \quad -d\omega_x \quad \dot{d} \,]^T \tag{14}$$

where $d = d_0 + d_2$ is the distance between the incision point and the organ.

Considering a pin-hole camera model, P and its perspective projection $p = (u_p, v_p)^T$ are related by

$$
z_p \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ z_p \\ 1 \end{bmatrix}
\tag{15}
$$

where $\mathbf{A} = (a_1, a_2, a_3)^T$ is the (3×3) upper triangular real matrix of the camera parameters. It follows that:

$$
z_p \begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} \dot{x}_p \\ \dot{y}_p \\ \dot{z}_p \end{bmatrix} - \dot{z}_p \begin{bmatrix} u_p \\ v_p \end{bmatrix}.
\tag{16}
$$

Substituting the expression (14) of the velocity of P in (16), one obtains the (2×3) interaction matrix $\mathbf{J}_S$ relating $\dot{\mathbf{S}}_p = (\dot{u}_p, \dot{v}_p)^T$ to the velocity screw $(\omega_x, \omega_y, \dot{d})^T$

$$
\mathbf{J}_S = \frac{1}{z_p} \left[\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \mathbf{R}_{CL} - \begin{bmatrix} u_p \\ v_p \end{bmatrix} r_3 \right] \begin{bmatrix} 0 & d & 0 \\ -d & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{17}
$$

REFERENCES

[1] A. Casals, J. Amat, D. Prats, and E. Laporte, "Vision guided robotic system for laparoscopic surgery," in *Proc. IFAC Int. Congr. Advanced Robotics*, Barcelona, Spain, 1995, pp. 33–36.

[2] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Automat.*, vol. 8, pp. 313–326, June 1992.

[3] M. Hayashibe and Y. Nakamura, "Laser-pointing endoscope system for intraoperative geometric registration," in *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 2, Seoul, Korea, May 2001, pp. 1543–1548.

[4] K. Hosoda and M. Asada, "Versatile visual servoing without knowledge of true Jacobian," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Munich, Germany, Sept. 1994, pp. 186–191.

[5] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Automat.*, vol. 12, pp. 651–670, Oct. 1996.

[6] D. P. Huttenlocher and S. Ullman, "Recognizing solid objects by alignment with an image," *Int. J. Comput. Vis.*, vol. 5, no. 2, pp. 195–212, 1990.

[7] A. Krupa, C. Doignon, J. Gangloff, M. de Mathelin, L. Soler, and G. Morel, "Toward semi-autonomy in laparoscopic surgery through vision and force feedback control," in *Experimental Robotics VII, Lecture Notes in Control and Information Sciences 271*, D. Rus and S. Singh, Eds. New York: Springer, 2001, pp. 189–198.

[8] A. Krupa, M. de Mathelin, C. Doignon, J. Gangloff, G. Morel, L. Soler, and J. Marescaux, "Development of semi-autonomous control modes in laparoscopic surgery using visual servoing," in *Proc. 4th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Utrecht, The Netherlands, Oct. 2001, pp. 1306–1307.

[9] A. Krupa, J. Gangloff, M. de Mathelin, C. Doignon, G. Morel, L. Soler, and J. Marescaux, "Autonomous retrieval and positioning of surgical instruments in robotized laparoscopic surgery using visual servoing and laser pointers," in *Proc. IEEE Int. Conf. Robotics and Automation*, Washington, DC, May 2002, pp. 3769–3774.

[10] E. Malis, F. Chaumette, and S. Boudet, "2-D 1/2 visual servoing," *IEEE Trans. Robot. Automat.*, vol. 15, pp. 238–250, Apr. 1999.

[11] P. Martinet and E. Cervera, "Combining pixel and depth information in image-based visual servoing," in *Proc. 9th Int. Conf. Advanced Robotics*, Tokyo, Japan, Oct. 1999, pp. 445–450.

[12] S. J. Maybank, "The cross-ratio and the j-invariant," in *Geometric Invariance in Computer Vision*, J. L. Mundy and A. Zisserman, Eds. Cambridge, MA: MIT Press, 1992, pp. 107–109.

[13] M. Jagersand, O. Fuentes, and R. Nelson, "Experimental evaluation of uncalibrated visual servoing for precision manipulation," in *Proc. IEEE Int. Conf. Robotics and Automation*, Albuquerque, NM, Apr. 1997, pp. 2874–2880.

[14] J. A. Piepmeier, G. V. McMurray, and H. Lipkin, "A dynamic quasi-Newton method for uncalibrated visual servoing," in *Proc. IEEE Int. Conf. Robotics and Automation*, Detroit, MI, May 1999, pp. 1595–1600.

[15] G.-Q. Wei, K. Arbter, and G. Hirzinger, "Real-time visual servoing for laparoscopic surgery," *IEEE Eng. Med. Biol. Mag.*, vol. 16, pp. 40–45, Jan. 1997.

[16] Y. F. Wang, D. R. Uecker, and Y. Wang, "A new framework for vision-enabled and robotically assisted minimally invasive surgery," *Comput. Med. Imaging and Graphics*, vol. 22, pp. 429–437, 1998.

[17] M. Ribo, "State of the art report on optical tracking," Vienna Univ. Technol., Vienna, Austria, Tech. Rep. 2001-25, 2001.

Alexandre Krupa (S'00–A'03) was born on April 20, 1976 in Strasbourg, France. He received the M.S. (DEA) and Ph.D. degrees in control science and signal processing from the National Polytechnic Institute of Lorraine, Nancy, France, in 1999 and 2003, respectively.

He is currently working in robotics in the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT) at the University of Strasbourg, Illkirch, France. His research interests include medical robotics, visual servoing of robotic manipulators, computer vision, and microrobotics.

Jacques Gangloff (M'97) graduated from the Ecole Nationale Supérieure de Cachan, Cachan, France, in 1995, and received the M.S. and Ph.D. degrees in robotics from the University Louis Pasteur, Strasbourg, France, in 1996 and 1999, respectively.

Since 1999, he has been Maître de Conférences with the University of Strasbourg, Illkirch, France, where he is a member of the EAVR team with the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT). His research interests mainly concern visual servoing of robotic manipulators, predictive control, and medical robotics.

Christophe Doignon (M'00) received the B.S. degree in physics in 1987 and the Engineer diploma in 1989, both from the Ecole Nationale Supérieure de Physique de Strasbourg, Strasbourg, France, and the Ph.D. degree in computer vision and robotics from Louis Pasteur University, Strasbourg, France in 1994.

In 1995 and 1996, he worked with the Department of Electronics and Computer Science at the University of Padua, Padua, Italy, for the European Community under HCM program "Model Based Analysis of Video Information." Since 1996, he has been an Associate Professor in Computer Engineering at the Louis Pasteur University, Strasbourg, France. His major research interests include computer vision, signal and image processing, visual servoing, and robotics.

Michel F. de Mathelin (S'86–M'87) received the Electrical Engineering degree from Louvain University, Louvain-La-Neuve, Belgium, in 1987, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie–Mellon University, Pittsburgh, PA, in 1988 and 1993, respectively.

During the 1991–1992 academic year, he was a Research Scientist with the Electrical Engineering Department, Polytechnic School of the Royal Military Academy, Brussels, Belgium. In 1993, he became Maître de Conférences with the Université Louis Pasteur, Strasbourg, France, where, since 1999, he has been a Professor with the Ecole Nationale Supérieure de Physique de Strasbourg (ENSPS). His research interests include adaptive and robust control, visual servoing, and medical robotics.

Dr. Mathelin is a Fellow of the Belgian American Educational Foundation.

Guillaume Morel (M'97) received the Ph.D. degree from the University of Paris 6, Fontenay-aux-Roses, France, in 1994.

He was a Postdoctoral Research Assistant in the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, in 1995-1996, and an Associate Professor at the University of Strasbourg, Illkirch, France, from 1997-2000. After a year spent as a Research Scientist for the French company of electricity (EDF), he joined the Laboratoire de Robotique de Paris, Paris, France. His research has concerned the sensor-based control of robots, with a particular focus on force-feedback control and visual servoing. Application of these techniques for medical robotics is now the main area of his research.

Joël Leroy was born on December 18, 1949. A graduate of the Medical University of Lille, Lille, France, he completed his residency in digestive surgery at the University Hospital of Lille.

In 1979, he was appointed Chief of the Department of General, Visceral and Digestive Surgery in a private surgical hospital in Bully les Mines, France. Toward the end of the 1980s, he participated in the development of gynecologic laparoscopic surgery. In 1991, he created and developed the first successful surgical division of minimally-invasive surgery specializing in colorectal surgery in France. In 1997, he was nominated Associate Professor of Digestive Surgery at the University of Lille, Lille, France. He is recognized worldwide as an innovative pioneer in the field of laparoscopic digestive surgery. He has been an expert contributor in the development of IRCAD/EITS (Institut de Recherche contre les Cancers de l'Appareil Digestif/Institute for Research into Cancer of the Digestive System) since its creation in 1994. In 1998, he became the Co-Director of IRCAD/EITS, and he was a key participant in the development of the Lindbergh Telesurgery Project, where his pivotal contribution was in the standardization of the ZEUS robotic surgical system assisted laparoscopic cholecystectomy.

Luc Soler was born on October 6, 1969. He received the Ph.D. degree in computer science in 1998 from the University of Paris II, INRIA, Orsay, France.

Since 1999, he has been the Research Project Manager in Computer Science at the Digestive Cancer Research Institute (IRCAD, Strasbourg, France). His principal areas of interest are computerized medical imaging and especially automated segmentation methods, discrete topology, automatic atlas definition, organ modeling, liver anatomy, and hepatic anatomical segmentation.

In 1999, Dr. Soler was Co-Laureate of a Computer World Smithsonian Award for his work on virtual reality applied to surgery.

Jacques Marescaux was born on August 4, 1948. He received the Ph.D. surgery award in 1977 from the Medical University Louis Pasteur, Strasbourg, France.

In 1977, he joined INSERM, the French Institute of Health and Medical Research, where he became a Professor in the digestive surgery department. In 1992, he founded the Institute for Research into Cancer of the Digestive System (IRCAD) and the European Institute of Telesurgery (EITS) in 1994, both in Strasbourg, France. Since 1989, he has been Head of Digestive and Endocrine Surgery Department at Strasbourg University Hospitals.

In June 2003, Dr. Marescaux received a ComputerWorld Honors award for the Lindbergh Transatlantic Telesurgery Project (September 7th, 2001) in partnership with Computer Motion, Inc. (Santa Barbara, CA) and France Telecom. He is a member of numerous scholarly societies.