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 3-D 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 to localize 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 multi-stages 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 IRCAD.

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.) or DaVinci (Intuitive Surgical, Inc.). With these systems, the surgeon tele-operates the robot 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 at expanding the potentialities of such systems by providing “automatic modes” using visual servoing (see [7], [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, towards their desired location.

Note that prior research was conducted on visual servoing techniques in laparoscopic surgery to automatically guide the camera towards the region of interest (see, e.g., [15], [1] 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 to do away with the practice of moving the endoscope in order to visualize 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

A. Krupa†, J. Gangloff†, C. Doignon†, M. de Mathelin†
G. Morel‡, J. Leroy ‡, L. Soler‡, J. Marescaux‡

† Laboratoire des Sciences de l’Image de l’Informatique et de la Télédétection (LSIIT - UMR CNRS 7005), Strasbourg I University, France. E-mail: krupa.gangloff.doignon.demathelin@lsiit.u-strasbg.fr
‡ G. Morel is now with the LRP (CNRS FRE 2705), Paris VI University, France.
† IRCAD (Institut de Recherche sur les Cancers de l’Appareil Digestif), Strasbourg, France. 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 AESOP medical robot. In particular, we would like to thank Dr. Moji Ghodoussi for his technical support.
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
3-D geometric registration system is presented. The authors
add a second endoscope with an optical galvano-scanner.
Then, a 955 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 (OptotrakTM 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 3D
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
synchronized with image acquisition [17]. The most famous
among these systems is the OptotrakTM from Northern Digital
Inc. which uses synchronized infrared LED markers tracked
by 3 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 3D measurement and
surgeon visual feedback. To do so, we propose a novel
method to extract efficiently, in real-time, with a high 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. The next section
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 the last section, we show experimental
results in real surgical conditions at the operating room of
IRCAD.

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 to move 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
device which provides the visual feedback and whose location
relatively 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 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 use 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 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 problems 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 fuzzy contours.

- **iii) 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 by 5 vertical high-pass filter mask shows that, the two markers can be easily detected with a simple thresholding procedure. 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 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 than 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 CMS (Component Mounted on Surface) LEDs 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 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.

![Image](image.png)

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).

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

![Image](image.png)

Fig. 6. (left) Detection of optical markers and laser spots (+) - (right) Contours detection of optical markers (odd frame).

For safety reasons, we’ve added the simple following test: the signal to noise 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.

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 do 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, 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’s 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, $\tau$, 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 a). Since a 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 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 $\Delta$ corresponding to the instrument axis is transformed, in the image, into a line $\delta = H(\Delta)$.

$$\tau = \frac{pp_2}{pp_3} = \frac{P_1 P_2}{P_1 P_3}$$

$$d_0 = \frac{P_1 P_3}{P_2} = (1 - \tau) \frac{P_1 P_3}{P_2}$$

where $\alpha$ and $\beta$ depend only on the known relative position of $P_1$, $P_2$ and $P_3$. Similar computations lead to the same relationship between $d_2$ and another cross-ratio $\mu$ 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 $\delta$ and $\delta'$ in the image. Then, $o_q$ is the intersection of these lines since $O_Q$ is motionless. Finally:

$$
\mu = \frac{\begin{vmatrix} P_1 & P_2 \\ P_3 & P_4 \end{vmatrix}}{\begin{vmatrix} P_0 & P_1 \\ P_0 & P_2 \end{vmatrix}} = \frac{\begin{vmatrix} z_Q \\ d_0 \end{vmatrix}}{\begin{vmatrix} z_Q \\ d_0 \end{vmatrix}}
$$

(3)

$$
d_2 = P_1 O_Q = \frac{\mu}{\mu + \frac{\beta}{1-\beta}}
$$

(4)

Fig. 8 shows the results of a sensitivity analysis on the depth estimation. The standard deviation $\sigma_{d_2}$ of the estimated depth is plotted as a function of the standard deviation $\sigma_p$ 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 $\alpha$ 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, $s_{\sigma} = \frac{d(\sigma_{d_2})}{d \sigma_p}$ in the Fig. 8) is proportional to the distance $d_C$ and $d_0$. The sensitivity, $s_{\sigma}$, varies in the interval [0.4, 3] corresponding to [0.4, 3] mm for $\sigma_{d_2}$ if $\sigma_p = 1$ pixel. Experimentally, $\sigma_p$ is typically 0.5 pixel, resulting in $\sigma_{d_2} = 1$ 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], [11]), however the depth, $d_0$, of concern here is independant 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 $S = [S_p^T \ d_0]^T$. 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 degrees of freedom $W_{op} = (\omega_x, \omega_y, v_z = \dot{d}_2)^T$ (see the appendix for more details):

$$
\begin{bmatrix}
\dot{u}_p \\
\dot{v}_p \\
\dot{\omega}_x \\
\dot{\omega}_y \\
\dot{v}_z
\end{bmatrix}
= \begin{bmatrix}
J_{11} & J_{12} & J_{13} \\
J_{21} & J_{22} & J_{23} \\
J_3 & \ddot{d}_0 + v_z
\end{bmatrix}
\begin{bmatrix}
\omega_x \\
\omega_y \\
v_z
\end{bmatrix}
$$

Even though all components of $J_S$ could be recovered from images of optical markers and camera parameters, $J_S$ is not invertible. Therefore, the velocity screw applied to the robot, $W_{op}^* = (\omega_x^*, \omega_y^*, v_z^*)^T$, cannot be directly computed without 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 splitted into three stages:
Fig. 9. The full visual servoing scheme.

**Instrument recovery and positioning procedure:**

- **Stage 1:** Positioning of the laser spot projection, $p$, at the center of the image by visual servoing of $S_p = (u_p, v_p)^T$. It means that only $\omega_x$ and $\omega_y$ are controlled. For safety reasons, $v_z^* = 0$ during this stage. Thus, from (5), it comes:

$$
\begin{bmatrix}
\dot{u}_p \\
\dot{v}_p
\end{bmatrix} =
\begin{bmatrix} J_{13} & J_{23} \end{bmatrix}
\begin{bmatrix}
\omega_x \\
\omega_y
\end{bmatrix}
\begin{bmatrix}
\omega_x^* \\
\omega_y^*
\end{bmatrix}
\begin{bmatrix} \dot{d} \\
J_{13} \dot{d} + J_{23} \dot{d}
\end{bmatrix}
$$

(6)

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

$$
\begin{bmatrix}
\omega_x^* \\
\omega_y^*
\end{bmatrix} = K^{-1}
\begin{bmatrix} u_p - u_p \\
v_p - v_p\end{bmatrix}
\begin{bmatrix} J_{13} \\
J_{23}\end{bmatrix}
\begin{bmatrix}
\omega_x \\
\omega_y
\end{bmatrix}
$$

(7)

where $K$ 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 $W_{op} = (0, 0, v_z^*)^T$.

- **Stage 3:** Full visual servoing.

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

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

(8)

where $k$ is a positive scalar and $d_0$ is a function of the cross-ratio $\tau$, $\alpha$ and $\beta$. The full servoing scheme is shown in Fig. 9.

**C. Implementation issues**

The signal $\dot{d}$ in equations (7) and (8) can be obtained by deriving equations (2) and (4). However, since $\dot{d}$ 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 $(2 \times 2)$ sub-matrix of $J_{s}$ must be computed even if the optical markers are not visible. When the instrument is out of the field of view, this sub-matrix is identified in an initial procedure. This identification consists in applying a constant rotational velocity reference $W_{op} = (\omega_x^*, \omega_y^*, 0)^T$ during a short time interval $\Delta T$ (see Fig. 11). Small variations of laser spot image coordinates are measured and the estimated $\hat{J}_w$, of the interaction matrix is given by:

$$
\hat{J}_w =
\begin{bmatrix}
\frac{\Delta u_p}{\Delta T} & \frac{\Delta v_p}{\Delta T} \\
\frac{\Delta u_y}{\Delta T} & \frac{\Delta v_y}{\Delta T}
\end{bmatrix}
$$

(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 ($v_z^* = 0$), no bias appears on the laser spot centering. Therefore, it is recommended in practice to choose the interaction matrix $M_I$, mapping $W_{op}$ into $\hat{S}$ with the following structure:

$$
M_I =
\begin{bmatrix}
\hat{J}_w & 0_{1 \times 2} \\
0_{1 \times 2} & -1
\end{bmatrix}
$$

(10)

This leads to the experimental control scheme shown in Fig. 10, with $K > 0$. The bandwidth of the visual control loop is directly proportional to $K$.

For the stability analysis, we consider an ellipsoid as a geometric model for the abdominal cavity, so that $\dot{d}$ is related to $\omega_x$ and $\omega_y$. In this case, the interactions matrix, in equation (5), is reduced to a $2 \times 2$ matrix $J_w$:

$$
\begin{bmatrix}
\dot{u}_p \\
\dot{v}_p
\end{bmatrix} =
\begin{bmatrix} J_{13} & J_{23} \end{bmatrix}
\begin{bmatrix}
\omega_x \\
\omega_y
\end{bmatrix}
$$

(11)

and the stability of the visual feedback loop is guaranteed as long as $J_w[\hat{J}_w]^{-1}$ remains positive definite [2]. In our application, if the camera and the incidence 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 modelled the organ as an ellipsoid. The estimated Jacobian $J_w$ is constant and correspond to a nominal configuration. We have then computed $J_w$ when the laser spot is moved across the organ surface, and computed the eigenvalues of $J_w[\hat{J}_w]^{-1}$ in the different configurations. Unsafe
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 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 ComputerMotion surgical robot. A standard 50fps 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 ($\omega^a_x$ and $\omega^a_y$) 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. Eq. (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 $\nu^a_z$ in open loop until it

![Image of experimental setup](image)

Fig. 13. The experimental setup.
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 towards a desired 3-D location by automatic visual servoing under the surgeon’s control. 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 3D positioning.

### C. Experimental measurements

Fig. 14 show experimental measurements of the laser image coordinates $u_p$ and $v_p$ during the identification stage of $J_p$. For the identification procedure, four positions have been considered to relate variations of the laser image positions and angular variations (see also Fig. 11). One can notice a significant perturbation due to the breathing during visual servoing. For robust identification purpose, we average several measurements of small displacements. This allows to reduce 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(a)(b) 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(c)(d). Note that the fact that the instrument seems to go slightly in the wrong direction at times $t = 10$ s and $t = 20$ s is due to a non perfect decoupling between $u_p$ and $v_p$ by the identified Jacobian matrix. With our experimental setup, the maximum achieved bandwidth is about 1 rad/sec. Fig. 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 3D 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 on-line identification of the interaction matrix. It does not require the knowledge of the initial respective position of the endoscope and the surgical instrument.

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). Responses (a)-(b) are obtained on living tissue. Responses (c)-(d) are obtained with an endo-trainer with no disturbance due to breathing. (1 mm ≈ 25 pixels)

APPENDIX - DERIVATION OF THE INTERACTION MATRIX
We derive here the relationship between the velocity screw $\mathbf{W}$ 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 degrees of freedom that can be controlled are the insertion velocity $\mathbf{d}_L$ and the rotational velocity of the instrument, $(\Omega_{L,Q})_{R_L} = (\omega_x, \omega_y, \omega_z)^T$, with respect to the incision point frame $R_Q$.

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

$$\begin{align*}
(V^O_{R_Q})_{R_L} &= [d_2 \omega_y - d_2 \omega_x \ d_2]^T
\end{align*}$$

(12)

On the other hand, the velocity $(V^P_{R_Q})_{R_L}$ 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:

$$\begin{align*}
(V^P_{R_C})_{R_L} &= (V^O_{R_C})_{R_L} + (V^P_{R_Q})_{R_L} + (\Omega_{L,C})_{R_L} \times \mathbf{O}_L \mathbf{P} \\
&= R^T_{CL} (x_p, y_p, z_p)^T_{R_C}
\end{align*}$$

(13)

where $R_{CL} = (r_{11}, r_{12}, r_{13})^T$ is the rotation matrix between the camera frame and the instrument frame. In the previous

REFERENCES


equation, \((V_p^p)^{R_L}) = d_0 \cdot \mathbf{z}^L\). Since there is no relative motion between \(R_Q\) and \(R_C\), \((V_Q^{O_L})/R_L = (V_Q^{O_L})/R_L\) and \((\Omega_{L,C})/R_L = (\Omega_{L,C})/R_L\). From equations (12) and (13) it comes:
\[
\begin{bmatrix}
\dot{x}_p \\
\dot{y}_p \\
\dot{z}_p \\
\end{bmatrix}
= R_{CL} \begin{bmatrix}
d \omega_y \\
-d \omega_x \\
\dot{d} \\
\end{bmatrix}
\] (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:
\[
\begin{bmatrix}
u_p \\
v_p \\
1 \\
\end{bmatrix}
= A \begin{bmatrix}
x_p \\
y_p \\
z_p \\
\end{bmatrix}
\] (15)

where \(A = (a_1, a_2, a_3)^T\) is the \((3 \times 3)\) upper triangular real matrix of the camera parameters. It follows that:
\[
\begin{bmatrix}
u_p \\
v_p \\
\end{bmatrix}
= R_{CL} \begin{bmatrix}
u_p \\
v_p \\
\end{bmatrix}
\] (16)

Substituting the expression (14) of the velocity of \(P\) in (16), one obtains the \((2 \times 3)\) interaction matrix \(J_S\) relating \(\dot{S}_p = (\dot{u}_p, \dot{v}_p)^T\) to the velocity screw \((\omega_x, \omega_y, \dot{d})^T\):
\[
J_S = \frac{1}{z_p} \begin{bmatrix}
a_1 \\
a_2 \\
\end{bmatrix} R_{CL} - \begin{bmatrix}
u_p \\
v_p \\
\end{bmatrix} r_3 \begin{bmatrix}
0 \\
-d \\
0 \\
0 \\
0 \\
1 \\
\end{bmatrix}
\] (17)

Alexandre Krupa was born on April 20, 1976 in Strasbourg, France. He received the M.S. degree (DEA) in Control Science and Signal Processing from the National Polytechnic Institute of Lorraine, Nancy, France in 1999. He is currently working toward the completion of his Ph.D. degree in robotics in the Laboratoire des Sciences de l’Image, de l’Informatique et de la Télédétection (LSIIT) at the University of Strasbourg. His research interests include medical robotics, visual servoing and robotic manipulators, computer vision and micro-robotics.

Jacques Gangloff graduated from the Ecole Nationale Supérieure de Cachan in 1995. He received the M.S. degree and the Ph.D. degree in robotics from the University Louis Pasteur, Strasbourg, France in 1996 and 1999 respectively. Since 1999, he is Maître de Conférences at the University of Strasbourg. He is member of the EAER team at the Laboratoire des Sciences de l’Image, de l’Informatique et de la Télédétection (LSIIT). His research interests include mainly visual servoing of robotic manipulators, predictive control, and medical robotics.

Christophe Doignon 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, France. He received 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 University of Padua, Italy, for the European Community under HCM program “Model Based Analysis of Video Information”. Since 1996, he’s 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 de Mathelin received the electrical engineering degree from Louvain University, Louvain-La-Neuve, Belgium, in 1987. He completed his M.S. and Ph.D. degrees in electrical and computer engineering at Carnegie Mellon University, Pittsburgh, PA, USA, in 1988 and 1993 respectively. He was a research scientist in the electrical engineering department of the Polytechnic School of the Royal Military Academy, Brussels, Belgium, during the 1991-1992 academic year. In 1993, he became Maître de Conférences at the Université Louis Pasteur (Strasbourg I University), France. Since 1999, he is Professor at the Ecole Nationale Supérieure de Physique de Strasbourg (ENSFPS). He is a fellow of the Belgian American Educational Foundation. His research interests include adaptive and robust control, visual servoing and medical robotics.

Guillaume Morel obtained his Ph.D. at the University of Paris 6 in 1994. He has been successively a postdoctoral research assistant at the dept. of Mech. Eng. of the Massachusetts Institute of Technology (’95-’96), and an associate professor at the University of Strasbourg, France (’97-’00). After a year passed in the industrial world, as a research scientist for the French company of electricity (EDF, ’00-’01) he is back to the academic world since sept 2001, at the Laboratoire de Robotique de Paris, France. During all these years, his research has concerned the sensor based control of robot, 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, France, Professor Leroy 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. During the end of the 1980's, he participated in the development of gynecologic laparoscopic surgery. In 1991, he created and developed the first successful surgical division of minimally-invasive surgery (MIS) specialising in colorectal surgery in France. In 1997, he was nominated Associated Professor of Digestive Surgery at the University of Lille, France. Professor Leroy is recognized world-wide 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 joined Professor Jacques Marescaux’s pioneering team full time as the Co-Director of IRCAD/EITS, and he was a key participant in the development of the Lindbergh Telesurgery Project, led by Professor Jacques Marescaux. Professor Leroy’s pivotal contribution was in the standardisation of the ZEUS Robotic Surgical System assisted laparoscopic cholecystectomy. This important standardisation helped enable Professor Marescaux’s performance of the world’s first telesurgery, the ZEUS Robotic Telesurgery Surgical System assisted laparoscopic cholecystectomy on September 7, 2001.

Luc Soler was born on October 6, 1969. In 1994, he was validictorian at the magister at the Computer Science High School of Paris. He obtained his Ph.D. degree in computer science in 1998. Since 1999, he is the research project manager in computer science at the Digestive Cancer Research Institute (IRCAD Strasbourg). The same year, he is co-laureate of a Computer World Smithsonian Award for his work on virtual reality applied to surgery. 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.

Jacques Marescaux was born on August 4, 1948. He quickly set his sights on a medical career, and his passion increased as he advanced in his medical studies. This commitment was apparent in his ranking at the top of his class each year, including the top score on the internship exam of the Medical University of Strasbourg, France, in 1971. He received a Ph.D. surgery award in 1977 and he joined a team of researchers at INSERM, the French Institute of Health and Medical Research. Thanks to this exceptional scientific collaboration, he was able to apply a position as university professor at the age of 32, in the digestive surgery department. In 1992, he decided to create a truly innovative structure for teaching and research. He was firmly convinced that surgeons could not only play a role in the information revolution, but act as high-profile ambassadors, so he founded the Institute for Research into Cancer of the Digestive System (IRCAD) and the European Institute of Telesurgery (EITS) in 1994. Professor Jacques Marescaux has since 1989 been Head of Digestive and Endocrine Surgery Department at Strasbourg University Hospitals. In June 2003, he received a ComputerWorld Honors award for the Lindbergh transatlantic Telesurgery project (September 7th, 2001) in partnership with Computer Motion Incorporation and France Telecom. He is a member of numerous scholarly societies.