

Combined image-based and depth visual servoing applied to robotized laparoscopic surgery

Alexandre Krupa, Christophe Daignon, Jacques Gangloff, Michel de Mathelin

LSIIT (UMR CNRS 7005) - Strasbourg I University, France
{alexandre.krupa,christophe.daignon,...}@ensps.u-strasbg.fr

Abstract

In this paper, we address the problem of controlling the motion of a surgical instrument closed to an unknown organ surface by visual servoing in the context of robotized laparoscopic surgery. To achieve this goal, a visual servoing algorithm is developed that combines feature errors in the image and errors in depth measurements. The relationship between the velocity screw of the surgical instrument, the depth and the motion field is defined and a two-stages servoing scheme is proposed. In order to measure the orientation and the depth of the instrument with respect to the organ, a laser dot pattern is projected on the organ surface and optical markers are stucked on the instrument. Our work has been successfully validated with a surgical robot by performing experiments on living tissues in the surgical training room of IRCAD.

1 Introduction

Robots have recently appeared in the field of laparoscopic surgery¹. There exist today several commercial systems, *e.g.* AESOP and ZEUS (Computer Motion, Inc.) or DaVinci (Intuitive Surgical, Inc.). These robotic systems allow to manipulate the surgical instrument as well as the endoscope, with increased confort and precision even from a long distance (see [9]). Our research in this field is about providing new functionalities to this type of system by using visual servoing technique to realize semi-autonomous tasks (see earlier work in [6],[7]).

There exist several difficulties in designing a visual servoing scheme in a laparoscopic surgical environment. One difficulty is the unknown relative position between the camera and the robot arm holding the instrument. Another difficulty may be the use of monocular vision that lacks depth information.

¹Laparoscopic surgery is a surgical procedure where small incisions are made in the body (usually, the abdomen) in order to introduce, through trocars, the surgical instruments as well as the endoscopic optical lens of a camera.

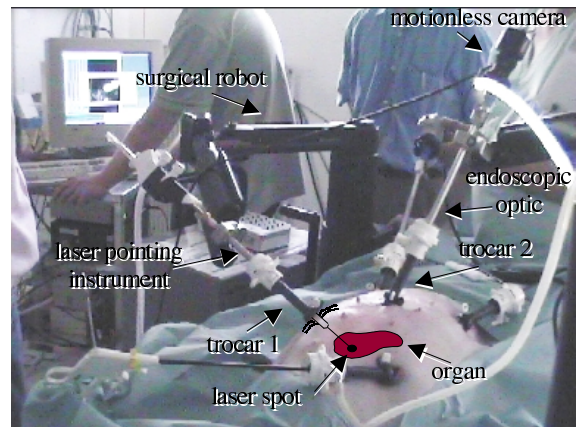


Figure 1: The surgical robotic visual servoing system in the IRCAD surgical training.

A third difficulty lies in the fact that the scene is highly unstructured with varying lighting conditions and with a moving background due to breathing and heart beating. Prior research has been conducted on visual servoing techniques in laparoscopic surgery (see *e.g.* [13], [1], [3] and [14]).

Casals *et al.* [1] use patterned marks on the instrument mounted on an industrial robot to realize an instrument tracking task. Projections of marks are approximated by straight lines for the image segmentation process. This guidance system works at a relatively slow sampling rate of 5 Hz with the aid of an assistant. Wei *et al.* [13] use a color stereovision system to realize a tracking task by the endoscope mounted on a robotic arm. They analyze the color histogram of typical laparoscopic images and they select the color with the lowest value in the histogram to mark the instrument. This spectral mark is used to control the robot motion at a sampling rate of 15 Hz. An interesting feature of the proposed method is the choice of the HSV color space for the segmentation, resulting in a good robustness with respect to lighting variations. In [3], an intraoperative 3-D geometric registration system is presented.



Figure 2: The endoscopic laser pointing instrument with optical markers (3 LEDs).

The authors put 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, their method requires to know the relative position between the laser-pointing endoscope and the camera which is estimated by adding an external camera watching the whole surgical scene (OPTOTRAK system).

In this paper, we present a visual servoing system to control the displacement and precise 3-D positioning with respect to the organ of the surgical instrument held by a robotic arm. Our laparoscopic surgical visual servoing system is shown in Figure 1. One arm is holding the endoscope and a second arm is holding the instrument which is mounted on a laser-pointer instrument holder. This instrument holder projects laser pattern on the organ surface in order to provide information about the relative orientation of the instrument with respect to organ, even if the instrument is not in the field of view. Optical markers have been added on the tip of the surgical instrument. These markers (composed of three circular LEDs) are directly projected onto the image and in conjunction with images of the laser pattern, they are used to guide the instrument. Because of the poor structuration of the scene and the variations of lighting, several optical markers are used (see Figure 2). We combine image feature coordinates and depth information while positioning the instrument with respect to the pointed organ. Note that there exist previous works about this combination (see *e.g.* [8], [10]). However here, the depth of concern is the distance between organ and instrument since this information can be extracted with an uncalibrated camera. Using the task-function approach, we derive the relationship between velocity screw of the instrument and time-derivative of the feature vector.

The paper is organized as follows. In the next section, the laser pattern and optical markers detection is briefly explained. Section 3 details the depth

estimation and the 2-stage visual servoing scheme used to position the surgical instrument. Section 4 is devoted to a discussion about experimental results in real surgical conditions at the operating room of IRCAD ("Institut de Recherche contre le Cancer de l'Appareil Digestif", Strasbourg, France - Director Prof. J. Marescaux).

2 Detection of image features

Due to the viscosity of organs, surfaces are bright and laser spots are diffused. In addition, complexity of the organ surface may lead to some occultations. Taking these phenomena into account, a robust detection of the projected laser spots in the image plane is suitable. The main difficulty is to localize projections of diffused laser spots in presence of images speckle. During experiments *in vivo*, we have found that the four laser spots barycentre is always detected with a small bias even if three spots are occulted. Our technique consists in synchronizing laser sources with the image capture board in such a way that laser sources are turned on during the even field of image frames and are turned off during the odd field. In this manner, laser spots are visible one line out of two and are detected with a high-pass spatial 2-D FIR filter, followed by an image thresholding and erosion (see Figure 3).



Figure 3:

<i>a</i>	<i>b</i>
<i>c</i>	<i>d</i>

 Detection of laser spots. - *a*) original interlaced image. - *b*) high-pass filtering and thresholding on even frame. - *c*) erosion. - *d*) localization of barycentre (square).

Optical markers are also detected in a similar fashion. However, they are turned on during the odd field and turned off during the even field. The same high-pass filter in combination with an edge detector is applied on the odd field. Edges detector always yields contours with many pixels of thickness. Thin-

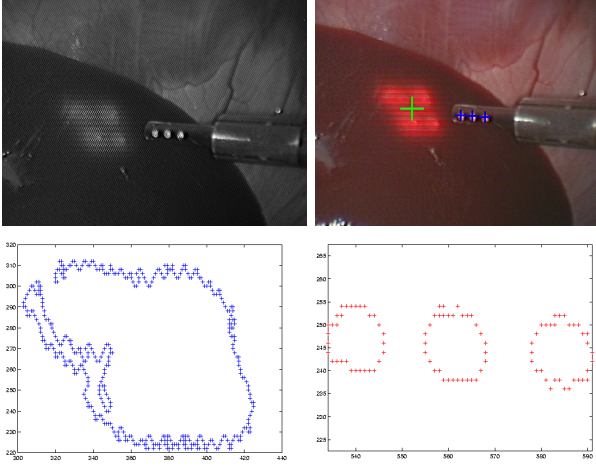


Figure 4:

a	b
c	d

 a) Original image. - b) Detection of optical markers and laser spots (+). - c) Binary contours of laser spot (even frame). - d) Contours detection of optical markers (odd frame).

ning operations are performed on the extracted set of pixels (non maxima suppression) producing a 1-pixel wide edge : this is a requirement in order to apply hysteresis thresholding and an edges tracking algorithm. Contours are merged by using a method called *mutual favorite pairing* [4] in order to merge neighboring contour chains into a single chain. Then, they are used to fit ellipses (see Figure 4).

3 Visual servoing

The objective of the proposed visual servoing is to guide and 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 strictly forbidden (see Figure 5). For practical convenience, rotation around the instrument axis is constrained in a way to keep optical markers visible. A slow visual servoing is performed, based on the ellipses minor/major semiaxes ratio. Since this motion does not contribute to position the tip of the instrument, it's not further considered. Next, we detail how the three remaining degrees of freedom are related to image features and depth.

3.1 Depth estimation

To perform the 3-D positioning, we need to estimate the distance between the instrument and the pointed organ (depth d_0 in Figure 5). Since the three optical markers centers P_1 , P_2 and P_3 are placed along the instrument axis, we assumed they are collinear with the laser spots barycentre P . Under this assumption,

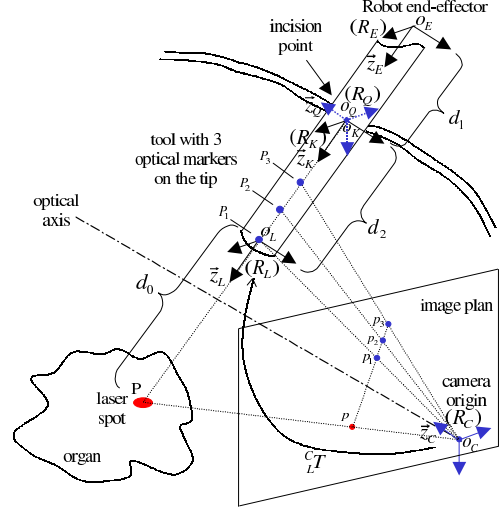


Figure 5: The basic geometry involved.

a cross ratio can be computed from this four points [11]. This projective invariant can also be computed in the image using their respective projections p_1 , p_2 , p_3 and p (see Figures 5 and 6) and can be used to estimate the depth d_0 . In other words, 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 cross ratio is a projective invariant built with the fourth point (P or p). Consequently, an homography H exists between these two bases, so that the straight line Δ corresponding to the instrument axis is transformed, in the image, into a line $\delta = H(\Delta)$ as shown in Fig. 6. The cross-ratio τ is given by:

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

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

where α and β depend only on the known relative position of P_1 , P_2 and P_3 . To simplify the computation of the cross-ratio in the image plane, it's necessary to characterize the straight line δ in order to relate the pixels coordinates of an image point $p = (u, v, 1)^t$ and its projective coordinates $(s\lambda, s)^t$ on δ . Let $(-b, a)^t$ be the normalized cosine directeur of δ and p_k a point on δ . It comes:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} -b & u_k \\ a & v_k \\ 0 & 1 \end{bmatrix}}_{\mathbf{F}} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \quad (3)$$

or

$$\lambda = \begin{bmatrix} -b & a \end{bmatrix} \begin{bmatrix} u - u_k \\ v - v_k \end{bmatrix} \quad (4)$$

with :

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \quad (5)$$

for a point on δ ($-c$ is the orthogonal distance from δ to the image origin). The computation of the cross-ratio is then:

$$\tau = \frac{\lambda_0 + \overline{p_1 p_2}}{\overline{p_1 p_2}} \frac{\overline{p_1 p_3}}{\lambda_0 + \overline{p_1 p_3}} \quad (6)$$

From equation (2), it's straightforward that \dot{d}_0 is a function of $\dot{\tau}$ which, in turn into the image plane, is a function of $\dot{\lambda}_0$, $\overline{p_1 p_2}$ and $\overline{p_1 p_3}$. Similar computations lead to the same relationship 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 given that O_Q is invariant. Finally :

$$\mu = \frac{\left(\frac{\overline{p_1 p_3}}{\overline{p_2 p_3}} \right)}{\left(\frac{\overline{p_1 o_q}}{\overline{p_2 o_q}} \right)} = \frac{\left(\frac{\overline{P_1 P_3}}{\overline{P_2 P_3}} \right)}{\left(\frac{\overline{P_1 O_Q}}{\overline{P_2 O_Q}} \right)} \quad (7)$$

$$d_2 = \overline{P_1 O_Q} = \frac{\frac{\alpha}{1-\beta}}{\mu + \frac{\beta}{1-\beta}} \quad (8)$$

3.2 Interaction matrix

A feature vector \mathbf{S} is built with image coordinates of the perspective projection of laser spot barycentre $p = (u_p, v_p)^t$ and the depth d_0 between the pointed organ and the instrument. Let R_L be a reference

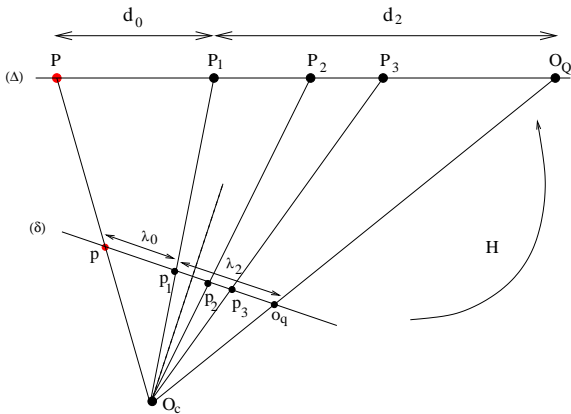


Figure 6: Markers P_1, P_2, P_3 on the tool axis Δ and their images p_1, p_2, p_3 on the straight line δ . Note that $o_q = H(O_Q)$ is invariant during experiments.

frame at the tip of the instrument, R_K the trocar reference frame, R_Q the incision point reference frame, and R_C the camera reference frame (see figure 5). The key issue is to express the interaction matrix relating the derivative of \mathbf{S} and the velocity screw of the surgical instrument $\{(\mathbf{V}_{R_Q}^{O_L})_{/R_L}, (\boldsymbol{\Omega}_{Q,L})_{/R_L}\}$. Here we derive this interaction matrix in the case where incision point frame R_Q and the camera frame R_C are motionless. The instrument can move with a translational velocity $(\mathbf{V}_{R_K}^{O_L})_{/R_L} = (0 \ 0 \ v_z)^t$ with respect to the trocar frame R_K and a rotational velocity $(\boldsymbol{\Omega}_{L,Q})_{/R_L} = (\omega_x \ \omega_y \ \omega_z)^t$ with respect to the incision point frame R_Q . That is :

$$\begin{aligned} (\boldsymbol{\Omega}_{L,Q})_{/R_L} &= (\boldsymbol{\Omega}_{L,C})_{/R_L} \\ &= -\mathbf{R}_{CL}^t (\boldsymbol{\Omega}_{C,L})_{/R_C} \end{aligned} \quad (9)$$

$\mathbf{R}_{CL} = (r_1, r_2, r_3)^t$ is the rotation matrix between camera frame and instrument frame, and the translational velocity with respect to R_Q is :

$$\begin{aligned} (\mathbf{V}_{R_Q}^{O_L})_{/R_L} &= (\mathbf{V}_{R_K}^{O_L})_{/R_L} + \underbrace{(\mathbf{V}_{R_Q}^{O_K})_{/R_L}}_{=0} \\ &+ (\boldsymbol{\Omega}_{Q,L})_{/R_L} \times \mathbf{O}_K \mathbf{O}_L \\ &= (d_2 \omega_y \ -d_2 \omega_x \ \dot{d}_2)^t \end{aligned} \quad (10)$$

On the other hand, the velocity of the laser spot barycentre, $P = (x_p, y_p, z_p)^t_{/R_C}$, with respect to the camera frame R_C can be readily expressed in the instrument frame R_L as follows :

$$\begin{aligned} (\mathbf{V}_{R_C}^P)_{/R_L} &= (\mathbf{V}_{R_C}^{O_L})_{/R_L} + (\mathbf{V}_{R_L}^P)_{/R_L} \\ &+ (\boldsymbol{\Omega}_{C,L})_{/R_L} \times \mathbf{O}_L \mathbf{P} \\ &= \mathbf{R}_{CL}^t \cdot (\dot{x}_p, \dot{y}_p, \dot{z}_p)^t_{/R_C} \end{aligned} \quad (11)$$

In the previous equation, $(\mathbf{V}_{R_L}^P)_{/R_L} = \dot{d}_0 \mathbf{z}^L$ and since there is no relative motion between R_Q and R_C , $(\mathbf{V}_{R_C}^{O_L})_{/R_L}$ is equal to $(\mathbf{V}_{R_Q}^{O_L})_{/R_L}$. Equations (9), (10), (11) provide the relationship between the velocity of P , the velocity screw and the distance between organ and incision point ($d = d_0 + d_2$) :

$$\begin{bmatrix} \dot{x}_p & \dot{y}_p & \dot{z}_p \end{bmatrix}^t = \mathbf{R}_{CL} \begin{bmatrix} d \omega_y & -d \omega_x & \dot{d} \end{bmatrix}^t \quad (12)$$

Considering a pin-hole camera model, P and its perspective projection p are related by :

$$z_p \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ z_p \\ 1 \end{bmatrix} \quad (13)$$

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

$$z_p \begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} \dot{x}_p \\ \dot{y}_p \\ \dot{z}_p \end{bmatrix} - \dot{z}_p \begin{bmatrix} u_p \\ v_p \end{bmatrix} \quad (14)$$

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

$$\mathbf{J}_S = \frac{1}{z_p} \left[\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \mathbf{R}_{CL} - \begin{bmatrix} u_p \\ v_p \end{bmatrix} r_3 \right] \begin{bmatrix} 0 & d & 0 \\ -d & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

$$\text{with } \begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \mathbf{J}_S \begin{bmatrix} \omega_x \\ \omega_y \\ \dot{d}_0 + v_z \end{bmatrix} \quad (16)$$

Similar relationships can be obtained from the time derivatives of p_1 , p_2 or p_3 :

$$\begin{bmatrix} \dot{u}_{p_i} \\ \dot{v}_{p_i} \end{bmatrix} = \mathbf{J}_{S_i} \begin{bmatrix} \omega_x \\ \omega_y \\ v_z \end{bmatrix} \quad (17)$$

Even though all components of $\mathbf{J}_S = (J_{ij})$ could be recovered from images of optical markers and camera parameters, \mathbf{J}_S is not inversible. To relate the velocity screw $(\omega_x, \omega_y, v_z)^t$ and $\dot{\mathbf{S}}$, equation (16) is rewritten as:

$$\begin{bmatrix} 1 & 0 & -J_{13} \\ 0 & 1 & -J_{23} \end{bmatrix} \begin{bmatrix} \dot{u}_p \\ \dot{v}_p \\ \dot{d}_0 \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ v_z \end{bmatrix} \quad (18)$$

Therefore, the velocity screw applied to the robot cannot be directly computed without some additional informations like, *e.g.* the estimation of the depth d_0 . In the previous section, we have seen that, by the use of the cross-ratio τ , $\dot{d}_0 = f(\tau, \alpha, \beta) \dot{\tau}$. Since $\dot{\tau} = f_{\lambda_0} \dot{\lambda}_0 + f_{12} \frac{\dot{p}_1 \dot{p}_2}{p_1 p_2} + f_{13} \frac{\dot{p}_1 \dot{p}_3}{p_1 p_3}$ and since $\frac{\dot{p}_i \dot{p}_j}{p_i p_j}$ are related to the velocity screw, \dot{d}_0 is a function of the velocity screw and $(\dot{u}_p, \dot{v}_p)^t$. Indeed, we have:

$$\begin{aligned} \frac{\dot{p}_i \dot{p}_j}{p_i p_j} = (\dot{\lambda}_j - \dot{\lambda}_i) &= \begin{bmatrix} -b & a \end{bmatrix} \begin{bmatrix} \dot{u}_{p_j} - \dot{u}_{p_i} \\ \dot{v}_{p_j} - \dot{v}_{p_i} \end{bmatrix} \\ &+ \begin{bmatrix} -\dot{b} & \dot{a} \end{bmatrix} \begin{bmatrix} u_{p_j} - u_{p_i} \\ v_{p_j} - v_{p_i} \end{bmatrix} \end{aligned} \quad (19)$$

It's clear that the first term of this expression directly depends on the velocity screw. For the second term, one has to consider time derivatives of equation (5) applied to points p_i and p_j , and time derivatives of equation $a^2 + b^2 = 1$. Then, vector $(\dot{a}, \dot{b})^t$, which represents the orientation variations of the straight line δ , can be expressed as a function of the velocity screw:

$$\begin{bmatrix} u_{p_j} - u_{p_i} & v_{p_j} - v_{p_i} \\ a & b \end{bmatrix} \begin{bmatrix} \dot{a} \\ \dot{b} \end{bmatrix} = \begin{bmatrix} -[a \ b] \begin{bmatrix} \dot{u}_{p_j} - \dot{u}_{p_i} \\ \dot{v}_{p_j} - \dot{v}_{p_i} \end{bmatrix} \\ 0 \end{bmatrix} = \begin{bmatrix} -[a \ b] (\mathbf{J}_{S_j} - \mathbf{J}_{S_i}) \begin{bmatrix} \omega_x \\ \omega_y \\ v_z \end{bmatrix} \\ 0 \end{bmatrix} \quad (20)$$

This proves that including \dot{d}_0 in the features vector leads to a (3×3) interaction matrix $\mathbf{J} = \mathbf{M}^{-1} \mathbf{N}$ (by expanding equation (18) with a third row) such that:

$$\underbrace{\begin{bmatrix} 1 & 0 & -J_{13} \\ 0 & 1 & -J_{23} \\ c_u & c_v & 1 \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} \dot{u}_p \\ \dot{v}_p \\ \dot{d}_0 \end{bmatrix} = \underbrace{\begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ c_{\omega_x} & c_{\omega_y} & c_{v_z} \end{bmatrix}}_{\mathbf{N}} \begin{bmatrix} \omega_x \\ \omega_y \\ v_z \end{bmatrix} \quad (21)$$

When the instrument is not in the camera field of view, d_0 and d_2 cannot be measured. Therefore, for safety reasons, we propose to decompose the visual servoing in two control loops which partially decouple the control of the pointed direction given by $(u_p, v_p)^t$ and the control of the depth d_0 (these two stages are shown together in Figure 7) :

- **Stage 1** : Positioning of the laser spot projection p at the desired position (u_p^*, v_p^*) . This means to control (ω_x, ω_y) only. For safety reasons, v_z is kept null during this control (however d_0 and d_2 are still estimated while optical markers are in the camera field of view). Thus, from (16), it comes :

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

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

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

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

- **Stage 2** : Controlling the depth d_0 , assuming that the pointed direction remains constant. Since strong deformations may be induced by breathing, an entire decoupling is not suitable. The first stage control 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:

$$v_z^* = \dot{d} - k (d_0^* - d_0(\tau(\lambda_0), \alpha, \beta)) \quad (24)$$

k is a positive scalar and $d_0(\tau(\lambda_0), \alpha, \beta)$ is a function of the cross-ratio τ , α and β .

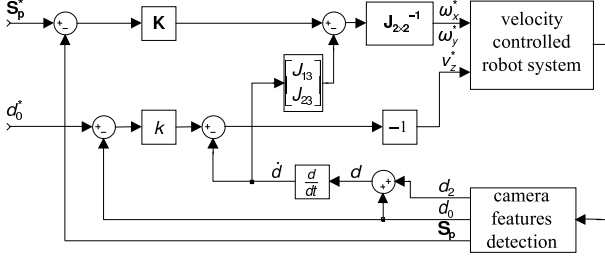


Figure 7: The complete servoing scheme.

For practical convenience, the upper (2×2) sub-matrix $\mathbf{J}_{2 \times 2}$ 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 sub-matrix is identified in an initial procedure. This identification consists in applying a constant velocity reference $(\omega_x^*, \omega_y^*)^t$ during a short time interval ΔT (see Fig. 8). Should the laser spot positioning performance degrades, it is always possible to start again the identification procedure (see also related work on on-line update of the interaction matrix [12]).

For the stability analysis during stage 1, we assume that the organ surface is approximately planar, then \dot{d} is linearly related to ω_x and ω_y . In this case, the interaction matrix is reduced to a 2×2 matrix :

$$\begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \mathbf{J}_\omega \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix} \quad (25)$$

and the stability of the visual feedback loop is guaranteed as long as $\mathbf{J}_\omega(\hat{\mathbf{J}}_\omega)^{-1}$ is positive definite [2].

4 Experiments

Experiments in real surgical conditions were conducted on living pigs in the operating room of IR-CAD. The experimental surgical robotic task consists in positioning the instrument at a 3-D desired position. We use a bi-processor PC computer (1.7 GHz) for image processing and for controlling, via a serial link, a Computer Motion surgical robot. A standard 50 fps PAL endoscopic camera hold by a second robot (at standstill) is linked to a PCI image capture board. For each image, the laser spots barycentre is detected in about 20 ms. Figure 8 shows an image-based visual tracking (a square) experiment with an initial identification procedure (around the image center) realized with the surgical robot-arm and an endo-trainer box. Figure 9 presents some simulations of the proposed image-based visual servoing and dotted curves represent image trajectories in the case \dot{d}_0 cannot be esti-

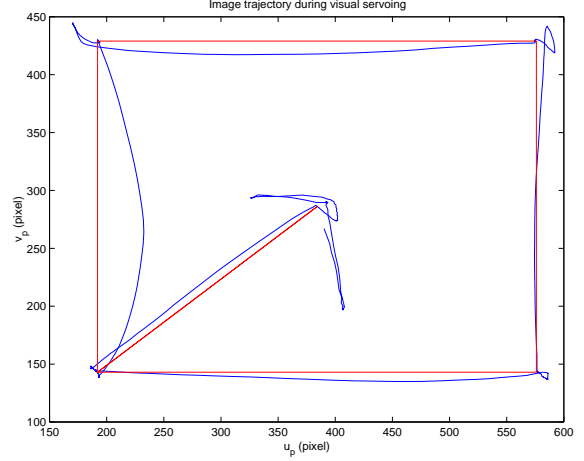


Figure 8: Initial on-line identification of matrix \mathbf{J}_ω (displacements around the image center). and image-based visual servoing along a square using this identification.

mated. During experiments, the orientation of the instrument is controlled by visual servoing in such a way that the laser spot image coordinates reach the desired position $p^* = (u_p^*, v_p^*)^t$ and the interaction matrix is provided by the initial identification procedure. Then, the instrument is brought down by a velocity reference v_z^* in open loop, until it is in the field of view. Finally, the depth d_0 is controlled (stage 2) to reach the 3-D position (see Figure 10). One can notice the significant perturbations due to the breathing during experiments.

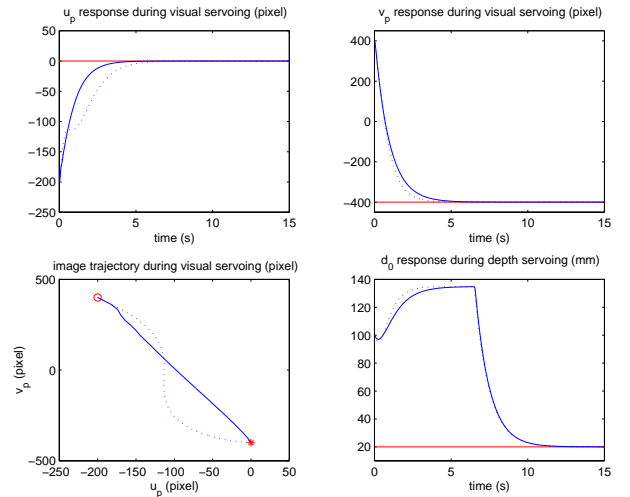


Figure 9: Image trajectories of laser spot (u_p, v_p) and d_0 responses time during the visual servoing (simulation). Dotted curves correspond to ignoring \dot{d}_0 in the first stage (second left member in eq. (22))

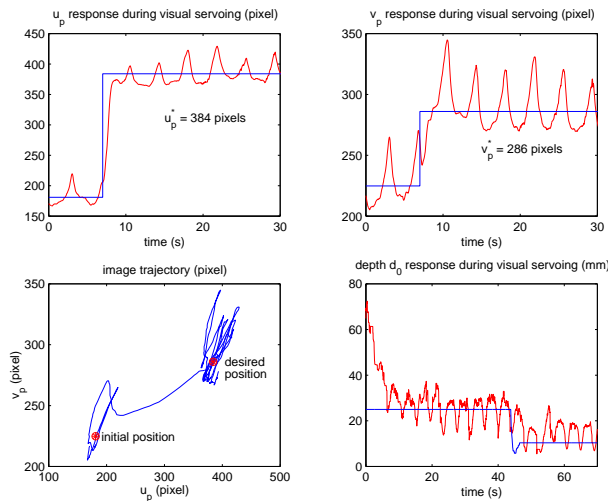


Figure 10: Image trajectories of laser spot (u_p, v_p) and d_0 responses time during the visual servoing (on living pigs). The effect of breathing is significant.

5 Conclusion

In this paper, we present a 3-D positioning system for laparoscopic surgical robots based on visual servoing. In our system, the scene perception is obtained by adding laser pointers and optical markers. To position the surgical instrument, we propose a visual servoing algorithm combining pixels coordinates of laser spots and the estimated distance between organ and instrument. We have shown that these informations are relevant for controlling the 3 dof of a symmetrical endoscopic surgical instrument. Successful experiments have been held with a surgical robot on living tissues in a surgical room. For safety reasons, during experiments, we have separated the instrument orientation control from depth control. The proposed technique does not require the knowledge of the initial respective position of the endoscope and the surgical instrument.

Acknowledgements

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. We also thank Prof. Marescaux, Leroy and Soler from IRCAD for their advices, as well as for the use of their facilities.

References

- [1] A. Casals, J. Amat, D. Prats and E. Laporte. "Vision Guided Robotic System for Laparoscopic Surgery". *Proc. of the IFAC International Congress on Advanced Robotics*. ICAR'95, pp. 33-36, Barcelona, Spain, 1995.
- [2] B. Espiau, F. Chaumette, P. Rives, "A New Approach to Visual Servoing in Robotics". *IEEE Trans. on Robotics and Automation*, 8(3), june 1992
- [3] M. Hayashibe, Y. Nakamura. "Laser-pointing endoscope system for intra-operative 3D geometric registration", *Proc. of the 2001 IEEE International Conference on Robotics and Automation*. Seoul, Korea, May 2001.
- [4] D.P. Huttenlocher, S. Ullman, "Recognizing solid objects by alignment with an image", *The International Journal of Computer Vision*, 5(2), p.195-212, 1990.
- [5] S. Hutchinson, G.D. Hager, P.I. Corke, "A tutorial on visual servo control", *IEEE Trans. on Robotics and Automation*, 12(5):p.651-670, October 1996.
- [6] A. Krupa, C. Doignon, J. Gangloff, M. de Mathelin, L. Soler, G. Morel. "Towards semi-autonomy in laparoscopic surgery through vision and force feedback control", *Proc. of the Seventh International Symposium on Experimental Robots (ISER)*, Hawaii, Dec. 2000.
- [7] A. Krupa, J. Gangloff, M. de Mathelin, C. Doignon, G. Morel, L. Soler, J. Leroy, J. Marescaux. "Autonomous retrieval and positioning of surgical instruments in robotized laparoscopic surgery using visual servoing and laser pointers". *To appear in the Proc. of the IEEE International Conference on Robotics and Automation*, Washington, D.C., May 2002.
- [8] E. Malis, F. Chaumette, S. Boudet, "2D 1/2 visual servoing", *IEEE Transactions on Robotics and Automation*, 15(2), pp.238-250, April 1999.
- [9] J. Marescaux, J. Leroy and M. Gagner, "Transatlantic robot-assisted telesurgery" *Nature*, 413,p.379-380, 2001.
- [10] P. Martinet, E. Cervera, "Combining Pixel and Depth Information in Image-Based Visual Servoing", *Proceedings of the 9th International Conference on Advanced Robotics*, pp. 445-450, Tokyo, Japan, October 1999.
- [11] S.J. Maybank. "The cross-ratio and the j-invariant", *Geometric invariance in computer vision.*, J. Mundy and A. Zisserman, pp. 107-109, MIT press, 1992.
- [12] H. Sutanto, R. Sharma, V. Varma Image based Autodocking without Calibration. *Proc. of the 1997 IEEE International Conference on Robotics and Automation*. Albuquerque, New Mexico, April 1997.
- [13] G.-Q. Wei, K. Arbter and G. Hirzinger. "Real-Time Visual Servoing for Laparoscopic Surgery". *IEEE Engineering in Medicine and Biology*, 16(1), pp. 40-45, 1997.
- [14] Y.F. Wang, D.R. Uecker, Y. Wang, "A new framework for vision-enabled and robotically assisted minimally invasive surgery", *Computerized Medical Imaging and Graphics*, vol.22,p.429-437,1998.