Optimizing Plane-to-Plane Positioning Tasks by Image-Based Visual Servoing and Structured Light

Jordi Pagès, Christophe Collewet, François Chaumette, *Member, IEEE*, and Joaquim Salvi

Abstract—This paper considers the problem of positioning an eye-in-hand system so that it becomes parallel to a planar object. Our approach to this problem is based on linking to the camera a structured light emitter designed to produce a suitable set of visual features. The aim of using structured light is not only for simplifying the image processing and allowing low-textured objects to be considered, but also for producing a control scheme with nice properties like decoupling, convergence, and adequate camera trajectory. This paper focuses on an image-based approach that achieves decoupling in all the workspace, and for which the global convergence is ensured in perfect conditions. The behavior of the image-based approach is shown to be partially equivalent to a 3-D visual servoing scheme, but with a better robustness with respect to image noise. Concerning the robustness of the approach against calibration errors, it is demonstrated both analytically and experimentally.

Index Terms—Convergence analysis, decoupled visual features, plane-to-plane task, structured light, visual servoing.

I. INTRODUCTION

VISUAL servoing is nowadays a widely used technique in robot control. The goal is to fulfill robotic tasks by using data provided by a vision sensor. Information extracted or calculated from the images are used in a closed-loop control law which leads to the execution of a task like positioning or target tracking [1].

This paper deals with the combination of visual servoing and structured light. For a long time, such a combination has been seen as a powerful option [2]. However, there are few visual servoing works exploiting structured light. The main contribution in this field is due to Motyl *et al.* [3], [4], who modeled the kinematics of simple visual features obtained when projecting laser planes onto planar objects and spheres. Andreff *et al.* [5] used a laser pointer in their approach based on 3-D lines in order to control the depth. Similarly, Krupa *et al.* [6] coupled a laser

pointer to a surgical instrument in order to control the pan-tilt and the depth of this instrument with respect to an organ, while both the organ and the laser are viewed from a static camera.

The main interest concerning the combination of visual servoing with structured light is that positioning tasks with respect to nontextured or nonstructured objects become feasible, and that the image processing is much simpler. Thus, tasks like docking, welding, or painting large surfaces lacking structure or texture can be faced. For example, Kondo and Tamaki [7] equipped an underwater robot with two laser pointers and a camera for avoiding obstacles and docking tasks. In [8], a glass-climbing robot uses two laser pointers and a camera for aligning the robot body with the glass surface. In both cases, the camera and lasers are accurately calibrated for fulfilling the positioning task by using 3-D data obtained by triangulation. In our opinion, these tasks could be robustly performed by using a visual servo control approach like the one presented in this paper.

Positioning tasks are still an issue in visual servoing. Indeed, during the last years, many works have focused on approaches for which the convergence of the system is ensured, even if the initial position is far from the desired one [5], [9], [10]. A suitable design strategy attempts to decouple visual features, so that each one is closely related to one degree of freedom (DOF). Of course, position-based visual servoing provides such a decoupling [11], [12], but the problem with such methods is with the stability of the pose-estimation algorithm with respect to image noise [13]. Hybrid techniques are based on controlling rotational DOFs in the Cartesian space while the translational ones are controlled from image data [9], [14]. However, they require partial pose estimation of the object at each iteration. On the other hand, some pure image-based techniques have succeeded to decouple rotational DOFs from translational ones near the desired state [15], [16]. Concerning the global stability analysis, for the most part, approaches for which analytical conditions have been found are hybrid approaches, as in [5] and [9], or the extended-2-D visual servoing [10], [17]. Usually, the global stability analysis of pure image-based techniques is too complex, even in absence of calibration errors.

Another important research topic in image-based visual servoing is to improve the camera trajectory in the Cartesian space. It is well known that even if an exponential decrease of the visual error is achieved, it does not necessarily imply a suitable camera trajectory [13]. This is mainly due to nonlinearities in the interaction matrix. Important efforts have been done in order to improve the mapping from the feature space to the camera velocities [16], [18].

J. Pagès is with Davantis Technologies, 08193 Bellaterra, Spain (e-mail: jordi.pages@davantis.com).

C. Collewet and F. Chaumette are with IRISA/INRIA Rennes, Campus Beaulieu, 35042 Rennes, France (e-mail: christophe.collewet@irisa.fr; francois.chaumette@irisa.fr).

J. Salvi is with the Computer Vision and Robotics Group, University of Girona, 17071 Girona, Spain (e-mail: qsalvi@eia.udg.es).

In this paper, we exploit the visual features provided by a structured light emitter based on laser pointers, in order to fulfill the classic plane-to-plane positioning task. Such a task consists of moving the camera linked to the robot end-effector to a pose where its image plane is parallel to a planar object at a given depth. For this task, three DOFs are constrained, while the remaining three are free. With this classical example, we demonstrate that the performance of the control loop can be optimized, thanks to an adequate modeling of the features obtained by using structured light. The main contribution of the paper is the formulation and analysis of an image-based approach with perfect decoupling properties in all of the workspace, thanks to the projected light pattern. The global convergence under ideal conditions is proven. Furthermore, its robustness against calibration errors is demonstrated analytically and experimentally. In addition to this, a linear map from the task function to the camera velocities is made, producing a suitable camera trajectory.

The paper is organized as follows. First, in Section II, the robotic task is formally defined. Secondly, the proposed sensor design and modeling are explained in Section III. Afterwards, a decoupled image-based approach for executing the task is proposed in Section IV. Then, Section V shows how to make the image-based approach robust against calibration errors concerning the relative camera-emitter pose. Our approach is compared with 3-D visual servoing through simulations in Section VI. Experiments validating the theoretical results are presented in Section VII. Finally, conclusions are presented in Section VIII.

II. DEFINITION AND REGULATION OF THE ROBOTIC TASK

The goal of the task is to control a robotic arm by using an eye-in-hand configuration, so that the camera linked to the robot end-effector becomes parallel to a planar object at a certain desired distance $Z^* > 0$. This kind of task achieves a plane-to-plane virtual link between the camera and the object. Let the camera kinematic screw be $\mathbf{v} = (\mathbf{V}, \mathbf{\Omega})$, where $\mathbf{V} = (V_x, V_y, V_z)$ are the translational velocities, and $\mathbf{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$ are the rotational ones. Then, once the camera is parallel to the object plane, camera motions according to V_x, V_y, and Ω_z do not change the plane-to-plane virtual link [19]. Therefore, as three DOFs of the camera remain free, a secondary task could be considered by using the redundancy framework [20].

In visual servoing, given a set of visual features stacked in a $\Re^k$ vector $\mathbf{s}$, its variation due to the camera velocity $\mathbf{v}$ is expressed by the well-known equation

$$\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v} \qquad (1)$$

where $\mathbf{L_s}$ is the interaction matrix. The rank of $\mathbf{L_s}$ determines the number m of DOFs that are controlled. If the rank is less than six, the null space of $\mathbf{L_s}$ determines the type of the virtual link that can be achieved by using $\mathbf{s}$.

Robotic tasks can be formally described by a function which must be regulated to zero [20]. The task function $\mathbf{e}$ is defined as a $\Re^m$ vector of the form

$$\mathbf{e} = \mathbf{C}(\mathbf{s} - \mathbf{s}^*) \qquad (2)$$

where $\mathbf{s}$ and $\mathbf{s}^*$ are, respectively, the visual features' values at the current and desired states, and $\mathbf{C}$ is an $m \times k$ combination matrix of full rank m. In the particular case of a plane-to-plane virtual link, we have $m = 3$, so that with $k = 3$ independent visual features, it is possible to set $\mathbf{C} = \mathbf{I}_3$. Thanks to the structured light and an adequate modeling, we will have this case where $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$.

The regulation of the task function $\mathbf{e}$ to $\mathbf{0}$ can be done by using a simple proportional control law of the form [1], [20]

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_s}}^+ \mathbf{e} \qquad (3)$$

with λ a positive gain, and $\widehat{\mathbf{L_s}}^+$ the pseudoinverse of a model or an approximation of $\mathbf{L_s}$. With this control law, the closed-loop equation of the system is

$$\dot{\mathbf{e}} = -\lambda \mathbf{L_s}\widehat{\mathbf{L_s}}^+ \mathbf{e} = -\lambda \mathbf{M}(\mathbf{e})\mathbf{e}. \qquad (4)$$

A key issue in visual servoing is to ensure that the desired state $\mathbf{e}^* = \mathbf{0}$ is reached for any initial pose.

III. A PROPOSAL FOR STRUCTURED LIGHT SENSOR

Nowadays, devices based on laser technology are the ones which have obtained the highest degree of compactness. Therefore, they are suitable to be used in an eye-in-hand configuration. By using different types of diffracting lenses, several patterns are available: from simple points and lines to more complex ones like grids, dot arrays, and circles. In visual servoing, laser planes have been suggested for positioning tasks [2]–[4].

In this paper, a structured light sensor is designed so that whenever the camera is parallel to the object, the projected pattern on the object is invariant to depth and the corresponding image is symmetric. The first property allows the visual features' variation to be independent of the depth whenever the camera is parallel to the object. The second property allows a partially decoupled control scheme to be obtained. Such a control scheme allows avoiding singularities, as well as making the stability analysis of the system easier [9], [15].

The structured light sensor that we propose consists of laser pointers, since very low-cost devices are easily available. Theoretically, three noncollinear points are enough to recover the equation of a planar object. Consequently, we have first thought of a sensor composed of three laser pointers. However, we have found that better decoupling properties are obtained when using four laser pointers. Concretely, the proposed structured light sensor is formed by four laser pointers attached to a cross structure, as shown in Fig. 1(a).

The coupling of the camera with the structured light emitter has been ideally modeled as follows.

- The frame {L} corresponding to the cross structure is perfectly aligned and has the same origin as the camera frame {C}, see Fig. 1.
- The four lasers have the same common direction. For modeling simplifications, it is set to $^L(0, 0, 1)$ which coincides with the camera optical axis. The projected pattern is thus invariant to the depth when the camera is parallel to the object.
- All the lasers are placed at the same distance L from the laser-cross intersection. According to this and the previous modeling constraints, the image of the pattern is
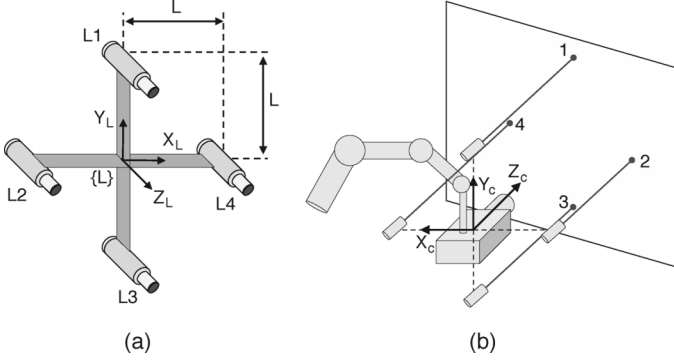
Fig. 1.   System architecture. (a) Proposed structured light sensor. (b) Ideal configuration of the robot manipulator, camera, and structured light sensor.
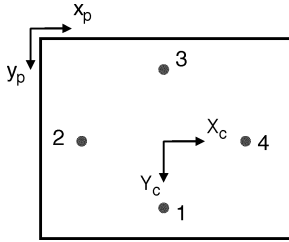

Fig. 2.   Camera image when it is parallel to the object at a given distance.

symmetric whenever the camera and the object are parallel (see Fig. 2).

Note that these assumptions have been only taken for modeling issues. In real conditions, it is very difficult to perfectly align the laser-cross with the camera frame because of the structure of the robot, or because the optical center position is not exactly known. That is why the study of the robustness against misalignments between the camera and the laser-cross, as done in Sections IV and V, is a key point when analyzing the approach presented in this paper.

A. Laser Pointer Modeling

This section presents the modeling of the interaction matrix of an image point corresponding to a projected laser point. Such an interaction matrix will be used later in the decoupled image-based approach presented in Section IV. We consider a planar object modeled according to the following equation:

$$\mathbf{N}^\top \mathbf{x} + D = 0 \tag{5}$$

with $\mathbf{N} = (A, B, C)$ the unitary normal vector to the plane, D its distance to the origin of the camera frame, and $\mathbf{X} = (X, Y, Z)$ a point belonging to the plane. The planar object can be also modeled by using the following minimal equation:

$$Z = \alpha X + \beta Y + \gamma \tag{6}$$

with

$$\alpha = -A/C, \quad \beta = -B/C, \quad \gamma = -D/C. \tag{7}$$

The laser pointer can be modeled according to a vectorial equation as follows:

$$\mathbf{X} = \mathbf{X}_0 + \mu \mathbf{U} \tag{8}$$

TABLE I
IDEAL MODEL PARAMETERS

Laser	X_0	Y_0	x	y	Z
1	0	L	0	L/Z_1	$\beta L + \gamma$
2	$-L$	0	$-L/Z_2$	0	$-\alpha L + \gamma$
3	0	$-L$	0	$-L/Z_3$	$-\beta L + \gamma$
4	L	0	L/Z_4	0	$\alpha L + \gamma$

where $\mathbf{U} = (U_x, U_y, U_z)$ is an unitary vector defining the laser direction, $\mathbf{X}_0 = (X_0, Y_0, 0)$ is the laser origin defined as the intersection of the laser direction with the plane $Z = 0$, and μ is the distance from $\mathbf{X}_0$ to $\mathbf{X}$.

The interaction matrix of a normalized image point $\mathbf{x} = (x, y) = (X/Z, Y/Z)$ corresponding to the 3-D intersection of a laser pointer and a planar object is [21], [22]

$$\mathbf{L_x} = \frac{1}{\Pi_0} \begin{bmatrix} \frac{-AX_0}{Z} & \frac{-BX_0}{Z} & \frac{-CX_0}{Z} & X_0\varepsilon_1 & X_0\varepsilon_2 & X_0\varepsilon_3 \\ \frac{-AY_0}{Z} & \frac{-BY_0}{Z} & \frac{-CY_0}{Z} & Y_0\varepsilon_1 & Y_0\varepsilon_2 & Y_0\varepsilon_3 \end{bmatrix} \tag{9}$$

with

$$\Pi_0 = \mathbf{N}^\top (\mathbf{x}_0 - (xZ, yZ, Z))$$
$$(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \mathbf{N} \times (x, y, 1). \tag{10}$$

A nonminimal representation for this matrix can be found in [3]. Note that the rank of $\mathbf{L_x}$ is 1. It is due to the particular epipolar geometry existing between the camera and the laser pointer [5], so that $\mathbf{x}$ always belongs a straight line.

B. Sensor Model

The sensor is composed of the camera and the structured light emitter. The camera observes the four projected points whose interaction matrices can be calculated from (9) and the following model parameters:
- reference point of each laser pointer $(X_0, Y_0, 0)$;
- normalized image point coordinates $\mathbf{x} = (x, y)$ of each projected point;
- depth Z of the projected points.

Given the equation of the straight line modeling a laser pointer (8) and the equation of the planar object (5), the 3-D coordinates $\mathbf{X}$ of the corresponding projected point are obtained. Afterwards, the computation of the normalized coordinates $\mathbf{x} = (X/Z, Y/Z)$ in the function of X_0, $\mathbf{N}$, and D is straightforward. The reference points are fixed according to the ideal configuration of the laser-cross shown in Fig. 1. Then, the model parameters obtained under this configuration and expressed in the camera frame are presented in Table I.

IV. DECOUPLED IMAGE-BASED APPROACH

A first image-based approach for our structured light sensor was proposed in [21]. However, decoupling was only reached near the desired state, and the selected visual features were too complex to prove the global convergence of the system, even under ideal conditions. This section presents a set of three decoupled visual features extracted from the image which partially decouples the controlled DOFs for any camera-object pose. Let us take a look at the interaction matrices of y_1^{-1}, y_3^{-1}, x_2^{-1}, and

x_4^{-1}. They are calculated taking into account $\dot{f}^{-1} = -\dot{f}/f^2$, the general interaction matrix in (9), the model parameters in Table I, and the relationships in (7). Then, noting $\mathbf{L_s} = [\mathbf{L_s^V} \mathbf{L_s^\Omega}]$, we obtain

$$
\begin{aligned}
\mathbf{L}_{y_1^{-1}}^V &= [& \alpha/L & & \beta/L & & -1/L &] \\
\mathbf{L}_{y_1^{-1}}^\Omega &= [& -\beta(\beta + \gamma/L) - 1 & & \alpha(\beta + \gamma/L) & & -\alpha &] \\
\mathbf{L}_{y_3^{-1}}^V &= [& -\alpha/L & & -\beta/L & & 1/L &] \\
\mathbf{L}_{y_3^{-1}}^\Omega &= [& -\beta(\beta - \gamma/L) - 1 & & \alpha(\beta - \gamma/L) & & -\alpha &] \\
\mathbf{L}_{x_2^{-1}}^V &= [& -\alpha/L & & -\beta/L & & 1/L &] \\
\mathbf{L}_{x_2^{-1}}^\Omega &= [& -\beta(\alpha - \gamma/L) & & \alpha(\alpha - \gamma/L) + 1 & & \beta &] \\
\mathbf{L}_{x_4^{-1}}^V &= [& \alpha/L & & \beta/L & & -1/L &] \\
\mathbf{L}_{x_4^{-1}}^\Omega &= [& -\beta(\alpha + \gamma/L). & & \alpha(\alpha + \gamma/L) + 1 & & \beta &].
\end{aligned}
$$

It is obvious that simple combinations of such features can lead to a decoupled system. To do that, we have chosen the following set of visual features:

$$
\mathbf{s} = \frac{1}{2} \left(y_1^{-1} - y_3^{-1}, y_1^{-1} + y_3^{-1}, x_2^{-1} + x_4^{-1} \right). \tag{11}
$$

Note that y_1, y_3, x_2, or x_4 are never equal to zero, except in the degenerate case when the camera is at infinity. The interaction matrix of $\mathbf{s}$ is

$$
\mathbf{L_s} = \begin{bmatrix} \alpha/L & \beta/L & -1/L & -\beta\gamma/L & \alpha\gamma/L & 0 \\ 0 & 0 & 0 & -(1+\beta^2) & \alpha\beta & -\alpha \\ 0 & 0 & 0 & -\alpha\beta & 1+\alpha^2 & \beta \end{bmatrix} \tag{12}
$$

which is always of rank 3 and never singular. Note that the last two features are independent to translational motions, which allows us to obtain a decoupled system. This happens because when the laser-cross is perfectly aligned with the camera frame, one can show from Table I that the visual features are related to the plane parameters as follows:

$$
\mathbf{s} = (\gamma/L, \beta, \alpha). \tag{13}
$$

Therefore, under ideal conditions, the image-based approach based on these features behaves as a position-based technique using α, β, and γ as features. Thus, a new way to implicitly estimate the object pose has been found from a nonlinear combination of the image-point coordinates. This relationship allows the interaction matrix in (12) to be expressed in terms of the visual features $\mathbf{s} = (s_1, s_2, s_3)$

$$
\mathbf{L_s}(\mathbf{s}) = \begin{bmatrix} s_3/L & s_2/L & -1/L & -s_1 s_2 & s_1 s_3 & 0 \\ 0 & 0 & 0 & -(1+s_2^2) & s_2 s_3 & -s_3 \\ 0 & 0 & 0 & s_2 s_3 & 1+s_3^2 & s_2 \end{bmatrix}. \tag{14}
$$

This allows us to decide which model of interaction matrix $\widehat{\mathbf{L_s}}$ can be used in the control law (3).

- *Nonconstant control law*: $\widehat{\mathbf{L_s}}$ is computed at each iteration. Note that in this case, all the elements of the interaction matrix can be obtained from the visual features.

- *Constant control law*: $\widehat{\mathbf{L_s}} = \mathbf{L_s^*}$ is the interaction matrix evaluated at the desired state, i.e., when $\alpha = 0, \beta = 0$, $\gamma = Z^*$ for $\mathbf{s^*} = (Z^*/L, 0, 0)$

$$
\mathbf{L_s^*} = \begin{bmatrix} 0 & 0 & -1/L & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{15}
$$

Note that this matrix is independent of the depth Z. This indicates that near the desired state, a linear map from camera velocities to visual features velocities is made, as in [18] and [21].

Note also that any desired orientation can be reached very easily, thanks to (13). However, we will focus in the remainder of the paper on the case of the plane-to-plane positioning task.

In the following sections, the analytic behavior of both control laws is analyzed for the ideal case, i.e., perfect alignment of the laser-cross and perfect camera calibration, and then for the case when the camera and the laser-cross are not aligned.

A. Analytic Behavior of the Nonconstant Control Law

The analytic behavior of the nonconstant control law under ideal conditions is straightforward. Note that in that case, a perfect estimation $\widehat{\mathbf{L_s}} = \mathbf{L_s}$ is available, so that the closed-loop equation of the system (4) becomes

$$
\dot{\mathbf{e}} = -\lambda \mathbf{e}. \tag{16}
$$

Therefore, a pure exponential decrease to $\mathbf{0}$ of the task function is achieved. Therefore, the system always converges to the desired position from any starting state.

B. Analytic Behavior of the Constant Control Law

Thanks to the linear and decoupled form of $\mathbf{L_s^*}$, it is also possible to demonstrate the global convergence of the system when using the constant control law. In addition to this, it is possible to obtain the equations describing the camera velocities and trajectory.

1) Global Convergence: When using $\mathbf{L_s^*}$ in the control law, the product of matrices $\mathbf{M} = \mathbf{L_s}\mathbf{L_s^{*+}}$ involved in (4) can be expressed as

$$
\mathbf{M}(\mathbf{e}) = \begin{bmatrix} 1 & e_2(e_1 + Z^*/L) & e_3(e_1 + Z^*/L) \\ 0 & 1 + e_2^2 & e_2 e_3 \\ 0 & e_2 e_3 & 1 + e_3^2 \end{bmatrix} \tag{17}
$$

using (14) and noting that $s_1 = e_1 + Z^*$, $s_2 = e_2$, and $s_3 = e_3$. The determinant of $\mathbf{M}(\mathbf{e})$ is $1 + e_2^2 + e_3^2$, which is always nonnull. Therefore, $\mathrm{Ker}(\mathbf{M}) = \emptyset$, yielding that the equilibrium point $\mathbf{e} = \mathbf{0}$ is unique. Furthermore, using (13), it is obvious to show that $\mathbf{e} = \mathbf{0}$ if and only if the object is parallel to the image plane ($\alpha = \beta = 0$) and at the desired distance Z^*.

Thanks to the decoupled form of the interaction matrix (14), we can solve the differential system in (4), which can be written using (17) as

$$
\begin{aligned}
\dot{e}_1(t) = &- \lambda \left[e_1(t) \left(1 + e_2(t)^2 + e_3(t)^2 \right) \right. \\
& \left. + \left(e_2(t)^2 + e_3(t)^2 \right) Z^*/L \right]
\end{aligned} \tag{18}
$$

$$\dot{e}_2(t) = -\lambda e_2(t)\left(1 + e_2(t)^2 + e_3(t)^2\right) \tag{19}$$

$$\dot{e}_3(t) = -\lambda e_3(t)\left(1 + e_2(t)^2 + e_3(t)^2\right). \tag{20}$$

The following solutions are obtained after some tedious developments [22]:

$$e_1(t) = \frac{e_1(0)}{a(t)} - \frac{bZ^* \arctan\left(u(t)\right)}{a(t)L} \tag{21}$$

$$e_2(t) = \frac{e_2(0)}{a(t)} \tag{22}$$

$$e_3(t) = \frac{e_3(0)}{a(t)} \tag{23}$$

with

$$a(t) = \sqrt{\left(e_2^2(0) + e_3^2(0)\right)\left(\exp^{2\lambda t} - 1\right) + \exp^{2\lambda t}} \tag{24}$$

$$b = \sqrt{e_2^2(0) + e_3^2(0)} \tag{25}$$

$$u(t) = \frac{b\left(a(t) - 1\right)}{b^2 + a(t)}. \tag{26}$$

Let us start by demonstrating the global convergence of the rotational subsystem defined by (19) and (20). The subsystem formed by $e_2(t)$ and $e_3(t)$ globally converges to the desired state if

$$\lim_{t \to \infty} e_2(t) = 0, \quad \lim_{t \to \infty} e_3(t) = 0. \tag{27}$$

Both functions clearly tend to zero when time approaches infinity, since $\lim_{t \to \infty} a(t) = \infty$. Moreover, it is easy to show that $e_2(t)$ and $e_3(t)$ are strictly monotonic functions by taking a look at their first derivative

$$\dot{e}_i(t) = -\frac{\lambda e_i(0)\exp^{2\lambda t}\left(1 + e_2^2(0) + e_3^2(0)\right)}{a(t)^3} \tag{28}$$

with $i = \{2, 3\}$. Note that the functions $e_2(t)$ and $e_3(t)$ are monotonic, since the sign of their derivatives never changes, and it only depends on the initial conditions. Furthermore, they are strictly monotonic, since their derivative only zeroes when $t \to \infty$ or when the function at $t = 0$ is already zero. Therefore, for any initial condition, $e_2(t)$ and $e_3(t)$ always tend towards zero strictly monotonically.

The global convergence of the translational subsystem depends on the behavior of $e_1(t)$. It is easy to show that $e_1(t)$ converges to zero for any initial state, since

$$\left.\begin{array}{c} \lim_{t \to \infty} u(t) = b \\ \lim_{t \to \infty} a(t) = \infty \end{array}\right\} \Rightarrow \lim_{t \to \infty} e_1(t) = 0. \tag{29}$$

In [22], it has been shown that $e_1(t)$ either is always monotonic, or it has a unique extremum before converging monotonically to zero. Furthermore, sufficient conditions are given so that it is possible to check, from the initial state of the system and the desired depth Z^*, whether either $e_1(t)$ will be monotonic during all the servoing or if it will have a peak.

2) Camera Trajectory: The control law based on $\mathbf{L}_\mathbf{s}^*$ maps the task function components $e_1(t), e_2(t)$, and $e_3(t)$ to the camera velocities as follows:

$$\mathbf{v} = -\lambda\mathbf{L}_\mathbf{s}^{*+}\mathbf{e} \tag{30}$$

so that using the pseudoinverse of (15), we obtain

$$\begin{cases} V_z(t) = \lambda L e_1(t) \\ \Omega_x(t) = \lambda e_2(t) \\ \Omega_y(t) = -\lambda e_3(t). \end{cases} \tag{31}$$

Note that $\Omega_x(t)$ and $\Omega_y(t)$ are strictly monotonic, while $V_z(t)$ is monotonic under the same conditions as $e_1(t)$.

Then, we can express the coordinates of a fixed point $\mathbf{X}$ in the camera frame at any instant of time t when the camera moves according to $\mathbf{v}(t)$, by using the well-known kinematic equation

$$\dot{\mathbf{X}}(t) = -\mathbf{V}(t) - \mathbf{\Omega}(t) \times \mathbf{X}(t). \tag{32}$$

Since the constant control law only generates velocities for V_z, Ω_x, and Ω_y, (32) can be rewritten as

$$\begin{cases} \dot{X}(t) = -\Omega_y(t)Z(t) \\ \dot{Y}(t) = +\Omega_x(t)Z(t) \\ \dot{Z}(t) = -V_z(t) + \Omega_y(t)X(t) - \Omega_x(t)Y(t) \end{cases} \tag{33}$$

where $V_z(t), \Omega_x(t)$, and $\Omega_y(t)$ are given by (31). If we choose as a fixed point the initial position of the camera $\mathbf{X}(0) = (0, 0, 0)$, the system of differential equations can be solved, obtaining

$$\begin{cases} X(t) = e_3(0)f(t) \\ Y(t) = e_2(0)f(t) \\ Z(t) = \dfrac{-\exp^{-\lambda t}}{1 + b^2}\left(-b^2 Z^*(\exp^{\lambda t} - 1)\right. \\ \qquad\qquad \left. + 4e_1(0)L\left(a(t) - 1\right)Z^*\left(a(t) - \exp^{\lambda t}\right)\right) \end{cases} \tag{34}$$

with

$$\begin{aligned} f(t) = \frac{\exp^{-\lambda t}}{(1 + b^2)b^3} & \left(\exp^{\lambda t}b^2 Z^*(1 + b^2)\arctan\left(u(t)\right)\right. \\ & \left. - e_1(0)Lb\left(\exp^{\lambda t}(1 + b^2) - b^2 - a(t)\right)\right. \\ & \left. + b^3 Z^*\left(a(t) - 1\right)\right). \end{aligned} \tag{35}$$

The expressions of $X(t)$ and $Y(t)$ have the same form. The only difference is that $X(t)$ depends on $e_3(0)$, while $Y(t)$ depends on $e_2(0)$. The study of the derivative of $X(t)$, and similarly for $Y(t)$, shows that both $X(t)$ and $Y(t)$ are monotonic functions [22].

Concerning $Z(t)$, its derivative can change sign, so its monotonicity is not ensured. Indeed, $Z(t)$ will be monotonic under the same conditions as $e_1(t)$ is monotonic. When $e_1(t)$ is not monotonic, a unique peak will appear also in $Z(t)$.

In summary, we can state that a complete analytic model describing the behavior of the ideal system when using the constant control law has been obtained.

C. Stability in Presence of Laser-Cross Misalignment

This section determines if the robotic task can be achieved when the camera and the laser-cross are not aligned, as assumed in the perfect model. As shown in [22], this type of calibration error is more important than calibration errors in the camera intrinsic parameters. Concretely, in [22], the stability against errors in the intrinsic parameters is demonstrated. The laser-cross
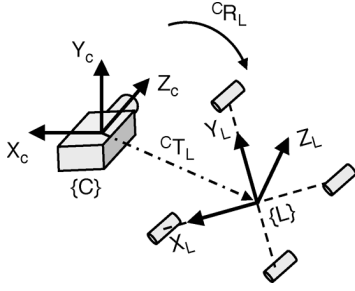
Fig. 3. Model of misalignment between the camera and the laser-cross.

TABLE II
MODEL PARAMETERS UNDER A TRANSLATIONAL MISALIGNMENT

Laser	X_0	Y_0	x	y	Z
1	t_x	$L+t_y$	t_x/Z_1	$(t_y+L)/Z_1$	$\alpha t_x + \beta(t_y+L) + \gamma$
2	$-L+t_x$	t_y	$(t_x-L)/Z_2$	t_y/Z_2	$\alpha(t_x-L) + \beta t_y + \gamma$
3	t_x	$-L+t_y$	t_x/Z_3	$(t_y-L)/Z_3$	$\alpha t_x + \beta(t_y-L) + \gamma$
4	$L+t_x$	t_y	$(t_x+L)/Z_4$	t_y/Z_4	$\alpha(t_x+L) + \beta t_y + \gamma$

misalignment is modeled according to a homogeneous frame transformation matrix of the form

$$^C\mathbf{M}_L = \begin{bmatrix} ^C\mathbf{R}_L & ^C\mathbf{T}_L \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \tag{36}$$

which transforms points from the structured light sensor frame to points in the camera frame (see Fig. 3). This transformation is used to calculate the actual lasers' orientation $\mathbf{U}$ and their reference points $\mathbf{X}_0$ in the camera frame. Afterwards, the normalized coordinates of the projected points, taking into account the actual pose of the laser-cross, can be obtained. The model parameters $\mathbf{X}_0$, $\mathbf{x}$, and Z of every laser pointer under different particular cases of $^C\mathbf{M}_L$ are given in [22].

Let $\tilde{\mathbf{e}}$ be the measured task function when the laser-cross is not aligned with the camera. Then, $\mathbf{L}_{\tilde{\mathbf{s}}}$ denotes the interaction matrix corresponding to the measured visual features $\tilde{\mathbf{s}}$ so that it takes into account the laser-cross misalignment. The closed-loop equation of the system can be noted as

$$\dot{\tilde{\mathbf{e}}} = -\lambda \mathbf{L}_{\tilde{\mathbf{s}}} \widehat{\mathbf{L}_{\mathbf{s}}}^{+} \tilde{\mathbf{e}}. \tag{37}$$

Analyzing the global convergence of this system is too difficult, because the interaction matrix $\mathbf{L}_{\tilde{\mathbf{s}}}$ is no longer partially decoupled. An alternative consists of studying the local asymptotic stability. This can be done by linearizing the equation around the desired state $\tilde{\mathbf{e}}^* = \mathbf{0}$ so that $\mathbf{N} = (0,0,1)$ and $D = -Z^*$. We thus consider

$$\dot{\tilde{\mathbf{e}}}^* = -\lambda \mathbf{L}_{\tilde{\mathbf{s}}}^* \mathbf{L}_{\mathbf{s}}^{*+} \tilde{\mathbf{e}}^* \tag{38}$$

where $\mathbf{L}_{\mathbf{s}}^{*+}$ is given by (15). The system is said to be locally asymptotically stable if and only if the eigenvalues of $\mathbf{M}^* = \mathbf{L}_{\tilde{\mathbf{s}}}^* \mathbf{L}_{\mathbf{s}}^{*+}$ have all positive real part. Note that this holds for both the constant and the nonconstant control law, since they are equivalent around the desired state.

Let us study the particular case when the laser-cross is displaced from the camera origin so that $^C\mathbf{R}_L = \mathbf{I}_3$ and $^C\mathbf{T}_L = (t_x, t_y, t_z)$. In this case, the model parameters are the ones shown in Table II.

The interaction matrix $\mathbf{L}_{\tilde{\mathbf{s}}}^*$ is then

$$\mathbf{L}_{\tilde{\mathbf{s}}}^* = \begin{bmatrix} 0 & 0 & \frac{L}{t_y^2-L^2} & 0 & -\frac{Lt_x}{t_y^2-L^2} & 0 \\ 0 & 0 & -\frac{t_y}{t_y^2-L^2} & -1 & \frac{t_x t_y}{t_y^2-L^2} & 0 \\ 0 & 0 & -\frac{t_x}{t_x^2-L^2} & -\frac{t_x t_y}{t_x^2-L^2} & 1 & 0 \end{bmatrix} \tag{39}$$

and the product of matrices $\mathbf{M}^* = \mathbf{L}_{\tilde{\mathbf{s}}}^* \mathbf{L}_{\mathbf{s}}^{*+}$ in the linearized closed-loop equation of the system is

$$\mathbf{M}^* = \begin{bmatrix} \frac{L^2}{L^2-t_y^2} & 0 & \frac{t_x L}{L^2-t_y^2} \\ -\frac{t_y L}{L^2-t_y^2} & 1 & -\frac{t_x t_y}{L^2-t_y^2} \\ -\frac{t_x L}{L^2-t_x^2} & -\frac{t_x t_y}{L^2-t_x^2} & 1 \end{bmatrix} \tag{40}$$

whose eigenvalues are

$$\begin{cases} \sigma_1 = \frac{L^2}{L^2-t_y^2} \\ \sigma_2 = \frac{L^2-t_x^2+\sqrt{t_x^2(t_x^2-L^2)}}{L^2-t_x^2} \\ \sigma_3 = \frac{L^2-t_x^2-\sqrt{t_x^2(t_x^2-L^2)}}{L^2-t_x^2} \end{cases} \tag{41}$$

so that, according to the demonstration given in [22], they are positive when

$$|t_x| < L, \quad |t_y| < L. \tag{42}$$

Note that the parameter L plays an important role in the stability. Concretely, it is necessary to maximize this parameter in order to enlarge the stability domain.

If the local asymptotic stability analysis is made for other types of misalignments modeled by (36), additional constraints are found [22]. Therefore, the image-based approach is quite sensitive to laser-cross misalignment, which can cause the system to diverge in presence of a large misalignment. In the following section, a way to enlarge the stability domain of the image-based approach is presented.

V. MAKING FEATURES ROBUST AGAINST LASER-CROSS MISALIGNMENT

In this section, we present a simple method to enlarge the robustness domain of the features against laser-cross misalignment. The goal is to define a corrected set $\mathbf{s}'$ which is analytically and experimentally robust against laser-cross misalignment. Fig. 4 shows the image-point distribution in the desired state under different types of misalignment (the four lasers have still the same relative orientation). As can be seen in Fig. 5(a), a complete misalignment of the laser-cross induces the polygon enclosing the four points in the desired image to appear misaligned and translated.

The idea consists of defining a planar transformation that minimizes the misalignment observed in the image. This image transformation is constrained as follows: in the absence of laser-cross misalignment, the corrected set of visual features $\mathbf{s}'$ must be equal to the uncorrected one $\mathbf{s}$. Therefore, in the ideal case, the results concerning the global convergence and camera trajectory concerning $\mathbf{s}$ will also hold for $\mathbf{s}'$.

First of all, we eliminate the misalignment exhibited by the polygon in Fig. 4(d), which is produced when the laser-cross
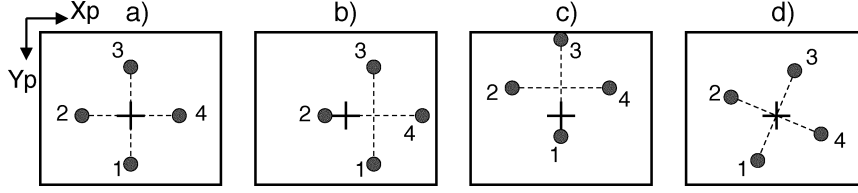
Fig. 4. Effects of laser-cross misalignment in the desired image. (a) Ideal image. (b) The laser-cross is horizontally displaced or rotated around Y_C. (c) The laser-cross is vertically displaced or rotated around X_C. (d) Laser-cross rotated around Z_C.
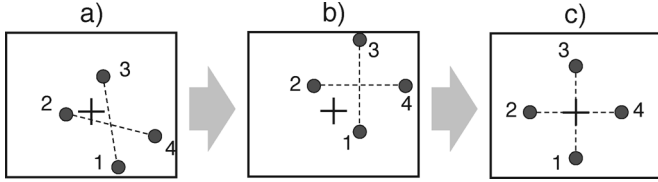

Fig. 5. Image points correction. (a) Desired image under a general misalignment of the laser-cross. (b) Image points after applying the transformation $\mathbf{T}^*$. (c) Image points after transformation $\mathbf{T}^*$ and translation $-\mathbf{x}_g$.

is rotated around the optical axis. Let us define the following unitary vectors:

$$
\mathbf{x}_{42}^* = \begin{bmatrix} x_{42}^* \\ y_{42}^* \end{bmatrix} = \frac{\mathbf{x}_4^* - \mathbf{x}_2^*}{\|\mathbf{x}_4^* - \mathbf{x}_2^*\|}, \quad \mathbf{x}_{13}^* = \begin{bmatrix} x_{13}^* \\ y_{13}^* \end{bmatrix} = \frac{\mathbf{x}_1^* - \mathbf{x}_3^*}{\|\mathbf{x}_1^* - \mathbf{x}_3^*\|}.
$$
(43)

Then, a simple 2-D transformation matrix of the form

$$
\mathbf{T}^* = \begin{bmatrix} \mathbf{x}_{24}^* & \mathbf{x}_{13}^* \end{bmatrix}^{-1} = \frac{1}{x_{42}^* y_{13}^* - x_{13}^* y_{42}^*} \begin{bmatrix} y_{13}^* & -x_{13}^* \\ -y_{42}^* & x_{42}^* \end{bmatrix}
$$
(44)

is defined so that $\mathbf{T}^*$ aligns the unitary vector corresponding to $\mathbf{x}_4 - \mathbf{x}_2$ with the image axis X_p and the unitary vector corresponding to $\mathbf{x}_1 - \mathbf{x}_3$ with the image axis Y_p. Let us note

$$
\mathbf{x}_i'' = \mathbf{T}^* \mathbf{x}_i.
$$
(45)

The result of applying $\mathbf{T}^*$ to the misaligned image points of Fig. 5(a) is shown in Fig. 5(b). Note that under ideal conditions, $\mathbf{T}^*$ is equal to the identity according to the normalized coordinates in Table I.

Then, it only remains to define a translation vector $\mathbf{x}_g$ which is able to center the polygon in the image [see Fig. 5(c)], and which is $\mathbf{0}$ for any object pose when there is no laser-cross misalignment. We propose to use the following vector:

$$
\mathbf{x}_g = \frac{1}{2} \begin{bmatrix} x_1'' + x_3'' \\ y_2'' + y_4'' \end{bmatrix}.
$$
(46)

Then, the corrected image points are obtained as follows:

$$
\mathbf{x}_i' = \mathbf{x}_i'' - \mathbf{x}_g = \mathbf{T}^* \mathbf{x}_i - \mathbf{x}_g
$$
(47)

so that the corrected set of visual features $\mathbf{s}'$ is, therefore

$$
\mathbf{s}' = \frac{1}{2} \left(y_1'^{-1} - y_3'^{-1}, y_1'^{-1} + y_3'^{-1}, x_2'^{-1} + x_4'^{-1} \right).
$$
(48)

The global convergence of the ideal model is also ensured when using $\mathbf{s}'$. In the following sections, the improvement in terms of robustness of $\mathbf{s}'$, with respect to laser-cross misalignment, is proved analytically. Furthermore, the corrected visual features avoid a potential problem of the uncorrected set $\mathbf{s}$. Since the definition of $\mathbf{s}$ involves the computation of $1/y_1, 1/x_2, 1/y_3$, and $1/x_4$, a division by zero may be reached due to the laser-cross misalignment. Note that this problem no longer appears in $\mathbf{s}'$, since the corrected image points are symmetrically distributed around the image center.

A. Stability in Presence of Laser-Cross Misalignment

In this section, we analyze the stability of the system based on $\mathbf{s}'$ in the presence of laser-cross misalignment. Even if the system has been stated to be very robust against this type of calibration error through simulations [22], the global convergence has not been proven, since the differential system is too complex. That is why the local asymptotic stability against different types of misalignment is presented as an analytic way to show the improvement in robustness.

1) Misalignment Consisting of a Translation: Let us first analyze the case when the laser-cross is aligned with the camera frame, but it is displaced from the camera origin, according to $^C\mathbf{T}_L = (t_x, t_y, t_z)$. The current and desired normalized coordinates of the laser points are obtained from Table II. The 2-D transformation components defined in (44) and in (46) are

$$
\mathbf{T}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad x_g = \frac{t_x}{Z^*}, \quad y_g = \frac{t_y}{Z^*}.
$$
(49)

Note that the laser-cross displacement is captured by $\mathbf{x}_g$. Similar to Section IV-C, the product of matrices in the linearized closed-loop equation of the system $\mathbf{M}^* = \mathbf{L}_{\mathbf{s}'}^* \mathbf{L}_{\mathbf{s}'}^{*+}$ must be calculated. After some developments, we obtain

$$
\mathbf{M}^* = \begin{bmatrix} 1 & 2t_y/L & t_x/L \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$
(50)

whose eigenvalues are the elements on the diagonal which are all equal to one. Therefore, the local asymptotic stability of the system in front of a displacement of the laser-cross is always ensured when using $\mathbf{s}'$. Note that the restriction imposed by the parameter L over the local asymptotic stability using $\mathbf{s}$ has been removed.

2) Misalignment Consisting of Individual Rotations: We now present the local asymptotic stability analysis when the laser-cross is centered in the camera origin, i.e., $^C\mathbf{T}_L =$

$(0, 0, 0)$, but it is rotated around one camera axis. Let us first consider a rotation ψ around the X axis. The model parameters under this type of misalignment can be found in [22]. In this case, the image transformation becomes

$$\mathbf{T}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad x_g = 0, \quad y_g = \frac{\sin \psi}{\cos \psi} \quad (51)$$

so that the product of matrices in the closed-loop equation of the system is

$$\mathbf{M}^* = \begin{bmatrix} \cos \psi & \frac{Z^* \cos \psi \sin \psi}{L \cos \psi} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (52)$$

Note that the eigenvalues are again the elements of the diagonal, and they are all positive if $\psi \in (-\pi/2, \pi/2)$, which imposes that the object must be in front of the camera.

In the case when the laser-cross is rotated around the Y camera axis by an angle θ, an equivalent result is obtained.

Finally, when the laser-cross is rotated by an angle ϕ around the optical axis of the camera, the image transformation is defined as [22]

$$\mathbf{T}^* = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix}, \quad x_g = 0, \quad y_g = 0 \quad (53)$$

obtaining the following matrix in the linearized closed-loop equation of the system:

$$\mathbf{M}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (54)$$

whose eigenvalues are

$$\begin{cases} \sigma_1 = 1 \\ \sigma_2 = \cos \phi + \sqrt{\cos^2 \phi - 1} \\ \sigma_3 = \cos \phi - \sqrt{\cos^2 \phi - 1}. \end{cases} \quad (55)$$

Note that the real part of σ_2 and σ_3 is $\cos \phi$, so that, in order to ensure their positiveness, it is only necessary that $\phi \in (-\pi/2, \pi/2)$.

In conclusion, the local asymptotic stability domain of the corrected features $\mathbf{s}'$ is practically unrestricted. Therefore, the improvement with respect to the uncorrected version is analytically proven. Furthermore, the local asymptotic stability of $\mathbf{s}'$ in the presence of camera-intrinsic errors is also ensured [22].

VI. COMPARING 2-D AND 3-D VISUAL SERVOING

As shown in Section IV, under ideal conditions, the decoupled visual features proposed in this paper are equivalent to the object plane parameters α, β, and γ in (6). A comparison of our method with a position-based approach can, thus, be performed. In this case, the plane parameters are reconstructed and directly used in the control loop. Let us denote the signal vector composed of the reconstructed 3-D parameters as

$$\mathbf{s}_r = (\gamma_r, \ \beta_r, \ \alpha_r) \quad (56)$$

so that the desired state vector is $\mathbf{s}_r^* = (Z^*, 0, 0)$ and the interaction matrix corresponding to $\mathbf{s}_r$ is [see (12)]

$$\mathbf{L}_{\mathbf{s}_r} = \begin{bmatrix} \alpha_r & \beta_r & -1 & -\gamma_r \beta_r & \gamma_r \alpha_r & 0 \\ 0 & 0 & 0 & -\left(1 + \beta_r^2\right) & \beta_r \alpha_r & -\alpha_r \\ 0 & 0 & 0 & -\beta_r \alpha_r & 1 + \alpha_r^2 & \beta_r \end{bmatrix}. \quad (57)$$

A simple way to reconstruct the 3-D parameters of the plane consists of using triangulation. However, this technique needs to accurately know the relative position of the lasers and the camera. Otherwise, bad reconstruction results are obtained. A better solution consists of defining the plane reconstruction problem in a nonlinear least-squares sense, taking advantage of the data provided by the current and desired images. As shown in [22], the following system of nonlinear equations can be built for each laser point:

$$\begin{cases} Z^* \left(x_i^* - U_{xz}\right)\left(\alpha_r x_i + \beta_r y_i - 1\right) + \gamma_r (x_i - U_{xz}) = 0 \\ Z^* \left(y_i^* - U_{yz}\right)\left(\alpha_r x_i + \beta_r y_i - 1\right) + \gamma_r (y_i - U_{yz}) = 0 \end{cases} \quad (58)$$

where $U_{xz} = U_x/U_z$ and $U_{yz} = U_y/U_z$ express the lasers' direction. Note that there are eight equations for five unknowns

$$\xi = (\alpha_r, \beta_r, \gamma_r, \ U_{xz}, \ U_{yz}). \quad (59)$$

Therefore, a possible misalignment between the laser-cross and the camera is taken into account in the equations. The system can be solved with an iterative algorithm like Gauss–Newton or Levenberg–Marquardt.

The performance of this position-based approach and the decoupled image-based approach have been compared through simulations. The following control law for the position-based approach has been used:

$$\mathbf{v} = -\lambda \mathbf{L}_{\mathbf{s}_r}^+ \mathbf{e}_r \quad (60)$$

$\mathbf{L}_{\mathbf{s}_r}$ being the interaction matrix in (57) estimated at each iteration using the reconstructed plane parameters $\mathbf{s}_r$, and $\mathbf{e}_r = (\gamma_r - Z^*, \beta_r, \alpha_r)$.

Concerning the decoupled image-based approach, the nonconstant control law has been used

$$\mathbf{v} = -\lambda \mathbf{L}_{\mathbf{s}}^+ \mathbf{e} \quad (61)$$

$\mathbf{L}_{\mathbf{s}}$ being the interaction matrix in (12) estimated at each iteration from the corrected visual features $\mathbf{s}'$ in (48), and $\mathbf{e} = \mathbf{s}' - \mathbf{s}'^*$.

A simulation example including large laser-cross misalignment and image noise is now presented. In the simulation, the laser-cross is displaced from the camera origin according to the translation vector ${}^C\mathbf{T}_L = (-4, -10, -9)$ cm. Furthermore, the laser-cross frame is rotated $12°$ around Z, $9°$ around Y, and $-15°$ around X. At each iteration of the simulation, Gaussian noise with a standard deviation of two pixels has been added to the image. The goal is to position the camera parallel to the object at a distance $Z^* = 60$ cm. The initial position of the camera is defined so that it is at a distance of 105 cm to the object, and their relative orientation is determined by $\alpha_x = -20°$
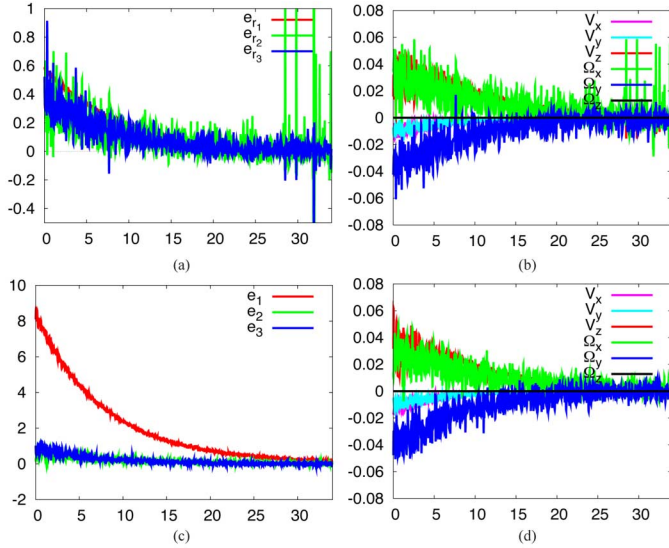
Fig. 6. Simulation results comparing the 3-D and 2-D approaches. First row: 3-D visual servoing. Second row: 2-D visual servoing. (a), (c) Task function versus time (in s). (b), (d) Camera kinematic screw versus time (m/s and rad/s).

and $\alpha_y = 20°$, α_x being the angle between Z_C and the projection of $\mathbf{N}$ to the plane $Y_C = 0$, and α_y the angle between Z_C and the projection of $\mathbf{N}$ to the plane $X_C = 0$.

The 3-D approach has been implemented by using the iterative Gauss–Newton algorithm in order to solve the nonlinear system of equations in (58) at each iteration. In the initial state of the simulation, the guess solution given to the iterative algorithm is $\xi = (0, 0, Z^*, 0, 0)$ which corresponds to the desired state and a perfect alignment of the laser-cross. During the remainder of the simulation, the solution provided by the previous state is given as an initial guess. Such a strategy has shown the best performance. The results obtained by the 3-D approach are plotted in the first row of Fig. 6. As can be seen, the task is realized, and therefore, the reconstruction results obtained by the proposed algorithm are satisfactory. However, we can observe that near the desired state, some eventual peaks appear in both the task function and the camera kinematic screw. Such peaks correspond to states where the nonlinear least-squares algorithm has not succeeded in converging to the right reconstruction. That confirms that this approach is sensitive to image noise, and that it is subject to falling into a local minimum. Let us note that if a simple triangulation algorithm is used to estimate the 3-D parameters α_r, β_r, and γ_r, the results are worse, since a very unstable behavior is obtained (not shown here). Furthermore, another drawback of this technique is that, like in all position-based approaches, the stability cannot be analytically proven, as it depends on the convergence of the reconstruction algorithm [11].

The second row of Fig. 6 shows the simulation results when using the image-based approach in the same conditions. As can be seen, the camera kinematic screw is similar to the one from the 3-D approach, but without any peak. Furthermore, as the visual features are directly computed from the image and not using an iterative algorithm, it is more suitable for a real-time robot platform.

VII. EXPERIMENTAL RESULTS

In order to validate the theoretical results presented in this paper, experiments with an eye-in-hand robot have been done. The experimental setup consists of a 6-DOF robot arm with a camera with focal length 8.5 mm fixed to its end-effector. The images are digitized at 782×582 pixels, and the pixel dimensions are about $8.3~\mu\text{m} \times 8.3~\mu\text{m}$. The normalized image coordinates (x, y) have been calculated from the pixel coordinates (u, v) by using the camera intrinsic parameters, as follows:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad (62)$$

where f is the focal length (in meters), (u_0, v_0) is the principal point (in pixels), and (k_u, k_v) are the conversion factors from meters to pixels. A rough approximation of the camera-intrinsic parameters has been used. In [22], the local asymptotic stability analysis shows that a coarse camera calibration does not affect the approach. This issue is here confirmed under real conditions.

The laser-cross has been built so that $L = 15$ cm. Such a parameter has been chosen taking into account the robot structure, so that the laser-cross can be approximately positioned according to the ideal model, i.e., aligned with the camera frame. Indeed, this parameter and the camera-intrinsic parameters constrain the minimum distance Z^* of the virtual link. In this configuration, the minimum Z^* to have all the laser points in the image bounds is about 0.5 m. If the real task requires smaller positioning distances, either L should be smaller, or another lens with smaller focal distance should be used.

The visual features corresponding to the desired state are calculated through the following learning stage. The camera is positioned with respect to a planar target containing structured landmarks by means of classic 2-D visual servoing. In our experiments, four points forming a square of known sides were used. Once the desired position is reached, the lasers are turned on, obtaining the desired image point distribution from which the desired visual features $\mathbf{s}^*$ are calculated. This target plane is only used once for obtaining the desired point distribution. Afterwards, the experiments are made with another planar object containing no visual marks.

A. Laser-Cross Approximately Aligned With the Camera

In the first experiment, the laser-cross has been approximately aligned with the camera. The direction of the four lasers is not exactly equal. Therefore, the robustness of the approach with respect to this kind of modeling error is also tested. The desired position is defined by $Z^* = 60$ cm, and the initial position is as in the simulation defined by $D(0) = -105$ cm, $\alpha_x = -20°$, and $\alpha_y = 20°$.

The image corresponding to the initial and the desired states is shown in Fig. 7(a). As can be seen in Fig. 7(b), the laser points do not exactly lie on the image axis, and their traces from the initial to the desired position, which shows us the epipolar line of each laser, are not perfectly parallel to the axis. Furthermore, it is not possible to ensure that all the four lasers have the same exact orientation, which causes the epipolar lines to not intersect at a unique point.
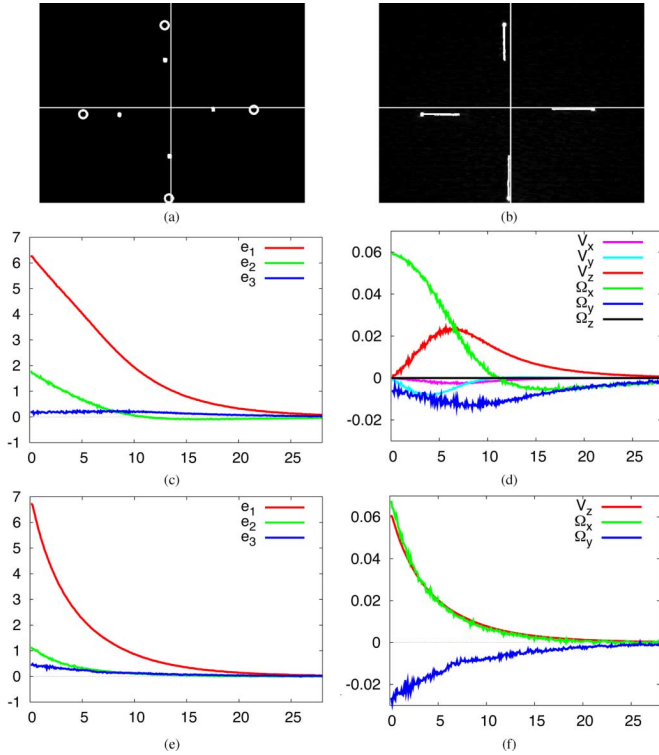
Fig. 7. Experiment with a coarse alignment. (a) Initial image (solid dots) including image axis and the desired position of each laser point (circles). (b) Final image with the trace of each laser point from its initial position to its final position. Nonconstant control law: (c) Task function versus time (in s). (d) Camera kinematic screw (in m/s and rad/s). Constant control law: (e) Task function versus time (in s). (f) Camera kinematic screw (in m/s and rad/s).

The results obtained in the experiment using a nonconstant matrix in the control law are plotted in Fig. 7(c) and (d). Note that even if the task function converges nicely to zero, the camera velocities are strongly nonmonotonic. The results of the experiment when using the constant control law are shown in Fig. 7(e) and (f). As can be seen, the monotonicity of both the task function and the camera velocities is preserved, as in the ideal case predicted by the analytic model.

B. Large Camera and Laser-Cross Misalignment

The same experiment has been repeated by introducing a large misalignment between the laser-cross and the camera. Concretely, the laser-cross has been displaced from the camera origin about 6 cm in the sense of the $-X$ axis of the camera frame. Furthermore, it has been rotated about $7°$ around the Z axis, and smaller rotations have been done around the X and Y axes. The initial image and the desired laser point distribution obtained during the learning stage are shown in Fig. 8(a). The large misalignment between the camera and the laser-cross is clearly observed in laser traces shown in Fig. 8(b). The image correction presented in Section V produces the laser points traces shown in Fig. 8(c). As can be seen, the corrected image minimizes with success the misalignment of the laser traces with respect to the image axis.

Fig. 8(d) and (e) presents the results when using $\mathbf{s}'$ and the constant control law. As can be seen, even with such a large
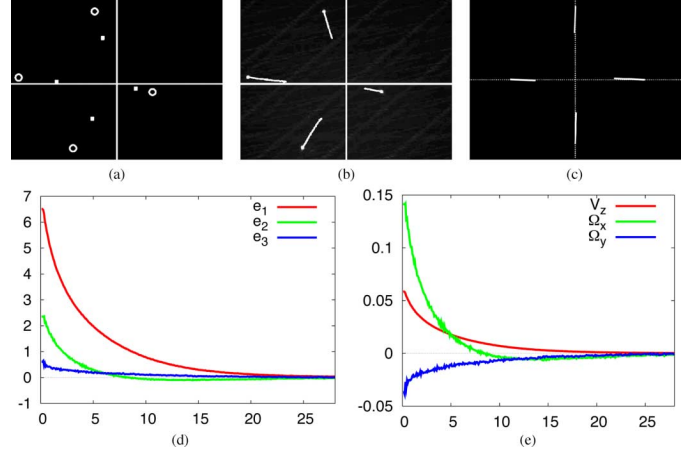

Fig. 8. Experiment with a large misalignment. (a) Initial image (solid dots) including image axis and the desired position of each laser point (circles). (b) Final image with the trace of each laser point from its initial position to its final position. (c) Corrected image from the initial to the desired position. Constant control law: (d) Task function versus time (in s). (e) Camera kinematic screw (m/s and rad/s).

misalignment, the approach still obtains almost a monotonic decrease in the task function, as well as in the camera velocities. Therefore, the robustness of this approach against laser-cross misalignment expected from the analytic results is confirmed. The nonconstant control law has also shown good performance. However, for this given example, the servo control has been stopped, since the robot has reached a joint limit. This is due to the nonmonotonic nature of the camera velocities generated by the control law, as observed in the previous experiment in Fig. 7(d).

VIII. CONCLUSION

This paper has presented a solution to the classic plane-to-plane positioning task when visual servoing and structured light are combined. The projection of structured light not only simplifies the image processing, but also enables dealing with low-textured objects. A structured light sensor composed of four laser pointers for eye-in-hand systems has been proposed.

An image-based approach decoupling the rotational part from the translational one in the whole workspace has been proposed. The new approach is similar to a position-based approach, but using only image data. Therefore, the new approach does not require reconstructing the object plane by using triangulation or a nonlinear minimization algorithm. The interaction matrix of the image-based approach can be also expressed in terms of the 2-D visual features. This allows two control laws to be used. The first is a nonconstant control law, where the online estimation of the matrix is used. The global convergence of the system in the absence of calibration errors is ensured. A constant control law based on the interaction matrix evaluated at the desired state can also be used. In this case, not only has the global convergence been proven under ideal conditions, but also, the equations describing the camera velocities and the camera trajectory have been obtained. Furthermore, the robustness, with respect to misalignment between the camera and the lasers, has been improved by defining an image transformation.

Experimental results have been carried out in order to prove the validity of our approach. Experiments with small and large calibration errors have been considered. The results show that both control laws are efficient and robust, with respect to modeling errors.

The high level of decoupling achieved in this work is due to the fact that the points are projected with structured light. Such decoupling has still not been reached with visual features extracted from the object itself. Future work will concern the control of additional DOFs by taking into account marks. In addition, we want to extend this work to nonplanar objects.

REFERENCES

[1] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[2] J. Albus, E. Kent, M. Nashman, P. Manabach, and L. Palombo, "Six-dimensional vision system," in *Proc. SPIE*, 1982, vol. 336, pp. 142–153.

[3] J. P. Urban, G. Motyl, and J. Gallice, "Realtime visual servoing using controlled illumination," *Int. J. Robot. Res.*, vol. 13, no. 1, pp. 93–110, 1994.

[4] D. Khadraoui, G. Motyl, P. Martinet, J. Gallice, and F. Chaumette, "Visual servoing in robotics scheme using a camera/laser-stripe sensor," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 743–750, Oct. 1996.

[5] N. Andreff, B. Espiau, and R. Horaud, "Visual servoing from lines," *Int. J. Robot. Res.*, vol. 21, no. 8, pp. 679–700, Aug. 2002.

[6] A. Krupa, J. Gangloff, C. Doignon, M. Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux, "Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing," *IEEE Trans. Robot. Autom.*, vol. 19, no. 5, pp. 842–853, Oct. 2003.

[7] H. Kondo and U. Tamaki, "Navigation of an AUV for investigation of underwater structures," *Control Eng. Practice*, vol. 12, no. 12, pp. 1551–1559, Dec. 2004.

[8] D. Sun, J. Zhu, C. Lai, and S. K. Tso, "A visual sensing application to a climbing cleaning robot on the glass surface," *Mechatron.*, vol. 14, no. 10, pp. 1089–1104, Dec. 2004.

[9] E. Malis and F. Chaumette, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," *IEEE Trans. Robot. Autom.*, vol. 18, no. 2, pp. 176–186, Apr. 2002.

[10] F. Schramm, G. Morel, A. Micaelli, and A. Lottin, "Extended-2D visual servoing," in *Proc. IEEE Int. Conf. Robot. Autom.*, New Orleans, LA, Apr. 26–May 1 2004, pp. 267–273.

[11] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, "Position based visual servoing: Keeping the object in the field of vision," in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington, DC, 2002, pp. 1624–1629.

[12] W. J. Wilson, C. C. W. Hulls, and G. S. Bell, "Relative end-effector control using Cartesian position based visual servoing," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 684–696, Oct. 1996.

[13] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Conference of Vision and Control*, ser. LNCIS. New York: Springer, 1998, pp. 66–78.

[14] K. Deguchi, "Direct interpretation of dynamic images and camera motion for visual servoing without image feature correspondence," *J. Robot. Mechatron.*, vol. 9, no. 2, pp. 104–110, 1997.

[15] P. I. Corke and S. A. Hutchinson, "A new partitioned approach to image-based visual servo control," *IEEE Trans. Robot. Autom.*, vol. 17, no. 4, pp. 507–515, Aug. 2001.

[16] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.

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

[18] R. Mahony, P. Corke, and F. Chaumette, "Choice of image features for depth-axis control in image-based visual servo control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Lausanne, Switzerland, Oct. 2002, vol. 1, pp. 390–395.

[19] F. Chaumette, P. Rives, and B. Espiau, "Classification and realization of the different vision-based tasks," in *Visual Servoing*, ser. Robotics and Automated Systems, K. Hashimoto, Ed. , Singapore: World Scientific, 1993, vol. 7, pp. 199–228.

[20] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 6, pp. 313–326, Dec. 1992.

[21] J. Pagès, C. Collewet, F. Chaumette, and J. Salvi, "Plane-to-plane positioning from image-based visual servoing and structured light," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sendai, Japan, Sep. 2004, vol. 1, pp. 1004–1009.

[22] ——, Visual Servoing by Means of Structured Light for Plane-to-Plane Positioning INRIA, Tech. Rep. 5579, May 2005.

Jordi Pagès was graduated in computer science from the Polytechnical School, University of Girona, Catalunya, Spain, in 2001, and received the Ph.D. degree in information technologies from the University of Girona and the University of Rennes, Rennes, France, in 2005.

He is currently with Davantis Technologies, Bellaterra, Spain, working on intelligent video surveillance based on computer vision. His research interests include visual control, 3-D reconstruction, and tracking.

Christophe Collewet was graduated in automatic control from the École Nationale Supérieure d'Ingénieurs de Caen, Caen, France, in 1986, and received the Ph.D. degree in signal processing from the University of Rennes, Rennes, France, in 1999.

From May 1988 to October 2005, he was with Cemagref, the French Institute of Agricultural and Environmental Engineering Research as "Chargé de Recherche." He is currently on secondment at IRISA/INRIA, Rennes, France, in the Lagadic group. His research interests include robotics, image processing, especially visual features tracking, and vision-based control.

François Chaumette (M'98) was born in Nantes, France, in 1963 and graduated from École Nationale Supérieure de Mécanique, Nantes, in 1987. He received the Ph.D degree and "Habilitation à Diriger des Recherches" in computer science from the University of Rennes, Rennes, France, in 1990 and 1998, respectively.

Since 1990, he has been with IRISA/INRIA, Rennes, where he is now "Directeur de Recherches" and head of the Lagadic Group. His research interests include robotics and computer vision, especially visual servoing and active perception.

Dr. Chaumette received the AFCET/CNRS Prize for the best French thesis in automatic control in 1991. He also received (with Ezio Malis) the 2002 King-Sun Fu Memorial Best IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION Paper Award. He was an Associate Editor for the IEEE TRANSACTIONS ON ROBOTICS from 2001 to 2005.

Joaquim Salvi was graduated in computer science from the Polytechnical University of Catalunya, Catalunya, Spain, in 1993, and received the D.E.A. degree in computer science in July 1996, and the Ph.D. degree in industrial engineering in 1998, from the Computer Vision and Robotics Group, University of Girona, Girona, Spain.

He is currently an Associate Professor with the Electronics, Computer Engineering and Automation Department, University of Girona. His current interests are in the field of computer vision and mobile robotics, focused on structured light, stereovision, and camera calibration. He is the Leader of the 3-D Perception Lab.

Dr. Salvi received the Best Thesis Award in Engineering for his Ph.D. dissertation.