Plane-to-plane positioning from image-based visual servoing and structured light

J. Pagès, Christophe Collewet, François Chaumette, J. Salvi

HAL Id: inria-00352034

https://inria.hal.science/inria-00352034

Submitted on 12 Jan 2009

Plane-to-plane positioning from image-based visual servoing and structured light

Jordi Pagès [†]   Christophe Collewet [‡]   François Chaumette [*]   Joaquim Salvi[†]

[†] Institut d'Informàtica i Aplicacions
University of Girona
Girona, Spain

[‡] Cemagref
17 Avenue de Cucillé
Rennes, France

[*] IRISA / INRIA Rennes
Campus Universitaire de Beaulieu
Rennes, France

Abstract— In this paper we face the problem of positioning a camera attached to the end-effector of a robotic manipulator so that it gets parallel to a planar object. Such problem has been treated for a long time in visual servoing. Our approach is based on linking to the camera several laser pointers so that its configuration is aimed to produce a suitable set of visual features. The aim of using structured light is not only for easing the image processing and to allow low-textured objects to be treated, but also for producing a control scheme with nice properties like decoupling, stability, well conditioning and good camera trajectory.

I. Introduction

2D or image-based visual servoing [4] is a robot control technique based on visual features extracted from the image of a camera. The goal consists of moving the robot to a desired position where the visual features contained in a k-dimensional vector $\mathbf{s}$ become $\mathbf{s}^*$. Therefore, $\mathbf{s}^*$ describes the features when the desired position is reached. The visual features velocity $\dot{\mathbf{s}}$ is related to the relative camera-object motion according to the following linear system

$$\dot{\mathbf{s}} = \mathbf{L_S} \cdot \mathbf{v} \tag{1}$$

where $\mathbf{L_S}$ is the so-called *interaction matrix* and $\mathbf{v} = (V_x, V_y, V_z, \Omega_x, \Omega_y, \Omega_z)$ is the relative camera-object velocity (*kinematic screw*) composed of 3 translational terms and 3 rotational terms. This linear relationship is usually used to design a control law whose aim is to cancel the following vision-based task function

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

where $\mathbf{C}$ is a combination matrix that is usually chosen as $\mathbb{I}_n$ when $n = k$, n being the number of controlled axes. Then, by imposing a decoupled exponential decrease of the task function

$$\dot{\mathbf{e}} = -\lambda \mathbf{e} \tag{3}$$

the following control law can be synthesised

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_S}}^+ \mathbf{e} \tag{4}$$

with λ a positive gain.

The key point on a 2D visual servoing task is to select a suitable set of visual features and to find their dynamics with respect to the camera-scene relative motion. Thanks to previous works concerning the study of improving the performance of 2D visual servoing, we can identify three main desirable conditions for a suitable set of visual features. First, the visual features should ensure the convergence of the system. A necessary condition for this is that the resulting interaction matrix must be not singular, or the number of cases when it becomes singular is reduced and can be analytically identified. A design strategy which can avoid singularities of $\mathbf{L_S}$ is to obtain decoupled visual features, so that each one only controls one degree of freedom. Even if such control design seems to be out of reach, some works concerning the problem of decoupling different subsets of degrees of freedom have been recently proposed [2], [8], [10]. Secondly, it is important to minimise the condition number of the interaction matrix. It is well known that minimising the condition number improves the robustness of the control scheme against image noise and increases the control stability [3]. Finally, a typical problem of 2D visual servoing is that even if an exponential decrease of the error on the visual features is achieved, the obtained 3D trajectory of the camera can be very unsatisfactory. This is usually due to the strong non-linearities in $\mathbf{L_S}$. Some recent works have shown that the choice of the visual features can reduce the non-linearities in the interaction matrix obtaining better 3D camera trajectories [7], [10].

In this paper we exploit the capabilities of structured light in order to improve the performance of visual servoing. A first advantage of using structured light is that the image processing is highly simplified [9] (no complex and computationally expensive point extraction algorithms are required) and the application becomes independent to the object appearance. Furthermore, the main advantage is that the visual features can be chosen in order to produce an optimal interaction matrix.

Few works exploiting the capabilities of structured light in visual servoing can be found in the literature. Andreff et al. [1] introduced on their control scheme a laser pointer in order to control the depth of the camera with respect to the objects once the desired relative orientation was already achieved. Similarly, Krupa et al. [6] coupled a laser pointer to a surgical instrument in order to control its depth to the organ surface, while both the organ and the laser are viewed from a static camera. In general, most of the applications only take profit of the emitted light in order to control one degree of freedom and to make easier the image processing. There are few works facing the issue of controlling several

degrees of freedom by using visual features extracted from the projected structured light. The main contribution in this field is due to Motyl et al. [5], who modelled the dynamics of the visual features obtained when projecting laser planes onto planar objects and spheres in order to fulfill positioning tasks.

In this paper we propose the first step for optimising a visual servoing scheme based on structured light, by focusing on a simple positioning task with respect to a planar object. The paper is structured as follows. Firstly, in Section II, the formulation of the interaction matrix of a projected point is developed. Secondly, the proposed structured light sensor based on laser pointers is presented in Section III. Afterwards, in Section IV, a set of decoupled visual features is proposed for the given sensor. Then, in Section V, our approach is analytically compared with the classic case of using directly image point coordinates. In Section VI, some experimental results using both approaches are shown. Finally, conclusions are discussed in Section VII.

II. PROJECTION OF A LASER POINTER ONTO A PLANAR OBJECT

The simplest case of visual servoing combined with structured light consists of using a laser pointer attached to the camera. Let us consider the case of a planar object as shown in Figure 1.
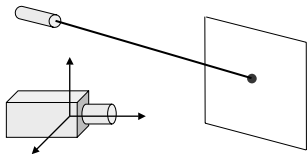

Fig. 1. Case of a laser pointer linked to the camera and a planar object.

The interaction matrix $\mathbf{L_X}$ corresponding to the image point coordinates of the projected laser pointer was firstly formulated by Motyl et al. who modelled the laser pointer as the intersection of two planes [5]. However, the resulting matrix was in function of up to 12 $3D$ parameters: 4 parameters for every one of the two planes defining the laser, and 4 additional parameters for the object plane. Even constraining the two planes modelling the laser to be orthogonal, the interpretation of the interaction matrix is not easy.

A. Proposed modelling

In order to reduce the number of 3D parameters involved in the interaction matrix, the laser pointer can be expressed in the following vectorial form (in the camera frame)

$$\mathbf{X} = \mathbf{X}_r + \lambda \underline{\mathbf{u}} \qquad (5)$$

where that $\underline{\mathbf{u}} = (u_x, u_y, u_z)$ is an unitary vector defining the laser direction while $\mathbf{X}_r = (X_r, Y_r, Z_r)$ is any reference point belonging to the line. The planar object is modelled according to the following equation

$$\Pi_3 : A_3 X + B_3 Y + C_3 Z + D_3 = 0 \qquad (6)$$

where $\underline{\mathbf{n}} = (A_3, B_3, C_3)$ is the unitary normal vector to the plane.

The expression of the depth Z corresponding to the intersection point between the object and the laser pointer can be obtained in a suitable form by solving the equations system built up with the planar object equation Π_3 (6) and the normalised perspective projection equations $\mathbf{x} = \mathbf{X}/Z$, yielding

$$Z = -\frac{D_3}{A_3 x + B_3 y + C_3} \qquad (7)$$

Applying the normalised perspective projection to the vectorial expression of the laser the following equations are obtained

$$\begin{cases} xZ & = & X_r + \lambda u_x \\ yZ & = & Y_r + \lambda u_y \\ Z & = & Z_r + \lambda u_z \end{cases} \qquad (8)$$

By summing the three equations we obtain

$$Z(x + y + 1) = X_r + Y_r + Z_r + \lambda(u_x + u_y + u_z) \quad (9)$$

Then, by applying the depth Z in equation (7) and solving the equation for λ, its expression in function of the 3D parameters of the object and the origin of the ray is obtained as follows

$$\lambda = -\frac{1}{\mu} \left(A_3 X_r + B_3 Y_r + C_3 Z_r + D_3 \right) \qquad (10)$$

with $\mu = \underline{\mathbf{n}}^T \underline{\mathbf{u}} \neq 0$. Note that μ would be 0 when the laser pointer did not intersect with the planar object, i.e. when the angle between $\underline{\mathbf{u}}$ and $\underline{\mathbf{n}}$ is 90°.

Thereafter, by taking into account that $\mathbf{X}_r$ and $\underline{\mathbf{u}}$ do not vary in the camera frame, the time derivative of λ can be calculated yielding

$$\dot{\lambda} = \eta_1 \dot{A}_3 + \eta_2 \dot{B}_3 + \eta_3 \dot{C}_3 + \eta_4 \dot{D}_3 \qquad (11)$$

with

$$\begin{cases} \eta_1 & = & -(\lambda u_x + X_r)/\mu & = & -xZ/\mu \\ \eta_2 & = & -(\lambda u_y + Y_r)/\mu & = & -yZ/\mu \\ \eta_3 & = & -(\lambda u_z + Z_r)/\mu & = & -Z/\mu \\ \eta_4 & = & -1/\mu \end{cases} \qquad (12)$$

From the time derivatives of A_3, B_3, C_3 and D_3 involved in (11) given in [5], we have that (11) depends only on $\underline{\mathbf{u}}$, $\mathbf{X}_r$, $\underline{\mathbf{n}}$, D_3 and $\mathbf{v}$. However, we can note that the unitary orientation vector $\underline{\mathbf{u}}$ can be expressed in terms of the two points belonging to the line $\mathbf{X}_r$ and $\mathbf{X}$ as follows

$$\underline{\mathbf{u}} = (\mathbf{X} - \mathbf{X}_r)/\|\mathbf{X} - \mathbf{X}_r\| \qquad (13)$$

Applying this equation in (11), the resulting expression no longer depends on the explicit orientation of the light beam. The orientation is then implicit in the reference point $\mathbf{X}_r$, the normalised point (x, y) and its corresponding depth Z.

Afterwards, the computation of $\dot{x}$ and $\dot{y}$ is straightforward. First, deriving the vectorial equation of the light beam (5), the following equation is obtained

$$\dot{\mathbf{X}} = \dot{\lambda} \underline{\mathbf{u}} \qquad (14)$$

Then, if the normalised perspective equations are derived and the previous relationship is applied we find

$$\dot{\mathbf{x}} = \frac{\dot{\mathbf{X}}}{Z} - \frac{\mathbf{X}}{Z^2}\dot{Z} \Rightarrow \dot{\mathbf{x}} = \frac{\dot{\lambda}}{Z}(\underline{\mathbf{u}} - \mathbf{x} \cdot u_z) \qquad (15)$$

After some developments, and choosing as reference point $\mathbf{X}_r = (X_0, Y_0, 0)$ we obtain the following interaction matrix

$$\mathbf{L_x} = \frac{1}{\Pi_0}\begin{pmatrix} \frac{-A_3 X_0}{Z} & \frac{-B_3 X_0}{Z} & \frac{-C_3 X_0}{Z} & X_0\varepsilon_1 & X_0\varepsilon_2 & X_0\varepsilon_3 \\ \frac{-A_3 Y_0}{Z} & \frac{-B_3 Y_0}{Z} & \frac{-C_3 Y_0}{Z} & Y_0\varepsilon_1 & Y_0\varepsilon_2 & Y_0\varepsilon_1 \end{pmatrix} \qquad (16)$$

where

$$\begin{aligned} \Pi_0 &= A_3(X_0 - xZ) + B_3(Y_0 - yZ) - C_3 Z \\ \varepsilon_1 &= B_3 - yC_3 \\ \varepsilon_2 &= C_3 x - A_3 \\ \varepsilon_1 &= A_3 y - B_3 x \end{aligned}$$

Note that with respect to the interaction matrix proposed by Motyl et al. in [5], the number of $3D$ parameters concerning the laser beam has been reduced from 8 to 3, i.e. X_0, Y_0 and Z. The orientation of the beam remains implicit in our equations. Concerning the planar object, the number of parameters has been reduced from 4 to 3 since D_3 has been expressed in function of the image coordinates (x, y), the corresponding depth Z of the point and the normal vector to the planar object.

From (16) we directly obtain that the rank of the interaction matrix $\mathbf{L_X}$ is always equal to 1, which means that the time variation of the x and y coordinates of the observed point are linked. Thus, the image point moves along a line as pointed out by Andreff et al. [1].

III. THE PROPOSED STRUCTURED LIGHT SENSOR

In this section we deal with the problem of positioning a camera parallel to a planar object by using only visual data based on some projected structured light. A plane has only 3 degrees of freedom, which means that only three axes of the camera will be controlled (or three combinations of the 6 axis).

In this section we propose a structured light sensor based on laser pointers which intends to achieve three main objectives: decoupling of the controlled degrees of freedom (at least near the desired position), robustness against image noise and control stability, and improving the camera trajectory by removing non-linearities in the interaction matrix.

The choice of a structured light sensor based on laser pointers implies the choice of the number of lasers and how they are positioned and oriented with respect to the camera.

In theory, three laser pointers are enough in order to control 3 degrees of freedom. The positioning of such lasers must be chosen in order to avoid three collinear image points, which would lead to a singularity in the interaction matrix. A simple way to avoid such situation is to position the lasers forming an equilateral triangle so that all the lasers point towards the same direction. However,

we propose a structured light sensor composed of four laser pointers. This choice has been taken because, as it will be demonstrated in the following section, a decoupled control scheme can be achieved by a suitable spatial configuration of the four lasers.

Concretely, we propose to position the lasers forming a symmetric cross centred in the focal point of the camera, so that the axis of the cross are aligned with the X and Y axis of the camera, as shown in Figure 2. Furthermore, we propose to set the direction of the lasers coinciding with the optical axis Z of the camera. The distance of every laser origin to the camera focus is called L. Such lasers
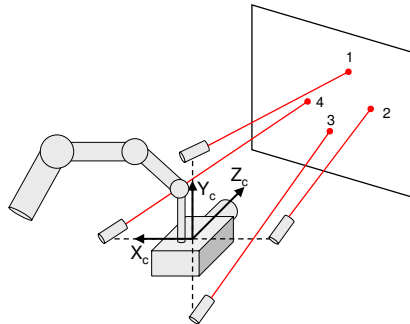

Fig. 2. Laser pointers configuration.

configuration produces a symmetric image like shown in Figure 3 when the image plane is parallel to the planar object, i.e. for the desired position. Given this lasers con-
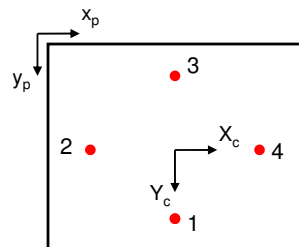

Fig. 3. Example of image when the camera and the object are parallel.

figuration, the coordinates of the reference points (origin) of every laser, and the normalised image coordinates of the projected points are shown in table I.

TABLE I

LASER REFERENCE POINT COORDINATES AND NORMALISED IMAGE
POINT COORDINATES (CAMERA FRAME)

Laser	X_0	Y_0	Z_0	x^*	y^*
1	0	L	0	0	L/Z
2	-L	0	0	-L/Z	0
3	0	-L	0	0	-L/Z
4	L	0	0	L/Z	0

IV. OPTIMISING THE INTERACTION MATRIX

Finding a set of visual features which produces a decoupled interaction matrix for any camera pose seems an

unreachable issue. However, we can find visual features which show a decoupled behaviour near the desired position. In this section we propose a set of 3 visual features which produce a decoupled interaction matrix with low condition number and which removes the non-linearities for all the positions where the image plane is parallel to the object. The first visual feature of the proposed set is the area enclosed in the region defined by the four image points. The area has been largely used to control the depth as in [2], [7], [10].

In our case, the area formed by the 4 image points can be calculated by summing the area of the triangle defined by the points 1, 2 and 3, and the area of the triangle defined by 3, 4, and 1. After some developments we obtain the following formula

$$a = \frac{1}{2}\left((x_3 - x_1)(y_4 - y_2) + (x_2 - x_4)(y_3 - y_1) \right) \quad (17)$$

that only depends on the point coordinates, whose interaction matrices are known.

After some developments, the interaction matrix of the area when the camera is parallel to the object can be obtained as follows

$$\mathbf{L}_a^{\parallel} = (\quad 0 \quad 0 \quad 4L^2/Z^3 \quad 0 \quad 0 \quad 0 \quad) \quad (18)$$

We can note that when the camera is parallel to the object, the $3D$ area enclosed by the 4 projected points is equal to $A = 2L^2$, independently of the depth Z. This is true because the laser pointers have the same direction than the optical axis. Then, knowing that the image area is $a = A/Z^2$ the interaction matrix $\mathbf{L}_a$ can be rewritten as

$$\mathbf{L}_a^{\parallel} = (\quad 0 \quad 0 \quad 2a/Z \quad 0 \quad 0 \quad 0 \quad) \quad (19)$$

The 2 visual features controlling the remaining degrees of freedom are selected from the 4 virtual segments defined according to the Figure 4.
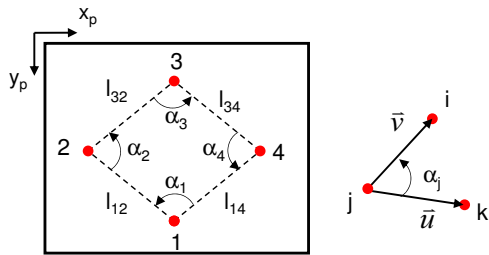

Fig. 4. At left side, virtual segments defined by the image points. At right side, definition of the angle α_j.

An interesting feature is the angle between each pair of intersecting virtual segments. The angle α_j corresponding to the angle between the segment l_{jk} and the segment l_{ji} (see Figure 4) is defined as

$$\sin \alpha_j = \frac{\| \overrightarrow{u} \times \overrightarrow{v} \|}{\| \overrightarrow{u} \| \| \overrightarrow{v} \|} \quad , \quad \cos \alpha_j = \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\| \overrightarrow{u} \| \| \overrightarrow{v} \|} \quad (20)$$

Then, developing the inner and outer product, the angle is obtained from the point coordinates as follows

$$\alpha_j = \arctan \frac{(x_k - x_j)(y_i - y_j) - (x_i - x_j)(y_k - y_j)}{(x_k - x_j)(x_i - x_j) + (y_k - y_j)(y_i - y_j)} \quad (21)$$

Knowing that the derivative of $f(x) = \arctan(x)$ is $\dot{f}(x) = \dot{x}/(1 + x^2)$, the interaction matrix of α_j can easily be calculated.

Then, by choosing the visual features $\alpha_{13} = \alpha_1 - \alpha_3$ and $\alpha_{24} = \alpha_2 - \alpha_4$, the following interaction matrices are obtained when the camera is parallel to the object.

$$\begin{array}{lll}
\mathbf{L}_{\alpha_{13}}^{\parallel} & = & (\quad 0 \quad 0 \quad 0 \quad 2L/Z \quad 0 \quad 0 \quad) \\
\mathbf{L}_{\alpha_{24}}^{\parallel} & = & (\quad 0 \quad 0 \quad 0 \quad 0 \quad 2L/Z \quad 0 \quad)
\end{array} \quad (22)$$

Note that by using the visual feature set $\mathbf{s} = (a, \alpha_{13}, \alpha_{24})$ the interaction matrix is diagonal (for the desired position) so that a decoupled control scheme is obtained with no singularities.

However, it can be also noted that the non-null terms of the interaction matrix are inversely proportional to the depth Z. This will indeed cause the camera trajectory not to be completely satisfactory. As pointed out by Mahony et al. [7], a good visual feature controlling one degree of freedom (dof) is that one whose error function should vary proportionally to the variation of the dof.

Let's start by searching a feature a_n whose time derivative only depends on constant values. Since the time derivative of a depends on the inverse of the depth, we can search a feature of the form $a_n = a^{\gamma}$ so that the depth is cancelled in its time derivative. Then, taking into account all this, the required power γ can be deduced as follows

$$a_n = a^{\gamma} \Rightarrow \dot{a}_n = \gamma a^{\gamma - 1} \dot{a} = \frac{2\gamma A^{\gamma}}{Z^{2\gamma + 1}} \cdot V_z \quad (23)$$

In order to cancel the depth it is necessary that

$$2\gamma + 1 = 0 \Rightarrow \gamma = -1/2 \quad (24)$$

Therefore, the interaction matrix of $a_n = 1/\sqrt{a}$ is

$$\mathbf{L}_{a_n}^{\parallel} = (\quad 0 \quad 0 \quad -1/(\sqrt{2}L) \quad 0 \quad 0 \quad 0 \quad) \quad (25)$$

Following the same methodology, it can be found that by choosing as visual features $\alpha_{n_{13}} = \alpha_{13}/\sqrt{a}$ and $\alpha_{n_{24}} = \alpha_{24}/\sqrt{a}$ the following constant matrices are obtained

$$\begin{array}{lll}
\mathbf{L}_{\alpha_{13_n}}^{\parallel} & = & (\quad 0 \quad 0 \quad 0 \quad \sqrt{2} \quad 0 \quad 0 \quad) \\
\mathbf{L}_{\alpha_{24_n}}^{\parallel} & = & (\quad 0 \quad 0 \quad 0 \quad 0 \quad \sqrt{2} \quad 0 \quad)
\end{array} \quad (26)$$

V. COMPARISON WITH THE IMAGE POINTS APPROACH

In this section the performance of the proposed set of visual features is compared with the set composed of the normalised image point coordinates. The comparison is made from an analytical point of view by calculating the local stability conditions around the desired position. The conditions of stability are expressed in function of the parameters describing the misalignment between the lasers-cross frame and the camera frame. This misalignment is modelled according to the following transformation matrix

$$^c\mathbf{M}_l = \begin{pmatrix} ^c\mathbf{R}_l & ^c\mathbf{T}_l \\ \mathbf{0}_3 & 1 \end{pmatrix} \quad (27)$$

The local stability analysis takes the closed-loop equation of the system in the desired state [8]

$$\dot{\mathbf{e}} = -\lambda \mathbf{C} \mathbf{L_S^*} (\widehat{\mathbf{C} \mathbf{L_S^*}})^{-1} \mathbf{e} \qquad (28)$$

where $\widehat{\mathbf{L_S^*}}$ is the model of the interaction matrix in the desired position where the laser-cross is supposed to be perfectly aligned with the camera frame. On the other hand, $\mathbf{L_S^*}$ is the real interaction matrix in the desired position taking into account the misalignment of the laser-cross described by the frame transformation in equation (27). Then, the local stability is ensured when the product of matrices in the closed-loop equation is definite positive. We remember that a matrix is definite positive if the eigenvalues of its symmetric part have all positive real part. In practice, if a whole model of misalignment is considered, the analytical computation of the eigenvalues becomes too complex, so we face the stability analysis by considering a simplified model where the laser-cross frame is only displaced with respect to the camera frame. Then, the transformation matrix between both frames in equation (27) is such that $^c\mathbf{R}_l = \mathbf{0}_3$ and $^c\mathbf{T}_l = (\epsilon_x, \epsilon_y, \epsilon_z)$.

Applying this simplified misalignment model to the laser-cross frame, the reference points and normalised image coordinates shown in table II are obtained when the camera is parallel to the object. These parameters are used to calculate $\mathbf{L_S^*}$, while the ideal parameters (assuming perfect alignment of both frames) in table I are used to calculate $\widehat{\mathbf{L_S^*}}$. Since both interaction matrices have null values for V_x, V_y and Ω_z, they can be reduced to 3×3 matrices, so that $\mathbf{C}$ can be chosen as the identity matrix.

TABLE II
LASER ORIGINS AND POINT COORDINATES

Laser	X_0	Y_0	Z_0	x^*	y^*
1	ϵ_x	$L+\epsilon_y$	ϵ_z	ϵ_x/Z	$(L+\epsilon_y)/Z$
2	$-L+\epsilon_x$	ϵ_y	ϵ_z	$(\epsilon_x - L)/Z$	ϵ_y/Z
3	ϵ_x	$-L+\epsilon_y$	ϵ_z	ϵ_x/Z	$(\epsilon_y - L)/Z$
4	$L+\epsilon_x$	ϵ_y	ϵ_z	$(L+\epsilon_x)/Z$	ϵ_y/Z

First of all we test the local stability conditions for the set of visual features composed of the normalised image point coordinates $\mathbf{s} = (x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$. Calculating the eigenvalues of the symmetric part of the product $\mathbf{L_S^*}(\widehat{\mathbf{L_S^*}})^{-1}$ and imposing their positivity, the following condition arises

$$\epsilon_x^2 + \epsilon_y^2 < 2L^2 \qquad (29)$$

which is nothing else that the equation of a circle with radius $\sqrt{2}L$. Note that ϵ_z does not affect the local stability, while a displacement of the laser-cross centre in the plane XY is tolerated if it is included in this circle.

If the local stability analysis is made for the proposed set of visual features $\mathbf{s} = (a_n, \alpha_{13_n}, \alpha_{24_n})$ all the eigenvalues are equal to 1. This means that the local stability when using this set of visual features is always ensured when the laser-cross is only displaced with respect to the focal point. Such analytical results demonstrate that the stability

domain when using the proposed visual features is larger than when using simple image points coordinates.

VI. EXPERIMENTAL RESULTS

In this section the performance of the proposed set of visual features is compared with the case when image points are directly used. The experimental setup consists of a six-degrees-of-freedom robot with a camera coupled to its end-effector. The laser-cross structure has been built so that each laser pointer is placed at $L = 15\ cm$ from the cross intersection. The goal of the task consisted of positioning the camera parallel to a plane placed in front of the robotic cell at a distance of $95\ cm$. The desired image was acquired by positioning the robot in such a desired position. The first experiment consisted of moving the camera far away from the desired position about $20\ cm$ and rotations of -40 and 20 degrees were applied with respect to the X and the Y axis, respectively. The results of the servoing by using the normalised image coordinates of the four laser points are shown in Figure 5a while the results when using the proposed visual features are presented in Figure 5b. As can be seen in the final images, the lasers are not perfectly aligned with the camera frame producing a non-symmetric image. However, the misalignment of the laser-cross is small enough to allow the convergence of the system even when using image point coordinates. As expected, the trajectories of the lasers points in the image are straight lines. The approach based on image points presents an almost exponential decrease of the error. However, the camera kinematics show a non-pure exponential behaviour. This fact is due to the non-linearities in the interaction matrix based on image points. On the other hand, as expected, a pure exponential decrease of the error in the visual features, as well as in the camera velocities, is observed in the case of the proposed set of visual features. In terms of numeric conditioning, the interaction matrix based on image point has a condition number equal to 11.17, while for our set of visual features the condition number is 3.3.

A second experiment was carried out in order to test the sensitivity of the system to the alignment between the camera and the lasers-cross. Concretely, the lasers-cross was displaced about $6\ cm$ along the X axis of the camera frame. Then, the camera was moved backwards from the desired position about $40\ cm$ and rotated -20 and 20 degrees around the X and Y axis respectively. With such a large misalignment of the lasers-cross, the approach based on image point coordinates rapidly failed. Meanwhile, the approach based on our set of visual features still converged as shown in Figure 5c. The image trajectories of points 1 and 3 are no longer parallel to the Y axis due to the misalignment. The control on depth and on Ω_x are almost unaffected by the cross misalignment. However, the feature α_{24_n} controlling Ω_y presents a non-pure exponential decrease because it is certainly affected by the large misalignment.

The robustness of the proposed visual features against large misalignment errors of the lasers-cross was already
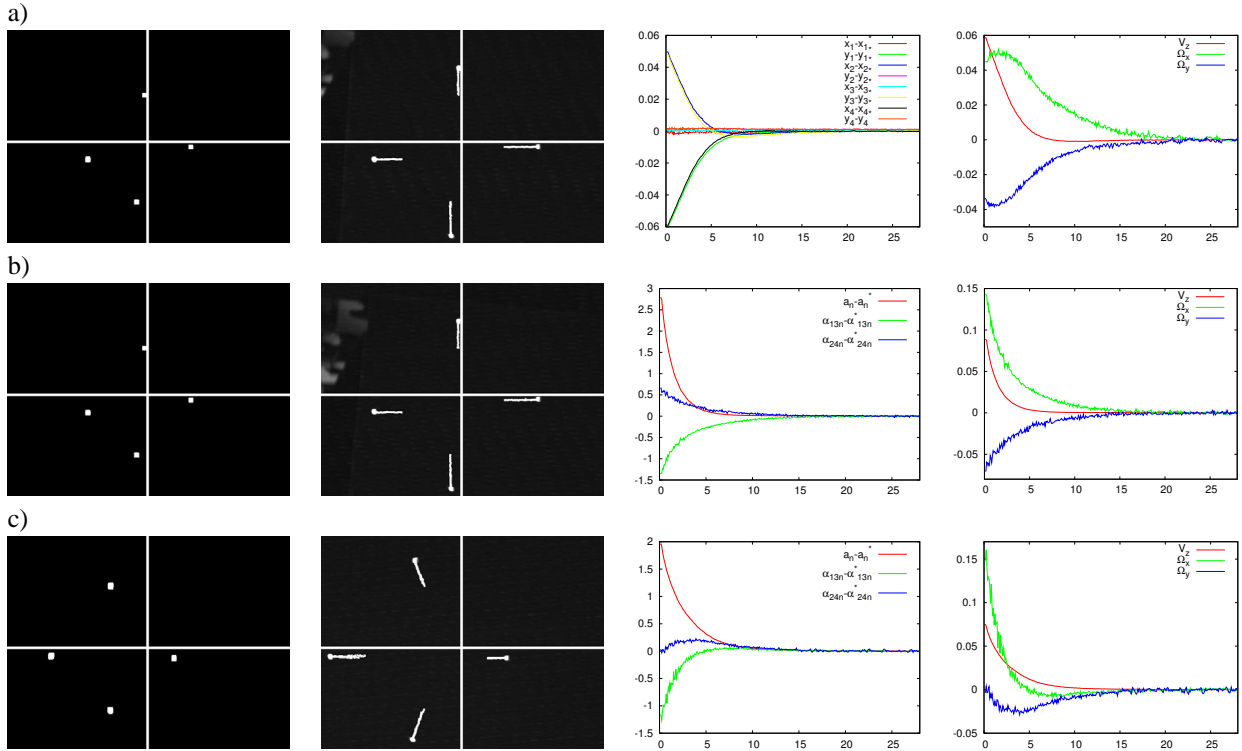
Fig. 5. a) 1st experiment using image point coordinates, b) 1st experiment using the proposed visual features, c) 2nd experiment. Each column shows (from left to right): initial image; final image (with the trajectories of the laser points); visual features $\mathbf{s} - \mathbf{s}^*$ versus time (in s); camera velocities (m/s and rad/s) versus time (in s)

expected from the local stability analysis. Furthermore, we can intuitively understand this robustness because the area and relative angles defined upon the 4 image points are invariant to $2D$ rotations, translations and scaling (when the image plane is parallel to the target and all the lasers have the same orientation).

VII. CONCLUSIONS

In this paper the task of positioning a camera parallel to a planar object has been faced. An approach based on visual servoing combined with structured light has been presented. Thanks to the flexibility of using structured light, the system behaviour has been optimised in three main aspects: control stability and decoupling, well-conditioning and better controlled camera velocities.

The structured light sensor proposed is composed of four lasers pointers which are placed symmetrically with respect to the camera and pointing towards the same direction than the optical axis. Such configuration is suitable for defining a set of 3 decoupled visual features (near the desired position) which guarantees a higher degree of robustness against the lasers calibration than when using image point coordinates. Furthermore, non-linearities in the interaction matrix are removed producing a better mapping from the feature space to the camera velocities.

The better performance of our approach in front of image point coordinates has been demonstrated through a local stability analysis and experimental results.

REFERENCES

[1] N. Andreff, B. Espiau, and R. Horaud. Visual servoing from lines. *Int. Journal of Robotics Research*, 21(8):679–700, August 2002.

[2] P. I. Corke and S. A. Hutchinson. A new partitioned approach to image-based visual servo control. *IEEE Trans. on Robotics and Automation*, 17(4):507–515, August 2001.

[3] J. T. Feddema, C. S. G. Lee, and O. R. Mitchell. Weighted selection of image features for resolved rate visual feedback control. *IEEE Trans. on Robotics and Automation*, 7(1):31–47, February 1991.

[4] S. Hutchinson, G. Hager, and P. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5):651–670, 1996.

[5] D. Khadraoui, G. Motyl, P. Martinet, J. Gallice, and F. Chaumette. Visual servoing in robotics scheme using a camera/laser-stripe sensor. *IEEE Trans. on Robotics and Automation*, 12(5):743–750, 1996.

[6] A. Krupa, J. Gangloff, C. Doignon, M. Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux. Autonomous 3d positionning of surgical instruments in rototized laparoscopic surgery using visual servoing. *IEEE Trans. on Robotics and Automation*, 19(5):842–853, 2003.

[7] R. Mahony, P. Corke, and F. Chaumette. Choice of image features for depth-axis control in image-based visual servo control. In *IEEE/RSJ Int. Conf. on Intellingent Robots and Systems*, Lausanne, Switzerland, September 2002.

[8] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):238–250, April 1999.

[9] J. Salvi, J. Pagès, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, 2004.

[10] O. Tahri and F. Chaumette. Image moments: Generic descriptors for decoupled image-based visual servo. In *IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 1185–1190, New Orleans, LA, April 2004.