Visual servoing by means of structured light for plane-to-plane positioning

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

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Visual servoing by means of structured light for plane-to-plane positioning

Jordi Pagès, Christophe Collewet, François Chaumette and Joaquim Salvi

N°5579

May 2005

Systèmes cognitifs

ISSN 0249-6399 ISRN INRIA/RR--5579--FR+ENG

Visual servoing by means of structured light for plane-to-plane positioning

Jordi Pagès*, Christophe Collewet†, François Chaumette‡ and Joaquim Salvi §

Systèmes cognitifs
Projet Lagadic

Rapport de recherche n ° 5579 — May 2005 — 99 pages

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, so-called plane-to-plane positioning, has been treated for a long time in visual servoing. Our approach is based on linking to the camera a structured light emitter composed of four parallel laser pointers so that its distribution is aimed to produce a suitable set of visual features. The aim of using structured lighting is not only for easing the image processing and allowing low-textured objects to be treated, but also for producing a control scheme with nice properties like decoupling, stability and good camera trajectory. This paper proposes several sets of visual features and analyzes their performance and robustness against different types of calibration errors both analytically and experimentally.

Key-words: Visual servoing, structured light, plane-to-plane positioning, stability analysis, decoupled visual features, laser pointers

(Résumé : tsvp)

* jpages@eia.udg.es
† christophe.collewet@cemagref.fr
‡ Francois.Chaumette@irisa.fr
§ qsalvi@eia.udg.es

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : 02 99 84 71 00 - International : +33 2 99 84 71 00
Télécopie : 02 99 84 71 71 - International : +33 2 99 84 71 71

Asservissement visuel basé sur la lumière structurée pour un positionnement plan à plan

Résumé : Nous traitons dans cet article de la réalisation d'une tâche de positionnement par asservissement visuel dans le cas d'une caméra embarquée. La tâche de positionnement, dite de plan à plan, consiste à faire coïncider le plan du capteur avec celui de l'objet supposé plan. Un tel problème n'est pas nouveau en soi, si ce n'est que notre approche repose sur l'emploi d'un éclairage structuré. Plus précisément, quatre pointeurs laser sont utilisés et le choix de leur distribution spatiale permet l'obtention d'informations visuelles pertinentes. Dans ce contexte, l'utilisation d'un tel éclairage ne permet pas seulement de faciliter le traitement de l'image ou de pouvoir prendre en compte des objets faiblement texturés mais d'élaborer une loi de commande présentant des propriétés remarquables de découplage et de stabilité tout en conduisant également à une bonne trajectoire de la caméra. Cet article propose plusieurs ensembles d'informations visuelles et analyse, analytiquement et expérimentalement, leurs comportements en terme de performance et de robustesse vis-à-vis de différents types d'erreur de calibration du dispositif laser.

Mots-clé : asservissement visuel, lumière structurée, tâche plan à plan, analyse de stabilité, découplage d'informations visuelles, pointeurs laser.

Contents

1 Introduction

Visual servoing is a largely used technique which intends to control a robot in order to fulfill a certain task by using data provided by visual sensors (usually cameras). Certain data extracted or calculated from the sensor signal are used in a closed-loop control law which leads to the execution of a task like positioning with respect to static objects or target tracking. A comprehensive survey on the different visual servoing approaches can be found, for example in [32].

This paper focuses on the combination of visual servoing and structured light, which is a type of approach which has not been very exploited up to date. The interest of combining both techniques is that positioning tasks with respect to non textured objects becomes feasible and the image processing is much simpler. In addition to this, with an appropriate choice of the structured light emitter configuration the control loop can be optimized in terms of decoupling and stability. In order to demonstrate this fact we have chosen the following classical task: achieving a plane-to-plane virtual link between a sensor attached to the robot end-effector and a planar object. The paper presents several sets of visual features exhibiting nice stability properties against calibration errors in both the camera intrinsic parameters and the structured light emitter, and different behaviors in terms of decoupling visual features. To do this we propose a set of visual features which decouples the rotational degrees of freedom (dof) from the translational ones and produces suitable camera velocities thanks to a proper configuration of the structured light emitter.

The paper is structured as follows. In Section 2 a brief introduction of the main visual servoing approaches is presented. The structured light sensor proposed to fulfill the plane-to-plane virtual link and its modelling is presented in Section 3. Section 4 reviews the task function definition used in visual servoing, the derivation of a proportional control law and the procedure used to analyze its stability. Afterwards, four visual servoing approaches exploiting the projected light are presented. First, Section 5 deals with several position-based approaches based on reconstructing the object pose by triangulation. Then, a simple $2D$ approach based on image points coordinates is shown in Section 6. After that, in Section 7 a $2D$ approach based on the area corresponding to the projected pattern and combinations of angles extracted from the image is analyzed. The last approach is based on a robust non-linear combination of image points which is presented in Section 8. Some simulations and experiments showing the behavior obtained with each one of the proposed approaches are shown in Section 9. The paper ends with conclusions.

2 An overview of the main visual servoing approaches

Visual servoing is based on controlling the motion of a robot by using a camera. Hereafter, we present a brief overview focusing on the most typical system configuration called *eye-in-hand*. In such systems, a camera is attached to the end-effector of the robot so that the control law computes at each iteration of the loop the camera velocities required to fulfill the

task. Then, the robot jacobian is used to compute the corresponding robot joints velocities. Several types of information can be used, which are detailed in the next sections.

2.1 Position-based visual servoing

In the case of position-based visual servoing, $3D$ information computed from the image(s) are used in the control law. Two cases can be distinguished depending of whether the object model is known or not. In the former case, 3D information can be obtained by using pose estimation algorithms based on points [22, 29, 30, 37, 55, 56], straight lines [24], polyhedric objects [25], conics [20] and some quadrics like spheres and cylinders [23]. In the latter case, when the object model is unknown, 3D information is obtained by reconstruction [19]. This can be done by using a passive stereovision system composed of two calibrated cameras [8] or by using different strategies based on a unique camera. In this case, two or more images must be considered and $3D$ reconstruction can be obtained by structure from motion [31], dynamic vision [3] or by active vision [11].

A classical advantage granted to position-based approaches refers to their ability to produce suitable camera trajectories since the control is made in the cartesian space [9]. In addition to this, the rotational velocities can be decoupled from the translational ones, even if this only holds when the system is perfectly calibrated. There are three main drawbacks inherent to position-based approaches. First, the pose estimation or $3D$ reconstruction algorithms are sensible to image noise. Secondly, since no control is made in the image space, the features used for the reconstruction can leave the image bounds. Finally, the stability analysis taking into account calibration errors is in most cases impossible to face.

2.2 Image-based visual servoing

$2D$ visual servoing is based on using visual features directly calculated from the images which are used as input in the control scheme. Thus, this type of approaches tend to avoid the use of any object model.

Former works started using image points which are still today one of the most popular primitives [7, 14, 26–28].Other $2D$ primitives have been modelled like straight lines [26, 39], segments, circles, ellipsis, cylinders and spheres [26, 42].

On the recent years, more complicated primitives have been taken into account. For example, complex contours [13, 16], the principal components of the image [21] and image moments [10].

Some other works tend to combine different visual features cited above in a unique control scheme. For example, in [17], point coordinates, the area enclosed by points and the angle between segments are used to improve the performance of the system.

Image-based visual servoing is traditionally robust against modelling errors of the system. Furthermore, since control is made in the image space, it is easier to design strategies to avoid image features going out the image bounds [17, 43]. However, since the features are usually strongly coupled, the generated camera trajectory cannot be very suitable (some efforts in order to decouple degrees of freedom can be found for example in [17, 38, 53]).

Other drawbacks are the possibility of reaching a singularity in the control law or falling into local minima [9].

2.3 Hybrid visual servoing

This approach combines $2D$ with $3D$ features. In case of knowing the model of the object, classic pose recovering algorithms can be used to estimate some $3D$ features as in position-based visual servoing. However, several model-free approaches have been presented. A lot of approaches have been done for the case when the desired image is known. Some of them are based on recovering the partial pose between the camera and the object from the desired and the current image [40,41,44]. The obtained homography is decomposed in a rotation matrix and a scaled translation. Note that if both displacement components are directly used in a control law, a model-free position-based visual servoing scheme arises as in [5]. However, the most usual choice is to combine part of the $3D$ information recovered with $2D$ features like an image reference point. Other approaches exist, like the one presented in [50], where the depth distribution of the object is explicitly included in the visual features. Another example is found in [4], where the binormalized plücker coordinates of $3D$ lines are used, so that the depth to the lines are estimated from a pose calculation algorithm assuming a partial knowledge of the object structure. When the desired image is unknown, the rotation to be executed can be calculated by doing a local reconstruction of the object normal in a certain point [3, 12, 15, 46, 51].

Typical advantages of hybrid approaches are: they are usually model-free (do not require to know the object model even if in most cases the desired image must be known); they allow control in the image since $2D$ information is included; they can exhibit decoupling between translational and rotational degrees of freedom; stability analysis in front of camera and robot calibration errors is often feasible [40,44]. On the other hand, the main drawback that can appear is the sensibility to image noise affecting the partial pose algorithm.

2.4 Dynamic visual servoing

The analysis of the $2D$ motion appearing in a sequence of images can be used to obtain geometric visual features which can be then used in a visual servoing scheme like the presented in the previous sections [3,12,15,51]. However, if the visual features are, for example, the parameters of the $2D$ motion model itself, a dynamic visual servoing scheme is defined [18,49].

Essentially, these techniques define the vision task in terms of dynamic visual features so that, for example, the system can be controlled in order to observe a desired $2D$ motion field along the sequence of acquired images.

2.5 Visual servoing by means of structured light

Although the large domain of applications that can be faced with classic visual servoing, there are still some open issues. For example, classic techniques can not cope with the simple problem of keeping a mobile robot running parallel to a wall containing no landmarks or

easily distinguishable visual features. This problem can be generalized to any task where it is necessary to position the camera with respect to a large object which presents a uniform appearance so that it is not possible to extract visual features like characteristic points, straight lines, contours, regions, etc. A possible solution to this problem is to use structured light emitters to project visual features in such objects. Thus, the application field of visual servoing can be enlarged to applications like painting, welding, sealing or cutting surfaces, or docking tasks with respect to large objects.

Two new configurations appear in an eye-in-hand systems when using structured light devices. If the structured light emitter remains static and separated from the camera and the robot, the projected marks remain static onto the object surface if it is also static (see Fig. 1b). The advantage of such approach is that classic visual servoing can be directly applied. However, for some applications it can be unappropriated due to the typical problems of classic $2D$ visual servoing (singularities, local minima, etc.) and some new ones related to the structured light emitter. For example, if the target object moves or its position is totally unknown it can fall outside the field of view of the structured light emitter. A possible solutions is to attach the structured light emitter to the robot so that it remains rigidly linked to the camera (see Fig. 1a). This approach must take into account that the projected visual marks do not remain static onto the object surface, but that they vary their position when the camera moves.

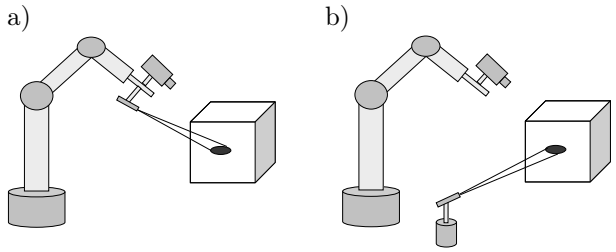

Figure 1: Two possible configurations of an eye-in-hand system and a structured light emitter

If the structured light emitter is assumed to be perfectly calibrated with respect to the camera, triangulation can be used in order to reconstruct the illuminated scene [1, 6, 33, 47] and use the $3D$ data in a visual servoing loop. However, there are no works providing results concerning the sensitivity to image noise and camera-sensor calibration of such type of approach.

Few works exploiting the combination of visual servoing and structured light can be found in the bibliography. Albus et al. proposed in [2] to attach two parallel laser planes to a camera in order to estimate the surface pose of planar objects. The vision system was intended to be used in a visual servoing eye-in-hand configuration. Nevertheless, no experimental results were reported and the suggested approach seems to rely on an accurate calibration of the laser planes. A similar approach was proposed by Kent et al. [34]. In this

case, however, the sensors linked to the camera of the robot end-effector are a unique laser plane and several infra-red proximity sensors.

Samson et al. [48] modelled a general thin-field rangefinder whose output is a function of the distance between the sensor and the planar object. The variation of this measure with respect to a change in the relative pose sensor-object was formulated so that it can be used in a control law.

Motyl et al. [35] properly investigated eye-in-hand systems with structured light emitters attached to the camera. They adapted the calculation method of the image jacobian corresponding to simple geometric primitives to the case when they are provided by the projection of laser planes onto simple objects. It was shown that in this type of configurations, both the projected light and the object surface must be modelled in the image jacobian. They modelled different cases where laser planes are projected to planar and spherical objects. Their study also focuses on fixing virtual links between the camera and such objects with the aid of laser planes. Different sets of visual features extracted from the projected light were formulated.

Some years later, Andreff et al. included depth control by using structured light in their $2D$ $1/2$ visual servoing approach based on $3D$ lines [4]. Their control scheme was however depth-invariant. They prevented this lack of depth-control by providing a laser pointer to the camera. They first formulated the variation of the distance between the laser and its projection onto the object due to robot motion. The interaction matrix of this $3D$ feature was formulated taking into account a planar object. The result was equivalent to the presented by Samson et al. [48]. Afterwards, they showed that the projected point lies always onto a line in the camera image, which is the unique epipolar line of the particular stereo system composed by the camera and the laser. Then, they chose as $2D$ visual feature the distance along the epipolar line between the current position of the point and a certain reference point. The variation of such feature was related to the variation of the distance between the laser and the object.

More recently, Krupa et al. [36] applied visual servoing and structured light to laparoscopic surgery. In their application, two incision points are made into the patient body: one for introducing an endoscopic camera and the other to introduce a surgical instrument held by the end-effector of a robot. The camera pose with respect to the robot frame is unknown. The task consists of moving the instrument to a certain point of the observed image keeping a desired depth to the underlying organ. Due to constraints of the surgical incision and the type of instrument, only three degrees of freedom must be controlled. The instrument holder is a laser pointer which has three LEDs which are collinear with the projected spot. The depth from the surgical instrument to the organ is estimated by using the cross ratio of the four physical points and the corresponding image points. The rotational degrees of freedom are controlled by an online identification of the image jacobian of the projected spot. On the other hand, the depth is directly controlled from the calculated estimation.

3 A proposal of structured light sensor for plane-to-plane positioning

In this paper we present an eye-in-hand system with a structured light sensor attached to the camera. The interest on combining visual servoing with structured light is not only to simplify the image processing, but also to treat low-textured objects containing no landmarks.

The goal of the task here faced consists on positioning the camera parallel to a planar object. Such type of task, namely plane-to-plane positioning, aims to fix a virtual link between the camera image plane and the object plane. With this classic task we aim to demonstrate that using a suitable structured light emitter, the performance of the visual servoing scheme can be optimized in terms of decoupling, stability and camera trajectory.

The structured light sensor is based on laser pointers since they are low-cost and easily available. Theoretically, three non-collinear points are enough to recover the parameters of the equation of a planar object. Consequently, we initially designed a sensor composed of three laser pointers. Nevertheless, we found that better results can be obtained by using four laser pointers in order to decouple visual features.

The structured light sensor that we propose consists of four laser pointers attached to a cross-shaped structure as shown in Fig. 2a. The direction of the lasers have been chosen to be equal so that four points are orthogonally projected to the planar object (see Fig. 2b). This causes that the projected shape enclosed by the four points is invariant to the distance between the object and the laser-cross. This invariance will be very useful as will be shown in following sections. The symmetric distribution of the lasers in the cross structure will be also useful for decoupling of visual features.

Consequently, the model of the proposed sensor is as follows. The laser-cross has its own frame $\{\mathrm{L}\}$ so that all the lasers have the same orientation vector ${}^{\mathrm{L}}\underline{\mathbf{u}} = (0, 0, 1)$. Furthermore, the lasers are placed symmetrically so that two of them lie on the X_L axis and the other two on the Y_L axis. All the lasers are positioned at a distance L from the origin of the laser-cross frame. The structured light sensor is modelled assuming that it is ideally attached to the camera of the robot manipulator as shown in Fig. 2b . As can be seen, in this model, the cross-laser frame perfectly coincides with the camera frame, so that the structured light sensor is placed just in the camera origin and the lasers point toward the direction of the camera optical axis. Whenever the camera and the planar object are parallel, the projected laser points exhibit a symmetric distribution onto the object surface and also in the camera image (see Fig. 3), which will allow us to find decoupled visual features.

These assumptions have been only taken for modelling issues. However, it is perhaps not always possible 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 will be a key point when analyzing the different approaches presented in this paper.

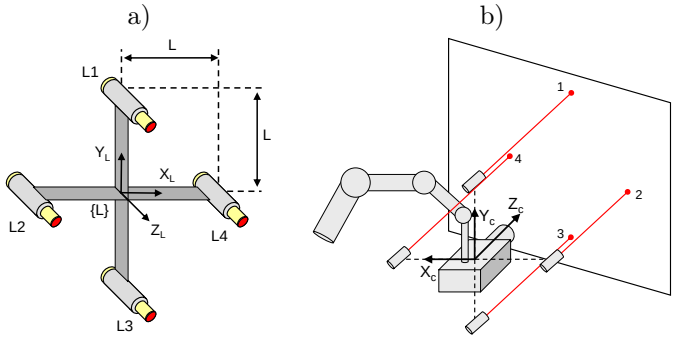

Figure 2: System architecture. a) The proposed structured light sensor. b) Ideal configuration of the robot manipulator, camera and structured light sensor.

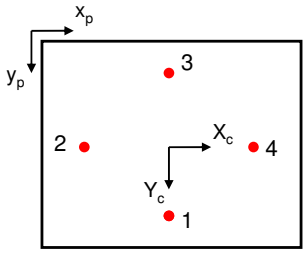

Figure 3: Camera image when it is parallel to the object at a given depth.

Next section presents the modelling of a laser pointer and the image jacobian corresponding to the projected point onto a planar object. Afterwards, the whole model of the proposed structured light emitter is presented under ideal conditions and different types of calibration errors.

3.1 Laser pointer modelling

In visual servoing, given a set of visual features $\mathbf{s}$ extracted from an image, its variation due to the relative camera-object velocity (*kinematic screw*) is expressed in the well known equation

$$\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v} \tag{1}$$

being $\mathbf{v} = (V_x, V_y, V_z, \Omega_x, \Omega_y, \Omega_z)$ the camera velocity screw assuming a static object, and $\mathbf{L_s}$ the image jacobian known as *interaction matrix*.

Given a $3D$ point $\mathbf{X} = (X, Y, Z)$ fixed to the observed object, its normalized coordinates $\mathbf{x} = (x, y)$ resulting of the perspective projection $\mathbf{x} = \mathbf{X}/Z$ are the most widely used features in image-based visual servoing. The interaction matrix of a fixed point of coordinates (x, y) is [26, 27]

$$\mathbf{L_x} = \begin{pmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \\ 0 & -1/Z & y/Z & 1+y^2 & -xy & -x \end{pmatrix} \tag{2}$$

note that the only $3D$ information included in the interaction matrix is the depth of the point which appears in the translational components.

The analog case when working with structured light consists of using a laser pointer so that the intersection of the laser with the object produces also a point $\mathbf{X}$ as shown in Fig. 4. When the laser pointer is linked to the camera in an eye-in-hand configuration, the time variation of the observed point $\mathbf{x}$ depends also on the geometry of the object since $\mathbf{X}$ is not a static physical point. Therefore, some modelling of the object surface must be included in the interaction matrix.

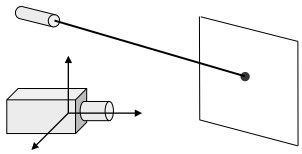

Figure 4: Case of a laser pointer and a planar object.

In this work we focus on the case of planar objects, which are modelled according to the following explicit equation

$$\underline{\mathbf{n}}^\top \mathbf{X} + D = 0 \tag{3}$$

being $\underline{\mathbf{n}} = (A, B, C)$ the unitary normal vector to the plane and D its distance to the origin of the camera frame. Hereafter, we take the convention that $C > 0$ which implies that $D < 0$ since the object is in front of the camera.

Motyl et al. were the first on formulating the interaction matrix of a projected point onto a planar object [35]. In their case, the projected point was the result of the intersection of two laser planes with the planar object. Note that a laser pointer (straight line in the space) can be modelled as the intersection of two planes, so that it is equivalent to projecting two intersecting laser planes. The interaction matrix proposed by Motyl et al. has the disadvantage of depending on 12 $3D$ parameters: 4 parameters for every one of the two laser planes (normal vector and distance to the origin) plus 4 parameters for the planar object. Furthermore, the explicit depth Z of the point does not appear in the interaction matrix.

In a more natural way, the laser pointer can be modelled with a vectorial equation as follows (all the parameters are expressed in the camera frame)

$$\mathbf{X} = \mathbf{X}_r + \mu\underline{\mathbf{u}} \tag{4}$$

where $\underline{\mathbf{u}} = (u_x, u_y, u_z)$ is an unitary vector defining the laser direction, $\mathbf{X}_r = (X_r, Y_r, Z_r)$ is any reference point belonging to the straight line, and μ is the distance from $\mathbf{X}_r$ to $\mathbf{X}$.

By deriving the above expression and taking into account that both $\mathbf{X}_r$ and $\underline{\mathbf{u}}$ do not change when the camera moves we find that the time derivative of the projected point is

$$\dot{\mathbf{X}} = \dot{\mu}\underline{\mathbf{u}} \tag{5}$$

then, deriving the normalized coordinates $\mathbf{x}$ and using the above result we can write

$$\dot{\mathbf{x}} = \frac{\dot{\mathbf{X}}}{Z} - \frac{\mathbf{X}}{Z^2}\dot{Z} = \frac{1}{Z}\dot{\mu}\underline{\mathbf{u}} - \frac{\mathbf{x}}{Z}\dot{Z} \tag{6}$$

from (5) we have that $\dot{Z} = \dot{\mu}u_z$ so that

$$\dot{\mathbf{x}} = \frac{\dot{\mu}}{Z}(\underline{\mathbf{u}} - \mathbf{x}u_z) \tag{7}$$

In order to calculate $\dot{\mu}$ let first express μ in function of the $3D$ parameters. By substituting $\mathbf{X}$ in (3) by its expression in (4) we have

$$\mu = -\frac{1}{\underline{\mathbf{n}}^\top\underline{\mathbf{u}}}(\underline{\mathbf{n}}^\top\mathbf{X}_r + D) \tag{8}$$

deriving this expression we obtain

$$\dot{\mu} = -\frac{1}{\underline{\mathbf{n}}^\top\underline{\mathbf{u}}}(\dot{\underline{\mathbf{n}}}^\top\mathbf{X}_r + \dot{D}) + \frac{\underline{\mathbf{n}}^\top\mathbf{X}_r + D}{(\underline{\mathbf{n}}^\top\underline{\mathbf{u}})^2}\dot{\underline{\mathbf{n}}}^\top\underline{\mathbf{u}} \tag{9}$$

which can be reduced to

$$\dot{\mu} = -\frac{1}{\underline{\mathbf{n}}^\top\underline{\mathbf{u}}}(\dot{\underline{\mathbf{n}}}^\top(\mathbf{X}_r + \mu\underline{\mathbf{u}}) + \dot{D}) \tag{10}$$

and finally, applying (4) the time derivative of μ is

$$\dot{\mu} = -\frac{1}{\underline{\mathbf{n}}^\top\underline{\mathbf{u}}}(\dot{\underline{\mathbf{n}}}^\top\mathbf{X} + \dot{D}) \tag{11}$$

Taking into account the time derivatives of the planar object parameters [54]

$$\begin{pmatrix} \dot{\underline{\mathbf{n}}} \\ \dot{D} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{3\times 3} & [\underline{\mathbf{n}}]_\times \\ \underline{\mathbf{n}}^\top & \mathbf{0}_{1\times 3} \end{pmatrix} \mathbf{v} \tag{12}$$

where $[\underline{\mathbf{n}}]_\times$ is the antisymmetric matrix associated to vector $\underline{\mathbf{n}}$, the interaction matrix of μ is

$$\mathbf{L}_\mu = -\frac{1}{\underline{\mathbf{n}}^\top \underline{\mathbf{u}}} \begin{pmatrix} \mathbf{n}^\top & (\mathbf{X} \times \underline{\mathbf{n}})^\top \end{pmatrix} \tag{13}$$

The equivalence of this formula to the one presented by Samson et al. [48] and the one provided by Andreff et al. [4] is shown in appendix A.

By using the time derivative of μ, Equation (7) can be rewritten as follows

$$\dot{\mathbf{x}} = -\frac{1}{\underline{\mathbf{n}}^\top \underline{\mathbf{u}}} (\underline{\mathbf{u}} - \mathbf{x} u_z) \frac{(\dot{\underline{\mathbf{n}}}^\top \mathbf{X} + \dot{D})}{Z} \tag{14}$$

note that the time derivative of $\mathbf{x}$ is expressed in function of 7 $3D$ parameters, namely Z, $\underline{\mathbf{n}}$ and $\underline{\mathbf{u}}$. The only parameters concerning the laser configuration are the components of its direction vector $\underline{\mathbf{u}}$. This result can be still improved by expressing $\underline{\mathbf{u}}$ as follows

$$\underline{\mathbf{u}} = (\mathbf{X} - \mathbf{X}_r)/\|\mathbf{X} - \mathbf{X}_r\| \tag{15}$$

applying this expression in (14) and after some developments $\dot{\mathbf{x}}$ becomes

$$\dot{\mathbf{x}} = \frac{(\mathbf{X}_r - \mathbf{x} Z_r)}{\underline{\mathbf{n}}^\top (\mathbf{X} - \mathbf{X}_r)} \left(\frac{\dot{D}}{Z} + \dot{\underline{\mathbf{n}}}^\top \mathbf{x} \right) \tag{16}$$

Note that the expression does not longer depend on the orientation of the laser $\underline{\mathbf{u}}$ but on its reference point $\mathbf{X}_r$. Furthermore, if the reference point $\mathbf{X}_r$ is chosen as $\mathbf{X}_r = \mathbf{X}_0 = (X_0, Y_0, 0)$, which corresponds to the intersection of the straight line modelling the laser and the plane $Z = 0$ of the camera frame, the expression simplifies to

$$\dot{\mathbf{x}} = \frac{\mathbf{X}_0}{\underline{\mathbf{n}}^\top (\mathbf{X} - \mathbf{X}_0)} \left(\frac{\dot{D}}{Z} + \dot{\underline{\mathbf{n}}}^\top \mathbf{x} \right) \tag{17}$$

Applying the time derivatives of the plane parameters in (12) into (17) the interaction matrix of a projected point is

$$\mathbf{L}_\mathbf{x} = \frac{1}{\Pi_0} \begin{pmatrix} \frac{-AX_0}{Z} & \frac{-BX_0}{Z} & \frac{-CX_0}{Z} & X_0\varepsilon_1 & X_0\varepsilon_2 & X_0\varepsilon_3 \\ \frac{-AY_0}{Z} & \frac{-BY_0}{Z} & \frac{-CY_0}{Z} & Y_0\varepsilon_1 & Y_0\varepsilon_2 & Y_0\varepsilon_3 \end{pmatrix} \tag{18}$$

$$\Pi_0 = \mathbf{n}^\top (\mathbf{X}_0 - \mathbf{x} Z)$$

$$(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \underline{\mathbf{n}} \times (x, y, 1) \tag{19}$$

note that Π_0 is the distance of the reference point $\mathbf{X}_0$ to the object. With respect to the interaction matrix given by Motyl et al. [35], the number of $3D$ parameters concerning the laser pointer has been reduced from 8 to 3, i.e. X_0, Y_0 and Z. The orientation $\underline{\mathbf{u}}$ of the laser remains implicit in our equations. Concerning the planar object, the number of parameters has been reduced from 4 to 3 since D has been expressed in function of the image coordinates (x, y), the corresponding depth Z, and the normal vector to the planar object $\underline{\mathbf{n}}$.

The rank of $\mathbf{L_x}$ is always equal to 1, which means that the time variation of the x and y coordinates are linked. As already pointed out by Andreff et al. [4], the image point $\mathbf{x}$ moves always along a straight line (hereafter called *epipolar line*). Andreff et al. did not specify the interaction matrix of a projected point, but the interaction matrix of the distance of the point to a certain origin of the epipolar line. Furthermore, the interaction matrix related to this feature was expressed in a frame centered in the laser reference point (similar to $\mathbf{X}_0$), and was expressed in function of the angles defining the normal of the planar object, the angle between the laser pointer and the camera optical axis, and the distance between the camera center and $\mathbf{X}_0$. The main problem concerning the feature used by Andreff et al. is that a convention must be taken to chose the sign of the distance from $\mathbf{x}$ to the origin of the epipolar line.

3.2 Model of the structured light sensor

This section presents the parametric model of the system composed by the camera and the structured light sensor composed of 4 laser pointers. The parameters of the model are the ones appearing in the interaction matrix of every projected point, which are here summarized

- Reference point of each laser pointer $(X_0, Y_0, 0)$.
- Normalized image point coordinates $\mathbf{x} = \mathbf{X}/Z$ of every projected point.
- Depth Z of the projected points.

The reference points $\mathbf{X}_0$ are determined by the actual pose of the laser-cross with respect to the camera frame. The $3D$ coordinates $\mathbf{X}$ of the projected points can be calculated from $\mathbf{X}_0$, the orientation of the lasers $\underline{\mathbf{u}}$ and the object pose. Concretely, from the equation of the line modelling a laser pointer and the equation of the planar object

$$\left\{ \begin{array}{rcl} \mathbf{X} & = & \mathbf{X}_0 + \mu\underline{\mathbf{u}} \\ 0 & = & \underline{\mathbf{n}}^\top \mathbf{X} + D \end{array} \right. \tag{20}$$

we can obtain the depth of the projected point

$$Z = -\frac{u_z(\underline{\mathbf{n}}^\top \mathbf{X}_0 + D)}{\underline{\mathbf{n}}^\top \underline{\mathbf{u}}} + Z_0 \tag{21}$$

and the real normalized coordinates of the image point

$$\begin{aligned} x_{real} &= \frac{u_x}{u_z} + \frac{X_0}{Z} \\ y_{real} &= \frac{u_y}{u_z} + \frac{Y_0}{Z} \end{aligned} \tag{22}$$

The following subsections present the values of the model parameters under different types of relative poses between the camera and the laser-cross. First of all, the ideal case is presented where the laser-cross frame is perfectly aligned with the camera frame. Afterwards, the parameters of the model are calculated under different types of misalignment between the camera and the laser-cross.

3.2.1 Ideal model

First of all, let us consider that the structured light sensor is perfectly attached to the camera so that the laser-cross frame perfectly coincides with the camera frame. In such a case the model parameters are shown in Table 1. The (x, y) and Z parameters have been calculated taking into account that the ideal orientation of the lasers coincides with the optical axis direction so that ${}^{C}\underline{\mathbf{u}} = {}^{L}\underline{\mathbf{u}} = (0, 0, 1)$.

Table 1: Ideal model parameters

Laser	X_0	Y_0	x	y	Z
1	0	L	0	L/Z_1	$-(BL+D)/C$
2	$-L$	0	$-L/Z_2$	0	$(AL-D)/C$
3	0	$-L$	0	$-L/Z_3$	$(BL-D)/C$
4	L	0	L/Z_4	0	$-(AL+D)/C$

3.2.2 Model considering laser-cross misalignment

In this case, we are interested in calculating the model parameters when the laser-cross is not perfectly aligned with the camera frame and not perfectly centered in the camera origin. Such a misalignment is represented in Fig. 5 and it can be modelled according to a frame transformation matrix ${}^{C}\mathbf{M}_{L}$ which passes from points expressed in the laser-cross frame to the camera frame. The model parameters under these conditions are developed in Appendix B.

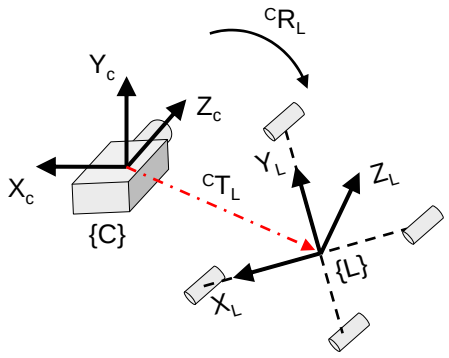

Figure 5: Model of misalignment of the laser-cross

4 Task function and control stability

This section reviews the typical formalism used to define a robotic task in visual servoing. This formalism allows us to construct a simple proportional control law and to derive its stability conditions. All this theoretical background will be exploited during the rest of the work in order to design the control law and to derive the stability analysis of different sets of visual features.

We remember that the goal of our task is to bring the camera to a position where it is parallel to the object. This task corresponds to fixing a plane-to-plane virtual link between the camera image plane and the planar object. Such a virtual link belongs to the class $N = 3$ since this is the number of degrees of freedom constrained by the link [26]. Concretely, 2 translational and 1 rotational degrees of freedom are constrained. This can be seen by stacking the interaction matrices of at least three projected points and evaluating it for $\underline{\mathbf{n}} = (0, 0, 1)$

$$\mathbf{L_x} = \begin{pmatrix} 0 & 0 & X_{01}/Z^2 & y_1 X_{01}/Z & -x_1 X_{01}/Z_1 & 0 \\ 0 & 0 & Y_{01}/Z^2 & y_1 Y_{01}/Z & -x_1 Y_{01}/Z_1 & 0 \\ 0 & 0 & X_{02}/Z^2 & y_2 X_{02}/Z & -x_2 X_{02}/Z_2 & 0 \\ 0 & 0 & Y_{02}/Z^2 & y_2 Y_{02}/Z & -x_2 Y_{02}/Z_2 & 0 \\ 0 & 0 & X_{03}/Z^2 & y_3 X_{03}/Z & -x_3 X_{03}/Z_3 & 0 \\ 0 & 0 & Y_{03}/Z^2 & y_3 Y_{03}/Z & -x_3 Y_{03}/Z_3 & 0 \end{pmatrix} \tag{23}$$

The rank of the above matrix is 3 if the points are not collinear. This means that there are three types of camera motion, namely V_x, V_y and Ω_z, which will produce no changes in the image. For the case of a general relative pose camera-object it can also be seen by expressing the above matrix in a frame attached to the object as explained in Appendix C. The interaction matrix expressed in the object frame has the following form (see the appendix

for the details)

$$ {}^{o}\mathbf{L}_{\mathbf{x}} = {}^{c}\mathbf{L}_{\mathbf{x}} \cdot {}^{c}\mathbf{T}_{o} = \begin{pmatrix} 0 & 0 & -X_{01}/(\Pi_{01}Z_1) & X_{01}\eta_1/\Pi_{01} & X_{01}\xi_1/\Pi_{01} & 0 \\ 0 & 0 & -Y_{01}/(\Pi_{01}Z_1) & Y_{01}\eta_1/\Pi_{01} & Y_{01}\xi_1/\Pi_{01} & 0 \\ 0 & 0 & -X_{02}/(\Pi_{02}Z_2) & X_{02}\eta_2/\Pi_{02} & X_{02}\xi_2/\Pi_{02} & 0 \\ 0 & 0 & -Y_{02}/(\Pi_{02}Z_2) & Y_{02}\eta_2/\Pi_{02} & Y_{02}\xi_2/\Pi_{02} & 0 \\ 0 & 0 & -X_{03}/(\Pi_{03}Z_3) & X_{03}\eta_3/\Pi_{03} & X_{03}\xi_3/\Pi_{03} & 0 \\ 0 & 0 & -Y_{03}/(\Pi_{03}Z_3) & Y_{03}\eta_3/\Pi_{03} & Y_{03}\xi_3/\Pi_{03} & 0 \end{pmatrix} \tag{24} $$

with

$$ \begin{aligned} \eta_i &= \frac{1-A^2}{C} y_i + \frac{A(Bx_i + ACy_i)}{C(1+C)} \\ \xi_i &= \frac{1-B^2}{C} x_i + \frac{B(Ay_i + BCx_i)}{C(1+C)} \end{aligned} $$

The rank of ${}^{o}\mathbf{L}_{\mathbf{x}}$ is 3 and the kernel is generated by the following base

$$ \{(1,0,0,0,0,0)\,,\ (0,1,0,0,0,0)\,,\ (0,0,0,0,0,1)\} \tag{25} $$

which clearly corresponds to a plane-to-plane virtual link. As can be seen, for any relative pose camera-object there are three degrees of freedom of the planar object which cannot be perceived by the camera.

In the following section the formulation of the task function is briefly reviewed. Afterwards, a simple control law, and how to analyze its stability, is presented.

4.1 Task function

A robotic task can be described by a function which must be regulated to 0. The task function $\mathbf{e}$ is defined as a $m-$dimensional vector of the form [26]

$$ \mathbf{e} = \mathbf{C}(\mathbf{s} - \mathbf{s}^*) \tag{26} $$

where $\mathbf{s}$ is a $k \times 1$ vector containing k visual features corresponding to the current state, while $\mathbf{s}^*$ denotes the visual features values in the desired state. $\mathbf{C}$ is a $m \times k$ combination matrix that must be of full rank $m \leq k$ in order to produce the m independent components of $\mathbf{e}$. The aim of visual servoing is to regulate the task function $\mathbf{e}$ to $\mathbf{0}$ so that $\mathbf{s} - \mathbf{s}^* = \mathbf{0}$.

The task $\mathbf{e}$ controls m degrees of freedom from a total of n. When $m < n$ it means that a virtual link of class N is fulfilled so that $m = n - N \leq k$. In our case $m = N = 3$.

A suitable choice of the combination matrix $\mathbf{C}$ is [26]

$$ \mathbf{C} = \mathbf{W}\widehat{\mathbf{L}_{\mathbf{s}}}^{+} \tag{27} $$

where $\widehat{\mathbf{L}_{\mathbf{s}}}^{+}$ is the pseudoinverse of a model of the interaction matrix and $\mathbf{W}$ is an $m \times 6$ matrix of full rank m having the same kernel that $\mathbf{L}_{\mathbf{s}}$. The choice of $\mathbf{W}$ depends on the number of visual features k. Since in our case $m < n$ there are two cases that can be considered:

- $\text{rank}(\mathbf{L_s}) = m = k \Rightarrow \mathbf{W} = \widehat{\mathbf{L_s}}$. In this particular case $\mathbf{C} = \mathbf{I}_m$.
- $\text{rank}(\mathbf{L_s}) = m < k \Rightarrow$ the rows of $\mathbf{W}$ are the m vectors forming the base of the row space generated by $\widehat{\mathbf{L_s}}$.

4.2 Control law

A simple control law can be defined in order to fulfill the task $\mathbf{e}$. We assume that the combination matrix $\mathbf{C}$ is constant so that the derivative of the task function (26) is

$$\dot{\mathbf{e}} = \mathbf{C}\dot{\mathbf{s}} \tag{28}$$

and taking into account that

$$\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v} \tag{29}$$

we have that

$$\dot{\mathbf{e}} = \mathbf{C}\mathbf{L_s}\mathbf{v} \tag{30}$$

then, by imposing an exponential decrease of the task function $\dot{\mathbf{e}} = -\lambda\mathbf{e}$ (being λ a positive gain) we find

$$-\lambda\mathbf{e} = \mathbf{C}\mathbf{L_s}\mathbf{v} \tag{31}$$

from this expression a proportional control law can be built by using a model of the interaction matrix $\widehat{\mathbf{L_s}}$

$$\mathbf{v} = -\lambda(\mathbf{C}\widehat{\mathbf{L_s}})^{+}\mathbf{e} \tag{32}$$

which is equal to

$$\mathbf{v} = -\lambda(\mathbf{C}\widehat{\mathbf{L_s}})^{+}\mathbf{C}(\mathbf{s} - \mathbf{s}^*) \tag{33}$$

If it is not possible to estimate all the parameters of the interaction matrix at each iteration, a typical choice is to set $\widehat{\mathbf{L_s}}$ as the interaction matrix evaluated at the desired position noted as $\mathbf{L}_\mathbf{s}^*$ or $\mathbf{L_s}(\mathbf{e}^*)$.

4.3 Stability analysis

One interesting aim in visual servoing is concerned on studying whether the control law is able to regulate the task function to the desired state or not. In other words, the aim is to study if the desired state $\mathbf{e}^* = \mathbf{0}$ is an stable equilibrium point which is reached when time approaches infinity. Let us remember the following basic definitions:

Equilibrium point: $\mathbf{e}^* = \mathbf{0}$ is said to be an equilibrium point if $\dot{\mathbf{e}} = \mathbf{0}$ when $\mathbf{e} = \mathbf{0}$.

Stability: the stability of an equilibrium point is classically defined in the Lyapunov sense. The equilibrium point in the origin is said to be stable if

$$\forall\epsilon > 0\ \exists\delta > 0\ s.t.\ \|\mathbf{e}(0)\| < \delta \Rightarrow \|\mathbf{e}(t)\| < \epsilon,\ \forall t \tag{34}$$

Asymptotic stability: the equilibrium point $\mathbf{e}^* = \mathbf{0}$ is asymptotically stable if it is stable and if it is attracting so that

$$\lim_{t\to\infty} \mathbf{e}(t) = \mathbf{e}^* = \mathbf{0} \tag{35}$$

Hereafter we will focus only on the asymptotic stability since it ensures that the equilibrium is reached. As explained in [40], the stability analysis of the control law allows us to determine whether $\mathbf{e}^* = 0$ is reached from any starting point (global asymptotic stability) or only when the initial state is nearby the equilibrium (local asymptotic stability).

In order to fulfill the stability analysis, it is necessary to obtain the closed-loop equation of the system, which is obtained by plugging the control law (33) into (30)

$$\dot{\mathbf{e}} = -\lambda \mathbf{C}\mathbf{L}_\mathbf{s}(\mathbf{C}\widehat{\mathbf{L}_\mathbf{s}})^+\mathbf{e} \tag{36}$$

which will be noted hereafter as

$$\dot{\mathbf{e}} = -\lambda \mathbf{M}(\mathbf{e})\mathbf{e} \tag{37}$$

Note that $\mathbf{L}_\mathbf{s}$ is the actual interaction matrix in a certain instant of time t, while $\widehat{\mathbf{L}_\mathbf{s}}$ is the value of the model used in the control law. Therefore, since $\mathbf{L}_\mathbf{s}$ depends on the state $\mathbf{e}$ so does $\mathbf{M}$. If the explicit expression of $\mathbf{e}$ in function of time can be obtained by solving the differential equation (37), then it can be checked if the task function zeroes when time approaches infinity. However, in most cases it is not possible to obtain such explicit solution. Alternatively, necessary and sufficient conditions for the local asymptotic stability, and sufficient conditions for the global asymptotic stability are hereafter recalled.

It is well known that the system in (37) is *locally asymptotically stable* if and only if the eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ (which represents the value of $\mathbf{M}$ evaluated in the desired state) have all positive real part.

When analyzing the global asymptotic stability a necessary condition is that $\mathbf{e}^* = \mathbf{0}$ is the only equilibrium in the task space, which can be ensured if and only if $\det(\mathbf{M}(\mathbf{e})) \neq 0$. If this is true, then a sufficient condition for the system to be *globally asymptotically stable* is that $\mathbf{M}(\mathbf{e})$ is a positive definite matrix.

5 Object plane parameters approach

The first visual servoing approach that we present is a pure position-based method. Indeed, we can use the triangulation capabilities of the system composed by the camera and the lasers in order to reconstruct up to 4 points of the object so that its pose can be recovered. In such case, the $3D$ parameters of the reconstructed plane can be directly used in the closed-loop of the control scheme so that a position-based approach is performed.

Let us consider that the four parameters of the planar object A, B, C and D can be precisely estimated at each iteration. The feature vector $\mathbf{s}$ could be built up by using these four $3D$ parameters. However, since the number of controlled dof is 3, a feature vector of the same dimension is going to be defined. The equation of the object can be noted as a

relationship between the depth of a point and its normalized image coordinates as follows

$$\frac{1}{Z} = P_1 + P_2 y + P_3 x \tag{38}$$

with $P_1 = -C/D$, $P_2 = -B/D$ and $P_3 = -A/D$. By using the time derivatives of $\underline{\mathbf{n}}$ and D in (12) the interaction matrix of $\mathbf{s} = (P_1,\ P_2,\ P_3)$ is calculated obtaining

$$\mathbf{L_s} = \begin{pmatrix} P_1P_3 & P_1P_2 & P_1^2 & -P_2 & P_3 & 0 \\ P_2P_3 & P_2^2 & P_1P_2 & P_1 & 0 & -P_3 \\ P_3^2 & P_2P_3 & P_1P_3 & 0 & -P_1 & P_2 \end{pmatrix} \tag{39}$$

which in the desired state it has the following value

$$\mathbf{L_s^*} = \begin{pmatrix} 0 & 0 & 1/Z^{*2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/Z^* & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 - 1/Z^* & 0 \end{pmatrix} \tag{40}$$

On the other hand, the depth of the points belonging to the planar object can be also expressed as

$$Z = \gamma + \beta Y + \alpha X \tag{41}$$

where $\gamma = -D/C$, $\beta = -B/C$ and $\alpha = A/C$. In this case, the interaction matrix of the parameters $\mathbf{s} = (\gamma,\ \beta,\ \alpha)$ is

$$\mathbf{L_s} = \begin{pmatrix} \alpha & \beta & -1 & -\gamma\beta & \gamma\alpha & 0 \\ 0 & 0 & 0 & -1-\beta^2 & \beta\alpha & -\alpha \\ 0 & 0 & 0 & -\beta\alpha & 1+\alpha^2 & \beta \end{pmatrix} \tag{42}$$

Note that in this case the level of decoupling between the $3D$ features is higher. Furthermore, if we look at the interaction matrix in the desired state

$$\mathbf{L_s^*} = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{43}$$

we can see that it does not depend on the depth as in the case of $\mathbf{s} = (P_1,\ P_2,\ P_3)$ shown in (40). Therefore, in this case the dynamics of the object parameters around the desired state vary linearly with respect to the camera motion. That is why we prefer to use the object plane representation based on $\mathbf{s} = (\gamma,\ \beta,\ \alpha)$. Since the dimension of $\mathbf{s}$ is 3 the control law is

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_s}}^{+} (\mathbf{s} - \mathbf{s}^*) \tag{44}$$

In order to estimate the object plane parameters it is necessary to reconstruct the four $3D$ points $\mathbf{X}$ projected by the lasers. Then, the equation of the plane best fitting the four points can be calculated by means of least squares. First of all, it is necessary to calculate the $3D$ point coordinates of every projected laser. The simplest way is to triangulate the

points by using the corresponding image normalized coordinates and the laser orientation $\mathbf{u}$ and the laser origin $\mathbf{X}_0$. Nevertheless, it is possible to reduce the number of parameters concerning the laser calibration by using the information provided by the desired image.

Remember that the $3D$ point $\mathbf{X}$ projected by a certain laser of orientation $\underline{\mathbf{u}}$ and origin $\mathbf{X}_0$ must accomplish the following relationship

$$\mathbf{X} = \mathbf{x}Z = \mu\underline{\mathbf{u}} + \mathbf{X}_0 \tag{45}$$

Given an image point $\mathbf{x}^*$ from the desired image the following relationships are extracted from the above equation

$$\left\{ \begin{array}{rcl} x^*Z^* & = & \mu^* u_x + X_0 \\ y^*Z^* & = & \mu^* u_y + Y_0 \\ Z^* & = & \mu^* u_z \end{array} \right. \tag{46}$$

Then, from the last equation we have that $\mu^* = Z^*/u_z$ so that plugging it onto the others we get

$$\begin{array}{rcl} x^*Z^* & = & Z^* u_x/u_z + X_0 \\ y^*Z^* & = & Z^* u_y/u_z + Y_0 \end{array} \tag{47}$$

so that the origin of the laser X_0 can be expressed as follows

$$\begin{array}{rcl} X_0 & = & Z^*(x^* - u_{xz}) \\ Y_0 & = & Z^*(y^* - u_{yz}) \end{array} \tag{48}$$

where $u_{xz} = u_x/u_z$ and $u_{yz} = u_y/u_z$.

By using the above definitions, the equations in (47) can be written for the current image as

$$\begin{array}{rcl} xZ & = & Zu_{xz} + Z^*(x^* - u_{xz}) \\ yZ & = & Zu_{yz} + Z^*(y^* - u_{yz}) \end{array} \tag{49}$$

From (41) the depth is related to the object parameters as

$$Z = \frac{\gamma}{1 - \alpha x - \beta y} \tag{50}$$

so that the equations in (49) can be expressed in terms of the object parameters as follows

$$\begin{array}{rcl} \alpha x Z^*(x^* - u_{xz}) + \beta y Z^*(x^* - u_{xz}) + \gamma(x - u_{xz}) - Z^*(x^* - u_{xz}) & = & 0 \\ \alpha x Z^*(y^* - u_{yz}) + \beta y Z^*(y^* - u_{yz}) + \gamma(y - u_{yz}) - Z^*(y^* - u_{yz}) & = & 0 \end{array} \tag{51}$$

Then, using these equations for every one of the four laser pointers the following system of non-linear equations is obtained

$$\left\{ \begin{array}{rcl} \alpha x_1 Z^*(x_1^* - u_{xz}) + \beta y_1 Z^*(x_1^* - u_{xz}) + \gamma(x_1 - u_{xz}) - Z^*(x_1^* - u_{xz}) & = & 0 \\ \alpha x_1 Z^*(y_1^* - u_{yz}) + \beta y_1 Z^*(y_1^* - u_{yz}) + \gamma(y_1 - u_{yz}) - Z^*(y_1^* - u_{yz}) & = & 0 \\ \vdots & = & \vdots \\ \alpha x_4 Z^*(x_4^* - u_{xz}) + \beta y_4 Z^*(x_4^* - u_{xz}) + \gamma(x_4 - u_{xz}) - Z^*(x_4^* - u_{xz}) & = & 0 \\ \alpha x_4 Z^*(y_4^* - u_{yz}) + \beta y_4 Z^*(y_4^* - u_{yz}) + \gamma(y_4 - u_{yz}) - Z^*(y_4^* - u_{yz}) & = & 0 \end{array} \right. \tag{52}$$

Note that there are 8 equations for 5 unknowns which are

$$\xi = (\alpha, \beta, \gamma,\ u_{xz},\ u_{yz}) \tag{53}$$

Therefore, all the four lasers are assumed to have the same orientation. The system can be numerically solved by a minimization algorithm based on non-linear least squares. Nevertheless, it cannot be analytically ensured that the algorithm always converges to the right solution. Under calibration errors and image noise, it is possible to reach local minima. Therefore, demonstrating analytically the global asymptotic stability of this position-based approach seems out of reach.

In the following subsection simulations using this position-based approach are presented.

5.1 Simulation results

The simulations have been performed by taking into account a sampling time of $\Delta t = 40$ ms and the camera intrinsic parameters obtained from the experimental setup (see Section 9). The laser-cross has been simulated using $L = 15$ cm according to the real experimental setup. The desired position has been chosen so that the camera is parallel to the plane at $Z^* = 60$ cm. The initial position the camera is at a distance of 105 cm from the plane and the relative orientation camera-object is defined by $\alpha_x = -30^\circ$ and $\alpha_y = 15^\circ$ according to the specification given in Appendix A. The gain λ has been set to 0.12.

5.1.1 Ideal system

A first simulation has been done by taking into account a perfect alignment of the laser-cross with the camera frame. Furthermore, it has been assumed that the camera intrinsic parameters are perfectly known and all the lasers have the same direction (which coincides in this case with the optical axis direction). The initial and desired image simulated under these conditions are shown in Fig. 6. As can be seen, the epipolar lines of the lasers $1-3$ and $2-4$ are perfectly orthogonal and intersect in the central point of the image.

non-constant control law: Fig. 7 shows the results when $\widehat{\mathbf{L_s}}$ is estimated at each iteration by using the reconstructed object plane parameters. Fig. 7c shows the coordinates of a fixed point expressed in the camera frame along the simulation. The fixed point has been set as the initial position of the camera origin. Note that the camera trajectory is almost a straight line in the cartesian space as can also be observed in Fig. 7d. This is possible since the object pose is perfectly reconstructed under the ideal conditions. Furthermore, the task function has a pure exponential decrease since $\mathbf{L_s} = \mathbf{L_s}$ and the closed-loop equation of the system becomes

$$\dot{\mathbf{e}} = -\lambda \mathbf{e} \tag{54}$$

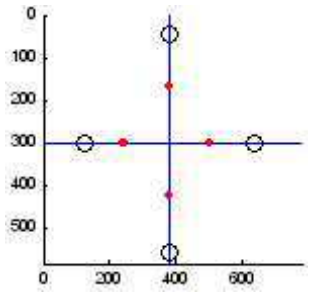

Figure 6: Simulation of the ideal system. The initial point distribution is shown with the red dots. The desired point distribution is depicted by the circles. The epipolar lines are painted in blue.

constant control law: the results when using the constant control law based on $\widehat{\mathbf{L}_\mathbf{s}} = \mathbf{L}_\mathbf{s}^*$ in (43) are plotted in Fig. 8. As can be seen, even if the camera trajectory is no longer almost a straight line, the lateral displacements of the camera are quite small. On the other hand, both the task function components and the camera velocities are strictly monotonic thanks to the linear link existing between them near the desired position (as can be seen in the form of $\mathbf{L}_\mathbf{s}^*$).

5.1.2 System including laser-cross misalignment and image noise

A second simulation including calibration errors and image noise has been performed. First, the laser-cross has been displaced from the camera origin according to the translation vector $(4, 10, 9)$ cm. Then, the laser-cross has been rotated $12°$ about its Z axis, $9°$ about Y and $-15°$ about X. The rest of model assumptions still fit (all the lasers have the same relative direction and perfect camera calibration). However, random gaussian noise with standard deviation of 0.5 pixels has been added to the images at each iteration.

The initial position of the camera is still at 105 cm from the object but their relative orientation is defined by $\alpha_x = -25°$ and $\alpha_y = 15°$. The initial and desired image are shown in Fig. 9. Note that the large misalignment of the laser-cross is evident in these images. However, note that all the epipolar lines intersect in a unique image point. This only happens when all the laser pointers have the same direction.

As can be seen in Fig. 10 and Fig. 11 the behavior of both control laws when using $\mathbf{s} = (\gamma,\ \beta,\ \alpha)$ is robust against large misalignment of the laser-cross. The image noise mainly affects the components of the task function e_2 and e_3 while e_1 remains almost insensitive to it.

Remark: the success of the position-based approach in front of large calibration errors relies on the iterative minimization of the non-linear system of equations which leads to a

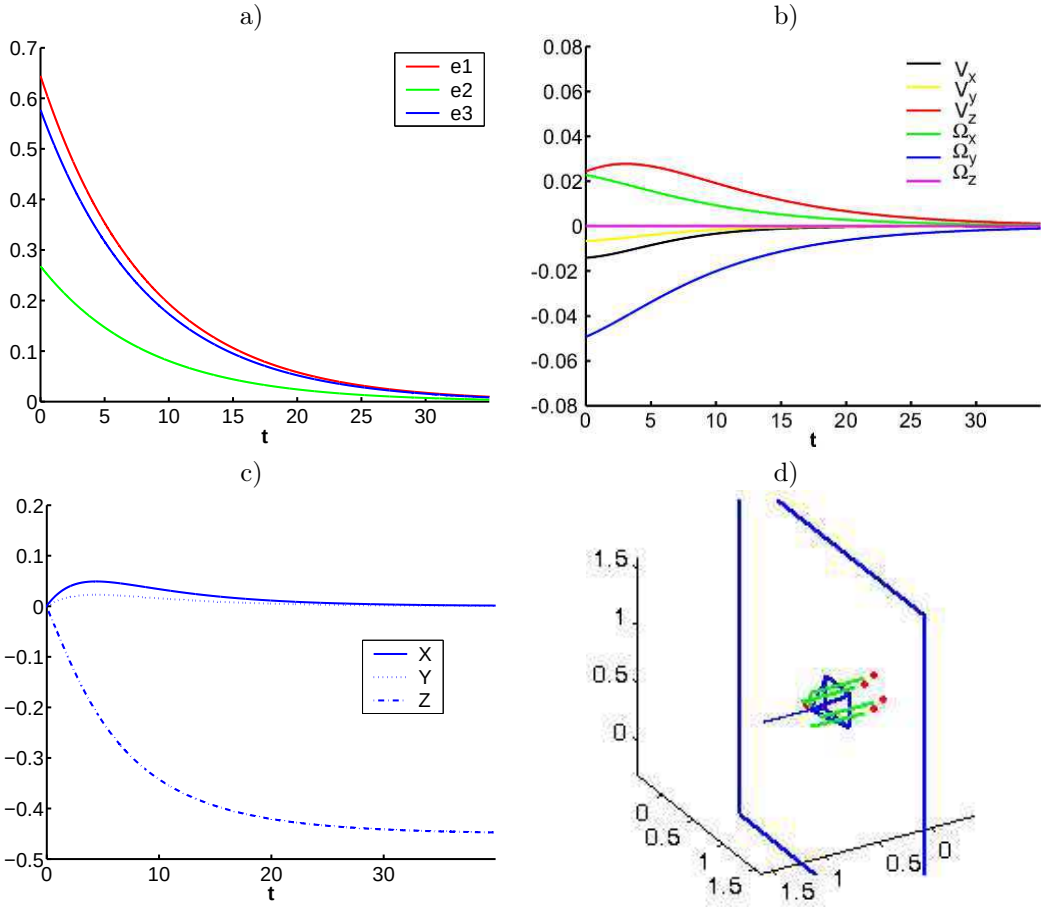

Figure 7: Ideal system: simulation using $\mathbf{s} = (\gamma, \beta, \alpha)$ and the non-constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

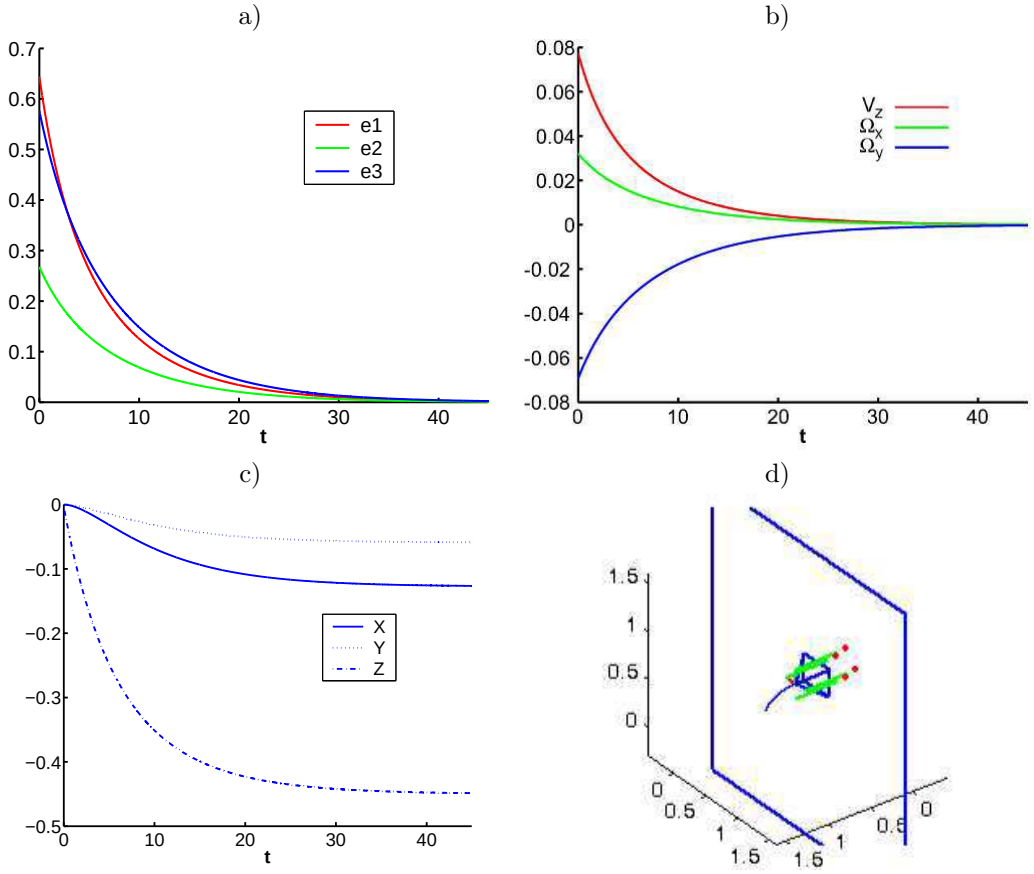

Figure 8: Ideal system: simulation using $\mathbf{s} = (\gamma, \beta, \alpha)$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

robust depth estimation of the four projected laser points. During this simulation we have detected certain sensitivity of the numeric algorithm in front of image noise. Therefore, a robust algorithm of minimization must be used.

6 Image points approach

The simplest $2D$ visual servoing approach that can be defined consists of using the image coordinates of the four projected points. According to the ideal model, if the laser-cross is

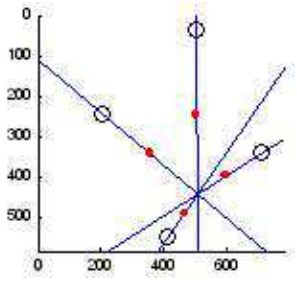

Figure 9: Simulation of the system including large laser-cross misalignment. The initial point distribution is shown with the red dots. The desired point distribution is depicted by the circles. The epipolar lines are painted in blue.

aligned with the camera frame, the coordinates x_1, y_2, x_3 and y_4 of the four points remain always to 0. Therefore, we can chose as visual features the following vector

$$\mathbf{s} = (y_1,\ x_2,\ y_3,\ x_4) \tag{55}$$

Since the number of visual features $k = 4$ is greater than the number of degrees of freedom that must be controlled ($m = 3$), matrix $\mathbf{W}$ is chosen so that its rows are the basis of the row space generated by $\widehat{\mathbf{L_s}}$

$$\mathbf{W} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{56}$$

Then, setting $\mathbf{C} = \mathbf{W}\widehat{\mathbf{L_s}}^{+}$, the control law in (33) becomes

$$\mathbf{v} = -\lambda \left(\mathbf{W}\widehat{\mathbf{L_s}}^{+}\widehat{\mathbf{L_s}}\right)^{+} \mathbf{W}\widehat{\mathbf{L_s}}^{+} (\mathbf{s} - \mathbf{s}^*) \tag{57}$$

and the closed-loop equation of the system in (36) when using image points is

$$\dot{\mathbf{e}} = -\lambda \mathbf{W}\widehat{\mathbf{L_s}}^{+}\mathbf{L_s} \left(\mathbf{W}\widehat{\mathbf{L_s}}^{+}\widehat{\mathbf{L_s}}\right)^{+} \mathbf{e} \tag{58}$$

A constant interaction matrix is used in the control law, which is obtained by using the parameters of the ideal model presented in Table 1 (see Section 3.2.1), evaluated in the desired position where $\underline{\mathbf{n}} = (0, 0, 1)$ and $D = -Z^*$.

$$\widehat{\mathbf{L_s}} = \mathbf{L_s^*} = \begin{pmatrix} 0 & 0 & L/Z^{*2} & L^2/Z^{*2} & 0 & 0 \\ 0 & 0 & -L/Z^{*2} & 0 & -L^2/Z^{*2} & 0 \\ 0 & 0 & -L/Z^{*2} & L^2/Z^{*2} & 0 & 0 \\ 0 & 0 & L/Z^{*2} & 0 & -L^2/Z^{*2} & 0 \end{pmatrix} \tag{59}$$

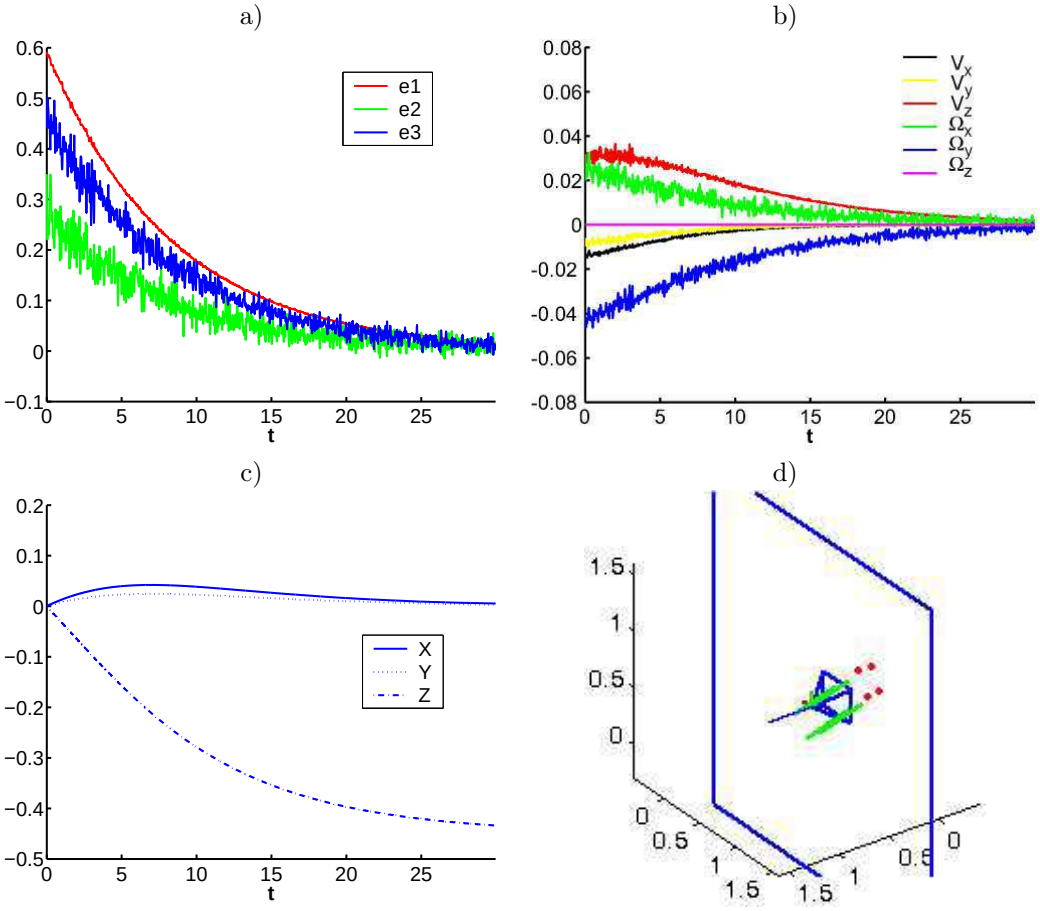

Figure 10: System including large laser-cross misalignment and image noise: simulation using $\mathbf{s} = (\gamma, \beta, \alpha)$ and the non-constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

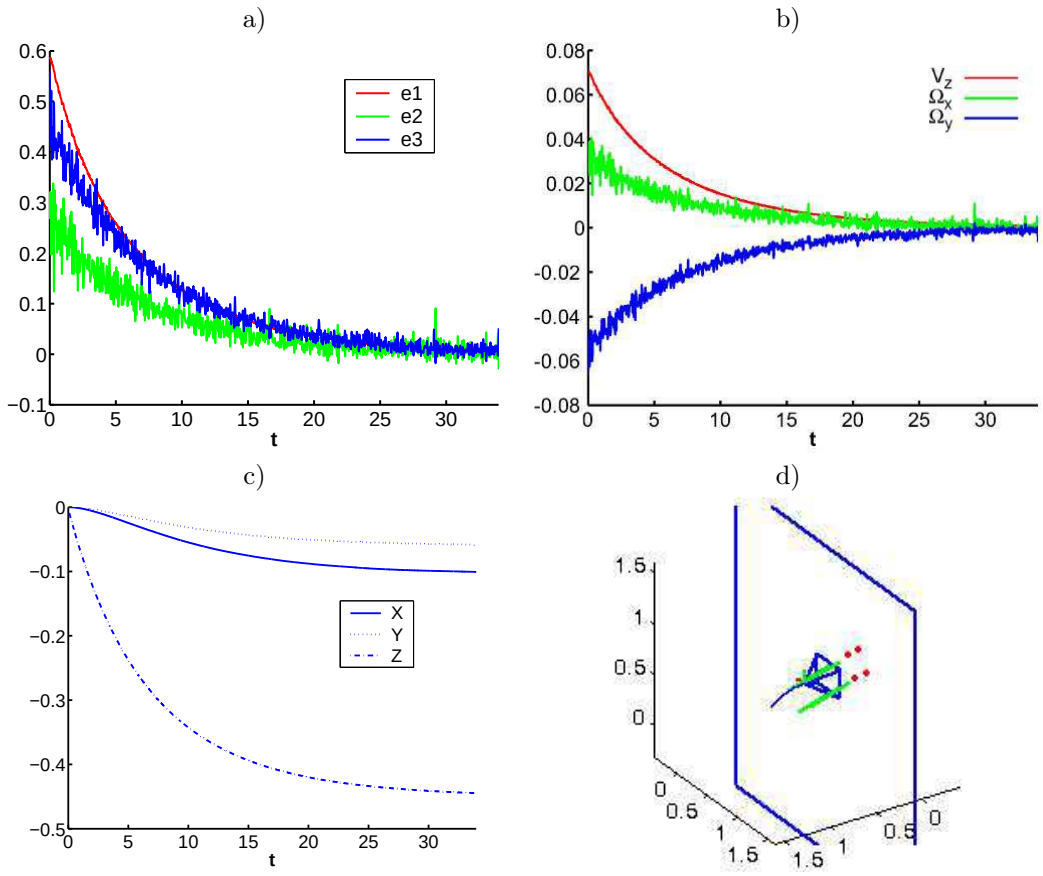

Figure 11: System including large laser-cross misalignment and image noise: simulation using $\mathbf{s} = (\gamma, \beta, \alpha)$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

In the following subsections three studies of stability are faced. First, we show that the global asymptotic stability of the ideal model cannot be proved. Afterwards, the local asymptotic stability of the system is analyzed taking into account certain types of laser-cross misalignment. Finally, the local asymptotic stability in front of errors in the camera intrinsic parameters is also studied.

6.1 Global asymptotic stability under perfect conditions

The general expression of $\mathbf{L_s}$ is obtained by using the ideal model parameters in Table 1 (see Section 3.2.1) which are expressed in function of the object parameters. The obtained matrix is

$$\mathbf{L_s}=\begin{pmatrix} \frac{LAC}{(BL+D)^2} & \frac{LBC}{(BL+D)^2} & \frac{LC^2}{(BL+D)^2} & \frac{L(B^2L+BD+LC^2)}{(BL+D)^2} & -\frac{LA}{BL+D} & -\frac{L^2AC}{(BL+D)^2} \\ -\frac{LAC}{(AL-D)^2} & -\frac{LBC}{(AL-D)^2} & -\frac{LC^2}{(AL-D)^2} & \frac{BL}{AL-D} & -\frac{L(LA^2+LC^2-AD)}{(AL-D)^2} & \frac{L^2BC}{(AL-D)^2} \\ -\frac{LAC}{(D-BL)^2} & -\frac{LBC}{(D-BL)^2} & -\frac{LC^2}{(D-BL)^2} & -\frac{L(BD-LB^2-LC^2)}{(D-BL)^2} & \frac{LA}{D-BL} & -\frac{L^2AC}{(D-BL)^2} \\ \frac{LAC}{(AL+D)^2} & \frac{LBC}{(AL+D)^2} & \frac{LC^2}{(AL+D)^2} & \frac{BL}{AL+D} & -\frac{L(LA^2+LC^2+AD)}{(AL+D)^2} & \frac{L^2BC}{(AL+D)^2} \end{pmatrix} \tag{60}$$

A sufficient condition for the system to be globally asymptotically stable is that the product of matrices $\mathbf{M}$ in the closed-loop equation is positive definite. The positiveness of $\mathbf{M}$ is ensured if all the eigenvalues of its symmetric part $\mathbf{S}$ are positive.

When using the constant control law based on $\mathbf{L_s^*}$ the analytic expression of the eigenvalues are too complex. On the other hand, if a non-constant law based on estimating $\mathbf{L_s}$ at each iteration is used (reconstructing the object pose parameters involved by triangulation), $\mathbf{M}$ is the identity so that the global asymptotic stability of the ideal model is ensured for any initial object pose.

In presence of calibration errors, the global asymptotic stability analysis becomes too complex. That is why we use instead the local asymptotic stability analysis.

6.2 Local asymptotic stability analysis under laser-cross misalignment

The local asymptotic stability analysis is based on studying the real part of the eigenvalues of the matrices product appearing in the closed-loop equation (58) evaluated in the desired state

$$\mathbf{M}(\mathbf{e}^*) = \mathbf{W}\widehat{\mathbf{L_s}}^+\mathbf{L_s}(\mathbf{e}^*)\left(\mathbf{W}\widehat{\mathbf{L_s}}^+\widehat{\mathbf{L_s}}\right)^+ \tag{61}$$

where $\mathbf{L_s}(\mathbf{e}^*)$ is the real interaction matrix at the desired state which takes into account the actual pose of the laser-cross. We take the model parameters presented in Appendix B in order to obtain such an interaction matrix. The study of stability when using an estimation of $\widehat{\mathbf{L_s}}$ at each iteration becomes too complex. The case of using the constant control law based on $\mathbf{L_s^*}$ is considered in the following sections.

6.2.1 Misalignment consisting of a translation

In this case we assume that the cross-laser frame has the same orientation as the camera frame, but that its center has been displaced according to the vector ${}^C\mathbf{T}_L = (t_x,\ t_y,\ t_z)$. In this case the parameters in Table 3 are evaluated for $A = B = 0$, $C = 1$ and $D = -Z^*$ in

order to obtain the expression of $\mathbf{L_s}(\mathbf{e}^*)$ which is

$$\mathbf{L_s}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & \frac{L+t_y}{Z^{*2}} & \frac{(L+t_y)^2}{Z^{*2}} & -\frac{(L+t_y)t_x}{Z^{*2}} & 0 \\ 0 & 0 & -\frac{L-t_x}{Z^{*2}} & -\frac{t_y(L-t_x)}{Z^{*2}} & -\frac{(L-t_x)^2}{Z^{*2}} & 0 \\ 0 & 0 & -\frac{L-t_y}{Z^{*2}} & \frac{(L-t_y)^2}{Z^{*2}} & \frac{(L-t_y)t_x}{Z^{*2}} & 0 \\ 0 & 0 & \frac{L+t_x}{Z^{*2}} & \frac{t_y(L+t_x)}{Z^{*2}} & -\frac{(L+t_x)^2}{Z^{*2}} & 0 \end{pmatrix} \tag{62}$$

The system is locally asymptotically stable if and only if the eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ have all positive real part. The eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are

$$\begin{aligned} \sigma_1 &= 1 \\ \sigma_2 &= \frac{t_x^2 + t_y^2 + 2L^2 + \sqrt{(t_x^2+t_y^2)^2 + 6L^2(t_x^2+t_y^2)}}{2L^2} \\ \sigma_3 &= \frac{t_x^2 + t_y^2 + 2L^2 - \sqrt{(t_x^2+t_y^2)^2 + 6L^2(t_x^2+t_y^2)}}{2L^2} \end{aligned}$$

Imposing their positivity the following constraint arises

$$t_x^2 + t_y^2 < 2L^2 \tag{63}$$

which means that the local asymptotic stability is only ensured when the projection of the laser-cross center into the camera plane $Z = 0$ is included in the circle of radius $\sqrt{2}L$ centered in the camera origin (see Fig. 12 for a schema). Note that the component t_z of the misalignment does not affect the local asymptotic stability. Therefore, a displacement of the laser-cross from the camera origin can strongly affect the global asymptotic stability of the system when using image points since even the local asymptotic stability is constrained.

6.2.2 Misalignment consisting of individual rotations

Let now test the stability of the system when the laser-cross is centered in the camera frame, but rotated with respect to one of the axis. If the three rotations are considered at the same time, too many parameters appear and no analytical results arise.

In the first case, the laser-cross is rotated an angle ψ around the X axis of the camera frame. The interaction matrix $\mathbf{L_s}(\mathbf{e}^*)$ is calculated according to the model parameters in Table 4. The eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are

$$\begin{aligned} \sigma_1 &= 1 \\ \sigma_2 &= \frac{1+\cos\psi}{2\cos\psi} \\ \sigma_3 &= \frac{1}{\cos^2\psi} \end{aligned}$$

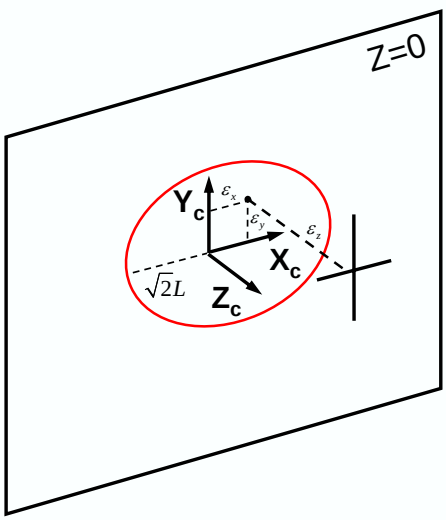

Figure 12: Region of local asymptotic stability for the projection of the laser-cross onto the image plane

it is easy to see that all the eigenvalues are positive if the rotation angle ψ is expressed in the interval $(-\pi/2,\ \pi/2)$. Note that there is a singularity for $\psi = -\pi/2$ and $\psi = \pi/2$, since in those configurations the lasers do not intersect the object and therefore the servoing is not possible.

When a rotation θ is done around the Y axis the same eigenvalues are obtained. Finally, the eigenvalues corresponding to the case of a rotation ϕ around the Z axis are

$$\begin{aligned} \sigma_1 &= \cos\phi \\ \sigma_2 &= \cos^2\phi + \sqrt{\cos^2\phi(\cos^2\phi - 1)} \\ \sigma_3 &= \cos^2\phi - \sqrt{\cos^2\phi(\cos^2\phi - 1)} \end{aligned}$$

The positivity of the first eigenvalue imposes that $\phi \in [-\pi/2,\ \pi/2]$. In the second and the third eigenvalue, the square root is always of a negative number, so that the real part of both eigenvalues is $\cos^2\phi$, which is always positive.

In conclusion, the approach based on image points is locally asymptotic stable with respect to individual rotations of the laser-cross around the principal axis of the camera frame.

6.3 Local asymptotic stability analysis in presence of camera calibration errors

This section presents the local asymptotic stability analysis of the system when the laser-cross is perfectly aligned with the camera frame but the calibrated intrinsic parameters of the camera are not the real ones.

We model the intrinsic parameters of the camera according to the following matrix

$$\mathbf{A} = \begin{pmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{64}$$

where (u_0, v_0) is the principal point in pixels, f the focal distance (in meters), and (k_u, k_v) the conversion factors from meters to pixels for the horizontal and vertical camera axis, respectively. This matrix expresses how the normalized coordinates $\mathbf{x} = \mathbf{X}/Z$ of a $3D$ point projects onto a certain pixel $\mathbf{x}_p$ of the image as follows

$$\mathbf{x}_p = \mathbf{A}\mathbf{x} \tag{65}$$

When only an estimation $\widetilde{\mathbf{A}}$ of the real intrinsic parameters is available, an estimation $\widetilde{\mathbf{x}}$ of the real normalized coordinates $\mathbf{x}$ is obtained from the pixel coordinates

$$\widetilde{\mathbf{x}} = \widetilde{\mathbf{A}}^{-1}\mathbf{x}_p \tag{66}$$

This estimation is related to the real normalized coordinates by

$$\widetilde{\mathbf{x}} = \widetilde{\mathbf{A}}^{-1}\mathbf{A}\mathbf{x} \tag{67}$$

Hereafter, the elements of $\widetilde{\mathbf{A}}^{-1}\mathbf{A}$ will be noted as follows

$$\widehat{\mathbf{A}}^{-1}\mathbf{A} = \begin{pmatrix} \frac{fk_u}{\widetilde{fk}_u} & 0 & \frac{u_0 - \widetilde{u}_0}{\widetilde{fk}_u} \\ 0 & \frac{fk_v}{\widetilde{fk}_v} & \frac{v_0 - \widetilde{v}_0}{\widetilde{fk}_v} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} K_u & 0 & U_0 \\ 0 & K_v & V_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{68}$$

We assume that $K_u > 0$ and $K_v > 0$ since f, k_u and k_v are positive by definition.

With this notation, the estimated normalized coordinates are related to the real ones as

$$\begin{aligned} \widetilde{x} &= K_u x + U_0 \\ \widetilde{y} &= K_v y + V_0 \end{aligned} \tag{69}$$

and therefore, its time derivatives are

$$\begin{aligned} \dot{\widetilde{x}} &= K_u \dot{x} \\ \dot{\widetilde{y}} &= K_v \dot{y} \end{aligned} \tag{70}$$

so that

$$\begin{aligned} \mathbf{L}_{\widetilde{x}} &= K_u \mathbf{L}_x \\ \mathbf{L}_{\widetilde{y}} &= K_v \mathbf{L}_y \end{aligned} \tag{71}$$

By using the equations above, it is easy to calculate the interaction matrix $\mathbf{L}_{\widetilde{\mathbf{s}}}$ corresponding to the visual features set $\widetilde{\mathbf{s}}$ measured under a bad camera calibration. Then, the local asymptotic stability analysis must be applied to the closed-loop equation of the measured task function $\widetilde{\mathbf{e}}$

$$\dot{\widetilde{\mathbf{e}}} = \mathbf{C}\mathbf{L}_{\widetilde{\mathbf{s}}}(\widetilde{\mathbf{e}}) \left(\mathbf{C}\widehat{\mathbf{L}_{\mathbf{s}}}\right)^{+} \widetilde{\mathbf{e}} \tag{72}$$

For the case of $\widetilde{\mathbf{s}} = (\widetilde{y}_1, \widetilde{x}_2, \widetilde{y}_3, \widetilde{x}_4)$ it can be found that the product of matrices $\mathbf{M}(\widetilde{\mathbf{e}})$ in the closed-loop equation of the system evaluated in the desired state becomes

$$\mathbf{M}(\widetilde{\mathbf{e}}^*) = \begin{pmatrix} \frac{K_u + K_v}{2} & 0 & 0 \\ 0 & K_v & 0 \\ 0 & 0 & K_u \end{pmatrix} \tag{73}$$

whose eigenvalues are in this case the elements of the main diagonal, which are always positive if $K_u > 0$ and $K_v > 0$, which is true if $\widetilde{f} > 0$, $\widetilde{k}_u > 0$ and $\widetilde{k}_v > 0$. Therefore, the local asymptotic stability of the system when using the image point coordinates is ensured if the elements of the main diagonal of $\widetilde{\mathbf{A}}$ are positive.

6.4 Simulation results

The system based on the set of visual features $\mathbf{s} = (y_1, x_2, y_3, x_4)$ has been simulated under the same conditions than the ones exposed in Section 5.1.

6.4.1 Ideal system

In Fig. 13 the results of the ideal system when using normalized image points is presented. In this case, the decrease of $\mathbf{s} - \mathbf{s}^*$ is not pure exponential and the rotational velocities generated by the constant control law based on $\mathbf{L}_{\mathbf{s}}^*$ are non-monotonic.

6.4.2 System including laser-cross misalignment and image noise

The system under the calibration errors described in Section 5.1.2 has rapidly diverged when using $\mathbf{s} = (y_1, x_2, y_3, x_4)$. This result was already expected from the local asymptotic stability analysis of this set of visual features when laser-cross misalignment occurs.

6.5 Linear combination of image points

As it has been said, in order to fulfill a plane-to-plane virtual link we only need $k = 3$ independent visual features. In the approach based on image points redundant information

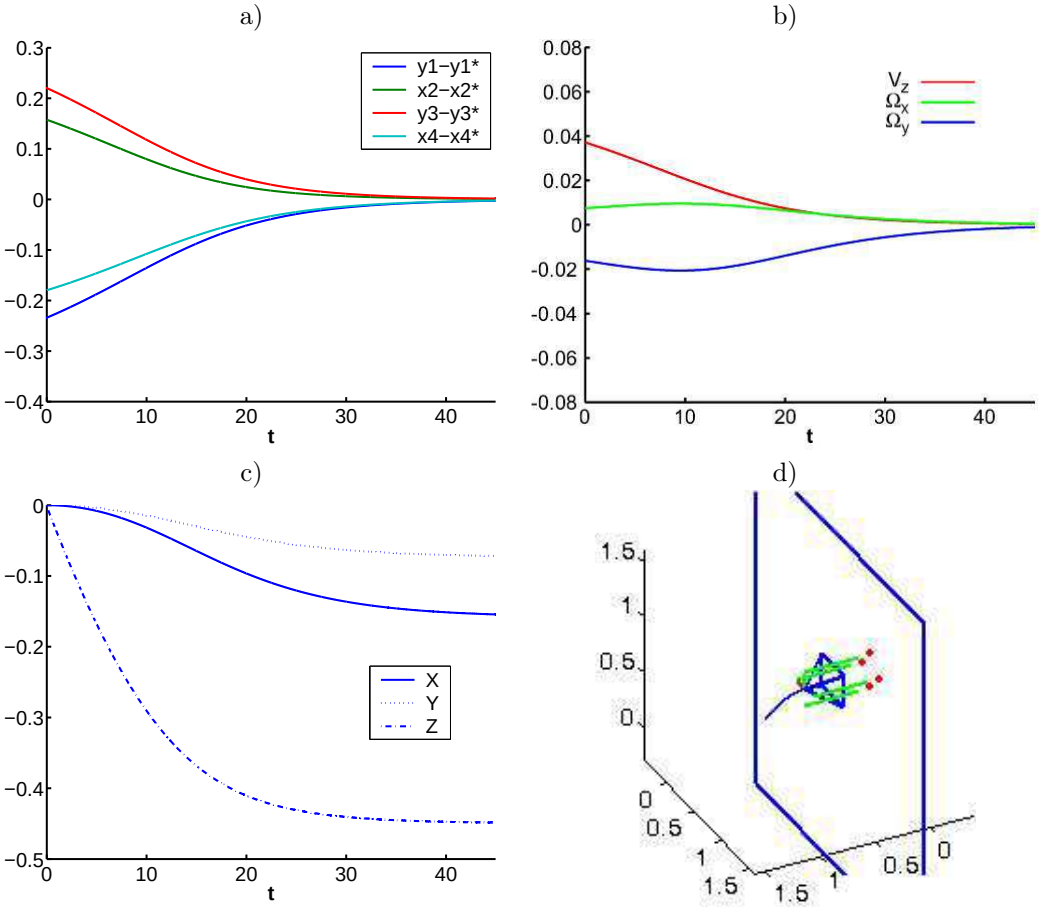

Figure 13: Ideal system: simulation using $\mathbf{s} = (y_1, x_2, y_3, x_4)$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

has been used since the number of visual features was greater than the number of controlled degrees of freedom. One might think about linearly combining the image points coordinates in order to obtain a set of 3 visual features.

For example, let us define the following set of visual features

$$\mathbf{f} = \begin{pmatrix} x_4 - x_2 & y_1 + y_3 & x_2 + x_4 \end{pmatrix} \tag{74}$$

The interaction matrix for the desired position is

$$\mathbf{L}_\mathbf{f}^* = \begin{pmatrix} 0 & 0 & 2L/Z^{*2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2L^2/Z^{*2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -2L^2/Z^{*2} & 0 \end{pmatrix} \tag{75}$$

which seems much more decoupled than the interaction matrix corresponding to the image points approach in (59).

Let us generalize the definition of the set of 3 visual features $\mathbf{f}$ as a linear combination of $\mathbf{s}$ of the form

$$\begin{aligned} \mathbf{f} &= \mathbf{Q}\mathbf{s} \\ \dot{\mathbf{Q}} &= \mathbf{0} \end{aligned} \tag{76}$$

so that the interaction matrix in the desired state has the general form

$$\mathbf{L}_\mathbf{f}^* = \begin{pmatrix} 0 & 0 & D_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & D_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & D_3 & 0 \end{pmatrix} \tag{77}$$

Note that $\mathbf{L}_\mathbf{f}^*$ can be decomposed as

$$\mathbf{L}_\mathbf{f}^* = \begin{pmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{pmatrix} \mathbf{W} = \mathbf{D}\mathbf{W} \tag{78}$$

On the other hand, by deriving (76) we have

$$\dot{\mathbf{f}} = \mathbf{Q}\dot{\mathbf{s}} = \mathbf{Q}\mathbf{L}_\mathbf{s}\mathbf{v} \tag{79}$$

so that

$$\mathbf{L}_\mathbf{f}^* = \mathbf{Q}\mathbf{L}_\mathbf{s}^* \tag{80}$$

Therefore, the following equality holds

$$\mathbf{D}\mathbf{W} = \mathbf{Q}\mathbf{L}_\mathbf{s}^* \tag{81}$$

and post-multiplying both sides for $\mathbf{L}_\mathbf{s}^{*+}$

$$\mathbf{D}\mathbf{W}\mathbf{L}_\mathbf{s}^{*+} = \mathbf{Q}\mathbf{L}_\mathbf{s}^*\mathbf{L}_\mathbf{s}^{*+} \tag{82}$$

since $\mathbf{L}_\mathbf{s}^* \mathbf{L}_\mathbf{s}^{*+} = \mathbf{I}_k$ we have

$$\mathbf{Q} = \mathbf{D}\mathbf{W}\mathbf{L}_\mathbf{s}^{*+} \tag{83}$$

The time derivative of the task function when using $\mathbf{s}$ can be expressed as

$$\mathbf{e_s} = \mathbf{C}(\mathbf{s} - \mathbf{s}^*) \Rightarrow \dot{\mathbf{e}}_\mathbf{s} = \mathbf{C}\dot{\mathbf{s}} = \mathbf{C}\mathbf{L_s}\mathbf{v} = \mathbf{W}\mathbf{L}_\mathbf{s}^{+}\mathbf{L_s}\mathbf{v} \tag{84}$$

Similarly, for the case of $\mathbf{f}$ we have

$$\mathbf{e_f} = \mathbf{f} - \mathbf{f}^* \Rightarrow \dot{\mathbf{e}}_\mathbf{f} = \dot{\mathbf{f}} = \mathbf{L_f}\mathbf{v} = \mathbf{Q}\mathbf{L}_\mathbf{s}^{+}\mathbf{L_s}\mathbf{v} = \mathbf{D}\mathbf{W}\mathbf{L}_\mathbf{s}^{+}\mathbf{L_s}\mathbf{v} \tag{85}$$

Note therefore that the dynamics of both tasks are related as follows

$$\dot{\mathbf{e}}_\mathbf{f} = \mathbf{D}\dot{\mathbf{e}}_\mathbf{s} \tag{86}$$

Therefore, the system dynamics of $\mathbf{e_f}$ are identical to the dynamics of $\mathbf{e_s}$ but including a constant factor. Therefore, using a linear combination of visual features which obtains a diagonal interaction matrix in the desired state does not affect the behavior of the system.

In the following sections new sets of visual features are proposed aiming to improve the performance of the system in terms of stability against calibration errors and decoupling. As it will be seen, the features are based on non-linear combinations of the image points coordinates. Thus, matrix $\mathbf{Q}$ will depend on the state so that it will be no longer constant and therefore, the dynamics of $\mathbf{e}$ will change.

7 Normalized area and angles approach

In this section we analyze the performance of a set of visual features consisting of non-linear combinations of the image points [45].

The first visual feature is based on the area of an element of the image. Such visual feature has been largely used for depth control [17,38,53]. In our case, we take into account the area enclosed by the four image points, which can be formulated as follows

$$a = \frac{1}{2}\left((x_3 - x_1)(y_4 - y_2) + (x_2 - x_4)(y_3 - y_1)\right) \tag{87}$$

The interaction matrix of the area can be easily derived by using the interaction matrices of the image point coordinates appearing in the formula above. The interaction matrix evaluated in any state where the camera is parallel to the object ($A = 0$, $B = 0$, $C = 1$) at a certain depth Z ($D = -Z$) will be hereafter denoted as $\mathbf{L}^{\parallel}$. For the case of the area this matrix is

$$\mathbf{L}_a^{\parallel} = \begin{pmatrix} 0 & 0 & 2a^{\parallel}/Z & 0 & 0 & 0 \end{pmatrix} \tag{88}$$

Note that the area $a^{\parallel}$ observed in any position where the camera is parallel to the object is known. According to the ideal model it depends on the lasers positions which are symmetrically placed around the camera and pointing towards the same direction than the optical axis. Concretely, we have

$$a^{\parallel} = \frac{A^{\parallel}}{Z^2} \tag{89}$$

where $A^{\parallel}$ is the $3D$ area enclosed by the four points onto the object surface whenever the camera is parallel to the object. Since the four laser pointers are orthogonally projected, the $3D$ area is constant for any position where the camera and the object are parallel. Concretely, we have that

$$A^{\parallel} = 2L^2 \tag{90}$$

Thus, the interaction matrix in (88) can be rewritten as

$$\mathbf{L}_a^{\parallel} = \begin{pmatrix} 0 & 0 & 4L^2/Z^3 & 0 & 0 & 0 \end{pmatrix} \tag{91}$$

Note that the dynamics of the area area strongly non-linear.

The 2 visual features controlling the remaining degrees of freedom are selected from the 4 virtual segments defined according to Fig. 14.

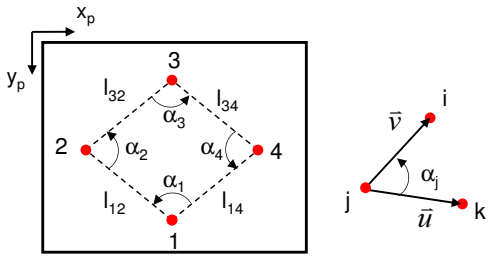

Figure 14: At left side, virtual segments defined by the image points. At right side, definition of the angle α_j.

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 Fig. 14) is defined as

$$\sin\alpha_j = \frac{\|\overrightarrow{u}\times\overrightarrow{v}\|}{\|\overrightarrow{u}\|\|\overrightarrow{v}\|} \quad , \quad \cos\alpha_j = \frac{\overrightarrow{u}\cdot\overrightarrow{v}}{\|\overrightarrow{u}\|\|\overrightarrow{v}\|} \tag{92}$$

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

$$\alpha_j = \arctan\frac{(x_k-x_j)(y_i-y_j)-(x_i-x_j)(y_k-y_j)}{(x_k-x_j)(x_i-x_j)+(y_k-y_j)(y_i-y_j)} \tag{93}$$

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 for the case whenever the camera is parallel to the object

$$\begin{aligned} \mathbf{L}_{\alpha_{13}}^{\parallel} &= \begin{pmatrix} 0 & 0 & 0 & 2L/Z & 0 & 0 \end{pmatrix} \\ \mathbf{L}_{\alpha_{24}}^{\parallel} &= \begin{pmatrix} 0 & 0 & 0 & 0 & 2L/Z & 0 \end{pmatrix} \end{aligned} \tag{94}$$

Note that by using the visual feature set $\mathbf{s} = (\ a,\ \alpha_{13},\ \alpha_{24})$ the interaction matrix is diagonal (for any state where the camera and the object are parallel) 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 or a power of the depth Z^3. This will cause the camera trajectory to be not completely satisfactory. As pointed out by Mahony et al. [38], a good visual feature controlling one degree of freedom (dof) is the one whose error function varies proportionally to the variation of the dof.

Let us 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

$$a_n = a^\gamma \Rightarrow \dot{a}_n = \gamma a^{\gamma-1}\dot{a} = \frac{2\gamma A^\gamma}{Z^{2\gamma+1}} \cdot V_z \tag{95}$$

In order to cancel the depth it is necessary that

$$2\gamma + 1 = 0 \Rightarrow \gamma = -1/2 \tag{96}$$

so that we find $a_n = 1/\sqrt{a}$ as in [38, 52]. The interaction matrix of a_n evaluated in the desired state is in fact valid for any camera position where it is parallel to the object since it only depends on constant values and not on the depth

$$\mathbf{L}^*_{a_n} = \mathbf{L}^{\parallel}_{a_n} = (\ 0\ \ 0\ \ -1/(\sqrt{2}L)\ \ 0\ \ 0\ \ 0\ ) \tag{97}$$

Following the same method we can find that defining

$$\alpha_{13n} = \alpha_{13n}/\sqrt{a} \tag{98}$$

$$\alpha_{24n} = \alpha_{24n}/\sqrt{a} \tag{99}$$

we obtain the following interaction matrices for the new normalized features

$$\begin{array}{rcl} \mathbf{L}^*_{\alpha_{13_n}} = \mathbf{L}^{\parallel}_{\alpha_{13_n}} & = & (\ 0\ \ 0\ \ 0\ \ \sqrt{2}\ \ 0\ \ 0\ ) \\ \mathbf{L}^*_{\alpha_{24_n}} = \mathbf{L}^{\parallel}_{\alpha_{24_n}} & = & (\ 0\ \ 0\ \ 0\ \ 0\ \ \sqrt{2}\ \ 0\ ) \end{array} \tag{100}$$

These normalized visual features are related to the object parameters (A, B, C, D) as follows

$$\begin{aligned} a_n &= -\frac{\sqrt{(A^2L^2 - D^2)(B^2L^2 - D^2)}}{\sqrt{2}LCD} \\ \alpha_{13n} &= a_n \arctan\left(\frac{4LBD(L^2(A^2+B^2) - 2D^2)}{L^4(A^2+B^2)^2 + 4D^2(D^2 - 2B^2L^2)}\right) \\ \alpha_{24n} &= a_n \arctan\left(\frac{4LAD(L^2(A^2+B^2) - 2D^2)}{L^4(A^2+B^2)^2 + 4D^2(D^2 - 2A^2L^2)}\right) \end{aligned} \tag{101}$$

Similarly, they are related to the object representation ($\gamma = -D/C, \beta = -B/C, \alpha = -A/C$) by

$$\begin{aligned} a_n &= -\frac{\sqrt{(\alpha^2L^2-\gamma^2)(\beta^2L^2-\gamma^2)}}{\sqrt{2}L\gamma} \\ \alpha_{13n} &= a_n \arctan\left(\frac{4L\beta\gamma(L^2(\alpha^2+\beta^2)-2\gamma^2)}{L^4(\alpha^2+\beta^2)^2+4\gamma^2(\gamma^2-2\beta^2L^2)}\right) \\ \alpha_{24n} &= a_n \arctan\left(\frac{4L\alpha\gamma(L^2(\alpha^2+\beta^2)-2\gamma^2)}{L^4(\alpha^2+\beta^2)^2+4\gamma^2(\gamma^2-2\alpha^2L^2)}\right) \end{aligned} \tag{102}$$

If a Taylor approximation of first order is made about $A = B = 0$ or $\alpha = \beta = 0$, the following relationships appear

$$\begin{aligned} a_n &\approx \frac{1}{\sqrt{2}L}\left(\frac{-D}{C}\right) = \frac{1}{\sqrt{2}L}\gamma \\ \alpha_{13n} &\approx -\sqrt{2}\left(\frac{-B}{C}\right) = -\sqrt{2}\beta \\ \alpha_{24n} &\approx \sqrt{2}\left(\frac{-A}{C}\right) = \sqrt{2}\alpha \end{aligned} \tag{103}$$

Therefore, when the camera is nearly parallel to the object, the features based on normalized area and angles are proportional to the object parameters (γ,β,α). That is why the features $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$ are decoupled in the desired state.

Given this set of visual features we have $m = k = 3$ so that $\mathbf{C} = \mathbf{I}_3$ and the control law is

$$\mathbf{v} = -\lambda\widehat{\mathbf{L}_\mathbf{s}}^{+}(\mathbf{s} - \mathbf{s}^*) \tag{104}$$

and therefore the closed-loop equation of the system is

$$\dot{\mathbf{e}} = -\lambda\mathbf{L}_\mathbf{s}\widehat{\mathbf{L}_\mathbf{s}}^{+}\mathbf{e} \tag{105}$$

When using a constant control law, the estimation of the interaction matrix is simply

$$\widehat{\mathbf{L}_\mathbf{s}} = \mathbf{L}_\mathbf{s}^* = \mathbf{L}_\mathbf{s}^{\parallel} = \begin{pmatrix} 0 & 0 & -1/(\sqrt{2}L) & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \end{pmatrix} \tag{106}$$

7.1 Global asymptotic stability under perfect conditions

Unfortunately, the interaction matrix in function of the object parameters corresponding to the ideal model is very complex. For example, we show the non-null elements of the general

interaction matrix for a_n expressed in the object frame (see Appendix C)

$$
\begin{aligned}
{}^{o}\mathbf{L}_{a_n} &= \begin{pmatrix} 0 & 0 & {}^{o}L_{a_n}(V_z) & {}^{o}L_{a_n}(\Omega_x) & {}^{o}L_{a_n}(\Omega_y) & 0 \end{pmatrix} \qquad (107)\\
{}^{o}L_{a_n}(V_z) &= 4\frac{C^2L^2D(A^2B^2L^4 - D^4)}{(A^2L^2 - D^2)^2(B^2L^2 - D^2)^2}\\
{}^{o}L_{a_n}(\Omega_x) &= \frac{4L^2CBD^2((1+C)(D^4 - L^2D^2) + A^2L^4(1 - A^2 + C(1+B^2)))}{(1+C)(A^2L^2 - D^2)^2(B^2L^2 - D^2)^2}\\
{}^{o}L_{a_n}(\Omega_y) &= -\frac{4L^2CAD^2((1+C)(D^4 - L^2D^2) + B^2L^4(1 - B^2 + C(1 - A^2)))}{(1+C)(A^2L^2 - D^2)^2(B^2L^2 + D^2)^2}
\end{aligned}
$$

The interaction matrices for α_{13n} and α_{24n} are still more complicated because of the definition of their time derivatives

$$
\begin{aligned}
\dot{\alpha}_{13n} &= a_n \cdot \dot{\alpha}_{13} + \alpha_{13} \cdot \dot{a}_n\\
\dot{\alpha}_{24n} &= a_n \cdot \dot{\alpha}_{24} + \alpha_{24} \cdot \dot{a}_n \qquad (108)
\end{aligned}
$$

Note that when the camera is not parallel to the object α_{13} and α_{24} are different to 0. Then, the general interaction matrices depend on arctan functions.

Trying to analyze the global asymptotic stability of the system when using the constant control law based on matrix (106) becomes too complex. We could only ensure the global asymptotic stability of the ideal system when using a non-constant control law based on perfectly estimating $\widehat{\mathbf{L_s}} = \mathbf{L_s}$ at each iteration. In this case, the closed-loop equation of the system is

$$\dot{\mathbf{e}} = -\lambda \mathbf{e} \qquad (109)$$

obtaining a pure exponential decrease of the task function.

Hereafter we focus on the image-based approach based on $\mathbf{s} = (a_n,\ \alpha_{13n},\ \alpha_{24n})$ and the constant control law. We show its robustness against to calibration errors through the local asymptotic stability analysis.

7.2 Local asymptotic stability analysis under laser-cross misalignment

7.2.1 Misalignment consisting of a translation

The interaction matrix for the desired position including a misalignment of the laser-cross consisting of a translation ${}^{C}\mathbf{T}_{L} = (t_x,\ t_y,\ t_z)$ has been computed taking into account the model parameters in Table 3 at Appendix B obtaining

$$
\mathbf{L_s}(\mathbf{e}^*) = \begin{pmatrix}
0 & 0 & -\frac{\sqrt{2}}{2L} & -\frac{3\sqrt{2}t_y}{4L} & \frac{3\sqrt{2}t_x}{4L} & 0\\
0 & 0 & 0 & \sqrt{2} & 0 & 0\\
0 & 0 & 0 & 0 & \sqrt{2} & 0
\end{pmatrix} \qquad (110)
$$

Note that the misalignment parameters only affect the normalized area a_n. On the other hand, α_{13n} and α_{24n} are invariant to such type of misalignment near the desired state.

Then, the product of matrices in the closed-loop equation (105) becomes $\mathbf{M}(\mathbf{e}^*) = \mathbf{L_s}(\mathbf{e}^*)\widehat{\mathbf{L_s}}^+$ that is

$$\mathbf{M}(\mathbf{e}^*) = \begin{pmatrix} 1 & -\frac{3t_y}{4L} & \frac{3t_x}{4L} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{111}$$

Note that the eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are equal to the elements of the main diagonal. Therefore, the eigenvalues are all equal to 1. It means that the local asymptotic stability of the system when using this set of visual features is not affected by a misalignment of the laser-cross consisting of a translation. Therefore, the stability domain of this set of visual features is much larger than the approach based on image points.

7.2.2 Misalignment consisting of individual rotations

Let us now consider how does a rotation of the laser-cross around one of the axis of the camera frame affect the local asymptotic stability of the system.

Given the case of a rotation ψ around the X axis the eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are

$$\begin{aligned} \sigma_1 &= \frac{2\cos\psi\sqrt{1/\cos\psi}}{\cos^2\psi + 1} \\ \sigma_2 &= \frac{2\cos^2\psi\sqrt{1/\cos\psi}}{\cos^2\psi + 1} \\ \sigma_3 &= \cos\psi\sqrt{1/\cos\psi} \end{aligned}$$

which are all positive and definite if $\psi \in (-\pi/2,\ \pi/2)$.

The same eigenvalues are obtained for the case of a rotation θ around the Y axis of the camera.

Finally, when a rotation ϕ is applied around the Z axis, the eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are

$$\begin{aligned} \sigma_1 &= 1 \\ \sigma_2 &= \cos\phi + \sqrt{\cos^2\phi - 1} \\ \sigma_3 &= \cos\phi - \sqrt{\cos^2\phi - 1} \end{aligned}$$

Note that the real part of the two last eigenvalues is $\cos\phi$ so in order to ensure their positivity we must have $\phi \in [-\pi/2,\ \pi/2]$.

It can be stated that the local asymptotic stability of the approach based on the normalized area and angles is not affected by individual rotations of the laser-cross around the camera axis.

7.3 Local asymptotic stability analysis in presence of camera calibration errors

We now present the local asymptotic stability analysis taking into account in the real interaction matrix both the real and the estimated intrinsic parameters of the camera. As explained in Section 6.3, it is necessary to study the dynamic behavior of the measured task function $\widetilde{\mathbf{e}}$ evaluated in the desired state

$$\dot{\widetilde{\mathbf{e}}}^* = -\lambda \mathbf{L}_{\widetilde{\mathbf{s}}}(\widetilde{\mathbf{e}}^*)\widehat{\mathbf{L}_{\mathbf{s}}}^{+}\widetilde{\mathbf{e}}^* \tag{112}$$

In this case, the interaction matrix in the desired state $\widetilde{\mathbf{e}}^*$ is

$$\mathbf{L}_{\widetilde{\mathbf{s}}}(\widetilde{\mathbf{e}}^*) = \begin{pmatrix} 0 & 0 & -\frac{\sqrt{2}(K_u+K_v)}{4L} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2LK_v}{\sqrt{K_u}Z^*} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2\sqrt{K_u}L}{Z^*} & 0 \end{pmatrix} \tag{113}$$

so that the product of matrices $\mathbf{M}(\widetilde{\mathbf{e}}^*) = \mathbf{L}_{\mathbf{s}}(\widetilde{\mathbf{e}}^*)\widehat{\mathbf{L}_{\mathbf{s}}}^{+}$ is

$$\mathbf{M}(\widetilde{\mathbf{e}}^*) = \begin{pmatrix} \frac{K_u+K_v}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}LK_v}{\sqrt{K_u}Z^*} & 0 \\ 0 & 0 & \frac{\sqrt{2K_u}L}{Z^*} \end{pmatrix} \tag{114}$$

Note that the eigenvalues are the elements of the main diagonal and are positive if $K_u > 0$ and $K_v > 0$, which is true if $\widetilde{f} > 0$, $\widetilde{k}_u > 0$ and $\widetilde{k}_v > 0$.

Therefore, the system based on the normalized area and angles is robust against camera calibration errors if the elements of the main diagonal of $\widetilde{\mathbf{A}}$ are positive.

7.4 Simulation results

The simulations described in Section 5.1 have been also done taking into account the set of visual features $\mathbf{s} = (a_n,\ \alpha_{13n},\ \alpha_{24n})$ and the constant control law based on $\mathbf{L}_{\mathbf{s}}^*$.

7.4.1 Ideal system

The results obtained according to the ideal system specification are shown in Fig.3 15. The decrease of the task function is strictly monotonic as well as the camera velocities.

Note that the camera velocities and the camera trajectory are pretty similar to the ones obtained by the position-based approach in Section 5.1.1 when using the constant control law. This similarity was expected from the analytic approximate behavior deduced in (103).

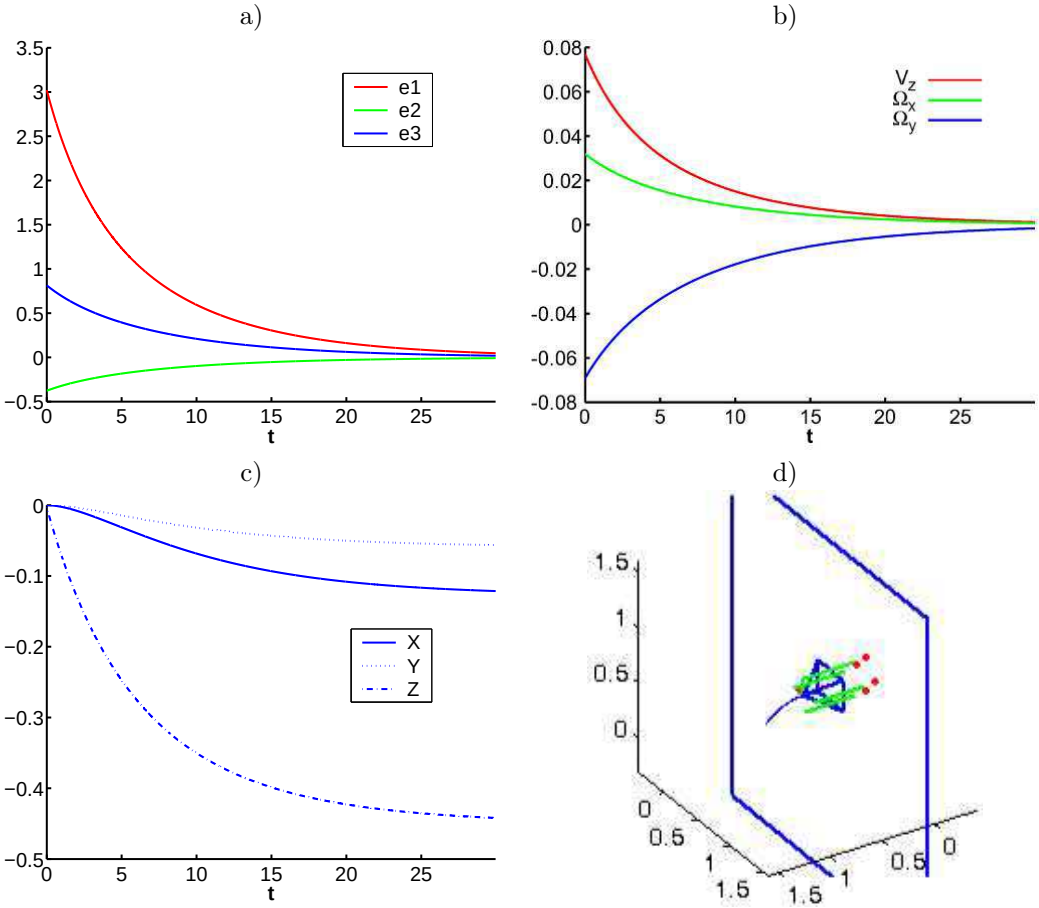

Figure 15: Ideal system: simulation using $\mathbf{s} = (a_n,\ \alpha_{13n},\ \alpha_{24n})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

7.4.2 System including laser-cross misalignment and image noise

Fig. 16 presents the results obtained by the current image-based approach for the simulation conditions described in Section 5.1.2. Note that the system is almost unaffected by the large laser-cross misalignment, as already expected from the local asymptotic stability analysis results in presence of this type of calibration errors. Again, the results are very similar to the ones obtained by the positioned-based approach when using the constant control law.

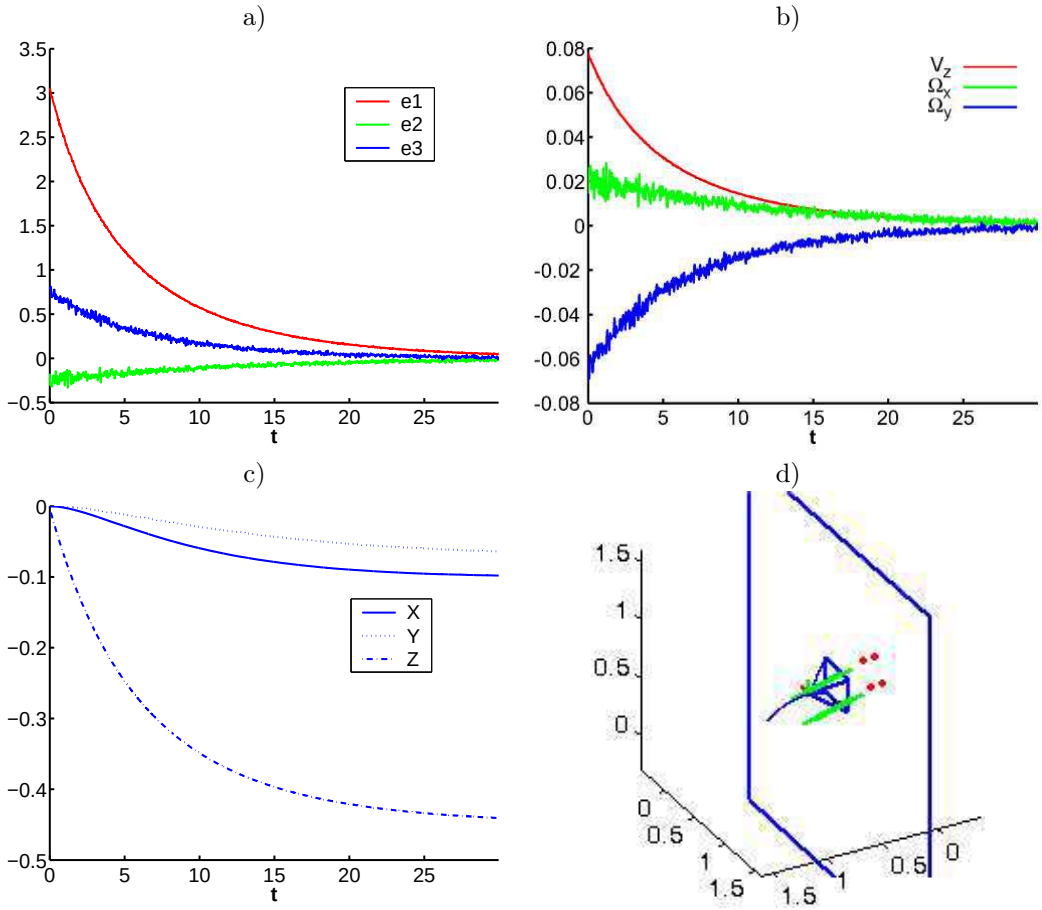

Figure 16: System including large laser-cross misalignment and image noise: simulation using $\mathbf{s} = (a_n,\ \alpha_{13n},\ \alpha_{24n})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

8 A decoupled image-based approach

The aim of this section is to obtain a set of 3 visual features which decouples as much as possible the controlled degrees of freedom not only near the desired position (as in the case of the normalized area and angles) but for any camera-object pose. With such a decoupling we also aim to demonstrate the global asymptotic stability of the system when using the constant control law based on $\mathbf{L}_\mathbf{s}^*$. Moreover, we expect a set of visual features which have the same robustness against calibration errors demonstrated for the previous features.

Let us take a look at the interaction matrices of y_1^{-1}, y_3^{-1}, x_2^{-1} and x_4^{-1}

$$\begin{aligned}
\mathbf{L}_{y_1^{-1}} &= \begin{pmatrix} -K_1 & -K_2 & -K_3 & -\frac{K_2(BL+D)}{C} - 1 & \frac{K_1(BL+D)}{C} & K_1 L \end{pmatrix} \\
\mathbf{L}_{y_3^{-1}} &= \begin{pmatrix} K_1 & K_2 & K_3 & -\frac{K_2(BL-D)}{C} - 1 & \frac{K_1(BL-D)}{C} & K_1 L \end{pmatrix} \\
\mathbf{L}_{x_2^{-1}} &= \begin{pmatrix} K_1 & K_2 & K_3 & -\frac{K_2(AL-D)}{C} & \frac{K_1(AL-D)}{C} + 1 & -K_2 L \end{pmatrix} \\
\mathbf{L}_{x_4^{-1}} &= \begin{pmatrix} -K_1 & -K_2 & -K_3 & -\frac{K_2(AL+D)}{C} & \frac{K_1(AL+D)}{C} + 1 & -K_2 L \end{pmatrix}
\end{aligned}$$

$$\text{with} \qquad K_1 = \frac{A}{LC} \quad K_2 = \frac{B}{LC} \quad K_3 = \frac{1}{L}$$

It is obvious that simple combinations of such features can lead to a decoupled system. We have chosen the following set of visual features

$$\begin{aligned}
\mathbf{s} &= \begin{pmatrix} y_1^{-1} - y_3^{-1}, & y_1^{-1} + y_3^{-1}, & x_2^{-1} + x_4^{-1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{y_3 - y_1}{y_1 y_3}, & \frac{y_1 + y_3}{y_1 y_3}, & \frac{x_2 + x_4}{x_2 x_4} \end{pmatrix}
\end{aligned} \tag{115}$$

whose interaction matrix is

$$\mathbf{L}_\mathbf{s} = \begin{pmatrix} -\frac{2A}{LC} & -\frac{2B}{LC} & -\frac{2}{L} & -\frac{2BD}{LC^2} & \frac{2AD}{LC^2} & 0 \\ 0 & 0 & 0 & -\frac{2(1-A^2)}{C^2} & \frac{2AB}{C^2} & \frac{2A}{C} \\ 0 & 0 & 0 & -\frac{2AB}{C^2} & \frac{2(1-B^2)}{C^2} & -\frac{2B}{C} \end{pmatrix} \tag{116}$$

which is always rank 3 unless for degenerated cases. Note that the rotational part is decoupled from the translational one for any camera-object pose. To our knowledge, there are no other image-based approach where the interaction matrix has such a high order of decoupling for any state.

In the ideal case, that is, when no calibration errors occur, the visual features are related to the object parameters as follows

$$\mathbf{s} = (s_1,\ s_2,\ s_3) = \left(-\frac{2D}{LC},\ -2\frac{B}{C},\ -2\frac{A}{C} \right) \tag{117}$$

Therefore, these features are proportional to the parameters of the object plane expressed as

$$Z = \gamma + \beta Y + \alpha X \tag{118}$$

so that

$$s_1 = \frac{2}{L}\gamma,\ s_2 = 2\beta,\ s_3 = 2\alpha \tag{119}$$

Furthermore, taking into account the Taylor developments in (103), the new visual features are related to $(a_n, \alpha_{13n}, \alpha_{24n})$ by

$$s_1 \approx 2\sqrt{2}a_n,\ s_2 \approx -\sqrt{2}\alpha_{13n},\ s_3 \approx \sqrt{2}\alpha_{24n} \tag{120}$$

Therefore, under ideal conditions, the image-based approach based on these features behaves as the position-based technique presented in Section 5 and very similar to the image-based approach based on $(a_n, \alpha_{13n}, \alpha_{24n})$. Thus, a new way to implicitly estimate the object pose has been found from a non-linear combination of the image point coordinates. In absence of calibration errors the equations in (119) could be used to obtain the object parameters (γ, β, α) and execute the position-based approach without need to solve the system of non-linear equations in (52).

Another interesting characteristic of these features is that the interaction matrix can be expressed in terms of the task function components. Usually, in most part of $2D$ visual servoing approaches this is unfeasible. By using the normalized image points coordinates for the ideal case (presented in Table 1), the components of $\mathbf{e} = \mathbf{s} - \mathbf{s}^* = (e_1, e_2, e_3)$ can be expressed as follows

$$e_1 = \frac{-2(D + CZ^*)}{LC},\ e_2 = -2\frac{B}{C},\ e_3 = -2\frac{A}{C} \tag{121}$$

We remember that $\underline{\mathbf{n}}$ is a unitary vector so that $C = \sqrt{1 - A^2 - B^2}$. Thus, we have a system of 3 equations and 3 unknowns (A, B and D) whose unique solution is

$$A = -\frac{e_3}{h},\ B = -\frac{e_2}{h},\ D = -\frac{e_1 L + 2Z^*}{h} \tag{122}$$

with

$$h = \sqrt{e_2^2 + e_3^2 + 4} \tag{123}$$

and therefore

$$C = \frac{2}{h} \tag{124}$$

Using these equivalences, the interaction matrix can be expressed in terms of the task function components as follows

$$\mathbf{L_s}(\mathbf{e}) = \begin{pmatrix} \frac{e_3}{L} & \frac{e_2}{L} & -\frac{2}{L} & -\frac{e_2(e_1 L + 2Z^*)}{2L} & \frac{e_3(e_1 L + 2Z^*)}{2L} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}e_2^2 - 2 & \frac{1}{2}e_2 e_3 & -e_3 \\ 0 & 0 & 0 & -\frac{1}{2}e_2 e_3 & 2 + \frac{1}{2}e_3^2 & e_2 \end{pmatrix} \tag{125}$$

Note that all the terms in the interaction matrix are known. This allows us to decide which model of interaction matrix is used in the control law obtained by setting $\mathbf{W} = \widehat{\mathbf{L_s}}$ so that $\mathbf{C} = \mathbf{I}_3$ and

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_s}}^{+} \mathbf{e} \tag{126}$$

- $\widehat{\mathbf{L_s}}$ estimated at each iteration. Note that in this case the elements of the interaction matrix can be obtained from the task function, without reconstructing the object parameters by triangulation. The main advantage of such a control law is that if the interaction matrix is perfectly estimated the task function will have a pure exponential decrease. However, the camera velocities may be inadequate due to the non-linearities visible in (125).
- $\widehat{\mathbf{L_s}} = \mathbf{L_s^*}$ being

$$\widehat{\mathbf{L_s}} = \mathbf{L_s^*} = \begin{pmatrix} 0 & 0 & -2/L & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix} \tag{127}$$

  In this case, the control law becomes simpler and can be calculated faster (it is not required to calculate the pseudoinverse at each iteration). Note that this matrix does not contain any non-linearities neither depth information, like in the set of visual features based on the normalized area and angles. In this case however, this result has been achieved without need to normalize the features as before. Then, since each visual feature varies proportionally to the dof which controls, if the task function has a good decreasing, which will be studied in the following subsection, suitable camera velocities will be produced [38].

8.1 Global asymptotic stability under perfect conditions

The closed-loop equation of the system is again

$$\dot{\mathbf{e}} = -\lambda \mathbf{L_s}(\mathbf{e}) \widehat{\mathbf{L_s}}^{+} \mathbf{e} \tag{128}$$

We now present the global asymptotic stability analysis when using two different control laws.

8.1.1 Non-constant control law

In this case we compute the real interaction matrix at each iteration from the task function value so that $\widehat{\mathbf{L_s}} = \mathbf{L_s}(\mathbf{e})$. Then the product of matrices $\mathbf{M}$ in the control law is the identity. Therefore, the equilibrium point is unique and the closed-loop equation becomes simply

$$\dot{\mathbf{e}} = -\lambda \mathbf{e} \tag{129}$$

which ensures a pure exponential decrease of the task function. However, the camera velocities produced by the non-constant control law can be not very suitable due to the strong non-linearities in $\mathbf{L_s}(\mathbf{e})^{+}$.

8.1.2 Constant control law

When using the constant diagonal matrix in (127) in the control law, the product of matrices $\mathbf{M} = \mathbf{L_s}\widehat{\mathbf{L_s}}^{+} = \mathbf{L_s}\mathbf{L_s^*}^{+}$ is the following 3×3 matrix

$$\mathbf{M} = \begin{pmatrix} 1 & \frac{BD}{LC^2} & \frac{AD}{LC^2} \\ 0 & \frac{B^2+C^2}{C^2} & \frac{AB}{C^2} \\ 0 & \frac{AB}{C^2} & \frac{A^2+C^2}{C^2} \end{pmatrix} = \begin{pmatrix} 1 & \frac{e_2(e_1L+2Z^*)}{4L} & \frac{e_3(e_1L+2Z^*)}{4L} \\ 0 & \frac{e_2^2}{4}+1 & \frac{e_2e_3}{4} \\ 0 & \frac{e_2e_3}{4} & \frac{e_3^2}{4}+1 \end{pmatrix} \tag{130}$$

whose determinant is

$$\det(\mathbf{M}) = \frac{1}{C^2} = \frac{h^2}{4} = \frac{\sqrt{e_2^2+e_3^2+4}}{4} \tag{131}$$

which is always non-null, and therefore, the equilibrium point $\mathbf{e} = 0$ is unique. The global asymptotic stability analysis can be done by using the Lyapunov method, but the sufficient conditions provided by this method are very restrictive as shown in Appendix D.1. However, thanks to the nice decoupled form of the interaction matrix, we can solve the differential system in function of time corresponding to the closed-loop equation of the system

$$\dot{\mathbf{e}}(t) = -\lambda \mathbf{L_s}(\mathbf{e}(t))\widehat{\mathbf{L_s}}^{+}\mathbf{e}(t) \tag{132}$$

This differential system can be decomposed as follows

$$\dot{e}_1(t) = -\frac{\lambda}{4L}\left(e_1(t)\left(4L + e_2(t)^2L + e_3(t)^2L\right) + 2Z^*\left(e_2(t)^2 + e_3(t)^2\right)\right) \tag{133}$$

$$\dot{e}_2(t) = -\frac{\lambda}{4}(e_2(t)^3 + 4e_2(t) + e_2(t)e_3(t)^2) \tag{134}$$

$$\dot{e}_3(t) = -\frac{\lambda}{4}(e_3(t)^3 + 4e_3(t) + e_3(t)e_2(t)^2) \tag{135}$$

The following solutions are obtained according to the developments presented in Appendix D.2

$$e_1(t) = \frac{2e_1(0)}{a(t)} - \frac{2bZ^*\arctan(u(t))}{a(t)L} \tag{136}$$

$$e_2(t) = \frac{2e_2(0)}{a(t)} \tag{137}$$

$$e_3(t) = \frac{2e_3(0)}{a(t)} \tag{138}$$

with

$$a(t) = \sqrt{(e_2^2(0)+e_3^2(0))\,(\exp^{2\lambda t}-1) + 4\exp^{2\lambda t}} \tag{139}$$

$$b = \sqrt{e_2^2(0)+e_3^2(0)} \tag{140}$$

$$u(t) = \frac{b(a(t)-2)}{b^2+2a(t)} \tag{141}$$

Let us start by demonstrating the global asymptotic stability of the rotational subsystem defined by (134) and (135). The subsystem formed by $e_2(t)$ and $e_3(t)$ is globally asymptotically stable if

$$\lim_{t\to\infty} e_2(t) = 0, \quad \lim_{t\to\infty} e_3(t) = 0 \tag{142}$$

Both functions clearly tend to 0 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

$$\dot{e}_i(t) \quad = \quad -\frac{2\lambda e_i(0)\exp^{2\lambda t}\left(e_2^2(0)+e_3^2(0)+4\right)}{a(t)^3} \tag{143}$$

with $i = \{2,3\}$. Note that the functions e_2 and e_3 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 0. Therefore, for any initial condition, $e_2(t)$ and $e_3(t)$ always tend towards 0 strictly monotonically.

The global asymptotic stability of the translational subsystem depends on the behavior of $e_1(t)$. It is easy to show that $e_1(t)$ converges to 0 for any initial state since

$$\left.\begin{array}{lcl} \lim_{t\to\infty} u(t) & = & \frac{b}{2} \\ \lim_{t\to\infty} a(t) & = & \infty \end{array}\right\} \Rightarrow \lim_{t\to\infty} e_1(t) = 0 \tag{144}$$

The monotonicity of $e_1(t)$ is not so easy to proof. In fact, depending on the initial conditions, $e_1(t)$ can be not monotonic showing some extrema. In Appendix D.3 it is shown that $e_1(t)$ either is always monotonic or it has a unique extremum before converging monotonically to 0. Furthermore, sufficient conditions are given so that it is possible to check from the initial state of the system and the desired depth Z^* if either $e_1(t)$ will be monotonic during all the servoing or it will have a peak.

8.2 Camera trajectory

Thanks to the decoupled form of the interaction matrix in (116) we have obtained the analytic functions describing the behavior of the task function $\mathbf{e}(t)$. Furthermore, when using the constant control law based on $\mathbf{L}_\mathbf{s}^*$ we can also obtain the functions describing the camera trajectory. In this case, the control law maps the task function components $e_1(t)$, $e_2(t)$ and $e_3(t)$ to the camera velocities as follows

$$\mathbf{v} = -\lambda \mathbf{L}_\mathbf{s}^{*+}\mathbf{e} \tag{145}$$

where

$$\mathbf{L}_\mathbf{s}^{*+} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -L/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 1/2 \\ 0 & 0 & 0 \end{pmatrix} \tag{146}$$

so that

$$\begin{cases} V_z(t) &= \lambda \dfrac{L}{2} e_1(t) \\ \Omega_x(t) &= \lambda \dfrac{1}{2} e_2(t) \\ \Omega_y(t) &= -\lambda \dfrac{1}{2} e_3(t) \end{cases} \tag{147}$$

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

$$\dot{\mathbf{X}}(t) = -\mathbf{V}(t) - \mathbf{\Omega}(t) \times \mathbf{X}(t) \tag{148}$$

Since the constant control law only generates velocities for V_z, Ω_x and Ω_y, the above equation can be rewritten as

$$\begin{cases} \dot{X}(t) &= -\Omega_y(t) Z(t) \\ \dot{Y}(t) &= +\Omega_x(t) Z(t) \\ \dot{Z}(t) &= -V_z(t) + \Omega_y(t) X(t) - \Omega_x(t) Y(t) \end{cases} \tag{149}$$

where $V_z(t)$, $\Omega_x(t)$ and $\Omega_y(t)$ are the expressions in (147). If we choose as fixed point the initial position of the camera ($\mathbf{X}(0) = (0,0,0)$), we can solve the system of differential equations obtaining

$$\begin{cases} X(t) &= \dfrac{e_3(0)\exp^{-\lambda t}}{h(0)^2 b^3} \left(\exp^{\lambda t} b^2 Z^* h(0)^2 \arctan\left(u(t)\right) - e_1(0) L b \left(\exp^{\lambda t} h(0)^2 - b^2 - 2a(t) \right) \right. \\ & \quad \left. + b^3 Z^* \left(a(t) - 2\right) \right) \\ Y(t) &= \dfrac{e_2(0)\exp^{-\lambda t}}{h(0)^2 b^3} \left(\exp^{\lambda t} b^2 Z^* h(0)^2 \arctan\left(u(t)\right) - e_1(0) L b \left(\exp^{\lambda t} h(0)^2 - b^2 - 2a(t) \right) \right. \\ & \quad \left. + b^3 Z^* \left(a(t) - 2\right) \right) \\ Z(t) &= \dfrac{-\exp^{-\lambda t}}{h(0)^2} \left(-b^2 Z^* \left(\exp^{\lambda t} - 1 \right) + e_1(0) L \left(a(t) - 2\right) 2 Z^* \left(a(t) - 2\exp^{\lambda t} \right) \right) \end{cases} \tag{150}$$

Note that $\mathbf{X}(t)$ describes how the camera moves farther from its initial position. 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 $Y(t)$ does 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. The demonstration is as follows. Let us look at for example at the derivative of $X(t)$

$$\dot{X}(t) = \frac{\lambda e_3(0)\exp^{\lambda t}}{h(0)^2 a(t)} \left(b^2 Z^* \left(\exp^{\lambda t} - 1 \right) + 2 Z^* \left(2\exp^{\lambda t} - a(t) \right) + e_1(0) L \left(2 - a(t) \right) \right) \tag{151}$$

noting that the sign depends on $e_3(0)$ and will not change if

$$b^2 Z^* \left(\exp^{\lambda t} - 1\right) + 2Z^* \left(2\exp^{\lambda t} - a(t)\right) + e_1(0)L\left(2 - a(t)\right) \geq 0 \tag{152}$$

By using the definition of $e_1(t)$ in (121) the above condition can be rewritten as

$$b^2 Z^* \left(\exp^{\lambda t} - 1\right) + 4Z^* \left(\exp^{\lambda t}\right) - \frac{2D(0)}{C(0)}\left(2 - a(t)\right) \geq 0 \tag{153}$$

which is always true since $D(0) < 0$, $C(0) > 0$, $Z^* > 0$, $\exp^{\lambda t} \geq 1$ and $a(t) \in [2, \infty)$ (as shown in Appendix D.2).

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

The final coordinates of $\mathbf{X}$ are obtained by calculating the limit when time approaches infinity

$$\lim_{t\to\infty} X(t) = e_3(0)G \tag{154}$$

$$\lim_{t\to\infty} Y(t) = e_2(0)G \tag{155}$$

$$\lim_{t\to\infty} Z(t) = \frac{e_1(0)L - 2Z^* + h(0)Z^*}{h(0)} \tag{156}$$

with

$$G = \frac{1}{2b^3 h(0)}\left(-2b^2 Z^*\left(h(0)\arctan\left(\frac{2}{b}\right) + b\right) + \pi b^2 Z^* h(0) + 2be_1(0)L\left(2 - h(0)\right)\right) \tag{157}$$

Note that by using the definition of e_1 in function of the object parameters in (121) and that $h = 2/C$ we have

$$\lim_{t\to\infty} Z(t) = D(0) + Z^* \tag{158}$$

where $D(0)$ is the initial distance between the object plane and the camera origin.

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

8.3 Local asymptotic stability analysis under laser-cross misalignment

The closed-loop equation of the system in presence of calibration errors becomes strongly coupled so that it is not possible to develop the global asymptotic stability analysis under these conditions. We present instead the local asymptotic stability analysis when the laser-cross is not aligned with the camera frame.

8.3.1 Misalignment consisting of a translation

By using the model parameters in Table 3 we can calculate the interaction matrices of the inverses of y_1, x_2, y_3 and x_4 for the desired state taking into account that the laser-cross is displaced ${}^{C}\mathbf{T}_{L} = (t_x, t_y, t_z)$ from the camera origin. Then, the interaction matrix of $\mathbf{s}$ can be computed obtaining

$$\mathbf{L_s}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & \frac{2L}{t_y^2-L^2} & 0 & -\frac{2Lt_x}{t_y^2-L^2} & 0 \\ 0 & 0 & -\frac{2t_y}{t_y^2-L^2} & -2 & \frac{2t_xt_y}{t_y^2-L^2} & 0 \\ 0 & 0 & -\frac{2t_x}{t_x^2-L^2} & -\frac{2t_xt_y}{t_x^2-L^2} & 2 & 0 \end{pmatrix} \tag{159}$$

The local asymptotic stability analysis of the system under this type of misalignment consists on studying the product of matrices $\mathbf{M}(\mathbf{e}^*) = \mathbf{L_s}(\mathbf{e}^*)\widehat{\mathbf{L_s}}^{+}$ which is

$$\mathbf{M}(\mathbf{e}^*) = \begin{pmatrix} \frac{L^2}{L^2-t_y^2} & 0 & \frac{t_xL}{L^2-t_y^2} \\ -\frac{t_yL}{L^2-t_y^2} & 1 & -\frac{t_xt_y}{L^2-t_y^2} \\ -\frac{t_xL}{L^2-t_x^2} & -\frac{t_xt_y}{L^2-t_x^2} & 1 \end{pmatrix} \tag{160}$$

The eigenvalues of $\mathbf{M}(\mathbf{e}^*)$ are

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

so that imposing the positiveness of σ_1 we have that $L^2 - t_y^2 > 0$ which means that

$$|t_y| < L \tag{161}$$

When imposing the positiveness of σ_2 and σ_3 we must deal with two hypothesis, one assuming $L^2 - t_x^2 > 0$ and the other $L^2 - t_x^2 < 0$. Let us develop both hypothesis:

- *hypothesis 1*: $L^2 - t_x^2 > 0$. Imposing the positiveness of σ_2 and σ_3 according to this assumption leads to

  $$L^2 - t_x^2 > 0 \Rightarrow \mathrm{Re}(\sigma_2) = \mathrm{Re}(\sigma_3) = 1 \tag{162}$$

  since

  $$\sqrt{t_x^2\,(t_x^2 - L^2)} = \sqrt{t_x^2\,(-t_x^2 + L^2)}\sqrt{i} \tag{163}$$

- *hypothesis 2*: $L^2 - t_x^2 < 0$. In this case, imposing the positiveness according to the second hypothesis we obtain

$$L^2 - t_x^2 < 0 \Rightarrow \sigma_2 > 0 \Leftrightarrow L^2 - t_x^2 + \sqrt{t_x^2\,(t_x^2 - L^2)} < 0 \tag{164}$$

which is never true as can be seen by developing the condition as follows

$$\begin{aligned} \sqrt{t_x^2\,(t_x^2 - L^2)} &< -L^2 + t_x^2 \\ t_x^2\,(t_x^2 - L^2) &< (t_x^2 - L^2)^2 \\ t_x^2 &< t_x^2 - L^2 \\ 0 &< -L^2 \end{aligned} \tag{165}$$

Therefore, the right hypothesis is

$$L^2 - t_x^2 > 0 \tag{166}$$

which imposes that

$$|t_x| < L \tag{167}$$

Note that the stability domain when using these visual features is a little bit more restricted than when using directly image point coordinates and therefore, than when using the normalized area and angles approach. Concretely, the misalignment of the center of the laser-cross projected onto the camera plane $Z = 0$ must be included in the square circumscribed by the circle $t_x^2 + t_y^2 < 2L^2$, which was the error tolerated when using image points (see Fig. 17).

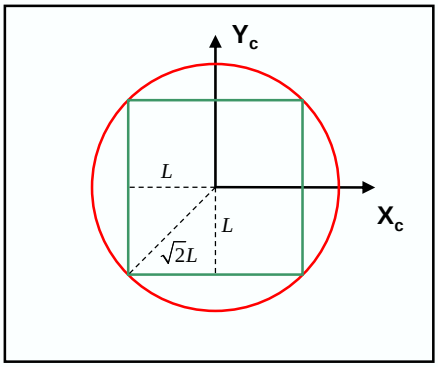

Figure 17: Local asymptotic stability areas for the projection of the laser-cross center onto the plane $Z = 0$.

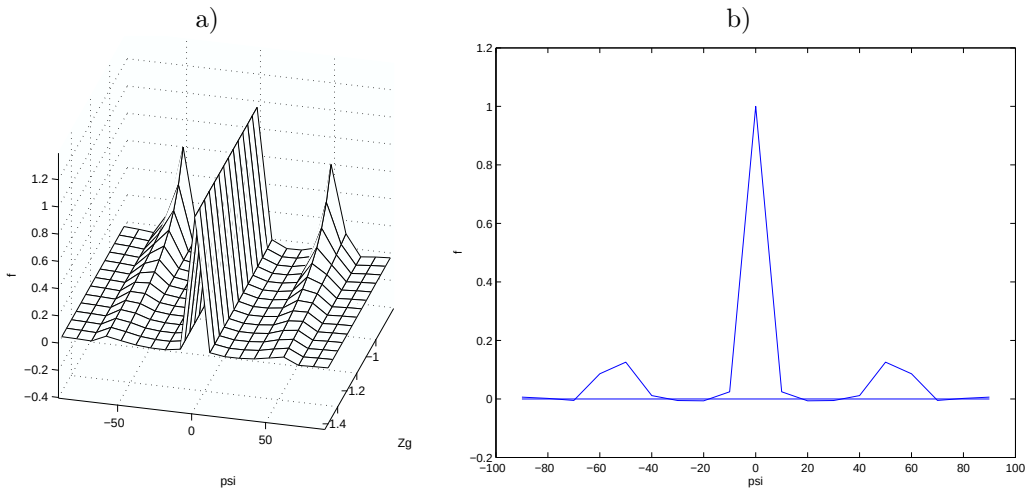

Figure 18: a) Rotation around X axis σ_2 in function of ψ (degrees) and $Zg = Z^*$ b) σ_2 in function of ψ (degrees) for the case $Zg = 1.1$ m

8.3.2 Misalignment consisting of individual rotations

First of all, let us consider a single rotation ψ of the laser-cross around the X axis of the camera frame. In this case, matrix $\mathbf{L_s}(\mathbf{e}^*)$ is calculated from the model parameters in Table 4. Then, the eigenvalues of $\mathbf{M} = \mathbf{L_s}(\mathbf{e}^*)\widehat{\mathbf{L_s}}^+$ are

$$\begin{aligned} \sigma_1 &= 1 \\ \sigma_2 &= f(L, Z^*, \psi) \\ \sigma_3 &= f(L, Z^*, \psi) \end{aligned}$$

the explicit expressions of σ_1 and σ_2 are too complex to be included here. In Fig. 18a the distribution of σ_2 in function of the rotation ψ and the depth to the object Z^* for $L = 0.15$ m is plotted. In Fig. 18b a particular case of σ_2 for $Z^* = 1.1$ m. As can be seen, the positiveness of σ_2 is ensured for almost all angle values. The same plots are shown in Fig. 19 for σ_3. As can be seen, the positiveness of this eigenvalue is not always ensured and depends on the rotation angle.

In the case of a rotation θ of the laser-cross around the Y axis of the camera, also two complex eigenvalues appear, whose distributions are plotted in Fig. 20 and Fig. 21, showing that the rotation around the Y axis is better tolerated by the system.

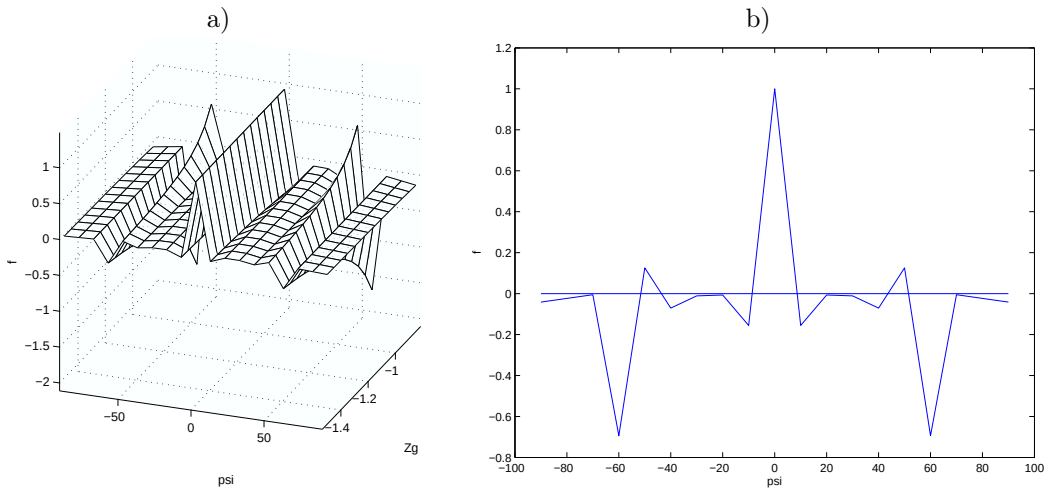

Figure 19: Rotation around X axis a) σ_3 in function of ψ (degrees) and $Zg = Z^*$ b) σ_3 in function of ψ (degrees) for the case $Zg = 1.1$ m

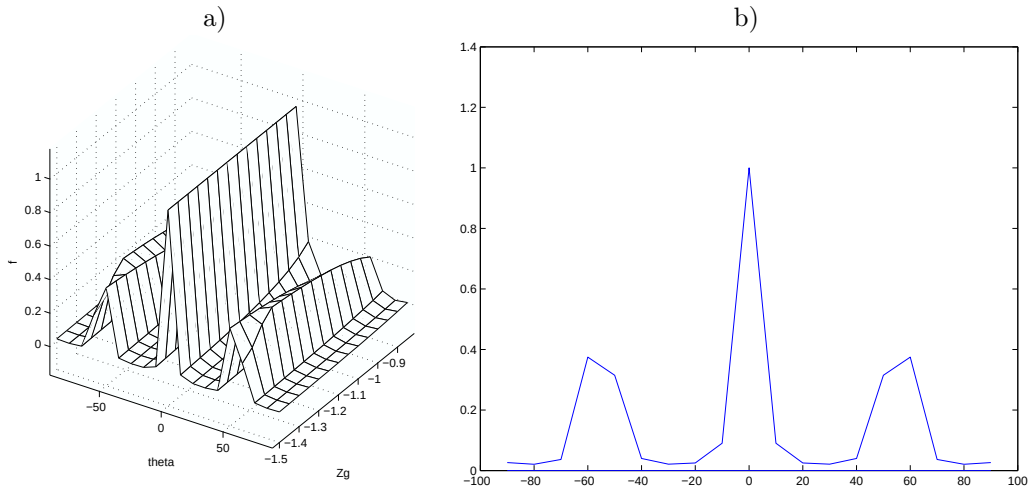

Figure 20: Rotation around Y axis a) σ_2 in function of θ (degrees) and $Zg = Z^*$ b) σ_2 in function of θ (degrees) for the case $Zg = 1.1$ m

Finally, if a rotation ϕ around the Z axis of the camera is applied to the laser-cross, the eigenvalues obtained from $\mathbf{M}$ are

$$\sigma_1 = \frac{1}{\cos\phi}$$

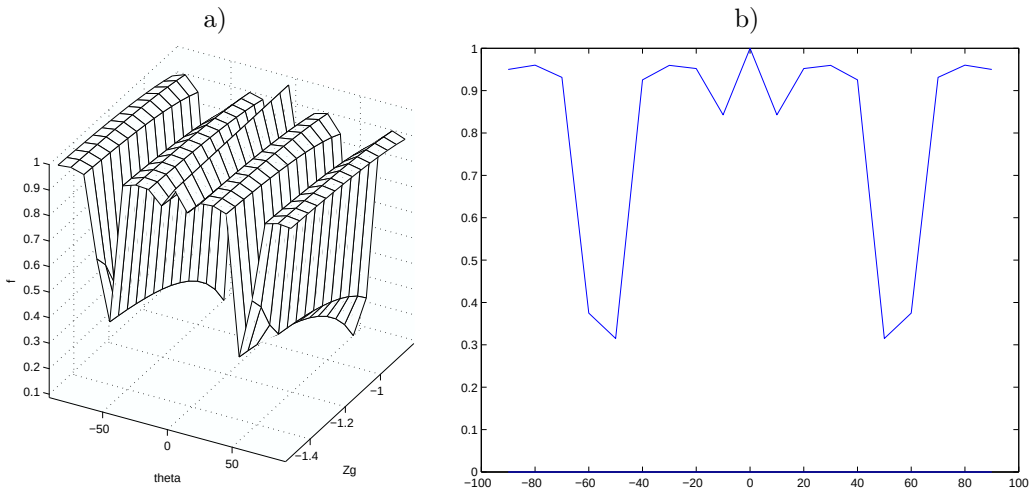

Figure 21: Rotation around Y axis a) σ_3 in function of θ (degrees) and $Zg = Z^*$ b) σ_3 in function of θ (degrees) for the case $Zg = 1.1$ m

$$\begin{aligned} \sigma_2 &= 1 + i \tan\phi \\ \sigma_2 &= 1 - i \tan\phi \end{aligned} \tag{168}$$

imposing the positiveness of the first one, we have that the rotation must be included in $\phi \in [-\pi/2,\ \pi/2]$. Note that the other eigenvalues are complex numbers and that their real part is always positive. Therefore, rotation of the laser-cross around the optical axis of the camera does not affect the local asymptotic stability of the system.

In summary, the approach based on this non-linear combination of image points is less robust against individual rotations of the laser-cross than the image points based approach and the normalized area and angles approach. In the following section we show how to overcome this problem by improving the set of visual features.

8.4 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 of visual features $\mathbf{s}'$ which is analytically and experimentally robust against laser-cross misalignment. Fig. 22 shows the image point distribution in the desired state when different types of misalignment take place (the 4 lasers have the same relative orientation). As can be seen, a general misalignment of the laser-cross produces that the polygon enclosing the 4 points in the desired image appears rotated and translated from the image center as shown in Fig. 23a.

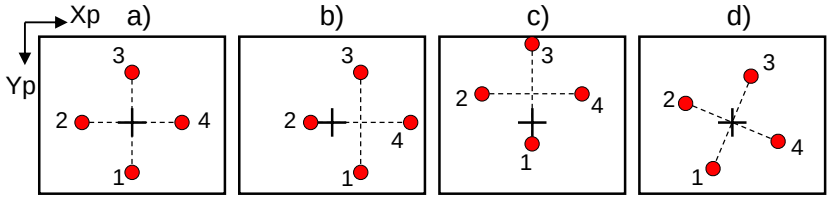

Figure 22: Effects of laser-cross misalignment in the desired image. a) Ideal image. b) The laser-cross is horizontally displaced or rotated around Y_C. c) The laser-cross is vertically displaced or rotated around Y_C. d) Laser-cross rotated around Z_C.

In fact, the set of visual features $\mathbf{s} = (a_n,\ \alpha_{13n},\ \alpha_{24n})$ is robust against laser-cross misalignment since both the area and the angles are invariant to the location and orientation of the polygon enclosing the 4 points in the image. Therefore, the corrected set of visual features $\mathbf{s}'$ must be also unaffected by this type of planar transformation. The idea consists of defining an image point transformation composed of a planar transformation and a translation which minimizes the misalignment observed in the image. This image transformation will be constrained as follows: in absence of laser-cross misalignment, the corrected set of visual features $\mathbf{s}'$ must be equal to the uncorrected one $\mathbf{s}$. Thus, in the ideal case the results concerning the global asymptotic stability and camera trajectory concerning $\mathbf{s}$ will also hold for $\mathbf{s}'$.

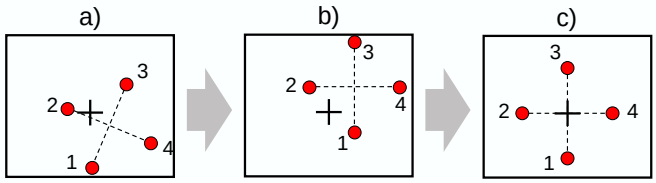

Figure 23: Image points correction. a) Desired image under a general misalignment of the laser-cross. b) Image points after applying the transformation $\mathbf{T}$. c) Image points after transformation $\mathbf{T}$ and translation $-\mathbf{x}_g$.

First of all, we eliminate the misalignment exhibited by the polygon in Fig. 22d which is produced when the laser-cross is rotated around the optical axis. Let us define the following unitary vectors

$$\mathbf{x}_{42} = \begin{pmatrix} x_{42} \\ y_{42} \end{pmatrix} = \frac{\mathbf{x}_4^* - \mathbf{x}_2^*}{\|\mathbf{x}_4^* - \mathbf{x}_2^*\|}, \quad \mathbf{x}_{13} = \begin{pmatrix} x_{13} \\ y_{13} \end{pmatrix} = \frac{\mathbf{x}_1^* - \mathbf{x}_3^*}{\|\mathbf{x}_1^* - \mathbf{x}_3^*\|} \tag{169}$$

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

$$\mathbf{T} = \begin{bmatrix} \mathbf{x}_{24} & \mathbf{x}_{13} \end{bmatrix}^{-1} = \frac{1}{x_{42}y_{13} - x_{13}y_{42}} \begin{pmatrix} y_{13} & -x_{13} \\ -y_{42} & x_{42} \end{pmatrix} \tag{170}$$

is defined so that $\mathbf{T}$ uses the desired image points in order to align 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 the transformed image points as follows

$$\mathbf{x}_i'' = \mathbf{T}\mathbf{x}_i \tag{171}$$

The result of applying the transformation matrix $\mathbf{T}$ to the misaligned image points of Fig. 23a is shown in Fig. 23b. Then, it only rests to define a translation vector which is able to center the polygon in the image (see Fig. 23c). The most intuitive choice is the center of gravity of the polygon $\mathbf{x}_g$. However, the choice of a suitable expression for $\mathbf{x}_g$ is not trivial as it could be supposed. First of all, $\mathbf{x}_g$ must be computed from the current image. Secondly, we remember that in absence of laser-cross misalignment $\mathbf{s}'$ must be equal to $\mathbf{s}$ so that $\mathbf{x}_i' = \mathbf{x}_i'' = \mathbf{x}_i$. Thus, under ideal conditions $\mathbf{T}$ must be the identity, which is true according to Eq. (170), and $\mathbf{x}_g$ must be $\mathbf{0}$. Thus, we could intuitively set $\mathbf{x}_g = (1/4)(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4)$. However, according to the ideal model parameters (see Table 1), the general expression of the laser image points in function of the object parameters are

$$\begin{array}{rclrcl} x_1 &=& 0 & y_1 &=& -\dfrac{LC}{BL + D} \\ x_2 &=& -\dfrac{LC}{AL - D} & y_2 &=& 0 \\ x_3 &=& 0 & y_3 &=& -\dfrac{LC}{BL - D} \\ x_4 &=& -\dfrac{LC}{AL + D} & y_4 &=& 0 \end{array} \tag{172}$$

Therefore, we have that

$$\frac{1}{4}(\mathbf{x}_1'' + \mathbf{x}_2'' + \mathbf{x}_3'' + \mathbf{x}_4'') = -\frac{L^2AC}{2\left(A^2L^2 - D^2\right)} \tag{173}$$

which is only zero when the camera is parallel to the object. Instead of this, we propose to use

$$\mathbf{x}_g = \frac{1}{2} \begin{pmatrix} x_1'' + x_3'' \\ y_2'' + y_4'' \end{pmatrix} \tag{174}$$

Note that this expression is also a measure of the polygon center of gravity according to Fig. 23b. Moreover, in the ideal case $\mathbf{x}_g$ is actually $\mathbf{0}$ for any object pose. Then, the corrected image points are obtained as follows

$$\mathbf{x}_i' = \mathbf{T}\mathbf{x}_i - \mathbf{x}_g \tag{175}$$

The corrected set of visual features $\mathbf{s}'$ is therefore

$$\mathbf{s}' = \begin{pmatrix} y_1'^{-1} - y_3'^{-1} & y_1'^{-1} + y_3'^{-1} & x_2'^{-1} + x_4'^{-1} \end{pmatrix} \tag{176}$$

The global asymptotic stability of the ideal model is also ensured when using $\mathbf{s}'$. In the following sections the 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 $\mathbf{s}$. Certainly, 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 0 may be reached due to the laser-cross misalignment. Note that this problem does not longer appear in $\mathbf{s}'$ since the corrected image points are symmetrically distributed around the image center.

8.5 Robust visual features: local asymptotic stability analysis under laser-cross misalignment

The study of the global asymptotic stability of the system when using the corrected set of visual features $\mathbf{s}'$ and the constant control law is again too complex. As in the previous approaches, we instead analyze the local asymptotic stability in front of different types of laser-cross misalignment. Thus, we intend to prove the robustness of the new set of visual features in front of such calibration errors.

8.5.1 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 real interaction matrix for this laser-cross pose evaluated in the desired state $\mathbf{L}_{\mathbf{s}'}(\mathbf{e}^*)$ must be calculated. First we evaluate the interaction matrices in the desired state of the point coordinates y_1, x_2, y_3 and x_4 using the model parameters in Table 3 evaluated according to the desired state $A = B = 0$, $C = 1$ and $D = -Z^*$. These parameters are also used in order to calculate the $2D$ transformation defined in (170) and in (174) in the desired state. The expressions obtained for $\mathbf{T}$ and $\mathbf{x}_g$ are

$$\mathbf{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad x_g = \frac{t_x}{Z^*}, \quad y_g = \frac{t_y}{Z^*} \tag{177}$$

The interaction matrix of the corrected set of visual features in the desired state is then

$$\mathbf{L}_{\mathbf{s}'}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & -\frac{2}{L} & -\frac{4t_x}{L} & \frac{2t_x}{L} & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix} \tag{178}$$

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

$$\mathbf{M} = \begin{pmatrix} 1 & 2\frac{t_x}{L} & \frac{t_x}{L} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{179}$$

whose eigenvalues are the elements on the main diagonal which are all equal to 1. Therefore, the local asymptotic stability of the system in front of a displacement of the laser-cross is always ensured when using $\mathbf{s}'$ and the constant control law.

8.5.2 Misalignment consisting of individual rotations

We now present the local asymptotic stability analysis when the laser-cross is centered in the camera origin but it is rotated around one of the camera axis. Let us first consider a rotation $\psi \in (-\pi/2, \pi/2)$ around the X axis. The $2D$ transformation based on $\mathbf{T}$ and $\mathbf{x}_g$ can be obtained from the model parameters in Table 4.

$$\mathbf{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad x_g = 0, \quad y_g = \frac{\sin\psi}{\cos\psi} \tag{180}$$

and the interaction matrix in the desired state taking into account this laser-cross pose is

$$\mathbf{L}_{\mathbf{s}'}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & -\frac{2\cos\psi}{L} & -\frac{-2Z^*\cos\psi\sin\psi}{L\cos\psi} & 0 & 0 \\ 0 & 0 & 0 & -\frac{2\cos\psi}{\cos\psi} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix} \tag{181}$$

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

$$\mathbf{M} = \begin{pmatrix} \cos\psi & \frac{Z^*\cos\psi\sin\psi}{L\cos\psi} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{182}$$

Note that all the eigenvalues (in this case the elements of the main diagonal) are positive since $\psi \in (-\pi/2, \pi/2)$.

In case that the laser-cross is rotated an angle $\theta \in (-\pi/2, \pi/2)$ around the Y camera axis (model parameters in Table 5), the $2D$ transformation in the desired state is

$$\mathbf{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad x_g = -\frac{\sin\theta}{\cos\theta}, \quad y_g = 0 \tag{183}$$

while the interaction matrix in the desired state taking into account this laser-cross pose is

$$\mathbf{L}_{\mathbf{s}'}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & -\frac{2}{L} & 0 & -\frac{-2Z^* \sin\theta}{L\cos\theta} & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2\cos\theta}{\cos\theta} & 0 \end{pmatrix} \tag{184}$$

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

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & -\frac{Z^* \sin\theta}{L\cos\theta} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{185}$$

Note that all the eigenvalues are also positive.

Finally, let us study the case when the laser-cross is rotated an angle ϕ around the optical axis of the camera (model parameters in Table 6. In this case we have that

$$\mathbf{T} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \quad x_g = 0, \quad y_g = 0 \tag{186}$$

$$\mathbf{L}_{\mathbf{s}'}(\mathbf{e}^*) = \begin{pmatrix} 0 & 0 & -\frac{2}{L} & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\cos\phi & 2\sin\phi & 0 \\ 0 & 0 & 0 & 2\sin\phi & 2\cos\phi & 0 \end{pmatrix} \tag{187}$$

while the product of matrices in the closed-loop equation of the system is

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix} \tag{188}$$

The eigenvalues are

$$\begin{aligned} \sigma_1 &= 1 \\ \sigma_2 &= \cos\phi + \sqrt{\cos^2\phi - 1} \\ \sigma_3 &= \cos\phi - \sqrt{\cos^2\phi - 1} \end{aligned}$$

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

Therefore, the system based on the corrected set of visual features $\mathbf{s}'$ is locally asymptotically stable if the laser-cross is not aligned with the camera frame.

8.6 Robust visual features: local asymptotic stability analysis in presence of camera calibration errors

We now present the local asymptotic stability analysis in presence of calibration errors in the intrinsic parameters of the camera. as explained in Section 6.3, it is necessary to study the closed-loop equation of the measured task function $\widetilde{\mathbf{e}}$ that in this case is

$$\dot{\widetilde{\mathbf{e}}}^* = -\lambda \mathbf{L}_{\widetilde{\mathbf{s}}}(\widetilde{\mathbf{e}}^*)\widehat{\mathbf{L}_{\mathbf{s}}}^{+}\widetilde{\mathbf{e}}^* \tag{189}$$

The interaction matrix in the desired state $\widetilde{\mathbf{e}}^*$ is

$$\mathbf{L}_{\widetilde{\mathbf{s}}}(\widetilde{\mathbf{e}}^*) = \begin{pmatrix} 0 & 0 & -\frac{2K_v}{L} & 0 & 0 & 0 \\ 0 & 0 & 0 & -2K_v & 0 & 0 \\ 0 & 0 & 0 & 0 & 2K_u & 0 \end{pmatrix} \tag{190}$$

so that the product of matrices $\mathbf{M}(\widetilde{\mathbf{e}}^*) = \mathbf{L}_{\mathbf{s}}(\widetilde{\mathbf{e}}^*)\widehat{\mathbf{L}_{\mathbf{s}}}^{+}$ is

$$\mathbf{M}(\widetilde{\mathbf{e}}^*) = \begin{pmatrix} K_v & 0 & 0 \\ 0 & K_v & 0 \\ 0 & 0 & K_u \end{pmatrix} \tag{191}$$

Note that the eigenvalues are the elements of the main diagonal and are positive if $K_u > 0$ and $K_v > 0$, which is true if and only if $\widetilde{f} > 0$, $\widetilde{k}_u > 0$ and $\widetilde{k}_v > 0$. Similarly to the previous image-based approaches, the system based on $\mathbf{s}'$ is also robust against camera calibration errors if the elements of the main diagonal of $\widetilde{\mathbf{A}}$ are positive.

8.7 Simulation results

We now present the simulation results obtained by the corrected set of visual features $\mathbf{s}'$ according to the conditions described in Section 5.1.

8.7.1 Ideal system

As has been shown, under ideal conditions, the interaction matrix of $\mathbf{s}'$ can be evaluated from the task function. The results obtained with the non-constant control law based on a perfect estimation of $\widehat{\mathbf{L}_{\mathbf{s}}}$ are plotted in Fig. 24. Both an exponential decrease of the task function and a monotonic behavior of the camera velocities are observed. Furthermore, the camera trajectory is almost a straight line in the space. Note that, as expected, the results coincide with the ones obtained by the position-based approach presented in Section 5.1.1. Unlike the position-based approach, this image-based approach does not require the minimization of the non-linear equations.

The behavior of the system when using $\mathbf{s}'$ and the constant control law based on $\mathbf{L}_{\mathbf{s}}^*$ is shown in Fig. 25. Note that both the task function components and the camera velocities are strictly monotonic as expected from the analytic results. We remark that these results

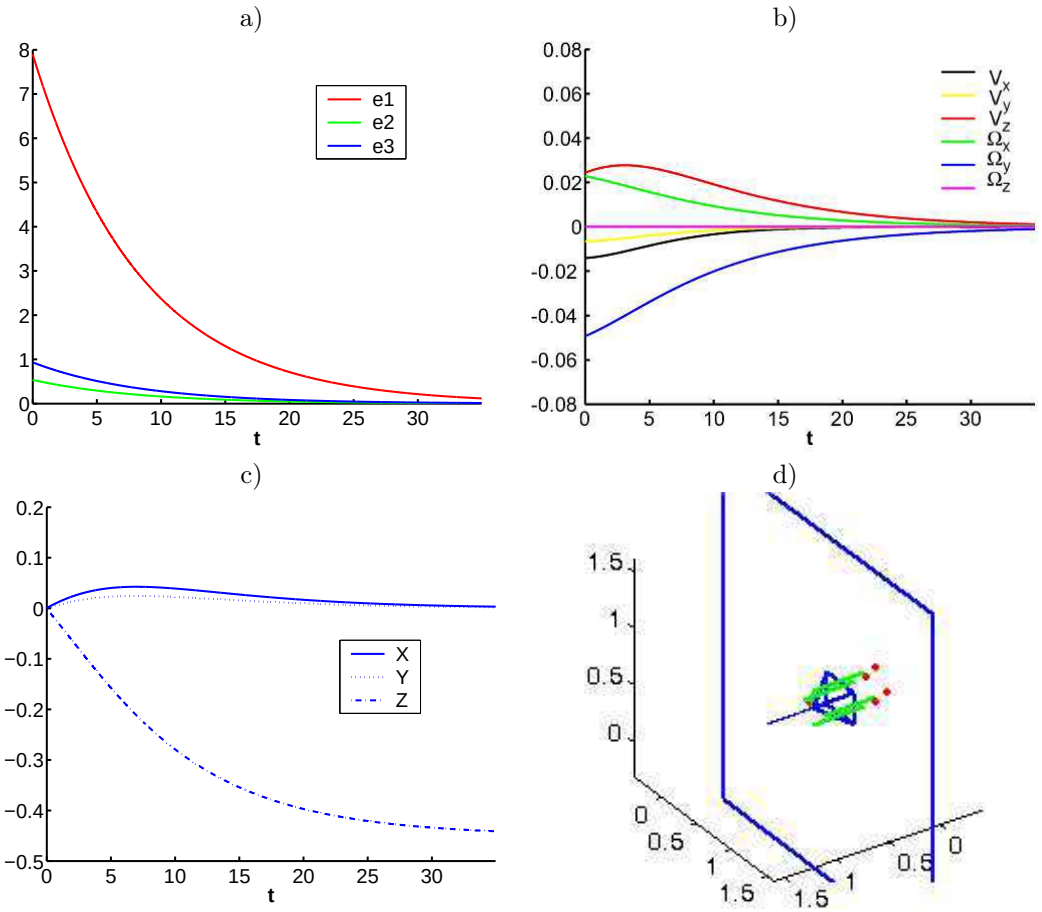

Figure 24: Ideal system: simulation using $\mathbf{s}'$ and the non-constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

also coincide with the ideal behavior of the position-based approach based on the constant control law presented in Section 5.1.1. In addition to this, the results of $\mathbf{s}'$ are also very similar to the ones given by the image-based approach using $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$. This result was already expected from the Taylor approximations shown in (120).

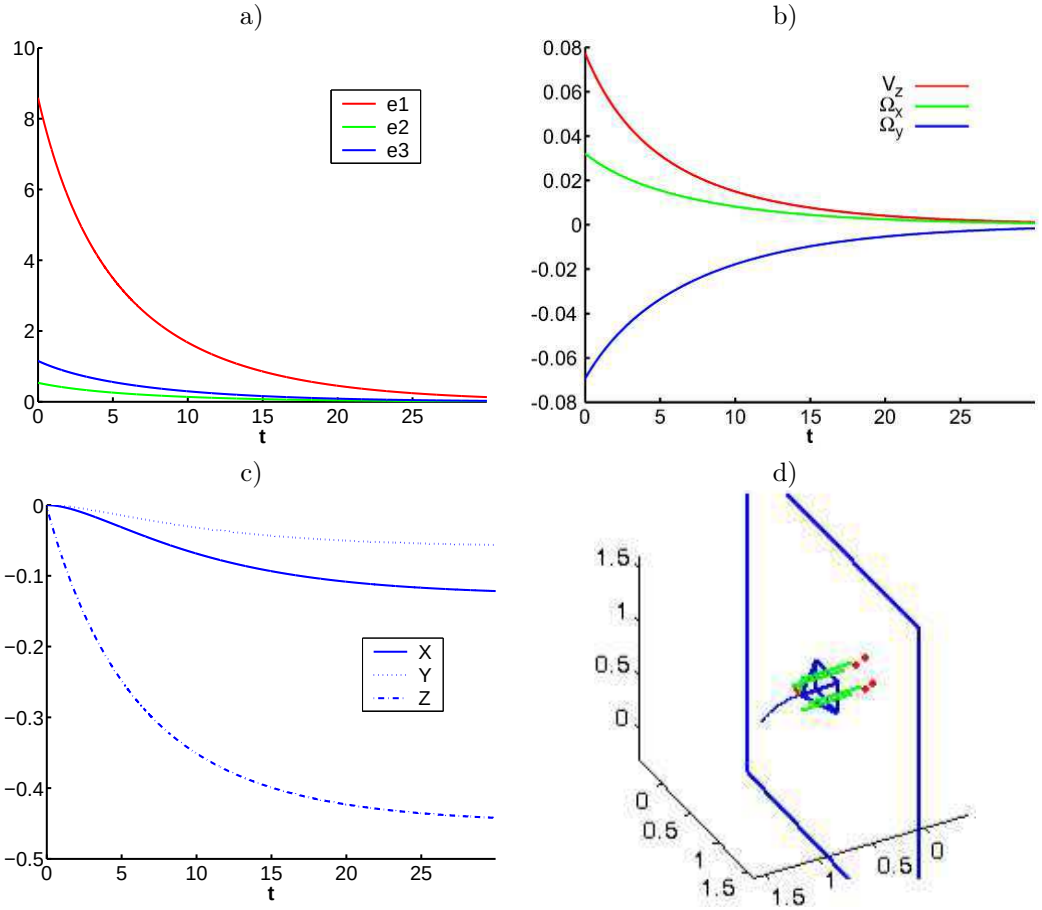

Figure 25: Ideal system: simulation using $\mathbf{s}'$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

Remember that the behavior of the system when using the constant control law under ideal conditions can also be obtained from the analytic expressions of the task function $\mathbf{e}(t)$, the camera velocities $V_z(t)$, $\Omega_x(t)$, $\Omega_y(t)$ and the trajectory $\mathbf{X}(t)$ developed in Appendix D.2

and in Section 8.2. Given the initial object pose parameters $\underline{\mathbf{n}}$ and D and the desired state defined by $\underline{\mathbf{n}} = (0, 0, 1)$, $D = -Z^*$, the task function components at $t = 0$ can be evaluated by using Equation (248), Equation (243) and Equation (244). The initial conditions found are the following

$$\begin{aligned} e_1(0) &= 8.5953 \\ e_2(0) &= 0.5359 \\ e_3(0) &= 1.1547 \end{aligned}$$

The functions $e_1(t)$, $e_2(t)$ and $e_3(t)$ and the corresponding camera velocities $V_z(t)$, $\Omega_x(t)$, $\Omega_y(t)$ in (147) have been evaluated in the interval $t \in [0, 30]$ s and are plotted in Fig. 26a-b. The coordinates of the initial position expressed in the camera frame obtained analytically in (150) have been also evaluated in the same interval. The resulting curves are plotted in Fig. 26c and the trajectory of this fixed point in the camera frame is shown in Fig. 26d. Note that the task function decrease, the camera velocities and the trajectory predicted by the analytic model coincide with the simulation results in Fig. 25.

8.7.2 System including laser-cross misalignment and image noise

The behavior of the system when using $\mathbf{s}'$ and the non-constant control law in presence of large calibration errors and image noise is shown in Fig. 27. Note that the system is almost unaffected by the calibration errors. We find again that the results are nearly the same that the ones obtained by the position-based approach.

When using the constant control law based on $\mathbf{L}_{\mathbf{s}}^*$, the system is also robust against the laser-cross misalignment, as expected from the local asymptotic stability analysis in presence of such errors. The results are plotted in Fig. 28. Note that under laser-cross misalignment the simulation results of this approach are still pretty similar to the ones obtained by $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$.

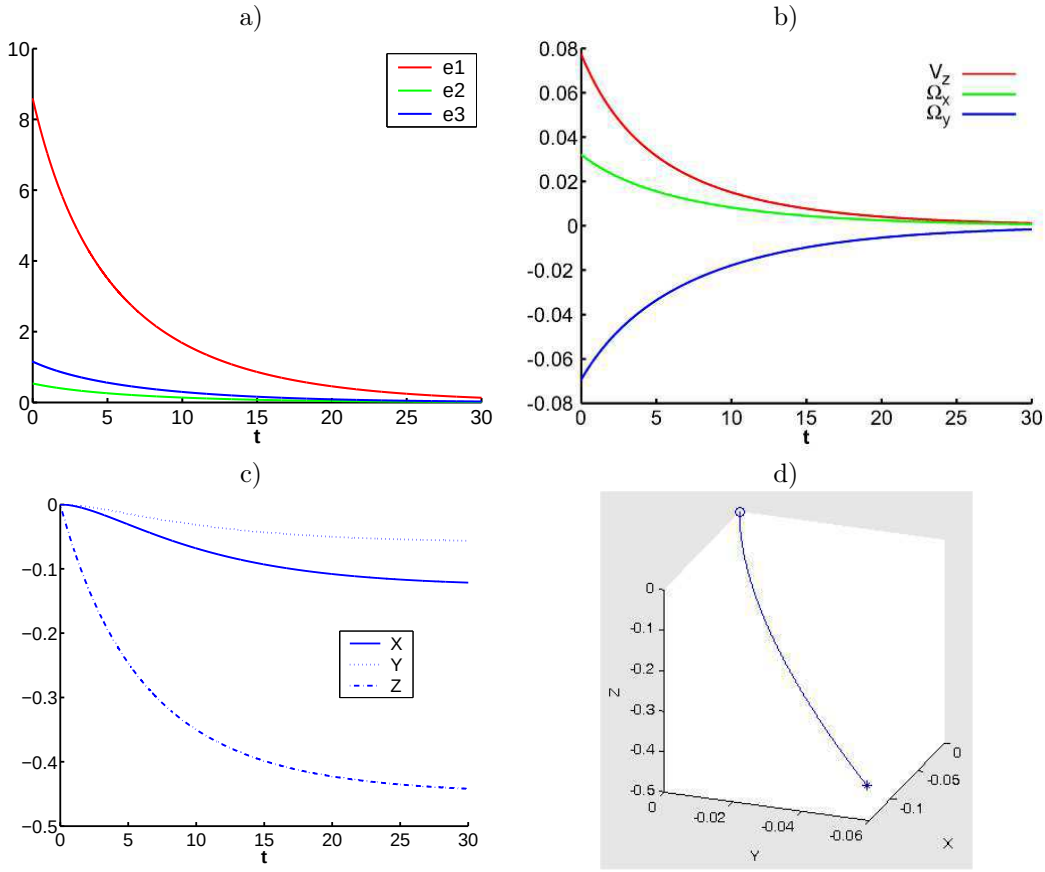

Figure 26: Ideal system: analytic behavior when using $\mathbf{s}'$ and the constant control law. a) $e_1(t)$, $e_2(t)$ and $e_3(t)$ evaluated at $t \in [0, 30]$ s. b) Camera velocities $V_z(t)$ (m/s), $\Omega_x(t)$ and $\Omega_y(t)$ (rad/s) evaluated at $t \in [0, 30]$ s. c) Coordinates of a fixed point (initial position) in the camera frame (in m). d) 3D plot of the same point ('o' and '*' are the initial and the final point respectively.

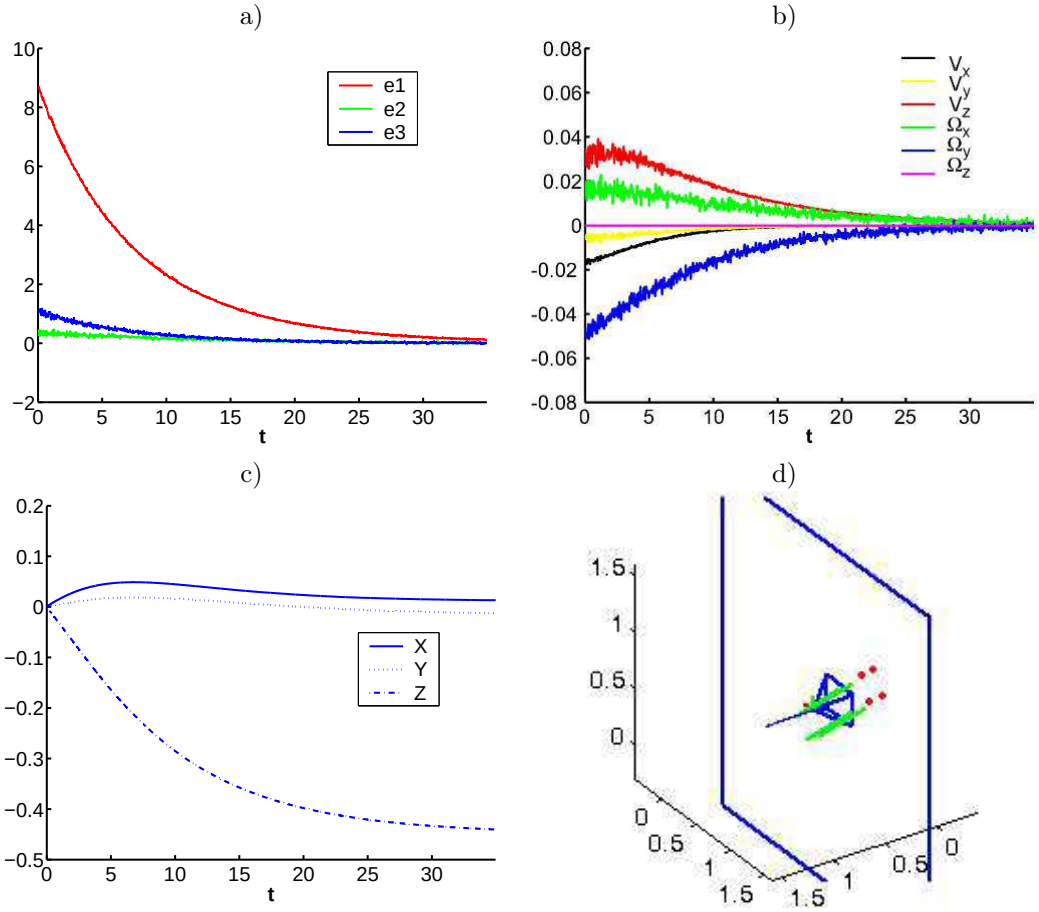

Figure 27: System including large laser-cross misalignment and image noise: simulation using $\mathbf{s}'$ and the non-constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

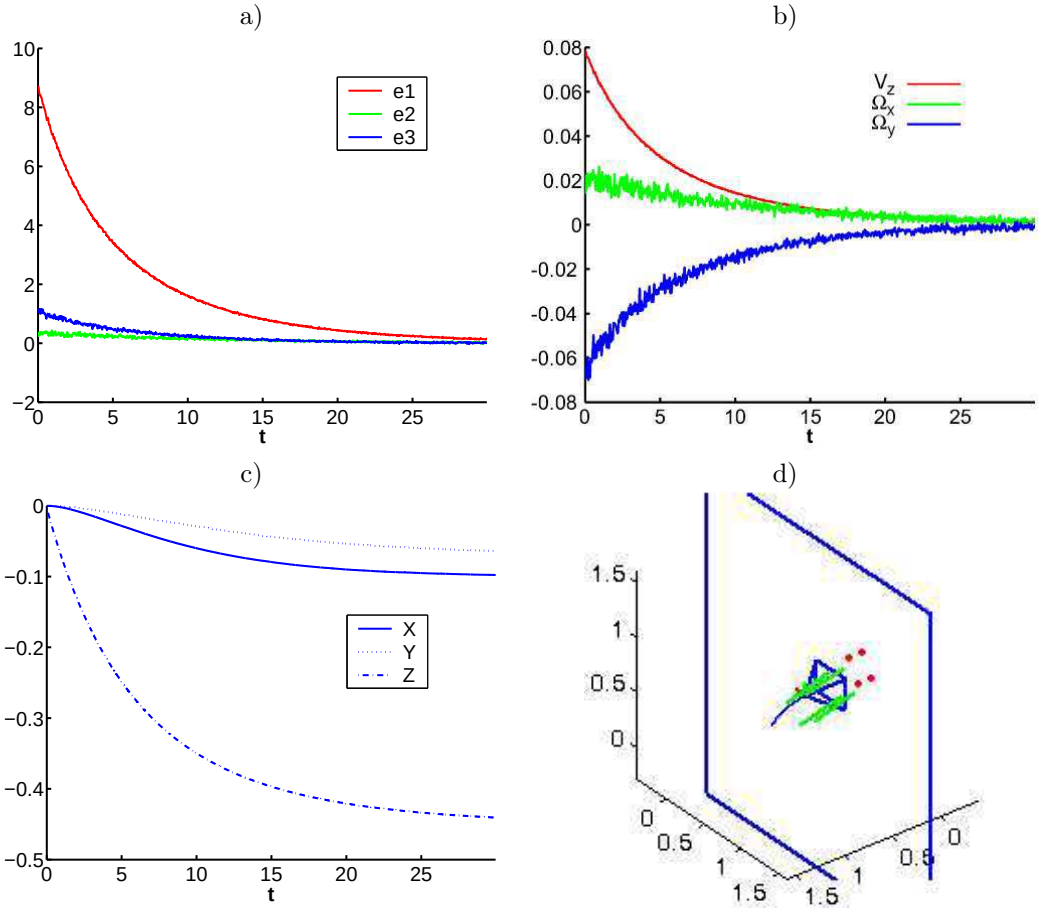

Figure 28: System including large laser-cross misalignment and image noise: simulation using $\mathbf{s}'$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Camera velocities (ms/s and rad/s) vs. time. c) Fixed point coordinates in the camera frame. d) Scheme of the camera trajectory.

9 Experimental results

In order to validate the theoretical results and to confirm the simulation results of the different approaches presented in this paper, real experiments have been carried out. The experimental setup consists of a six-degrees-of-freedom robot manipulator with a camera with focal length 8.5 mm coupled 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 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.

The aim of the experiments is to test the behavior of the control loop when both the laser-cross is positioned according to the ideal model specifications and when a large misalignment between the camera and the laser-cross takes place. Furthermore, during the real experiments a coarse calibration of the camera intrinsic parameters has been used and the direction of all the lasers is not exactly equal, so that the robustness of the approaches against this kind of modelling errors is also tested.

9.0.3 Laser-cross coarsely aligned with the camera

In the first experiment, the laser-cross has been approximately aligned with the camera. The desired and initial position of the camera is defined by $Z^* = 60$ cm, $D(0) = 105$ cm, $\alpha_x = -20°$ and $\alpha_y = 20°$. The image corresponding to the initial state is shown in Fig. 29a. On the other hand, Fig. 29b shows the trace of the laser points in the image from the initial state to the desired one.

Indeed, a perfect alignment of the camera and the laser-cross is not possible since we do not exactly know neither the camera origin location nor the orientation of its axis. This is evident by looking at the initial and desired images. As can be seen, the laser points do not exactly lie onto 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 neither possible to ensure that all the 4 lasers have the same exact orientation (which causes that the epipolar lines do not intersect in a unique point).

The position-based approach has not been implemented since it is equivalent to the image-based based approach based on $\mathbf{s}'$. Furthermore, the latter has the advantage that it is less time consuming since no numerical minimization of non-linear equations is required.

First of all, Fig. 30 presents the system response when using image points as visual features. As can be seen, since the laser-cross misalignment is small enough, the system converges showing a nice decrease on the visual feature errors and the norm of the task function (even if it is not a pure exponential decrease, as expected from the expression of $\mathbf{L}_\mathbf{s}^*$ which depends on $1/Z^*$). On the other hand, note that the camera velocities generated by the constant control law are not monotonic, specially the rotational ones.

Fig. 31 shows the results when using the set of visual features based on the normalized area and angles. As expected, both the task function and the camera velocities better fit

an exponential decrease. Furthermore, a linear mapping from task function space to camera velocities is almost exhibited.

Very similar results are obtained with the corrected version of the decoupled set of visual features $\mathbf{s}'$, which are presented in Fig. 32. In this case, the results when using the constant control law are plotted. We can observe the monotonic decrease of the task function and the camera velocities as predicted by the analytic model. No major differences are appreciated with respect to the approach based on normalized area and angles.

In Fig. 33 the results when using $\mathbf{s}'$ and the non-constant control law are shown. In this case, a pure exponential decrease of the task function is expected. Note however, that the actual behavior is not monotonic, which implies that this type of control law is a bit more sensitive to the lasers directions and the camera calibration errors. Nevertheless, the system converges with no major problems. Note also the non-monotonic camera velocities generated by the control law. We must also mention that when using the non-constant control law the computation time required at each iteration is higher since the pseudoinverse of the estimated $\widehat{\mathbf{L_s}}$ must be calculated.

9.0.4 Large misalignment between the camera and the laser-cross

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 (the rotation introduced about the X and Y axis are much smaller). Such a large misalignment is clearly observed in the initial and desired images shown in Fig. 34.

Under these conditions, only the approaches based on normalized area and angles, and the corrected version of the decoupled set of visual features have succeeded. On the other hand, the image-based approach based on $\mathbf{s} = (y_1, x_2, y_3, x_4)$ has diverged as expected from the simulation results. In Fig. 35 the results when using the constant control law based on $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$ are shown. On the other hand, Fig. 36 presents the results when using $\mathbf{s}'$. As can be seen, even with such a large misalignment of the laser-cross, both approaches still obtain almost a monotonic decrease in the task function as well as an almost monotonic decrease in the camera velocities. Therefore, the large convergence domain of these approaches expected from the analytic results is here confirmed.

When using $\mathbf{s}'$ and the non-constant control law, the system has not been able to converge to the desired position since the robot has reached a joint limit. This fact has been also observed when using other initial positions. It seems that the non-linearities in the camera velocities produced by such a control law become stronger due to errors in the lasers directions when the laser-cross is largely misaligned. Thus, some of the demanded robot motions are unfeasible or usually bring the robot very close to some joint limits. Therefore, we confirm that designing decoupled visual features which vary proportionally to the corresponding controlled degree of freedom is a good strategy to obtain suitable camera trajectories [38].

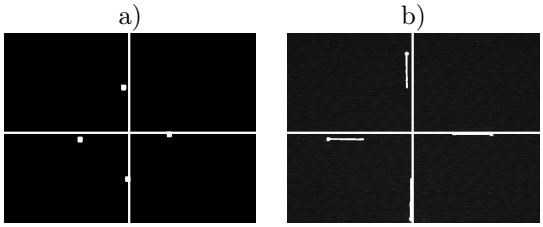

Figure 29: Experiment using a coarse alignment. a) Initial image. b) Final image with the trace of each laser point from its initial position to its final position.

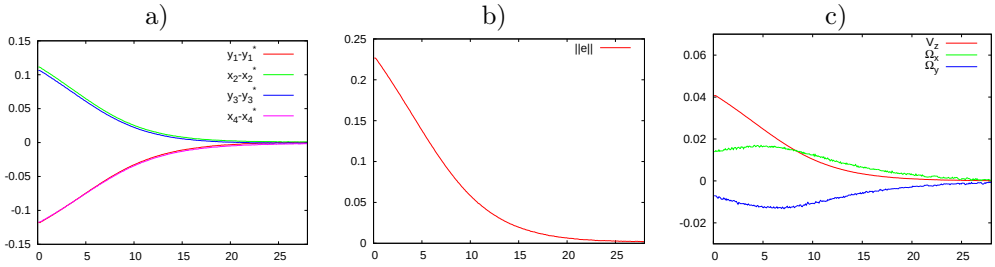

Figure 30: Coarse alignment: experiment using $\mathbf{s} = (y_1, x_2, y_3, x_4)$ and the constant control law. a) $\mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

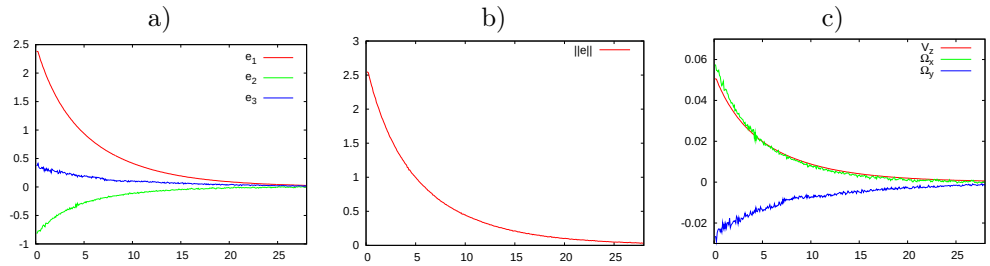

Figure 31: Coarse alignment: experiment using $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

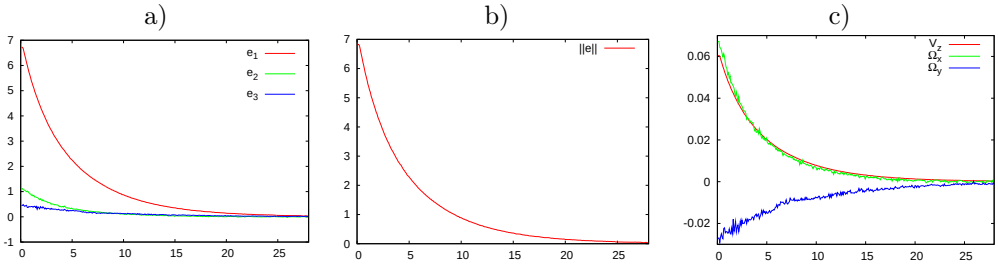

Figure 32: Coarse alignment: experiment using $\mathbf{s} = ({y'_1}^{-1} - {y'_3}^{-1}, {y'_1}^{-1} + {y'_3}^{-1}, {x'_2}^{-1} + {x'_4}^{-1})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

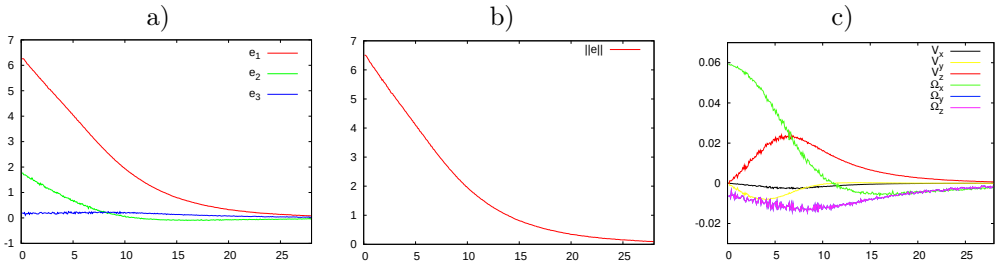

Figure 33: Coarse alignment: experiment using $\mathbf{s} = ({y'_1}^{-1} - {y'_3}^{-1}, {y'_1}^{-1} + {y'_3}^{-1}, {x'_2}^{-1} + {x'_4}^{-1})$ and the non-constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

10 Summary

In this section we briefly summarize the different approaches presented along the paper. The analytic results concerning the stability under ideal conditions and under calibration errors are remembered in Table 2. The results provided in the table take into account the control based on the interaction matrix evaluated in the desired state $\mathbf{L}_\mathbf{s}^*$. When using a non-constant control law, the global asymptotic stability has only been proven when the estimation of $\mathbf{L}_\mathbf{s}$ at each iteration is perfect, that is, when the conditions of the ideal model hold and there are no calibration errors.

Position-based approach: the input of the control law are the parameters of the object plane equation γ, β and α. The interaction matrix of such parameters shows a nice decoupling from the rotational to the translational part. Calculating these parameters requires to triangulate the depths of the image points assuming that all lasers have the same direction so

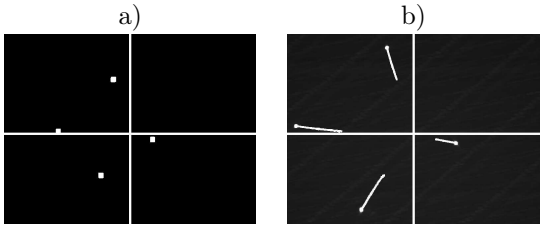

Figure 34: Experiment using a large misalignment. a) Initial image. b) Final image with the trace of each laser point from its initial position to its final position.

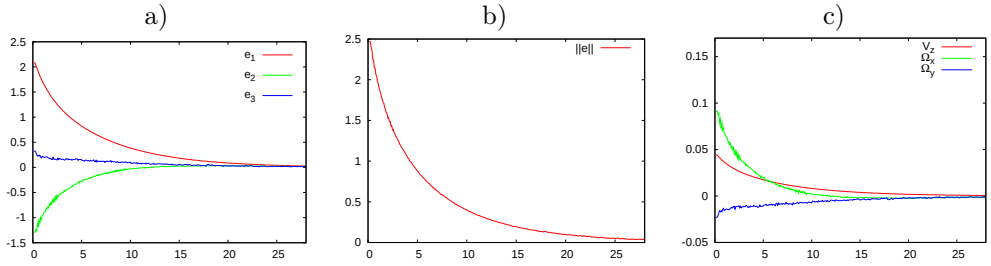

Figure 35: Large misalignment: experiment using $\mathbf{s} = (a_n, \alpha_{13n}, \alpha_{24n})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

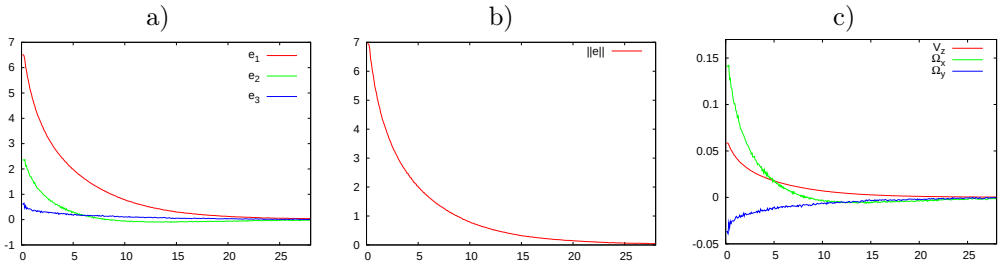

Figure 36: Large misalignment: experiment using $\mathbf{s} = ({y'_1}^{-1} - {y'_3}^{-1}, {y'_1}^{-1} + {y'_3}^{-1}, {x'_2}^{-1} + {x'_4}^{-1})$ and the constant control law. a) $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ vs. time (in s). b) Norm of the task function vs. time. c) Camera velocities (ms/s and rad/s) vs. time.

that the equation of the plane can be fitted at each iteration. The reconstructed depths can be used to estimate the interaction matrix at each iteration. In order to robustly estimate the depths of the projected points it is necessary to solve a system of non-linear equations. This can be done numerically with a minimization algorithm which can be sensible to fall into local minima. Nevertheless, the camera trajectory obtained is almost a straight line and the task function shows an exponential decrease under ideal conditions.

Image points approach: the inputs of the control law are the image point coordinates (y_1,x_2,y_3,x_4). The stability of the system is not ensured in presence of a displacement of the laser-cross with respect to the camera center. Furthermore, the camera velocities are non-monotonic.

Normalized area and angles approach: the input of the control law are the geometric image-based features $(a_n, \alpha_{13n}, \alpha_{24n})$, which have shown a strong robustness against calibration errors both analytically (through the local asymptotic stability analysis) and experimentally. Simulations and experiments have shown that both the task function and the camera velocities are monotonic. However, due to the complexity of the features no analytic results concerning the global asymptotic stability have been provided.

Decoupled features approach: the input of the control law are the image-based features

$$\mathbf{s} = \left(y_1^{-1} - y_3^{-1}, y_1^{-1} + y_3^{-1}, x_2^{-1} + x_4^{-1}\right) \tag{192}$$

This set of features decouples the rotational from the translational dof in all the workspace and the interaction matrix can be entirely expressed in terms of the task function approach. This is possible since these features are proportional to the object parameters γ, β and α. It has been possible to demonstrate the global asymptotic stability under ideal conditions when using both a constant control law based on $\mathbf{L}_\mathbf{s}^*$ and when using an estimation of $\mathbf{L}_\mathbf{s}$ (which does not require to reconstruct the object). When using the non-constant control law a pure exponential decrease of the task function is obtained under ideal conditions. In the case of the constant control law, a monotonic decrease is obtained for the visual features controlling the rotational subsystem, and the behavior of the feature controlling the depth is either monotonic or it presents a unique peak. The system is however quite sensitive to laser-cross misalignment.

Corrected decoupled features approach: a simple planar transformation applied to the image allows the robustness domain in presence of calibration errors to be enlarged. The corrected set of visual features $\mathbf{s}'$ has nice robustness properties and obtains the same results than the uncorrected version under ideal conditions.

Table 2: Stability analysis of the different approaches

Visual Features	Ideal system: stability		Laser-cross misalignment: local asymptotic stability				intrinsic errors: local asymptotic stability
	Global	Local	(t_x, t_y, t_z)	Rot(X,ψ)	Rot(Y,θ)	Rot(Z,ϕ)	
(γ, β, α)	$\surd$	$\surd$	?	?	?	?	?
(y_1, x_2, y_3, x_4)	?	$\surd$	$t_x^2 + t_y^2 < 2L^2$	$\surd$	$\surd$	$\surd$	$\surd$
$(a_n, \alpha_{13n}, \alpha_{24n})$	?	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$\mathbf{s}$	$\surd$	$\surd$	$\lvert t_{x,y} \rvert < L$	restricted	$\surd$	$\surd$	?
$\mathbf{s}'$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$

11 Conclusions

This paper presents a solution to the classic plane-to-plane positioning task from the combination of visual servoing and structured light. The projection of structured light not only simplifies the image processing but also allows the system to deal with low-textured objects lacking of visual features. A structured light sensor for eye-in-hand systems has been proposed. The sensor is based on four laser pointers attached to a cross-structure. Such a configuration has been chosen in order to obtain an optimal distribution of image points which is invariant to depth once the camera is parallel to the object. A position-based approach and several image-based approaches have been presented. The former is based on reconstructing the parameters of the plane equation of the object and has shown pretty decoupling and robustness against calibration errors. However, it requires to robustly reconstruct the parameters by solving a system of non-linear equations at each iteration. This process may be sensible to image noise depending on the numeric algorithm used. On the other hand, the image-based approaches have been analytically compared through stability analysis in front of different types of misalignments and camera calibration errors.

Two of the image-based approaches have shown a pretty robustness against calibration errors. The first is based on the area and several angles extracted from the polygon containing the four points in the image. Such features have been normalized in order to obtain a linear mapping from the task function space to the camera velocities near the desired state. Furthermore, they exhibit a nice decoupling near the desired state. The good performance of these features has been experimentally demonstrated. However, the high complexity of the analytic expressions of the features has avoided to obtain analytic results concerning either the global asymptotic stability or the camera trajectory generated by the constant control law based on the interaction matrix evaluated in the desired state.

On the other hand, a set of visual features based on non-complex non-linear combinations of the image point coordinates has also obtained very good performance. The advantage of these features is that they decouple rotational from translational dof in all the workspace

(not only around the desired state as in the previous approach). This decoupling is possible because these image-based features are proportional to the object plane parameters used in the position-based approach. By evaluating the interaction matrix for the desired state it can be seen that these features also produce a linear mapping from the task function space to the camera velocities around the desired state, without needing any normalization. Thanks to the decoupled form of the general interaction matrix, it has been possible to prove the global asymptotic stability under ideal conditions (so that it is sure that the system converges for any initial camera-object relative pose where the visibility constraint holds), and the analytic expression of the camera trajectory in the space. The advantage of this control law is that both the task function and the camera velocities are monotonic producing good camera trajectories. The main drawback of this approach is its sensitivity to large calibration errors like a large misalignment between the camera and the laser-cross. However, its robustness has been improved (the analytic proof has been provided) by defining a corrected version of the features based on a planar transformation applied to the image. The corrected visual features have obtained the same robustness against large laser-cross misalignment that the features based on the area and the angles. The camera velocities produced by these two techniques in the simulations are almost the same, since both sets of visual features are proportional when the camera is nearly parallel to the object. In the real experiments under large calibration errors similar performances have been obtained.

An interesting characteristic of the image-based approach based on the decoupled features and the position-based approach is that the interaction matrix can be estimated at each iteration from the feature vector. Thus, a non-constant law based on the estimated interaction matrix can be also used, obtaining a camera trajectory very similar to a straight line even in presence of small calibration errors. The image-based approach has the advantage that the feature vector is calculated only from image data and, unlike the position-based approach, it does not require to solve a non-linear system of equations. Experiments have shown that in presence of large calibration errors the robot is not always capable of reaching the desired position since the velocities produced by the control law become strongly non-linear. Therefore, it seems preferable to use the constant control law which produces almost strictly monotonic velocities in presence of large calibration errors.

Finally, we remark that the level of decoupling achieved in this work is due to the fact that the points are projected. Such decoupling has not been reached with visual features extracted from the object itself.

A Interaction matrix of μ

Samson et al. [48] modelled a general thin-field rangefinder which obtains a measure of the depth to an object along a straight line. According to our notation, the case studied by Samson et al. is represented in Fig. 37.

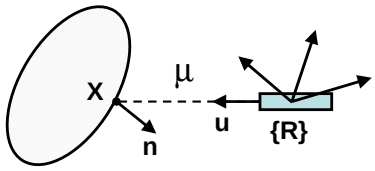

Figure 37: Thin-field rangefinder schema

Note that $\underline{\mathbf{n}}$ is the normal to the object at point $\mathbf{X}$. The measure of the depth that the sensor obtains is expressed in the rangefinder frame denoted as $\{R\}$. From the variation of the distance μ due to the sensor motion found by Samson et al., we can extract the following interaction matrix

$$ {}^{R}\mathbf{L}_{\mu} = -\frac{1}{\underline{\mathbf{n}}^{\top}\underline{\mathbf{u}}} \left(\underline{\mathbf{n}}^{\top} \mid \mu \left(\underline{\mathbf{u}} \times \underline{\mathbf{n}} \right)^{\top} \right) \tag{193} $$

where both $\underline{\mathbf{n}}$ and $\underline{\mathbf{u}}$ are expressed in the sensor frame.

In our case, when the system is composed by a camera and a laser pointer as shown in Fig. 38, both $\underline{\mathbf{n}}$ and $\underline{\mathbf{u}}$ are expressed in the camera frame, as well as the interaction matrix of μ.

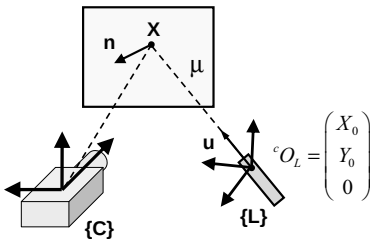

Figure 38: Our camera-laser system modelling

As shown in (13), the interaction matrix of μ expressed in the camera frame is

$$ {}^{C}\mathbf{L}_{\mu} = -\frac{1}{\underline{\mathbf{n}}^{\top}\underline{\mathbf{u}}} \left(\mathbf{n}^{\top} \mid (\mathbf{X} \times \underline{\mathbf{n}})^{\top} \right) \tag{194} $$

In order to demonstrate the equivalence with the interaction matrix by Samson et al. it is necessary to express $\mathbf{X}$, $\underline{\mathbf{n}}$ and $\underline{\mathbf{u}}$ in a laser frame with origin equal to $\mathbf{X}_0$. Then from (4) we have that point $\mathbf{X}$ in the laser frame is

$$\mathbf{X} = \mu \underline{\mathbf{u}} \tag{195}$$

so that (194) expressed in the laser frame becomes the same than the interaction matrix by Samson et al.

Andreff et al. also formulated the interaction matrix of μ [4]. In their case, the laser frame was chosen so that the Z axis coincides with the laser direction as shown in Fig. 39.

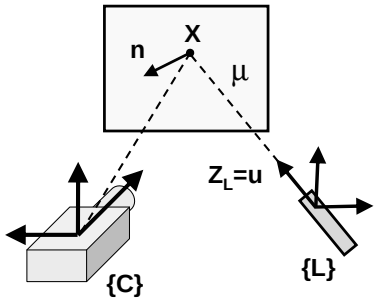

Figure 39: Camera-laser system modelling by Andreff et al.

The interaction matrix presented by Andreff et al. was expressed in function of α_x and α_y. The former is the angle between Z_L and $\mathbf{n}_x = (A, 0, C)$, being $\mathbf{n}_x$ the projection of $\underline{\mathbf{n}}$ to the plane $Y_L = 0$. Similarly, α_y is the angle between Z_L and $\mathbf{n}_y = (0, B, C)$ which is the projection of $\underline{\mathbf{n}}$ to the plane $X_L = 0$. The geometric interpretation of α_x and α_y is shown in Fig. 40. Taking into account the sign conventions of the angles and the constraint $C > 0$, we have that

$$\begin{aligned} A &= C \tan \alpha_x \\ B &= -C \tan \alpha_y \end{aligned} \tag{196}$$

so that expressing $C = \sqrt{1 - A^2 - B^2}$ the following relationship arise

$$\begin{aligned} A &= \frac{\tan \alpha_x}{\sqrt{\tan^2 \alpha_x + \tan^2 \alpha_y + 1}} \\ B &= -\frac{\tan \alpha_y}{\sqrt{\tan^2 \alpha_x + \tan^2 \alpha_y + 1}} \end{aligned} \tag{197}$$

The interaction matrix by Andreff et al. can be derived from the one by Samson et al. in (193). If the parameters of this interaction matrix are expressed in the laser frame

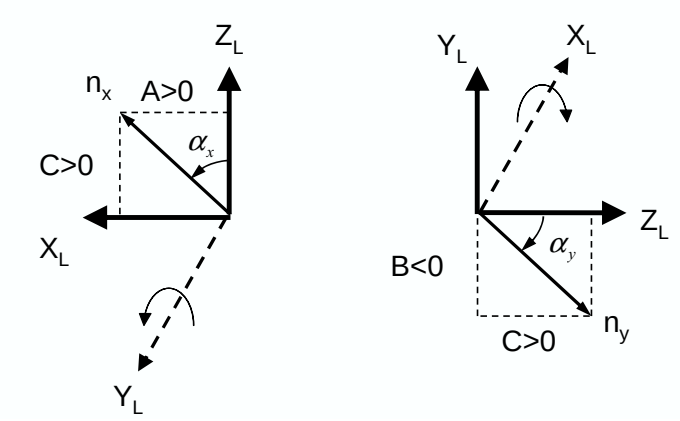

Figure 40: Geometric interpretation of α_x and α_y when they are positive.

proposed by Andreff et al. we have that $\underline{\mathbf{u}} = (0, 0, 1)$ and

$$\begin{aligned} \underline{\mathbf{u}} \times \underline{\mathbf{n}} &= \begin{vmatrix} \underline{\mathbf{i}} & \underline{\mathbf{j}} & \underline{\mathbf{k}} \\ 0 & 0 & 1 \\ A & B & C \end{vmatrix} = \begin{pmatrix} -B \\ A \\ 0 \end{pmatrix} \\ \underline{\mathbf{n}}^\top \underline{\mathbf{u}} &= C \end{aligned} \tag{198}$$

Then, the formula by Samson becomes

$$\mathbf{L}_\mu = \begin{pmatrix} -A/C & -B/C & -1 & \mu B/C & -\mu A/C & 0 \end{pmatrix} \tag{199}$$

so that taking into account the relationships in (196) we obtain

$$\mathbf{L}_\mu = \begin{pmatrix} -\tan\alpha_x & \tan\alpha_y & -1 & -\mu\tan\alpha_y & \mu\tan\alpha_x & 0 \end{pmatrix} \tag{200}$$

which is the form of the interaction matrix proposed by Andreff et al. [4].

B Model considering laser-cross misalignment

This appendix presents the model parameters when the laser-cross center is displaced from the camera origin and the orientation of the laser-cross frame is not the same that the camera frame. The camera intrinsic parameters are supposed to be perfectly calibrated. The laser-cross misalignment is modelled according to a homogeneous frame transformation matrix of the form

$${}^C\mathbf{M}_L = \begin{pmatrix} {}^C\mathbf{R}_L & {}^C\mathbf{T}_L \\ \mathbf{0}_3 & 1 \end{pmatrix} \tag{201}$$

which passes from points expressed in the structured light sensor frame to the camera frame (see Fig. 5).

First of all, the orientation vector and the reference point of every laser in the camera frame must be calculated taking into account the misalignment. We start from the values of these parameters in the laser-cross frame, which coincide with the ideal parameters shown in Table 1. Then, in the camera frame we have that for every laser

$$
\begin{aligned}
{}^{C}\underline{\mathbf{u}} &= {}^{C}\mathbf{R}_{L}\,{}^{L}\underline{\mathbf{u}} && (202)\\
{}^{C}\mathbf{X}_{r} &= {}^{C}\mathbf{R}_{L}\,{}^{L}\mathbf{X}_{0} + {}^{C}\mathbf{T}_{L} && (203)
\end{aligned}
$$

where ${}^{L}\underline{\mathbf{u}} = (0, 0, 1)$. Note that ${}^{C}\mathbf{X}_{r}$ is a point belonging to the laser direction but it is not the reference point of the laser according to our definition. Remember that the reference point ${}^{C}\mathbf{X}_{0}$ must lie on the plane $Z_C = 0$. The equation of the line corresponding to the laser direction can be expressed in function of ${}^{C}\mathbf{X}_{0}$ as follows

$$
{}^{C}\mathbf{X}_{r} = \mu\,{}^{C}\underline{\mathbf{u}} + {}^{C}\,\mathbf{X}_{0} \tag{204}
$$

so that since we impose that ${}^{C}Z_{0} = 0$ then

$$
\mu = \frac{{}^{C}Z_{r}}{{}^{C}u_{z}} \tag{205}
$$

The calculation of the ${}^{C}\mathbf{X}_{0}$ is then straightforward

$$
{}^{C}\mathbf{X}_{0} = -\frac{{}^{C}Z_{r}}{{}^{C}u_{z}}\,{}^{C}\underline{\mathbf{u}} + {}^{C}\mathbf{X}_{r} \tag{206}
$$

The model parameters taking into account the whole model of misalignment become too complicated. Instead of this, we present the model parameters under individual types of misalignment, namely a simple displacement of the laser-cross with respect to the camera origin, and individual rotations of the laser-cross around the X, Y and Z axis, respectively.

Displacement In this case, the laser-cross frame has the same orientation that the camera frame (${}^{C}\mathbf{R}_{L} = \mathbf{I}_{3}$), but its origin has been displaced according to the vector

$$
{}^{C}\mathbf{T}_{L} = (t_x,\ t_y,\ t_z) \tag{207}
$$

The model parameters are then the ones shown in Table 3.

Rotation around the X_C axis The laser-cross is centered in the camera origin, but the laser-cross frame is rotated a certain angle ψ with respect to the X axis of the camera frame. The rotation is restricted to the interval $\psi \in (-\pi/2, \pi/2)$, otherwise the lasers projection is out of the camera field of view. The rotation matrix ${}^{C}R_{L}$ is then

$$
\mathrm{Rot}(X, \psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{c}\psi & -\mathrm{s}\psi \\ 0 & \mathrm{s}\psi & \mathrm{c}\psi \end{pmatrix} \tag{208}
$$

Table 3: Model parameters under a translational misalignment of the laser-cross.

Laser	X_0	Y_0	x	y	Z
1	t_x	$L+t_y$	t_x/Z_1	$(t_y+L)/Z_1$	$-\frac{At_x+BL+Bt_y+D}{C}$
2	$-L+t_x$	t_y	$(t_x-L)/Z_2$	t_y/Z_2	$-\frac{At_x-AL+Bt_y+D}{C}$
3	t_x	$-L+t_y$	t_x/Z_3	$(t_y-L)/Z_3$	$-\frac{At_x-BL+Bt_y+D}{C}$
4	$L+t_x$	t_y	$(t_x+L)/Z_4$	t_y/Z_4	$-\frac{At_x+AL+Bt_y+D}{C}$

where $c\psi = \cos\psi$ and $s\psi = \sin\psi$. The model parameters under this type of misalignment are shown in Table 4.

Table 4: Model parameters when the laser-cross is rotated around X_C.

Laser	X_0	Y_0	x	y	Z
1	0	$\frac{L}{c\psi}$	0	$-\frac{Ds\psi+LC}{BL+Dc\psi}$	$\frac{BL+Dc\psi}{Bs\psi-Cc\psi}$
2	$-L$	0	$-\frac{L(Bs\psi-Cc\psi)}{c\psi(D-AL)}$	$-\frac{s\psi}{c\psi}$	$\frac{c\psi(D-AL)}{Bs\psi-Cc\psi}$
3	0	$-\frac{L}{c\psi}$	0	$\frac{LC-Ds\psi}{Dc\psi-BL}$	$\frac{-BL+Dc\psi}{Bs\psi-Cc\psi}$
4	L	0	$\frac{L(Bs\psi-Cc\psi)}{c\psi(AL+D)}$	$-\frac{s\psi}{c\psi}$	$\frac{c\psi(AL+D)}{Bs\psi-Cc\psi}$

Rotation around the Y_C axis. Let us present now the case where the laser-cross is centered in the camera origin, but it is rotated an angle $\theta \in (-\pi/2, \pi/2)$ with respect to the Y axis of the camera frame. The rotation matrix ${}^C R_L$ has the following form

$$\mathrm{Rot}(Y,\theta) = \begin{pmatrix} \mathrm{c}\theta & 0 & \mathrm{s}\theta \\ 0 & 1 & 0 \\ -\mathrm{s}\theta & 0 & \mathrm{c}\theta \end{pmatrix} \tag{209}$$

where $c\theta = \cos\theta$ and $s\theta = \sin\theta$. The model parameters are then the ones shown in Table 5.

Table 5: Model parameters when the laser-cross is rotated around Y_C.

Laser	X_0	Y_0	x	y	Z
1	0	L	$\frac{s\theta}{c\theta}$	$-\frac{L\,(As\theta + Cc\theta)}{c\theta\,(BL + D)}$	$-\frac{c\theta\,(BL + D)}{As\theta + Cc\theta}$
2	$-\frac{L}{c\theta}$	0	$\frac{Ds\theta + LC}{-AL + Dc\theta}$	0	$-\frac{-AL + Dc\theta}{As\theta + Cc\theta}$
3	0	$-L$	$\frac{s\theta}{c\theta}$	$\frac{L\,(As\theta + Cc\theta)}{c\theta\,(-BL + D)}$	$-\frac{c\theta\,(-BL + D)}{As\theta + Cc\theta}$
4	$\frac{L}{c\theta}$	0	$\frac{Ds\theta - LC}{AL + Dc\theta}$	0	$-\frac{AL + Dc\theta}{As\theta + Cc\theta}$

Rotation around the Z_C axis In case that a rotation of ϕ occurs around the Z axis of the camera frame, the rotation matrix is

$$\mathrm{Rot}(Z,\phi) = \begin{pmatrix} \mathrm{c}\phi & -\mathrm{s}\phi & 0 \\ \mathrm{s}\phi & \mathrm{c}\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{210}$$

where $c\phi = \cos\phi$ and $s\phi = \sin\phi$. In this case, the model parameters are shown in Table 6.

C Kinematic screw frame transformation

The objective of this appendix is to define a frame transformation which allows the kinematic screw typically expressed in the camera frame to be expressed in a frame attached to the object. This can be done by using a transformation like

$$^c\mathbf{v} = {}^c\mathbf{T}_o\,{}^o\mathbf{v} \tag{211}$$

where $^c\mathbf{v}$ and $^o\mathbf{v}$ are the kinematic screw expressed in the camera and the object frames, respectively, and $^c\mathbf{T}_o$ is the 6×6 transformation changing the basis frame. This transformation is then useful to express the interaction matrix $\mathbf{L_x}$ in the object frame, which can be used to check which type of object motions can be detected in the camera image by using a certain set of visual features. The time variation of the visual features can be expressed in

Table 6: Model parameters when the laser-cross is rotated around Z_C.

Laser	X_0	Y_0	x	y	Z
1	$-s\phi L$	$c\phi L$	$-\frac{Ls\phi}{Z_1}$	$\frac{Lc\phi}{Z_1}$	$\frac{As\phi L - Bc\phi L - D}{C}$
2	$-c\phi L$	$-s\phi L$	$-\frac{Lc\phi}{Z_2}$	$-\frac{Ls\phi}{Z_2}$	$\frac{Ac\phi L + Bs\phi L - D}{C}$
3	$s\phi L$	$-c\phi L$	$\frac{Ls\phi}{Z_3}$	$-\frac{Lc\phi}{Z_3}$	$-\frac{As\phi L - Bc\phi L + D}{C}$
4	$c\phi L$	$s\phi L$	$\frac{Lc\phi}{Z_4}$	$\frac{Ls\phi}{Z_4}$	$-\frac{Ac\phi L + Bs\phi L + D}{C}$

two ways

$$\dot{\mathbf{s}} = {}^c\mathbf{L_s}\,{}^c\mathbf{v} \tag{212}$$

$$\dot{\mathbf{s}} = {}^o\mathbf{L_s}\,{}^o\mathbf{v} \tag{213}$$

where ${}^c\mathbf{L_s}$ and ${}^o\mathbf{L_s}$ are the interaction matrices expressed in the camera and object frames, respectively. Then, by plugging (211) into (212), we can write

$$\dot{\mathbf{s}} = {}^c\mathbf{L_s}\,{}^c\mathbf{T}_o\,{}^o\mathbf{v} \tag{214}$$

so that according to (213) the interaction matrix expressed in the object frame is

$${}^o\mathbf{L_s} = {}^c\mathbf{L_s}\,{}^c\mathbf{T}_o \tag{215}$$

The transformation matrix has the following form

$${}^c\mathbf{T}_o = \begin{pmatrix} {}^c\mathbf{R}_o & [{}^c\mathbf{P}_o]_\times\,{}^c\mathbf{R}_o \\ \mathbf{0}_3 & {}^c\mathbf{R}_o \end{pmatrix} \tag{216}$$

where ${}^c\mathbf{R}_o$ is the rotation matrix from the camera frame to the object frame, and ${}^c\mathbf{P}_o$ is the origin of the frame {O} expressed in the object frame {C}. $[{}^c\mathbf{P}_o]_\times$ is the antisymmetric matrix associated to the vector ${}^c\mathbf{P}_o$. In Fig. 41 the frame transformation is represented.

Let us now present how to obtain the analytic expression of the rotation matrix ${}^c\mathbf{R}_o$ and the origin of the object frame ${}^c\mathbf{P}_o$.

C.1 Rotation matrix

The constraint that is fixed to calculate the rotation matrix ${}^c\mathbf{R}_o$ is that $\mathbf{Z}_o$ must be equal to $\mathbf{Z}_c$ after applying the rotation (both $\mathbf{Z}_o$ and $\mathbf{Z}_c$ are expressed in the camera frame). Note

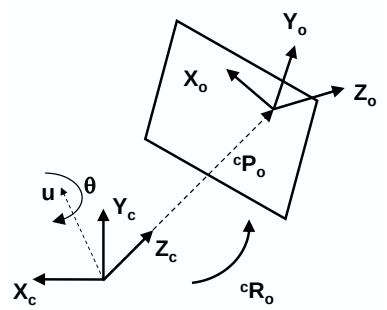

Figure 41: Frame transformation schema

that this constraint implies that a single rotation θ is made around an unitary axis $\underline{\mathbf{u}}$ which is orthogonal to $\mathbf{Z}_c$ and $\mathbf{Z}_o$. The rotation axis $\mathbf{u}$ can be calculated as follows

$$\mathbf{u} = \mathbf{Z}_c \times \mathbf{Z}_o = (0,0,1) \times (A,B,C) = (-B,A,0) \tag{217}$$

Therefore, the unitary vector $\underline{\mathbf{u}}$ is

$$\underline{\mathbf{u}} = \frac{1}{\sqrt{A^2+B^2}} \begin{pmatrix} -B \\ A \\ 0 \end{pmatrix} \tag{218}$$

The rotation θ can be calculated as follows

$$\begin{array}{rclcrcl} \sin(\theta) & = & \|\mathbf{Z}_c \times \mathbf{Z}_o\| & \Rightarrow & \sin(\theta) & = & \sqrt{A^2+B^2} \\ \cos(\theta) & = & \mathbf{Z}_c^\top \cdot \mathbf{Z}_o & \Rightarrow & \cos(\theta) & = & C \end{array} \tag{219}$$

According to the Rodrigues formula, a rotation matrix can be written as

$${}^c\mathbf{R}_o = \cos(\theta)\mathbf{I}_3 + \sin(\theta)[\underline{\mathbf{u}}]_\times + (1-\cos(\theta))\underline{\mathbf{u}} \cdot \underline{\mathbf{u}}^\top \tag{220}$$

with

$$[\underline{\mathbf{u}}]_\times = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix} \tag{221}$$

which leads to the rotation matrix

$${}^c\mathbf{R}_o = \begin{pmatrix} 1-\frac{A^2}{1+C} & -\frac{AB}{1+C} & A \\ -\frac{AB}{1+C} & 1-\frac{B^2}{1+C} & B \\ -A & -B & C \end{pmatrix} \tag{222}$$

C.2 Origin of the object frame

The origin of the object frame has been chosen to be equal to the projection of the focal point along the optical axis of the camera onto the planar object. Such a point expressed in the camera frame is

$$ {}^c\mathbf{P}_o = \begin{pmatrix} 0 \\ 0 \\ -D/C \end{pmatrix} \tag{223} $$

C.3 Frame transformation

Given the chosen ${}^c\mathbf{P}_o$ in (223) and the calculated rotation matrix ${}^c\mathbf{R}_o$ in (222), the frame transformation ${}^c\mathbf{T}_o$ has the following expression

$$ {}^c\mathbf{T}_o = \begin{pmatrix} 1-\frac{A^2}{1+C} & -\frac{AB}{1+C} & A & -\frac{ABD}{C(1+C)} & -\frac{D(B^2-1-C)}{C(1+C)} & \frac{DB}{C} \\ -\frac{AB}{1+C} & 1-\frac{B^2}{1+C} & B & \frac{D(A^2-1-C)}{C(1+C)} & \frac{ABD}{C(1+C)} & -\frac{AD}{C} \\ -A & -B & C & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\frac{A^2}{1+C} & -\frac{AB}{1+C} & A \\ 0 & 0 & 0 & -\frac{AB}{1+C} & 1-\frac{B^2}{1+C} & B \\ 0 & 0 & 0 & -A & -B & C \end{pmatrix} \tag{224} $$

Then, the interaction matrix $\mathbf{L_x}$ expressed in the object frame is

$$ {}^o\mathbf{L_x} = {}^c\mathbf{L_x} \cdot {}^c\mathbf{T}_o = \frac{1}{\Pi_0} \begin{pmatrix} 0 & 0 & -\frac{X_0}{Z} & X_0\eta & X_0\xi & 0 \\ 0 & 0 & -\frac{Y_0}{Z} & Y_0\eta & Y_0\xi & 0 \end{pmatrix} \tag{225} $$

where

$$ \begin{aligned} \eta &= \frac{1-A^2}{C}y + \frac{A(Bx+ACy)}{C(1+C)} \\ \xi &= \frac{1-B^2}{C}x + \frac{B(Ay+BCx)}{C(1+C)} \\ \Pi_0 &= A(X_0 - xZ) + B(Y_0 - yZ) - CZ \end{aligned} $$

Therefore, according to $\dot{\mathbf{x}} = {}^o\mathbf{L_x}{}^o\mathbf{v}$, if the object moved on ${}^oV_x$, ${}^oV_y$ or ${}^o\Omega_z$, the image coordinates of the projected points would not change.

D Stability analysis of the decoupled approach with a constant control law

This appendix presents different issues concerning the stability analysis of the set of visual features

$$\mathbf{s} \;=\; \left(\; y_1^{-1} - y_3^{-1}, \quad y_1^{-1} + y_3^{-1}, \quad x_2^{-1} + x_4^{-1} \; \right) \tag{226}$$

when using the constant control law based on the interaction matrix evaluated in the desired state $\mathbf{L}_{\mathbf{s}}^{*}$.

D.1 Sufficient conditions through the Lyapunov method

The global asymptotic stability analysis through the Lyapunov method is based on the product of matrices $\mathbf{M}$ in the closed-loop equation of the system $\mathbf{M} = \mathbf{L}_{\mathbf{s}}\widehat{\mathbf{L}_{\mathbf{s}}}^{+}$ which is the following 3×3 matrix

$$\mathbf{M} = \begin{pmatrix} 1 & \frac{BD}{LC^2} & \frac{AD}{LC^2} \\ 0 & \frac{B^2+C^2}{C^2} & \frac{AB}{C^2} \\ 0 & \frac{AB}{C^2} & \frac{A^2+C^2}{C^2} \end{pmatrix} \tag{227}$$

We remember that a necessary condition for the global asymptotic stability is that $\det(\mathbf{M}) \neq 0$. In this case, we have that

$$\det(\mathbf{M}) = \frac{1}{C^2} \tag{228}$$

so that it is never 0. Then, a sufficient condition for the global asymptotic stability is that the symmetric part $\mathbf{S}$ of $\mathbf{M}$ is positive definite.

Before calculating the eigenvalues of $\mathbf{S}$ let us reduce the number of parameters by using the following change of variables

$$a = \frac{A}{C}, \quad b = \frac{B}{C}, \quad d = \frac{D}{CL} \tag{229}$$

Then, $\mathbf{M}$ becomes

$$\mathbf{M} = \begin{pmatrix} 1 & bd & ad \\ 0 & b^2+1 & ab \\ 0 & ab & a^2+1 \end{pmatrix} \tag{230}$$

The symmetric part of $\mathbf{M}$ is

$$\mathbf{S} = \begin{pmatrix} 1 & \frac{bd}{2} & \frac{ad}{2} \\ \frac{bd}{2} & b^2+1 & ab \\ \frac{ad}{2} & ab & a^2+1 \end{pmatrix} \tag{231}$$

whose eigenvalues are

$$\begin{aligned}\sigma_1 &= 1\\ \sigma_2 &= \frac{2+a^2+b^2+\sqrt{(a^2+b^2)(a^2+b^2+d^2)}}{2}\\ \sigma_2 &= \frac{2+a^2+b^2-\sqrt{(a^2+b^2)(a^2+b^2+d^2)}}{2}\end{aligned}$$

we can see that the two first eigenvalues are always positive real numbers. The third one is also real, and imposing its positiveness, the following equation arises

$$4+(a^2+b^2)(4-d^2)>0 \tag{232}$$

by naming $X^2=(a^2+b^2)$ and $Y^2=d^2$ we can write

$$V=4+X^2(4-Y^2)>0 \tag{233}$$

Note that $X\in[-1,1]$ while the domain of Y is unrestricted. The plot of V is shown in Fig. 42. In order to see which part of the surface corresponds to positive values of V, the

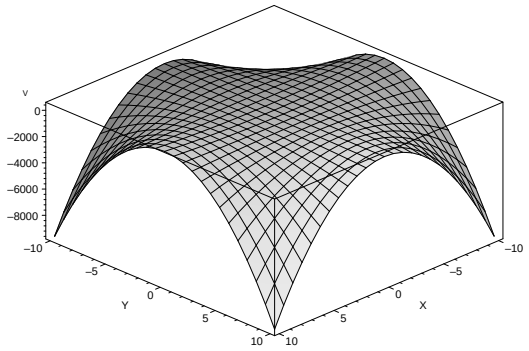

Figure 42: Plot of V in function of X and Y.

intersections of the surface with the plane $V=0$ are shown in Fig. 43. The positive part is the one included between the intersections and the X and Y axis. Then, the interpretation of the result is straightforward. The global asymptotic stability is ensured when $X\approx 0$ or $Y\approx 0$, which, expressed in terms of the planar object pose

$$\begin{aligned}X=0 \quad &\Leftrightarrow \quad a^2+b^2=0\Leftrightarrow\frac{A^2+B^2}{C^2}=0\Leftrightarrow\frac{1}{C^2}-1=0\Leftrightarrow|C|=1\\ Y=0 \quad &\Leftrightarrow \quad d^2=0\Leftrightarrow\frac{D^2}{C^2L^2}=0\end{aligned}$$

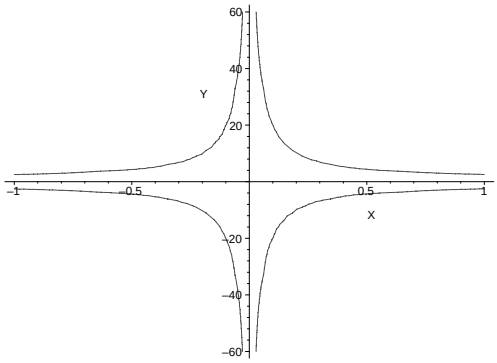

Figure 43: Intersection of the surface with the plane $V = 0$.

Then, the global asymptotic stability is ensured when the camera and the object are near parallel or when the distance between both is small. However, the relative orientation of the camera and the object has more weight than its distance, as can be seen in Fig. 43. Furthermore, increasing the parameter L the convergence domain also increases since Y becomes smaller.

As can be seen, these sufficient conditions are very restrictive even when no calibration errors have been taken into account. In the following subsection, we show that through the direct analysis of the solutions of the differential system the exact convergence domain is proved to be much larger.

D.2 Solution to the differential system

The closed-loop equation of the system corresponding to the set of visual features (115) when using the constant control law based on $\mathbf{L}_\mathbf{s}^*$ can be written as the following differential system

$$\dot{e}_1(t) = -\frac{\lambda}{4L}\left(e_1(t)\left(4L + e_2(t)^2 L + e_3(t)^2 L\right) + 2Z^*\left(e_2(t)^2 + e_3(t)^2\right)\right) \tag{234}$$

$$\dot{e}_2(t) = -\frac{\lambda}{4}(e_2(t)^3 + 4e_2(t) + e_2(t)e_3(t)^2) \tag{235}$$

$$\dot{e}_3(t) = -\frac{\lambda}{4}(e_3(t)^3 + 4e_3(t) + e_3(t)e_2(t)^2) \tag{236}$$

The solutions of $e_1(t)$, $e_2(t)$ and $e_3(t)$ can be found as follows. As expected, both $\dot{e}_2(t)$ (235) and $\dot{e}_3(t)$ (236) are not affected by $e_1(t)$ (234). Hence, we can start by searching the solutions of $e_2(t)$ and $e_3(t)$, which control the rotational dof of the system. First, we obtain the

expression of $\ddot{e}_2(t)$ by deriving (235)

$$\ddot{e}_2(t) = -\frac{\lambda}{4}\left(\dot{e}_2(t)\left(3e_2^2(t) + e_3^2(t) + 4\right) + 2e_2(t)e_3(t)\dot{e}_3(t)\right) \tag{237}$$

Afterwards, plugging $\dot{e}_3(t)$ from (236) into (237)

$$\begin{aligned}\ddot{e}_2(t) \;=\; & \frac{\lambda}{8}\left(-\dot{e}_2(t)\left(6e_2^2(t) + 2e_3^2(t) + 8\right) + \right. \\ & \left. +\lambda e_3^2(t)e_2(t)\left(e_2^2(t) + e_3^2(t) + 4\right)\right)\end{aligned} \tag{238}$$

From (235) the expression of $e_3^2(t)$ can be expressed in function of $\dot{e}_2(t)$ and $e_2(t)$

$$e_3^2(t) = \frac{-4\dot{e}_2(t) - \lambda e_2^3(t) - 4\lambda e_2(t)}{\lambda e_2(t)} \tag{239}$$

then, by plugging it into (238) and after some developments we obtain

$$\ddot{e}_2(t) - \dot{e}_2(t)\left(2\lambda + 3\frac{\dot{e}_2(t)}{e_2(t)}\right) = 0 \tag{240}$$

which is a second order Liouville differential equation with two symmetric solutions

$$e_2(t) = \pm\frac{\sqrt{\lambda}}{\sqrt{-(C_1\exp^{2\lambda t} + 2C_2\lambda)}} \tag{241}$$

where C_1 and C_2 are integration constants. By plugging (241) into (239) the solutions for $e_3(t)$ are directly obtained

$$e_3(t) = \mathrm{sgn}(e_2(t))\frac{\sqrt{\lambda(8C_2 - 1)}}{\sqrt{-(C_1\exp^{2\lambda t} + 2\lambda C_2}} \tag{242}$$

where $\mathrm{sgn}(x)$ returns the sign of the given value. Then, by evaluating any of the two pairs of solutions at time $t = 0$, C_1 and C_2 can be expressed in terms of the initial conditions $e_2(0)$ and $e_3(0)$ leading finally to

$$e_2(t) \;=\; \frac{2e_2(0)}{a(t)} \tag{243}$$

$$e_3(t) \;=\; \frac{2e_3(0)}{a(t)} \tag{244}$$

with

$$a(t) = \sqrt{\left(e_2^2(0) + e_3^2(0)\right)\left(\exp^{2\lambda t} - 1\right) + 4\exp^{2\lambda t}} \tag{245}$$

Finally, let us plug (243) in the definition of $\dot{e}_1(t)$ in (234) and solving a first order differential equation with non-constant coefficients the following solution arises

$$e_1(t) = \frac{2e_1(0)}{a(t)} - \frac{2bZ^* \arctan\left(\frac{b(a(t)-2)}{b^2+2a(t)}\right)}{a(t)L} \tag{246}$$

with $b = \sqrt{e_2^2(0) + e_3^2(0)}$.

We note that function $a(t)$ has the following properties

$$\begin{aligned} a(0) &= 2 \\ \lim_{t\to\infty} a(t) &= \infty \end{aligned}$$

Furthermore, by looking at its derivative

$$\dot{a}(t) = \frac{\lambda \exp(2\lambda t)\left(e_2(0)^2 + e_3(0)^2 + 4\right)}{a(t)} \tag{247}$$

it is always positive. Therefore $a(t)$ is monotonic in $t \in [0,\ \infty)$ and is bounded in the interval $[2,\ \infty)$.

D.3 Study of the behavior of the depth vs. time

We now study the behavior of the depth control which depends on $e_1(t)$. We are interested on identifying under which analytic conditions it becomes a monotonic function. In order to achieve it, we are going to identify the extrema of $e_1(t)$ by studying when the first derivative zeroes. Then, we will search for sufficient conditions which ensure that the derivative never zeroes so that $e_1(t)$ is monotonic.

We remember that the expression of $e_1(t)$ is given by

$$e_1(t) = \frac{2e_1(0)}{a(t)} - \frac{2bZ^* \arctan\left(\frac{b(a(t)-2)}{b^2+2a(t)}\right)}{a(t)L} \tag{248}$$

whose derivative

$$\dot{e}_1(t) = -\frac{\lambda}{4L}\left(e_1(t)\left(4L + e_2(t)^2 L + e_3(t)^2 L\right) + 2Z^*\left(e_2(t)^2 + e_3(t)^2\right)\right) \tag{249}$$

From this expression it is evident that when $b = 0$ ($e_2(0) = 0$ and $e_3(0) = 0$ so that the camera is already parallel to the object) a linear differential equation is obtained so that $e_1(t)$ is monotonic. Otherwise, it is necessary to study the derivative of $e_1(t)$ which can be rewritten as

$$\dot{e}_1(t) = -\frac{-2\dot{a}(t)\left(-bZ^* \arctan\left(\frac{b(a(t)-2)}{b^2+2a(t)}\right)\right)\left(b^2 + a(t)^2\right) + e_1(0)L\left(b^2 + a(t)^2\right) + b^2 Z^* a(t)\Big)}{L\left(b^2 + a(t)^2\right)a(t)^2} \tag{250}$$

As shown in Appendix D.2, $\dot{a}(t)$ is always positive and never zeroes, while the denominator of $\dot{\mathbf{e}}_1(t)$ is also positive. Therefore, $\dot{e}_1(t)$ only zeroes when

$$-bZ^* \arctan\left(\frac{b\,(a(t)-2)}{b^2+2a(t)}\right)\left(b^2+a(t)^2\right)+e_1(0)L\left(b^2+a(t)^2\right)+b^2Z^*a(t)=0 \tag{251}$$

By setting the following change of variable

$$u(t)=\frac{b\,(a(t)-2)}{b^2+2a(t)} \tag{252}$$

the expression in (251) can be rewritten as

$$\arctan\left(u(t)\right)=f(u(t)) \tag{253}$$

with

$$f(u(t))=\frac{u(t)^2\left(e_1(0)L\left(b^2+4\right)-2b^2Z^*\right)+u(t)bZ^*\left(b^2-4\right)e_1(0)L\left(b^2+4\right)+2b^2Z^*}{bZ^*\left((b^2+4)(\,u(t)^2+1\,)\right)} \tag{254}$$

Note that the derivative of $e_1(t)$ only zeroes if and only if $\arctan\left(u(t)\right)$ intersects with $f(u(t))$. Therefore, if we can find analytical conditions which avoid both functions to intersect, $e_1(t)$ will be monotonic under those conditions since its derivative never will zero.

We first study the behavior of $u(t)$. The following properties hold

$$\dot{u}(t) = \dot{a}(t)\frac{b\left(b^2+4\right)}{\left(b^2+2a(t)\right)^2} \tag{255}$$

$$u(0) = 0 \tag{256}$$

$$\lim_{t\to\infty} u(t) = b/2 \tag{257}$$

and since $a(t)$ is strictly monotonic increasing then $u(t)$ is also strictly monotonic increasing when $b\neq 0$ and it is bounded in the interval $[0,b/2]$ for $t\in[0,\infty)$.

The derivative of $f(u(t))$ is

$$\dot{f}(u(t))=-\dot{u}(t)\frac{u(t)^2\left(b^2-4\right)+u(t)8b+4-b^2}{\left(u(t)+1\right)\left(\left(b^2+4\right)\left(u(t)^2+1\right)\right)} \tag{258}$$

since $\dot{u}(t)$ is always positive and so is the denominator of $\dot{f}(u(t))$, the sign of $\dot{f}(u(t))$ depends on the following polynomial

$$p(u(t))=u(t)^2(b^2-4)+u(t)8b+4-b^2 \tag{259}$$

which can be written as

$$\left(b^2-4\right)\left(u-u_1\right)\left(u-u_2\right) \tag{260}$$

with u_1 and u_2 the roots of the polynomial

$$\begin{aligned} u_1 &= -\frac{b+2}{b-2} \\ u_2 &= \frac{b-2}{b+2} \end{aligned} \tag{261}$$

When $p(u(t)) < 0$ then $f(u(t))$ increases and inversely. Note that depending on if $b < 2$ or $b > 2$ the sign of $p(u(t))$ is affected. Let us study the sign depending on these conditions.

b<2: in this case $b^2 - 4$ is negative, and $u_1 > b/2$ and $u_2 < 0$ as shown hereafter

$$u_1 > \frac{1}{2} \Leftrightarrow \frac{1}{2} - u_1 < 0 \equiv \frac{1}{2} + \frac{b+2}{b-2} < 0 \equiv \frac{b^2+4}{2(b-2)} < 0 \tag{262}$$

$$u_2 < 0 \Leftrightarrow \frac{b-2}{b+2} < 0 \tag{263}$$

which means that when $b < 2$ there are no zero-crossings in the interval $u \in [0, b/2]$. Furthermore, the sign of the polynomial in this interval is always

$$\underbrace{(b^2-4)}_{<0}\underbrace{(u-u_1)}_{<0}\underbrace{(u-u_2)}_{>0} > 0 \tag{264}$$

b>2: in this case $b^2 - 4$ is positive, and $u_1 < 0$ and $u_2 \in [0, b/2]$ since

$$u_1 < 0 \Leftrightarrow -\frac{b+2}{b+2} < 0 \tag{265}$$

$$\begin{cases} u_2 > 0 \Leftrightarrow \dfrac{b-2}{b+2} > 0 \\ u_2 < \dfrac{b}{2} \Leftrightarrow \dfrac{b}{2} - u_2 > 0 \equiv \dfrac{b^2+4}{2(b+2)} > 0 \end{cases} \tag{266}$$

Therefore $p(u(t))$ has a unique zero-crossing in the interval $u \in [0, b/2]$ so that $p(u(t))$ is always increasing in such interval because

$$\begin{aligned} p(u=0) &= \underbrace{(b^2-4)}_{>0}\underbrace{(u_1)}_{<0}\underbrace{(u_2)}_{>0} < 0 \\ p(u=b/2) &= \underbrace{(b^2-4)}_{>0}\underbrace{(b/2-u_1)}_{>0}\underbrace{(b/2-u_2)}_{>0} > 0 \end{aligned} \tag{267}$$

Then, the behavior of $f(u(t))$ is determined by the following two cases (we remember that $b >= 0$ and when $b = 0$ $e_1(t)$ is always monotonic)

- $b \in (0, 2]$: $f(u(t))$ is monotonic decreasing when $u(t) \in [0, b/2]$.

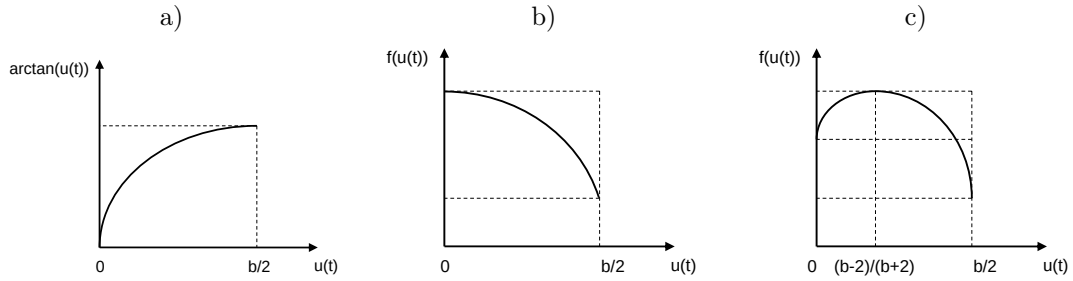

Figure 44: Schema of $\arctan(u(t))$ and $f(u(t))$ when $u(t) \in [0, b/2]$. a) $\arctan(u(t))$. b) $f(u(t))$ when $b \in (0, 2]$. c) $f(u(t))$ when $b > 2$.

- $b > 2$: $f(u(t))$ has a global maximum in the interval $[0, b/2]$ when $u(t) = (b-2)/(b+2)$.

In summary, the behavior of $\arctan(u(t))$ and $f(u(t))$ in the interval $[0, b/2]$ is represented in Fig. (44).

Note that two sufficient conditions can be defined in order to avoid the intersection of $\arctan(u(t))$ and $f(u(t))$:

- $\min(f(u(t))) > \max(\arctan(u(t)))$
- $\max(f(u(t))) < \min(\arctan(u(t)))$

According to the first condition we need the following expressions

$$\min(f(u(t))) = f(u = b/2) = \frac{e_1(0)L}{b \arctan(b/2)} \tag{268}$$

$$\max(\arctan(u(t))) = \arctan(b/2) \tag{269}$$

so that the condition is

$$Z^* < -\frac{e_1(0)L}{b \arctan(b/2)} \tag{270}$$

The second condition requires to distinguish between two cases

- $b \in (0, 2]$: $\max(f(u(t))) = f(u = 0) = \dfrac{e_1(0)L(b^2+4) + 2b^2 Z^*}{bZ^*(b^2+4)}$
- $b > 2$: $\max(f(u(t))) = f\left(u = \frac{b-2}{b+2}\right) = \dfrac{bZ^* + 2Le_1(0)}{2bZ^*}$

while

$$\min(\arctan(u(t))) = 0 \tag{271}$$

Therefore, the second sufficient condition is

$$\begin{aligned} b \in (0,2] \quad &\rightarrow \quad Z^* < -\frac{e_1(0)L(b^2+4)}{2b^2} \\ b > 2 \quad &\rightarrow \quad Z^* < -\frac{2Le_1(0)}{b} \end{aligned} \tag{272}$$

If these sufficient conditions are true then no intersection between $\arctan(u(t))$ and $f(u(t))$ will occur, and therefore, $e_1(t)$ will be monotonic.

The sufficient conditions in (270) and in (272) are expressed in terms of the initial state of the task function. They can be rewritten in terms of the initial object pose in the camera frame obtaining two different cases. The first case corresponds to the condition in (270) which is valid when the camera must go forward ($Z^* < -D(0)$). In such a case, the sufficient condition to ensure that $e_1(t)$ is monotonic is

$$Z^* < \frac{-2D(0)}{C(0)\left(2+\sqrt{1-C(0)^2}\arctan\left(\frac{\sqrt{1-C(0)^2}}{2}\right)\right)} \tag{273}$$

On the other hand, from (272) we obtain the sufficient condition valid when the camera must go backwards since $Z^* > -D(0)$

$$\begin{aligned} b \in (0,2] \quad &\rightarrow \quad Z^* > -\frac{D(0)}{C(0)^3} \\ b > 2 \quad &\rightarrow \quad Z^* > \frac{-2D(0)}{2C(0)-\sqrt{1-C(0)^2}} \end{aligned} \tag{274}$$

In summary, we have obtained sufficient conditions depending on the initial state (or the initial camera-object pose) which ensure that $e_1(t)$ will be monotonic. Even if these sufficient conditions are not ensured, we can al least ensure that $e_1(t)$ will present a unique peak since a unique intersection of $\arctan(u(t))$ and $f(u(t))$ occurs.

References

[1] G.J. Agin. Calibration and use of a light stripe range sensor mounted on the hand of a robot. In *IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 680–685, St. Louis, Missouri, USA, Mars 1985.

[2] J. Albus, E. Kent, M. Nashman, P. Manabach, and L. Palombo. Six-dimensional vision system. In *Proc. SPIE*, volume 336 of *Robot Vision*, pages 142–153, 1982.

[3] A. Alhaj, C. Collewet, and F. Chaumette. Visual servoing based on dynamic vision. In *IEEE Int. Conf. on Robotics and Automation*, pages 3055–3060, Taipei, Taiwan, 2003.

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

[5] R. Basri, E. Rivlin, and I. Shimshoni. Visual homing: surfing on the epipoles. In *IEEE Int. Conf. on Computer Vision*, pages 863–869, Bombay, India, January 1998.

[6] Paul J. Besl. *Advances in Machine Vision: Architectures and Applications*, chapter Active Optical Range Imaging Sensors, pages 1–53. Springer-Verlag, 1988.

[7] Z. Bien, W. Jang, and J. Park. *Visual servoing*, chapter Characterization and use of feature-jacobian matrix for visual servoing, pages 317–363. World Scientific, 1993.

[8] E. Cervera, F. Berry, and P. Martinet. Is 3d useful in stereo visual control? In *IEEE Int. Conf. on Robotics and Automation*, pages 1630–1635, Washington, D.C., USA, 2002.

[9] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In *The Conference of Vision and Control*, volume 237, pages 66–78. LNCIS Series, 1998.

[10] F. Chaumette. Image moments: a general and useful set of features for visual servoing. *IEEE Trans. on Robotics*, 20(4):713–723, August 2004.

[11] F. Chaumette, S. Boukir, P. Bouthemy, and D. Juvin. Structure from controlled motion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(5):492–504, May 1996.

[12] C. Collewet, A. Alhaj, and F. Chaumette. Model-free visual servoing on complex images based on 3d reconstruction. In *IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 751–756, April 26 - May 1 2004.

[13] C. Collewet and F. Chaumette. A contour approach for image-based control of objects with complex shape. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 1, pages 751–756, Takamatsu, Japan, October-November 2000.

[14] C. Collewet and F. Chaumette. Positioning a camera with respect to planar objects of unknown shape by coupling 2-d visual servoing and 3-d estimations. *IEEE Trans. on Robotics and Automation*, 3(18):322–333, June 2002.

[15] C. Colombo, B. Allota, and P. Dario. Affine visual servoing: a framework for relative positionning with a robot. In *IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 21–27, Nagoya, Japan, May 1995.

[16] C. Colombo and B. Allotta. Image-based robot task planning and control using a compact visual representaion. *IEEE Trans. on Systems, Man, and Cybernetics*, 29(1):92–100, January 1999.

[17] 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.

[18] A. Crétual and F. Chaumette. Visual servoing based on image motion. *Int. Journal of Robotics Research*, 20(11):857–877, 2001.

[19] N. Daucher, M. Dhome, J. T. Lapresté, and G. Rives. Speed command of a robotic system by monocular pose estimate. In *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pages 55–62, Grenoble, France, September 1997.

[20] S. De Ma. Conics-based stereo, motion estimation and pose determination. *Int. Journal of Computer Vision*, 10(1):7–25, 1993.

[21] K. Deguchi. Direct interpretation of dynamic images and camera motion for visual servoing without image feature correspondence. *Journal of Robotics and Mechatronics*, 9(2):104–110, 1997.

[22] D. Dementhon and L. S. Davis. Model-based object pose in 25 lines of code. *Int. Journal of Computer Vision*, 15(1/2):123–141, 1995.

[23] M. Dhome, J. T. Lapresté, G. Rives, and M. Richetin. Spatial localization of modelled objects of revolution in monocular perspective vision. In *European Conf. on Computer Vision*, volume 427, pages 475–485, Antibes, 1990.

[24] M. Dhome, M. Richetin, J.T. LaPresté, and G. Rives. Determination of the attitude of 3-d objects from a single perspective view. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(12):1265–1278, 1989.

[25] T. Drummond and R. Cipolla. Real-time tracking of complex structures with on-line camera calibration. *Image and Vision Computing*, 20(5-6):427–433, 2002.

[26] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(6):313–326, June 1992.

[27] 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.

[28] G. D. Hager. A modular system for robust positioning using feedback from stereo vision. *IEEE Trans. on Robotics and Automation*, 13(4):582–595, 1997.

[29] R.M. Haralick, H. Joo, C.N. Lee, X. Zhuang, V.G. Vaidya, and M.B. Kim. Pose estimation from corresponding point data. *IEEE Trans. on Systems, Man and Cybernetics*, 19(6):1426–1446, 1989.

[30] R. Horaud, B. Conio, O. Leboulleux, and B. Lacolle. An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics, and Image processing*, 47(1):33–44, 1989.

[31] T. S. Huang and A. N. Netravali. Motion and structure from feature correspondences: a review. *Proceedings of the IEEE*, 82(2):252–268, 1994.

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

[33] R. Jarvis. Range sensing for computer vision. *Advances in Image Communications. Elsevier Science Publishers, Amsterdam*, pages 17–56, 1993.

[34] E. Kent, T. Wheatley, and M. Nashman. Real-time cooperative interaction between structured-light and reflectance ranging for robot guidance. *Robotica*, 3:7–11, 1985.

[35] 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.

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

[37] D. G. Lowe. Fitting parameterized three-dimensional models to images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(5):441–450, 1991.

[38] 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.

[39] E. Malis, J.J. Borrely, and P. Rives. Intrinsics-free visual servoing with respect to straight lines. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 1, pages 384–389, Laussane, Switzerland, September 2002.

[40] E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *IEEE Trans. on Robotics and Automation*, 18(2):176–186, April 2002.

[41] 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.

[42] E. Marchand and F. Chaumette. Active vision for complete scene reconstruction and exploration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(1):65–72, 1999.

[43] Y. Mezouar and F. Chaumette. Path planning for robust image-based control. *IEEE Trans. on Robotics and Automation*, 18(4):534–549, August 2002.

[44] G. Morel, T. Leibezeit, J. Szewczyk, S. Boudet, and J. Pot. Explicit incorporation of 2d constraints in vision-based control of robot manipulators. In *Int. Symp. on Experimental Robotics*, volume 250 of *LNCIS*, pages 99–108, Sidney, Australia, 2000. Springer Verlag.

[45] J. Pages, C. Collewet, F. Chaumette, and J. Salvi. Plane-to-plane positioning from image-based visual servoing and structured light. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 1, pages 1004–1009, Sendai, Japan, September 2004.

[46] P. Questa, E. Grossmann, and G. Sandini. Camera self orientation and docking maneuver using normal flow. In *SPIE AeroSense*, Orlando, Florida, 1995.

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

[48] C. Samson, M. Le Borgne, and B. Espiau. *Robot control: the task function approach*. Clarendon Press, Oxford, 1991.

[49] J. Santos-Victor and G. Sandini. Visual behaviors for docking. *Computer Vision and Image Understanding*, 67(3):223–238, 1997.

[50] F. Schramm, G. Morel, A. Micaelli, and A. Lottin. Extended-2d visual servoing. In *IEEE Int. Conf. on Robotics and Automation*, pages 267–273, New Orleans, USA, April 26-May 1 2004.

[51] V. Sundareswaran, P. Bouthemy, and F. Chaumette. Exploiting image motion for active vision in a visual servoing framework. *Int. Journal of Robotics Research*, 15(6):629–645, 1996.

[52] O. Tahri and F. Chaumette. Application of moment invariants to visual servoing. In *IEEE Int. Conf. on Robotics and Automation*, Taipei, Taiwan, September 2003.

[53] 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.

[54] J. P. Urban, G. Motyl, and J. Gallice. Realtime visual servoing using controlled illumination. *Int. Journal of Robotics Research*, 13(1):93–10, 1994.

[55] W. J. Wilson, C. C. W. Hulls, and G. S. Bell. Relative end-effector control using cartesian position based visual servoing. *IEEE Trans. on Robotics and Automation*, 12:684–696, October 1996.

[56] J. S. C. Yuan. A general photogrammetric method for determining object position and orientation. *IEEE Trans. on Robotics and Automation*, 5(2):129–142, 1989.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399