Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods

E. Malis, François Chaumette

▶ To cite this version:

Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods

Ezio Malis and François Chaumette

Abstract— **This paper concerns the stability analysis of a new class of model-free visual servoing methods. These methods are "model-free" since they are based on the estimation of the relative camera orientation between two views of an object without knowing its 3D model. The visual servoing is decoupled by controlling the rotation of the camera separately from the rest of the system. The way the remaining degrees of freedom are controlled differentiates the methods within the class. For all the methods of the class, the robustness with respect to both camera and hand-eye calibration errors can be analytically studied. In some cases, necessary and sufficient conditions can be found not only for the local asymptotic stability but also for the global asymptotic stability. In the other cases, simple conditions on the calibration errors are sufficient to ensure the global asymptotic stability of the control law. In addition to the theoretical proof of the stability, the experimental results prove the validity of the control strategy proposed in the paper.**

Index Terms— **Visual servoing, stability analysis, robustness, model-free reconstruction.**

I. Introduction

Visual servoing is a promising method to control dynamic systems using the information provided by visual sensors. In this paper, we focus on the control of a single camera mounted on the end-effector of a robot manipulator. The computer vision system controls the robot's end-effector in order to position it with respect to an object. Similarly to most of the vision-based control approaches, we consider only robot kinematics. Indeed, most manufacturers only provide robots with access to kinematics controllers. The robot dynamics is generally taken into account in a low-level controller which cannot be modified by users. Even if robots dynamics is important when tracking object at high speed, kinematics is fundamental to robot control and cannot be ignored. Since the design of the vision system presented in the paper is general, the robot dynamics can be easily taken into account afterwards, as it was done in [11] and [4] for a standard image-based control.

Many different visual servoing approaches have been proposed [13] [14] and the corresponding control laws are generally stable in absence of calibration errors on the parameters of the system. In general, it can be observed experimentally that the control laws are robust in presence of calibration errors especially when the initial camera displacement is not too large. However, the theoretical analysis of the stability and robustness of a control law is generally impossible. For example, if we consider the standard image-based approach proposed in [8], it is extremely difficult to find analytical stability conditions, which can be exploited in practice, since the system is highly coupled [9]. Even if the behavior of the image features is generally good, it is possible to reach local minima (which means that the final robot position does not correspond to the desired one) and/or the image Jacobian can become singular during the servoing (which of course leads to an unstable behavior) [3]. Position-based approaches are naturally decoupled since they are based on the reconstruction of the camera position with respect to the object. In this case the control of the rotation is decoupled from the control of the translation and the stability analysis is simplified. However, when the position of the camera is computed using a CAD model of the object [23], it is extremely difficult to have an analytical expression of the estimated information as a function of the calibration errors. Even if it is possible to provide a stable control law in the presence of small calibration errors [24], it seems hard to know how small they should be in order to ensure the convergence of the control law. Furthermore, these approaches may be completely unstable with respect to small image measurement errors [3]. Decoupled hybrid approaches [16] [6] [20] [18] are a possible solution to some drawbacks of image-based and position-based methods.

Thanks to recent results in computer vision, it is possible to estimate the camera displacement between two views without knowing any CAD model of the considered object [17]. Visual servoing approaches exploiting these results will be called model-free in order to distinguish them from previous ones which are model-dependent. With model-free approaches it is sufficient to know that several features correspond in the two images to recover the rotation of the camera and the translation up to a scale factor. As a consequence, the control of the rotation of the camera can be decoupled from the control of the translation. This important fact characterizes the new class of model-free visual servoing methods analyzed in this paper. The translation of the camera can be controlled directly, as it was proposed in [1] (using thus a model-free position-based approach), or using an hybrid approach as we proposed in [16] (it consists in combining visual features obtained directly from the image with features expressed in the Euclidean space). Another interesting property of the model-free approaches is that it is possible to prove the stability and the robustness of the control law and it is possible to have the analytical expression of the robustness domain (i.e. the exact knowledge of the amount of calibration errors that can be tolerated by

E. Malis was with the University of Cambridge, Cambridge, U.K. He is now with INRIA Sophia Antipolis, 06902 Sophia Antipolis Cedex, France (e-mail: Ezio.Malis@sophia.inria.fr).

F. Chaumette is with the IRISA/INRIA Rennes, 35042 Rennes, France.

the system). More particularly, in [16] we proved the necessary and sufficient conditions for local asymptotic stability in presence of camera calibration errors. Local asymptotic stability is very important since it guarantees that the system will converge by starting with a small initial error. If the initial error is large, it can be sampled and a reference trajectory can be built from a sequence of images [19]. Therefore, at each iteration it is possible to work with small errors and, if the error remains small, the servoing will always converge. Furthermore, in [15] sufficient conditions for global asymptotic stability were proposed not only in the presence of camera calibration errors but also in the presence of calibration errors in the hand/eye transformation. Global asymptotic stability is important since it guarantees that the system will converge even if the initial error is very large. Recently, Taylor et al. [22] proposed a position-based approach, very similar to the one proposed by Basri [1], finding local stability conditions resembling to those proposed in [15]. As it will be shown in this paper, this is not surprising, since model-free position-based approaches are particular cases of the class we analyze in this paper. We will show that the calibration error that can be tolerated by the schemes of the class can be considerably larger than the one proposed in [16]. Indeed, in this paper we considerably improve the stability analysis of the model-free visual servoing systems still taking into account both the calibration errors in the camera parameters and in the camera/hand-eye transformation. In particular, we prove simple necessary and sufficient conditions for the global asymptotic convergence of a new hybrid approach and of a position-based approach. It must be also noticed that it is not sufficient to prove the convergence of the control-loop to ensure that the visual servoing will not fail. Indeed, in position-based approaches, the lack of control of the behavior of the image features implies that the target may get out of the camera field of view during the servoing (leading of course to its failure). Even if a complex strategy could be used to incorporate image constraints, we think that hybrid approaches, combining a good behavior of the features in the image with a good behavior of the camera in the Cartesian space, are preferable with respect to position-based approaches. In this paper, we propose a simple control strategy in order to keep at least the centroid of the observed object in the image. In such case, it is sure that some parts of the object will be always visible by the camera and it will be more likely that there will always be a sufficient number of visual features to perform the visual servoing.

The paper is organized as follows. In Section II we review the theoretical background of the model-free visual servoing using an eye-in-hand camera. Section III describes in details the class of visual servoing system we want to analyze. In Section IV we provide the stability analysis of the hybrid approaches (a new method is proposed) and we prove that the similar conditions hold for the stability of a new position-based approach. Finally, in Section V we give some experimental results using a simple control strategy.

II. Theoretical background

In this section we first review some properties of the projective geometry that are used to recover the camera displacement from two views (a detailed description can be found in [17]). Then, we give a general description of the visual servoing within the task function approach [21].
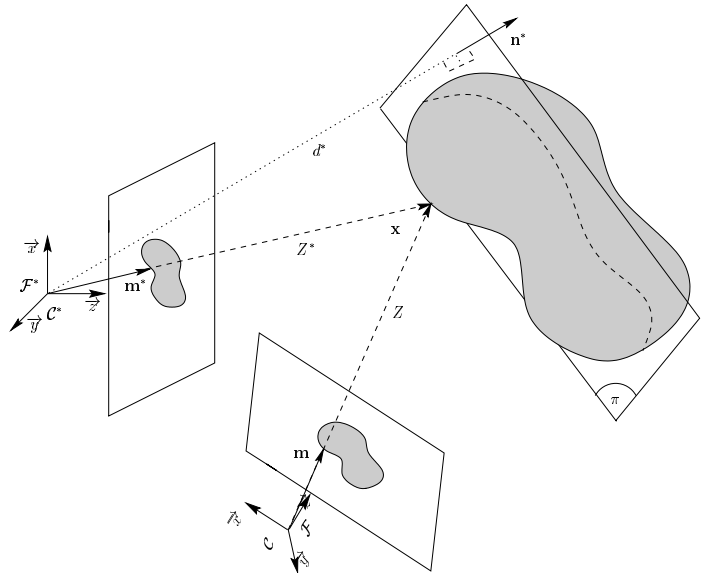

Fig. 1.   Geometry of two views of an unknown object.

A. Recovering the camera displacement from two views

Consider two views of an object with a generic shape (see Figure 1). The reference view, corresponding to frame $\mathcal{F}^*$, is learn during an off-line step. The current view acquired at each iteration of the control scheme corresponds to frame $\mathcal{F}$. A 3D point $\mathbf{x} = (X, Y, Z, 1)$ projects to the point $\mathbf{m}^* = (x^*, y^*, 1)$ in the reference frame $\mathcal{F}^*$ and to the point $\mathbf{m} = (x, y, 1)$ in the current frame $\mathcal{F}$. From the knowledge of several matched points, it is possible to recover the camera displacement between the two views without knowing the 3D model of the object. Indeed, the relative displacement of the camera can be extracted from a homography relative to a virtual plane attached to the object [10] [17]:

$$\mathbf{H} = \mathbf{R} + \frac{\mathbf{t}}{d^*}\mathbf{n}^*$$

where $\mathbf{R}$ and $\mathbf{t}$ are respectively the rotation matrix and the translation vector between frames $\mathcal{F}$ and $\mathcal{F}^*$, $\mathbf{n}^*$ is the unit vector normal to π expressed in $\mathcal{F}^*$ and d^* is the distance of π to $\mathcal{F}^*$. In general, there are two possible solutions to decompose the homography [10], but the right one can be found considering another reference plane if the object is not planar or considering more images if the object is planar. Since the measured points in the image are in pixels, the homography can be computed knowing the matrix $\mathbf{A}$ which contains the camera intrinsic parameters:

$$\mathbf{A} = \begin{bmatrix} fk_u & -fk_u\cot(\varphi) & u_0 \\ 0 & fk_v/\sin(\varphi) & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where u_0 and v_0 are the coordinates of the principal point (in pixels), f is the focal length (in meters), k_u and k_v are the magnifications respectively in the $\overrightarrow{u}$ and $\overrightarrow{v}$ direction (in pixels/meters), and φ is the angle between these axes. In general, the camera internal parameters are not perfectly known. An approximated matrix $\widehat{\mathbf{A}}$ should be used. Thus, the estimated homography matrix $\widehat{\mathbf{H}}$ is:

$$\widehat{\mathbf{H}} = \tilde{\mathbf{A}}\mathbf{H}\tilde{\mathbf{A}}^{-1}$$

where $\tilde{\mathbf{A}} = \widehat{\mathbf{A}}^{-1}\mathbf{A}$ is a triangular matrix which contains the calibration errors (if $\widehat{\mathbf{A}} = \mathbf{A}$ then $\tilde{\mathbf{A}} = \mathbf{I}$ and $\widehat{\mathbf{H}} = \mathbf{H}$). A particular homography is the homography of the plane at infinity: $\mathbf{H}_\infty = \mathbf{R}$. This matrix can be recovered from several views or even from two views if we add some additional information (i.e. we need at least four vanishing points). Then, the estimated translation and rotation are respectively: $\widehat{\mathbf{t}} \propto \tilde{\mathbf{A}}\mathbf{t}$ and $\widehat{\mathbf{R}} \propto \tilde{\mathbf{A}}\mathbf{R}\tilde{\mathbf{A}}^{-1}$. Since $\widehat{\mathbf{R}}$ is similar to $\mathbf{R}$, it is possible to estimate the angle of rotation $\widehat{\theta} = \theta$ and the axis of rotation $\widehat{\mathbf{u}} = \tilde{\mathbf{A}}\mathbf{u}/\|\tilde{\mathbf{A}}\mathbf{u}\|$.

B. The task function approach for visual servoing

The aim of the visual servoing method we are going to describe is to position a camera, mounted on the end-effector of a robot, with respect to an object. Visual servoing can be achieved by regulating to zero a task function built from the information issued by the camera. The time derivative of the task function is:

$$\dot{\mathbf{e}} = \mathbf{L}\,\mathbf{v}_c = \mathbf{L}\,\mathbf{W}\,\mathbf{v}_r \qquad (2)$$

where $\mathbf{L}$ is called the image Jacobian matrix and links the variation of the task function with respect to the camera velocity $\mathbf{v}_c$. It depends on the camera parameters and some geometric parameters of the observed scene (like for example the distance between the camera and the object). $\mathbf{W}$ is the matrix transforming the end-effector velocity $\mathbf{v}_r$ in the camera velocity: $\mathbf{v}_c = \mathbf{W}\,\mathbf{v}_r$. This matrix is upper block-triangular and contains the rotation ${}^c\mathbf{R}_r$ and the translation ${}^c\mathbf{t}_r$ between camera and end-effector frames :

$$\mathbf{W} = \begin{bmatrix} {}^c\mathbf{R}_r & [{}^c\mathbf{t}_r]_\times\,{}^c\mathbf{R}_r \\ 0 & {}^c\mathbf{R}_r \end{bmatrix} \qquad (3)$$

where $[{}^c\mathbf{t}_r]_\times$ is the skew symmetric matrix associated to vector ${}^c\mathbf{t}_r$ (such that $[\mathbf{a}]_\times \mathbf{b}$ is the cross product $\mathbf{a} \times \mathbf{b}$). In our previous work [16], ${}^c\mathbf{R}_r$ and ${}^c\mathbf{t}_r$ were supposed to be exactly known. In this paper, we will only suppose that the Jacobian of the robot is perfectly known and full rank. However, even in the presence of small errors in the model of the robot, the camera will be correctly repositioned if the error in the image decreases to zero. As already mentioned in the introduction, we suppose that the low-level controller of the robot (which generally takes into account the robot dynamics) is able to provide the velocity screw we use as control input. If this is not the case, one must design a control law which considers the robot dynamics.

C. Control law

A simple control law can be obtained imposing the exponential decrease of the task function $\dot{\mathbf{e}} = -\mathbf{D}\,\mathbf{e}$, where $\mathbf{D}$

is a positive diagonal matrix whose elements tune the exponential decrease. In the absence of calibration errors, a decoupling and linearizing control law would be:

$$\mathbf{v}_r = -\mathbf{W}^{-1}\,\mathbf{L}^{-1}\,\mathbf{D}\,\mathbf{e}$$

where $\mathbf{v}_r$ is the end-effector velocity sent to the robot controller. Due to the presence of calibration errors only approximations or estimates of the parameters of the system are available. Thus, the end-effector velocity $\mathbf{v}_r$ sent to the robot controller is:

$$\mathbf{v}_r = -\widehat{\mathbf{W}}^{-1}\,\widehat{\mathbf{L}}^{-1}\,\mathbf{D}\,\widehat{\mathbf{e}} \qquad (4)$$

where $\widehat{\mathbf{L}}^{-1}$, $\widehat{\mathbf{W}}^{-1}$ and $\widehat{\mathbf{e}}$ are respectively approximations of $\mathbf{L}^{-1}$, $\mathbf{W}^{-1}$ and $\mathbf{e}$. In general the matrix $\mathbf{L}$ and vector $\mathbf{e}$ depends on the matrix $\mathbf{A}$ containing the camera internal parameters. These parameters are unknown and only an approximated matrix $\widehat{\mathbf{A}}$ can be used in the control law.

D. The stability of the closed-loop system

Assuming that the intrinsic camera parameters do not change after the off-line acquisition of the desired image, we will have $\mathbf{e} = 0$ if and only if the target points have reached their desired position in the image. The estimated task function $\widehat{\mathbf{e}}$ is measured from the current and desired images and can be expressed as a function of the real $\mathbf{e}$: $\widehat{\mathbf{e}} = \mathbf{f}(\mathbf{e})$. It is important that the function $\mathbf{f}()$ computing $\widehat{\mathbf{e}}$ from the image points is bijective in order to have $\widehat{\mathbf{e}} = 0$ if and only if $\mathbf{e} = \mathbf{0}$. Thus, the estimated task function can always be written as follow:

$$\widehat{\mathbf{e}} = \mathbf{E}(\mathbf{e})\mathbf{e} \qquad (5)$$

where $\mathbf{E}(\mathbf{e})$ is a full rank matrix which form depends on the chosen task function. The closed-loop system taking into account all the calibration errors can thus be written:

$$\dot{\mathbf{e}} = -\mathbf{Q}(\mathbf{e})\mathbf{e} \qquad (6)$$

where:

$$\mathbf{Q}(\mathbf{e}) = \mathbf{L}\,\tilde{\mathbf{W}}\,\widehat{\mathbf{L}}^{-1}\,\mathbf{D}\,\mathbf{E} \qquad (7)$$

and $\tilde{\mathbf{W}}$ is an upper block triangular square matrix :

$$\tilde{\mathbf{W}} = \mathbf{W}\,\widehat{\mathbf{W}}^{-1} = \begin{bmatrix} \tilde{\mathbf{R}} & \tilde{\mathbf{R}}\,[\tilde{\mathbf{t}}]_\times \\ 0 & \tilde{\mathbf{R}} \end{bmatrix} \qquad (8)$$

with $\tilde{\mathbf{R}} = {}^c\mathbf{R}_e\,{}^c\widehat{\mathbf{R}}_e^T$ and $\tilde{\mathbf{t}} = (\tilde{\mathbf{R}}^{T\,c}\mathbf{t}_e - {}^c\widehat{\mathbf{t}}_e)$. In the next section, it will be shown that there is no singularity in the workspace (i.e. $\det(\mathbf{Q}(\mathbf{e})) \neq 0$). Thus, if the task function decreases, it decreases towards the equilibrium point $\mathbf{e} = 0$.

This paper is concerned with the stability of the equilibrium point. An equilibrium point is asymptotically stable if all solutions starting at nearby points tend to the equilibrium point as time approaches infinity. If the starting point is very close to the equilibrium point we can study the local stability. An equilibrium point is locally stable if all solutions starting in a small neighborhood tend to the equilibrium point as time approaches infinity [12]. By definition,

local stability only provides information about the behavior of the system near the convergence. From a practical point of view, it is more interesting to study global stability. An equilibrium point is globally stable when the equilibrium is reached from any starting point [12].

Stability of equilibrium points is usually characterized in the sense of Lyapunov. In the case of linear systems (i.e. when the matrix $\mathbf{Q}$ does not depend on the state $\mathbf{e}$), local and global stability coincide. The necessary and sufficient condition for the global (and thus local) asymptotic stability of a linear system is that the matrix $\mathbf{Q}$ has all eigenvalues with positive real part. In the case of non-linear systems, the matrix $\mathbf{Q}$ is a function of the state $\mathbf{e}$. The system can be linearized around the equilibrium point $\mathbf{e} = 0$ and the local stability can be studied as for linear system. For non-linear systems, the study of the global stability is complex and, in general, only sufficient conditions can be given. For example, a sufficient condition to ensure the global asymptotic stability is $\mathbf{Q}(\mathbf{e}) > 0$. In this case, it can be also shown that $\|\mathbf{e}\|$ always decreases, which is very interesting in practice. Even if, under particular conditions, the error decreases at each iteration, it is not always possible to ensure that the observed features always remain in the field of view of the camera. Thus, one constraint must be added to the system which can be called the "visibility constraint": a minimal number of visual features *must* remain in the camera field of view in order to avoid the failure of the servoing. To deal with global stability we should not only ensure that the system is stable but satisfies also the visibility constraint.

III. Characterizing the new class

The new class of visual servoing methods is mainly characterized by the fact that the chosen task function is decomposed into two parts $\mathbf{e} = [\ \mathbf{e}_\nu \quad \mathbf{e}_\omega\ ]$ where $\mathbf{e}_\omega$ is used to control the 3 rotational camera d.o.f. (and depends only on it) and $\mathbf{e}_\nu$ is used to control the 3 remaining degrees of freedom. This decoupling is possible since it is possible to estimate the relative orientation $\mathbf{R}$ between the current and reference camera frames [17]. As a consequence, the matrix $\mathbf{L}$ is an upper block triangular square matrix:

$$\mathbf{L} = \left[\begin{array}{cc} \mathbf{L}_\nu & \mathbf{L}_{\nu\omega} \\ \mathbf{0}_3 & \mathbf{L}_\omega \end{array} \right] \tag{9}$$

The structure of the matrix simplifies the analysis of the stability of the control law. In order to keep the control law decoupled, we will set two independent positive gains:

$$\mathbf{D} = \left[\begin{array}{cc} \lambda_\nu \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \lambda_\omega \mathbf{I}_3 \end{array} \right] \tag{10}$$

A. Control of the rotation

In order to control the orientation of the camera, it is natural to use the 3D estimated rotation $\mathbf{R}$ between $\mathcal{F}$ and $\mathcal{F}^*$ (which has to reach the identity matrix). Let $\mathbf{u}$ and θ be respectively the axis and the angle of rotation obtained from $\mathbf{R}$. Many task functions have been proposed to be used in visual servoing: $\mathbf{e}_\omega = \mathbf{u}\sin(\theta)$ [22], $\mathbf{e}_\omega = \mathbf{u}\sin(\frac{\theta}{2})$ [24] or $\mathbf{e}_\omega = \mathbf{u}\theta$ [16]. In general, the task function can be written

as $\mathbf{e}_\omega = \mathbf{u}f(\theta)$ where $f(\theta)$ is a strictly increasing function in the domain $0 \leq \theta < \theta_{max}$ [2]. The derivative of $\mathbf{e}_\omega$ depends only on the rotational velocity of the camera:

$$\dot{\mathbf{e}}_\omega = \mathbf{L}_\omega \boldsymbol{\omega} \tag{11}$$

where:

$$\mathbf{L}_\omega = \frac{\partial f(\theta)}{\partial \theta}\mathbf{I}_3 - \frac{f(\theta)}{2}[\mathbf{u}]_\times + \left(\frac{\partial f(\theta)}{\partial \theta} - \frac{f(\theta)}{2\tan(\frac{\theta}{2})} \right)[\mathbf{u}]_\times^2$$

$[\mathbf{u}]_\times$ being the (3×3) skew matrix associated to the (3×1) vector $\mathbf{u}$. The determinant of $\mathbf{L}_\omega$ is:

$$\Delta = \det(\mathbf{L}_\omega) = \left(\frac{f(\theta)/2}{\frac{\partial f(\theta)}{\partial \theta}\sin(\frac{\theta}{2})} \right)^2 \tag{12}$$

Table I shows three different common choices for $f(\theta)$ and gives the singularity of the determinant Δ which corresponds to θ_{max}.

$f(\theta)$	Δ	singularity
θ	$(\mathrm{sinc}(\frac{\theta}{2}))^{-2}$	2π
$\sin(\frac{\theta}{2})$	$(\cos(\frac{\theta}{2}))^{-2}$	π
$\sin\theta$	$\left(\dfrac{\cos(\theta)}{\cos(\frac{\theta}{2})} \right)^2$	$\dfrac{\pi}{2}$

TABLE I
Common choices for $f(\theta)$ and its corresponding singularity

Thus, it is preferable to choose $\mathbf{e}_\omega = \mathbf{u}\theta$ since the singularity is in 2π and it provides the largest possible domain for the angle of rotation: $0 \leq \theta < 2\pi$. Moreover, we have in that case the following nice property: $\mathbf{L}_\omega^{-1}\mathbf{e}_\omega = \mathbf{e}_\omega$.

B. Stability and robustness of the rotation control

It is important to notice that the estimated task function $\widehat{\mathbf{e}}_\omega$ can be written as a function of $\mathbf{e}_\omega$:

$$\widehat{\mathbf{e}}_\omega = \mathbf{E}_\omega \mathbf{e}_\omega = \mu\tilde{\mathbf{A}}\mathbf{e}_\omega \tag{13}$$

where $\tilde{\mathbf{A}} = \widehat{\mathbf{A}}^{-1}\mathbf{A}$ and $\mu = 1/\|\tilde{\mathbf{A}}\mathbf{u}\|$ (see Section II-A). Choosing $f(\theta) = \theta$ we can set $\widehat{\mathbf{L}}_\omega^{-1} = \mathbf{I}$ in the control law without loss of generality, since $\widehat{\mathbf{L}}_\omega^{-1}\widehat{\mathbf{e}}_\omega = \widehat{\mathbf{e}}_\omega$. The closed loop equation for the rotational subsystem is thus:

$$\dot{\mathbf{e}}_\omega = -\lambda_\omega\mu\mathbf{L}_\omega\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{e}_\omega \tag{14}$$

The stability of this system can be analyzed regardless on the choice of $\mathbf{e}_\nu$ and its robustness domain is given by:

Theorem 1: The equilibrium point $\mathbf{e}_\omega = 0$ of the differential system (14) is locally asymptotically stable *if and only if* $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. A simple sufficient condition to ensure that is $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$. The equilibrium point is globally asymptotically stable if

$\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$. In that case, $\|\mathbf{e}_\omega\|$ decreases at each iteration of the control law.

The proof of the theorem is given in the Appendix. The necessary conditions imply that one must be careful in approximating the camera intrinsic (i.e. matrix $\widehat{\mathbf{A}}$) and extrinsic (i.e. matrix $\widehat{\mathbf{R}}$) parameters. A huge error on these parameters would imply an unstable behavior of the system. On the other hand, the conditions are very easily satisfied in practice since $\tilde{\mathbf{R}} > 0$ if the error on the angle of rotation is $\tilde{\phi} < \pi/2$ and $\tilde{\mathbf{A}} > 0$ if the conditions given in [16] are satisfied. The physical interpretation of the conditions tells us that large errors on camera intrinsic and extrinsic parameters can be tolerated. Finally, note that the condition $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$ is stronger than the conditions $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$ (i.e. if $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$, it is not sure that $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$). The conditions to ensure $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$ are [15]:

$$|\tilde{\phi}| + |\tilde{\alpha}| < \pi/2$$

where $\tilde{\alpha} = \arccos\left(\rho^{-1}((\tilde{\mathbf{A}} + \tilde{\mathbf{A}}^T)^{-1})/\sqrt{4\rho(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})}\right)$, $\rho(\mathbf{M})$ being the spectral radius of matrix $\mathbf{M}$. These conditions are in general easily satisfied using, for example, the camera parameters given by the manufacturer instead of the real ones. With respect to previous visual servoing approaches, the knowledge of the exacts conditions under which the system is stable provides new insights on the capabilities of the system. For example, if the necessary conditions are satisfied we can be sure that our visual servoing will converge. On the other hand, other visual servoing approaches could have unstable behavior (or not) in the presence the same calibration errors. The only way to know it is to test the control law in real conditions. That could be catastrophic if the control law diverges.

C. A new closed-loop system

Since $\mu \geq 1/\|\mathbf{A}\|$, the solution of the differential equation (14) can be bounded as follow:

$$\|\mathbf{e}_\omega(t)\| \leq \|\mathbf{e}_\omega(0)\|e^{-\lambda'_\omega t} \tag{15}$$

where $\lambda'_\omega = \lambda_\omega\sigma/\|\tilde{\mathbf{A}}\|$, σ being the unknown minimum singular value of $\frac{1}{2}(\tilde{\mathbf{R}}\tilde{\mathbf{A}} + \tilde{\mathbf{A}}^T\tilde{\mathbf{R}}^T)$. Thus, $\|\mathbf{e}_\omega(t)\|$ will converge exponentially to zero. Finally, since $\mathbf{e}_\omega(t)$ is completely defined by the differential system (14) and by its initial condition $\mathbf{e}_\omega(0)$, we can compute $\widehat{\mathbf{e}}_\omega(t)$ from equation (13) and plug it in equation (6) obtaining a new closed-loop system for $\mathbf{e}_\nu$. The new closed-loop equation for $\mathbf{e}_\nu$ is the sum of two nonlinear terms:

$$\dot{\mathbf{e}}_\nu = -\lambda_\nu\mathbf{f}_1(\mathbf{e}_\nu) + \lambda_\omega\mathbf{f}_2(\mathbf{e}_\nu, t) \tag{16}$$

where:

$$\mathbf{f}_1 = \mathbf{L}_\nu\tilde{\mathbf{R}}\widehat{\mathbf{L}}_\nu^{-1}\mathbf{E}_\nu\mathbf{e}_\nu \tag{17}$$

$$\mathbf{f}_2 = \mathbf{L}_\nu\tilde{\mathbf{R}}\left([\tilde{\mathbf{t}}]_\times + \tilde{\mathbf{R}}^T\mathbf{L}_\nu^{-1}\mathbf{L}_{\nu\omega}\tilde{\mathbf{R}} - \widehat{\mathbf{L}}_\nu^{-1}\widehat{\mathbf{L}}_{\nu\omega}\right)\widehat{\mathbf{e}}_\omega(t) \tag{18}$$

Equation (16) is a differential equation where $\mathbf{e}_\nu$ has to be determined and $\mathbf{e}_\omega(t)$ is a function of time. As we will see in the next sections, the form of the functions $\mathbf{f}_1$ and $\mathbf{f}_2$ depends on the choice of $\mathbf{e}_\nu$.

IV. The different approaches within the class

Since the rotation is completely controlled by $\mathbf{e}_\omega$, the choice of $\mathbf{e}_\nu$ will determine the control of the translation. Depending on the form of $\mathbf{e}_\nu$ we will have different approaches within the class. For example, one could directly use image features as proposed by [7]. However, even if the decoupling translation and rotation improves the behavior of the image-based visual servoing, some problems such as local minima and Jacobian singularity may still remain. Therefore, it is preferable to use a different $\mathbf{e}_\nu$, as we are going to show in this section.

A. Hybrid control

Hybrid control consists in using information available directly at the image level jointly with reconstructed information. It is "hybrid" in the sense that we are using the estimated rotation but we are not using the translation of the camera (even if it is available from the reconstruction). In [16] we proposed a hybrid approach called $2\frac{1}{2}$D visual servoing. The translation of the camera is not directly controlled in the Cartesian space. It is controlled by the error on the image coordinates of a reference point and the logarithm of the ratio between the current depth Z and the depth Z^* at the final position:

$$\mathbf{e}_\nu = \begin{bmatrix} x - x^* & y - y^* & \log(\frac{Z}{Z^*}) \end{bmatrix}^T \tag{19}$$

The corresponding image Jacobian is given by :

$$\mathbf{L}_\nu = \begin{bmatrix} \frac{-1}{Z} & 0 & \frac{x}{Z} \\ 0 & \frac{-1}{Z} & \frac{y}{Z} \\ 0 & 0 & \frac{-1}{Z} \end{bmatrix} \tag{20}$$

$$\mathbf{L}_{\nu\omega} = \begin{bmatrix} xy & -1-x^2 & y \\ 1+y^2 & -xy & -x \\ -y & x & 0 \end{bmatrix} \tag{21}$$

The estimated task function can be written as a function of the real task function: $\widehat{\mathbf{e}}_\nu = \mathbf{E}_\nu\mathbf{e}_\nu$. It is easy to show that $\widehat{\mathbf{L}}_\nu^{-1}\mathbf{E}_\nu = \eta\tilde{\mathbf{A}}\mathbf{L}_\nu^{-1}$ and $\mathbf{L}_\nu^{-1}\mathbf{L}_{\nu\omega} = Z[\mathbf{m}]_\times$ [16], where $\eta = \widehat{Z}/Z$. Then, equation (16) can be rewritten as follow:

$$\dot{\mathbf{e}}_\nu = \lambda_\nu\eta\mathbf{L}_\nu\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{L}_\nu^{-1}\mathbf{e}_\nu + \lambda_\omega\mathbf{L}_\nu\tilde{\mathbf{R}}\left[\tilde{\mathbf{t}} + (Z\tilde{\mathbf{R}}^T - \widehat{Z}\tilde{\mathbf{A}})\mathbf{m}\right]_\times\widehat{\mathbf{e}}_\omega \tag{22}$$

The closed-loop system is non-linear. However, it is easy to prove the local asymptotic stability of the control law. The following theorem, proved in the Appendix, improves the results given in [15] and [16]:

Theorem 2: The equilibrium point $\mathbf{e}_\nu = 0$ of the differential system (22) is locally asymptotically stable *if and only if* $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. A sufficient condition to ensure that is $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$.

As we already noticed, the sufficient conditions are very easily satisfied in practice. On the other hand, the global asymptotic stability of the system controlled with a simple proportional law is not easy to prove. Indeed, the system is non-linear and an adaptive control law was necessary in [16]

to prove that it exists a large region of attraction around the equilibrium point provided that $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$. In order to provide a larger stability domain we propose in the next section a new hybrid control scheme for which it is possible to provide *necessary and sufficient conditions* to ensure the global asymptotic stability.

B. A new hybrid control scheme

Even if the approach described above has a large stability domain and can tolerate large calibration errors, things can be improved using a new hybrid control scheme. In order to obtain a larger stability domain we propose the following different task function:

$$\mathbf{e}_\nu = \left[\begin{array}{ccc} \frac{Z}{Z^*}x - x^* & \frac{Z}{Z^*}y - y^* & \frac{Z}{Z^*} - 1 \end{array} \right]^T \qquad (23)$$

The corresponding image Jacobian is again an upper block-triangular matrix as in equation (9). The blocks $\mathbf{L}_v$ and $\mathbf{L}_{\nu\omega}$ of the matrix $\mathbf{L}$ are now:

$$\mathbf{L}_\nu = -\frac{1}{Z^*}\mathbf{I} \qquad (24)$$

$$\mathbf{L}_{\nu\omega} = [\mathbf{e}_\nu + \mathbf{m}^*]_\times \qquad (25)$$

where $\mathbf{m}^* = \left[\begin{array}{ccc} x^* & y^* & 1 \end{array} \right]^T$ is the vector containing the homogeneous coordinates of the reference point at the desired position (which are thus fixed during the servoing). In this case we have $\mathbf{E}_\nu = \tilde{\mathbf{A}}$. The closed-loop system is an *eventually time-varying linear system* (i.e. a system with a constant component and a time-varying component):

$$\dot{\mathbf{e}}_\nu = -\mathbf{M}(t)\mathbf{e}_\nu + \mathbf{l}(t) \qquad (26)$$

where $\mathbf{M}(t)$ is the sum of a constant component and a time-varying component:

$$\mathbf{M}(t) = \lambda_\nu \eta^* \tilde{\mathbf{R}}\tilde{\mathbf{A}} - \lambda_\omega (\eta^* [\hat{\mathbf{e}}_\omega(t)]_\times \tilde{\mathbf{A}} + [\tilde{\mathbf{R}}\hat{\mathbf{e}}_\omega(t)]_\times) \qquad (27)$$

with $\eta^* = \frac{\hat{Z}^*}{Z^*}$ (which is unknown but constant) and:

$$\mathbf{l}(t) = \left(\eta^* \tilde{\mathbf{R}}[\tilde{\mathbf{A}}\tilde{\mathbf{m}}^*]_\times - [\tilde{\mathbf{m}}^*]_\times \tilde{\mathbf{R}} - \frac{1}{Z^*}\tilde{\mathbf{R}} [\tilde{\mathbf{t}}]_\times \right) \hat{\mathbf{e}}_\omega(t) \qquad (28)$$

The global stability of the system is proved by the following theorem (a short proof is given in the Appendix):

Theorem 3: The equilibrium point $\mathbf{e}_\nu = 0$ of the differential system (26) is globally asymptotically stable *if and only if* the matrix $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. A simple sufficient condition to ensure the stability is $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$.

We have to note that these conditions do not ensure that $\|\mathbf{e}_\nu\|$ decreases at each iteration. Since $\|\mathbf{e}_\nu\|$ can temporarily increase, it is not sufficient to ensure the stability of the control law to avoid the failure of the servoing. We have to prove that it is possible to provide a control law for which at least the reference point stays always in the camera field of view. The time derivative of the norm of the task function is:

$$\frac{d\|\mathbf{e}_\nu\|^2}{dt} = -\frac{1}{2}\mathbf{e}_\nu^T\mathbf{M}(t)\,\mathbf{e}_\nu - \frac{1}{2}\mathbf{e}_\nu^T\mathbf{l}(t) \qquad (29)$$

Supposing $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$, and denoting $\sigma > 0$ the minimum singular value of $\frac{1}{2}(\tilde{\mathbf{R}}\tilde{\mathbf{A}} + \tilde{\mathbf{A}}^T\tilde{\mathbf{R}}^T)$, this equation can be bounded as follow:

$$\frac{d\|\mathbf{e}_\nu\|^2}{dt} \leq -\varphi \|\mathbf{e}_\nu\|^2 + \psi \|\mathbf{e}_\nu\| \qquad (30)$$

where $\varphi = \lambda_\nu a - \lambda_\omega b$ and $\psi = \lambda_\omega$, and:

$$a = \eta^*\sigma \qquad (31)$$

$$b = (1 + \eta^*\|\tilde{\mathbf{A}}\|)\|\hat{\mathbf{e}}_\omega(0)\| \qquad (32)$$

$$c = \left((1 + \eta^*\|\tilde{\mathbf{A}}\|)\|\mathbf{m}^*\| + \frac{1}{Z^*}\|\tilde{\mathbf{t}}\|\right) \|\hat{\mathbf{e}}_\omega(0)\| \qquad (33)$$

Setting $x(t) = \|\mathbf{e}\|^2$, we obtain the following Bernoulli differential equation:

$$\dot{x} = -\varphi\,x + \psi x^{\frac{1}{2}} \qquad (34)$$

whose solution is:

$$x(t) = e^{-\int \varphi dt}\left(x(0) + \frac{1}{2}\int \psi e^{\frac{1}{2}\int \varphi dt}dt\right)^2 \qquad (35)$$

The norm of the task function can be bounded by a maximum value:

$$\|\mathbf{e}_\nu(t)\| \leq \|\mathbf{e}_\nu(0)\| + \frac{\psi}{\varphi} \leq \|\mathbf{e}_\nu\|_{\max} \qquad (36)$$

Posing $d = \|\mathbf{e}_\nu\|_{\max} - \|\mathbf{e}_\nu(0)\|$ (and supposing that $\|\mathbf{e}_\nu\|_{\max} > \|\mathbf{e}_\nu(0)\|$), the solution of the last inequality is:

$$\lambda_\omega/\lambda_\nu \leq (c + b\,d)/(a\,d) \qquad (37)$$

In conclusion, it is always possible to tune $\lambda_\omega/\lambda_\nu$ such that $\|\mathbf{e}_\nu\|$ decreases at each iteration and the reference point remains in the image. For example, one can temporarily choose $\lambda_\omega = 0$, even if in this case $\mathbf{e}_\omega$ will not decrease. Indeed, choosing $\lambda_\omega(0) = 0$ the inequality (37) is satisfied and the norm of $\mathbf{e}_\nu$ decreases.

Since a, b, c and d are unknown, we have to use an adaptive control law and increase the ratio $\lambda_\omega/\lambda_\nu$ in order to decrease the norm of $\mathbf{e}_\omega$ at the same time. In Section V, we describe more in details such a control strategy.

C. Comparison between the two hybrid schemes

We show now that, in absence of calibration errors, the two hybrid control schemes produce the same behavior in the image for the reference point (i.e. a straight line). Indeed, in the case of $2\frac{1}{2}D$ visual servoing, imposing that $\dot{\mathbf{e}}_\nu = -\lambda\mathbf{e}_\nu$, with the initial conditions $x(0) = x_0$, $y(0) = y_0$ and $Z(0) = Z_0$, we have to solve the following differential system :

$$\left\{ \begin{array}{rcl} \dot{x} & = & -\lambda(x - x^*) \\ \dot{y} & = & -\lambda(y - y^*) \\ d(\ln(\frac{Z}{Z^*}))/dt & = & -\lambda\ln(\frac{Z}{Z^*}) \end{array} \right. \qquad (38)$$

The solution of this system is :

$$\left\{ \begin{array}{rcl} x(t) & = & x^* - (x^* - x_0)e^{-\lambda t} \\ y(t) & = & y^* - (y^* - y_0)e^{-\lambda t} \end{array} \right. \qquad (39)$$

The trajectory of the point in the image plane is a straight line while the corresponding point in the 3D space has the following trajectory :

$$\begin{cases} X(t) &=& (X^* - (X^* - X_0\frac{Z^*}{Z_0})e^{-\lambda t})\left(\frac{Z_0}{Z^*}\right)^{e^{-\lambda t}} \\ Y(t) &=& (Y^* - (Y^* - Y_0\frac{Z^*}{Z_0})e^{-\lambda t})\left(\frac{Z_0}{Z^*}\right)^{e^{-\lambda t}} \\ Z(t) &=& Z^*\left(\frac{Z_0}{Z^*}\right)^{e^{-\lambda t}} \end{cases} \quad (40)$$

The differential system of the new hybrid control scheme can be reduced to the following one :

$$\begin{cases} \dot{x} &=& -\lambda\frac{Z^*}{Z}(x - x^*) \\ \dot{y} &=& -\lambda\frac{Z^*}{Z}(y - y^*) \\ d(\ln(\frac{Z}{Z^*}))/dt &=& -\lambda(1 - \frac{Z^*}{Z}) \end{cases} \quad (41)$$

Once again the trajectory of the point in the image is a straight line. However the velocity of the point on the line is different :

$$\begin{cases} x(t) &=& x^* - (x^* - x_0)\dfrac{\exp^{-\lambda t}}{\frac{Z^*}{Z_0} + (1 - \frac{Z^*}{Z_0})\exp^{-\lambda t}} \\ y(t) &=& y^* - (y^* - y_0)\dfrac{\exp^{-\lambda t}}{\frac{Z^*}{Z_0} + (1 - \frac{Z^*}{Z_0})\exp^{-\lambda t}} \end{cases} \quad (42)$$

On the other hand, the trajectory of the 3D point in the camera frame is now a straight line :

$$\begin{cases} X(t) &=& X^* - (X^* - X_0)\exp^{-\lambda t} \\ Y(t) &=& Y^* - (Y^* - Y_0)\exp^{-\lambda t} \\ Z(t) &=& Z^* - (Z^* - Z_0)\exp^{-\lambda t} \end{cases} \quad (43)$$

In conclusion, both schemes will have a very similar behavior since the trajectory of the point in the image is the same in the ideal case. This is especially true near the convergence when $\frac{Z}{Z^*} \approx 1$. In this case, we have $\ln(\frac{Z}{Z^*}) \approx \frac{Z}{Z^*} - 1$, and both methods must produce the same results. The main advantage in using the new hybrid control law is that it has been proved to tolerate larger amount of calibration errors. On the other hand, the new hybrid control could be more sensible to measurement errors since the task function is not directly computed as a difference of image data but it contains the ratio $\frac{Z}{Z^*}$ which is obtained from the projective reconstruction. However, it is possible to switch between the two control laws near the convergence. This can be done without introducing any discontinuity in the control law since when $\ln(\frac{Z}{Z^*}) \approx \frac{Z}{Z^*} - 1$ the two control laws are practically the same.

D. Model-free position-based control

In this section, we analyze the stability of a new model-free position-based control scheme. Position-based control consists in using the estimated rotation and translation (up to a scale factor). Instead of using the task function proposed by Basri [1] or Taylor [22], we choose the following task function:

$$\mathbf{e}_\nu = \mathbf{t}/d^* \quad (44)$$

which can be computed from the homography matrix. The Jacobian of the task is still triangular and its blocks are:

$$\mathbf{L}_\nu = -\frac{1}{d^*}\mathbf{I} \quad (45)$$

$$\mathbf{L}_{\nu\omega} = [\mathbf{e}_\nu]_\times \quad (46)$$

If we compare these matrices with those given in equation (20), we can see that they are extremely similar. Since $\widehat{\mathbf{e}}_\nu = \mathbf{E}_\nu\mathbf{e}_\nu$ with $\mathbf{E}_\nu = \tilde{\mathbf{A}}$, the closed-loop equation is again an eventually time-varying system:

$$\dot{\mathbf{e}}_\nu = -\mathbf{M}(t)\mathbf{e}_\nu + \mathbf{l}(t) \quad (47)$$

where:

$$\mathbf{M}(t) = \lambda_\nu\eta^*\tilde{\mathbf{R}}\tilde{\mathbf{A}} - \lambda_\omega(\eta^*[\widehat{\mathbf{e}}_\omega(t)]_\times\tilde{\mathbf{A}} + [\tilde{\mathbf{R}}\widehat{\mathbf{e}}_\omega(t)]_\times) \quad (48)$$

with $\eta^* = \frac{\widehat{d}^*}{d^*}$ (which is unknown but constant) and:

$$\mathbf{l}(t) = -\frac{1}{d^*}\tilde{\mathbf{R}}\,[\tilde{\mathbf{t}}]_\times\widehat{\mathbf{e}}_\omega(t) \quad (49)$$

Equation (47) is very similar to equation (26) and the global stability can be proved exactly as for Theorem 3:

Theorem 4: The equilibrium point $\mathbf{e}_\nu = 0$ of the differential system (47) is globally asymptotically stable *if and only if* the matrix $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. A simple sufficient condition to ensure that is $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$.

This theorem is not sufficient to ensure the convergence of the servoing. Indeed, nothing is said about the constraint that the image features must stay in the camera field of view. It could be possible to use a complex strategy incorporating this constraint with position-based visual servoing. However, a simpler solution is to use the new hybrid control scheme proposed in this paper which is not only stable under the same necessary and sufficient conditions but also guarantees that at least the reference point will stay in the image during the servoing. Consequently, choosing as reference one point close to the centroid of the set of points, there will be several points always visible in the image.

V. Experimental results

In order to validate the new hybrid visual servoing approach described in section IV-B, a series of experiments were carried out at the University of Cambridge. The experimental setup consists of a camera mounted on a 6 d.o.f. robot manipulator. The vision-based controller rate is $0.08s$ per iteration. In order to position the robot's end-effector with respect to a contour (which shape is not "a priori" known), we use a teaching-by-showing technique. Initially, a reference image is stored. Figure 2(a) displays two contours segmented from the reference image. The contour on the left is selected as the reference contour. The robot is displaced such that the camera performs a large rotation and the initial centroid of the contour is near to the reference one (see Figure 2(b)). Following, the initial and the refer-

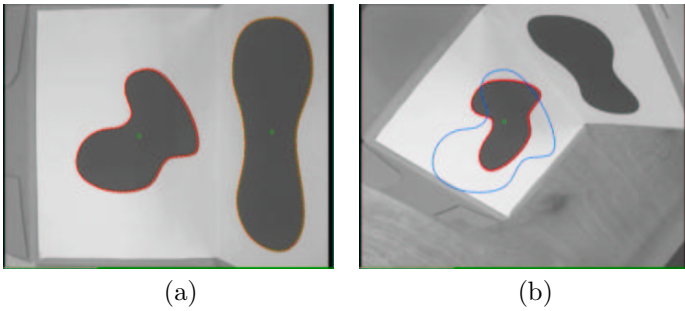(a)   (b)

Fig. 2.   Images of the contours for the desired and the initial camera position. The reference contour is superposed in the right image.

ence contours are automatically matched and the homography **H** between the contours is computed using the method described in [5]. The centroid of the contour is used as reference point to build the task function from equation (23) (the ratio Z/Z^* is obtained from the homography). This adds an error in the control law as the projection of the 3D contour's centroid does not generally coincide with the centroid of the projected 2D contour. The rotation is extracted from **H** and the task function $\mathbf{e}_\omega$ is set to $\mathbf{e}_\omega = \mathbf{u}\theta$. The control law sent to the robot is given by equation (4). Calibration errors in the control law are introduced for the hand-eye transformation, the camera parameters, the reference distance of the camera to the centroid of contour. For example, the real distance was $d^* = 40$ cm but we used $\widehat{d^*} = 60$ cm in the calculation of the control law. The camera parameters used to compute the control law where: $fk_u = 780$, $fk_v = 810$, $u_0 = 370$, $v_0 = 230$, $\varphi = \pi/2$. Although all the calibration errors, the visual servoing system is exponentially stable even if the task function is not strictly decreasing (i.e. the error can increase temporarily). However, the convergence is obtained only if the error is bounded in order to maintain the contour (or at least a part of it) in the camera field of view.

A. Constant gains

The first experiment shows that using a control law with constant gains (λ_ν and λ_ω are both fixed to 0.02) the contour may leave the camera field of view. Indeed, Figure 3(b) shows that the norm of $\mathbf{e}_\omega$ decreases at each iteration since the matrix $\widetilde{\mathbf{R}}\widetilde{\mathbf{A}}$ is definite positive as required by Theorem 2. However, the norm of $\mathbf{e}_\nu$, plotted in Figure 3(a), increases at the same time. Details on the task function ($\mathbf{e}_\nu$, $\mathbf{e}_\omega$) are shown in Figure 3(c) and (d). The x component of $\mathbf{e}_\nu$ (which is related to the u coordinates of the centroid) decreases (the centroid moves to the left of the image). Even if the contour goes out of the camera field of view, the control law is not unstable. There is only an overshoot due to the rotation acting on the translational control. Had the image plane been infinite, after a finite amount of time which could be computed from equation (29) if all the parameters were known, the error would start to decrease. Figures 4(a) to (d) display four images of the video taken during the experiment. The contour goes out of the image to the left and the visual servoing is
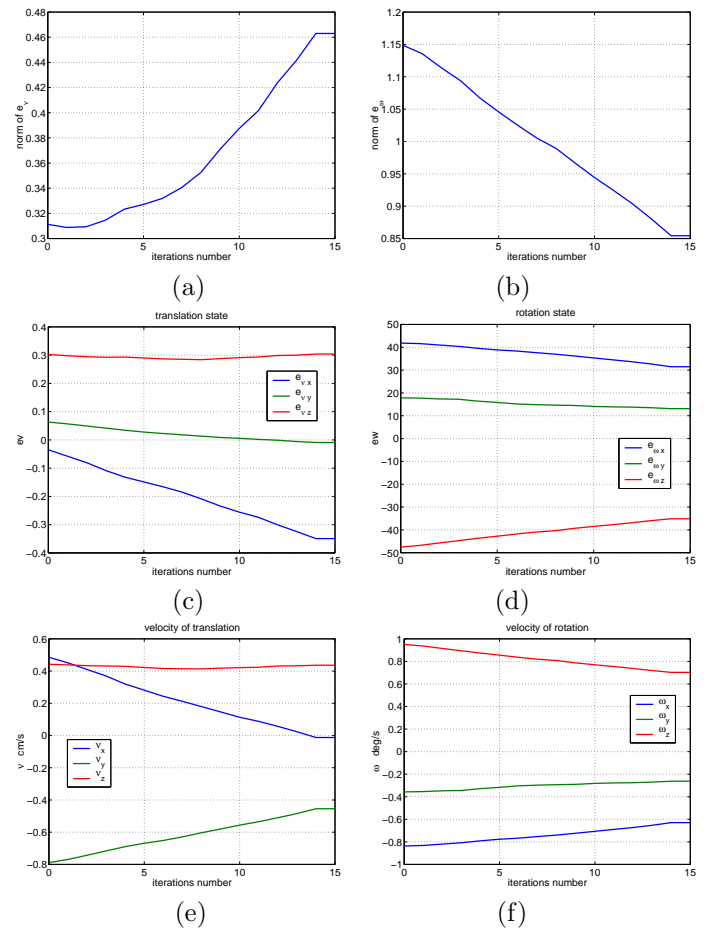

(a)   (b)

(c)   (d)

(e)   (f)

Fig. 3.   Results of the experiment with fixed λ_ν and λ_ω. (a) $\|\mathbf{e}_\nu\|$. (b) $\|\mathbf{e}_\omega\|$. (c) $\mathbf{e}_\nu$. (d) $\mathbf{e}_\omega$. (e) Velocity of translation. (f) Velocity of rotation.
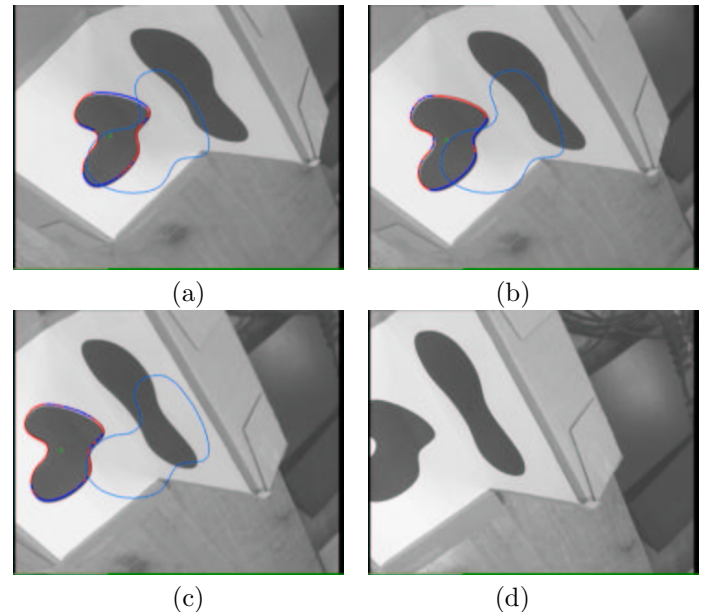

(a)   (b)

(c)   (d)

Fig. 4.   Four images of the experiment with fixed λ_ν and λ_ω. (a) image I.1 (iteration 3). (b) image I.2 (iteration 6). (c) image I.3 (iteration 9). (d) image I.4 (iteration 15).

stopped when the centroid is too close to the border of

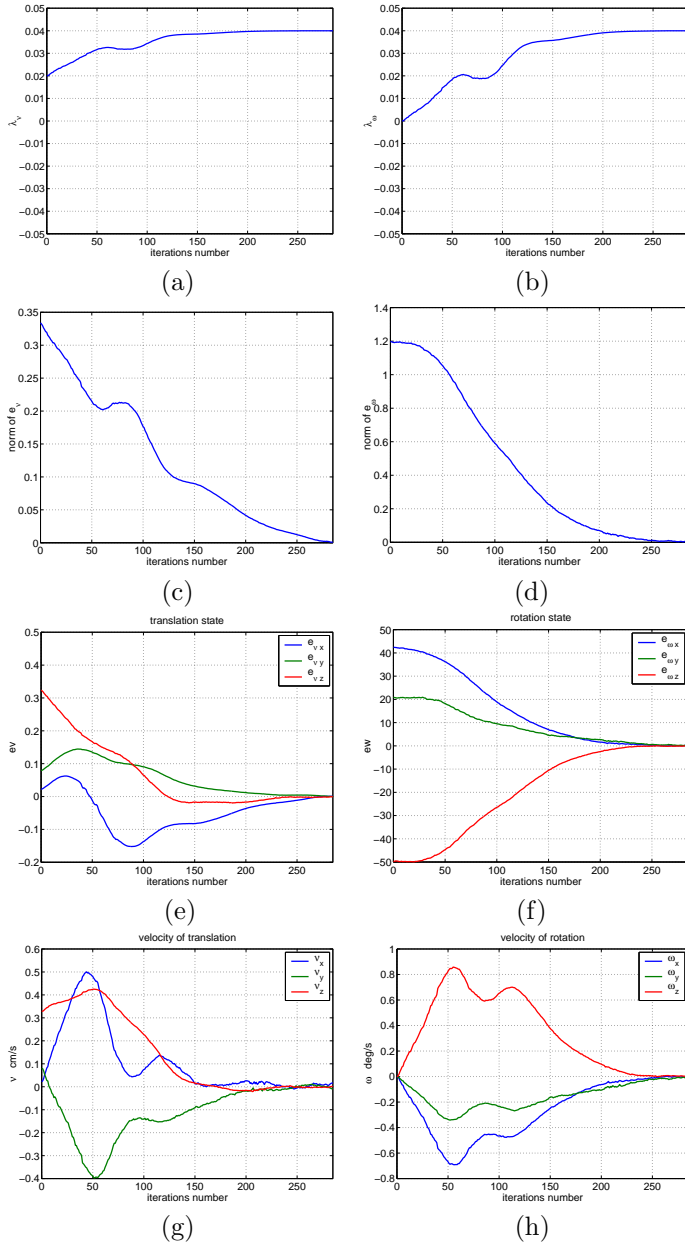the image (iterations 13, 1 second after starting).

Fig. 5. Results of the experiment with varying λ_ν and λ_ω. (a) λ_ν. (b) λ_ω. (c) $\|\mathbf{e}_\nu\|$. (d) $\|\mathbf{e}_\omega\|$. (e) $\mathbf{e}_\nu$. (f) $\mathbf{e}_\omega$. (g) Velocity of translation. (h) Velocity of rotation.

B. Varying gains

In this experiment we prove that, instead of fixing the gains, a more complex control law can be chosen in order to keep the centroid of the contour in the camera field of view. The gains in the control law are computed as follows (such that $\lambda \leq \lambda_\nu \leq 2\lambda$ and $0 \leq \lambda_\omega \leq 2\lambda$, where $\lambda = 0.02$):

$$\lambda_\nu = \begin{cases} \lambda + \lambda \frac{\|\mathbf{e}_\nu(0)\| - \|\mathbf{e}_\nu\|}{\|\mathbf{e}_\nu(0)\|} & \text{if} \quad \|\mathbf{e}_\nu\| \leq \|\mathbf{e}_\nu(0)\| \\ \lambda & \text{if} \quad \|\mathbf{e}_\nu\| \geq \|\mathbf{e}_\nu(0)\| \end{cases} \quad (50)$$

$$\lambda_\omega = \begin{cases} \lambda_\nu \frac{\|\mathbf{e}_\nu(0)\| - \|\mathbf{e}_\nu\|}{\|\mathbf{e}_\nu(0)\|} & \text{if} \quad \|\mathbf{e}_\nu\| \leq \|\mathbf{e}_\nu(0)\| \\ 0 & \text{if} \quad \|\mathbf{e}_\nu\| \geq \|\mathbf{e}_\nu(0)\| \end{cases} \quad (51)$$
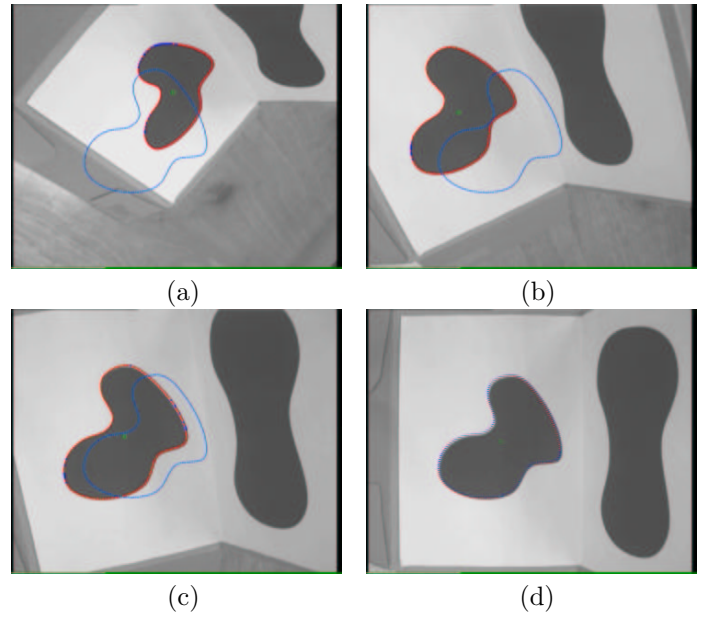

Fig. 6. Four images of the experiment with varying λ_ν and λ_ω. (a) image II.1 (iteration 50). (b) image II.2 (iteration 150). (c) image II.3 (iteration 200). (d) image II.4 (iteration 280).

Note that, tuning gains is very easy since we simply set $\lambda = 0.02$. Then, λ_ν and λ_ω automatically change as follow. At the first iteration ($t = t_0 = 0$ s) we have $\|\mathbf{e}_\nu\| = \|\mathbf{e}_\nu(0)\|$ and $\lambda_\nu(0) = \lambda$ and $\lambda_\omega(0) = 0$. Therefore, the ratio $\lambda_\omega/\lambda_\nu = 0$ and the system is stable if $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$. Since $\lambda_\omega(0) = 0$, the norm of $\mathbf{e}_\nu$ must decrease and the norm of $\mathbf{e}_\omega$ must stay constant. Consequently, at the next iteration ($t = t_1 = 0.08$ s) we have $\|\mathbf{e}_\nu(t_1)\| < \|\mathbf{e}_\nu(t_0)\|$ and both λ_ν and λ_ω start to increase. Figure 5(a) and 5(b) shows the behavior of λ_ν and λ_ω during the servoing. Contrarily to the previous experiment, now the norm of $\mathbf{e}_\omega$ and $\mathbf{e}_\nu$ both decrease at the same time. Increasing the ratio $\lambda_\omega/\lambda_\nu$ too much (i.e. if equation (37) is not verified) could cause the norm of $\mathbf{e}_\nu$ to stop decreasing (see Figure 5(c) after iteration 50, i.e. 4 s). For example, at iteration 60 the norm of $\mathbf{e}_\nu$ starts to increase. The value of $\lambda_\omega/\lambda_\nu$ corresponds to 0.625 which is the maximum value. However, as soon as $\|\mathbf{e}_\nu\|$ increases, λ_ω decreases faster than λ_ν such that $\|\mathbf{e}_\nu\|$ starts to decrease again. Both gains are increased near the convergence in order to speed up the servoing when the error becomes small. The servoing is stopped when the maximal error on all the points of the contour is smaller that 0.5 pixels. Figures 6(a) to 6(d) show four images of the sequence taken during the servoing. Since the error on the reference point is kept small, it is more likely that the remaining points of the contour stay in the camera field of view (especially if the reference point is chosen near the image center).

C. Behavior when necessary conditions are not satisfied

In order to verify the necessity of the stability conditions given by the theorems, we present now an example when the system is not stable. For safety reasons, we prefer to perform a simulation instead of use an unstable control law with the real robot. In this example, we use a very bad approximation $\widehat{\mathbf{A}}$ of the true camera intrinsic parameters $\mathbf{A}$

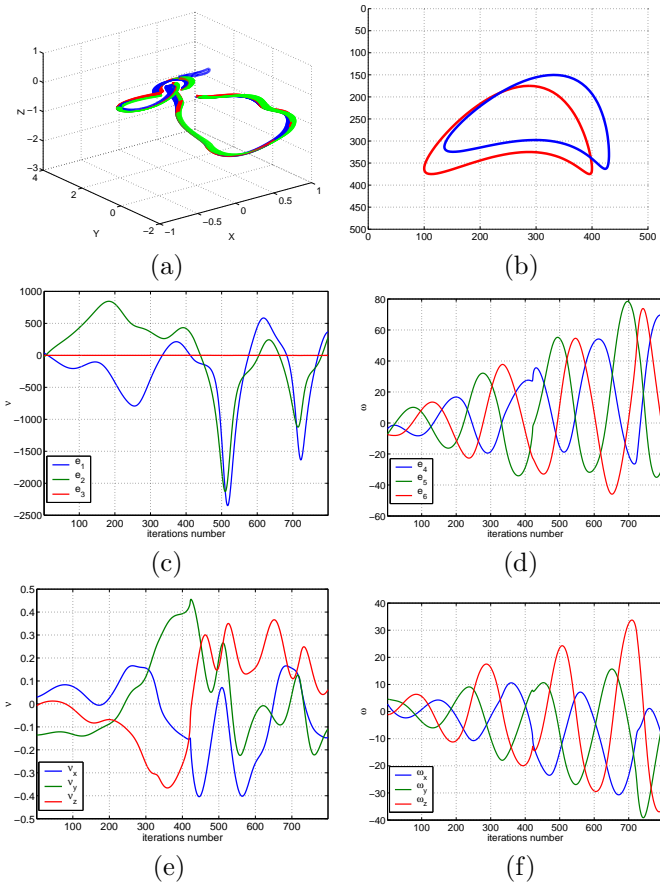(a)

(b)

(c)

(d)

(e)

(f)

Fig. 7. The system is not stable when the necessary conditions are not satisfied. Unstable behavior is oscillatory and divergent. Camera frame trajectory: $\vec{x}$ axis (green) , $\vec{y}$ axis (blue), $\vec{z}$ axis (red). (b) Small displacement between initial (blue) and reference (red) contours. (c) Task function $\mathbf{e}_\nu$. (d) Task function $\mathbf{e}_\omega$. (e) Velocity of translation. (f) Velocity of rotation.

(we have 50 % error on the focal length and principal point). We also use a very bad approximation $\widehat{\mathbf{R}}$ such that the rotation matrix $\tilde{\mathbf{R}} = \widehat{\mathbf{R}}^{-1}\mathbf{R}$ has an angle $\tilde{\phi} = 100°$. In that case, the eigenvalues of the matrix $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ are $\lambda_1 = -0.15+i0.8$ $\lambda_2 = -0.15 - i0.8$ and $\lambda_3 = 0.68$. The two complex conjugate eigenvalues have negative real part. Therefore, the necessary conditions are not satisfied and the system is not stable. Figure 7 shows the behavior of the system when the initial error is small (Figure 7(b)). The error grows (Figure 7(c) and (d)) with an oscillatory behavior since the camera turns around the contour (Figure 7(a)). Note that, with such a behavior, in a real experiment the contour would rapidly get out of the camera field of view. However, in the simulation we can suppose that the camera CCD is large enough.

VI. Conclusion

In this paper, we proved the stability of a new class of visual servoing systems and its robustness to camera and hand/eye calibration errors. The most important result is that, using a new hybrid control scheme, it is possible to find necessary and sufficient conditions on the calibration errors for the global stability of the control law. In addition, we proposed a simple adaptive control strategy in order to keep

the centroid of object always in the camera field of view. It would be interesting to extend the analysis to the robustness with respect to measurements errors. Indeed, the new class is based on the reconstruction of the camera displacement between two views. This reconstruction step generally amplifies the effects of the image noise on the control law, even if the experimental results we have obtained were not very sensitive to noise due to the quite simple images we have considered.

Appendix

Proof: [**Theorem 1**] The local asymptotic stability of the non-linear system (14) can be deduced from the stability analysis of the linearized system. In this case we have $\mathbf{L}_\omega|_{\mathbf{e}_\omega=0} = \mathbf{I}$ and the linear system is:

$$\dot{\mathbf{e}}_\omega = -\lambda_\omega \mu \tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{e}_\omega \qquad (52)$$

It is well known that $\forall \lambda_\omega > 0$ and $\forall \mu > 0$ the system is stable if and only if the matrix $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. Simple sufficient conditions to ensure that $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part can be obtained as follows. The eigenvalues λ of $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ and their correspondent eigenvectors $\mathbf{x}$ are related by the equation:

$$\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{x} = \lambda \mathbf{x} \qquad (53)$$

Multiplying the left and right side of this equation by $\mathbf{x}^T\tilde{\mathbf{R}}^T$, we obtain:

$$\mathbf{x}^T\tilde{\mathbf{R}}^T\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{x} = \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} = \lambda\mathbf{x}^T\tilde{\mathbf{R}}^T\mathbf{x} = \lambda\mathbf{x}^T\tilde{\mathbf{R}}\mathbf{x}$$

and the eigenvalue λ can be computed as:

$$\lambda = \frac{\mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x}}{\mathbf{x}^T\tilde{\mathbf{R}}\mathbf{x}} \qquad (54)$$

which is positive $\forall \mathbf{x} \neq 0$ if $\tilde{\mathbf{A}} > 0$ and $\tilde{\mathbf{R}} > 0$. Finally, $\tilde{\mathbf{R}} > 0$ if the error on the estimated rotation angle $\tilde{\phi} < \pi/2$ and $\tilde{\mathbf{A}} > 0$ if the conditions given in [16] are satisfied.

The sufficient conditions for the global asymptotic stability of the non-linear system (14) can be obtained from the Lyapunov method. Consider the following Lyapunov function:

$$\mathbf{V}(\mathbf{e}_\omega) = \frac{1}{2}\mathbf{e}_\omega^T\mathbf{e}_\omega$$

which is positive $\forall \mathbf{e}_\omega \neq 0$. The derivative of $\mathbf{V}(\mathbf{e}_\omega)$ is:

$$\dot{\mathbf{V}}(\mathbf{e}_\omega) = \mathbf{e}_\omega^T\dot{\mathbf{e}}_\omega$$

Using equation (14), knowing that $\mathbf{e}_\omega^T\mathbf{L}_\omega = \mathbf{e}_\omega^T$ we have:

$$\dot{\mathbf{V}}(\mathbf{e}_\omega) = -\lambda_\omega \mu \mathbf{e}_\omega^T\mathbf{L}_\omega\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{e}_\omega = -\lambda_\omega \mu \mathbf{e}_\omega^T\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{e}_\omega \qquad (55)$$

Then $\dot{\mathbf{V}}(\mathbf{e}_\omega) < 0$ if $\tilde{\mathbf{R}}\tilde{\mathbf{A}} > 0$ $\forall \lambda_\omega > 0$ and $\forall \mu > 0$.

Proof: [**Theorem 2**] The local stability analysis is obtained linearizing the nonlinear system. In this case we have $\mathbf{L}_\nu|_{\mathbf{e}_\nu=0} = \mathbf{L}_\nu^*$ and the linear system is:

$$\dot{\mathbf{e}}_\nu = -\lambda_\nu \eta^*\mathbf{L}_\nu^*\tilde{\mathbf{R}}\tilde{\mathbf{A}}\mathbf{L}_\nu^{*-1}\mathbf{e}_\nu \qquad (56)$$

The matrix $\mathbf{L}_\nu^* \tilde{\mathbf{R}} \, \tilde{\mathbf{A}} \mathbf{L}_\nu^{*-1}$ is similar to the matrix $\tilde{\mathbf{R}} \, \tilde{\mathbf{A}}$ and so they have the same eigenvalues. Since $\eta^* = \widehat{Z}^*/Z^* > 0$ (which is always true), the matrix has positive eigenvalues if $\tilde{\mathbf{R}} > 0$ and $\tilde{\mathbf{A}} > 0$.

Proof: [**Theorem 3**] The closed loop system is an *eventually time varying linear system*:

$$\dot{\mathbf{e}}_\nu = -\mathbf{M}(t)\mathbf{e}_\nu + \mathbf{l}(t)$$

The term $\mathbf{l}(t)$ does not depend on $\mathbf{e}_\nu$ and:

$$\lim_{t \to \infty} \|\mathbf{l}(t)\| \leq \left((1 + \eta^*\|\tilde{\mathbf{A}}\|)\|\mathbf{m}^*\| + \frac{1}{Z^*}\|\tilde{\mathbf{t}}\| \right) \|\widehat{\mathbf{e}}_\omega(t)\| = 0$$

Therefore, the stability and the decreasing of $\mathbf{e}_\nu$ towards 0 only depend on the matrix $\mathbf{M}(t)$. This matrix is the sum of a constant part and a perturbation term:

$$\mathbf{M}(t) = \lambda_\nu \mathbf{C} - \lambda_\omega \mathbf{B}(t) \tag{57}$$

where $\mathbf{C} = \eta^* \tilde{\mathbf{R}}\tilde{\mathbf{A}}$ and $\mathbf{B}(t) = \eta^*[\widehat{\mathbf{e}}_\omega(t)]_\times \, \tilde{\mathbf{A}} + [\tilde{\mathbf{R}}\widehat{\mathbf{e}}_\omega(t)]_\times$. The perturbation term satisfies the following properties:

$$\|\mathbf{B}(t)\mathbf{e}_\nu\| \leq \|\mathbf{B}(t)\|\|\mathbf{e}_\nu\|$$

$$\lim_{t \to \infty} \|\mathbf{B}(t)\| \leq (1 + \eta^*\|\tilde{\mathbf{A}}\|)\|\widehat{\mathbf{e}}_\omega(t)\| = 0$$

The system is stable if and only if $\tilde{\mathbf{R}}\tilde{\mathbf{A}}$ has eigenvalues with positive real part. A more detailed proof is given in [12].

ACKNOWLEDGMENTS

REFERENCES

[1] R. Basri, E. Rivlin, and I. Shimshoni, "Visual homing: Surfing on the epipoles," in *IEEE Int. Conf. on Computer Vision*, pp. 863–869, Bombay, India, January 1998.
[2] F. Caccavale, C. Natale, B. Siciliano and L. Villani, "6-DOF Impedance control based on angle/axis representation," *IEEE Trans. Robot. Automat.*, 15(2):289-300, Apr. 1999.
[3] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The confluence of vision and control*, D. Kriegman, G. Hager, A. Morse, Eds., vol. 237, *LNCIS Series*, pp. 66-78. Springer Verlag, 1998.
[4] C. C. Cheah, S. Kawamura and S. Arimoto, "Feedback control for Robotic Manipulators with uncertain kinematics and dinamics," *IEEE Int. Conf. Robotics and Automation*, vol. 4, pp. 3607-3612, Leuven, Belgium, May 1998.
[5] G. Chesi, E. Malis, and R. Cipolla, "Automatic segmentation and matching of planar contours for visual servoing," in *IEEE Int. Conf. on Robotics and Automation*, April 2000.
[6] P. I. Corke, and S. Hutchinson, "A new partitioned approach to image-based visual servo control," in *IEEE Int. Conf. Decision and Control*, pp. 2521–2526, Sydney, December 2000.
[7] K. Deguchi, "Optimal motion control for image-based visual servoing by decoupling translation and rotation," in *IEEE Int. Conf. Intell. Rob. and Syst.*, vol.2, pp. 705-711, Oct. 1998.
[8] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Automat.*, 8(3):313-326, June 1992.
[9] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," In *3rd Int. Symp. Experimental Robotics*, Kyoto, Japan, Oct. 1993.
[10] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," *Int. J. Pattern Recognit. Artif. Intell.*, 2(3):485-508, 1988.
[11] R. Kelly, R. Carelli, O. Nasisi, B. Kuchen and F. Reyes, "Stable Visual Servoing of Camera-in-hand robotic systems," *IEEE/ASME Trans. Mechatron.*, 5(1):39-48, Mar. 2000.
[12] H.K. Khalil, "Nonlinear Systems," *Prentice Hall*, 1996.
[13] K. Hashimoto, Ed., *Visual Servoing: Real Time Control of Robot manipulators based on visual sensory feedback*, vol. 7 of *World Scientific Series in Robotics and Automated Systems*, World Scientific Press, Singapore, 1993.
[14] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Automat.*, 12(5):651-670, Oct. 1996.
[15] E. Malis, "Contributions à la modélisation et à la commande en asservissement visuel," PhD thesis, IRISA, Rennes, Nov. 1998.
[16] E. Malis, F. Chaumette, and S. Boudet, "2 1/2D visual servoing," *IEEE Trans. Robot. Automat.*, 15(2):234-246, Apr. 1999.
[17] E. Malis, F. Chaumette, "2 1/2D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *Int. J. Computer Vision*, 37(1):79-97, June 2000.
[18] E. Malis, "Hybrid vision-based robot control robust to large calibration errors on both intrinsic and extrinsic camera parameters," in *Eur. Control Conf.*, pp. 2898-2903, Porto, Portugal, Sept. 2001.
[19] Y. Mezouar, F. Chaumette, "Path planning in image space for robust visual servoing," *IEEE Int. Conf. Robotics Automation*, vol. 3, pp. 2759-2764, San Francisco, CA, April 2000.
[20] P.Y. Oh, P.K. Allen, "Visual Servoing by partitioning degrees of fredom," *IEEE Trans. Robot. Automat.*, 17(1):1-17, Feb. 2001.
[21] C. Samson, M. Le Borgne, and B. Espiau, *Robot Control: the Task Function Approach*, vol. 22 of *Oxford Engineering Science Series*. Clarendon Press, Oxford, UK, 1991
[22] C.J. Taylor, J.P. Ostrowski, and S.H. Jung, "Robust Vision-based Pose Control," *Int. Conf. Robotics Automation*, vol. 3, pp. 2734-2740, San Francisco, CA, Apr. 2000.
[23] W. J. Wilson, C. C. W. Hulls, and G. S. Bell, "Relative end-effector control using cartesian position-based visual servoing," *IEEE Trans. Robot. Automat.*, 12(5):684-696, Oct. 1996.
[24] P. Zanne, G. Morel, and F. Plestan, "Robust vision based 3D trajectory tracking using sliding mode control," *IEEE Int. Conf. Robotics and Automation*, Vol. 3, pp. 2088-2093, CA, Apr. 2000.

Ezio Malis was born in Gorizia, Italy, in 1970. He graduated both from the University Politecnico di Milano and from the Ecole Supérieure d'Electricité (Supélec), Paris, in 1995, and the PhD degree from the University of Rennes, France, in 1998.

He performed three years of joint research work with IRISA/INRIA Rennes and with the national French company of electricity power (EDF). He spent two years as research associate at the University of Cambridge, Cambridge, UK, where he participated in the ESPRIT project VIGOR (visually guided robots using uncalibrated cameras). He joined the ICARE team at INRIA Sophia-Antipolis in 2000 as Research Scientist. His research interests include automatics, robotics, computer vision, and in particular vision-based control.

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

Since 1990, he has been with IRISA/INRIA in Rennes. His research interests include robotics, computer vision, and especially the coupling of these two research domains (vision-based control, active vision and purposive vision).

Dr Chaumette received the AFCET/CNRS Prize for the best French thesis in Automatic Control in 1991.