HAL open science

Visual Servoing Invariant to Changes in Camera Intrinsic Parameters

Ezio Malis

▶ To cite this version:

Ezio Malis. Visual Servoing Invariant to Changes in Camera Intrinsic Parameters. RR-4309, INRIA. 2001. inria-00072278

HAL Id: inria-00072278

https://inria.hal.science/inria-00072278v1

Submitted on 23 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Visual servoing invariant to changes in camera intrinsic parameters

Ezio Malis

N° 4309

Novembre 2001

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4309--FR+ENG

Visual servoing invariant to changes in camera intrinsic parameters

Ezio Malis

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Icare

Rapport de recherche n° 4309 — Novembre 2001 — 31 pages

Abstract: This research report presents a new visual servoing scheme which is invariant to changes in camera intrinsic parameters. Current visual servoing techniques are based on the learning of a reference image with the same camera used during the servoing. With the new method it is possible to position a camera (with eventually varying intrinsic parameters), with respect to a non-planar object, given a "reference image" taken with a completely different camera. The necessary and sufficient conditions for the local asymptotic stability show that the control law is robust in the presence of large calibration errors. Local stability implies that the system can accurately track a path in the invariant space. The path can be chosen such that the camera follows a straight line in the Cartesian space. Simple sufficient conditions are given in order to keep the tracking error bounded. This promising approach has been successfully tested with an eye-in-hand robotic system.

Key-words: visual servoing, projective invariance, uncalibrated camera, path planning

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Asservissement visuel independant des parametres intrinseques de la caméra

Résumé : Toutes les méthodes d'asservissement visuel non calibré qui n'utilisent pas de connaissances a priori sur l'environnement sont basées sur l'apprentissage préalable d'une image d'un objet par rapport auquel on veut se positionner. A partir de cette image on extrait des informations visuelles qui sont utilisées comme "référence" à atteindre dans l'image. Après avoir amenée la caméra à une position différente, son mouvement est contrôlé, d'une manière ou d'une autre, afin que les informations visuelles observées coïncident avec les informations visuelles de "référence". Ceci étant accompli, la caméra se trouve, par rapport à l'objet observé, dans la même position que lors de l'apprentissage. De manière générale, indépendamment de la méthode d'asservissement visuel employée, ceci est vrai si et seulement si les paramètres intrinsèques de la caméra sont, après convergence, les mêmes que lors de l'apprentissage de l'image de référence. Cette condition est cependant très contraignante. En effet, dans certaines applications il est intéressant de pouvoir changer les paramètres intrinsèques de la caméra pendant l'asservissement (ou utiliser pour l'asservissement une caméra différente de celle utilisée pendant l'apprentissage). Dans ce travail, nous proposons d'éliminer la contrainte sur les paramètres intrinsèques afin d'augmenter considérablement les domaines d'application de l'asservissement visuel. Plus précisément, on a montré qu'il est possible de définir un erreur, à partir des données extraites de l'image, qui ne dépend pas des paramètres intrinsèques de la caméra avec laquelle l'image a été prise. L'invariance est obtenue grâce a une transformation projective qui peut être calculée à partir des information extraites dans l'image. Afin de contrôler le mouvement de la caméra et d'asservir l'erreur dans l'espace invariant à zéro, on propose une technique d'asservissement visuel basée sur l'approche par fonction de tâche. L'analyse de la stabilité montre que la loi de commande est robuste aux erreurs de calibration du système. De nombreuses expériences ont été effectuées sur le robot Cartésien AFMA de l'IRISA afin de valider la loi de commande. Les résultats montrent qu'il est possible de repositionner une caméra par rapport à un objet même si l'image de référence a été apprise avec une caméra différente et/ou les paramètres intrinsèques de la caméra varient pendant l'asservissement..

Mots-clés : asservissement visuel, invariance projective, caméra non-calibrée, planification de trajectoire

Contents

1 Introduction **4**

2 Theoretical background **5**
2.1 Notations 5
2.2 Camera Model 6
2.3 Invariance to camera parameters 7

3 Path planning in the invariant space **8**

4 Control in the invariant space **9**
4.1 Control of five d.o.f. of the camera 9
4.2 Control of the last d.o.f. of the camera 13

5 Experimental Results **14**
5.1 Stationary zooming camera 14
5.2 Visual Servoing without path planning 16
5.3 Path planning in the invariant space 22

6 Conclusion **27**

1 Introduction

The aim of visual servoing is to control the movement of a robot using the information provided by vision sensors. A typical task is to reposition an eye-in-hand system with respect to an observed object. Vision-based control has been widely investigated in the last few years [7, 8]. Despite the diversity of the approaches proposed to accomplish a positioning task, most visual servoing techniques are based on a "teaching-by-showing" approach. With this approach, the robot is moved to a goal position, the camera is shown the target view and a "reference image" of the object is stored (i.e. a set of features of the reference image). The position of the camera with respect to the object will be called the "reference position". After the camera and/or the object has been moved, several visual servoing schemes have been proposed in order to reposition the camera with respect to the object: image-based visual servoing [4], position-based visual servoing [15], hybrid visual servoing [11]. Each scheme has its own advantages and drawbacks [1] [2], but for all schemes if the visual features currently observed in the image coincide with the features extracted from the reference image, the camera is back to the reference position with respect to the object. Generally speaking, whatever is the visual servoing method used to achieve the task, that will be true if and only if the camera intrinsic parameters at the convergence are the same parameters of the camera used for learning. Indeed, if the camera intrinsic parameters change during the servoing (or the camera used during the servoing is different from the camera used to learn the reference image), even if the current image coincide with the reference image, the position of the camera with respect to the object will be completely different from the reference position. The aim of this paper is to increase the versatility of visual servoing by extending the teaching-by-showing technique to the case when different cameras are used for learning the reference image and for servoing. Indeed, intrinsic parameters may significantly vary during the life of the vision system and/or they can be changed intentionally when using zooming cameras. If so, with current visual servoing techniques the reference image must be shown again. In some applications, learning again the reference image could be very difficult. On the other hand, the visual servoing technique proposed in this paper allows us to learn the reference image once and for all. The basic idea is to use projective invariance so as to build an error function, from only measured image features, which is invariant on the intrinsic parameters of the camera. Projective invariance has been used in [6] to define set-points for stereo visual control that are independent on the viewing location. The system proposed in [6] is made up of two cameras (not mounted on the robot) observing both the object and the robot end-effector. Thus, it is possible to move a point on the manipulator to a point on the object independently on the cameras used to realize the task. In this paper, the proposed approach is completely different since the system is made up of only one camera mounted on a robot manipulator. Projective invariance is used to define an error which is dependent on the viewing location but it is independent on the camera intrinsic parameters. The projective transformation proposed in [10] is used to define a projective space only depending on the position of the camera with respect to an object. Thus, the camera can be repositioned with respect to a non-planar object given a reference image learned with a different camera and even if the camera parameters change during

the servoing [10]. The visual servoing scheme proposed in the paper has been completely validate by experiment. Experiments have shown that the teaching-by-showing technique can be extended to the case when different uncalibrated cameras are used for learning the reference image and for servoing [9]. Even if the measured task function does not depend on the camera intrinsic parameters, they are needed to estimate the Jacobian matrix which links the camera velocity to the displacements of the features in the invariant space [9]. Thus, calibration errors can affect the stability of the control law. For this reason, it is important to study the stability of the proposed method and formally prove its robustness with respect to camera calibration errors. In particular, the necessary and sufficient conditions for the local asymptotic stability in the presence of large calibration errors are obtained. Local stability implies that the system can accurately follow a path in the invariant space. Simple sufficient conditions are given in order to keep the tracking error bounded. The path can be chosen such that the robot follows a straight line even if the camera parameters are unknowns and the camera is not directly controlled in the Cartesian space. Contrarily to [12], path planning is not used here to solve the problem of the target visibility during the servoing. Indeed, the scheme proposed in this paper can be used with a zooming camera and the visibility problem can be easily solved by decreasing the focal length in order to obtain a larger field of view.

2 Theoretical background

2.1 Notations

Let the point $\mathcal{C}^*$ of the 3D Cartesian space, be the center of projection (see Figure 1). Suppose that $\mathcal{C}^*$ coincide with the origin $\mathcal{O}^*$ of the absolute frame $\mathcal{F}^*$. Let the plane of projection be parallel to the plane $(\overrightarrow{x}, \overrightarrow{y})$. The 3D point, with homogeneous coordinates $\boldsymbol{\mathcal{X}} = (X, Y, Z, 1)$ is projected in $\mathcal{F}^*$ to the point $\mathbf{m}^*$:

$$\zeta^* \mathbf{m}^* = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \end{bmatrix} \boldsymbol{\mathcal{X}} \tag{1}$$

where ζ^* is the positive depth. Let the point $\mathcal{C}$ be a different center of projection and the origin of frame $\mathcal{F}$. Let $\mathbf{t}$ and $\mathbf{R}$ be respectively the translation and the rotation between $\mathcal{F}^*$ and $\mathcal{F}$. Let $\mathbf{r} = \theta \mathbf{u}$ be the (3×1) vector containing the axis of rotation $\mathbf{u}$ and the angle of rotation θ $(0 \leq \theta < 2\pi)$. If $[\mathbf{r}]_\times$ is the skew symmetric matrix associated to vector $\mathbf{r}$, then $\mathbf{R} = \exp([\mathbf{r}]_\times)$. Let $\boldsymbol{\xi} = (\mathbf{t}, \mathbf{r})$ be the (6×1) vector containing global coordinates of an open subset $\mathcal{S} \subset \mathbb{R}^3 \times SO(3)$. Then, the reference position is $\boldsymbol{\xi}^* = 0$. The 3D point projects in $\mathcal{F}$ to the point $\mathbf{m}$:

$$\zeta(\boldsymbol{\xi})\, \mathbf{m}(\boldsymbol{\xi}) = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \boldsymbol{\mathcal{X}} \tag{2}$$

where $\zeta(\boldsymbol{\xi})$ is the positive depth.

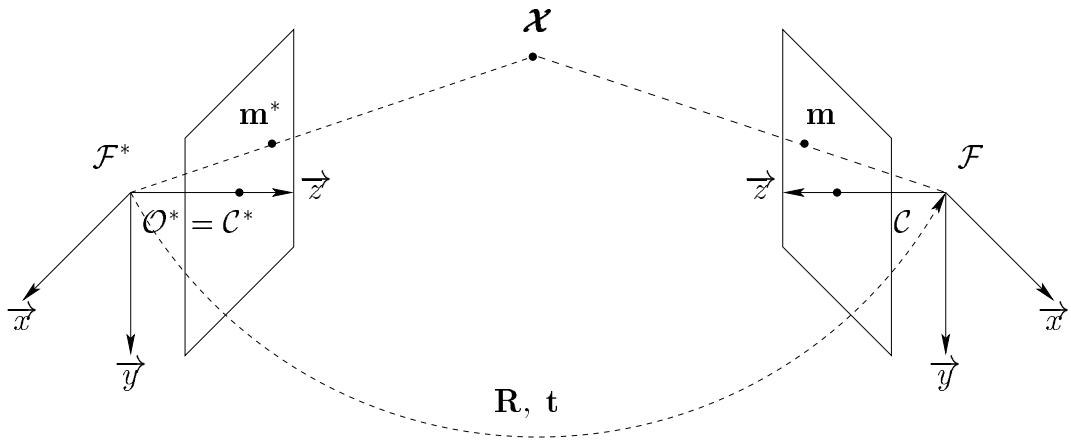


Figure 1: Perspective projections $\mathbf{m}^*$ and $\mathbf{m}$ of the same 3D point $\boldsymbol{\mathcal{X}}$ from two different points of view $\mathcal{C}^*$ and $\mathcal{C}$. The displacement of current $\mathcal{F}$, with respect to the reference frame $\mathcal{F}^*$, is represented by the translation $\mathbf{t}$ and the rotation $\mathbf{R}$.

2.2 Camera Model

Pinhole cameras perform a perspective projection of a 3D point. However, vectors $\mathbf{m}$ and $\mathbf{m}^*$ are not directly measured by the cameras. Indeed, the point $\mathbf{p} = (u, v, 1)$ observed in the image $\mathcal{I}$, taken at the position $\mathcal{F}$, depends on the camera internal parameters:

$$\mathbf{p}(\boldsymbol{\xi}, \mathbf{K}) = \mathbf{K}\mathbf{m}(\boldsymbol{\xi}) \qquad \mathbf{p} \in \mathcal{I}(\boldsymbol{\xi}, \mathbf{K}) \tag{3}$$

where:

$$\mathbf{K} = \begin{bmatrix} fk_u & -fk_u \cot(\theta) & u_0 \\ 0 & fk_v/\sin(\theta) & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

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 $\vec{u}$ and $\vec{v}$ direction (in pixels/meters), and θ is the angle between these axes. On the other hand, the point $\mathbf{p}^*$ observed in the image $\mathcal{I}$, taken at the reference position $\mathcal{F}^*$, depends on maybe different camera parameters:

$$\mathbf{p}^*(\boldsymbol{\xi}^*, \mathbf{K}^*) = \mathbf{K}^*\mathbf{m}^*(\boldsymbol{\xi}^*) \qquad \mathbf{p}^* \in \mathcal{I}(\boldsymbol{\xi}^*, \mathbf{K}^*) \tag{5}$$

where $\mathbf{K}^*$ has the same form of $\mathbf{K}$ given in equation (4). The image $\mathcal{I}(\boldsymbol{\xi}^*, \mathbf{K}^*)$ will be called the reference image since it is taken at the reference position $\mathcal{F}^*$. The objective of vision-based control is to drive a camera, mounted on the end-effector of a robot, to the reference position using the information provided by the image $\mathcal{I}(\boldsymbol{\xi}, \mathbf{K})$ currently observed. It is clear that both images depend on the intrinsic parameters of the cameras. Thus, even if $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ the reference and the current image will be different if $\mathbf{K} \neq \mathbf{K}^*$.

2.3 Invariance to camera parameters

As already mentioned in the introduction, most visual servoing techniques are generally based on the hypothesis that the camera frame $\mathcal{F}$ will coincide to the reference frame $\mathcal{F}^*$ if $\mathbf{p}_k = \mathbf{p}_k^*$, $\forall\ k \in \{1, 2, ..., n\}$ (supposing that a sufficient number of corresponding points are observed in the images). However, this hypothesis is valid *if and only if* $\mathbf{K} = \mathbf{K}^*$ at the convergence. In order to control the robot regardless to the camera used during the visual servoing it is necessary to build an error function which is independent on the camera intrinsic parameters. That is possible by using the simple projective transformation proposed in [10]. Suppose that n non coplanar points $\boldsymbol{\mathcal{X}}_k$ ($k \in \{1, 2, ..., n\}$) are visible on the observed object and consider three non-collinear 3D points $\boldsymbol{\mathcal{X}}_1$, $\boldsymbol{\mathcal{X}}_2$ and $\boldsymbol{\mathcal{X}}_3$. These three points project to the points $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$ in the current frame and to the points $\mathbf{m}_1^*, \mathbf{m}_2^*, \mathbf{m}_3^*$ in the reference frame. The corresponding image points in pixel coordinates $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ and $\mathbf{p}_1^*, \mathbf{p}_2^*, \mathbf{p}_3^*$ are obtained using equation (3) and equation (5) as follows:

$$\mathbf{Q}(\boldsymbol{\xi}, \mathbf{K}) = \mathbf{K}\mathbf{M}(\boldsymbol{\xi}) \tag{6}$$

$$\mathbf{Q}^*(\boldsymbol{\xi}^*, \mathbf{K}^*) = \mathbf{K}^*\mathbf{M}^*(\boldsymbol{\xi}^*) \tag{7}$$

where $\mathbf{Q} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}$, $\mathbf{M} = \begin{bmatrix} \mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 \end{bmatrix}$, $\mathbf{Q}^* = \begin{bmatrix} \mathbf{p}_1^* & \mathbf{p}_2^* & \mathbf{p}_3^* \end{bmatrix}$ and $\mathbf{M}^* = \begin{bmatrix} \mathbf{m}_1^* & \mathbf{m}_2^* & \mathbf{m}_3^* \end{bmatrix}$. The matrices $\mathbf{Q}$ and $\mathbf{Q}^*$ are non-singular (3×3) matrices and thus can be used to define two projective spaces $\mathcal{Q}$ and $\mathcal{Q}^*$ in both the current and reference images. The transformed points $\mathbf{q}$ and $\mathbf{q}^*$ are:

$$\begin{aligned} \mathbf{q}(\boldsymbol{\xi}) &= \mathbf{Q}^{-1}\mathbf{p} = \mathbf{M}^{-1}\mathbf{K}^{-1}\mathbf{K}\mathbf{m} = \mathbf{M}^{-1}\mathbf{m} \\ \mathbf{q}^*(\boldsymbol{\xi}^*) &= \mathbf{Q}^{*-1}\mathbf{p}^* = \mathbf{M}^{*-1}\mathbf{K}^{*-1}\mathbf{K}^*\mathbf{m}^* = \mathbf{M}^{*-1}\mathbf{m}^* \end{aligned}$$

Both $\mathbf{q}$ and $\mathbf{q}^*$ do not depend on the internal parameter of the cameras but they only depend on the position of the camera with respect to the observed object and on its three-dimensional structure. From now on we will refer to the transformed points $\mathbf{q} \in \mathcal{Q}(\boldsymbol{\xi})$ and $\mathbf{q}^* \in \mathcal{Q}(\boldsymbol{\xi}^*)$ as two "invariant" points where the invariance is related to camera internal parameters. However, $\mathbf{q}$ and $\mathbf{q}^*$ are not only invariant to camera intrinsic parameters. The invariance holds up to a collineation of the form:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Indeed, the matrix $\mathbf{Q}' = \mathbf{G}\mathbf{Q}$ is also a possible change of coordinates in the projective space. The matrix $\mathbf{G}$ must have the form given in equation (8) since, by definition the last row of $\mathbf{Q}'$ must be $(1, 1, 1)$ ($\mathbf{Q}'$ can be build from three image points). Thus, if $\mathbf{p}' = \mathbf{G}\mathbf{p}$ then:

$$\mathbf{q}' = \mathbf{Q}'^{-1}\mathbf{p}' = \mathbf{Q}^{-1}\mathbf{G}^{-1}\mathbf{G}\mathbf{p} = \mathbf{q}$$

As a consequence, it will not be possible using only the error $(\mathbf{q} - \mathbf{q}^*)$ to position the camera with respect to a planar target. Indeed, starting from a general collineation between the

two images, we will obtain $\mathbf{q} = \mathbf{q}^*$ up to $\mathbf{G}$. If the object is non planar, the only possible collineation is the rotation around the $\overrightarrow{z}$ axis:

$$\mathbf{G} = \mathbf{R}_z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, the transformed coordinates $\mathbf{q}$ and $\mathbf{q}^*$ are invariant to a rotation r_z around the $\overrightarrow{z}$ axis. Thus, if $\mathbf{q} = \mathbf{q}^*$ $\forall \mathbf{q}, \mathbf{q}^*$ then $t_x = t_y = t_z = r_x = r_y = 0$ but r_z can be any. It turns out that it is not a drawback at all. Indeed, the rotation r_z can be controlled separately, as explained in Section 4. Thus, there are no problems coming from the control of r_z when the displacement is large, as for image-based visual servoing [1, 3].

3 Path planning in the invariant space

The epipolar geometry in the invariant space is equivalent to a plane + parallax factorization [10]. The three points chosen for the projective transformation define a virtual plane attached to the object. Let $\mathbf{n}^*$ be the normal to the plane in the absolute frame $\mathcal{F}^*$ and let d^* be the distance of $\mathcal{C}^*$ from the plane. Let $\mu_k^* = \frac{1}{\zeta_k^*} - \frac{1}{d^*}\mathbf{n}^{*T}\mathbf{m}_k^*$ be a constant scalar such that if $\mu_k^* = 0$ then the 3D point $\boldsymbol{\mathcal{X}}_k$, projecting to $\mathbf{q}_k$ and $\mathbf{q}_k^*$, lies on the plane defined by the points $\boldsymbol{\mathcal{X}}_1$, $\boldsymbol{\mathcal{X}}_2$ $\boldsymbol{\mathcal{X}}_3$. The fundamental equation linking $\mathbf{q}_k$ and $\mathbf{q}_k^*$ is [10]:

$$\gamma_k \, \mathbf{q}_k = \mathbf{\Gamma}(\mathbf{q}_k^* + \mu_k^* \mathbf{e}^*) \tag{9}$$

where $\gamma_k = \frac{\zeta_k}{\zeta_k^*}$, $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1, \gamma_2, \gamma_3) = \mathrm{diag}(\frac{\zeta_1}{\zeta_1^*}, \frac{\zeta_2}{\zeta_2^*}, \frac{\zeta_3}{\zeta_3^*})$ and $\mathbf{e}^*$ is the epipole in the reference invariant space. From the current and reference image, one can measure $\mathbf{\Gamma}$, $\mathbf{e}^*$, γ_k and μ_k^* ($k \in \{4, 5, ..., n\}$) up to a scale factor. In order to eliminate the scale factor, the equation (9) can be divided by the last entry of the diagonal matrix $\mathbf{\Gamma}$ ($\gamma_3 > 0$) without loss of generality. After estimating all the parameters, one can build a function $\mathbf{q}_k^*(t)$ ($t \in [0; T]$) such that $\mathbf{q}_k^*(0) = \mathbf{q}_k$, $\mathbf{q}_k^*(T) = \mathbf{q}_k^*$ and such that the camera approximatively follows a straight line in the Cartesian space (even if the camera internal parameters are completely unknown). Indeed, if the camera follows the line $(\mathcal{C}, \mathcal{C}^*)$, then the epipole in the reference image is the same at each iteration. The epipole in the reference image is thus set to $\mathbf{e}^*/\|\mathbf{e}^*\|$ and determine the direction of translation (i.e. 2 d.o.f.). The norm of the translation can be fixed by a function $\beta_k(t)$ such that $\beta_k(0) = \frac{\mu_k^*}{\gamma_3}\|\mathbf{e}^*\|$ ($\beta_k(0)$ can be measured from equation (9)) and $\beta_k(T) = 0$:

$$\beta_k(t) = \beta_k(0) + \frac{t}{T}(\beta_k(T) - \beta_k(0)) = \frac{\mu_k^*}{\gamma_3}\|\mathbf{e}^*\| \left(1 - \frac{t}{T}\right)$$

The rotational d.o.f. r_x and r_y can be fixed with a diagonal matrix $\Delta(t)$ such that $\Delta(0) = \text{diag}(\frac{\gamma_1}{\gamma_3}, \frac{\gamma_2}{\gamma_3}, 1)$ and $\Delta(T) = \text{diag}(1, 1, 1)$:

$$\begin{aligned} \Delta(t) &= \Delta(0) + \frac{t}{T}(\Delta(T) - \Delta(0)) \\ &= \begin{bmatrix} \frac{\gamma_1}{\gamma_3} + \frac{t}{T}(1 - \frac{\gamma_1}{\gamma_3}) & 0 & 0 \\ 0 & \frac{\gamma_1}{\gamma_3} + \frac{t}{T}(1 - \frac{\gamma_1}{\gamma_3}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Since by definition $[1\ 1\ 1]\,\mathbf{q}_k(t) = 1, \forall t$, the functions $\mathbf{q}_k^*(t)$ ($k \in \{4, 5, ..., n\}$) can be obtained as follows:

$$\mathbf{q}_k^*(t) = \frac{\Delta(t)(\mathbf{q}_k^*(T) + \beta_k(t)\mathbf{e}^*/\|\mathbf{e}^*\|)}{[1\ 1\ 1]\,\Delta(t)(\mathbf{q}^*(T) + \beta(t)\mathbf{e}^*/\|\mathbf{e}^*\|)} \tag{10}$$

It is easy to verify that $\mathbf{q}_k^*(0) = \mathbf{q}_k(0)$ and $\mathbf{q}_k^*(T) = \mathbf{q}_k^*$. Note again that if the current invariant points are such that $\mathbf{q}_k(t) = \mathbf{q}_k^*(t)\ \forall t$ then the epipole in the reference image is constant and the camera exactly follows a straight line in the 3D space. The rotational d.o.f. r_z has not been considered here since it does not have any influence on the path in the invariant space.

4 Control in the invariant space

The control the camera in the invariant space is divided into two different parts since the points $\mathbf{q}_k$ are invariant on the rotation around the $\vec{z}$ axis and they can only be used to control five d.o.f. of the camera. Another information must be extracted from the images in order to control the last camera d.o.f.

4.1 Control of five d.o.f. of the camera

Suppose that n matched points are available in both images. Since three points are used to define the projective transformations, only the remaining $n-3$ points can be used to control five d.o.f. of the camera $(t_x, t_y, t_z, r_x, r_y)$. Let $\mathbf{s}(\boldsymbol{\xi}) = (\mathbf{q}_4, \mathbf{q}_5, ..., \mathbf{q}_n)$ be the $(3(n-3) \times 1)$ vector containing the current invariant points. The time derivative of the vector is:

$$\dot{\mathbf{s}} = \mathbf{J}(\boldsymbol{\xi})\ \boldsymbol{\mu} \tag{11}$$

where $\boldsymbol{\mu} = (\nu_x, \nu_y, \nu_z, \omega_x, \omega_y)$ and $\mathbf{J}(\boldsymbol{\xi})$ is the Jacobian matrix which can be built from the Jacobian matrix $\mathbf{J}_k(\boldsymbol{\xi})$ relative to each invariant point: $\dot{\mathbf{q}}_k = \mathbf{J}_k(\boldsymbol{\xi})\boldsymbol{\mu}$. Each point $\mathbf{q}_k$ gives a (3×5) interaction matrix $\mathbf{J}_k$ with only 2 independent rows since $q_{1k} + q_{2k} + q_{3k} = 1$ and thus $\dot{q}_{1k} + \dot{q}_{2k} + \dot{q}_{3k} = 0$. Consequently, in order to control five d.o.f. of the camera (except for r_z) we need at least $n = 6$ points. Differentiating $\mathbf{q}_k$, we have:

$$\dot{\mathbf{q}}_k = \frac{d\mathbf{Q}^{-1}}{dt}\mathbf{p}_k + \mathbf{Q}^{-1}\dot{\mathbf{p}}_k = \frac{d\mathbf{Q}^{-1}}{dt}\mathbf{Q}\mathbf{q}_k + \mathbf{Q}^{-1}\dot{\mathbf{p}}_k \tag{12}$$

Note that:

$$\frac{d\mathbf{Q}^{-1}}{dt}\mathbf{Q} = -\mathbf{Q}^{-1}\frac{d\mathbf{Q}}{dt}$$

Plugging this equation in equation (12) we obtain:

$$\dot{\mathbf{q}}_k = \mathbf{P}^{-1}(\dot{\mathbf{p}}_k - \dot{\mathbf{P}}\mathbf{q}_k) = \mathbf{P}^{-1}(\dot{\mathbf{p}}_k - q_{1k}\dot{\mathbf{p}}_1 - q_{2k}\dot{\mathbf{p}}_2 - q_{3k}\dot{\mathbf{p}}_3)$$

The derivative of the current image point is:

$$\dot{\mathbf{p}}_k = \mathbf{L}_k(\zeta_k, \mathbf{K}, \mathbf{p}_k)\begin{bmatrix} \boldsymbol{\mu} \\ \omega_z \end{bmatrix}$$

where $\mathbf{L}_k$ is obtained from the standard interaction matrix for an image point [4] adding a third row of zeros. Finally, we obtain:

$$\dot{\mathbf{q}}_k = \mathbf{P}^{-1}(\mathbf{L}_k - q_{1k}\mathbf{L}_1 - q_{2k}\mathbf{L}_2 - q_{3k}\mathbf{L}_3)\begin{bmatrix} \boldsymbol{\mu} \\ \omega_z \end{bmatrix}$$

Since $\mathbf{q}_k$ is invariant on r_z the last column of the Jacobian matrix is null and $\mathbf{J}_k$ is obtained from the first five columns of matrix $\mathbf{P}^{-1}(\mathbf{L}_k - q_{1k}\mathbf{L}_1 - q_{2k}\mathbf{L}_2 - q_{3k}\mathbf{L}_3)$. It must be emphasised that $\mathbf{J}_k$ can be written as follows:

$$\mathbf{J}_k = \begin{bmatrix} \boldsymbol{\Phi}_k(\zeta_1, \zeta_2, \zeta_3, \zeta_k) & \boldsymbol{\Psi}_k \end{bmatrix}\begin{bmatrix} \mathbf{K} & 0 \\ 0 & \mathbf{F}(\mathbf{K}) \end{bmatrix}$$

where:

$$\boldsymbol{\Phi}_k = \mathbf{Q}^{-1}\begin{bmatrix} \sum_{i=1}^{3}\frac{q_{ik}}{\zeta_i} - \frac{q_{ik}}{\zeta_k} & 0 & \frac{u_k}{\zeta_k} - \sum_{i=1}^{3}\frac{u_i q_{ik}}{\zeta_i} \\ 0 & \sum_{i=1}^{3}\frac{q_{ik}}{\zeta_i} - \frac{q_{ik}}{\zeta_k} & \frac{v_k}{\zeta_k} - \sum_{i=1}^{3}\frac{v_i q_{ik}}{\zeta_i} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{\Psi}_k = \mathbf{Q}^{-1}\begin{bmatrix} u_k v_k - \sum_{i=1}^{3} u_i v_i q_{ik} & \sum_{i=1}^{3} u_i^2 q_{ik} - u_k^2 \\ v_k^2 - \sum_{i=1}^{3} v_i^2 q_{ik} & \sum_{i=1}^{3} u_i v_i q_{ik} - u_k v_k \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{F}(\mathbf{K}) = \begin{bmatrix} \frac{\sin(\theta)}{fk_v} & \frac{-\cos(\theta)}{fk_v} \\ 0 & \frac{1}{fk_u} \end{bmatrix}$$

Therefore, the Jacobian can be decomposed as follows:

$$\mathbf{J}(\boldsymbol{\xi}, \boldsymbol{\zeta}, \mathbf{K}) = \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\zeta}) & \boldsymbol{\Psi} \end{bmatrix}\begin{bmatrix} \mathbf{K} & 0 \\ 0 & \mathbf{F}(\mathbf{K}) \end{bmatrix}$$

This matrix depends on the depths distribution $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, ..., \zeta_n)$ and on the camera intrinsic parameters $\mathbf{K}(t)$ which can eventually vary during the servoing. Since the structure of the object is rigid, one can estimate the depth distribution up to a scalar factor using the epipolar geometry $\widehat{\boldsymbol{\zeta}} = \kappa(t)\boldsymbol{\zeta}$ (where $\kappa(t) > 0\ \forall t$), as explained in Appendix A. On the other hand, the camera parameters can eventually vary during the servoing when using a zooming camera. The stability analysis will show that it is not necessary to use a self-calibration algorithm to estimate the camera parameters but only a rough approximation $\widehat{\mathbf{K}}$ can be used to compute the Jacobian matrix. Since $\boldsymbol{\Phi}(\widehat{\boldsymbol{\zeta}}) = \frac{1}{\kappa}\boldsymbol{\Phi}(\boldsymbol{\zeta})$, the estimated Jacobian matrix can be written as:

$$\widehat{\mathbf{J}} = \begin{bmatrix} \frac{1}{\kappa}\boldsymbol{\Phi}(\boldsymbol{\zeta}) & \boldsymbol{\Psi} \end{bmatrix} \mathbf{F}(\widehat{\mathbf{K}}) \tag{13}$$

Let us suppose that the Jacobian matrix $\mathbf{J}$ is full rank. If $\mathbf{J}$ is full rank, then also estimated matrix $\widehat{\mathbf{J}}$ is full rank. In order to control the five d.o.f. of the camera I use the task function approach [14]. Consider the following (5×1) task function:

$$\boldsymbol{\varepsilon}(\boldsymbol{\xi}) = \widehat{\mathbf{J}}^{+}(\boldsymbol{\xi})(\mathbf{s}(\boldsymbol{\xi}) - \mathbf{s}^*(t)) \tag{14}$$

where $\mathbf{s}^*(t) = (\mathbf{q}_4^*(t), \mathbf{q}_5^*(t), ..., \mathbf{q}_n^*(t))$ and $\mathbf{q}_k^*(t)$ is given by equation (10). Obviously, if $\mathbf{s} = \mathbf{s}^*$ then $\boldsymbol{\varepsilon} = 0$. On the other hand, sufficient conditions are given in Appendix B such that if $\|\mathbf{s} - \mathbf{s}^*\|$ is sufficiently small then $\boldsymbol{\varepsilon} = 0$ only if $\mathbf{s} = \mathbf{s}^*$ (i.e. $\mathbf{s} - \mathbf{s}^*$ never belongs to $\mathrm{Ker}(\widehat{\mathbf{J}}^{+})$). Differentiating equation (14) we obtain:

$$\dot{\boldsymbol{\varepsilon}} = \frac{d\widehat{\mathbf{J}}^{+}}{dt}(\mathbf{s} - \mathbf{s}^*(t)) + \widehat{\mathbf{J}}^{+}\dot{\mathbf{s}} - \widehat{\mathbf{J}}^{+}\frac{\partial \mathbf{s}^*(t)}{\partial t} \tag{15}$$

Using equation (11) and since the vector $\frac{d\widehat{\mathbf{J}}^{+}}{dt}(\mathbf{s} - \mathbf{s}^*(t))$ can be written as $\frac{d\widehat{\mathbf{J}}^{+}}{dt}(\mathbf{s} - \mathbf{s}^*(t)) = \mathbf{O}(\mathbf{s} - \mathbf{s}^*(t))\boldsymbol{\mu}$ (where $\mathbf{O}(\mathbf{s} - \mathbf{s}^*(t)) \to 0$ if $\mathbf{s} \to \mathbf{s}^*(t)$), we obtain:

$$\dot{\boldsymbol{\varepsilon}} = \left(\mathbf{O}(\mathbf{s} - \mathbf{s}^*(t)) + \widehat{\mathbf{J}}^{+}\mathbf{J}\right)\boldsymbol{\mu} - \widehat{\mathbf{J}}^{+}\frac{\partial \mathbf{s}^*(t)}{\partial t} \tag{16}$$

Consider the control law:

$$\boldsymbol{\mu} = -\lambda\boldsymbol{\varepsilon} + \widehat{\mathbf{J}}^{+}\frac{\partial \mathbf{s}^*(t)}{\partial t} \tag{17}$$

where λ is a positive gain. From equations (17) and (16) we obtain the following closed-loop equation:

$$\begin{aligned} \dot{\boldsymbol{\varepsilon}} &= -\lambda\left(\mathbf{O}(\mathbf{s} - \mathbf{s}^*(t)) + \widehat{\mathbf{J}}^{+}\mathbf{J}\right)\boldsymbol{\varepsilon} \\ &+ \left(\mathbf{O}(\mathbf{s} - \mathbf{s}^*(t)) + \widehat{\mathbf{J}}^{+}\mathbf{J} - \mathbf{I}\right)\widehat{\mathbf{J}}^{+}\frac{\partial \mathbf{s}^*(t)}{\partial t} \end{aligned} \tag{18}$$

In order to study the behavior of the task function during the path tacking, consider the system (18) linearised for $\mathbf{s} = \mathbf{s}^*(t)$:

$$\dot{\boldsymbol{\varepsilon}} = -\lambda\ \mathbf{A}(t)\ \boldsymbol{\varepsilon} + \mathbf{b}(t) \tag{19}$$

where:

$$\begin{aligned}
\mathbf{A}(t) &= \widehat{\mathbf{J}}^{+}\mathbf{J}\Big|_{\mathbf{s}=\mathbf{s}^*(t)} = \begin{bmatrix} \kappa(t)\widehat{\mathbf{K}}^{-1}\mathbf{K}(t) & 0 \\ 0 & \widehat{\mathbf{G}}^{-1}\mathbf{F}(t) \end{bmatrix} \\
\mathbf{b}(t) &= \left(\widehat{\mathbf{J}}^{+}\mathbf{J} - \mathbf{I}\right)\widehat{\mathbf{J}}^{+}\Big|_{\mathbf{s}=\mathbf{s}^*(t)} \frac{\partial \mathbf{s}^*(t)}{\partial t}
\end{aligned}$$

Let us consider first the case when the starting point $\boldsymbol{\varepsilon}(0)$ is in a neighbourhood of $\boldsymbol{\varepsilon} = 0$ and suppose that $\partial \mathbf{s}^*(t)/\partial t = 0$ and that the camera parameters do not vary during the servoing $\partial \mathbf{K}(t)/\partial t = 0$.

Proposition 1 *The equilibrium point $\boldsymbol{\varepsilon} = 0$, of the time varying linear system (19) when $\mathbf{b}(t) = 0$ and $\partial \mathbf{K}(t)/\partial t = 0$, is locally asymptotically stable if and only if $\kappa(t) > 0$, $\widehat{f} > 0$, $\widehat{k}_u > 0$, $\widehat{k}_v > 0$ and $0 < \widehat{\theta} < \pi$.*

The proof is detailed in the Appendix B. This proposition means that the control law is extremely robust to calibration errors since any positive approximation of the camera parameters is sufficient to stabilise the system. In the case when the camera parameters vary during the servoing, we have:

Proposition 2 The equilibrium point $\boldsymbol{\varepsilon} = 0$, of the time varying linear system (19) when $\mathbf{b}(t) = 0$, is locally asymptotically stable if $\kappa(t) > 0$, $\widehat{\mathbf{K}}^{-1}\mathbf{K}(t) > 0$ and $\widehat{\mathbf{F}}^{-1}\mathbf{F}(t) > 0$.

The proof is given in the Appendix B. The sufficient conditions are the same found in [11] and they are very large. This means that one can choose once and for all an estimate $\widehat{\mathbf{K}}$ of the camera parameters and keep it constant even if the real parameters $\mathbf{K}(t)$ are changing during the servoing. Consider now the general case when a path computed from equation (10) is tracked (i.e. $\mathbf{b}(t) \neq 0$) and the camera parameters are eventually varying during the servoing.

Proposition 3 If the system is locally stable (see Proposition 2) and $\mathbf{b}(t)$ is bounded $\|\mathbf{b}(t)\| \leq \psi/2$, during the path tracking (with $\boldsymbol{\varepsilon}(0) = 0$ since $\mathbf{s}(0) = \mathbf{s}^*(0)$) the norm of the task function can be bounded by:

$$\|\boldsymbol{\varepsilon}(t)\| \leq \frac{\psi}{\varphi} \tag{20}$$

where $\varphi = \lambda\sigma$ (σ is the minimum singular value of the symmetric positive matrix $\mathbf{A}(t) + \mathbf{A}^T(t)$). If $\|\boldsymbol{\varepsilon}(t)\|$ is bounded, then $\|\mathbf{s}(t) - \mathbf{s}^*(t)\|$ is also bounded.

The proof is presented in the Appendix B. In order to reduce the tracking error, one can increase the gain λ (i.e. φ increases) or decrease the velocity $\frac{\partial \mathbf{s}^*(t)}{\partial t}$ (i.e. ψ decreases). Obviously, the bound on $\|\boldsymbol{\varepsilon}(t)\|$ depends also on the calibration errors. Indeed, if all the parameters of th system are perfectly known, then $\widehat{\mathbf{J}} = \mathbf{J}$ and $\psi = 0$. Thus the tracking is perfect since $\mathbf{b}(t) = 0$. On the other hand, even if the bound on the tracking error increases when calibration errors increase, the misestimation of $\mathbf{b}(t)$ has a small influence on the stability of the servoing (see the experimental results).

4.2 Control of the last d.o.f. of the camera

As already mentioned, the remaining d.o.f. of the camera cannot be controlled using $\mathbf{q}$. Therefore, it is necessary to find a parameter depending on the rotation around the $\overrightarrow{z}$ axis. Let $\mathbf{T}$ be the following matrix:

$$\mathbf{T} = \mathbf{Q}\mathbf{Q}^{*-1} = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ 0 & 0 & 1 \end{bmatrix} \tag{21}$$

The matrix $\mathbf{T}$ must be triangular at the convergence (i.e. when the camera is back at the reference position $\mathcal{F} = \mathcal{F}^*$). Indeed, if $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ then $\mathbf{M} = \mathbf{M}^*$ and from equations (7) and (6) we obtain:

$$\mathbf{T} = \mathbf{Q}\mathbf{Q}^{*-1}|_{\boldsymbol{\xi}=\boldsymbol{\xi}^*} = \mathbf{K}(t)\mathbf{K}^{*-1} \tag{22}$$

At the convergence, matrix $\mathbf{T}$ must be upper triangular for any matrices $\mathbf{K}(t)$ and $\mathbf{K}^*$ (i.e. $\tau_{21} = 0$) and have positive eigenvalues (i.e. $\tau_{11} > 0$ and $\tau_{22} > 0$). Therefore, one must impose the constraints $\tau_{21} = 0$, $\tau_{11} > 0$ and $\tau_{22} > 0$ in order to have $r_z = 0$. Note that $\tau_{21} = 0$ if $r_z = \pm\frac{\pi}{2}$ or $r_z = \pm\pi$. However, in that cases $\tau_{11} < 0$ and/or $\tau_{22} < 0$. For simplicity, I consider now the case when $-\frac{\pi}{2} < r_z < \frac{\pi}{2}$ but few changes have to be made in order to consider the general case. The remaining d.o.f. is thus controlled by a second scalar task function:

$$\epsilon = \det(\mathbf{Q}^*)\ \tau_{21} = v_1^*(v_3 - v_2) + v_2^*(v_1 - v_3) + v_3^*(v_2 - v_1)$$

Since $\det(\mathbf{Q}^*) \neq 0$, if $\epsilon = 0$ then $\tau_{21} = 0$ and at the convergence $\mathcal{F} = \mathcal{F}^*$ (since $\boldsymbol{\varepsilon} \to 0$ thanks to the previous control law). The derivative of ϵ can be written:

$$\dot{\epsilon} = a(\epsilon, t)\omega_z + \mathbf{c}^T(t)\boldsymbol{\mu}(t) \tag{23}$$

where $a(\epsilon, t)$ is a scalar, $\mathbf{c}(t)$ is a (5×1) vector and $\boldsymbol{\mu}(t)$ is the control law (17). Note that if $-\frac{\pi}{2} < r_z < \frac{\pi}{2}$ then $a(\epsilon, t) \neq 0$. Imposing an exponential convergence of ϵ to zero the control law for r_z is:

$$\omega_z = -\frac{\eta\,\epsilon}{\widehat{a}(\epsilon, t)} - \frac{\widehat{\mathbf{c}}^T\boldsymbol{\mu}(t)}{\widehat{a}(\epsilon, t)} \tag{24}$$

where η is a positive scalar tuning the speed of convergence, $\widehat{a}$ and $\widehat{\mathbf{c}}$ are approximations of a and $\mathbf{c}$. The closed loop equation is the following:

$$\dot{\epsilon} = -\eta \frac{a(\epsilon,t)}{\widehat{a}(\epsilon,t)}\epsilon + \left(\mathbf{c}^T(t) - \frac{a(\epsilon,t)}{\widehat{a}(\epsilon,t)}\widehat{\mathbf{c}}^T(t)\right)\boldsymbol{\mu}(t) \tag{25}$$

Note that $\boldsymbol{\mu}(t) \to 0$ and $\mathbf{c}(t)$ is bounded. Thus, the control law is stable if and only if $\frac{a(\epsilon,t)}{\widehat{a}(\epsilon,t)} > 0$. In the ideal case when $\widehat{a} = a$ and $\widehat{\mathbf{c}} = \mathbf{c}$ the convergence is exponential. As it is shown by the experiments, the control law is stable even in the presence of calibration errors.

5 Experimental Results

The visual servoing scheme proposed in the paper has been tested on a six d.o.f. Cartesian robot AFMA (at IRISA/INRIA Rennes). The robot is very well calibrated and it provides a ground truth in order to measure the positioning precision of visual servoing.

5.1 Stationary zooming camera

In order to position a robot with respect to an object regardless to the camera used during the visual servoing it is necessary to build an error function which is independent on the camera intrinsic parameters. To test whether the theory presented in the paper is a reasonable approximation it is important to prove the invariance of the new projective space $\mathcal{Q}$ to changes in camera intrinsic parameters. In this experiment, the focal length of a stationary camera changes six times approximatively from 2250 pixels to 1550 pixels (the corresponding six images are given in Figure 2). For each image, 16 points (corresponding to the 4 corners of the border of the 4 calibration grids) are extracted. The three points defining the reference plane are chosen spread in the image (points number 1, 8 and 11). The first image at the top to the left in Figure 2 is chosen as the reference image. Figure 3(a) shows that the error $\|\mathbf{p}_k - \mathbf{p}_k^*\|$ $(k = 1, 2, ..., 16)$ in the image space increases while the camera is zooming. On the other hand, Figure 3(b) shows that the error $\|\mathbf{q}_k - \mathbf{q}_k^*\|$ $(k = 1, 2, ..., 16)$ in the invariant space is not only close to zero but also practically constant. Obviously, the error is not exactly null because of noise in features extraction.

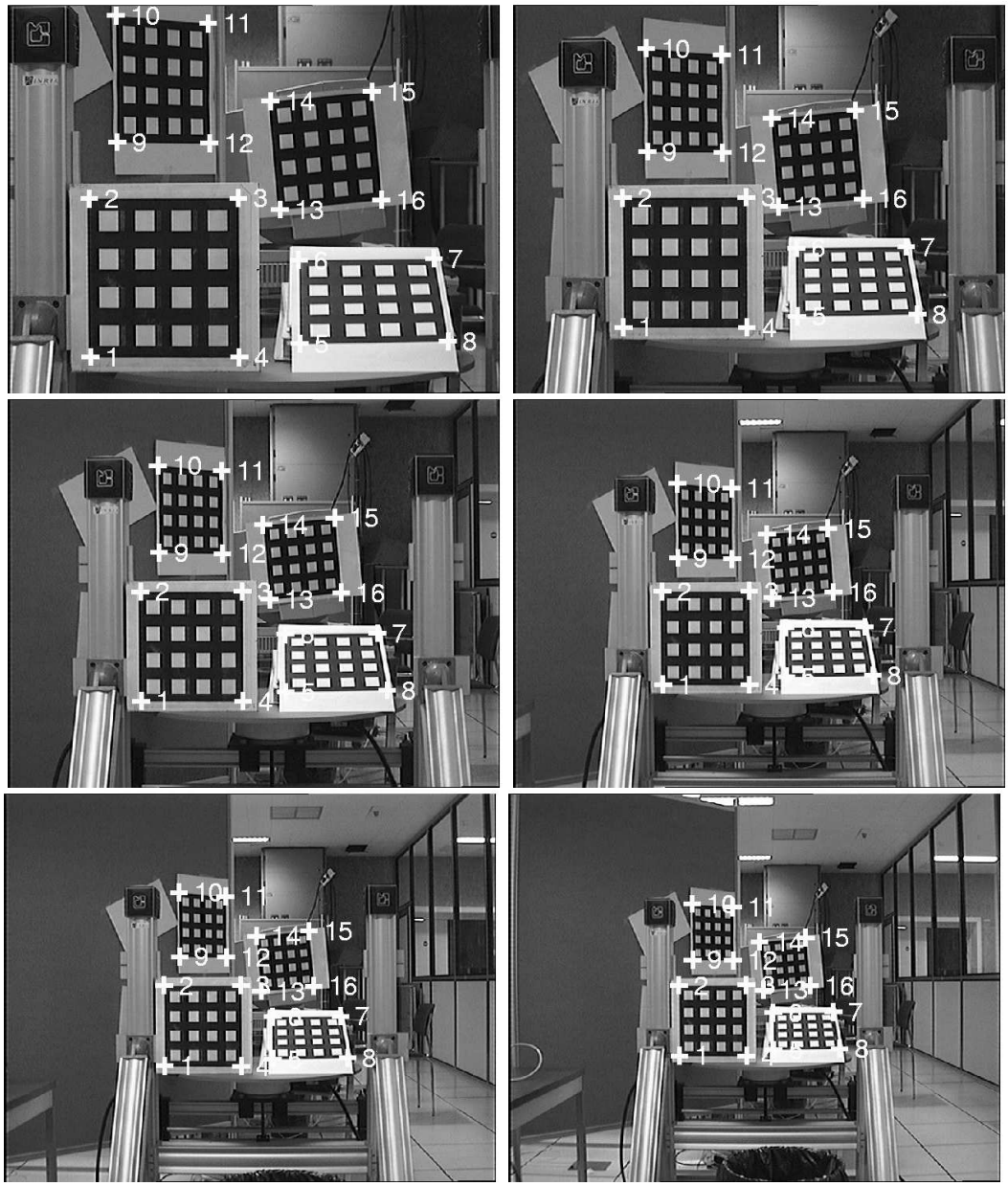
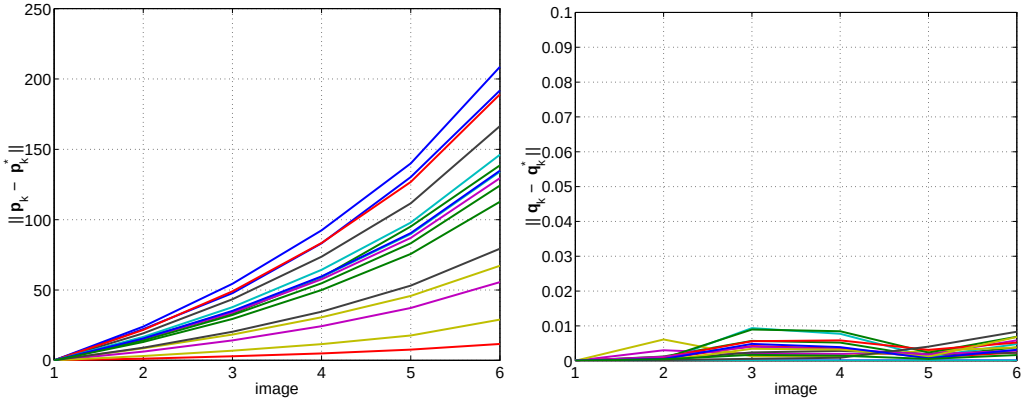


Figure 2: Six images taken with a stationary camera which is zooming



(a) Error $\|\mathbf{p}_k - \mathbf{p}_k^*\|$ (b) Error $\|\mathbf{q}_k - \mathbf{q}_k^*\|$

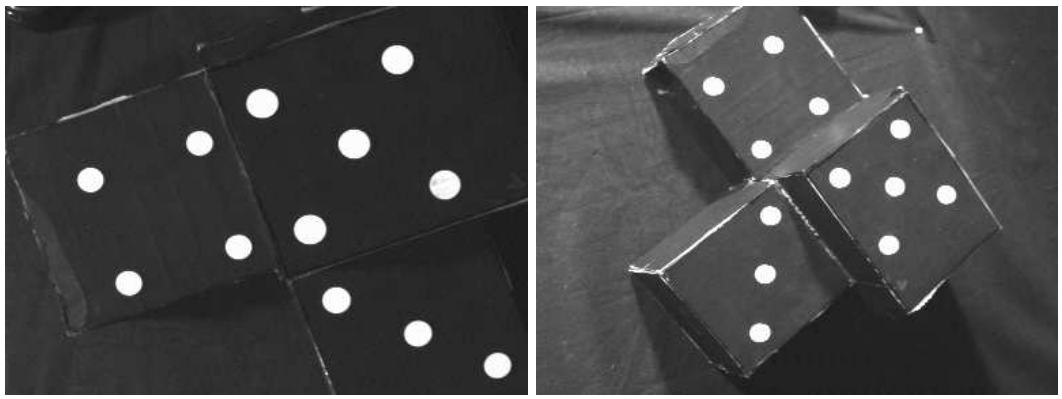
Figure 3: Since the camera is stationary, the error in the invariant space (b) is close to zero and practically constant (despite the camera is zooming) while the error in the image space (a) obviously increases.

5.2 Visual Servoing without path planning

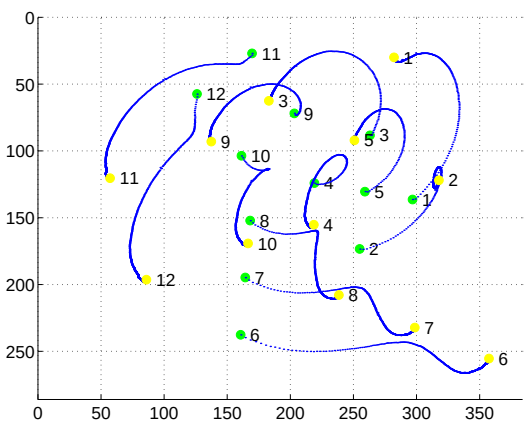
In these experiments, I show that the proposed control scheme can work even if the error $\mathbf{q}-\mathbf{q}^*$ is large and it is not constrained to follow a path in the invariant space. The reference is fixed to $\mathbf{q}_k^*(t) = \mathbf{q}_k^*\ \forall t, \forall k$, thus $\partial \mathbf{s}^*(t)/\partial t = 0$. For both experiments, a camera (with a 12mm lens) is initially mounted on the robot end-effector and observe an "object" composed by 12 points. An approximation $\widehat{\mathbf{K}}$ of the current camera parameters is used in the control law. In order to show that the visual servoing technique is robust with respect to large errors on the estimation of the depth of the observed points, the current depth $\boldsymbol{\zeta}$ is fixed to the depth $\boldsymbol{\zeta}^*$ measured in the reference frame (this corresponds to the image-based visual servoing setup proposed in [4]).

In the first experiment (Figure 4), the 12mm lens used for learning the reference image (see Figure 4(a)) is the same lens used for servoing. After the reference image has been learned the robot is displaced to its initial position (Figure 4(b) shows the initial image of the object). The initial displacement of the robot is quite large: 300 mm for the translation and 80 *degrees* for the rotation (see the initial values of the curves in Figures 5(i) and (j)). The control laws computed from equations (17) and (24) is plotted in Figures 5(g) and (h). The control laws are stable despite the large intial displacement. Thus, the task function converges to zero (see Figures 5(e) and (f)) as well as the error in the invariant space (see Figures 4(d)). Consequently, the position error of the robot end-effector converges to zero (see Figures 5(i) and (j)). At the convergence, the positioning precision is less than 1 mm for the translation and 0.1 *degrees* for the rotation. The servoing is stopped when the error for each point is smaller than 0.5 pixels. Figure 4(c) plots the trajectory of the points in the

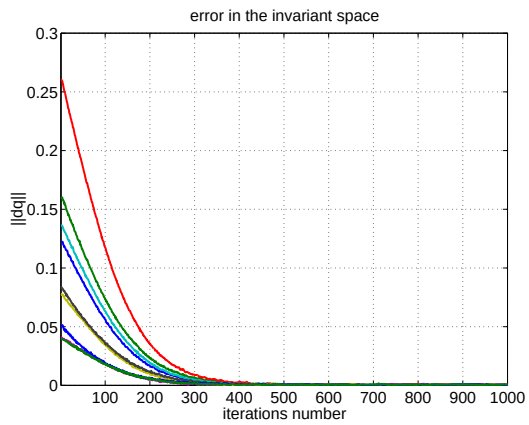
image (the blue curves) and shows that the points go from their initial position (the green points) to their reference position (the yellow points). The first experiment proves that the new scheme can perform positioning tasks exactly as standard visual servoing techniques. Since the robot is not controlled in the image space, some of the tracked features could get out of the image. Whith the proposed visual servoing scheme, this problem can be simply avoided by changing the focal length (zooming out in order to have a larger field of view).



(a) Reference image (b) Initial image

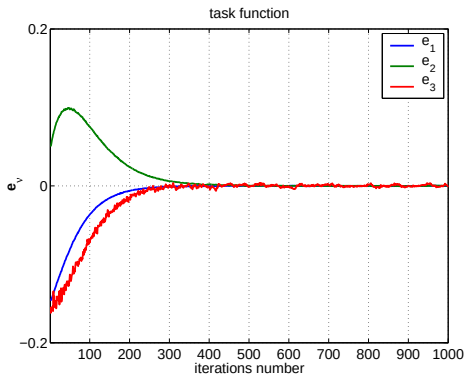


(c) Points trajectory

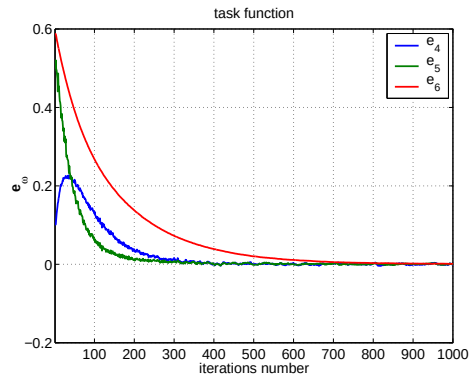


(d) Invariant space error

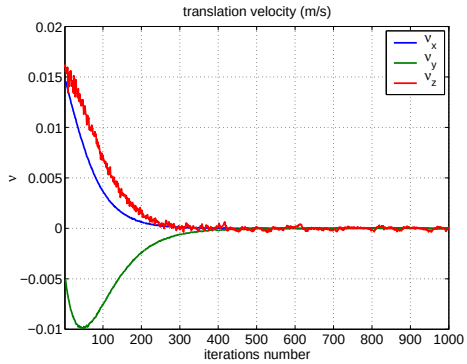
Figure 4: Experiment using the same 12mm lens for learning and servoing. The camera-independent vision-based control works as well as standard visual servoing.



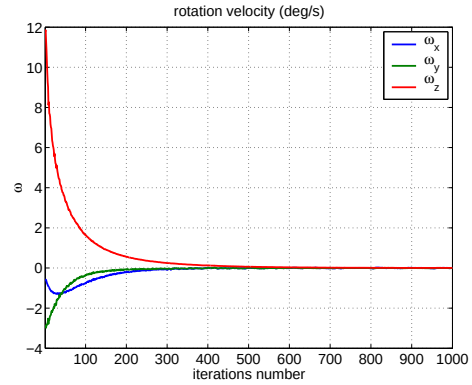
(a) Task function: $\mathbf{e}_\nu$



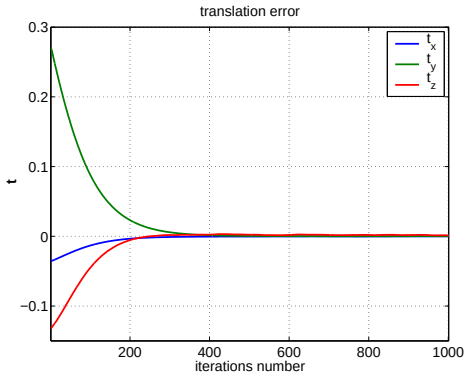
(b) Task function: $\mathbf{e}_\omega$



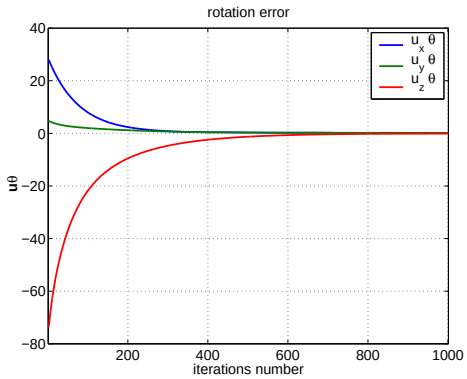
(c) Translation speed ($\frac{m}{s}$)



(d) Rotation speed ($\frac{deg}{s}$)



(e) Translation error (m)



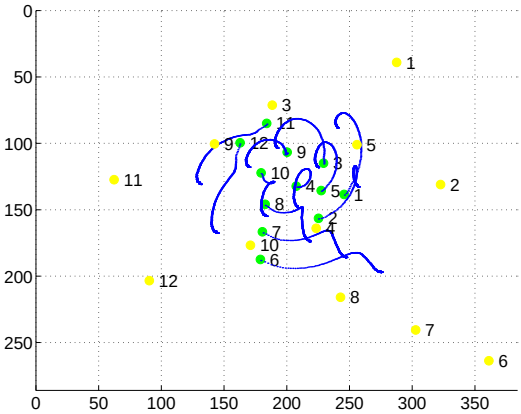
(f) Rotation error (deg)

Figure 5: Experiment using the same 12mm lens for learning and servoing. The camera-independent vision-based control works as well as standard visual servoing.

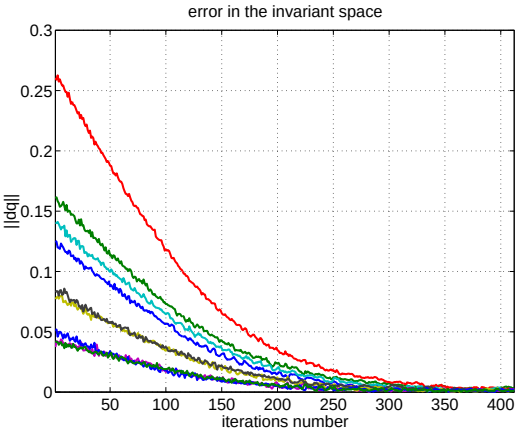
The proof that the camera-independent vision-based control increases the versatility of visual servoing is provided by the second experiment (Figure 6). In this second experiment, the reference image is identical to the previous one (see Figure 4(a) and Figure 6(a)). The robot starts from the same initial position (the initial displacement of the robot is the same in Figures 5(i) and (j) and Figures 7(i) and (j)) but now the lens of the camera is changed. A 6mm lens is used for servoing, thus the initial image in Figure 6(b) is completely different from the image in Figure 5(b). On the other hand, the curves plotted in Figure 5(d) and Figure 7(d) have identical initial values since the error in the invariant space is independent on the camera intrinsic parameters and the initial and reference positions are the same in both experiments. Even if the task function has similar behaviours in Figures 5(e) and (f) and in Figures 7(e) and (f), the curves are slightly different since different approximations of the camera parameters were used to compute the Jacobian matrix. Figure 6(c) shows the trajectory of the points in the image. The final image is obviously different from the reference image (the yellow points). However, Figures 7(i) and (j) show that the error on the position of the end-effector converges to zero. With respect to the first experiment, the number of iterations is reduced by a factor of two. Indeed, since the error in the image is not meaningful, the servoing is now stopped when the error in the invariant space converges to zero. Figures 7(i) and (j) plot the control law sent to the robot. The control law is robust to noise and ensures again the convergence of the end-effector to the reference position. When compared to the previous experiment, the higher level of noise (especially on e_3, as is it shown by Figure 7(e)) is explained by the fact that the size of the object is now small in the image due to the change of the lens. Indeed, the extraction of the centroids of the white spots is less precise. Furthermore, the main failure mode for the method occurs when the observed object is planar. The points of the object used in the experiments are close to a planar configuration. Consequently, the precision obtained in this experiment is around 2 mm for the translation and 0.2 degrees for the rotation. The experiment was repeated several times with similar results. The error is mainly on the translation along the z axis since a change of the lens causes a small translation of the center of projection. Let us remark that switching the lens in the opposite manner (from a 6mm to a 12mm lens) slightly worsen the results because the reference image is at a lower resolution and some points are not visible in the final image. In that case, a zooming camera should be used in order to zoom out near the converge enlarging the camera field of view. Finally, similar results have been obtained changing the lens from 25mm to 12mm.

(a) Reference image

(b) Initial image

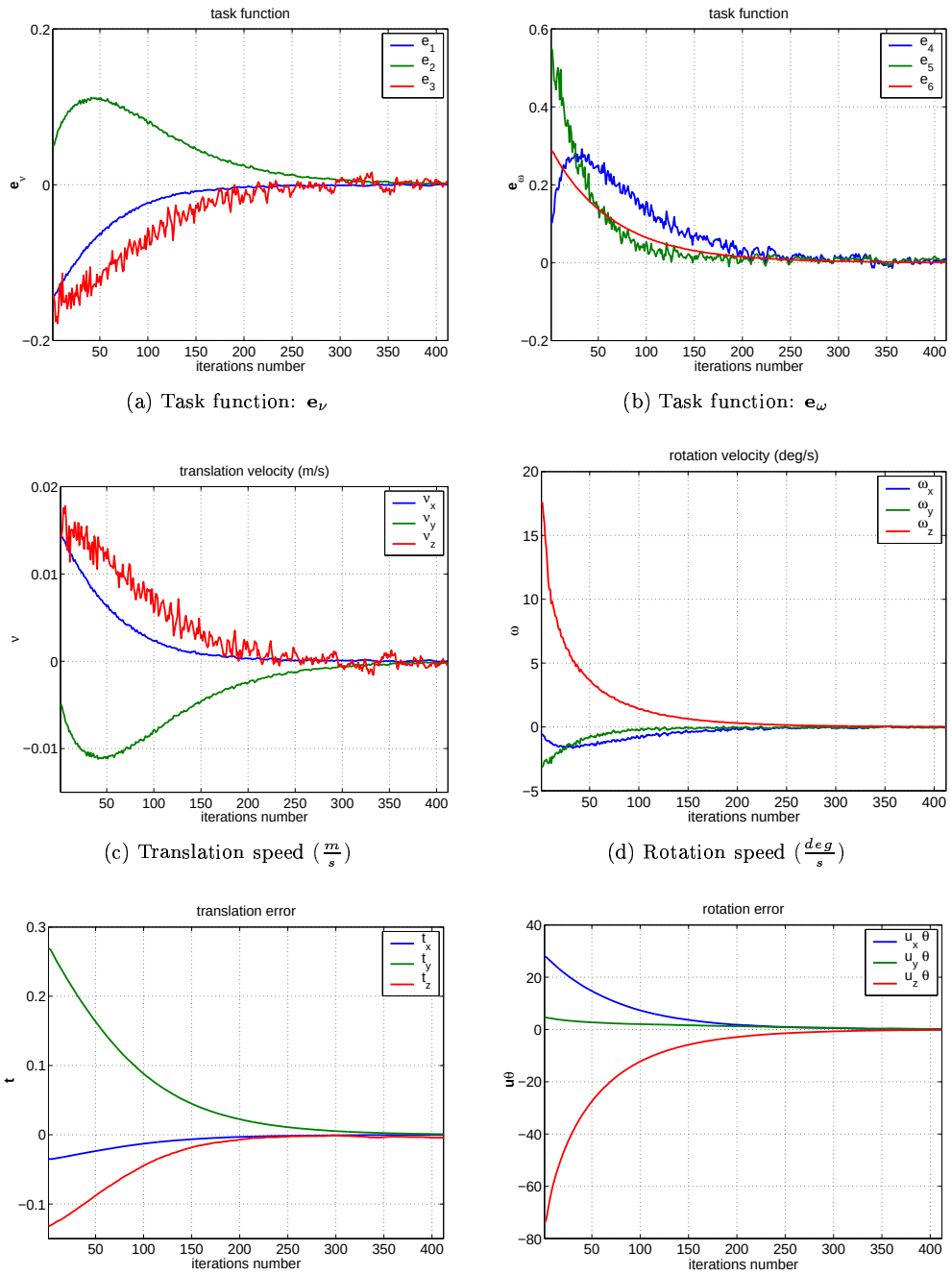


(c) Points trajectory



(d) Invariant space error

Figure 6: Experiment using a 12mm lens for learning and a 6mm lens for servoing. The visual servoing scheme is independent on the camera used for servoing.



(a) Task function: $\mathbf{e}_\nu$

(b) Task function: $\mathbf{e}_\omega$

(c) Translation speed ($\frac{m}{s}$)

(d) Rotation speed ($\frac{deg}{s}$)

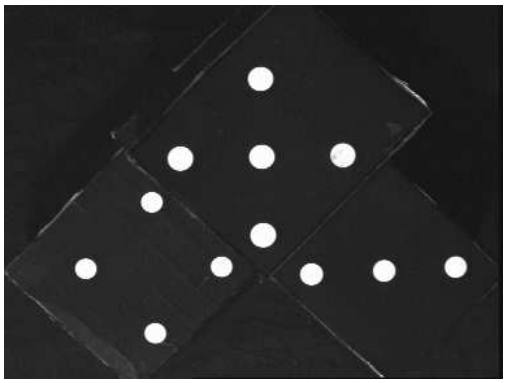
(e) Translation error (m)

(f) Rotation error (deg)

Figure 7: Experiment using a 12mm lens for learning and a 6mm lens for servoing. The visual servoing scheme is independent on the camera used for servoing.

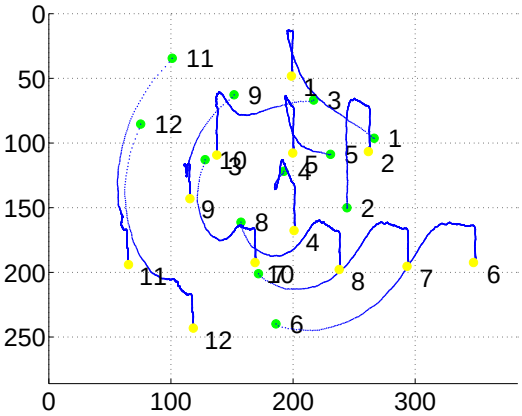
5.3 Path planning in the invariant space

The local stability analysis presented if the paper is valid if the error $\mathbf{q} - \mathbf{q}^*$ is sufficiently small at each iteration of the control law. As explained in Section 3, one can build a path in the invariant space such that the camera follows a straight line in the Cartesian space. Simple sufficient conditions on calibration errors have been obtained such that the tracking error is bounded (thus $\mathbf{q} - \mathbf{q}^*$ can be maintained small). In the following experiments, I show that the sufficient conditions for the local stability found in the theoretical analysis are very large. Indeed, a very bad approximation $\widehat{\mathbf{K}}$ of the current camera parameters is used in the control law (50% error on each parameter). On the other hand, the current depth distribution is computed, at each iteration of the control law, using the epipolar geometry as explained in Appendix A. For both experiments, a camera (with a 12mm lens) is mounted on the robot end-effector. After the reference image (Figure 8(a)) has been learned the robot is displaced to its initial position (Figure 8(b) shows the initial image of the object). The initial displacement of the robot is approximately 30 *cm* for the translation and 80 *deg* for the rotation (see the initial values of the curves in Figures 9(i) and (j)). A path $\mathbf{s}^*(t)$ is planned in the invariant space such that the camera follows a straight line in the 3D space. As already mentioned the bad estimation of $\mathbf{b}(t)$ has no influence on the stability of the servoing as far as $\mathbf{b}(t)$ is bounded. To show that, I set $\widehat{\mathbf{J}}^+ \partial \mathbf{s}^*(t)/\partial t = 0$ in the control law used in the experiments even if in reality $\partial \mathbf{s}^*(t)/\partial t \neq 0$. In the first experiment, the path is sampled $N = 500$ times. The first part of the control law corresponding to the translational velocity ν_x, ν_y, ν_z is plotted in Figure 9(g) while the second part corresponding to the translational velocity $\omega_x, \omega_y, \omega_z$ is plotted in Figure 9(h). The task function converges to zero (see Figures 9(e) and (f)) as well as the error in the invariant space (see Figures 8(d)). Consequently, the position error of the robot end-effector converges to zero (see Figures 9(i) and (j)). Figure 8(c) plots the trajectory of the points in the image (the blue curves) and shows that the points go from their initial position (green points) to their reference position (yellow points). This experiment proves that the new scheme can perform positioning tasks exactly as standard visual servoing techniques. Although the camera is controlled in the invariant space, the behaviour of the robot in the Cartesian space is very nice. Indeed, Figure (10) shows that the camera approximately follows a straight line despite the large calibration errors. As expected, the tracking error is bounded. Since the control law ν_z is very noisy, the error is mainly on t_z as is it shown by Figures (10)(c) and (d). In the second experiment, the robot starts approximately from the same initial position. The results of the visual servoing are therefore similar to the results shown in Figure 9. On the other hand, the path is sampled $N = 3000$ times. Thus, the speed of the reference trajectory is reduce by a factor 6. As a consequence, the tracking error presented in Figure 11 is considerably smaller than the tracking error presented in Figure 10.

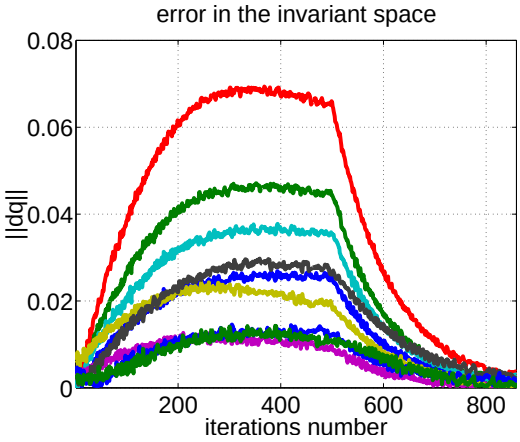


(a) Reference image

(b) Initial image

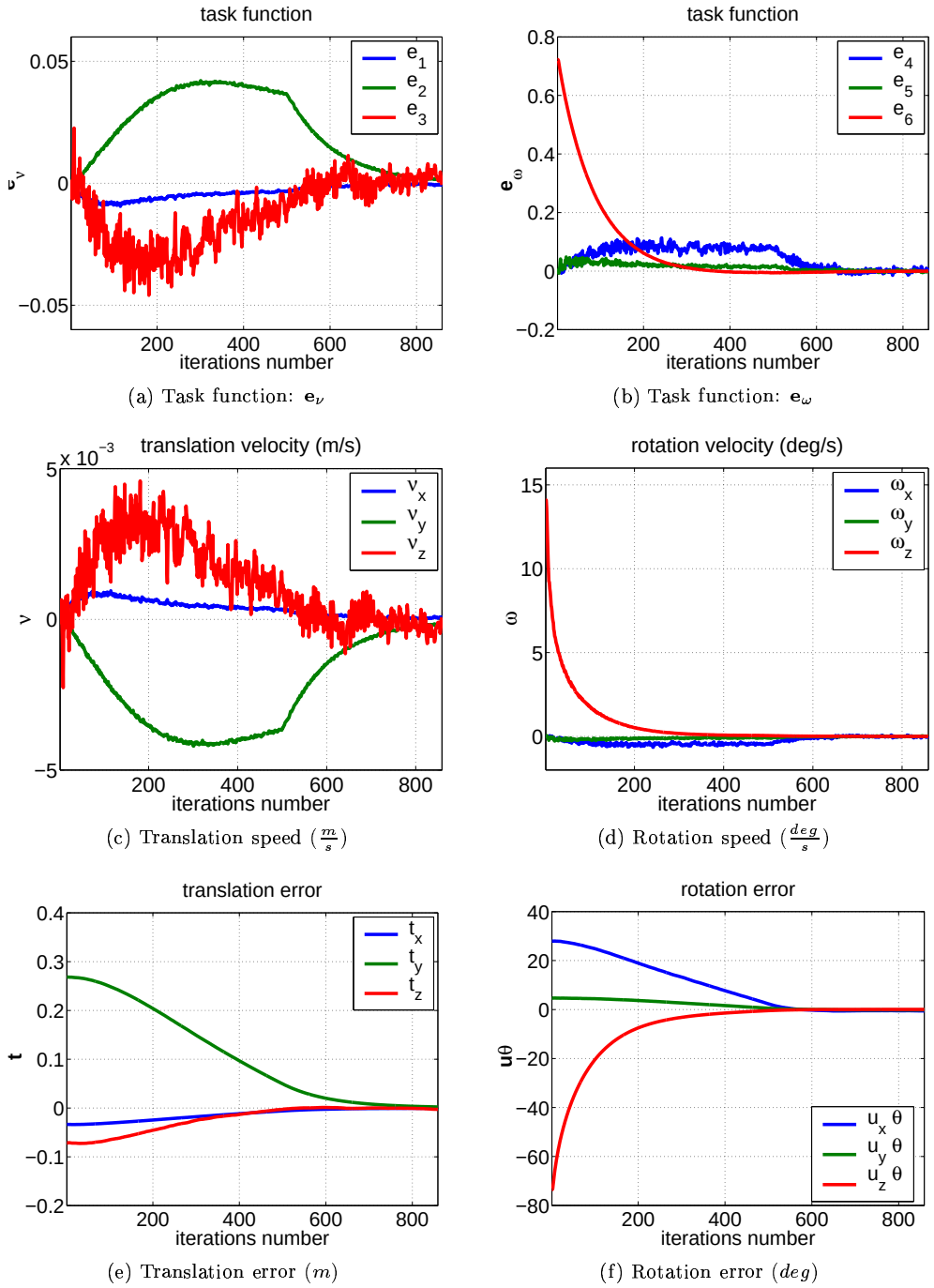


(c) Points trajectory



(d) Invariant space error

Figure 8: Path tracking in the invariant space. The error in the invariant space stay small. The final image coincide to the reference image since the camera is not zooming in this experiment.

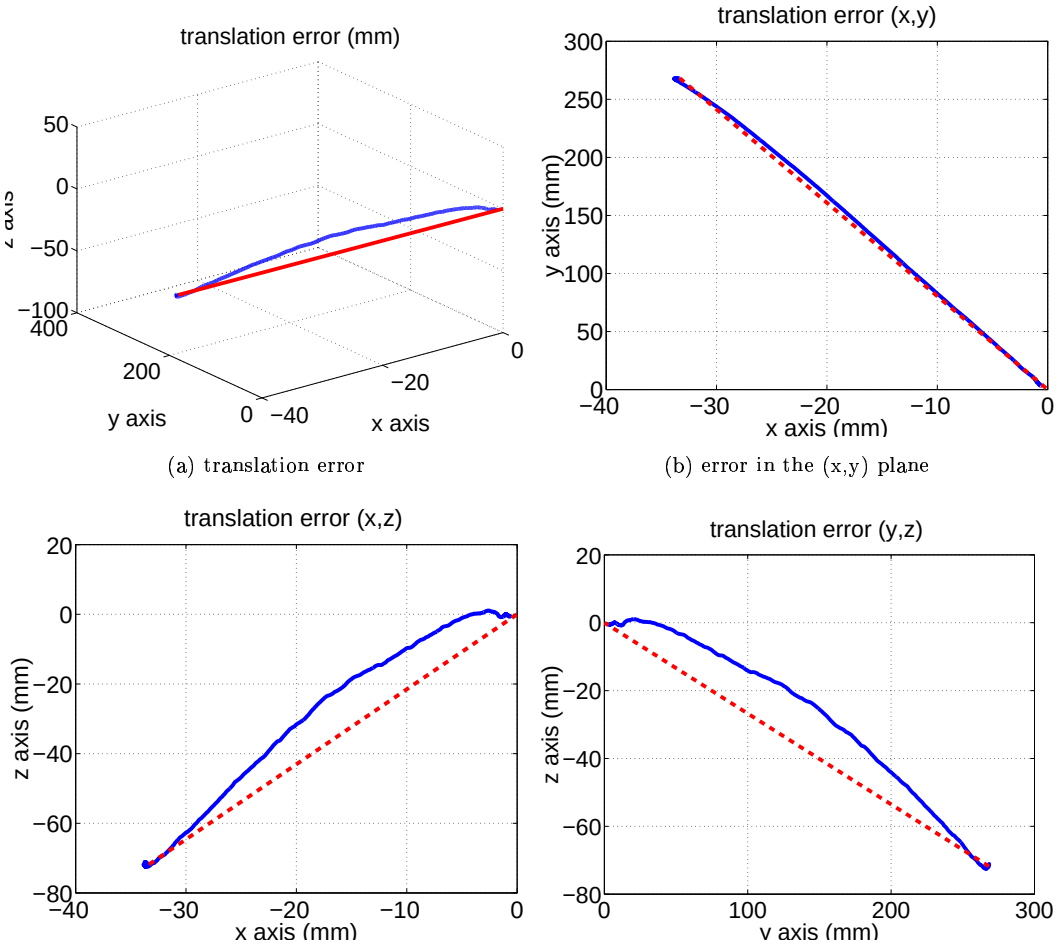


(a) Task function: $\mathbf{e}_\nu$ (b) Task function: $\mathbf{e}_\omega$

(c) Translation speed ($\frac{m}{s}$) (d) Rotation speed ($\frac{deg}{s}$)

(e) Translation error (m) (f) Rotation error (deg)

Figure 9: Path tracking in the invariant space. The invariant vision-based control is stable and, at the convergence, the camera is back to the reference position.



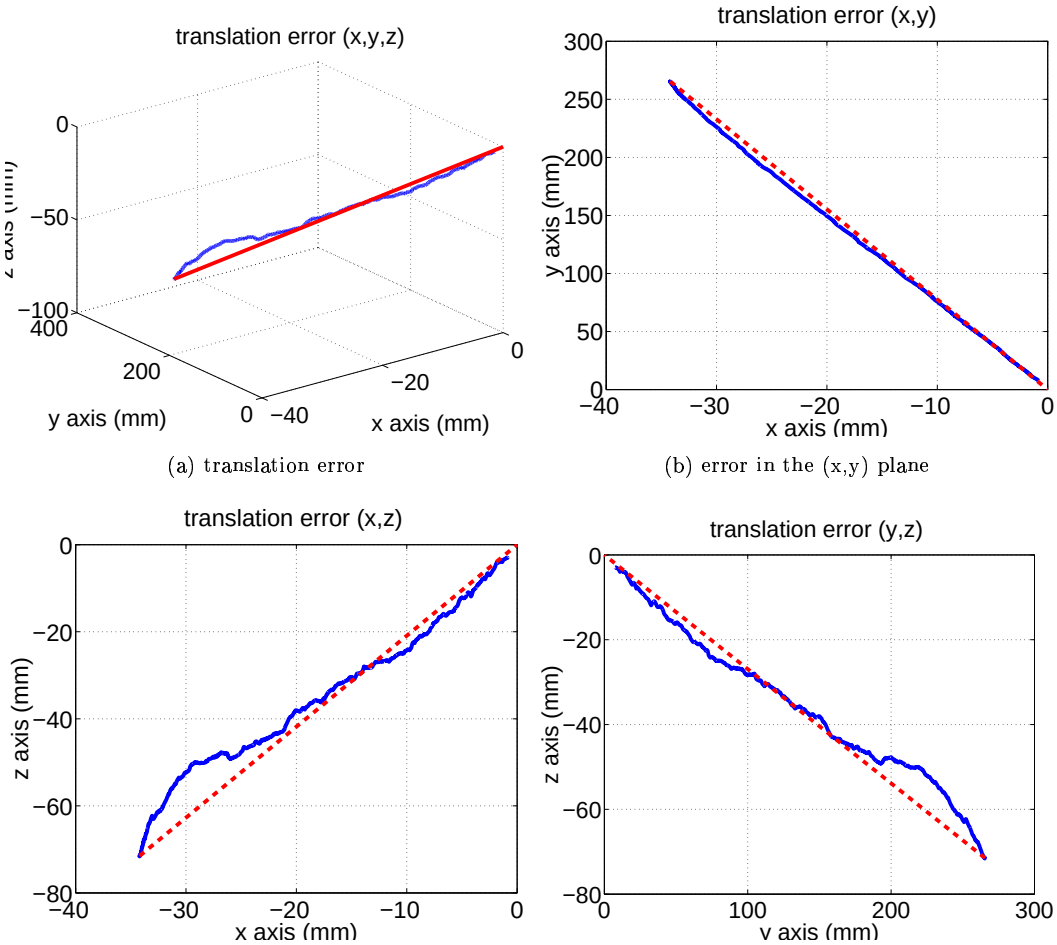
(a) translation error

(b) error in the (x,y) plane

(c) error in the (x,z) plane

(d) error in the (y,z) plane

Figure 10: The path is sampled with N=500. The camera (blue line) approximately follows a straight line (in red).



(a) translation error

(b) error in the (x,y) plane

(c) error in the (x,z) plane

(d) error in the (y,z) plane

Figure 11: The path is sampled with N=3000. The tracking error is reduced and the camera (blue line) follows more accurately straight the line (in red).

Acknowledgements

I would like to thank Eric Marchand (IRISA/INRIA Rennes, VISTA project) for his help in programming the experiments with the robot and François Gaspard (INRIA Sophia-Antipolis, ROBOTVIS) for providing the images of the zooming camera.

6 Conclusion

This work shows how to position a camera, with respect to a non-planar object, even if the intrinsic parameters change. The new visual servoing scheme will be useful especially when a zooming camera is mounted on the end-effector of the robot. The zoom can be used to enlarge the field of view of the camera and to bound the size of the object observed in the image (this can improve the robustness of features extraction from the images). The stability analysis of the visual servoing invariant to camera intrinsic parameters has shown that the control law proposed in the paper is robust to camera calibration errors. Despite the camera is controlled in a projective space, it is possible to follow a straight line in the Cartesian space.

Appendix A

Several approaches have been proposed in order to estimate the depths distribution on-line [5] [13]. These approaches requires some knowledge on the parameters of the system (camera calibration, robot motion,...). A different solution to the problem is to use projective geometry. Consider two images of the same 3D object taken with the same camera having intrinsic parameters $\overline{\mathbf{K}}$. Suppose that the displacement between the two images taken at the positions $\boldsymbol{\xi}^I$ and $\boldsymbol{\xi}^{II}$ is an unknown pure translation. A pure translation of the camera frame is very easy to obtain (for example one can mount the camera on a rail). In that case, the equation linking a point $\mathbf{p}_k^{II}$ in the second image to the point $\mathbf{p}_k^I$ in the first image is:

$$\frac{\zeta_k^{II}}{\zeta_k^I}\mathbf{p}_k^{II} = \mathbf{p}_k^I + \frac{\|\mathbf{e}\|}{\zeta_k^I}\frac{\mathbf{e}}{\|\mathbf{e}\|} \tag{26}$$

From this equation one can measure $\frac{\zeta_k^{II}}{\zeta_k^I}$, $\frac{\|\mathbf{e}\|}{\zeta_k^I}$ and $\frac{\mathbf{e}}{\|\mathbf{e}\|}$. Thus, without knowing the translation of the camera and without knowing its intrinsic parameters $\overline{\mathbf{K}}$, one can measure the depths distribution $\boldsymbol{\zeta}^I = (\zeta_1^I, \zeta_2^I, ..., \zeta_n^I)$ in frame $\mathcal{F}^{\mathcal{I}}$ up to a scale factor. Similarly, one can measure the depths distribution $\boldsymbol{\zeta}^{II} = (\zeta_1^{II}, \zeta_2^{II}, ..., \zeta_n^{II})$ in frame $\mathcal{F}^{\mathcal{I}\mathcal{I}}$. This can be done once and for all off-line. Note that the camera parameters $\overline{\mathbf{K}}$ can be completely different from the parameters used to learn the reference image $\mathbf{K}^* \neq \overline{\mathbf{K}}$ or those used during the servoing $\mathbf{K} \neq \overline{\mathbf{K}}$. Similarly, the two images can be taken at any position and generally not only $\boldsymbol{\xi}^I \neq \boldsymbol{\xi}^*$ and $\boldsymbol{\xi}^{II} \neq \boldsymbol{\xi}^*$ but also $\boldsymbol{\xi}^I \neq \boldsymbol{\xi}$ and $\boldsymbol{\xi}^{II} \neq \boldsymbol{\xi}$. Knowing the depths distribution $\boldsymbol{\zeta}^I$ or the depths distribution $\boldsymbol{\zeta}^{II}$, it is possible to use the epipolar geometry to measure the depths distribution $\boldsymbol{\zeta}$ and the depth distribution $\boldsymbol{\zeta}^*$. For example, between the image taken at frame $\mathcal{F}^{\mathcal{I}}$ and the current image taken at frame $\mathcal{F}$ we have:

$$\frac{\zeta_k^I}{\zeta_k}\mathbf{p}_k^I = (\overline{\mathbf{K}}\mathbf{R}^I\mathbf{K}^{-1})\mathbf{p}_k + \frac{\|\overline{\mathbf{K}}\mathbf{t}^I\|}{\zeta_k}\frac{\overline{\mathbf{K}}\mathbf{t}^I}{\|\overline{\mathbf{K}}\mathbf{t}^I\|} \tag{27}$$

where $\mathbf{R}^I$ and $\mathbf{t}^I$ are respectively the rotation and the translation between frames $\mathcal{F}^{\mathcal{I}}$ and $\mathcal{F}$. Since from equation (27) it is possible to measure the ratio $\frac{\zeta_k^I}{\zeta_k}$ and since ζ_k^I is known up

to a scale factor, one can measure all the ζ_k up to the same scale factor: $\widehat{\boldsymbol{\zeta}} = \kappa(t)\boldsymbol{\zeta}$. All the depths ζ_k must be positives, therefore $\kappa(t) > 0\ \forall t$.

Appendix B

Proof: [**Proposition 1**] The time varying linear system (19) when $\mathbf{b}(t) = 0$ and $\partial\mathbf{K}/\partial t = 0$ is:

$$\dot{\varepsilon} = -\lambda\mathbf{A}(t)\varepsilon = -\lambda \begin{bmatrix} \kappa(t)\widehat{\mathbf{K}}^{-1}\mathbf{K} & 0 \\ 0 & \widehat{\mathbf{G}}^{-1}\mathbf{F} \end{bmatrix} \varepsilon$$

The equilibrium point $\varepsilon = 0$ is locally asymptotically stable if and only if each block on the diagonal of $\mathbf{A}(t)$ has eigenvalues with positive real part. Indeed, if $\widehat{\mathbf{K}}^{-1}\mathbf{K} = \mathbf{C}_1^{-1}\mathbf{D}_1\mathbf{C}_1$ and $\widehat{\mathbf{G}}^{-1}\mathbf{F} = \mathbf{F}_2^{-1}\mathbf{D}_2\mathbf{C}_2$ then:

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{C}_1^{-1} & 0 \\ 0 & \mathbf{C}_2^{-1} \end{bmatrix} \begin{bmatrix} \kappa(t)\mathbf{D}_1 & 0 \\ 0 & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{C}_1 & 0 \\ 0 & \mathbf{C}_2 \end{bmatrix}$$

The nonsingular constant matrix $\mathbf{C} = \mathrm{diag}(\mathbf{C}_1, \mathbf{C}_2)$ defines a change of coordinates $\boldsymbol{\rho} = \mathbf{C}\varepsilon$. In the new coordinate system, the differential equation becomes:

$$\dot{\boldsymbol{\rho}} = -\lambda \begin{bmatrix} \kappa(t)\mathbf{D}_1 & 0 \\ 0 & \mathbf{D}_2 \end{bmatrix} \boldsymbol{\rho}$$

This system is asymptotically stable if and only if $\lambda > 0$, $\kappa(t) > 0$, $\mathbf{D}_1 > 0$ and $\mathbf{D}_2 > 0$. The entries on the diagonal of $\mathbf{D}_1 = \mathrm{diag}\left(\frac{fk_u}{\widehat{f}\widehat{k}_u}, \frac{fk_v \sin(\widehat{\theta})}{\widehat{f}\widehat{k}_v \sin(\theta)}, 1\right)$ and $\mathbf{D}_2 = \mathrm{diag}\left(\frac{\widehat{f}\widehat{k}_u}{fk_u}, \frac{\widehat{f}\widehat{k}_v \sin(\theta)}{fk_v \sin(\widehat{\theta})}\right)$ are positive if and only if $\widehat{f} > 0$, $\widehat{k}_u > 0$, $\widehat{k}_v > 0$ and $0 < \widehat{\theta} < \pi$.

Proof: [**Proposition 2**] The time varying linear system (19) when $\mathbf{b}(t) = 0$ is:

$$\dot{\varepsilon} = -\lambda\mathbf{A}(t)\varepsilon = -\lambda \begin{bmatrix} \kappa(t)\widehat{\mathbf{K}}^{-1}\mathbf{K}(t) & 0 \\ 0 & \widehat{\mathbf{G}}^{-1}\mathbf{F}(t) \end{bmatrix} \varepsilon$$

If $\kappa(t) > 0$, $\widehat{\mathbf{K}}^{-1}\mathbf{K}(t) > 0$ and $\widehat{\mathbf{G}}^{-1}\mathbf{F}(t)\ \forall t$ then $\mathbf{A}(t) > 0\ \forall t$. Consider the following positive Lyapounov function:

$$\mathcal{V}(\varepsilon) = \frac{1}{2}\varepsilon^T\varepsilon$$

The derivative of this function is:

$$\dot{\mathcal{V}}(\varepsilon) = \varepsilon^T\dot{\varepsilon} = -\lambda\varepsilon^T\mathbf{A}(t)\varepsilon$$

which is always negative if $\varepsilon \neq 0$, since $\lambda > 0$ and $\mathbf{A}(t) > 0$. Thus the system is locally stable. To prove the asymptotic convergence of ε to zero, we need also to show that it exists a neighbourhood $\mathcal{U}$ of $\boldsymbol{\xi}^* = 0$ such that $\varepsilon = \widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) \neq 0$, $\forall\boldsymbol{\xi} \in \mathcal{U}$ (i.e. $\varepsilon = 0$

only if $\mathbf{s}(\boldsymbol{\xi}) = \mathbf{s}^*$). Let us suppose that $\mathbf{s}(\boldsymbol{\xi}) \neq \mathbf{s}^*$ and therefore $\boldsymbol{\xi} \neq \boldsymbol{\xi}^* = 0$. The Taylor development of $\mathbf{s}(\boldsymbol{\xi})$ in a neighbourhood of $\boldsymbol{\xi}^* = 0$ is the following:

$$\mathbf{s} - \mathbf{s}^* = \mathbf{J}(\boldsymbol{\xi}^*)\,\boldsymbol{\xi} + O^2(\boldsymbol{\xi}) \tag{28}$$

Multiplying by $\boldsymbol{\xi}^T\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)$ (where $\boldsymbol{\xi}^T\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*) \neq 0$ since $\boldsymbol{\xi} \neq 0$ and $\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)$ is full rank) both sides of equation (28) we obtain:

$$\boldsymbol{\xi}^T\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) = \boldsymbol{\xi}^T\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)\mathbf{J}(\boldsymbol{\xi}^*)\boldsymbol{\xi} + O^3(\boldsymbol{\xi})$$

remember that if $\mathbf{A}(t) = \widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)\mathbf{J}(\boldsymbol{\xi}^*) > 0$ then $\boldsymbol{\xi}^T\mathbf{A}(t)\boldsymbol{\xi} \geq 2\sigma\|\boldsymbol{\xi}\|^2$, where $\sigma > 0$ is the minimum singular value of the positive definite matrix $\mathbf{A}(t) + \mathbf{A}^T(t)$. If $\widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) = 0$ then:

$$0 \geq 2\sigma\|\boldsymbol{\xi}\|^2 + O^3(\boldsymbol{\xi})$$

that means:

$$\|\boldsymbol{\xi}\|^2 \leq |O^3(\boldsymbol{\xi})|$$

which is impossible since, by definition of $O^3(\boldsymbol{\xi})$, it exists a neighbourhood of $\boldsymbol{\xi}^*$ in which:

$$\|\boldsymbol{\xi}\|^2 > |O^3(\boldsymbol{\xi})|$$

Therefore, $\varepsilon = \widehat{\mathbf{J}}^+(\boldsymbol{\xi}^*)(\mathbf{s} - \mathbf{s}^*) \neq 0$ in a neighbourhood of $\boldsymbol{\xi}^*$ and the system is locally asymptotically stable.

Proof: [**Proposition 3**] The time derivative of the norm of the task function is:

$$\frac{d\|\varepsilon\|^2}{dt} = \frac{d(\varepsilon^T\varepsilon)}{dt} = \varepsilon^T\dot{\varepsilon} + \dot{\varepsilon}^T\varepsilon$$

Using equation (19), we obtain:

$$\frac{d\|\varepsilon\|^2}{dt} = -\lambda\varepsilon^T(\mathbf{A}(t) + \mathbf{A}^T(t))\varepsilon + 2\mathbf{b}(t)^T\varepsilon \tag{29}$$

From Proposition 2, if the system is locally stable when $\mathbf{b}(t) = 0$, then $\mathbf{A}(t) > 0$. If $\sigma > 0$ is the minimum singular value of $(\mathbf{A} + \mathbf{A}^T)$, then $\varepsilon^T(\mathbf{A} + \mathbf{A}^T)\varepsilon \geq \sigma\|\varepsilon\|^2$. Suppose that $\frac{\partial \mathbf{s}^*(t)}{\partial t}$ is bounded (i.e. $\|\mathbf{b}(t)\| \leq \frac{1}{2}\psi$), equation (29) can be bounded as follow:

$$\frac{d\|\varepsilon\|^2}{dt} \leq -\varphi\,\|\varepsilon\|^2 + \psi\,\|\varepsilon\| \tag{30}$$

where $\varphi = \lambda\sigma$. Setting $\boldsymbol{\chi}(t) = \|\varepsilon\|^2$, consider the following Bernoulli differential equation:

$$\dot{\chi} = -\varphi\,\chi + \psi\chi^{\frac{1}{2}} \tag{31}$$

whose solution is:

$$\chi(t) = e^{-\int_0^t \varphi d\tau} \left(\chi(0) + \frac{1}{2} \int_0^t \psi e^{\frac{1}{2}\int_0^t \varphi d\tau} d\tau \right)^2 \tag{32}$$

Remember that $\|\boldsymbol{\varepsilon}(0)\| = 0$, thus $\chi(0) = 0$ and:

$$\chi(t) = \left(\frac{\psi}{\varphi}\right)^2 \left(1 - e^{-\frac{1}{2}\varphi t}\right)^2 \tag{33}$$

Finally, since $\|\boldsymbol{\varepsilon}(t)\| = \sqrt{\boldsymbol{\chi}(t)}$ we obtain:

$$\|\boldsymbol{\varepsilon}(t)\| \leq \frac{\psi}{\varphi} \tag{34}$$

From proposition 2, we know that in a neighbourhood of $\boldsymbol{\xi}^*$ we have $\|\boldsymbol{\varepsilon}\| = \|\widehat{\mathbf{J}}^+(\mathbf{s}-\mathbf{s}^*)\| \neq 0$ if $\mathbf{s} \neq \mathbf{s}^*$. Then, $\exists \delta > 0$ such that $\|\boldsymbol{\varepsilon}\| = \|\widehat{\mathbf{J}}^+(\mathbf{s}-\mathbf{s}^*)\| \geq \delta \|\mathbf{s}-\mathbf{s}^*\|$. Thus, from equation (34) one obtain:

$$\delta \|\mathbf{s} - \mathbf{s}^*\| \leq \|\boldsymbol{\varepsilon}(t)\| \leq \frac{\psi}{\varphi} \tag{35}$$

from which we can deduce that if $\|\boldsymbol{\varepsilon}\|$ is bounded, then error $\|\mathbf{s} - \mathbf{s}^*\|$ is also bounded:

$$\|\mathbf{s} - \mathbf{s}^*\| \leq \frac{\psi}{\delta\varphi} \tag{36}$$

References

[1] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. Kriegman, G. Hager, and A. Morse, editors, *The confluence of vision and control*, volume 237 of *LNCIS Series*, pages 66–78. Springer Verlag, 1998.

[2] F. Chaumette and E. Malis. 2 1/2 d visual servoing: a possible solution to improve image-based and position-based visual servoings. In *IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 630–635, San Francisco, USA, April 2000.

[3] P. I. Corke and S. A. Hutchinson. A new hybrid image-based visual servo control scheme. In *IEEE Conference on Decision and Control*, volume 3, pages 2521–2526, Sydney, NSW, Australia, December 2000.

[4] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3):313–326, June 1992.

[5] J. T. Feddema and O. R. Mitchell. Vision-guided servoing with feature-based trajectory generation. *IEEE Trans. on Robotics and Automation*, 5(5):691–700, October 1989.

[6] G. D. Hager. Calibration-free visual control using projective invariance. In *IEEE Int. Conf. on Computer Vision*, pages 1009–1015, MIT, Cambridge (USA), June 1995.

[7] K. Hashimoto. *Visual Servoing: Real Time Control of Robot manipulators based on visual sensory feedback*, volume 7 of *World Scientific Series in Robotics and Automated Systems*. World Scientific Press, Singapore, 1993.

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

[9] E. Malis. Vision-based control using different cameras for learning the reference image and for servoing. In *IEEE/RSJ International Conference on Intelligent Robots Systems*, pages xx–yy, Maui, Hawaii, November 2001.

[10] E. Malis. Visual servoing invariant to changes in camera intrinsic parameters. In *International Conference on Computer Vision*, volume 1, pages 704–709, Vancouver, Canada, July 2001.

[11] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):234–246, April 1999.

[12] Y. Mezouar and F. Chaumette. Path planning in image space for robust visual servoing. In *IEEE Int. Conf. on Robotics and Automation*, volume 3, pages 2759–2764, San Francisco, CA, April 2000.

[13] N. P. Papanikolopoulos and P. K. Kosla. Adaptive robot visual tracking: Theory and experiments. *IEEE Trans. on Robotics and Automation*, 38(3):429–445, March 1993.

[14] C. Samson, M. Le Borgne, and B. Espiau. *Robot Control: the Task Function Approach*, volume 22 of *Oxford Engineering Science Series*. Clarendon Press, Oxford, UK, 1991.

[15] 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(5):684–696, October 1996.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399