



HAL
open science

Improvements on visual servoing from spherical targets using a spherical projection model

R. Tatsambon Fomena, François Chaumette

► **To cite this version:**

R. Tatsambon Fomena, François Chaumette. Improvements on visual servoing from spherical targets using a spherical projection model. *IEEE Transactions on Robotics*, 2009, 25 (4), pp.874-886. inria-00436920

HAL Id: inria-00436920

<https://inria.hal.science/inria-00436920>

Submitted on 27 Nov 2009

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.

Improvements on Visual Servoing From Spherical Targets Using a Spherical Projection Model

Romeo Tatsambon Fomena, *Member, IEEE*, and François Chaumette, *Member, IEEE*

Abstract—This paper is concerned with improvements to visual feature modeling using a spherical projection model. Three spherical targets are considered: a sphere, a sphere marked with a tangent vector to a point on its surface, and a sphere marked with two points on its surface. A new minimal and decoupled set of visual features is proposed for each target using any central catadioptric camera. Using the proposed set for a sphere, a classical control law is proved to be globally asymptotically stable in the presence of modeling errors and locally asymptotically stable in the presence of calibration errors, considering that perspective and paracatadioptric cameras were used. Simulation and experimental results with perspective, paracatadioptric, and fish-eye cameras validate the proposed theoretical results.

Index Terms—Central catadioptric systems, spherical projection, visual servoing.

I. INTRODUCTION

VISUAL servoing consists of using data provided by a vision sensor to control the motion of a dynamic system [1]. A vision sensor provides a large spectrum of potential visual features. However, if no planning strategy is developed, the use of some visual features as input of the control scheme may lead to stability problems if the displacement that the robot has to achieve is very large [2]. For this reason, we need to design ideal visual features for visual servoing. Ideally, satisfaction of the following criteria is expected: local and—as far as possible—global stability of the system, robustness to calibration and modeling errors, nonsingularity, local minima avoidance, satisfactory trajectory of the system and of the features in the image, and, finally, a maximal decoupled and linear link (the ultimate goal) between the visual features and the degrees of freedom (DOFs) taken into account.

Manuscript received July 11, 2008; revised January 19, 2009 and April 27, 2009. First published June 5, 2009; current version published July 31, 2009. This paper was recommended for publication by Associate Editor P. Rives and Editor L. Parker upon evaluation of the reviewers' comments. This paper was presented in part at the IEEE International Conference on Robotics and Automation, Rome, Italy, April 2007, in part at the International Conference on Advanced Robotics, Jeju, Korea, August 2007, and in part at the IEEE/Robotics Society of Japan International Conference on Intelligent Robots and Systems, Nice, France, September 2008.

R. T. Fomena is with IRISA, Université de Rennes 1, 35042 Rennes Cedex, France (e-mail: rtatsamb@irisa.fr).

F. Chaumette is with INRIA Rennes-Bretagne Atlantique, IRISA, 35042 Rennes Cedex, France (e-mail: chaumett@irisa.fr).

This paper has supplementary downloadable multimedia material available at <http://ieeexplore.ieee.org> provided by the author. This material includes a video, *Tatsambon_Chaumette_video_TRO.avi*. The video is an mpeg4 format and is playable using QuickTime, RealPlayer, or Microsoft Windows Media Player. The size of the video is not specified. Contact rtatsamb@irisa.fr for further questions about this work.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2009.2022425

Several methods have been proposed to approach an ideal system behavior. These methods are usually classified into 3-D, hybrid, and 2-D visual servoings. In 3-D visual servoing, features in the 3-D Cartesian space are used as inputs to the control scheme [3]. Those features are obtained from the relative pose of the object with respect to (w.r.t.) the camera. This pose can be recovered using the geometric model of the object [4], [5]. Pose estimation is the key issue of this approach, since small amount of image noise could lead to large errors in the pose estimation. When the pose is correctly estimated, this class of visual servoing is known to provide adequate system motion in the Cartesian space either in the moving camera frame or in a fixed frame [6]. In the latter case, the fact that there is no control in the image might cause visual servoing to fail, since visual measurements used for pose estimation can leave the field of view (FOV). Using coordinates of several 3-D points that are selected on the target can potentially keep the target in the FOV [7].

Another solution to the problem of ideal features modeling is to use hybrid methods that combine 3-D and 2-D data. This solution has been exploited in 2-1/2-D visual servoing, where a decoupled control scheme with no singularity in the whole task space and tolerating a coarsely calibrated system has been designed [8]. However, there are some drawbacks to this method: It is more sensitive to image noise (like 3-D visual servoing) than 2-D visual servoing, which uses directly features extracted in the image as control inputs.

In fact, 2-D visual servoing is appealing because of its robustness to camera calibration errors [9] and to image noise. It is well known that by using only simple 2-D data, such as image-point coordinates, there is no control over the system motion in the Cartesian space, and possible singularities or local minima may be reached, thus leading to a limited convergence domain [2]. This is one of the reasons why a lot of work has been (and is still being) done to improve the system behavior using only 2-D data. For a satisfactory motion of the system in the Cartesian space, the z -axis translational and rotational motions can be decoupled from the other DOFs through a partitioned approach [10]. Another option for the decoupling of the optical axis motions is to use cylindrical coordinates [11]. A generic and intuitive representation of the image of a target can be obtained using 2-D moments. Recently, moment invariants theory has been used to determine specific combinations of 2-D moments whose interaction with the system presents linear and decoupling properties when planar objects are considered [12].

All the previously mentioned studies are based on a perspective projection model. Other models are, however, also suitable. For instance, a spherical projection model has already been used in [13] and [14]. Considering as a target a sphere marked with a

tangent vector to a point on its surface (which we refer to as CC target), this model has also been used in [15] to propose a set of features based on a diffeomorphism. This diffeomorphism has been coupled with navigation functions in [15] to design a global controller in a well-defined domain while taking into account occlusions and an FOV boundary.

In this paper, the CC target is revisited, and the diffeomorphism proposed in [15] is improved to decouple the orientation control from the camera translation control. The obtained decoupling allows significant improvements on the system behavior using a simple control law, i.e., in this paper, we have not considered high-level control strategies, such as navigation functions [16] or path-planning techniques [17], to take into account constraints, such as the visibility constraint or the avoidance of occlusion, joint limits, and singularities. This could, of course, be done in future work.

We also consider a more realistic target that is composed of a sphere marked with two points on its surface (which we refer to as special sphere). Indeed, from a practical point of view, the special sphere is more natural than the CC target. In addition, the special sphere can be used as an alternative passive target for spacecraft autonomous docking, as it has been done with three spheres in [18].

In the next section, we first consider a simple sphere as a target for which we propose a new minimal set of three independent features. The interaction matrix related to those features is maximally decoupled and presents a linear link between the visual features and the camera translational velocities. In Section III, this set is shown to be general in the sense that it can be computed from the image of a sphere on any central catadioptric system. Due to their large FOV, visual servoing with such a sensor is indeed appealing, while the interaction matrix of features extracted in their image plane is usually complex [19], [20]. For paracatadioptric systems, the general set of features is not always suitable because these features mostly draw a straight-line trajectory in the image space that may cross the dead angle in the center of the image of such cameras. Therefore, a specific set is designed for such cameras using cylindrical coordinates of the general set. For both sets of features (general and specific), a theoretical analysis of the stability and the robustness of a classical control law w.r.t. modeling error is given in Section IV. In addition, the effects of camera calibration errors are analyzed with regard to perspective and paracatadioptric cameras. We then consider the CC target and the special sphere in Section V. For these two targets, we design a new minimal set of six visual features to control the 6 DOFs of the system. For the CC target, the new set is compared with the previous set, as proposed in [15]. In Section VI, we validate the theoretical results obtained for each considered target in simulations and on real experiments with perspective, paracatadioptric, and fish-eye cameras. Finally, in Section VII, we discuss the position of our method w.r.t. classical visual servoing approaches.

II. VISUAL FEATURES DESIGN

In this section, a spherical projection model is used to design a new set of three independent features for visual servoing from

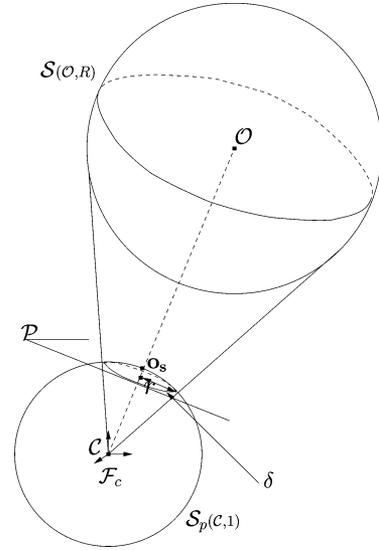


Fig. 1. Spherical projection of a sphere.

spheres. We first present the spherical projection of a sphere, and then, we select a new minimal set of three visual features.

A. Spherical Projection of a Sphere

Let $\mathcal{F}_c = (\mathcal{C}, \mathbf{i}, \mathbf{j}, \mathbf{k})$ be the center of projection frame, and $\mathcal{S}_p(\mathcal{C}, 1)$ be the unit sphere of center \mathcal{C} . In addition, let $\mathcal{S}(\mathcal{O}, R)$ be a sphere of radius R and center \mathcal{O} , and let \mathbf{O} be the vector coordinates of \mathcal{O} in \mathcal{F}_c (see Fig. 1). The spherical projection of $\mathcal{S}(\mathcal{O}, R)$ on $\mathcal{S}_p(\mathcal{C}, 1)$ is a dome hat [15]. This dome hat can be characterized by the contour δ of its base. We now determine the analytical form of δ . The equation that defines $\mathcal{S}(\mathcal{O}, R)$ is given by

$$\|\mathbf{M} - \mathbf{O}\|^2 - R^2 = 0 \quad (1)$$

where \mathbf{M} is the vector coordinates of a point $\mathcal{M} \in \mathcal{S}(\mathcal{O}, R)$. Let $\mathbf{m}_s = \pi_s(\mathbf{M}) = \mathbf{M}/\|\mathbf{M}\|$ be the spherical projection of \mathbf{M} . Multiplying (1) by $1/\|\mathbf{M}\|^2$ gives

$$\left\| \mathbf{m}_s - \frac{1}{\|\mathbf{M}\|} \mathbf{O} \right\|^2 - \frac{R^2}{\|\mathbf{M}\|^2} = 0 \quad (2)$$

which can be rewritten in a $1/\|\mathbf{M}\|$ polynomial function

$$\frac{K_O^2}{\|\mathbf{M}\|^2} - \frac{2}{\|\mathbf{M}\|} \mathbf{O}^\top \mathbf{m}_s + 1 = 0 \quad (3)$$

where $K_O^2 = \|\mathbf{O}\|^2 - R^2$. The contour δ is such that the discriminant of (3) is equal to 0 [only one solution exists for (3)]

$$(\mathbf{O}^\top \mathbf{m}_s)^2 - K_O^2 = 0. \quad (4)$$

Since δ is in front of the projection center, we obtain from (4) the equation of the plane \mathcal{P} that supports δ (see Fig. 1)

$$\mathbf{O}^\top \mathbf{m}_s = K_O. \quad (5)$$

To sum up, δ is the intersection of the unit sphere $\mathcal{S}_p(\mathcal{C}, 1)$ and plane \mathcal{P} , which is given by

$$\delta = \left\{ \begin{array}{l} \|\mathbf{m}_s\| = 1 \\ \mathbf{O}^\top \mathbf{m}_s = K_O. \end{array} \right. \quad (6)$$

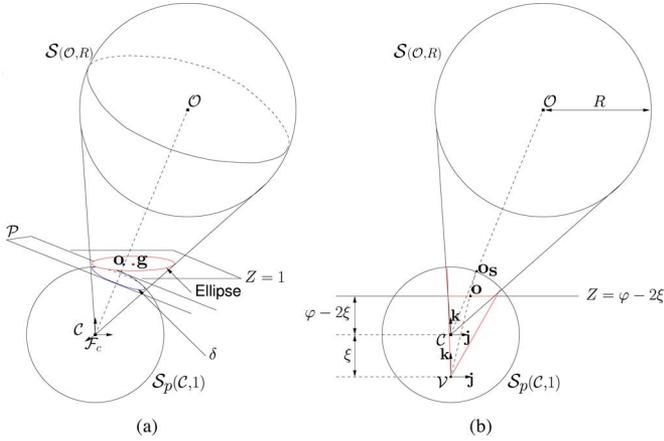


Fig. 2. Central catadioptric image of a sphere. (a) Perspective projection ($\varphi = 1$, $\xi = 0$). (b) General case.

The contour δ is, therefore, a circle from which it is easier to extract ideal visual features than from the ellipse observed on the image plane of a perspective camera [see Fig. 2(a)] or of an omnidirectional camera [see Fig. 2(b)].

B. Visual Features Selection

We recall that the interaction matrix \mathbf{L}_f related to a set of features $\mathbf{f} \in \mathbb{R}^n$ is defined such that $\dot{\mathbf{f}} = \mathbf{L}_f \mathbf{v}_c$, where $\mathbf{v}_c = (\mathbf{v}, \boldsymbol{\omega}) \in \mathfrak{se}(\mathbf{3})$ is the instantaneous camera velocity [21], \mathbf{v} and $\boldsymbol{\omega}$ are the translational and the rotational velocities of the camera, respectively, and $\mathfrak{se}(\mathbf{3}) \simeq \mathbb{R}^3 \times \mathbb{R}^3$ is the Lie algebra of the Lie group of displacements $\mathbf{SE}(\mathbf{3})$.

Second-order 2-D moments have been used previously to design a set of three features for visual servoing from spheres [21]. However, the corresponding interaction matrix is highly coupled and nonlinear. The set $\mathbf{s}_r = (r, \mathbf{o}_s)$ has been recently proposed in [15], where $r = R/\|\mathbf{O}\|$ is the radius of δ and $\mathbf{o}_s = \mathbf{O}/\|\mathbf{O}\|$ is the dome hat summit (see Fig. 1). However, this set is not minimal since three parameters are sufficient to characterize the image of a sphere [21]. This is why we prefer the minimal set of features

$$\mathbf{s} = \frac{1}{r} \mathbf{o}_s = \frac{1}{R} \mathbf{O}. \quad (7)$$

Since $\|\mathbf{o}_s\| = 1$, it is easy to show the one-to-one mapping between \mathbf{s} and \mathbf{s}_r . On one hand, we can note that \mathbf{s} is a pure 2-D feature set, since $\mathbf{s} = (1/r)\mathbf{o}_s$ is directly obtained from the spherical image. On the other hand, $\mathbf{s} = (1/R)\mathbf{O}$ is also a pure 3-D visual feature set, since it is nothing but the coordinates of a 3-D point up to a constant scalar value. The set \mathbf{s} has many advantageous characteristics: Its interaction matrix \mathbf{L}_s is simple, maximally decoupled, and always of full rank 3 as soon as $R \neq 0$ and $R \neq +\infty$, which means there is no singularity in the whole task space. Indeed, we have [22] the following:

$$\mathbf{L}_s = \left[-\frac{1}{R} \mathbf{I}_3 \quad [\mathbf{s}]_{\times} \right] \quad (8)$$

where $[\mathbf{s}]_{\times}$ is the skew matrix related to \mathbf{s} . In addition to the decoupling property, \mathbf{L}_s presents the same dynamic $(1/R)$ in the translational velocities. Since R is a constant, there is a linear

link between the visual features and the camera translational velocities. We can also see that \mathbf{s} presents the passivity property [14] (i.e., $\|\mathbf{s}\|$ is independent of the rotational velocity).

The only unknown 3-D parameter in \mathbf{L}_s is the constant R . In practice, \hat{R} (an estimated value of R) is used instead. In Section IV, we will analyze the stability and the robustness of a simple control law w.r.t. error on the estimation \hat{R} .

We now present the computation of \mathbf{s} from the image plane of any central catadioptric system.

III. APPLICATION TO CENTRAL CATADIOPTIC SYSTEMS

A central catadioptric system consists of a couple (camera, mirror) that has a single viewpoint [23]. In this section, we first present the analytical form of the image of the sphere on any central catadioptric system. Then, we show how to compute \mathbf{s} from this analytical form using 2-D moments. Finally, we present a specific set of features for cameras that have a dead angle in the center of their FOV.

A. Equation of the Catadioptric Image of a Sphere

Let (φ, ξ) be the mirror parameters of a central catadioptric system. The catadioptric image of a sphere is an ellipse. Image formation can be decomposed in two steps [see Fig. 2(b)] while we consider the unified model proposed in [24]. This model is also valid for perspective projection, which is a particular case of central catadioptric projection with $\xi = 0$ and $\varphi = 1$, as shown in Fig. 2(a). Let \mathcal{F}_v be the sensor frame ($\mathcal{V}, \mathbf{i}, \mathbf{j}, \mathbf{k}$). Note that \mathcal{C} is the unique viewpoint of the system.

The first step is the spherical projection of $\mathcal{S}(\mathcal{O}, R)$, which has already been presented in Section II-A which and is given by (6)

$$\delta = \begin{cases} m_{sx}^2 + m_{sy}^2 + m_{sz}^2 = 1 \\ m_{sx}O_x + m_{sy}O_y + m_{sz}O_z = K_O \end{cases} \quad (9)$$

where $\mathbf{m}_s = (m_{sx}, m_{sy}, m_{sz})$, and $\mathbf{O} = (O_x, O_y, O_z)$. The second step is the perspective projection of δ onto the image plane. For a point, the corresponding equations are given by

$$m_x = \frac{m_{sx}}{m_{sz} + \xi}, \quad m_y = \frac{m_{sy}}{m_{sz} + \xi} \quad (10)$$

where $\mathbf{m} = (m_x, m_y)$ is the vector coordinates of the image of \mathbf{m}_s . Substituting (10) into (9) gives

$$\begin{cases} \frac{1}{m_{sz} + \xi} = \frac{m_x O_x + m_y O_y + O_z}{K_O + \xi O_z} & (11a) \\ m_x^2 + m_y^2 + 1 - 2 \frac{\xi}{m_{sz} + \xi} + \frac{\xi^2 - 1}{(m_{sz} + \xi)^2} = 0. & (11b) \end{cases}$$

Finally, putting (11a) in (11b) leads to the equation of an ellipse defined by

$$k_0 m_x^2 + k_1 m_y^2 + 2k_2 m_x m_y + 2k_3 m_x + 2k_4 m_y + k_5 = 0 \quad (12)$$

with

$$\begin{cases} k_0 = (K_O + \xi O_z)^2 + (\xi^2 - 1) O_x^2 \\ k_1 = (K_O + \xi O_z)^2 + (\xi^2 - 1) O_y^2 \\ k_2 = (\xi^2 - 1) O_x O_y \\ k_3 = O_x ((\xi^2 - 1) O_z - \xi (K_O + \xi O_z)) \\ k_4 = O_y ((\xi^2 - 1) O_z - \xi (K_O + \xi O_z)) \\ k_5 = (K_O + \xi O_z)^2 + (\xi^2 - 1) O_z^2 - 2\xi O_z (K_O + \xi O_z). \end{cases}$$

B. Visual Feature Computation

Now, we present the computation of $\mathbf{s} = (s_x, s_y, s_z)$, which is defined in (7) using the ellipse moments $\boldsymbol{\mu} = (g_x, g_y, n_{20}, n_{11}, n_{02})$ measured on the catadioptric image plane. More precisely, $\mathbf{g} = (g_x, g_y)$ is the coordinates of the center of gravity of the ellipse, and n_{20} , n_{11} , and n_{02} are the normalized centered moments of order 2. After some computations presented in [25], we obtain from (12)

$$\begin{cases} s_x = \frac{g_x h_2}{\sqrt{h_2 + (1 - \xi^2)}} \\ s_y = \frac{g_y h_2}{\sqrt{h_2 + (1 - \xi^2)}} \end{cases} \quad (13)$$

where $h_2 = h_2(\boldsymbol{\mu}) = (g_x^2 + g_y^2) / (4n_{20}g_y^2 + 4n_{02}g_x^2 - 8n_{11}g_x g_y)$. It is possible to demonstrate that $h_2(\boldsymbol{\mu})$ is continuous even when $g_x = g_y = 0$, in which case, $h_2(\boldsymbol{\mu}) = 1/4n_{20}$.

In the case of paracatadioptric systems (where $\xi = 1$, which corresponds to the coupling of a parabolic mirror with an orthographic camera), we have [25]

$$s_z = \frac{h_2 - (s_x^2 + s_y^2 - 1)}{2\sqrt{h_2}} \quad (14)$$

and for all other catadioptric systems ($\xi \neq 1$), we have

$$s_z = \left[h_1 - \xi \sqrt{h_1^2 + (1 - \xi^2)(s_x^2 + s_y^2 - 1)} \right] / (1 - \xi^2) \quad (15)$$

where $h_1 = \sqrt{h_2 + (1 - \xi^2)}$. Note that, for the classical perspective projection, where $\xi = 0$ and $\varphi = 1$, we simply have

$$\begin{cases} s_x = \frac{g_x h_2}{\sqrt{h_2 + 1}} \\ s_y = \frac{g_y h_2}{\sqrt{h_2 + 1}} \\ s_z = \sqrt{h_2 + 1}. \end{cases} \quad (16)$$

To conclude, the set \mathbf{s} is general in the sense that it can be computed from the image of a sphere on any central catadioptric system. As we will see in Section VI-D4, it can also be obtained from fish-eye cameras, even if such sensors do not satisfy the unique viewpoint constraint [26].

From (13), we can note that s_x and s_y are almost linearly linked to g_x and g_y . Consequently, the ellipse center will almost draw a straight-line trajectory in the image plane of any catadioptric system using the simple control law used in Section IV. This is not always suitable for paracatadioptric cameras since there is a dead angle in the center of their image (see Fig. 3). For

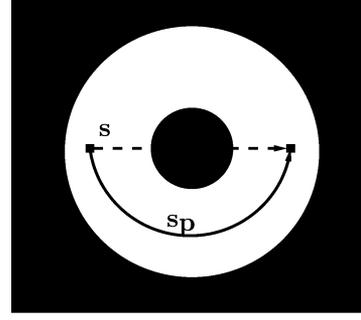


Fig. 3. Coordinate system dependence of the feature motion.

this reason, we present now a specific set that is more suitable for such cameras.

C. Visual Features for Omnidirectional Cameras

We propose to use the cylindrical coordinates of \mathbf{s} that will prevent the loss of the target in the dead angle by enforcing a circular feature motion (see Fig. 3). The new specific visual features $\mathbf{s}_p = (\rho, \theta, s_z)$ computed from \mathbf{s} are given by [27]

$$\begin{cases} \rho = \sqrt{s_x^2 + s_y^2} \\ \theta = \arctan\left(\frac{s_y}{s_x}\right). \end{cases} \quad (17)$$

In addition to the better features motion in the image, it is important to note that the feature ρ can never be 0 on a paracatadioptric image plane since the region where the target is visible does not include the center of the image. Thus, θ is always defined.

The interaction matrix related to \mathbf{s}_p is given by

$$\mathbf{L}_{\mathbf{s}_p} = \begin{bmatrix} -c/R & -s/R & 0 & s s_z & -c s_z & 0 \\ s/(\rho R) & -c/(\rho R) & 0 & c s_z/\rho & s s_z/\rho & -1 \\ 0 & 0 & -1/R & -\rho s & \rho c & 0 \end{bmatrix} \quad (18)$$

where $c = \cos \theta$, and $s = \sin \theta$. From this interaction matrix, we can see that s_z is the only feature that is sensitive to the z -translation, while θ is the only feature related to the rotation around the optical axis. This constrains the features motion to avoid the dead angle. Once again, the only unknown 3-D parameter in $\mathbf{L}_{\mathbf{s}_p}$ is the constant R . As mentioned before, an estimated value \hat{R} of R is used in practice. In the next section, the robustness domain of a classical control law in the presence of error on the estimation \hat{R} will be given.

IV. STABILITY ANALYSIS

In this section, we analyze separately the stability to modeling errors and to calibration errors. Considering either the general features \mathbf{s} or the specific features \mathbf{s}_p , a necessary-and-sufficient condition is given for the stability w.r.t. modeling error. In addition, the influence of the camera calibration errors is analyzed, and a necessary-and-sufficient condition is given for the local stability using perspective and paracatadioptric cameras.

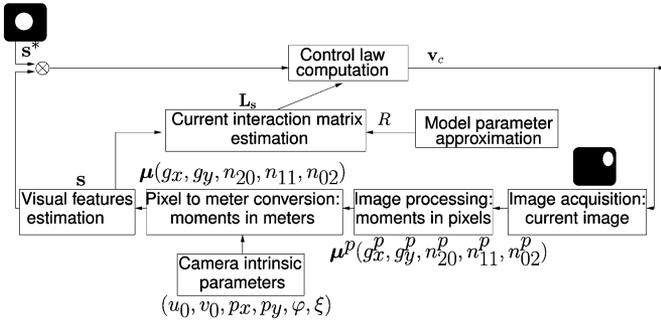


Fig. 4. Visual servoing closed loop.

Let us first consider the set s using a central catadioptric camera. We use the classical control law

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}}_s^+ (\mathbf{s} - \mathbf{s}^*) \quad (19)$$

where \mathbf{v}_c is the camera velocity sent to the low-level robot controller, λ is a positive gain, and $\widehat{\mathbf{L}}_s^+$ is the pseudoinverse of an approximation of \mathbf{L}_s .

Fig. 4 describes the visual servoing closed-loop system. For a catadioptric system with mirror parameter (φ, ξ) , the pixel-to-meter conversion is given by [24]

$$\begin{cases} g_x = \frac{g_x^p - u_0}{p_x(\varphi - \xi)}, & g_y = \frac{g_y^p - v_0}{p_y(\varphi - \xi)} \\ n_{20} = \frac{n_{20}^p}{p_x^2(\varphi - \xi)^2}, & n_{11} = \frac{n_{11}^p}{p_x p_y(\varphi - \xi)^2}, & n_{02} = \frac{n_{02}^p}{p_y^2(\varphi - \xi)^2} \end{cases} \quad (20)$$

where $u_0, v_0, p_x, p_y, \varphi$, and ξ are the sensor intrinsic parameters (and we recall that $\varphi = 1, \xi = 0$ for a classical perspective camera). From Fig. 4, we can see that there are three potential sources of errors in the system: modeling error arising from the approximation of R , calibration errors affecting the pixel-to-meter conversion, and image processing errors.

A. Stability to Modeling Error

As is usually done for classical stability analysis (but in [15] where navigation functions are designed to take into account the visibility of the target), we assume that the features can always be computed, which means the sphere is always in the FOV, and that there is no joint limits problems, which would make the realization of the control law impossible.

Assuming that we have neither image processing errors nor calibration errors, the closed-loop system equation can be written as follows:

$$\dot{\mathbf{s}} = -\lambda \mathbf{L}_s \widehat{\mathbf{L}}_s^+ (\mathbf{s} - \mathbf{s}^*) \quad (21)$$

with

$$\widehat{\mathbf{L}}_s^+ = \begin{bmatrix} -\frac{\widehat{R}r^2}{r^2 + \widehat{R}^2} (\widehat{R}^2 \mathbf{s} \mathbf{s}^\top + \mathbf{I}_3) \\ -\frac{\widehat{R}^2 r^2}{r^2 + \widehat{R}^2} [\widehat{\mathbf{s}}]_\times \end{bmatrix}$$

where $r = 1/\|\mathbf{s}\|$. Since $\text{rank } \mathbf{L}_s = 3$ in the whole task space, a sufficient condition for the global asymptotic stability

(GAS) is $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ > 0$. The eigenvalues of the symmetric matrix $\mathbf{L}_s \widehat{\mathbf{L}}_s^+$ can be computed and are given by \widehat{R}/R and $\widehat{R}(r^2 + R\widehat{R})/(R(r^2 + \widehat{R}^2))$ (which is a double eigenvalue). We thus have

$$\mathbf{L}_s \widehat{\mathbf{L}}_s^+ > 0 \iff \frac{\widehat{R}(r^2 + R\widehat{R})}{R(r^2 + \widehat{R}^2)} > 0 \text{ and } \frac{\widehat{R}}{R} > 0 \iff \widehat{R} > 0.$$

Therefore, the system is GAS if $\widehat{R} > 0$, which is also a necessary condition. Indeed, if $\widehat{R} \leq 0$ then $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ \leq 0$. Consequently, the system would diverge by choosing $\widehat{R} < 0$.

It is also possible to show that control law (19) is GAS using \mathbf{s}_p if and only if $\widehat{R} > 0$ [27]. Therefore, considering either \mathbf{s} or \mathbf{s}_p , the robustness domain w.r.t. modeling error is extremely large, i.e., $\widehat{R} \in]0, +\infty[$. From a practical point of view, a rough estimate of R will thus be sufficient.

B. Stability to Calibration Errors

In this section, the stability to calibration errors is presented using \mathbf{s} for perspective cameras and \mathbf{s}_p for paracatadioptric cameras. As usual, we consider that the desired, initial, and current images have been acquired with the same camera.

1) *Perspective Cameras:* Let us consider only calibration errors; then, (21) becomes

$$\dot{\mathbf{e}} = -\lambda \mathbf{L}_s \widehat{\mathbf{L}}_s^+ \mathbf{E} \mathbf{e} \quad (22)$$

with $\widehat{\mathbf{s}} = \mathbf{E}(\mathbf{s})$, $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$, and

$$\widehat{\mathbf{L}}_s^+ = \begin{bmatrix} -\frac{R\widehat{r}^2}{\widehat{r}^2 + R^2} (R^2 \widehat{\mathbf{s}} \widehat{\mathbf{s}}^\top + \mathbf{I}_3) \\ -\frac{R^2 \widehat{r}^2}{\widehat{r}^2 + R^2} [\widehat{\mathbf{s}}]_\times \end{bmatrix}$$

where $\widehat{r} = 1/\|\widehat{\mathbf{s}}\|$. The nonlinearity of (22) makes it complicated to analyze the GAS. The analysis is simplified by considering the local stability (i.e., for $\mathbf{s} = \mathbf{s}^*$) and by linearizing \mathbf{E} around $\mathbf{s}^* = (0, 0, s_z)$.

By differentiating (16) and using (20) (with $\xi = 0$ and $\varphi = 1$), we obtain [25]

$$\begin{cases} ds_x = -du_0 (p_x^2 + p_y^2) s_z / (p_x k) \\ ds_y = -dv_0 (p_x^2 + p_y^2) s_z / (p_y k) \\ ds_z = (p_x dp_x + p_y dp_y) s_z / k \end{cases}$$

with $k = p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}$, from which we deduce

$$\mathbf{E}^* = \begin{bmatrix} 1 & 0 & \alpha_{u_0} \\ 0 & 1 & \alpha_{v_0} \\ 0 & 0 & 1 + \alpha_{p_x p_y} \end{bmatrix}$$

with

$$\begin{cases} \alpha_{u_0} = \frac{-\Delta u_0 (p_x^2 + p_y^2)}{p_x k} \\ \alpha_{v_0} = \frac{-\Delta v_0 (p_x^2 + p_y^2)}{p_y k} \\ \alpha_{p_x p_y} = \frac{p_x \Delta p_x + p_y \Delta p_y}{k} \end{cases}$$

where $\Delta u_0 = \widehat{u}_0 - u_0$, $\Delta v_0 = \widehat{v}_0 - v_0$, $\Delta p_x = \widehat{p}_x - p_x$, and $\Delta p_y = \widehat{p}_y - p_y$. Therefore

$$\mathbf{L}_{\mathbf{s}^*} \widehat{\mathbf{L}}_{\mathbf{s}^*}^+ \mathbf{E}^* = \frac{\widehat{r}^{*2}}{\widehat{r}^{*2} + R^2} \left(R^2 \mathbf{E}^* \mathbf{s}^* \mathbf{s}^{*\top} \mathbf{E}^{*\top} + \mathbf{I}_3 \right) \mathbf{E}^* - \frac{R^2 \widehat{r}^{*2}}{\widehat{r}^{*2} + R^2} \left(\mathbf{E}^* \mathbf{s}^* \mathbf{s}^{*\top} - \mathbf{s}^{*\top} \mathbf{E}^* \mathbf{s}^* \mathbf{I}_3 \right) \mathbf{E}^*. \quad (23)$$

By computing the eigenvalues γ_i of $\mathbf{L}_{\mathbf{s}^*} \widehat{\mathbf{L}}_{\mathbf{s}^*}^+ \mathbf{E}^*$, it is possible to show that a sufficient condition to ensure $Re(\gamma_i) > 0$ is $\widehat{p}_x > 0$ and $\widehat{p}_y > 0$ [25]. This condition is also necessary since if $\widehat{p}_x \leq 0$ or $\widehat{p}_y \leq 0$, the system diverges.

2) *Paracatadioptric Cameras* ($\xi = 1$): On paracatadioptric systems, the intrinsic parameters are given by (u_0, v_0, f, f) , where f is related to the focal length of the mirror on the x -direction and y -direction, as it is generally assumed for such sensors [28]. Considering calibration errors using (18), we have

$$\dot{\mathbf{e}} = -\lambda \mathbf{L}_{\mathbf{s}_p} \widehat{\mathbf{L}}_{\mathbf{s}_p}^+ \mathbf{E} \mathbf{e} \quad (24)$$

with \mathbf{E} such that $\widehat{\mathbf{s}}_p = \mathbf{E}(\mathbf{s}_p) \mathbf{s}_p$, $\mathbf{e} = \mathbf{s}_p - \mathbf{s}_p^*$, and

$$\widehat{\mathbf{L}}_{\mathbf{s}_p}^+ = \frac{1}{\widehat{d}} \begin{bmatrix} -\widehat{c}R(\widehat{\rho}^2 R^2 + 1) & \widehat{\rho} \widehat{s} R & -\widehat{\rho} \widehat{c} \widehat{s}_z R^3 \\ -\widehat{s} R(\widehat{\rho}^2 R^2 + 1) & -\widehat{\rho} \widehat{c} R & -\widehat{\rho} \widehat{s} \widehat{s}_z R^3 \\ -\widehat{\rho} \widehat{s}_z R^3 & 0 & -R(1 + \widehat{s}_z^2 \widehat{r}^2) \\ \widehat{s} \widehat{s}_z R^2 & \widehat{\rho} \widehat{c} \widehat{s}_z R^2 & -\widehat{\rho} \widehat{s} R^2 \\ -\widehat{c} \widehat{s}_z R^2 & \widehat{\rho} \widehat{s} \widehat{s}_z R^2 & \widehat{\rho} \widehat{c} R^2 \\ 0 & -\widehat{\rho}^2 R^2 & 0 \end{bmatrix}$$

where $\widehat{d} = 1 + (\widehat{s}_z^2 + \widehat{\rho}^2) R^2$, $\widehat{c} = \cos \widehat{\theta}$, and $\widehat{s} = \sin \widehat{\theta}$.

Once again, due to the nonlinearity of the closed-loop system (24), we will only study the local stability at the desired position $\mathbf{s}_p = \mathbf{s}_p^* = (1, 0, s_z)$. From (17) and (20), we obtain

$$d\rho = \frac{-du_0 s_z}{f}, \quad d\theta = \frac{-dv_0 s_z}{f}, \quad ds_z = \frac{du_0 \rho}{2f} + \frac{df s_z}{f}$$

from which we deduce

$$\mathbf{E}^* = \begin{bmatrix} 1 & 0 & \alpha_{u_0} \\ 0 & 1 & \alpha_{v_0} \\ -\alpha_{u_0}/2 & 0 & 1 + \alpha_f \end{bmatrix}$$

with $\alpha_{u_0} = -\Delta u_0/f$, $\alpha_{v_0} = -\Delta v_0/f$, and $\alpha_f = \Delta f/f$. Let us consider the error only on the focal length f (i.e., $\Delta u_0 = 0$ and $\Delta v_0 = 0$) and the linearization of $\widehat{\mathbf{L}}_{\mathbf{s}_p}^+$ (i.e., $\widehat{c} = \cos \widehat{\theta}^* = 1$ and $\widehat{s} = \sin \widehat{\theta}^* = \widehat{\theta}^*$), the eigenvalues γ_i of $\mathbf{L}_{\mathbf{s}_p^*} \widehat{\mathbf{L}}_{\mathbf{s}_p^*}^+ \mathbf{E}^*$ can be obtained. They are given by $(fr^2(f + \Delta f) + 4n_{20}^{p*2}(1 + r^2))/(fr^2(f + \Delta f)^2 + 4n_{20}^{p*2}(1 + r^2))$ (double eigenvalue) and $(f + \Delta f)/f$. Not surprisingly, a sufficient condition for the local stability is $\widehat{f} > 0$. Of course, this is also a necessary condition.

V. CONTROL OF THE SIX DOFS

In the previous part of the paper, we have considered the image of a sphere that induces only three constraints. We are now interested in controlling the six DOFs, and for that, the CC target is first revisited, i.e., the original set of features is described, and

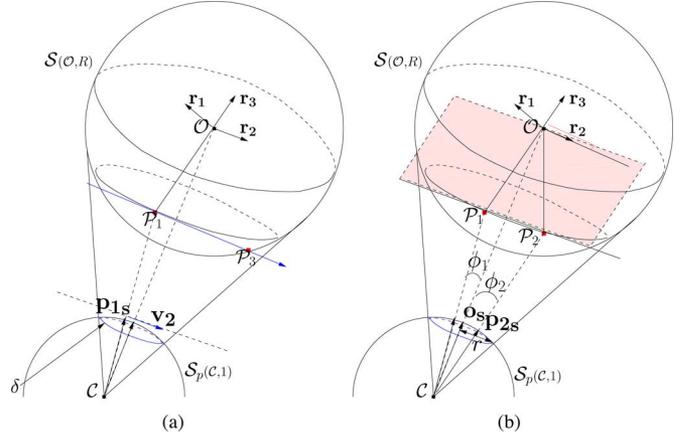


Fig. 5. Spherical projection of targets. (a) CC target. (b) Special sphere.

then a new minimal set of six features that improves the system behavior is designed. Second, it is proved that the new minimal set can also be used for visual servoing w.r.t. a special sphere.

A. CC Target

1) *Previous Approach*: As in [15], let us mark the sphere $\mathcal{S}(O,R)$ with a tangent vector $\mathbf{a}_{13} = \mathbf{P}_1 \mathbf{P}_3$ to a feature point \mathcal{P}_1 on its surface. We obtain the CC target, where $\mathcal{O} \mathcal{P}_1 \perp \mathcal{P}_1 \mathcal{P}_3$ [see Fig. 5(a)].

The set of seven features proposed in [15] is given by $\mathbf{s}_{cc} = (r, \mathbf{o}_s, \zeta)$, where ζ is equivalent to the $\theta \mathbf{u}$ representation (here, θ is the angle of rotation and \mathbf{u} is the unitary axis of rotation) of the rotation matrix $\mathbf{V} \mathbf{V}^{*-1}$, and matrix $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ is built as follows [15]:

$$\begin{cases} \mathbf{v}_1 = \mathbf{p}_{1s} = \pi_s(\mathbf{P}_1) \\ \mathbf{v}_2 = \frac{\Phi_{\mathbf{v}_1} \mathbf{a}_{13}}{\|\Phi_{\mathbf{v}_1} \mathbf{a}_{13}\|} = \frac{\Phi_{\mathbf{v}_1} (\pi_s(\mathbf{P}_3) - \pi_s(\mathbf{P}_1))}{\|\Phi_{\mathbf{v}_1} (\pi_s(\mathbf{P}_3) - \pi_s(\mathbf{P}_1))\|} \\ \mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 \end{cases} \quad (25)$$

with $\Phi_{\mathbf{v}_1} = \mathbf{I}_3 - \mathbf{v}_1 \mathbf{v}_1^\top$. More precisely, \mathbf{v}_2 is a tangent vector to $\mathcal{S}_p(C,1)$ at the point vector \mathbf{v}_1 . Matrix \mathbf{V}^* is built similarly using the desired position of \mathcal{P}_1 and \mathcal{P}_3 on the unit sphere.

The interaction matrix corresponding to \mathbf{s}_{cc} is given by [15]

$$\mathbf{L}_{\mathbf{s}_{cc}} = \begin{bmatrix} \frac{r^2}{R} \mathbf{o}_s^\top & \mathbf{0} \\ -\frac{r}{R} (\mathbf{I}_3 - \mathbf{o}_s \mathbf{o}_s^\top) & [\mathbf{o}_s]_\times \\ -\frac{1}{R\beta} (\alpha \mathbf{v}_1 \mathbf{v}_3^\top - \mathbf{v}_2 \mathbf{v}_3^\top + \mathbf{v}_3 \mathbf{v}_2^\top) & -\mathbf{I}_3 \end{bmatrix} \quad (26)$$

where α and β are scalar functions of r , \mathbf{o}_s , \mathbf{v}_1 , and \mathbf{v}_2 .

The key point of the original approach proposed in [15] is the definition of a diffeomorphism that maps a visible subset of the target pose $\mathbf{p} = ({}^c \mathbf{t}_o, {}^c \mathbf{R}_o)$ to the image features \mathbf{s}_{cc} . However, the set \mathbf{s}_{cc} is not minimal and its interaction matrix $\mathbf{L}_{\mathbf{s}_{cc}}$ presents a coupling between its orientation and translation terms (note the first term in the third row of $\mathbf{L}_{\mathbf{s}_{cc}}$). The following section presents a new set of features that overcomes the two problems mentioned earlier.

2) *New Approach*: We propose to use the minimal set of six features $\mathbf{s}_n = (\mathbf{s}, \theta\mathbf{u})$ in which \mathbf{s} has been defined by (7), and $\theta\mathbf{u}$ is computed from the relative orientation ${}^{c*}\mathbf{R}_c$ of the current camera frame w.r.t. the desired camera frame. Indeed, the orientation ${}^c\mathbf{R}_o = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$ of the CC target w.r.t. the camera can be recovered using the diffeomorphism proposed in [15]. The interaction matrix related to \mathbf{s}_n has the following form [29]:

$$\mathbf{L}_{\mathbf{s}_n} = \begin{bmatrix} -\frac{1}{R}\mathbf{I}_3 & [\mathbf{s}]_{\times} \\ \mathbf{0} & \mathbf{L}_{\omega} \end{bmatrix} \quad (27)$$

in which \mathbf{L}_{ω} is given by [8]

$$\mathbf{L}_{\omega} = \mathbf{I}_3 + \frac{\theta}{2}[\mathbf{u}]_{\times} + \left(1 - \frac{\text{sinc}\theta}{\text{sinc}^2(\theta/2)}\right) [\mathbf{u}]_{\times}^2 \quad (28)$$

where $\text{sinc}(x) = \sin x/x$. The orientation control using $\theta\mathbf{u}$ is completely decoupled from the translational motions, which was not the case when using ζ . In addition, $\mathbf{L}_{\mathbf{s}_n}$ is an upper block triangular square matrix. As will be shown in Section V-D, it allows to simplify the stability analysis. Finally, let us mention that features \mathbf{s}_{cc} and \mathbf{s}_n can easily be computed on any central catadioptric system using the image of \mathcal{P}_1 and \mathcal{P}_3 and the equations given in Section III-B.

B. Special Sphere

A special sphere is obtained by gluing two feature points \mathcal{P}_1 and \mathcal{P}_2 on its surface. Fig. 5(b) presents the spherical projection of a special sphere.

Once again, the three features \mathbf{s} that characterize the image of $\mathcal{S}(\mathcal{O}, R)$ can be used to control the camera position. The camera orientation can be controlled by using features $\theta\mathbf{u}$ that are selected from ${}^{c*}\mathbf{R}_c$. We now show how to determine ${}^c\mathbf{R}_o = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$.

Features vector $\mathcal{P}_1\mathcal{P}_2$ is defined such that

$$\mathbf{a}_{12} = \mathbf{P}_1\mathbf{P}_2 = \lambda_1\mathbf{r}_3 + \lambda_2\mathbf{r}_2 \quad (29)$$

where $(\lambda_1, \lambda_2) \in \mathbb{R}^2$. As shown in Fig. 5(b), the feature point \mathcal{P}_1 is such that [15]

$$\mathbf{r}_3 = \frac{1}{R}({}^c\mathbf{t}_o - \mathbf{P}_1) \quad (30)$$

where $(1/R)\mathbf{P}_1 = (\|\mathbf{P}_1\|/R)\mathbf{p}_{1s}$. It is possible to compute $\|\mathbf{P}_1\|/R$ by applying the cosine rule to the triangle $(\mathcal{C}, \mathcal{P}_1, \mathcal{O})$ in \mathcal{F}_c [15]. We obtain the second-degree equation in $\|\mathbf{P}_1\|$, which is given by

$$\|\mathbf{P}_1\|^2 + \|\mathbf{O}\|^2 - 2\|\mathbf{P}_1\|\|\mathbf{O}\|\cos\phi_1 = R^2 \quad (31)$$

where $\phi_1 = \mathbf{o}_s^\top \mathbf{p}_{1s}$. The two solutions of (31) are given by

$$\|\mathbf{P}_1\| = \frac{R}{r} \left(\cos\phi_1 + \sigma\sqrt{r^2 - \sin^2\phi_1} \right) \quad (32)$$

where $\sigma = \pm 1$, from which it is easy to deduce $\|\mathbf{P}_1\|/R$. The solution with $\sigma = -1$ is chosen due to the visibility condition defined in [15].

At this point, we can conclude that \mathbf{r}_3 can be computed from the image of the special sphere. Indeed, since ${}^c\mathbf{t}_o/R$ is equal to

\mathbf{s} , we obtain from (30) and (32)

$$\mathbf{r}_3 = \mathbf{s} - \frac{1}{r}K_1\mathbf{p}_{1s} \quad (33)$$

where $K_1 = \cos\phi_1 - \sqrt{r^2 - \sin^2\phi_1}$. Similar to $\|\mathbf{P}_1\|/R$, we obtain

$$\frac{\|\mathbf{P}_2\|}{R} = \frac{K_2}{r} \quad (34)$$

where $K_2 = \cos\phi_2 - \sqrt{r^2 - \sin^2\phi_2}$, with $\phi_2 = \mathbf{o}_s^\top \mathbf{p}_{2s}$. It is, therefore, possible to compute

$$\frac{1}{R}\mathbf{a}_{12} = \frac{1}{R}(\mathbf{P}_2 - \mathbf{P}_1) \quad (35)$$

which can be expressed using (32) and (34) as

$$\frac{1}{R}\mathbf{a}_{12} = \frac{K_1}{r}\mathbf{p}_{1s} - \frac{K_2}{r}\mathbf{p}_{2s}. \quad (36)$$

From (36) and (29), we obtain

$$\frac{\lambda_2}{R}\mathbf{r}_2 = \left(\frac{1}{R}\mathbf{a}_{12} - \left(\left(\frac{1}{R}\mathbf{a}_{12} \right)^\top \mathbf{r}_3 \right) \mathbf{r}_3 \right) \quad (37)$$

from which we easily deduce \mathbf{r}_2 . Finally, we have $\mathbf{r}_1 = \mathbf{r}_2 \times \mathbf{r}_3$ that gives us ${}^c\mathbf{R}_o$. The desired orientation matrix ${}^{c*}\mathbf{R}_o$ can be obtained using the same calculation steps as for ${}^c\mathbf{R}_o$.

To conclude, it is possible to compute the partial pose $({}^c\mathbf{t}_o/R, \theta\mathbf{u})$ of the special sphere and, therefore, to compute the set of six features $\mathbf{s}_n = (\mathbf{s}, \theta\mathbf{u})$ from the image of a special sphere.

C. Discussion

From the set $\mathbf{s}_n = (\mathbf{s}, \theta\mathbf{u})$, we can design another set $(\mathbf{s}_t, \theta\mathbf{u})$, which fully decouples the control of the translation from the camera rotational motions. Indeed, if we use

$$\mathbf{s}_t = (\mathbf{s}^* - {}^{c*}\mathbf{R}_c\mathbf{s}) \quad (38)$$

we obtain

$$\mathbf{L} = \begin{bmatrix} \frac{1}{R}{}^{c*}\mathbf{R}_c & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{\omega} \end{bmatrix} \quad (39)$$

Let ${}^{c*}\mathbf{t}_c$ be the relative position of the current camera frame w.r.t. the desired camera frame. We can note that $\mathbf{s}_t = (1/R){}^{c*}\mathbf{t}_c$, which corresponds to a classical 3-D visual servoing method (up to a scale factor) where there is no control of the target in the image [6]. In addition, there is no more a linear relationship between the visual features for translation \mathbf{s}_t and the camera translational velocities \mathbf{v} . Therefore, we prefer the set $\mathbf{s}_n = (\mathbf{s}, \theta\mathbf{u})$, which is linearly linked w.r.t. \mathbf{v}_c [see (27)] while enabling a partial control of the target in the image, as discussed in Section III-B.

Finally, it is important to note that when either \mathbf{P}_1 , \mathbf{P}_3 , or \mathbf{P}_2 is occluded or out of the camera FOV, feature $\theta\mathbf{u}$ can no longer be computed. It is also the case if either $\mathbf{p}_{1s} = \mathbf{p}_{3s}$ or $\mathbf{p}_{1s} = \mathbf{p}_{2s}$ [see (25) and (36)].

D. Stability Analysis of Modeling Error

As before, we use the classical control law given by (19). We suppose that all special marks of each target are visible and assume that $\mathcal{OP}_1 \perp \mathcal{P}_1\mathcal{P}_3$ for the CC target. In this case, the new set \mathbf{s}_n can be computed and the interaction matrix $\mathbf{L}_{\mathbf{s}_n}$ is of full rank six. In the ideal case, where the radius R is known, it is clear that the classical control law (19) is GAS since the closed-loop system equation is given by $\dot{\mathbf{e}} = -\lambda\mathbf{e}$. Similarly, as for \mathbf{s} (see Section IV-A), incorporating an error on the estimation of R does not affect the set \mathbf{s}_n [which is directly computed in the image from \mathbf{o}_s , r , \mathbf{p}_{1_s} , and \mathbf{p}_{2_s} , as shown by (13)–(15), (33), and (37)], which leads to the following closed-loop system equation:

$$\dot{\mathbf{e}} = -\lambda\mathbf{L}_{\mathbf{s}_n}\widehat{\mathbf{L}}_{\mathbf{s}_n}^{-1}\mathbf{e} \quad (40)$$

with $\mathbf{e} = (\mathbf{s} - \mathbf{s}^*, \theta\mathbf{u})$, and

$$\widehat{\mathbf{L}}_{\mathbf{s}_n} = \begin{bmatrix} -\frac{1}{\widehat{R}}\mathbf{I}_3 & [\mathbf{s}]_{\times} \\ 0 & \mathbf{L}_{\omega} \end{bmatrix}.$$

The local asymptotic stability is obtained if and only if the eigenvalues of the constant matrix $\mathbf{L}_{\mathbf{s}_n^*}\widehat{\mathbf{L}}_{\mathbf{s}_n^*}^{-1}$ are positive. Therefore, we have

$$\mathbf{L}_{\mathbf{s}_n^*}\widehat{\mathbf{L}}_{\mathbf{s}_n^*}^{-1} = \begin{bmatrix} \frac{\widehat{R}}{R}\mathbf{I}_3 & \left(-\frac{\widehat{R}}{R}[\mathbf{s}^*]_{\times} + [\mathbf{s}^*]_{\times}\right) \\ 0 & \mathbf{I}_3 \end{bmatrix} \quad (41)$$

from which we easily deduce the triple eigenvalues \widehat{R}/R and 1. Finally, we get that the system is locally asymptotically stable, even in the presence of modeling errors if and only if $\widehat{R} > 0$.

Even if the robustness domain is theoretically limited near the desired pose, the convergence domain is very large in practice, as will be shown in Section VI-D2.

VI. RESULTS

In this section, we begin with the results concerning a simple spherical target: First, the new general set \mathbf{s} is validated on a real robotic system equipped with a perspective camera. Second, we validate this set using a paracatadioptric system. Third, we show in simulation that, for some configurations, this general set draws a highly undesirable straight-line trajectory in the image plane of paracatadioptric cameras. In addition, we validate in simulation the new set \mathbf{s}_p specific to such cameras. Fourth, we prove, in both random simulations and experimental results on a perspective camera, that the new set designed to control the six DOFs using the CC target is better than the set proposed in [15]. Finally, results obtained for controlling the six DOFs using the special sphere are presented with perspective and fish-eye cameras.

A. Experimental Results With a Perspective Camera

These experiments have been carried out with a conventional perspective camera mounted on the end-effector of a 6-DOF robotic system. The experiments consist of a centering task. The target is a 2 cm radius ping-pong ball. Using such a simple

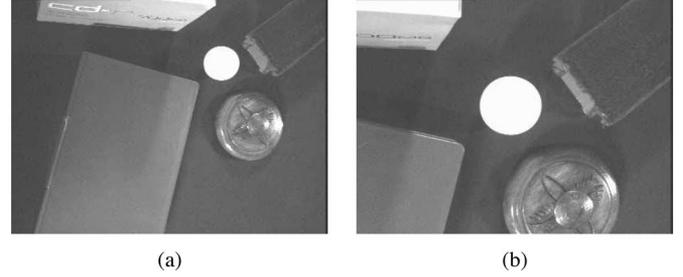


Fig. 6. (a) Initial image. (b) Final image.

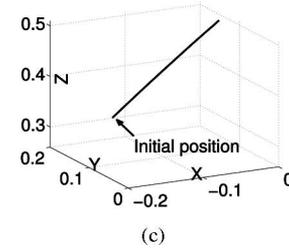
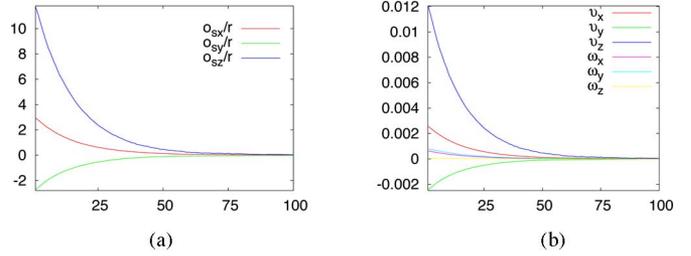


Fig. 7. Ideal case. (a) \mathbf{s} error. (b) Computed camera velocities (in meters per second and in radians per second). (c) Camera trajectory (in meters) in the Cartesian space.

object allows us to easily compute the ellipse moments $\boldsymbol{\mu}$ at video rate with no image processing problems. The variable \mathbf{s}^* has been selected as $(0, 0, \tau^*)$, where τ^* allows us to define the depth between the camera and the ball. Fig. 6 shows the initial and the final images used for each experiment. For all the experiments, the same gain $\lambda = 0.5$ has been used.

1) *Ideal Case:* In order to validate the selected visual features, we first consider the ideal case where we have correct camera calibration values and $\widehat{R} = R$. Indeed, when $\widehat{R} = R$, we should have a perfect system behavior, since $\mathbf{L}_s\widehat{\mathbf{L}}_s^+ = \mathbf{I}_3$. As expected, a pure exponential decrease of the error on the visual features can be observed in Fig. 7(a), and the camera velocities also decrease exponentially [see Fig. 7(b)]. Finally, the camera trajectory in the Cartesian space is almost a pure straight line in this experiment [see Fig. 7(c)], since the rotational velocities are small [see Fig. 7(b)], due to the fact that the camera is relatively far away from the target. We can also note that no noise appears at all on the presented plots, which will be true for all the experiments presented in this paper. This is due in part to the well-contrasted images, but probably also to the use of image moments to compute the features.

2) *Modeling Error:* The GAS to modeling error has been proved in this paper. We have validated this proof with two experiments. The results in the case where $\widehat{R} = 10R$ and $\widehat{R} = 0.1R$

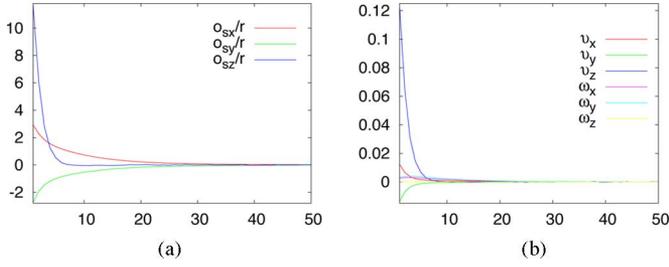


Fig. 8. Modeling error $\widehat{R} = 10R$. (a) s error. (b) Computed camera velocities (in meters per second and in radians per second).

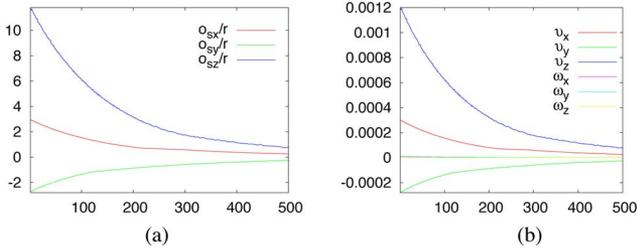


Fig. 9. Modeling error $\widehat{R} = 0.1R$. (a) s error. (b) Computed camera velocities (in meters per second and in radians per second).

are depicted, respectively, in Figs. 8 and 9. We can note that in both cases the system converges. Fig. 8(b) shows a high speed on the translational velocities, while Fig. 9(b) shows a low speed on the same components. Indeed, from (19) [where $\widehat{\mathbf{L}}_s^+$ is defined in (21)], we easily obtain

$$\begin{cases} \mathbf{v} = \lambda \frac{\widehat{R}r^2}{r^2 + \widehat{R}^2} (\widehat{R}^2 \mathbf{s} \mathbf{s}^\top + \mathbf{I}_3) (\mathbf{s} - \mathbf{s}^*) \\ \boldsymbol{\omega} = \lambda \frac{\widehat{R}^2 r^2}{r^2 + \widehat{R}^2} [\mathbf{s}]_{\times} (\mathbf{s} - \mathbf{s}^*). \end{cases} \quad (42)$$

When \widehat{R} tends to $+\infty$, (42) tends to

$$\begin{cases} \mathbf{v} = \infty \\ \boldsymbol{\omega} = \lambda r^2 [\mathbf{s}]_{\times} (\mathbf{s} - \mathbf{s}^*) \end{cases}$$

which explains the fast convergence observed in Fig. 8 (50 iterations) when $\widehat{R} = 10R$. When \widehat{R} tends to 0, from (42), we can get that \mathbf{v} and $\boldsymbol{\omega}$ tend to 0. This explains the slow convergence observed in Fig. 9 (500 iterations) when $\widehat{R} = 0.1R$. In practice, the behavior could be easily improved by using a higher gain λ to deal with under-approximation of \widehat{R} and by saturating \mathbf{v}_c when needed (to deal with an over-approximation of \widehat{R}).

3) *Calibration Errors*: Finally, we validate the results obtained for the stability analysis w.r.t. calibration errors. The necessary and sufficient condition $\widehat{p}_x > 0$ and $\widehat{p}_y > 0$ is validated by introducing the following errors on the camera intrinsic parameters: 35% p_x , -47% p_y , -25% u_0 , and 57% v_0 . The obtained results are shown in Fig. 10. Once again, the system converges (while only local stability has been demonstrated). The rapid decrease of the error on s_y compared with the other features is simply due to the biggest error introduced on p_y . Of course, we have observed that if $\widehat{p}_x \leq 0$ or $\widehat{p}_y \leq 0$, the system diverges, but this is not surprising.

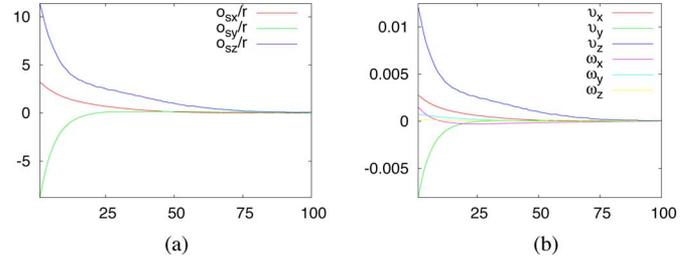


Fig. 10. Calibration errors. (a) s error. (b) Computed camera velocities (in meters per second and in radians per second).

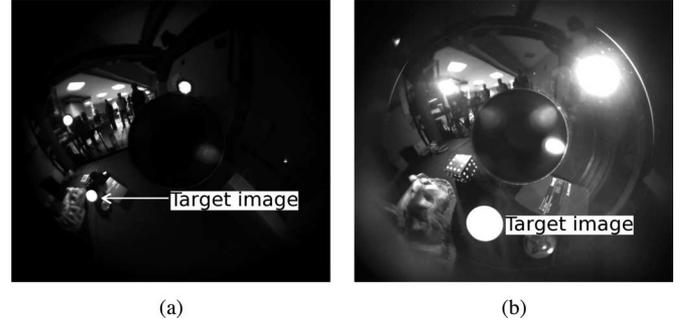


Fig. 11. (a) Initial image. (b) Final image.

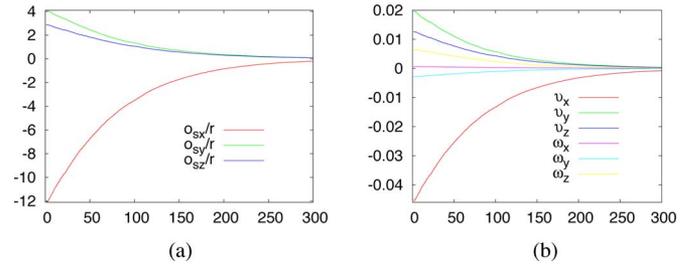


Fig. 12. Ideal case. (a) s error. (b) Computed camera velocities (in meters per second and in radians per second).

B. Experimental Results With Paracatadioptric Cameras

The following experiments have been carried out with a paracatadioptric camera mounted on the end-effector of a 6-DOF robotic system. This type of camera corresponds to a central catadioptric system with $\xi = 1$. The target is a 4 cm radius polystyrene white ball. Fig. 11 depicts the initial and the desired images used for each experiment. For all the experiments, the same gain $\lambda = 0.1$ has been used.

1) *Ideal Case*: It corresponds to the case where we have $\widehat{R} = R$ and exact camera calibration values. As expected, a pure exponential decrease of the error on the visual features can be observed in Fig. 12(a), and the camera velocities also decrease exponentially [see Fig. 12(b)]. The first part of the attached video illustrates this case.

2) *Modeling Error*: The stability w.r.t. modeling error using \mathbf{s} has been formally proved. For paracatadioptric systems, we have validated this proof with one experiment. The results in the case where $\widehat{R} = 5R$ is depicted in Fig. 13. The same comments as in Section VI-A2 can be made.

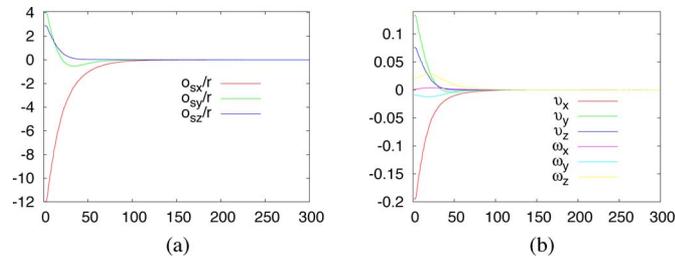


Fig. 13. Modeling error $\widehat{R} = 5R$. (a) s error. (b) Computed camera velocities (in meters per second and in radians per second).

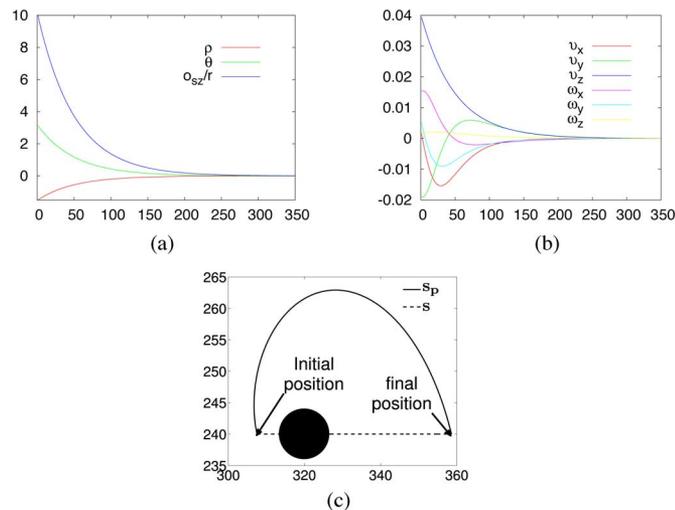


Fig. 14. Ideal case. (a) s_p error. (b) Computed camera velocities (in meters per second and in radians per second). (c) Image-plane trajectories (in meters) of the center of gravity of the sphere image.

C. Simulation Results With Paracatadioptric Cameras

In this section, it is shown that for the general features s , the motion in the image is not always suitable with paracatadioptric cameras, particularly when the initial position and desired position in the image space are each other's mirror image. In addition, the features s_p specific to paracatadioptric systems are validated.

We consider the case where we have $\widehat{R} = R$ and correct camera calibration values. In this case, we have $\mathbf{L}_{s_p} \widehat{\mathbf{L}}_{s_p}^+ = \mathbf{I}_3$, and thus a perfect system behavior. Fig. 14(a) plots the features error trajectory, while Fig. 14(b) shows the camera velocities. The image-plane trajectory of the center of gravity of the sphere image is drawn in Fig. 14(c). In this figure, we can see that the general features s generate a straight line passing through the center of the image, which means that in case of a real sensor, the target would get lost in the dead angle. As can be seen in Fig. 14(c), using the new features s_p leads to a circular trajectory, as expected, which means that with a real sensor, it is possible to constrain ρ to avoid the dead angle.

Other simulation results in the presence of modeling and calibration errors can be found in [27].

TABLE I
EMPIRICAL COMPARISON

	s_{cc}	s_n
Number of success	51	100
Camera average length path for the 51 success	1.05 m	0.56 m

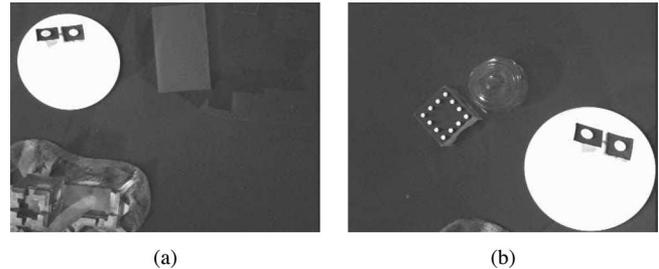


Fig. 15. CC target. (a) Initial image. (b) Desired image.

D. Control of the Six DOFs Using Special Targets

In this section, we validate the choice of visual features proposed in Section V through simulations and experimental results obtained on a 6-DOF robotic system equipped with either perspective or fish-eye cameras. First, we use the CC target to compare the two sets $s_{cc} = (r, \mathbf{o}_s, \zeta)$ and $s_n = ((1/r)\mathbf{o}_s, \theta\mathbf{u})$. Second, we use a special sphere and the set s_n , whereby the robustness of the classical control law (19) w.r.t. both modeling and calibration errors is validated on a perspective camera. Third, we use a nonspherical decoration balloon to validate the robustness of the control law to shape-modeling error using s_n . Finally, we use a special soccer ball and a fish-eye camera to validate the robustness w.r.t. camera modeling errors using the set s_n .

1) CC Target—Previous Set Versus New Set:

a) *Simulation results:* Here, the two sets are empirically compared. The method of comparison consists of analyzing the system behavior between a fixed desired pose and 100 randomly chosen initial poses. In some cases, the initial pose was near the visibility boundaries.

Table I summarizes the result. The simple control law (19) has failed almost half the time using s_{cc} although it has succeeded all the time using s_n . The high percentage failure using s_{cc} is due to the coupling between the orientation control and the translation control (see the first term in the third row of $\mathbf{L}_{s_{cc}}$) that leads to a large side motion of the camera (see Fig. 16(a) as a typical example) and, thus, causes feature point \mathcal{P}_1 to leave the camera FOV. For the 51 successful cases using s_{cc} , the average length path of the camera using s_{cc} is 88%, which is greater than the mean 0.56 m obtained using s_n . Finally, this result also shows an empirically large convergence domain of control law (19) using s_n .

b) *Experimental results:* The CC target is a white soccer ball of 9.5 cm radius. Fig. 15 depicts the initial and desired images used for each experiment. The same gain $\lambda = 0.1$ has been used.

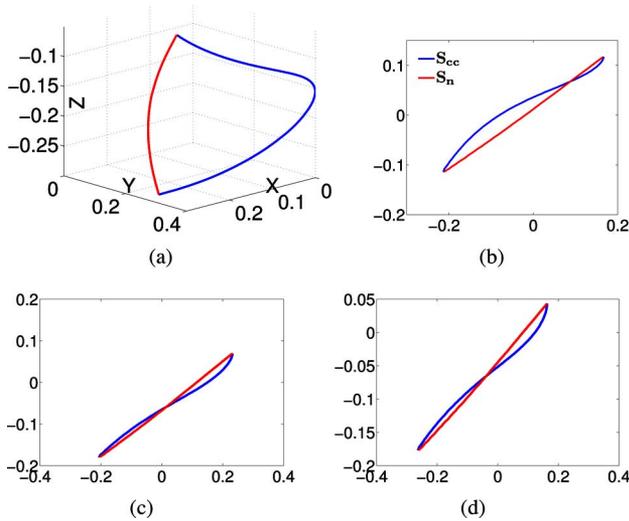


Fig. 16. Six DOFs control using the CC target. (a) Camera trajectory (in meters) in the Cartesian space. (b) Dome hat summit \mathbf{o}_s trajectory (in meters) in the image. (c) Special mark \mathbf{P}_1 trajectory (in meters) in the image. (d) Special mark \mathbf{P}_3 trajectory (in meters) in the image.

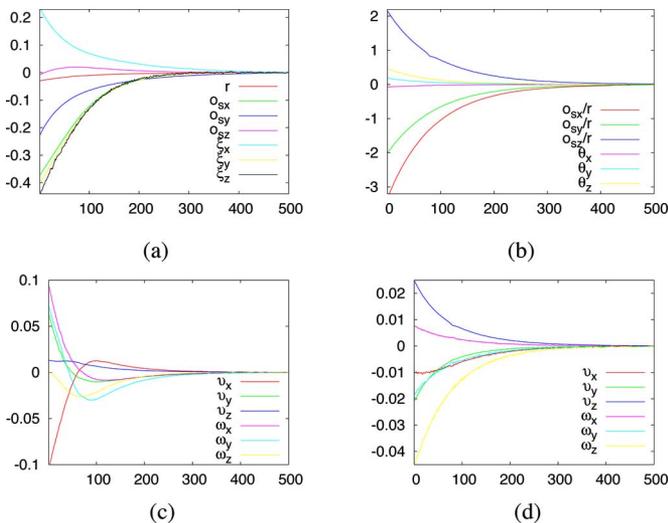


Fig. 17. Ideal case. (a) s_{cc} error. (b) s_n error. (c) and (d) Computed camera velocities (in meters per second and in radians per second) using s_{cc} and s_n .

We first consider the exact value of the sphere radius and correct camera calibration values. In this case, the camera and feature trajectories are shown in Fig. 16. As expected, we can see that, either in the image plane or in the Cartesian space, the trajectories using s_n are shorter than the one using s_{cc} . This result confirms those obtained in simulation.

In addition, and as expected, the new set s_n also provides an exponential decrease of the camera rotational velocities with fewer overshoots in the translational velocities [compare Fig. 17(c) and (d)]. The first shot of the second part of the attached video illustrates the differences between s_{cc} and s_n .

2) *Special Sphere—Modeling and Calibration Errors:* Using a special sphere, we verify the stability of the classical control law using the new set s_n w.r.t. to both modeling and calibration errors. The target is a white soccer ball of radius 9.5 cm

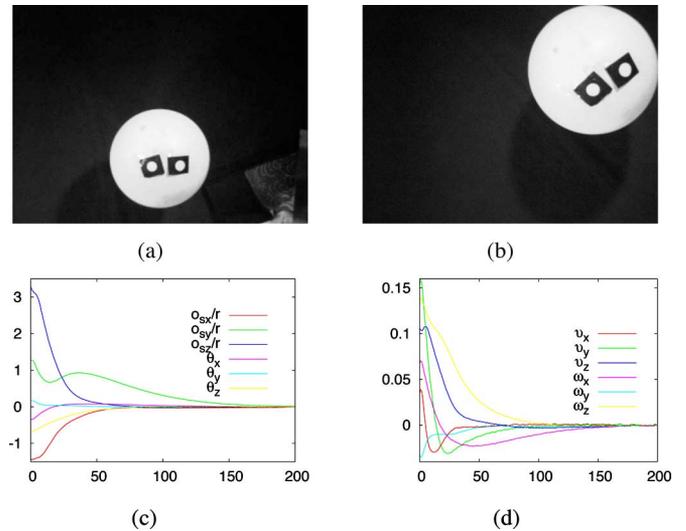


Fig. 18. Robustness w.r.t. both modeling and calibration errors. (a) Initial image. (b) Desired image. (c) $(s, \theta u)$ error. (d) Computed camera velocities (in meters per second and in radians per second).

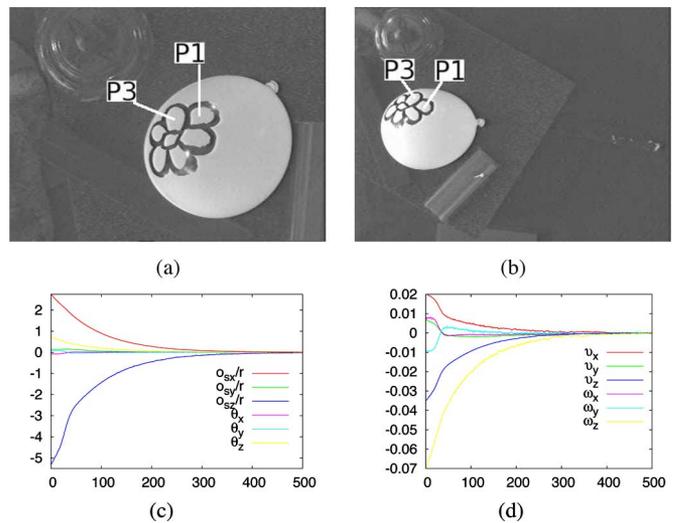


Fig. 19. Application to a decoration balloon. (a) Initial image. (b) Desired image. (c) $(s, \theta u)$ error. (d) Computed camera velocities (in meters per second and in radians per second).

marked with points \mathbf{P}_1 and \mathbf{P}_2 on its surface. Here, we consider the case where $\hat{R} = 2R$ and introduce the following errors on the camera intrinsic parameters: $-25\%p_x$, $25\%p_y$, $-47\%u_0$, and $20\%v_0$. Fig. 18 plots the obtained results, where we can see that the system converges.

3) *Application to a Decoration Balloon:* Here, we validate the robustness of the control law to strong modeling errors. The target is a white nonspherical decoration balloon marked with a flower picture in black. Fig. 19(a) and (b) shows the initial and desired images of the target. Taking advantage of the stability to modeling error, we used an approximated radius of 6.5 cm. In addition, the special marks \mathbf{P}_1 and \mathbf{P}_3 have been selected as the center of gravity of two leaves on the flower picture. Fig. 19(c) and (d) plots the results where we can see that

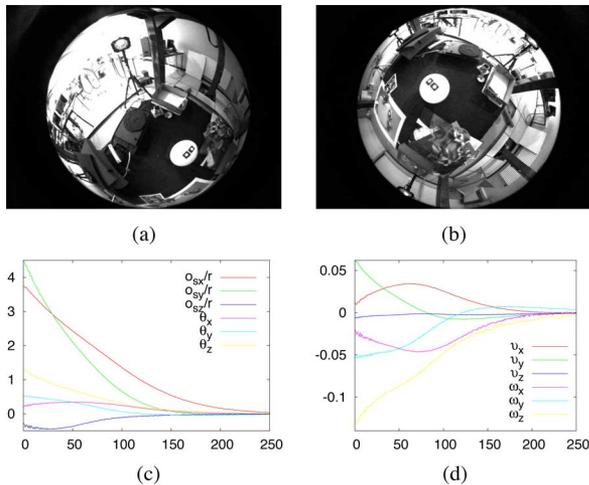


Fig. 20. New approach applied to a special soccer ball using a fish-eye camera. (a) Initial image. (b) Desired image. (c) (s, θ) error. (d) Computed camera velocities (in meters per second and in radians per second).

system converges. The last shot of the second part of the attached video illustrates this application.

4) *Application to a Special Soccer Ball:* We finally validate the robustness of the control w.r.t. camera modeling errors using the new set s_n . The target is a white soccer ball of radius 9.5 cm marked with points P_1 and P_2 on its surface. We use a fish-eye camera [see Fig. 20(a) and (b)]. Although this omnidirectional sensor is not a central catadioptric system, it is possible to map a fish-eye image on a sphere using an adequate projection model [30], and we can see that the system converges [see Fig. 20(c) and (d)].

VII. DISCUSSION

This section tries to position our method w.r.t. classical approaches. On one hand, it is clear that it is a pure 2-D visual servoing, since the visual features are directly computed in the single current image from image moments and the image of two points (as well as, to be complete, from the camera intrinsic parameters). On the other hand, it is also clear that our method is almost a pure 3-D visual servoing, since the visual features represent nothing but the pose up to a constant scalar factor as for the translation, which is exactly similar to a scheme presented in [31] called hybrid visual servoing. Finally, our method could easily become a pure 3-D visual servoing by just multiplying the first three components of the features by the value of the radius of the sphere (or an approximated value of this radius). The control law output would be exactly the same, since the radius would disappear from the interaction matrix. We are thus in the same situation as the method proposed in [32] that is based on straight lines and simultaneously gathers 2-D and 3-D information. From this short analysis, we conclude that the standard classification of visual servoing is clearly outdated.

VIII. CONCLUSION

In this paper, improvements on modeling features for visual servoing using a spherical projection model have been proposed.

Three different spherical targets have been considered: a simple sphere, the CC target, and a special sphere. For each of these targets, a new minimal set of visual features has been designed, which are decoupled and nearly linear linked to the sensor velocities; they can be computed on any central catadioptric image plane using 2-D moments. In addition, these sets draw adequate trajectories both in the image and in the Cartesian space. For each newly proposed set of features, a classical control law has been analytically proved to be asymptotically stable w.r.t. modeling error. The effects of camera calibration errors have been analyzed for spheres using perspective and paracatadioptric sensors. The obtained theoretical results have been confirmed through simulation and experimental results with either perspective, paracatadioptric, or fish-eye cameras mounted on a 6-DOF robotic system. As future works, it would be interesting to explicitly take into account the visibility of the marks of the special sphere, as well as joint limits constraints, to define a global convergence domain. For this, navigation functions [15], [16] or path-planning techniques [17] could be designed.

ACKNOWLEDGMENT

The authors would like to thank N. J. Cowan for fruitful discussions and F. Janabi-Sharifi, S. Hutchinson, and the reviewers for their comments on this research and earlier drafts of this paper.

REFERENCES

- [1] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 3, pp. 651–670, Oct. 1996.
- [2] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Confluence of Vision and Control* (Lecture Notes In Control and Information Sciences, no. 237), D. Kriegman, G. Hager, and A. S. Morse, Eds. New York: Springer-Verlag, 1998, pp. 66–78.
- [3] W. Wilson, C. Hulls, and G. Bell, "Relative end-effector control using cartesian position-based visual servoing," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 684–696, Oct. 1996.
- [4] D. Dementhon and L. Davis, "Model-based object pose in 25 lines of code," *Int. J. Comput. Vis.*, vol. 15, no. 1/2, pp. 123–141, Jun. 1995.
- [5] C.-P. Lu, G. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 6, pp. 610–622, Nov. 2000.
- [6] L. Deng, W. Wilson, and F. Janabi-Sharifi, "Dynamic performance of the position-based visual servoing method in the cartesian and image spaces," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Las Vegas, Nevada, Oct. 2003, vol. 1, pp. 510–515.
- [7] P. Martinet, J. Gallice, and D. Khadraoui, "Vision based control law using 3d visual features," in *Proc. World Autom. Congr., Robot. Manuf. Syst.*, vol. 3, Montpellier, France, May 1996, pp. 497–502.
- [8] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 d visual servoing," *IEEE Trans. Robot. Autom.*, vol. 15, no. 2, pp. 238–250, Apr. 1999.
- [9] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," presented at the 3rd Int. Symp. Exp. Robot., Kyoto, Japan, Oct. 1993.
- [10] P. Corke and S. Hutchinson, "A new partitioned approach to image-based visual servo control," *IEEE Trans. Robot. Autom.*, vol. 17, no. 4, pp. 507–515, Aug. 2001.
- [11] M. Iwatsuki and N. Okiyama, "A new formulation for visual servoing based on cylindrical coordinate system," *IEEE Trans. Robot.*, vol. 21, no. 2, pp. 266–273, Apr. 2005.
- [12] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.

- [13] D. Burschka and G. Hager, "Vision-based control of mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, Seoul, Korea, May 2001, pp. 1707–1713.
- [14] T. Hamel and R. Mahony, "Visual servoing of an under actuated dynamic rigid-body system: An image-based approach," *IEEE Trans. Robot. Autom.*, vol. 18, no. 2, pp. 187–198, Apr. 2002.
- [15] N. Cowan and D. Chang, "Geometric visual servoing," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1128–1138, Dec. 2005.
- [16] N. Cowan, J. Weingarten, and D. Koditschek, "Visual servoing via navigation functions," *IEEE Trans. Robot. Autom.*, vol. 18, no. 4, pp. 521–533, Aug. 2002.
- [17] Y. Mezouar and F. Chaumette, "Path planning for robust image-based control," *IEEE Trans. Robot. Autom.*, vol. 18, no. 4, pp. 534–549, Aug. 2002.
- [18] D. Miller and A. Wright, "Autonomous spacecraft docking using multi-color targets," presented at the 6th Top. Meet. Robot., Monterey, CA, Feb. 1995.
- [19] J. Barreto, F. Martin, and R. Horaud, "Visual servoing/tracking using central catadioptric images," in *Proc. Int. Symp. Exp. Robot.*, Ischia, Italy, Jul. 2002, pp. 245–254.
- [20] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet, "2/1/2 visual servoing with central catadioptric cameras," in *Proc. IEEE Int. Conf. Robot. Autom.*, Barcelona, Spain, Apr. 2005, pp. 3572–3577.
- [21] F. Chaumette, P. Rives, and B. Espiau, "Classification and realization of the different vision-based tasks," in *Visual Servoing*, vol. 7, K. Hashimoto, Ed. Singapore: World Sci. Ser. Robot. Autom. Syst., 1993, pp. 199–228.
- [22] R. T. Fomena and F. Chaumette, "Visual servoing from spheres using a spherical projection model," in *Proc. IEEE ICRA*, Rome, Italy, Apr. 2007, pp. 2080–2085.
- [23] S. Baker and S. Nayar, "A theory of catadioptric image formation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jan. 1998, pp. 35–42.
- [24] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *Proc. Eur. Conf. Comput. Vis.*, 2000, vol. 29, pp. 159–179.
- [25] R. T. Fomena, "Asservissement visuel par projection sphérique," Ph.D. dissertation, Univ. Rennes 1, Rennes, France, Nov. 2008.
- [26] M. Born and E. Wolf, *Principles of Optics*. New York: Pergamon, 1965.
- [27] R. T. Fomena and F. Chaumette, "Visual servoing from spheres with paracatadioptric cameras," presented at the 13th Int. Conf. Adv. Robot., Jeju, Korea, Aug. 2007.
- [28] B. Vanderportaele, M. Cattoen, P. Marthon, and P. Gurdjos, "A new linear calibration method for paracatadioptric cameras," in *Proc. ICPR*, 2006, pp. 647–651.
- [29] R. T. Fomena and F. Chaumette, "Visual servoing from two special compounds of features using a spherical projection model," in *Proc. IEEE/RSI IROS*, Nice, France, Sep. 2008, pp. 3040–3045.
- [30] J. Kannala and S. Brandt, "A generic camera model and calibration method for conventional, wide angle and fish-eye lenses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1335–1340, Aug. 2006.
- [31] E. Malis and F. Chaumette, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," *IEEE Trans. Robot. Autom.*, vol. 18, no. 2, pp. 176–186, Apr. 2002.
- [32] N. Andreff, B. Espiau, and R. Horaud, "Visual servoing from lines," *Int. J. Robot. Res.*, vol. 21, pp. 679–700, Aug. 2002.



Romeo Tatsambon Fomena (M'09) received the Bachelor's degree in mathematics and computer science from the University of Ngaoundéré, Ngaoundéré, Cameroon, in 2002 and the Master's and the Ph.D. degrees in computer science from the Université de Rennes 1, Rennes, France, in 2005 and 2008, respectively.

For one month, he was with the University of Illinois at Urbana-Champaign in the framework of a collaborative work between the Beckman Institute and INRIA in 2008. He is currently a Teaching Associate with the Université de Rennes 1, where he is also a Research Associate with IRISA in the Lagadic Group. His current research interests include robotics, computer vision, and omnidirectional vision, particularly visual servoing using omnidirectional vision systems.



François Chaumette (M'98) received the Graduate degree from École Nationale Supérieure de Mécanique, Nantes, France, in 1987 and the Ph.D. degree in computer science from the University of Rennes 1, Rennes, France, in 1990.

Since 1990, he has been with INRIA Rennes-Bretagne Atlantique, IRISA, where he is the Head of the Lagadic Group. His current research interests include robotics and computer vision, especially visual servoing and active perception.

Dr. Chaumette was an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS from 2001 to 2005 and is currently on the Editorial Board of the *International Journal of Robotics Research*. He was the recipient of the AFCET/CNRS Prize for the Best French Thesis on Automatic Control in 1991 and the King-Sun Fu Memorial Best IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION Paper Award, with E. Malis, in 2002.