Visual servoing from spheres using a spherical projection model

R. Tatsambon Fomena, François Chaumette

Visual Servoing from Spheres using a Spherical Projection Model

Romeo Tatsambon Fomena and François Chaumette

Abstract— **In this paper, we investigate the use of a spherical projection model to search for optimal visual features for visual servoing. A new minimal set of three visual features is proposed for visual servoing from spheres using any central catadioptric system. Using this set of features, a classical control method is proved to be globally stable even in the presence of modeling error and locally stable to calibration errors on perspective cameras. Using this type of cameras, experimental results are presented and validate the proposed theoretical results.**

I. INTRODUCTION

Visual servoing consists of controlling the motion of a dynamic system using data provided by a vision sensor [1], [2]. A vision sensor provides a large spectrum of potential visual features. However, the use of some visual features may lead to stability problems if the displacement that the robot has to realize is very large [3]. Therefore, there is a need to model optimal visual features for visual servoing. The optimality criteria being: local and -as far as possible- global stability of the system, robustness to calibration and to modeling errors, none singularity nor local mimima, satisfactory motion of the system and of the measures in the image, and finally linear link and maximal decoupling between the visual features and the degrees of freedom (DOFs) taken into account.

Several approaches have been proposed to try to reach an optimal system behaviour using only 2D data (by lack of space, we do not recall here the properties of pose-based visual servoing and 2 1/2 D visual servoing). A visual feature proportionally linked to the depth of the observed object has been designed for a good system behaviour in the *z*-axis [4]. Similarly, for 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 [5]. The cylindrical coordinates of points can also be used to obtain a nice system motion in the *z* direction [6]. A generic and intuitive representation of the image of a target can be obtained using 2D moments. Recently, moment invariants theory have been used to determine specific combinations of 2D moments whose interaction with the system presents linear and decoupling properties when planar objects are considered [7].

All the above studies use a perspective projection model which is not always suitable to control a system. Indeed, for certain under-actuated systems, the use of a spherical projection model for visual servoing is suitable since this projection of an image point conserves the passivity property in the image space [8], this passivity property being important to control such systems.

Although much work on modeling issues has been done to date, very little is concerned with the use of a spherical projection model. This paper is concerned with the use of a spherical projection model to search for optimal visual features for visual servoing from spheres. Another high motivation to use this projection model is its simplicity compared to the complex equations generated by an omnidirectional projection model. We will see that using a spherical projection model, determining optimal features is quite easy and intuitive, while it would have been very difficult to obtain the same results from perspective or omnidirectional projection models.

In the next section, the spherical projection of a sphere is presented. It is also shown that the spherical projection of a sphere can be obtained from a perspective image of a sphere. This projection yields several potential visual features whose interaction matrices are presented in section III. In this same section, we select three independent visual features to control the image of a sphere. The interaction matrix related to those features is maximally decoupled and presents a linear link between the visual features and the system translational velocities. The proposed result is generalized to any central catadioptric system in section IV. For the proposed visual features, a theoretical analysis of the stability and the robustness of a classical control law with respect to modeling error and to calibration errors on perspective cameras is given in section V. Finally, using this sort of cameras, experimental results are given in section VI.

II. SPHERES SPHERICAL PROJECTION

Let $S_{(O,R)}$ be a sphere of radius R and center O with coordinates (X_O, Y_O, Z_O) in the camera frame. Let $S_{p(C,1)}$ be the unit sphere of center the camera optical center C. The spherical projection of $S_{(O,R)}$ onto $S_{p(C,1)}$ is a dome hat. This dome hat can be characterized by the contour Γ of its base. This contour is pictured in Fig. 1(a). We will now determine the analytical form that characterizes Γ.

A. Γ equation

The equation that defines $S_{(O,R)}$ is given by

$$(X - X_O)^2 + (Y - Y_O)^2 + (Z - Z_O)^2 - R^2 = 0. \quad (1)$$

Let $d = \sqrt{X^2 + Y^2 + Z^2}$. A point M of $S_{(O,R)}$ with coordinates (X, Y, Z) is projected onto $S_{p(C,1)}$ in M_S with coordinates (X_S, Y_S, Z_S) such that: $X_S = \frac{X}{d}$, $Y_S = \frac{Y}{d}$ and

R. T. Fomena is with IRISA / Universite de Rennes 1, Campus de Beaulieu, 35 042, Rennes-cedex, France, `rtatsamb@irisa.fr`

F. Chaumette is with IRISA / INRIA, Campus de Beaulieu, 35 042, Rennes-cedex, France, `chaumett@irisa.fr`

$Z_S = \frac{Z}{d}$. Multiplying (1) by $\frac{1}{d^2}$ leads to a $\frac{1}{d}$ polynomial function

$$\frac{K_O^2}{d^2} - \frac{2}{d}(X_O X_S + Y_O Y_S + Z_O Z_S) + 1 = 0 \quad (2)$$

where $K_O^2 = X_O^2 + Y_O^2 + Z_O^2 - R^2$.

The contour Γ is such that the discriminant Δ of (2) is equal to 0 (only one solution to (2)):

$$(X_O X_S + Y_O Y_S + Z_O Z_S)^2 - K_O^2 = 0. \quad (3)$$

Since Γ is in front of the projection center, from (3) we obtain the equation of the plane $\mathcal{P}$ supporting Γ (see Fig. 1(a))

$$X_O X_S + Y_O Y_S + Z_O Z_S = K_O. \quad (4)$$

To sum up, Γ is the intersection of the unit sphere and plane $\mathcal{P}$:

$$\Gamma = \begin{cases} X_S^2 + Y_S^2 + Z_S^2 = 1 \\ X_O X_S + Y_O Y_S + Z_O Z_S = K_O. \end{cases} \quad (5)$$

The contour Γ is therefore a circle from which it is easier to extract optimal visual features than from the ellipse observed on the perspective plane.

B. Γ radius and center

Let $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$ be a new coordinates system where $(\mathbf{u}, \mathbf{v})$ is the orthonormal system on $\mathcal{P}$ and $\mathbf{n}$ is the normal vector of $\mathcal{P}$ (see Fig.1(a)). In this new coordinates system, (5) becomes

$$\Gamma = \begin{cases} X'^2_S + Y'^2_S + Z'^2_S = 1 \\ Z'_S = \sqrt{1 - \frac{R^2}{d_O^2}} \end{cases} \quad (6)$$

with $d_O = \sqrt{X_O^2 + Y_O^2 + Z_O^2}$ and $M'_S = \mathbf{R} M_S$ where $\mathbf{R}$ is the transformation from the $(C, \mathbf{x}, \mathbf{y}, \mathbf{z})$ frame to the $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$ frame.

From (6) we easily obtain the radius of Γ

$$r_s = R/d_O, \quad (7)$$

and the center A of Γ in the $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$ frame

$$A\left(0, 0, \sqrt{1 - r_s^2}\right). \quad (8)$$

Using $\mathbf{R}^\top$, we obtain the coordinates of A in the camera frame

$$\begin{cases} X_A = X_O \sqrt{1 - r_s^2}/d_O \\ Y_A = Y_O \sqrt{1 - r_s^2}/d_O \\ Z_A = Z_O \sqrt{1 - r_s^2}/d_O. \end{cases} \quad (9)$$

C. Potential visual features

In addition to A and r_s, the dome hat summit B (see Fig. 1(b)) can also be considered as potential visual features. The coordinates of B in the camera frame are given by

$$\begin{cases} X_B = X_O/d_O \\ Y_B = Y_O/d_O \\ Z_B = Z_O/d_O. \end{cases} \quad (10)$$

The value of r_s and the coordinates of A and B can be computed using the moments $\mu = (x_g, y_g, n_{20}, n_{11}, n_{02})$ of the ellipse observed on the image plane of a usual camera
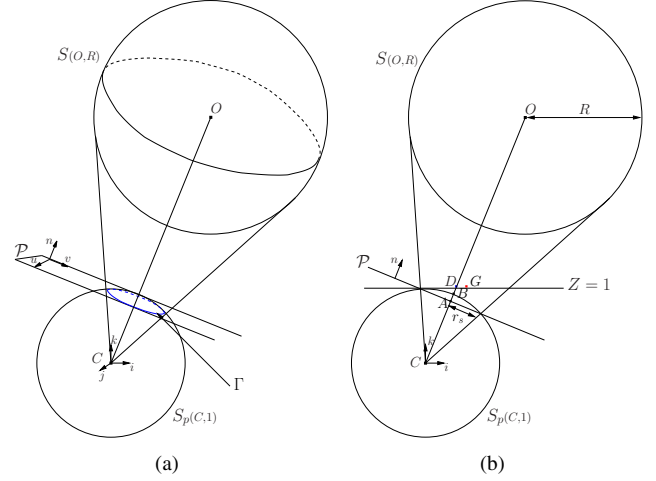

Fig. 1. Spherical projection of a sphere. (a) Contour of the dome hat base. (b) Cut made perpendicular to $\mathcal{P}$.

modeled by a perspective projection. The proof of this assertion is decomposed in two steps.

Let $r_O = \frac{R}{Z_O}$. Let $D(x_O, y_O, 1)$ be the perspective projection of $O(X_O, Y_O, Z_O)$, which means that $x_O = \frac{X_O}{Z_O}$ and $y_O = \frac{Y_O}{Z_O}$.

In the first step, it is showed that (A, B, r_s) can be expressed using x_O, y_O and r_O. First, we obtain immediately from (10)

$$\begin{cases} X_B = x_O/\rho_O \\ Y_B = y_O/\rho_O \\ Z_B = 1/\rho_O \end{cases} \quad (11)$$

where $\rho_O = \sqrt{x_O^2 + y_O^2 + 1}$. Then it is easy to obtain from (7)

$$r_s = \frac{r_O}{\rho_O}. \quad (12)$$

Plugging (12) in (9), we finally get

$$\begin{cases} X_A = x_O \sqrt{\rho_O^2 - r_O^2}/\rho_O^2 \\ Y_A = y_O \sqrt{\rho_O^2 - r_O^2}/\rho_O^2 \\ Z_A = \sqrt{\rho_O^2 - r_O^2}/\rho_O^2. \end{cases} \quad (13)$$

In the second step, we show that x_O, y_O and r_O can be computed using the ellipse moments μ. The ellipse parameters are indeed expressed using X_O, Y_O, Z_O and R as follow [9]:

$$\begin{cases} x_g = -\frac{X_O Z_O}{R^2 - Z_O^2} \\ y_g = -\frac{Y_O Z_O}{R^2 - Z_O^2} \\ 4n_{20} = -\frac{(R^2 - X_O^2 - Z_O^2)R^2}{(R^2 - Z_O^2)^2} \\ 4n_{11} = \frac{X_O Y_O R^2}{(R^2 - Z_O^2)^2} \\ 4n_{02} = -\frac{(R^2 - Y_O^2 - Z_O^2)R^2}{(R^2 - Z_O^2)^2}. \end{cases} \quad (14)$$

From (14) it is easy to obtain

$$
\begin{cases}
x_g = -\dfrac{x_O}{r_O^2 - 1} \\[2mm]
y_g = -\dfrac{y_O}{r_O^2 - 1} \\[2mm]
4n_{20} = -\dfrac{\left(r_O^2 - x_O^2 - 1\right) r_O^2}{\left(r_O^2 - 1\right)^2} \\[2mm]
4n_{11} = \dfrac{x_O y_O r_O^2}{\left(r_O^2 - 1\right)^2} \\[2mm]
4n_{02} = -\dfrac{\left(r_O^2 - y_O^2 - 1\right) r_O^2}{\left(r_O^2 - 1\right)^2}.
\end{cases}
\tag{15}
$$

After tedious computations we get:

$$
\frac{r_O^2}{1 - r_O^2} = f(\mu)
\tag{16}
$$

where $f(\mu) = \frac{4n_{20} y_g^2 + 4n_{02} x_g^2 - 8n_{11} x_g y_g}{x_g^2 + y_g^2}$. It is possible to demonstrate that $f(\mu)$ is continuous even when $x_g = y_g = 0$ in which case $f(\mu) = 4n_{20}$. From (16) we obtain

$$
r_O = \sqrt{\frac{f(\mu)}{1 + f(\mu)}},
\tag{17}
$$

from which we deduce using the first two relations in (15):

$$
\begin{cases}
x_O = \dfrac{x_g}{1 + f(\mu)} \\[2mm]
y_O = \dfrac{y_g}{1 + f(\mu)}.
\end{cases}
\tag{18}
$$

Finally, by plugging (17) and (18) in (11), (12) and (13) we obtain r_s, B and A expressed as functions of μ.

III. INTERACTION MATRICES

In this section we present the interaction matrices related to r_s, A and B. In addition, we propose a new set of three independent visual features for visual servoing from a sphere.

We recall that the interaction matrix $\mathbf{L_s}$ related to $\mathbf{s}$ is defined such that $\dot{\mathbf{s}} = \mathbf{L_s} \mathbf{v}$ where $\mathbf{v}$ is the instantaneous camera velocity [1].

A. Radius r_s

From (7) we obtain after few developments

$$
\dot{r_s} = -\frac{r_s^2}{R} \mathbf{B}^\top \dot{\mathbf{O}}
\tag{19}
$$

where $\dot{\mathbf{O}}$ is given by the well-known equation

$$
\dot{\mathbf{O}} = \begin{bmatrix} -\mathbf{I}_3 & [\mathbf{O}]_\times \end{bmatrix} \mathbf{v}
$$

where $[\mathbf{O}]_\times$ is the skew matrix related to $\mathbf{O}$. Therefore we have

$$
\mathbf{L_{r_s}} = \begin{bmatrix} \dfrac{r_s^2}{R} \mathbf{B}^\top & \mathbf{0}_3 \end{bmatrix}.
\tag{20}
$$

B. Center A - Summit B

From (9), after few developments, the time derivative of the coordinates of A is given by

$$
\dot{\mathbf{A}} = \begin{bmatrix} \dfrac{K r_s}{R} \left(\mathbf{I}_3 + \left(\dfrac{r_s^2}{K^2} - 1 \right) \mathbf{B}\mathbf{B}^\top \right) \end{bmatrix} \dot{\mathbf{O}}
\tag{21}
$$

where $K = \sqrt{1 - r_s^2}$. Therefore we have

$$
\mathbf{L_A} = \begin{bmatrix} -\dfrac{K r_s}{R} \left(\mathbf{I}_3 + \left(\dfrac{r_s^2}{K^2} - 1 \right) \mathbf{B}\mathbf{B}^\top \right) & K[\mathbf{B}]_\times \end{bmatrix}
\tag{22}
$$

Similarly, we get from (10)

$$
\mathbf{L_B} = \begin{bmatrix} \dfrac{r_s}{R} \left(\mathbf{B}\mathbf{B}^\top - \mathbf{I}_3 \right) & [\mathbf{B}]_\times \end{bmatrix}.
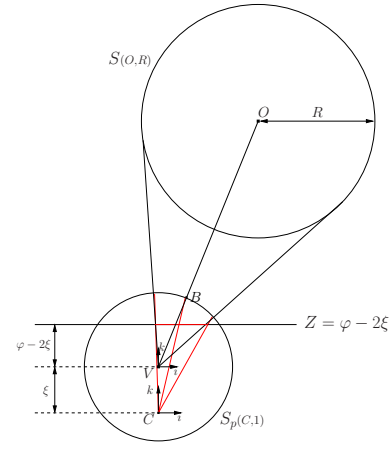\tag{23}
$$

Fig. 2.   Catadioptric image of a sphere.

C. Selection of visual features

Three parameters are sufficient to characterize the spherical projection of a sphere. Therefore, we need to select a combination of three visual features among $\{X_A, Y_A, Z_A, X_B, Y_B, Z_B, r_s\}$.

Among the possible combinations, the decoupling between $(\boldsymbol{v}_x, \boldsymbol{\omega}_y)$ and $(\boldsymbol{v}_y, \boldsymbol{\omega}_x)$ is induced by $\mathbf{s} = (\frac{X_B}{rs}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s})$ since we have in that case

$$
\mathbf{L_s} = \begin{bmatrix}
-\dfrac{1}{R} & 0 & 0 & 0 & -\dfrac{Z_B}{r_s} & \dfrac{Y_B}{r_s} \\[2mm]
0 & -\dfrac{1}{R} & 0 & \dfrac{Z_B}{r_s} & 0 & -\dfrac{X_B}{r_s} \\[2mm]
0 & 0 & -\dfrac{1}{R} & -\dfrac{Y_B}{r_s} & \dfrac{X_B}{r_s} & 0
\end{bmatrix}.
\tag{24}
$$

In addition to the decoupling property, $\mathbf{L_s}$ presents the same dynamic ($\frac{1}{R}$) in the translational velocities. Since R is a constant, there is a linear link between the visual features and the camera translational velocities. For these reasons, we propose the combination $(\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s})$ for visual servoing from spheres.

The only unknown 3D parameter in $\mathbf{L_s}$ is the constant R. In practice, $\hat{R}$ (estimated value of R) is used instead. From the stability analysis to modeling error given in section V, a robustness domain of $\hat{R}$ will be given.

IV. GENERALIZATION TO CENTRAL CATADIOPRTIC SYSTEMS

Considering a catadiotric system with (φ, ξ) as the mirror parameter, we show in this section that we can compute the visual features $\mathbf{s} = (\frac{X_B}{rs}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s})$ from the catadioptric image of a sphere.

A. Equation of the catadioptric image of a sphere

The catadioptric image of a sphere is an ellipse whose formation can be decomposed in two steps considering the unified model of catadiotric image formation [10]. After tedious computations, we obtain the ellipse equation

$$
k_0 x_o^2 + k_1 y_o^2 + 2k_2 x_o y_o + 2k_3 x_o + 2k_4 y_o + k_5 = 0
\tag{25}
$$

$$\text{with}\begin{cases} k_0 = (K_O + \xi Z_O)^2 + \left(\xi^2 - 1\right) X_O \\ k_1 = (K_O + \xi Z_O)^2 + \left(\xi^2 - 1\right) Y_O \\ k_2 = \left(\xi^2 - 1\right) X_O Y_O \\ k_3 = X_O \left(\left(\xi^2 - 1\right) Z_O - \xi \left(K_O + \xi Z_O\right)\right) \\ k_4 = Y_O \left(\left(\xi^2 - 1\right) Z_O - \xi \left(K_O + \xi Z_O\right)\right) \\ k_5 = \left(\left(K_O + \xi Z_O\right)^2 + \left(\xi^2 - 1\right) Z_O^2 - 2\xi Z_O \left(K_O + \xi Z_O\right)\right). \end{cases}$$

B. Visual features computation

Now, we show how to compute **s** using the ellipse moments μ measured on the catadiotric image plane. First of all, we recall that:

$$\begin{cases} \frac{X_B}{r_s} = \frac{X_O}{R} \\ \frac{Y_B}{r_s} = \frac{Y_O}{R} \\ \frac{Z_B}{r_s} = \frac{Z_O}{R} \end{cases} \tag{26}$$

From (25), the ellipse moments on the catadioptric image plane can be expressed using the 3D parameters:

$$\begin{cases} x_g = X_O H_1 / H_2 \\ y_g = Y_O H_1 / H_2 \\ 4n_{20} = \left(H_2 - \left(\xi^2 - 1\right) X_O^2\right) R^2 / H_2^2 \\ 4n_{11} = -X_O Y_O \left(\xi^2 - 1\right) R^2 / H_2^2 \\ 4n_{02} = \left(H_2 - \left(\xi^2 - 1\right) Y_O^2\right) R^2 / H_2^2 \end{cases} \tag{27}$$

$$\text{with}\begin{cases} H_1 = Z_O + \xi K_O \\ H_2 = H_1^2 + \left(\xi^2 - 1\right) R^2. \end{cases}$$

After very tedious computations, we obtain using (27)

$$\begin{cases} \frac{X_B}{r_s} = x_g \frac{h_2}{\sqrt{h_2 + (1 - \xi^2)}} \\ \frac{Y_B}{r_s} = y_g \frac{h_2}{\sqrt{h_2 + (1 - \xi^2)}} \end{cases} \tag{28}$$

where $h_2 = 1/f(\mu)$.

In the case of paracatadioptric systems ($\xi = 1$), we have:

$$\frac{Z_B}{r_s} = \frac{h_1^2 - \left(\frac{X_B^2}{r_s^2} + \frac{Y_B^2}{r_s^2} - 1\right)}{2h_1} \tag{29}$$

where $h_1 = \sqrt{h_2 + (1 - \xi^2)}$, and for all other catadioptric systems ($\xi \neq 1$)

$$\frac{Z_B}{r_s} = \frac{h_1 - \xi\sqrt{h_1^2 + (1 - \xi^2)\left(\frac{X_B^2}{r_s^2} + \frac{Y_B^2}{r_s^2} - 1\right)}}{(1 - \xi^2)}. \tag{30}$$

Finally, we can note that, when $\xi = 0$, (28) and (30) become

$$\begin{cases} \frac{X_B}{r_s} = \frac{x_g}{\sqrt{f(\mu)}\sqrt{1 + f(\mu)}} \\ \frac{Y_B}{r_s} = \frac{y_g}{\sqrt{f(\mu)}\sqrt{1 + f(\mu)}} \\ \frac{Z_B}{r_s} = \frac{\sqrt{1 + f(\mu)}}{\sqrt{f(\mu)}} \end{cases} \tag{31}$$

which exactly corresponds, as expected, to the result obtained using a perspective camera by combining (17) and (18). The results obtained in this section are thus a generalization of the results presented in section II.
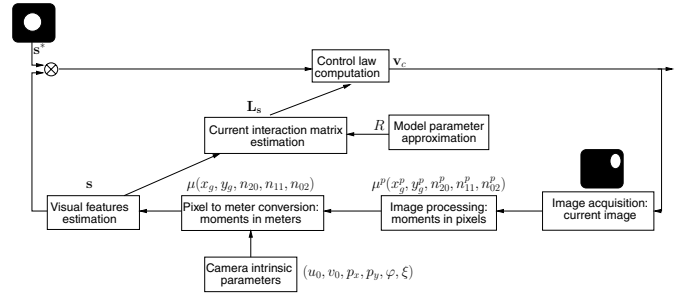
Fig. 3. Visual servoing closed-loop.

V. STABILITY ANALYSIS

A. Stability to modeling error

Let us consider visual servoing from spheres with the combination $\mathbf{s} = \left(\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s}\right)$.

We use the classical control law

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}_\mathbf{s}}^{+} (\mathbf{s} - \mathbf{s}^*) \tag{32}$$

where $\mathbf{v}_c$ is the camera velocity sent to the low level robot controller, λ is a positive gain and $\widehat{\mathbf{L}_\mathbf{s}}^{+}$ is the pseudo-inverse of an approximation of the interaction matrix related to **s**.

Fig. 3 describes the visual servoing closed-loop system. For a catadioptric system with mirror parameter (φ, ξ), the pixel to meter conversion is given by

$$\begin{cases} x_g = \frac{x_g^p - u_0}{p_x(\varphi - \xi)} \\ y_g = \frac{y_g^p - v_0}{p_y(\varphi - \xi)} \\ n_{20} = \frac{n_{20}^p}{p_x^2(\varphi - \xi)^2} \\ n_{11} = \frac{n_{11}^p}{p_x p_y(\varphi - \xi)^2} \\ n_{02} = \frac{n_{02}^p}{p_y^2(\varphi - \xi)^2} \end{cases} \tag{33}$$

where u_0, v_0, p_x, p_y, φ and ξ are the camera intrinsic parameters ($\varphi = 1$, $\xi! = 0$ for a perspective camera).

From Fig. 3, 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. Assuming first that we do not have neither image processing errors nor calibration errors, the closed-loop system equation can be written as:

$$\dot{\mathbf{s}} = -\lambda \mathbf{L_s} \widehat{\mathbf{L}_\mathbf{s}}^{+} (\mathbf{s} - \mathbf{s}^*) \tag{34}$$

with

$$\widehat{\mathbf{L}_\mathbf{s}}^{+} = \begin{bmatrix} -\frac{\hat{R}}{r_s^2 + \hat{R}^2}\left(\hat{R}^2 \mathbf{B}\mathbf{B}^\top + r_s^2 \mathbf{I}_3\right) \\ -\frac{\hat{R}^2 r_s}{r_s^2 + \hat{R}^2}\left[\mathbf{B}\right]_\times \end{bmatrix}.$$

A necessary and sufficient condition for the global asymptotic stability to modeling error is $\mathbf{L_s}\widehat{\mathbf{L}_\mathbf{s}}^{+} > 0$. The eigenvalues of $\mathbf{L_s}\widehat{\mathbf{L}_\mathbf{s}}^{+} > 0$ can be computed. They are given by $\frac{\hat{R}}{R}$ and $\frac{\hat{R}(r_s^2 + R\hat{R})}{R(r_s^2 + \hat{R}^2)}$ (which is a double eigenvalue). Thus we have:

$$\mathbf{L_s}\widehat{\mathbf{L}_\mathbf{s}}^{+} > 0 \iff \hat{R} > 0.$$

Therefore the robustness domain with respect to modeling error is extremely large: $\hat{R} \in]0, +\infty[$. From a practical point of view, a rough estimate of R will be sufficient.

B. Stability to calibration errors

Considering only calibration errors, (34) becomes

$$\dot{\mathbf{e}} = -\lambda \mathbf{L_s} \widehat{\mathbf{L_s}}^+ \mathbf{E} \mathbf{e} \qquad (35)$$

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

$$\widehat{\mathbf{L_s}}^+ = \begin{bmatrix} -\frac{R\widehat{r_s}^2}{\widehat{r_s}^2 + R^2}\left(R^2\left(\widehat{\frac{\mathbf{B}}{r_s}}\right)\left(\widehat{\frac{\mathbf{B}}{r_s}}\right)^\top + \mathbf{I}_3\right) \\ -\frac{R^2\widehat{r_s}^2}{\widehat{r_s}^2 + R^2}\left[\left(\widehat{\frac{\mathbf{B}}{r_s}}\right)\right]_\times \end{bmatrix}.$$

The non-linearity of (35) makes it complex to analyse the global stability to calibration errors. The analysis can be simplified by considering the local stability (that is for $\mathbf{s} = \mathbf{s}^*$) and by linearizing $\mathbf{E}$. $\mathbf{E}$ can be linearized using the derivation of $\mathbf{s}$. Indeed, from (31) and (33) (with $\xi = 0$ and $\varphi = 1$), we obtain after few developments

$$\begin{cases} \left(\frac{X_B}{r_s}\right)^{\cdot} = -\frac{\dot{u}_0\left(p_x^2 + p_y^2\right)}{p_x\left(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}\right)}\frac{Z_B}{r_s} \\ \left(\frac{Y_B}{r_s}\right)^{\cdot} = -\frac{\dot{v}_0\left(p_x^2 + p_y^2\right)}{p_y\left(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}\right)}\frac{Z_B}{r_s} \\ \left(\frac{Z_B}{r_s}\right)^{\cdot} = \frac{p_x\dot{p}_x + p_y\dot{p}_y}{p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}}\frac{Z_B}{r_s}, \end{cases}$$

from which we deduce

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

$$\text{with} \begin{cases} \alpha_{u_0} = -\frac{\Delta u_0\left(p_x^2 + p_y^2\right)}{p_x\left(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}\right)} \\ \alpha_{v_0} = -\frac{\Delta v_0\left(p_x^2 + p_y^2\right)}{p_y\left(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}\right)} \\ \alpha_{p_x p_y} = \frac{p_x\Delta p_x + p_y\Delta p_y}{p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}}. \end{cases}$$

The eigenvalues of $\mathbf{L_{s^*}}\widehat{\mathbf{L_{s^*}}}^+\mathbf{E}$ can be obtained after complex computations. They are given by

$$\begin{cases} \gamma_1 = \frac{\widehat{r_s^*}^2}{r_s^{*2}}\frac{r_s^{*2} + R^2(1 + \alpha_{p_x p_y})}{\widehat{r_s^*}^2 + R^2} \\ \gamma_2 = a + b^{\frac{1}{2}} \\ \gamma_3 = a - b^{\frac{1}{2}} \end{cases} \qquad (36)$$

where a and b are real values which are too complex to be written here. From (36) it is possible to show that a sufficient condition to ensure $Re(\gamma_i) > 0$ is $\Delta p_x > -p_x$ and $\Delta p_y > -p_y$; that is $\hat{p}_x > 0$ and $\hat{p}_y > 0$ which is, once again, an excellent result.

VI. EXPERIMENTAL RESULTS

The experiments have been carried out with a conventional perspective camera mounted on the end of effector of a six DOFs robotic system. The experiments consist of a centering task. The target is a $2\,cm$ radius ping-pong ball. Using such simple object allows to easily compute the ellipse moments at video rate without any image processing problem. $\mathbf{s}^*$ has

(a)   (b)

Fig. 4.   (a) Desired image. (b) Initial image.

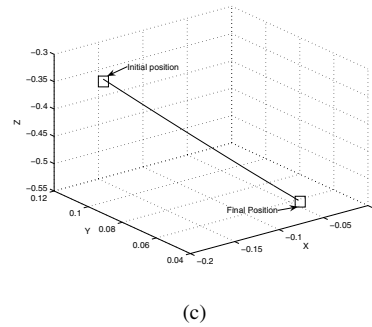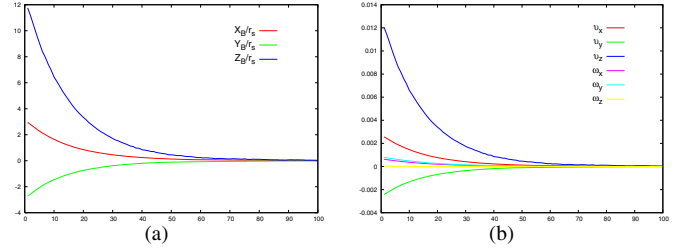(a)   (b)

(c)

Fig. 5.   Ideal case. (a) $\frac{B}{r_s}$ error. (b) Computed camera velocities (m/s and dg/s). (c) Camera trajectory in the cartesian space.

been selected as $(0, 0, a^*)$ where a^* allows to fix the depth range between the camera and the ball. Fig. 4 pictures the desired and the initial images used for each experiment. For all the experiments, the same gain $\lambda = 0.5$ has been used.

A. Ideal case

In oder to validate the selected visual features, we first consider the ideal case where $\hat{R} = R$. Indeed, when $\hat{R} = R$ we have a perfect system behaviour 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 on Fig. 5(a) while the camera velocities are plotted on Fig. 5(b). Finally, the camera trajectory in the cartesian space is a pure straight line as expected also (see Fig. 5(c)).

B. Modeling error

The stability to modeling error has been proved in this paper. We have validated this proof, with two experiments. The results in the case where $\hat{R} = 10R$ and $\hat{R} = 0.1R$ are depicted respectively in Fig. 6 and Fig. 7. We can note that in both cases the system still converges.

Fig. 6(b) shows a high speed on the system translational velocities while Fig. 7(b) shows a low speed on the same
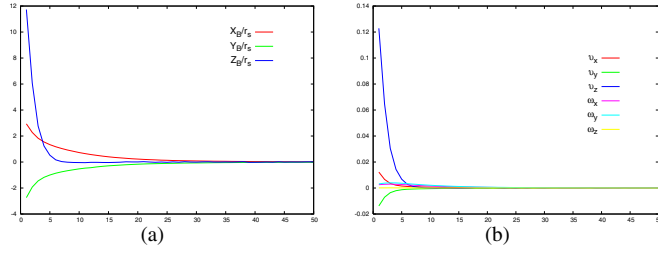
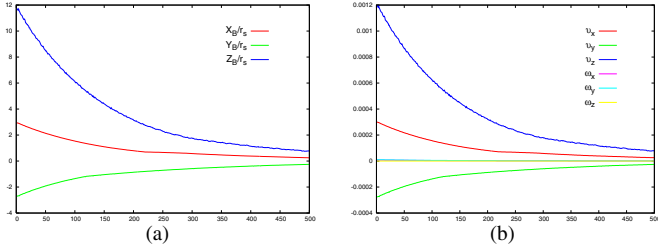Fig. 6. Modeling error $\hat{R} = 10R$. (a) $\frac{B}{r_s}$ error. (b) Computed camera velocities (m/s and dg/s).

Fig. 7. Modeling error $\hat{R} = 0.1R$. (a) $\frac{B}{r_s}$ error. (b) Computed camera velocities (m/s and dg/s).

components. Indeed, from (32) (where $\widehat{\mathbf{L_s}}^+$ is defined in (34)), we easily obtain

$$\begin{cases} \mathbf{v} = \lambda \frac{\hat{R}}{r_s^2+\hat{R}^2}\left(\hat{R}^2 \mathbf{B}\mathbf{B}^\top + r_s^2 \mathbf{I}_3\right)\left(\frac{\mathbf{B}}{r_s} - \frac{\mathbf{B}^*}{r_s^*}\right) \\ \boldsymbol{\omega} = \lambda \frac{\hat{R}^2 r_s}{r_s^2+\hat{R}^2}[\mathbf{B}]_\times \left(\frac{\mathbf{B}}{r_s} - \frac{\mathbf{B}^*}{r_s^*}\right). \end{cases} \quad (37)$$

where $\mathbf{v}$ and $\boldsymbol{\omega}$ are respectively the translational and the rotational velocities of the camera. When $\hat{R}$ tends to $+\infty$, (37) tends to

$$\begin{cases} \mathbf{v} = \infty \\ \boldsymbol{\omega} = \lambda r_s [\mathbf{B}]_\times \left(\frac{\mathbf{B}}{r_s} - \frac{\mathbf{B}^*}{r_s^*}\right) \end{cases}$$

which explains the fast convergence observed in Fig. 6 (50 iterations) when $\hat{R} = 10R$. When $\hat{R}$ tends to 0, from (37) we have: $\mathbf{v}$ and $\boldsymbol{\omega}$ tend to 0. This explains the slow convergence observed in Fig. 7 (500 iterations) when $\hat{R} = 0.1R$. In practice, the behaviour could be easily improved, by using a higher gain λ (to deal with under approximation of $\hat{R}$) and by saturating $\mathbf{v}_c$ when needed (to deal with an over approximation of $\hat{R}$).

C. Calibration errors

Finally we validate the results obtained for the stability analysis to calibration errors. The sufficient condition $\hat{p_x} > 0$ & $\hat{p_y} > 0$ is validated by introducing errors on the camera intrinsic parameters: $35\% p_x$, $-47\% p_y$, $-25\% u_0$ and $57\% v_0$. The results obtained are given on Fig. 8. Once again the system converges (while only local stability has been considered). The rapid decrease of the error on $\frac{Y_B}{r_s}$ compared to the other features is simply due to the biggest error introduced on p_y. Of course we have observed that if $\hat{p_x} < 0$ or $\hat{p_y} > 0$, the system diverges, but that is not a surprise at all.
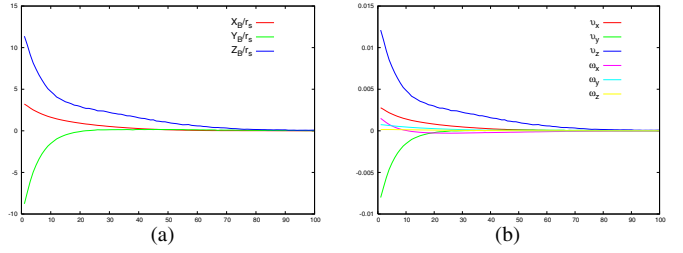

Fig. 8. Calibration errors $\hat{p_x} > 0$ & $\hat{p_y} > 0$. (a) $\frac{B}{r_s}$ error. (b) Computed camera velocities (m/s and dg/s).

VII. CONCLUSIONS AND FUTURE WORKS

In this paper, a spherical projection model has been used to search for optimal visual features for visual servoing. A new combination of three independent visual features for visual servoing from spheres has been presented. This combination can be computed from the image of a sphere on any catadioptric system. The interaction matrix related to this combination is linear on the translational components and is maximally decoupled. Using this new combination, a classical control law has been analytically proved to be globally stable with respect to modeling error and locally stable to calibration errors on perpective cameras. These strong and nice theoretical results have been validated experimentally with a conventional perspective camera mounted on a robotic system.

Future works will be devoted to search for similar results for other geometric primitives such as straight line for instance.

REFERENCES

[1] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 3, pp. 313–326, June 1992.

[2] 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.

[3] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Confluence of Vision and Control*. LNCIS Series, No 237, Springer-Verlag, 1998, pp. 66–78.

[4] R. Mahony, P. Corke, and F. Chaumette, "Choice of image features for depth-axis control in image-based visual servo control," in *IROS'02*, vol. 1, Lausanne, Switzerland, Oct. 2002, pp. 390–395.

[5] P. Corke and S. Hutchinson, "A new partioned approach to image-based visual visual servo control," *IEEE Trans. Robot. Autom.*, vol. 17, no. 4, pp. 507–515, Aug. 2001.

[6] 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.

[7] 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.

[8] 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.

[9] F. Chaumette, P. Rives, and B. Espiau, "Classification and realization of the different vision-based tasks," in *Visual Servoing*, K. Hashimoto, Ed. Singapore: World Scientific Series in Robotics and Automated Systems, 1993, vol. 7, pp. 199–228.

[10] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," *ECCV*, vol. 29, pp. 159–179, May 2000.