

# Visual servoing from three points using a spherical projection model

Romeo Tatsambon Fomena, Omar Tahri, François Chaumette

► **To cite this version:**

Romeo Tatsambon Fomena, Omar Tahri, François Chaumette. Visual servoing from three points using a spherical projection model. IEEE Int. Conf. on Robotics and Automation, ICRA'10, 2010, Anchorage, Alaska, United States. pp.5537-5542, 2010. <inria-00544794>

**HAL Id: inria-00544794**

**<https://hal.inria.fr/inria-00544794>**

Submitted on 9 Dec 2010

**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.

# Visual Servoing from Three Points using a Spherical Projection Model

Romeo Tatsambon Fomena, Omar Tahri and François Chaumette

**Abstract**—This paper deals with visual servoing from three points. Using the geometric properties of the spherical projection of points, a new decoupled set of six visual features is proposed. The main originality lies in the use of the distances between spherical projection of points to define three features that are invariant to camera rotations. The three other features present a linear link with respect to camera rotations. In comparison with the classical perspective coordinates of points, the new decoupled set does not present more singularities. In addition, using the new set in its non-singular domain, a classical control law is proven to be ideal for rotational motions. These theoretical results as well as the robustness to errors of the new decoupled control scheme are illustrated through simulation results.

## I. INTRODUCTION

Visual servoing consists in 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 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. By ideal, satisfaction of the following criteria is meant: local and -as far as possible- global stability of the system, robustness to calibration and to modeling errors, non-singularity, local minima avoidance, satisfactory trajectory of the system and of the features in the image, and finally maximal decoupled and linear link (the ultimate goal) between the visual features and the degrees of freedom (DOFs) taken into account.

Points are the most simple features that can be extracted from an image. That is why most of visual servoing experiments use the image of points. For example the image of points of interest is sometimes used in mobile robotic applications in natural environment [3]. The image of points can also be used to regulate the position of an aerial vehicle [4], [5].

Regarding the image of points, lots of works have been dedicated to approach an ideal system behaviour using 3D data [6], [7]; hybrid data [8]; and 2D data [9], [10], [11]. Moment invariants theory has been used to determine specific combinations of 2D moments whose interaction with the system presents linear and decoupling properties when planar objects are considered [12]. Recently spherical moments

invariant to camera rotations (e.g. the area of a surface) have been used to design decoupled sets of features for objects defined by points cloud [13], [14].

Even if it is well known that the same image of three points corresponds to four different camera poses [15], it is possible to control a six degrees of freedom (DOFs) robot using only three points in a local neighborhood of the desired pose. Inspired by the above-mentioned last works, this paper exploits the geometric properties of the spherical projection of points and proposes a new decoupled set of six features. The originality presented in this paper in comparison with the earlier works [13], [14], is that the decoupling is obtained by three features invariant to rotations, which are the distances between the spherical projection of three points. The three other features present a linear link w.r.t. the camera rotational velocities and generalizes the work proposed in [16], where the target is a sphere marked with a tangent vector to a point on its surface. The features modeling is given in Section II. In Section III we compare our approach with the classical perspective coordinates of three points, which have been proven (a long time ago) to present a singularity domain defined by a cylinder [17]: the cylinder of singularities is defined by the circumcircle passing through the three points (when they are not aligned) and the normal to the plane on which the points lie. In addition the perspective coordinates of points are not suited for camera rotations [2]. A key contribution of this paper is the formal demonstration that our approach does not present more singularities and is well suited for camera rotations in comparison with the classical perspective coordinates of three points. Finally, these theoretical results and the robustness of the control scheme are validated in simulation in Section IV.

## II. FEATURE MODELING

In this section, using a spherical projection model, we design a decoupled set  $\mathbf{s}_{\text{sp}} = (\mathbf{s}_t, \zeta)$  of six features to control the image of three points. The set of features  $\mathbf{s}_t$  is invariant to camera rotations while the set  $\zeta$  is linearly linked to camera rotations.

Let  $\mathcal{S}(\mathcal{C}, 1)$  be the unit sphere of projection center in  $\mathcal{C}$ ;  $\mathcal{F}_c = (\mathcal{C}, \mathbf{x}, \mathbf{y}, \mathbf{z})$  be the frame attached to the camera unique projection center  $\mathcal{C}$ ;  $\mathcal{P}_0, \mathcal{P}_1$  and  $\mathcal{P}_2$  be a set of three points; and  $\mathbf{P}_i = (P_{ix}, P_{iy}, P_{iz})$  be the vector coordinates of  $\mathcal{P}_i$  in  $\mathcal{F}_c$ .

We first 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 [18];

R. Tatsambon Fomena is with Université de Rennes 1, IRISA, Campus de Beaulieu, 35042, Rennes-cedex, France [rtatsamb@irisa.fr](mailto:rtatsamb@irisa.fr)

O. Tahri is with ISR, University of Coimbra, Polo II PT-3030-290 Coimbra, Portugal [omartahri@isr.uc.pt](mailto:omartahri@isr.uc.pt)

F. Chaumette is with INRIA Rennes-Bretagne Atlantique, IRISA, Campus de Beaulieu, 35042, Rennes-cedex, France, [chaumett@irisa.fr](mailto:chaumett@irisa.fr)

$\mathbf{v}$  and  $\boldsymbol{\omega}$  are respectively the translational and the rotational velocities of the camera.

Let  $\mathbf{p}_{0s}$ ,  $\mathbf{p}_{1s}$  and  $\mathbf{p}_{2s}$  be the spherical projection of points  $\mathcal{P}_0$ ,  $\mathcal{P}_1$  and  $\mathcal{P}_2$  respectively. We recall that this projection is defined by  $\mathbf{p}_{is} = \mathbf{P}_i / \|\mathbf{P}_i\|$ ,  $i = 0, 1, 2$ . It is clear that the distance  $d_{12}$  between the spherical projections  $\mathbf{p}_{1s}$  and  $\mathbf{p}_{2s}$  (see Fig. 1(a)) is invariant to camera rotations. The analytical expression of  $d_{12}$  is given by

$$d_{12} = \|\mathbf{p}_{1s} - \mathbf{p}_{2s}\| = \left( (\mathbf{p}_{1s} - \mathbf{p}_{2s})^\top (\mathbf{p}_{1s} - \mathbf{p}_{2s}) \right)^{\frac{1}{2}}. \quad (1)$$

From the time variation of (1), we obtain after some developments the expression of the interaction matrix related to  $d_{12}$ :

$$\mathbf{L}_{d_{12}} = \left[ -\frac{1}{d_{12}} (\mathbf{p}_{1s} - \mathbf{p}_{2s})^\top \left( \frac{1}{\|\mathbf{P}_1\|} \Gamma_{\mathbf{P}_1} - \frac{1}{\|\mathbf{P}_2\|} \Gamma_{\mathbf{P}_2} \right) \quad \mathbf{0}_{\{1 \times 3\}} \right] \quad (2)$$

where  $\Gamma_{\mathbf{P}_i} = \mathbf{I}_3 - \mathbf{p}_{is} \mathbf{p}_{is}^\top$ ,  $i = 1, 2$ . The bloc  $\mathbf{0}_{\{1 \times 3\}}$  on the rotation component of the interaction matrix  $\mathbf{L}_{d_{12}}$  clearly shows the invariance property.

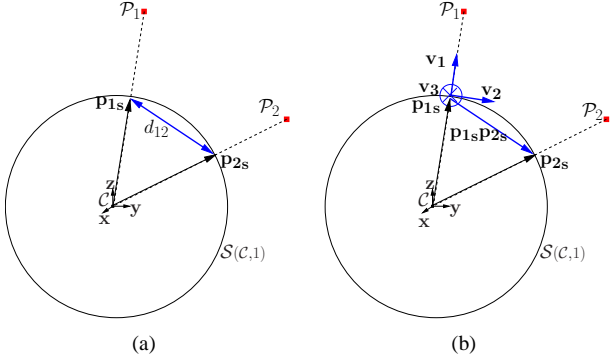


Fig. 1. Spherical projection of two points: (a) distance between the spherical images of the two points, (b) components of the rotation matrix.

From the spherical projections  $\mathbf{p}_{1s}$  and  $\mathbf{p}_{2s}$ , it is possible to determine a set of three features  $\zeta$  such that the interaction matrix has the form

$$\mathbf{L}_\zeta = \begin{bmatrix} \mathbf{L}_{\omega, v} & -\mathbf{I}_3 \end{bmatrix}, \quad (3)$$

where  $\mathbf{L}_{\omega, v}$  will be defined later. The set  $\zeta$  has been originally proposed for visual servoing from a sphere marked with a tangent vector in [16], where the configuration between the two points  $\mathcal{P}_1$  and  $\mathcal{P}_2$  in the object frame  $\mathcal{F}_O = (\mathcal{O}, \mathbf{x}, \mathbf{y}, \mathbf{z})$  is such that  $\mathcal{O}\mathcal{P}_1 \perp \mathcal{P}_1\mathcal{P}_2$  (see Fig. 2(a)). The set  $\zeta$  can be seen as the  $\theta \mathbf{u}$  representation ( $\theta$  is the angle of rotation and  $\mathbf{u}$  the unitary axis of rotation) of the rotation matrix  $\mathbf{V}\mathbf{V}^{*-1}$ , where  $\mathbf{V}^*$  is the desired value of the matrix  $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$  (see Fig. 1(b)) defined as follows:

$$\mathbf{v}_1 = \mathbf{p}_{1s}, \quad \mathbf{v}_2 = \frac{\Gamma_{\mathbf{P}_1}(\mathbf{P}_2 - \mathbf{P}_1)}{\|\Gamma_{\mathbf{P}_1}(\mathbf{P}_2 - \mathbf{P}_1)\|}, \quad \mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2. \quad (4)$$

In that specific configuration ( $\mathcal{O}\mathcal{P}_1 \perp \mathcal{P}_1\mathcal{P}_2$ ), the expression of matrix  $\mathbf{L}_{\omega, v}$  is given by:

$$\mathbf{L}_{\omega, v} = \frac{1}{\|\mathbf{P}_1\|} (-\delta \mathbf{v}_1 \mathbf{v}_3^\top + \mathbf{v}_2 \mathbf{v}_3^\top - \mathbf{v}_3 \mathbf{v}_2^\top) \quad (5)$$

with  $\delta = (\mathbf{o}_s^\top \mathbf{v}_2) / \sqrt{r^2 - 1 + (\mathbf{o}_s^\top \mathbf{v}_1)^2}$ ,  $r^2 = \|\mathbf{P}_1 - \mathbf{O}\|^2 / \|\mathbf{O}\|^2$ ,  $\mathbf{o}_s = \mathbf{O} / \|\mathbf{O}\|$  where  $\mathbf{O} = (O_x, O_y, O_z)$  is the vector coordinates of the center of object frame  $\mathcal{O}$  in the camera frame.

In this paper, we propose to compute  $\zeta$  for a general configuration between the points as pictured on Fig. 2(b). With that generalization, matrix  $\mathbf{L}_{\omega, v}$  is now given by (see [19] for the details of the developments)

$$\mathbf{L}_{\omega, v} = \frac{1}{\|\Gamma_{\mathbf{P}_1}(\mathbf{P}_2 - \mathbf{P}_1)\|} \mathbf{v}_1 \mathbf{v}_3^\top \mathbf{M}_{\mathbf{P}_1 \mathbf{P}_2} + \frac{1}{\|\mathbf{P}_1\|} (\mathbf{v}_2 \mathbf{v}_3^\top - \mathbf{v}_3 \mathbf{v}_2^\top), \quad (6)$$

where

$$\mathbf{M}_{\mathbf{P}_1 \mathbf{P}_2} = -\frac{1}{\|\mathbf{P}_2\|} \Gamma_{\mathbf{P}_1} \Gamma_{\mathbf{P}_2} + \frac{1}{\|\mathbf{P}_1\|} ((\mathbf{P}_1^\top \mathbf{P}_2) \mathbf{I}_3 + \mathbf{P}_1 \mathbf{P}_2^\top) \Gamma_{\mathbf{P}_1}. \quad (7)$$

In the case of the specific configuration  $\mathcal{O}\mathcal{P}_1 \perp \mathcal{P}_1\mathcal{P}_2$  (see Fig. 2(a)), the value of the general expression (6) of the block matrix  $\mathbf{L}_{\omega, v}$  of  $\mathbf{L}_\zeta$  (see (3)) has been numerically verified to be equal to the value of specific expression (5).

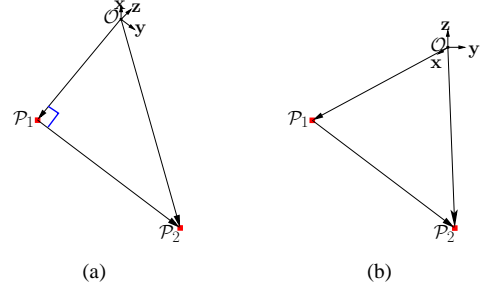


Fig. 2. Configuration of the two points in object frame  $\mathcal{F}_O = (\mathcal{O}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ : (a) specific case presented in [16], (b) general case.

To sum up, the new set  $\mathbf{s}_{sp} = (\mathbf{s}_t, \zeta)$  is such that the decoupling is obtained by the set  $\mathbf{s}_t = (d_{01}, d_{02}, d_{12})$  where  $d_{01}$  and  $d_{02}$  are the distances between the spherical projection couples  $(\mathbf{p}_{0s}, \mathbf{p}_{1s})$  and  $(\mathbf{p}_{0s}, \mathbf{p}_{2s})$  respectively; the linear link with the camera rotation motions is obtained by the set  $\zeta$  which is proposed here for a general configuration between two points. Of course, we recall that the spherical projection of points can easily be obtained from its perspective projection, or from its projection on any omnidirectional sensor [20].

### III. CONTROL ANALYSIS

The new interaction matrix, obtained by stacking the two interaction matrices  $\mathbf{L}_{s_t}$  and  $\mathbf{L}_\zeta$ , is a lower block triangular square matrix:

$$\mathbf{L}_{\mathbf{s}_{sp}} = \begin{bmatrix} \mathbf{L}_v & \mathbf{0} \\ \mathbf{L}_{\omega, v} & -\mathbf{I}_3 \end{bmatrix}, \quad (8)$$

where matrix  $\mathbf{L}_{\omega,v}$  is given by (6) and

$$\mathbf{L}_v = \begin{bmatrix} -\frac{1}{d_{01}} (\mathbf{p}_{0s} - \mathbf{p}_{1s})^\top \left( \frac{1}{\|\mathbf{cP}_0\|} \Gamma_{\mathbf{p}_{0s}} - \frac{1}{\|\mathbf{cP}_1\|} \Gamma_{\mathbf{p}_{1s}} \right) \\ -\frac{1}{d_{02}} (\mathbf{p}_{0s} - \mathbf{p}_{2s})^\top \left( \frac{1}{\|\mathbf{cP}_0\|} \Gamma_{\mathbf{p}_{0s}} - \frac{1}{\|\mathbf{cP}_2\|} \Gamma_{\mathbf{p}_{2s}} \right) \\ -\frac{1}{d_{12}} (\mathbf{p}_{1s} - \mathbf{p}_{2s})^\top \left( \frac{1}{\|\mathbf{cP}_1\|} \Gamma_{\mathbf{p}_{1s}} - \frac{1}{\|\mathbf{cP}_2\|} \Gamma_{\mathbf{p}_{2s}} \right) \end{bmatrix}.$$

With this property the determination of the singularity domain of  $\mathbf{L}_{s_{sp}}$  will be shown to be easier than in the case of the perspective projection of three points. In this section we also analyse the stability of the control law.

We use the classical control law

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}}_{s_{sp}}^{-1} (\mathbf{s}_{sp} - \mathbf{s}_{sp}^*) \quad (9)$$

where  $\mathbf{v}_c = (\mathbf{v}, \boldsymbol{\omega})$  is the camera velocity sent to the low level robot controller,  $\lambda$  is a positive gain and  $\widehat{\mathbf{L}}_{s_{sp}}^{-1}$  is the inverse of an approximation of the interaction matrix related to  $\mathbf{s}_{sp}$ .

Taking the inverse of the interaction matrix (8) and plugging it into (9) leads to the ideal control law

$$\mathbf{v} = -\lambda \mathbf{L}_v^{-1} (\mathbf{s}_t - \mathbf{s}_t^*), \quad \boldsymbol{\omega} = \mathbf{L}_{\omega,v} \mathbf{v} + \lambda \boldsymbol{\zeta}. \quad (10)$$

The domain of singularity of the above control is given by the following theorem:

*Theorem 1:* the classical control method (10) is singular if and only if:

- the three points are aligned; or
- the camera optical center  $\mathcal{C}$  lies on the cylinder of singularities, that is the cylinder which is defined by the circumcircle passing through the three points and the normal to the plane on which the point lie (see Fig. 3).

The proof of *theorem 1* is given in the appendix.

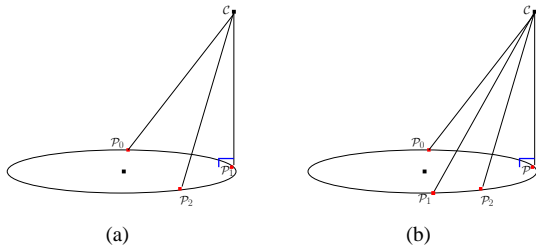


Fig. 3. Cylinder of singularities (a) particular case, (b) general case.

*Theorem 1* ensures that the new set  $\mathbf{s}_{sp}$  does not present more singularities than the perspective coordinates of the three points  $\mathbf{s}_{pp} = (p_{0x}, p_{0y}, p_{1x}, p_{1y}, p_{2x}, p_{2y})$  where the interaction matrix  $\mathbf{L}_{s_{pp}}$  is obtained by stacking the interaction matrix of single point given in [9]:

$$\mathbf{L}_{s_{pp}} = \begin{bmatrix} \mathbf{L}_t & \mathbf{L}_r \end{bmatrix}, \quad (11)$$

with  $\mathbf{L}_t \neq \mathbf{0}$  and  $\mathbf{L}_r \neq \mathbf{0}$ . Indeed, using  $\mathbf{s}_{pp}$  the classical control law (9) has been shown to be singular when the three points are aligned or when the camera optical center belongs to the cylinder of singularities [17].

With the decoupling property of  $\mathbf{s}_{sp}$  it is also easier to determine the kernel of  $\mathbf{L}_{s_{sp}}$  that characterizes the set of camera motions  $\mathbf{v}_c \neq \mathbf{0}$  which leaves the image unchanged

i.e.  $\dot{\mathbf{s}}_{sp} = \mathbf{0}$ . Indeed, from the expression of the interaction matrix given in (8), we have

$$\ker(\mathbf{L}_{s_{sp}}) = \{\mathbf{v}_c = (\mathbf{v}, \boldsymbol{\omega}) \in \mathfrak{se}(3), \mathbf{L}_v \mathbf{v} = \mathbf{0}, \boldsymbol{\omega} = \mathbf{L}_{\omega,v} \mathbf{v}\}, \quad (12)$$

which shows that we have to deal only with the  $3 \times 3$  matrix  $\mathbf{L}_v$ , contrary to the perspective projection of three points where we deal with the larger  $6 \times 6$  matrix  $\mathbf{L}_{s_{pp}}$ .

The new interaction matrix  $\mathbf{L}_{s_{sp}}$  given in (8) depends on the depths of the points which are unknown in practice. Indeed we use an estimated value  $\|\widehat{\mathbf{P}}_i\|$ ,  $i = 0, 1, 2$  which can be expressed as follows

$$\|\widehat{\mathbf{P}}_i\| = |\widehat{P}_{iz}| \rho_i \quad (13)$$

with  $\rho_i = \sqrt{(P_{ix}/P_{iz})^2 + (P_{iy}/P_{iz})^2 + 1}$  where  $P_{ix}/P_{iz} = p_{s_{ix}}/p_{s_{iz}}$  and  $P_{iy}/P_{iz} = p_{s_{iy}}/p_{s_{iz}}$  can be measured from the spherical image of the points.

We assume in the following stability analysis that the interaction matrix never loses its rank during the servoing, i.e. the camera never crosses the cylinder of singularities pictured on Fig. 3. We also suppose that we have neither image processing errors nor vision system calibration errors. Under these assumptions the closed-loop system equation (using the control law (9)) can be written as:

$$\dot{\mathbf{e}} = -\lambda \mathbf{L}_{s_{sp}} \widehat{\mathbf{L}}_{s_{sp}}^{-1} \mathbf{e} \quad (14)$$

with  $\mathbf{e} = \mathbf{s}_{sp} - \mathbf{s}_{sp}^*$ ,

$$\widehat{\mathbf{L}}_{s_{sp}}^{-1} = \begin{bmatrix} \widehat{\mathbf{L}}_v^{-1} & \mathbf{0} \\ \widehat{\mathbf{L}}_{\omega,v} \widehat{\mathbf{L}}_v^{-1} & -\mathbf{I}_3 \end{bmatrix},$$

where  $\widehat{\mathbf{L}}_v^{-1}$  and  $\widehat{\mathbf{L}}_{\omega,v}$  depends on  $\widehat{P}_{iz}$ ,  $i = 0, 1, 2$ . The stability of the system (14) can be analysed: in the ideal case (no depths errors), expression (14) becomes  $\dot{\mathbf{e}} = -\lambda \mathbf{e}$  which means that the system is locally asymptotically stable. In the case of errors on the estimation of the points depths  $\widehat{P}_{iz}$ ,  $i = 0, 1, 2$ , the robustness domain of the control law given in (9) is complex to establish. But simulation results given in Section IV will demonstrate the robustness of the classical control (9) in the case of depths errors.

For a pure rotation motion, the value of  $\mathbf{s}_t$  is constant, i.e.  $\mathbf{s}_t = \mathbf{s}_t^*$ . From the expression of the control law given in (9) we immediately obtain

$$\mathbf{v} = \mathbf{0}, \quad \boldsymbol{\omega} = \lambda \boldsymbol{\zeta}, \quad (15)$$

which shows that  $\boldsymbol{\zeta}$  is well suited to control camera rotations. From (15), it is clear that, in the case of pure rotations, the classical control is perfectly adequate even in the case of errors on points depths.

#### IV. SIMULATION RESULTS

In this section, we compare the new decoupled set  $\mathbf{s}_{sp}$  with the classical perspective coordinates of three points  $\mathbf{s}_{pp}$  using the ViSP simulator [21]. We first present the case of camera rotation motions only; then we consider only translation motions; and finally we present the case where we consider both camera rotation and translation motions.

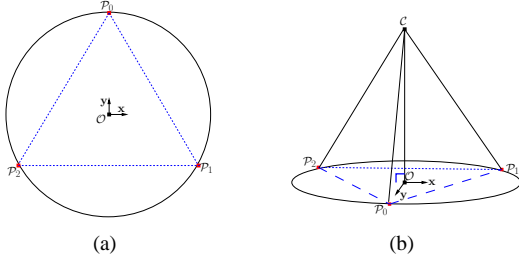


Fig. 4. Configuration of three points in the 3D space: (a) equilateral triangle and circumcircle, (b) desired pose of the camera.

In the object frame  $\mathcal{F}_o = (\mathcal{O}, \mathbf{x}, \mathbf{y}, \mathbf{z})$  the coordinates of the three points are given by  $\mathbf{OP}_0 = r(0, 1, 0.30)$ ,  $\mathbf{OP}_1 = r(\sqrt{3}/2, -1/2, 0.30)$  and  $\mathbf{OP}_2 = r(-\sqrt{3}/2, -1/2, 0.30)$  with  $r = 0.5\text{m}$ ; the configuration of the three points describes an equilateral triangle as shown on Fig. 4(a); and the circumcircle (of center  $\mathcal{O}$  and radius  $r$ ) of the equilateral triangle is the circular cross-section of the cylinder of singularities. The pose of the object frame  $\mathcal{F}_o$  w.r.t. the desired pose of the camera frame  $\mathcal{F}_{c^*}$  is set to the values  ${}^{c^*}\mathbf{t}_o = (0, 0, 2.35)$  (m) and  $\theta_{\mathbf{u}}({}^{c^*}\mathbf{R}_o) = (0, 0, 0)$  (rad), which means that the desired pose of the camera is inside the cylinder of singularities as shown on Fig. 4(b).

Now we present two experiments where we consider only camera rotation motions. In the first experiment we highlight both the decoupling of the control using the new set  $\mathbf{s}_{\text{SP}}$  and the coupling of the control using the classical set  $\mathbf{s}_{\text{PP}}$ . The orientation of the initial camera frame w.r.t. the desired camera frame has been set to  $\theta_{\mathbf{u}}({}^{c^*}\mathbf{R}_c) = (-0.20, 0.17, 0.79)$  (rad). As expected in the Cartesian space, because of the decoupling, the new set  $\mathbf{s}_{\text{SP}}$  does not cause any translation displacement of the camera while the classical set  $\mathbf{s}_{\text{PP}}$  does (see Fig. 5(a)). Indeed the coupling in the control law using the classical set  $\mathbf{s}_{\text{PP}}$  (see (11)) generates undesired translation velocities with oscillations (compare Fig. 5(b) and Fig. 5(c)). Using the classical set, all the features vary which is not the case when using the new set where only the subset  $\zeta$  varies since the subset  $\mathbf{s}_t = (d_{01}, d_{02}, d_{12})$  is invariant to rotations (compare Figs. 5(d) and 5(e)). As expected also, the control using the new set  $\mathbf{s}_{\text{SP}}$  is ideal since the camera rotation velocities are linearly linked to the set of features  $\zeta$ , and both decrease exponentially (compare Figs. 5(e) and 5(c)).

In the second experiment we validate the robustness of the control in the case of modeling errors, i.e. errors on points depths. We have introduced the following error of the depth estimation of points:  $\widehat{P}_{iz} = 0.5P_{iz}^*$ ,  $i = 0, 1, 2$ . The relative orientation of the initial camera frame w.r.t. the desired camera frame is set to the same value as in the first experiment above. As expected, the robot displays exactly the same ideal behaviour using  $\mathbf{s}_{\text{SP}}$ , as shown on Figs. 5(e) and 5(c) where only rotation motions are generated. Indeed, even in the case of modeling errors, the control scheme generates rotation motions only, as given by the expression (15).

Now we validate the new set for a translation motion and compare it with the classical set. The relative pose of the

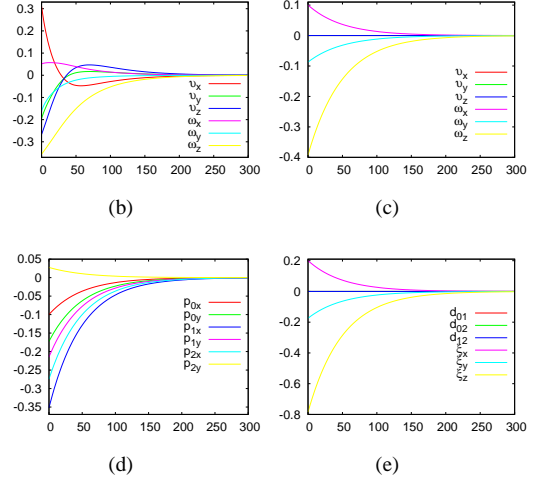
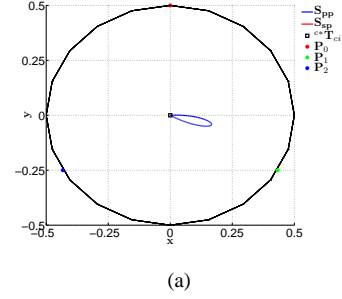


Fig. 5. Decoupled vs coupled control: (a) camera Cartesian trajectories, (b) and (c) computed camera velocities (m/s and rad/s) using  $\mathbf{s}_{\text{PP}}$  and  $\mathbf{s}_{\text{SP}}$ , (d) and (e) errors on  $\mathbf{s}_{\text{PP}}$  and  $\mathbf{s}_{\text{SP}}$ .

initial camera frame w.r.t. the desired camera frame is set to the value  ${}^{c^*}\mathbf{t}_c = (0.21, 0.31, -0.5)$  (m). Using the new set, even if there is a little oscillation on  $v_x$  (at the beginning of the servoing) which does not appear with the control using the classical set (compare Figs. 6(b) and 6(c)), the robot Cartesian trajectory is satisfactory (see Fig. 6(a)).

In the next two experiments we consider complex motions, i.e. motions made up of both rotation and translation displacements. In the first experiment we validate the new set in the case of a motion where the relative pose of the initial camera frame w.r.t. the desired camera frame is set to the value:  ${}^{c^*}\mathbf{t}_c = (0.29, 0.16, -0.52)$  (m) and  $\theta_{\mathbf{u}}({}^{c^*}\mathbf{R}_c) = (0.21, -0.21, 0.30)$  (rad). In this case we have a better Cartesian trajectory (straight line) using the new set (see Fig. 7(a)). In addition, camera velocities present no oscillation contrary to the velocities using the classical set (compare Figs. 7(b) and 7(c)).

In the second experiment, in order to validate the larger convergence domain of the control using the new set, we consider a relatively large displacement where the camera initial pose is very close to the boundary of the cylinder of singularities. The relative initial pose of the camera frame w.r.t. the desired camera frame is set to the value:  ${}^{c^*}\mathbf{t}_c = (-0.29, -0.37, -0.48)$  (m) and  $\theta_{\mathbf{u}}({}^{c^*}\mathbf{R}_c) = (-0.26, 0.17, -0.52)$  (rad). The coupled control induced by the classical set causes the control to cross the cylinder of singularities and to converge towards another global minimum (see Fig. 8), while the new decoupled

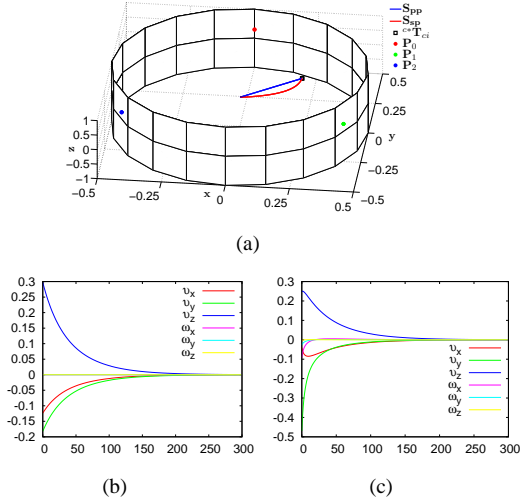


Fig. 6. Comparison in the case of a translation motion: (a) camera Cartesian trajectories, (b) and (c) computed camera velocities (m/s and rad/s) using  $\mathbf{s}_{pp}$  and  $\mathbf{s}_{sp}$

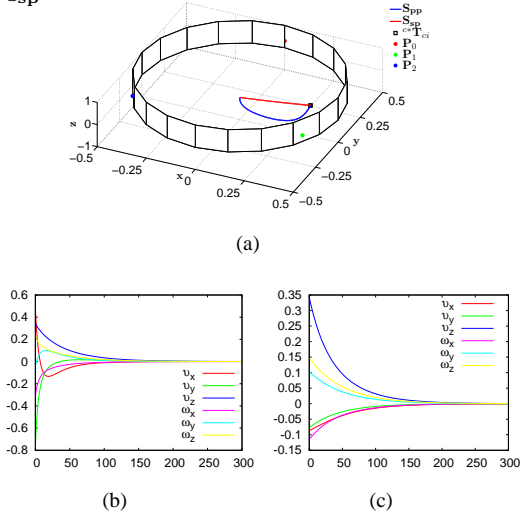


Fig. 7. Comparison in the case of a complex motion: (a) camera Cartesian trajectories, (b) and (c) computed camera velocities (m/s and rad/s) using  $\mathbf{s}_{pp}$  and  $\mathbf{s}_{sp}$ .

control shows satisfactory Cartesian trajectory and converges towards the desired pose.

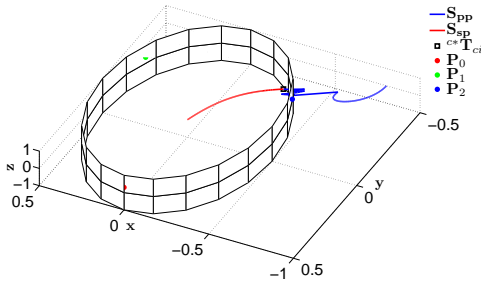


Fig. 8. Larger convergence domain for the new set: the cylinder of singularities is crossed by the classical set  $\mathbf{s}_{pp}$ .

To conclude, simulation results have shown that the new set is well suited to camera rotation motions; more importantly, using the new set, the convergence domain has been

shown to be larger than in the case of the classical perspective coordinates.

## V. CONCLUSIONS AND FUTURE WORKS

In this paper we have proposed a new decoupled set of features for visual servoing from three points. Using a spherical projection model, the new set consists in three features invariant to camera rotations and three other features which vary linearly w.r.t. camera rotation motions. The three invariants to camera rotation are the three distances between the spherical projection of points. In comparison with the classical perspective coordinates of three points we have demonstrated that the new set does not have more singularities. Indeed, the singularity domain of the new set has been theoretically characterized: the singularities appear when either the three points are aligned or the camera optical center lies on the well-known cylinder of singularities. The new decoupled set has been formally shown to be well suited to camera rotation motions. These theoretical results have been successfully validated in simulation, where the convergence domain of a classical control method using the new set has been shown to be larger than with the classical set. As future works, it would be interesting to investigate the existence of image features invariant to camera translations in order to design a totally decoupled control scheme.

## APPENDIX

Here we give a proof of *theorem 1*.

*Proof:* [Theorem 1] The key element to the determination of the singularities of the classical control given in (10) is the factorization of the determinant of the interaction matrix given in (8).

From (8), since  $\mathbf{L}_{\mathbf{s}_{sp}}$  is a square triangular matrix, it is immediate to show that

$$|\mathbf{L}_{\mathbf{s}_{sp}}| = -|\mathbf{L}_v|. \quad (16)$$

Using  $\mathbf{P}_0$ ,  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , the block matrix  $\mathbf{L}_v$  of the interaction matrix (8) can be rewritten as

$$\mathbf{L}_v = \begin{bmatrix} \alpha_{01}(k_{10}\mathbf{P}_0^\top + k_{01}\mathbf{P}_1^\top) \\ \alpha_{02}(k_{20}\mathbf{P}_0^\top + k_{02}\mathbf{P}_2^\top) \\ \alpha_{12}(k_{21}\mathbf{P}_1^\top + k_{12}\mathbf{P}_2^\top) \end{bmatrix}, \quad (17)$$

where  $\alpha_{ij} = \frac{1}{d_{ij}\|\mathbf{P}_i\|\|\mathbf{P}_j\|}$  and  $k_{ij} = 1 - \frac{\|\mathbf{P}_i\|}{\|\mathbf{P}_j\|} \cos(\mathbf{P}_i, \mathbf{P}_j)$ .

The determinant of  $\mathbf{L}_v$  can be easily computed from the determinant of its transpose  $\mathbf{L}_v^\top$ . Indeed, from the fact that  $|\mathbf{L}_v| = |\mathbf{L}_v^\top|$ , we have

$$|\mathbf{L}_v| = |\mathbf{L}_{vx}^\top, \mathbf{L}_{vy}^\top, \mathbf{L}_{vz}^\top|, \quad (18)$$

where

$$\begin{cases} \mathbf{L}_{vx}^\top = \alpha_{01}(k_{10}\mathbf{P}_0 + k_{01}\mathbf{P}_1) \\ \mathbf{L}_{vy}^\top = \alpha_{02}(k_{20}\mathbf{P}_0 + k_{02}\mathbf{P}_2) \\ \mathbf{L}_{vz}^\top = \alpha_{12}(k_{21}\mathbf{P}_1 + k_{12}\mathbf{P}_2). \end{cases}$$

Using the multilinearity property of the determinant application, from (18) we obtain after some developments,

$$|\mathbf{L}_v| = \alpha_{01}\alpha_{02}\alpha_{12}(k_{10}k_{02}k_{21} + k_{01}k_{20}k_{12}) |\mathbf{P}_1, \mathbf{P}_0, \mathbf{P}_2|. \quad (19)$$



By plugging (19) into (16), we easily obtain a factorization of the determinant of  $\mathbf{L}_{\text{ssp}}$ :

$$|\mathbf{L}_{\text{ssp}}| = -\alpha_{01}\alpha_{02}\alpha_{12}(k_{10}k_{02}k_{21} + k_{01}k_{20}k_{12}) |\mathbf{P}_1, \mathbf{P}_0, \mathbf{P}_2|, \quad (20)$$

where  $\alpha_{01}\alpha_{02}\alpha_{12} \neq 0$  since  $\alpha_{ij} \neq 0$ ,  $i = 0, 1, j = 1, 2$ . From (20), it is clear that, if the three points are aligned then  $|\mathbf{P}_1, \mathbf{P}_0, \mathbf{P}_2| = 0$ .

Now we show that in the case where the three points are not aligned, the expression

$$k_{10}k_{02}k_{21} + k_{01}k_{20}k_{12} = 0 \quad (21)$$

characterizes the cylinder of singularities (see Fig. 3) defined in [17]. Using the expressions of  $k_{ij}$  given in (17), after some developments, it is possible to show that (21) is equivalent to

$$(\mathbf{P}_1\mathbf{P}_0^\top\mathbf{C}\mathbf{P}_0)(\mathbf{P}_0\mathbf{P}_2^\top\mathbf{C}\mathbf{P}_2)(\mathbf{P}_2\mathbf{P}_1^\top\mathbf{C}\mathbf{P}_1) + (\mathbf{P}_0\mathbf{P}_1^\top\mathbf{C}\mathbf{P}_1)(\mathbf{P}_2\mathbf{P}_0^\top\mathbf{C}\mathbf{P}_0)(\mathbf{P}_1\mathbf{P}_2^\top\mathbf{C}\mathbf{P}_2) = 0 \quad (22)$$

Expression (22) is easily verified for the particular configuration of the cylinder of singularities illustrated on Fig. 3(a) where  $(\mathbf{P}_2\mathbf{P}_1^\top\mathbf{C}\mathbf{P}_1) = (\mathbf{P}_0\mathbf{P}_1^\top\mathbf{C}\mathbf{P}_1) = 0$ .

Let the point  $\mathcal{P}$  be the orthogonal projection of the camera optical center  $\mathcal{C}$  onto the plane defined by the three points  $\mathcal{P}_0, \mathcal{P}_1$  and  $\mathcal{P}_2$  (see Fig. 3(b)). Let  $\mathcal{F}'_c$  be a frame centered in  $\mathcal{C}$  and oriented such that  $\mathbf{z} = \mathbf{C}\mathbf{P}/\|\mathbf{C}\mathbf{P}\|$ . In  $\mathcal{F}'_c$ , the coordinates of  $\mathcal{P}$  are given by  $(0, 0, P'_z)$  and points  $\mathcal{P}_0, \mathcal{P}_1$  and  $\mathcal{P}_2$  have all the same z-component  $P'_{iz} = P'_z$ . We denote  $\mathbf{P}'_i = (P'_{ix}, P'_{iy}, P'_z)$  the vector coordinates of  $\mathcal{P}_i$  in  $\mathcal{F}'_c$ .

Equation (22) still holds in  $\mathcal{F}'_c$  since rotation preserves dot product. By expressing (22) in  $\mathcal{F}'_c$ , expanding and simplifying, we get the product of two determinants

$$\begin{vmatrix} P'_{0x} & P'_{0y} & 1 \\ P'_{1x} & P'_{1y} & 1 \\ P'_{2x} & P'_{2y} & 1 \end{vmatrix} \cdot \begin{vmatrix} P'_{0x} & P'_{0y} & P'^2_{0x} + P'^2_{0y} \\ P'_{1x} & P'_{1y} & P'^2_{1x} + P'^2_{1y} \\ P'_{2x} & P'_{2y} & P'^2_{2x} + P'^2_{2y} \end{vmatrix} = 0.$$

Since the three points are not aligned, we have on one hand

$$\begin{vmatrix} P'_{0x} & P'_{0y} & 1 \\ P'_{1x} & P'_{1y} & 1 \\ P'_{2x} & P'_{2y} & 1 \end{vmatrix} \neq 0.$$

On the other hand, we must have

$$\begin{vmatrix} P'_{0x} & P'_{0y} & P'^2_{0x} + P'^2_{0y} \\ P'_{1x} & P'_{1y} & P'^2_{1x} + P'^2_{1y} \\ P'_{2x} & P'_{2y} & P'^2_{2x} + P'^2_{2y} \end{vmatrix} = 0. \quad (23)$$

The camera optical center  $\mathcal{C}$  belongs to the cylinder of singularities iff  $\mathcal{P}$  belongs to the circumcircle defined by the three points. The property that the point  $\mathcal{P}$  of coordinates  $(0, 0, P'_z)$  belongs to the circumcircle defined by the three points can be expressed by the following three point formula for the circle [22]

$$\begin{vmatrix} 0 & 0 & 0 & 1 \\ P'^2_{0x} + P'^2_{0y} & P'_{0x} & P'_{0y} & 1 \\ P'^2_{1x} + P'^2_{1y} & P'_{1x} & P'_{1y} & 1 \\ P'^2_{2x} + P'^2_{2y} & P'_{2x} & P'_{2y} & 1 \end{vmatrix} = 0 \quad (24)$$

which is clearly equivalent to expression (23). The cylinder of singularities described on Fig. 3 is thus characterized by the expression (23) which is equivalent to (21). ■

## REFERENCES

- [1] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. on Robotics and Automation*, 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*, D. Kriegman, G. Hager, and A. S. Morse, Eds. LNCIS Series, No 237, Springer-Verlag, 1998, pp. 66–78.
- [3] A. Remazeilles, F. Chaumette, and P. Gros, "Robot motion control from a visual memory," in *ICRA'04*, vol. 4, Avril 2004, pp. 4695–4700.
- [4] T. Hamel and R. Mahony, "Visual servoing of an under-actuated dynamic rigid-body system: an image-based approach," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 2, pp. 187–198, Apr. 2002.
- [5] O. Bourquardez, R. Mahony, N. Guenard, F. Chaumette, T. Hamel, and L. Eck, "Image-based visual servo control of the translation kinematics of a quadrotor aerial vehicle," *IEEE Trans. on Robotics*, vol. 25, no. 3, pp. 743–749, Jun. 2009.
- [6] W. Wilson, C. Hulls, and G. Bell, "Relative end-effector control using cartesian position-based visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 5, pp. 684–696, Oct. 1996.
- [7] P. Martinet, J. Gallice, and D. Khadraoui, "Vision based control law using 3d visual features," in *World Automation Congress, WAC, Robotics and manufacturing systems*, 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. on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, Apr. 1999.
- [9] J. T. Feddema, C. S. G. Lee, and O. R. Mitchell, "Automatic selection of image features for visual servoing of a robot manipulator," in *ICRA '89*, Mai 1989, pp. 832–837.
- [10] P. Corke and S. Hutchinson, "A new partitioned approach to image-based visual servo control," *IEEE Trans. on Robotics and Automation*, 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. on Robotics*, 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. on Robotics*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.
- [13] O. Tahri, F. Chaumette, and Y. Mezouar, "New decoupled visual servoing scheme based on invariants projection onto a sphere," in *ICRA'08*, May 2008.
- [14] O. Tahri, Y. Mezouar, F. Chaumette, and P. Corke, "Generic decoupled image-based visual servoing for cameras obeying the unified projection model," in *ICRA'09*, May 2009, pp. 1116–1121.
- [15] R. P. Horaud, "New methods for matching 3d objects with single perspective views," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 9, no. 3, pp. 401–412, 1987.
- [16] N. Cowan and D. Chang, "Geometric visual servoing," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1128–1138, Dec. 2005.
- [17] H. Michel and P. Rives, "Singularities in the determination of the situation of a robot," in *Technical Report, INRIA*, Feb. 1993.
- [18] 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.
- [19] R. T. Fomena, "Asservissement visuel par projection sphérique," Ph.D. dissertation, Université de Rennes 1, November 2008.
- [20] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," *European Conference on Computer Vision*, vol. 29, pp. 159–179, May 2000.
- [21] E. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robotics and Automation Magazine*, vol. 12, no. 4, pp. 40–52, Dec. 2005.
- [22] H. S. M. Coxeter, *Introduction to geometry*. Wiley, New York, 1969.