

OMNIDIRECTIONAL VISION-BASED CONTROL FROM HOMOGRAPHY

Youcef Mezouar, Hicham Hadj Abdelkader, Philippe Martinet
LASMEA / Blaise Pascal University
UMR6602 CNRS
24, avenue des Landais
France

1. Introduction

Vision-based servoing schemes are flexible and effective methods to control robot motions from cameras observations (Hutchinson et al 1996). They are traditionally classified into three groups, namely position-based, image-based and hybrid-based control (Espiau et al 1992), (Hutchinson et al 1996), (Malis et al 1999). These three schemes make assumptions on the link between the initial, current and desired images since they require correspondences between the visual features extracted from the initial image with those obtained from the desired one. These features are then tracked during the camera (and/or the object) motion. If these steps fail the visually based robotic task can not be achieved. Typical cases of failure arise when matching joint images features is impossible (for example when no joint features belongs to initial and desired images) or when some parts of the visual features get out of the field of view during the servoing. Some methods have been investigated to resolve this deficiency based on path planning (Mezouar et al 2002), switching control (Chesi et al 2003), zoom adjustment (Benhimane et al 2003). However, such strategies are sometimes delicate to adapt to generic setup.

Conventional cameras suffer thus from restricted field of view. There is significant motivation for increasing the field of view of the cameras. Many applications in vision-based robotics, such as mobile robot localization (Blaer et al 2002) and navigation (Winter et al 2000), can benefit from panoramic field of view provided by omnidirectional cameras. In the literature, there have been several methods proposed for increasing the field of view of cameras systems (Benosman et al 2000). One effective way is to combine mirrors with conventional imaging system. The obtained sensors are referred as catadioptric imaging systems. The resulting imaging systems have been termed central catadioptric when a single projection center describes the world-image mapping. From a theoretical and practical view point, a single center of projection is a desirable property for an imaging system (Baker et al 1999). Baker and Nayar in (Baker et al 1999) derive the entire class of catadioptric systems with a single viewpoint. Clearly, visual servoing applications can also benefit from such sensors since they naturally overcome the visibility constraint. Vision-based control of robotic arms, single mobile robot or formation of mobile robots appear thus in the literature with omnidirectional cameras (refer for example to (Barreto et al 2002), (Vidal et al 2003), (Mezouar et al 2004). Image-based visual servoing with central catadioptric cameras using points has been studied by in (Barreto et al 2002). The use of straight lines has also been

investigated in (Mezouar et al 2004).

This paper is concerned with homography-based visual servo control techniques with central catadioptric cameras. This framework (called 2 1/2 D visual servoing) has been first proposed by Malis and Chaumette in (Malis et al 1999). The 2 1/2 D visual servoing scheme exploits a combination of reconstructed Euclidean information and image-space information in the control design. The 3D informations are extracted from a homography matrix relating two views of a reference plane. As a consequence, the 2 1/2 D visual servoing scheme does not require any 3D model of the target. The resulting interaction matrix is triangular with interesting decoupling properties and it has no singularity in the whole task space. Unfortunately, in such approach the image of the target is not guaranteed to remain in the camera field of view. Motivated by the desire to overcome this deficiency, we extend in this paper homography-based visual servo control techniques to an entire class of omnidirectional cameras. We describe how to obtain a generic homography matrix related to a reference plane for central catadioptric cameras using imaged points or lines. Then the 3D informations obtained from the homography is used to develop 2 1/2 D visual servoing schemes based on points and lines features. Simulations as well as experimental results on a six degrees of freedom robotic arm illustrate the efficiency of omnidirectional vision-based control with homography.

2. Central catadioptric imaging model

The central catadioptric projection can be modelled by a central projection onto a virtual unitary sphere, followed by a perspective projection onto an image plane. This virtual unitary sphere is centered in the principal effective view point and the image plane is attached to the perspective camera. In this model, called unified model and proposed by Geyer and Daniilidis in (Geyer et al 2000), conventional perspective camera appears as a particular case.

2.1 Projection of point

Let F_c and F_m be the frames attached to the conventional camera and to the mirror respectively. In the sequel, we suppose that F_c and F_m are related by a simple translation along the Z-axis (F_c and F_m have the same orientation as depicted in Figure 1). The origins C and M of F_c and F_m will be termed optical center and principal projection center respectively. The optical center C has coordinates $[0 \ 0 \ -\xi]^T$ with respect to F_m and the image plane $Z=f.(\varphi-2\xi)$ is orthogonal to the Z-axis where f is the focal length of the conventional camera and ξ and φ describe the type of sensor and the shape of the mirror, and are function of mirror shape parameters (refer to (Barreto et al 2002b)).

Consider the virtual unitary sphere centered in M as shown in Fig. 1 and let X be a 3D point with coordinates $X = [X \ Y \ Z]^T$ with respect to F_m . The world point X is projected in the image plane into the point of homogeneous coordinates $x_i = [x_i \ y_i \ 1]^T$. The image formation process can be split in three steps as:

- **First step:** The 3D world point X is first projected on the unit sphere surface into a point of coordinates in F_m :

$$\mathbf{X}_m = \frac{1}{\|\mathbf{X}\|} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

The projective ray \mathbf{X}_m passes through the principal projection center \mathbf{M} and the world point \mathbf{X} .

- **Second step:** The point \mathbf{X}_m lying on the unitary sphere is then perspectively projected on the normalized image plane $Z=1-\xi$. This projection is a point of homogeneous coordinates $\underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix}^T = \mathbf{f}(\mathbf{X})$ (where $\mathbf{x} = [x \ y]^T$):

$$\underline{\mathbf{x}} = \mathbf{f}(\mathbf{X}) = \begin{bmatrix} \frac{X}{Z+\xi\|\mathbf{X}\|} & \frac{Y}{Z+\xi\|\mathbf{X}\|} & 1 \end{bmatrix}^T \quad (1)$$

- **Third step:** Finally the point of homogeneous coordinates \mathbf{x}_i in the image plane is obtained after a plane-to-plane collineation \mathbf{K} of the 2D projective point \mathbf{x} :

$$\mathbf{x}_i = \mathbf{K}\mathbf{x}$$

The matrix \mathbf{K} can be written as $\mathbf{K}=\mathbf{K}_c \mathbf{M}$ where the upper triangular matrix \mathbf{K}_c contains the conventional camera intrinsic parameters, and the diagonal matrix \mathbf{M} contains the mirror intrinsic parameters:

$$\mathbf{M} = \begin{bmatrix} \varphi - \xi & 0 & 0 \\ 0 & \varphi - \xi & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{K}_c = \begin{bmatrix} \varphi - \xi & 0 & 0 \\ 0 & \varphi - \xi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that, setting $\xi=0$, the general projection model becomes the well known perspective projection model.

In the sequel, we assume that $Z \neq 0$. Let us denote $\eta = s\|\mathbf{X}\|/|Z| = s\sqrt{1+X^2/Z^2+Y^2/Z^2}$, where s is the sign of Z . The coordinates of the image point can be rewritten as:

$$x = \frac{X/Z}{1+\xi\eta} \text{ and } y = \frac{Y/Z}{1+\xi\eta}$$

By combining the two previous equations, it is easy to show that η is the solution of the following second order equation:

$$\eta^2 - (x+y)^2(1+\xi\eta) - 1 = 0$$

Noticing that the sign of η is equal to the sign of Z , it can be shown that the exact solution is:

$$\eta = \frac{-\gamma - \xi(x^2 + y^2)}{\xi(x^2 + y^2) - 1} \quad (2)$$

where $\gamma = \sqrt{1 + (1 - \xi)(x^2 + y^2)}$. Equation (2) shows that η can be computed as a function of image coordinates \mathbf{x} and sensor parameter ξ . Noticing that:

$$\mathbf{X}_m = (\eta^{-1} + \xi)\bar{\mathbf{x}} \quad (3)$$

where $\bar{\mathbf{x}} = \begin{bmatrix} \mathbf{x}^T & \frac{1}{1 + \xi\eta} \end{bmatrix}^T$, we deduce that \mathbf{X}_m can also be computed as a function of image coordinates \mathbf{x} and sensor parameter ξ .

2.2 Projection of lines

Let \mathbf{L} be a 3D straight line in space lying on the interpretation plane which contains the principal projection center \mathbf{M} (see Figure 1). The binormalized Euclidean Plücker

coordinates (Andreff et al 2002) of the 3D line are defined as: $\mathbf{L} = \begin{bmatrix} \bar{\mathbf{h}}^T & \bar{\mathbf{u}}^T & \mathbf{h} \end{bmatrix}^T$. The unit vectors $\bar{\mathbf{h}} = (h_x \ h_y \ h_z)^T$ and $\bar{\mathbf{u}} = (u_x \ u_y \ u_z)^T$ are respectively the orthogonal vector to the interpretation plane and the orientation of the 3D line \mathbf{L} and are expressed in the mirror frame F_m . h is the distance from \mathbf{L} to the origin of the definition frame. The unit vectors $\bar{\mathbf{h}}$ and $\bar{\mathbf{u}}$ are orthogonal, thus verify $\bar{\mathbf{h}}^T \bar{\mathbf{u}} = 0$. If the 3D line is imaged with a perspective camera then the unit vector $\bar{\mathbf{h}}$ contains the coefficient of the 2D line \mathbf{l} in the image plane, i.e the homogeneous coordinates \mathbf{x} of the perspective projection of any world point lying on \mathbf{L} verifies:

$$(\mathbf{K}^{-T} \bar{\mathbf{h}})^T \mathbf{x} = \mathbf{l}^T \mathbf{x} = 0 \quad (4)$$

If the line is imaged with a central catadioptric camera then the 3D points on the 3D line \mathbf{L} are mapped into points \mathbf{x} in the catadioptric image lying on a conic curve:

$$\mathbf{x}^T \mathbf{K}^{-T} \bar{\mathbf{u}} \mathbf{K}^{-1} \mathbf{x} = \mathbf{x}^T \bar{\mathbf{u}}_1 \mathbf{x} = 0 \quad (5)$$

Where $\bar{\mathbf{u}}_1 = \mathbf{K}^{-T} \bar{\mathbf{u}} \mathbf{K}^{-1}$ and:

$$\Omega \propto \begin{bmatrix} h_x^2 - \xi(1-h_y^2) & h_x h_y(1-\xi^2) & h_x h_z \\ h_x h_y(1-\xi^2) & h_y^2 - \xi(1-h_x^2) & h_y h_z \\ h_x h_z & h_y h_z & h_z^2 \end{bmatrix}$$

2.3 Polar lines

The quadratic equation (5) is defined by five coefficients. Nevertheless, the catadioptric image of a 3D line has only two degrees of freedom. In the sequel, we show how we can get a minimal representation using polar lines.

Let Φ and \mathbf{A} be respectively a 2D conic curve and a point in the definition plane of Φ . The polar line \mathbf{l} of \mathbf{A} with respect to Φ is defined by $\mathbf{l} \propto \Phi \mathbf{A}$. Now, consider the principal point $\mathbf{O}_i = [u_0 \ v_0 \ 1]^T = \mathbf{K}[0 \ 0 \ 1]^T$ and the polar line \mathbf{l}_i of \mathbf{O}_i with respect to Φ : $\mathbf{l}_i \propto \Phi \mathbf{O}_i$, then:

$$\begin{aligned} \mathbf{l}_i &\propto \mathbf{K}^{-T} \Omega \mathbf{K}^{-1} \mathbf{O}_i = \mathbf{K}^{-T} \Omega \mathbf{K}^{-1} \mathbf{K} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \\ &\propto \mathbf{K}^{-T} \bar{\mathbf{h}} \end{aligned} \quad (6)$$

Moreover, equation (6) yields:

$$\bar{\mathbf{h}} = \frac{\mathbf{K}^T \mathbf{l}_i}{\|\mathbf{K}^T \mathbf{l}_i\|} \quad (7)$$

It is thus clear that the polar line \mathbf{l}_i contains the coordinates of the projection of the 3D line \mathbf{L} in an image plane of an equivalent (virtual) perspective camera defined by the frame $\mathbf{F}_v = \mathbf{F}_m$ (see Figure 2) with internal parameters chosen equal to the internal parameters of the catadioptric camera (*i.e.* $\mathbf{K}_v = \mathbf{K}_c \mathbf{M}$). This result is fundamental since it allows us to represent the physical projection of a 3D line in a catadioptric camera by a simple (polar) line in a virtual perspective camera rather than a conic. Knowing only the optical center \mathbf{O}_i , it is thus possible to use the linear pin-hole model for the projection of a 3D line instead of the non linear central catadioptric projection model.

3. Scaled Euclidean reconstruction

Several methods were proposed to obtain Euclidean reconstruction from two views (Faugeras et al 1988). They are generally based on the estimation of the fundamental matrix (Faugeras et al 96) in pixel space or on the estimation of the essential matrix (Longuet and Higgins 1981) in normalized space. However, for control purposes, the methods based on the essential matrix are not well suited since degenerate configurations can occur (such as pure rotational motion). Homography matrix and Essential matrix based approaches do not share the same degenerate configurations, for example pure rotational motion is not a

degenerate configuration when using homography-based method. The epipolar geometry of central catadioptric system has been more recently investigated (Geyer et al 2003, Svoboda et al 1998). The central catadioptric fundamental and essential matrices share similar degenerate configurations that those observed with conventional perspective cameras, it is why we will focus on homographic relationship. In the sequel, the collineation matrix \mathbf{K} and the mirror parameter ξ are supposed known. To estimate these parameters the algorithm proposed in Barreto et al 2002 can be used. In the next section, we show how we can compute homographic relationships between two central catadioptric views of co-planar points and co-planar lines.

Let \mathbf{R} and \mathbf{t} be the rotation matrix and the translation vector between two positions F_m and F_m^* of the central catadioptric camera (see Figs. 1 and 2). Consider a 3D reference plane (π) given in F_m^* by the vector $\mathbf{n}^T = [\mathbf{n}^* - d^*]$, where \mathbf{n}^* is its unitary normal in F_m^* and d^* is the distance from (π) to the origin of F_m^* .

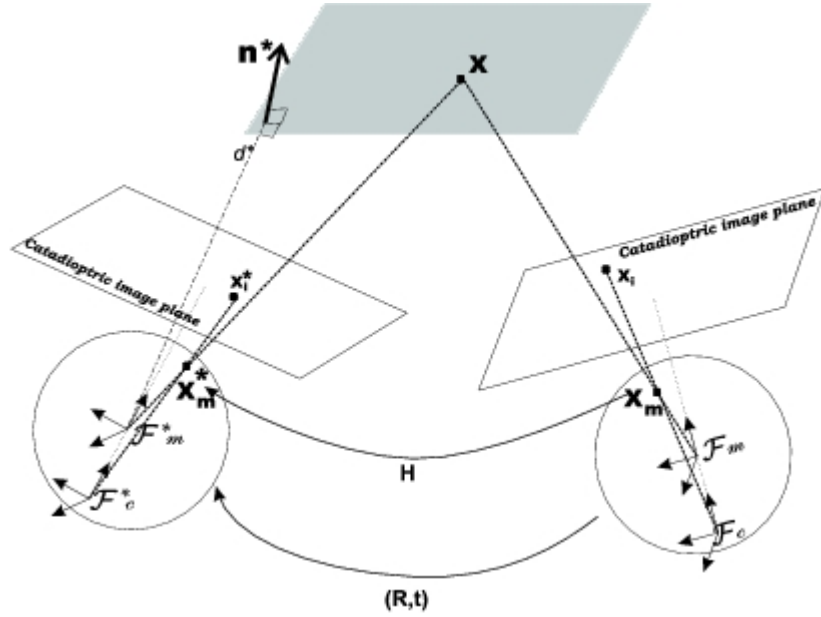


Fig. 1: Geometry of two views, the case of points

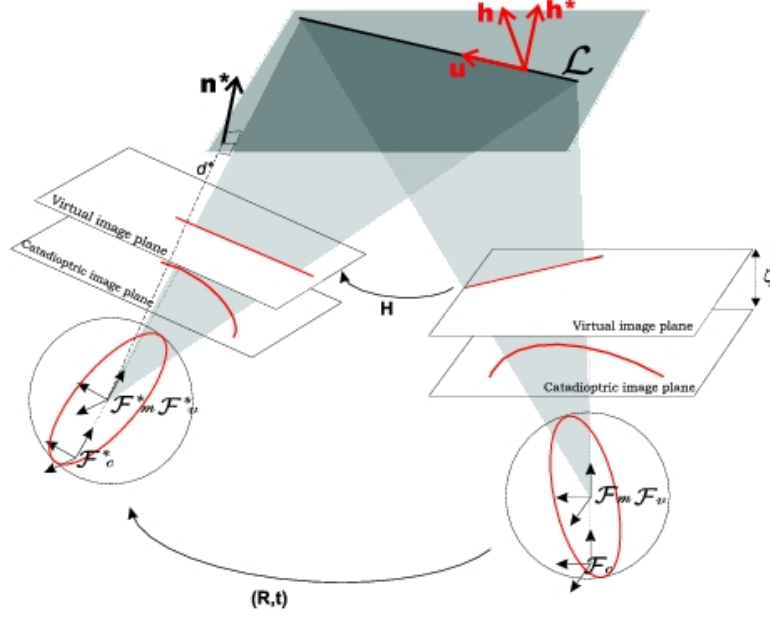


Fig. 2: Geometry of two views, the case of lines

3.1 Homography matrix from points

Let \mathbf{X} be a 3D point with coordinates $\mathbf{X} = [X \ Y \ Z]^T$ with respect to \mathcal{F}_m and with coordinates $\mathbf{X}^* = [X^* \ Y^* \ Z^*]^T$ with respect to \mathcal{F}_m^* . Its projection in the unit sphere for the two camera positions is:

$$\mathbf{X}_m = (\eta^{-1} + \xi) \bar{\mathbf{x}} = \frac{1}{\rho} [X \ Y \ Z]^T$$

$$\mathbf{X}_m^* = (\eta^{*-1} + \xi) \bar{\mathbf{x}}^* = \frac{1}{\rho} [X^* \ Y^* \ Z^*]^T$$

Using the homogenous coordinates $\underline{\mathbf{X}} = [X \ Y \ Z \ H]^T$ and , we can write:

$$\rho(\eta^{-1} + \xi) \bar{\mathbf{x}} = [\mathbf{I}_3 \ 0] \underline{\mathbf{X}} = [\mathbf{R} \ \mathbf{t}] \underline{\mathbf{X}}^* \quad (8)$$

The distance $d(\mathbf{X}, \pi)$ from the world point \mathbf{X} to the plane (π) is given by the scalar product $\pi^{*T} \cdot \underline{\mathbf{X}}^*$ and:

$$d(\mathbf{X}^*, \pi^*) = \rho^* (\eta^{*-1} + \xi) \mathbf{n}^{*T} \cdot \bar{\mathbf{x}}^* - d^* H^*$$

As a consequence, the unknown homogenous component H^* is given by:

$$H^* = \frac{\rho^*(\eta^{*-1} + \xi)}{d^*} \mathbf{n}^{*T} \bar{\mathbf{x}} - \frac{d(\mathbf{X}^*, \pi^*)}{d^*} \quad (9)$$

The homogeneous coordinates of \mathbf{X} with respect to F_m^* can be rewritten as:

$$\underline{\mathbf{X}}^* = \rho^*(\eta^{*-1} + \xi) \begin{bmatrix} \mathbf{I}_3 & 0 \end{bmatrix}^T \bar{\mathbf{x}} + \begin{bmatrix} 0_{1 \times 3} & H^* \end{bmatrix} \quad (10)$$

By combining the Equations (9) and (10), we obtain:

$$\underline{\mathbf{X}}^* = \rho^*(\eta^{*-1} + \xi) \begin{bmatrix} \mathbf{I}_3 & 0 \end{bmatrix}^T \mathbf{A} \bar{\mathbf{x}} + \mathbf{b}^* \quad (11)$$

Where

$$\mathbf{A}^* = \begin{bmatrix} \mathbf{I}_3 & \frac{\mathbf{n}^*}{d^*} \end{bmatrix}^T \text{ and } \mathbf{b}^* = \begin{bmatrix} 0_{1 \times 3} & -\frac{d(\mathbf{X}, \pi)}{d^*} \end{bmatrix}^T$$

According to (11), the expression (8) can be rewritten as:

$$\rho(\eta^{-1} + \xi) \bar{\mathbf{x}} = \rho^*(\eta^{*-1} + \xi) \mathbf{H} \bar{\mathbf{x}} + \alpha \mathbf{t} \quad (12)$$

With $\mathbf{H} = \mathbf{R} + \frac{\mathbf{t}}{d} \mathbf{n}^{*T}$ and $\alpha = -\frac{d(\mathbf{X}, \pi)}{d^*}$. \mathbf{H} is the Euclidean homography matrix written as a function of the camera displacement and of the plane coordinates with respect to F_m^* . It has the same form as in the conventional perspective case (it is decomposed into a rotation matrix and a rank 1 matrix). If the world point \mathbf{X} belongs to the reference plane (π) (i.e $\alpha = 0$) then Equation (12) becomes:

$$\bar{\mathbf{x}} \propto \mathbf{H} \bar{\mathbf{x}} \quad (13)$$

Note that the Equation (13) can be turned into a linear homogeneous equation $\bar{\mathbf{x}} \times \mathbf{H} \bar{\mathbf{x}} = 0$ (where \times denotes the cross-product). As usual, the homography matrix related to (π), can thus be estimated up to a scale factor, using four couples of coordinates $(\mathbf{x}_k; \mathbf{x}_k^*)$, $k = 1 \dots 4$, corresponding to the projection in the image space of world points \mathbf{X}_k belonging to (π). If only three points belonging to (π) are available then at least five supplementary points are necessary to estimate the homography matrix by using for example the linear algorithm proposed in (Malis et al 2000). From the estimated homography matrix, the camera motion

parameters (that is the rotation \mathbf{R} and the scaled translation $\mathbf{t}_{d^*} = \frac{\mathbf{t}}{d^*}$, and the structure of the observed scene (for example the vector \mathbf{n}^*) can thus be determined (refer to (Faugeras et al 1988)). It can also be shown that the ratio $\sigma = \frac{\rho}{\rho^*}$ can be estimated as follow:

$$\sigma = \frac{\rho}{\rho^*} = (1 + \mathbf{n}^{*T} \mathbf{R} \mathbf{t}_{d^*}) \frac{(\eta^{*-1} + \xi) \mathbf{n}^{*T} \mathbf{x}}{(\eta^{-1} + \xi) \mathbf{n}^{*T} \mathbf{R}^T \mathbf{x}} \quad (14)$$

This parameter is used in our 2 1/2 D visual servoing control scheme from points.

3.2 Homography matrix from lines

Let \mathbf{L} be a 3D straight line with binormalized Euclidean Plücker coordinates $\begin{bmatrix} -\mathbf{T} & -\mathbf{T} \\ \mathbf{h} & \mathbf{u} & h \end{bmatrix}^T$ with respect to F_m and with coordinates $\begin{bmatrix} -^{*}\mathbf{T} & -^{*}\mathbf{T} & h^{*} \end{bmatrix}^T$ with respect to F_m^* . Consider that the 3D line \mathbf{L} lies in a 3D reference plane (π) as defined below. Let \mathbf{X}_1 and \mathbf{X}_2 be two points in the 3D space lying on the line \mathbf{L} . The central catadioptric projection of the 3D line \mathbf{L} is fully defined by the normal vector to the interpretation plane $\bar{\mathbf{h}}$. The vector $\bar{\mathbf{h}}$ can be defined by two points in the 3D line as $\bar{\mathbf{h}} = \frac{\mathbf{X}_1 \times \mathbf{X}_2}{\|\mathbf{X}_1 \times \mathbf{X}_2\|}$. Noticing that $[\mathbf{H}\mathbf{X}_i^*]_x = \det(\mathbf{H})\mathbf{H}^{-T}[\mathbf{X}_i^*]_x \mathbf{H}^{-1}$ ($[\mathbf{H}\mathbf{X}_i^*]_x$ being the skew-symmetric matrix associated to the vector $\mathbf{H}\mathbf{X}_i^*$) and according to (3) and (13), $\bar{\mathbf{h}}$ can be written as follow:

$$\bar{\mathbf{h}} \propto \frac{\det(\mathbf{H})\mathbf{H}^{-T}\mathbf{X}_1^* \times \mathbf{X}_2^*}{\|\mathbf{X}_1^* \times \mathbf{X}_2^*\|} \quad (18)$$

Since $\bar{\mathbf{h}}^* = \frac{\mathbf{X}_1^* \times \mathbf{X}_2^*}{\|\mathbf{X}_1^* \times \mathbf{X}_2^*\|}$ is the normal vector to the interpretation plane expressed in the frame F_m^* , the relationship between two views of the 3D line can be written as:

$$\bar{\mathbf{h}} \propto \mathbf{H}^{-T} \bar{\mathbf{h}}^* \quad (15)$$

The expression of the homography matrix in the pixel space can be derived hence using the polar lines. As depicted below, each conic, corresponding to the projection of a 3D line in the omnidirectional image, can be explored through its polar line. Let \mathbf{l}_i and \mathbf{l}_i^* be the polar lines of the image center \mathbf{O}_i with respect to the conics Ω_i and Ω_i^* respectively in the two

positions F_m and F_m^* of the catadioptric camera. From equation (6), the relationship given in equation (15) can be rewritten as:

$$\mathbf{I}_i \propto \mathbf{G}^{-T} \mathbf{I}_i^* \quad (16)$$

Where $\mathbf{G} = \mathbf{K}\mathbf{H}\mathbf{K}^{-1} = \mathbf{K}(\mathbf{R} + \frac{\mathbf{t}}{d^*} \mathbf{n}^{*T})\mathbf{K}^{-1}$. As in the case of points the homography matrix related to (π) can be linearly estimated. Equation (16) can be rewritten as: $\mathbf{I}_i \times \mathbf{G}^{-T} \mathbf{I}_i^* = 0$ and \mathbf{G} can thus be estimated using at least four couples of coordinates $((\mathbf{I}_{ik}, \mathbf{I}_{ik}^*), k=1 \dots 4)$. The homography matrix is then computed as $\mathbf{K}^{-1}\mathbf{H}\mathbf{K} = \mathbf{G}$. From \mathbf{H} , the camera motion parameters (that is the rotation \mathbf{R} and the scaled translation $\mathbf{t}_{d^*} = \frac{\mathbf{t}}{d^*}$, and the structure of the observed scene (for example the vector \mathbf{n}^*) can thus be determined. It can also be shown that the ratio $r = \frac{h}{h^*}$ (ratio of the lines depth) can be computed as follow:

$$r = \frac{h}{h^*} = (1 + \mathbf{t}_{d^*}^T \mathbf{R}^T \mathbf{n}^*) \frac{\|\mathbf{n}^{*T} \times \mathbf{K}^T \mathbf{I}_i^*\|}{\|\mathbf{R} \mathbf{n}^* \times \mathbf{K}^T \mathbf{I}_i\|} \quad (17)$$

These parameters are important since they are used in the design of our control scheme with imaged lines. In the next section, we propose a vision control scheme which allows to fully decouple rotational and translational motions.

4. Control schemes

In order to design an hybrid visual servoing scheme, the features used as input of the control law combine 2D and 3D informations. We propose to derive these informations from imaged points or polar lines and the homography matrix computed and decomposed as depicted in the last section. Let us first define the input of the proposed hybrid control scheme as follow:

$$\mathbf{s} = [\mathbf{s}_v^T \quad \mathbf{s}_\omega^T]^T \quad (18)$$

The vector \mathbf{s}_v depends of the chosen image features. The vector \mathbf{s}_ω is chosen as $\mathbf{s}_\omega = \mathbf{u}\theta$ where \mathbf{u} and θ are respectively the axis and the rotation angle extracted from \mathbf{R} (*i.e* the rotation matrix between the mirror frame when the camera is in these current and desired positions). The task function \mathbf{e} to regulate to $\mathbf{0}$ is then given by:

$$\mathbf{e} = [\mathbf{s} - \mathbf{s}^*] = \begin{bmatrix} \mathbf{s}_v - \mathbf{s}_v^* \\ \mathbf{s}_\omega - \mathbf{s}_\omega^* \end{bmatrix} = \begin{bmatrix} \mathbf{s}_v - \mathbf{s}_v^* \\ \mathbf{u}\theta \end{bmatrix} \quad (19)$$

where \mathbf{s}^* is the desired value of \mathbf{s} . Note that the rotational part of the task function can be estimated using partial Euclidean reconstruction from the homography matrix derived in Section 3). The exponential convergence of \mathbf{e} can be obtained by imposing $\dot{\mathbf{e}} = -\lambda \mathbf{e}$, the corresponding control law is:

$$\boldsymbol{\tau} = -\lambda \mathbf{L}^{-1}(\mathbf{s} - \mathbf{s}^*) \quad (20)$$

where $\boldsymbol{\tau} = [\mathbf{v}^T \ \boldsymbol{\omega}^T]^T$ is the central catadioptric camera velocity (\mathbf{v} and $\boldsymbol{\omega}$ denote respectively the linear and angular velocities), λ tunes the convergence rate and \mathbf{L} is the interaction matrix which links the variation of feature vector \mathbf{s} to the camera velocity $\dot{\mathbf{s}} = \mathbf{L}\boldsymbol{\tau}$.

The time derivative of $\mathbf{s}_\omega = \mathbf{u}\theta$ can be expressed as a function of the camera velocity as:

$$\frac{d(\mathbf{u}\theta)}{dt} = [0_3 \ \mathbf{L}_\omega] \boldsymbol{\tau}$$

Where \mathbf{L}_ω is given in (Malis et al 1999):

$$\mathbf{L}_\omega = \mathbf{I}_3 - \frac{\theta}{2} [\mathbf{u}]_\times + \left(1 - \frac{\sin(\theta)}{\sin^2\left(\frac{\theta}{2}\right)} \right) [\mathbf{u}]_\times^2 \quad (21)$$

Where $\text{sinc}(\theta) = \frac{\sin(\theta)}{\theta}$ and $[\mathbf{u}]_\times$ being the antisymmetric matrix associated to the rotation axis \mathbf{u} .

4.1 Using points to define \mathbf{s}_v

To control the 3 translational degrees of freedom, the visual observations and the ratio σ expressed in (14) are used:

$$\mathbf{s}_v = [x \ y \ \delta]^T \quad (22)$$

Where x and y are the current coordinates of a chosen catadioptric image point given by Equation (1), $\delta = \log(\rho)$. The translational part of the task function is thus:

$$\mathbf{e} = \mathbf{s}_v - \mathbf{s}_v^* = [x - x^* \ y - y^* \ \Gamma]^T \quad (23)$$

Where $\Gamma = \log\left(\frac{\rho}{\rho^*}\right) = \log(\sigma)$. The first two components of $\mathbf{s}_v - \mathbf{s}_v^*$ are computed from the normalized current and desired catadioptric images, and its last components can be estimated using Equation (14).

Consider a 3-D point \mathbf{X} , lying on the reference plane π , as the reference point. The time derivative of its coordinates, with respect to the current catadioptric frame F_{nv} is given by:

$$\dot{\mathbf{X}} = \begin{bmatrix} -\mathbf{I}_3 & [\mathbf{X}]_{\times} \end{bmatrix} \tau \quad (24)$$

The time derivative of \mathbf{s}_v can be written as:

$$\dot{\mathbf{s}}_v = \frac{\partial \mathbf{s}_v}{\partial \mathbf{X}} \dot{\mathbf{X}} \quad (25)$$

With:

$$\frac{\partial \mathbf{s}_v}{\partial \mathbf{X}} = \frac{1}{\rho(Z+\xi\rho)^2} \begin{bmatrix} \rho Z + \xi(Y^2 + Z^2) & \xi XY & -X(\rho + \xi Z) \\ -\xi XY & \rho Z + \xi(X^2 + Z^2) & -Y(\rho + \xi Z) \\ \frac{X(Z + \xi\rho)^2}{\rho} & \frac{Y(Z + \xi\rho)^2}{\rho} & \frac{Z(Z + \xi\rho)^2}{\rho} \end{bmatrix}$$

By combining the equations (24), (25) and (14), it can be shown that:

$$\dot{\mathbf{s}}_v = [\mathbf{A} \quad \mathbf{B}] \tau \quad (26)$$

With

$$\mathbf{A} = \frac{1}{\rho\rho^*} \begin{bmatrix} -\frac{1+x^2(1-\xi(\gamma_x+\xi))+y^2}{\gamma_x+\xi} & \xi xy & x\gamma_x \\ \xi xy & -\frac{1+y^2(1-\xi(\gamma_x+\xi))+x^2}{\gamma_x+\xi} & y\gamma_x \\ \eta_x x & \eta_x y & (\eta_x - 1)\xi \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} xy & \frac{(1+x^2)\gamma_x - \xi y^2}{\gamma_x + \xi} & y \\ \frac{(1+y^2)\gamma_x - \xi x^2}{\gamma_x + \xi} & -xy & -x \\ 0 & 0 & 0 \end{bmatrix}$$

Where: $\gamma_x = \sqrt{1+(1-\xi^2)(x^2+y^2)}$ and $\eta_x = \frac{\xi^2 + \sqrt{(1-\xi^2)(x^2+y^2)+\xi^2}}{x^2+y^2+\xi^2}$. The task function \mathbf{e} (see Equation (19)) can thus be regulated to $\mathbf{0}$ using the control law (Equation (20)) with the following interaction matrix \mathbf{L} :

$$\mathbf{L} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ 0_3 & \mathbf{L}_\omega \end{bmatrix} \quad (27)$$

In practice, an approximated interaction matrix $\hat{\mathbf{L}}$ is used. The parameter ρ^* can be estimated only once during an off-line learning stage.

4.2 Using imaged lines to define \mathbf{s}_v

To control the 3 translational degrees of freedom with imaged lines, the chosen visual observation vector is:

$$\mathbf{s}_v = [\log(h_1) \quad \log(h_2) \quad \log(h_3)]^T \quad (28)$$

Where h_1 , h_2 and h_3 are the depth of three co-planar lines. From the time derivative of the line depth expressed as a function of the camera velocity (Andreff et al 2002), given by $\dot{h}_k = (\bar{\mathbf{u}}_k \times \bar{\mathbf{h}}_k)^T \mathbf{v}$, it can be shown that:

$$\frac{d(\log(h_k))}{dt} = \left[\frac{1}{h_k} (\bar{\mathbf{u}}_k \times \bar{\mathbf{h}}_k)^T \quad 0_3 \right] \mathbf{v} \quad (29)$$

According to (6) and (29), the time derivative of the vector \mathbf{s}_v is thus given by:

$$\dot{\mathbf{s}}_v = [\mathbf{L}_v \quad 0_3]^T \boldsymbol{\tau}$$

Where:

$$\mathbf{L}_v = \begin{bmatrix} \|\mathbf{K}^T \mathbf{l}_{i1}\| h_1 & 0 & 0 \\ 0 & \|\mathbf{K}^T \mathbf{l}_{i2}\| h_2 & 0 \\ 0 & 0 & \|\mathbf{K}^T \mathbf{l}_{i3}\| h_3 \end{bmatrix} \begin{bmatrix} (\bar{\mathbf{u}}_1 \times \mathbf{K}^T \mathbf{l}_{i1})^T \\ (\bar{\mathbf{u}}_2 \times \mathbf{K}^T \mathbf{l}_{i2})^T \\ (\bar{\mathbf{u}}_3 \times \mathbf{K}^T \mathbf{l}_{i3})^T \end{bmatrix} \quad (30)$$

Note that the time derivative of \mathbf{s}_v does not depend of the camera angular velocity. It is also clear that \mathbf{L}_v is singular only if the principal point \mathbf{M} of the mirror frame lies in the 3D reference plane (π). The task function \mathbf{e} can thus be regulated to zero using the control law (20) with the following square block-diagonal interaction matrix:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_v & 0 \\ 0 & \mathbf{L}_\omega \end{bmatrix} \quad (31)$$

As can be seen on equation (30), the unknown depth h_i and the unitary orientations \mathbf{u}_i with respect to the catadioptric camera frame have to be introduced in the interaction matrix. Noticing that $\bar{\mathbf{u}}_i = (\bar{\mathbf{h}}_i \times \mathbf{R}\bar{\mathbf{h}}_i^*) / \|\bar{\mathbf{h}}_i \times \mathbf{R}\bar{\mathbf{h}}_i^*\|$ and using equation (6), the orientation can be estimated as follow:

$$\bar{\mathbf{u}} = \frac{\mathbf{K}^T \mathbf{I}_i \times \mathbf{R} \mathbf{K}^T \mathbf{I}_i^*}{\|\mathbf{K}^T \mathbf{I}_i \times \mathbf{R} \mathbf{K}^T \mathbf{I}_i^*\|}$$

Furthermore, if the camera is calibrated and \hat{h}_i is chosen to approximate h_i , then it is clear that $\hat{\mathbf{L}}_v^{-1} \mathbf{L}_v$ is a diagonal matrix with $\frac{\hat{h}_i}{h_i}$ for $i=1, 2, 3$ as entries. The only point of equilibrium is thus \mathbf{s}^* and the control law is asymptotically stable in the neighbourhood of \mathbf{s}^* if \hat{h}_i is chosen positive. In practice, an approximated matrix $\hat{\mathbf{L}}^{*-1}$ at the desired position is used to compute the camera velocity vector and the rotational part of the interaction matrix can be set to $\mathbf{L}_\omega^{-1} = \mathbf{I}_3$ (Malis et al 1999). Finally, the control law is thus given by:

$$\boldsymbol{\tau} = -\lambda \begin{bmatrix} \hat{\mathbf{L}}_v^{*-1} & 0 \\ 0 & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} \mathbf{s}_v - \mathbf{s}_v^* \\ \boldsymbol{\theta} \mathbf{u} \end{bmatrix} \quad (32)$$

5 Results

5.1 Simulation Results

We present now results concerning a positioning task of a six degrees of freedom robotic arm with a catadioptric camera in eye-in-hand configuration. The catadioptric camera used is an hyperbolic mirror combined with a perspective camera (similar results are obtained with a catadioptric camera combining a parabolic mirror and an orthographic lens, these results are not presented in this paper). From an initial position, the catadioptric camera has to reach the desired position. This means that the task function (refer to equation (19)), computed from the homography matrix between the current and desired images, converges to zero. To be close to a real setup, image noise has been added (additive noise with maximum amplitude of 2 pixels) to the image and the interaction matrix is computed using erroneous internal camera parameters. The first simulation concerns imaged points while the second simulation concerns imaged lines.

5.1.a Imaged points

The initial and desired attitudes of the catadioptric camera are plotted in the Figure 3. This figure also shows the 3-D camera trajectory from its initial position to the desired one. Figure 3 shows the initial (blue *) and desired (red *) images of the observed target. It shows

also the trajectory of the point (green trace) in the image plane (the controlled image point has a black trace trajectory). The norm of the error vector is given in Figure 4(b). As can be seen in the Figures 4(c) and 4(d) showing the errors between desired and current observation vectors the task is correctly realized. The translational and rotational camera velocities are given in Figures 4(e) and 4(f) respectively.

5.1.b Imaged lines

Figure 2 shows the spatial configuration of the 3D lines as well as the 3D trajectory of the central catadioptric. The images corresponding to the initial and desired positions are shown by figures 5(c) and 5(d). These figures show the projected 3D lines (conics) and the associated polar lines. The trajectories of the conics and of the corresponding polar lines in the image plane are given in Figures 5(a) and 5(b) respectively. These trajectories confirm that the initial images (conics and polar lines) reach the desired images. Figures 5(e) and 5(f) show respectively the translational and rotational velocities of the catadioptric camera. As shown in Figures 5(g) and 5(h), the error vector \mathbf{e} between the current and desired observation vectors are well regulated to zeros, and thus the positioning task is correctly realized.

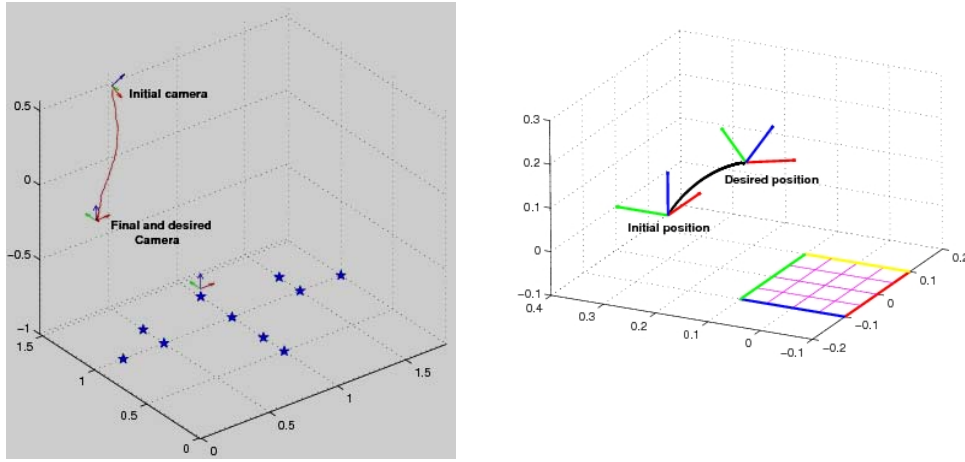


Fig. 3 : 3D trajectories of the catadioptric camera [meters]: (left) the case of points, (right) the case of lines

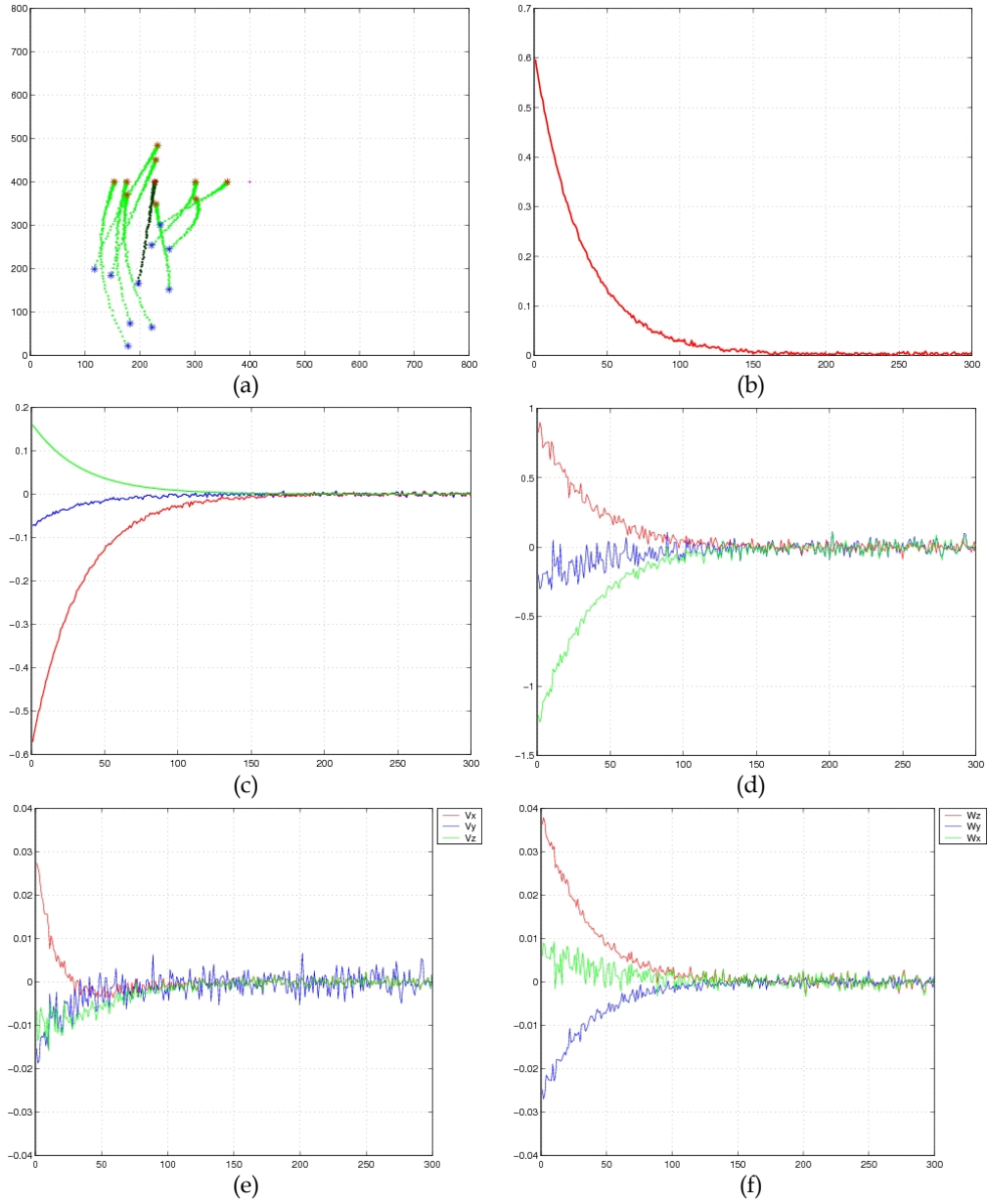


Fig. 4: (a) Trajectories in the image of the target points [pixels]. (b) norm of the error vector, (c) error vector: [meters], (d) rotation vector [rad], (e) Translational velocity [m/s], (f) rotational velocity [rad/s].

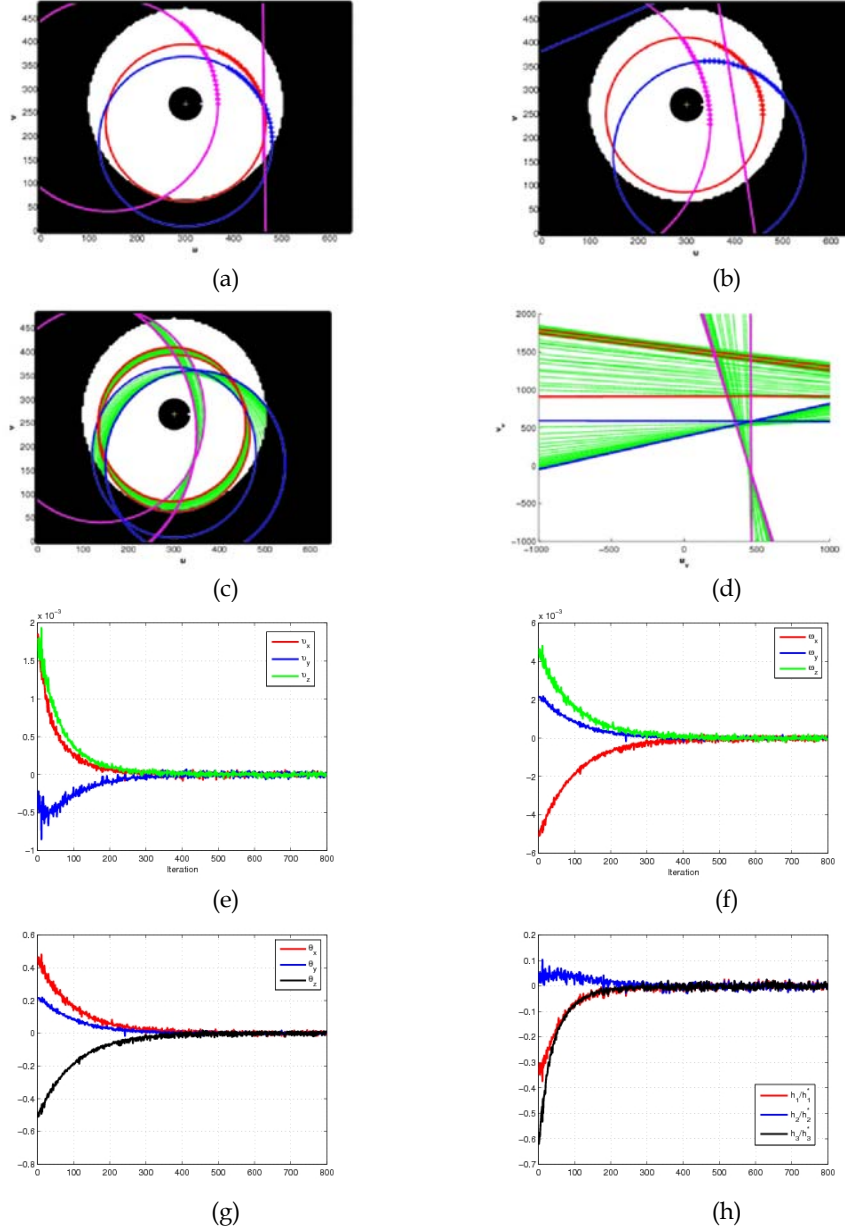


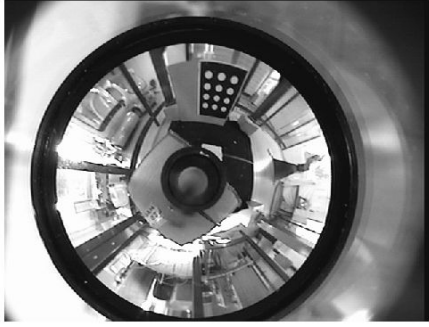
Fig. 5: Visual servoing with para-catadioptric camera: (a) initial image, (b) desired image (c) trajectory of the conics in the image plane, (d) trajectory of the polar line, (e) translation velocities [m/s], (f) rotational velocities [rad/s], (g) $\mathbf{u}\theta$ errors [rad], (h) $\mathbf{s}_v - \mathbf{s}_v^*$ vector errors

5.1 Experimental Results

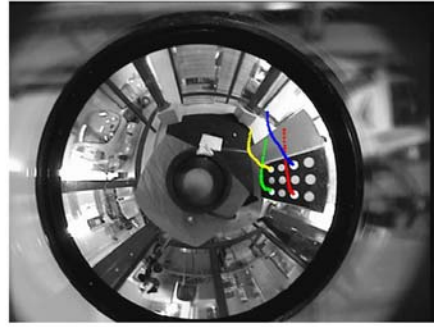
The proposed control law has been validated on a six d-o-f eye-to-hand system (refer to Fig. 6). Since we were not interested in image processing in this paper, the target is composed of white marks (see Figure 6) from which points or straight lines can be defined. The coordinates of these points (the center of gravity of each mark) are extracted and tracked using the VISP library (Marchand et al 2005). The omnidirectional camera used is a parabolic mirror combined with an orthographic lens ($\xi=1$). Calibration parameters of the camera are: $f.(\phi-\xi)=161$ and the coordinates of the principal point are $[300\ 270]^T$. From an initial position the robot has to reach a desired position known as a desired 2D observation vector \mathbf{s}^* . Two experiments are presented. In the first one whose results are depicted in Fig. 7, the point-based visual servoing has been used.. The error on the visual features is plotted on Fig. 7(e) while the camera velocities are plotted on Fig. 7(c)-(d). These results confirm that the positioning task is correctly achieved. The second experiment has been conducted using the line-based visual servoing. The corresponding results are depicted on Fig. 8. We can note that the system still converges.



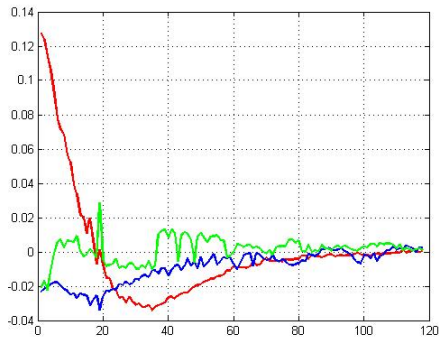
Fig. 6: Experimental setup : eye-to-hand configuration



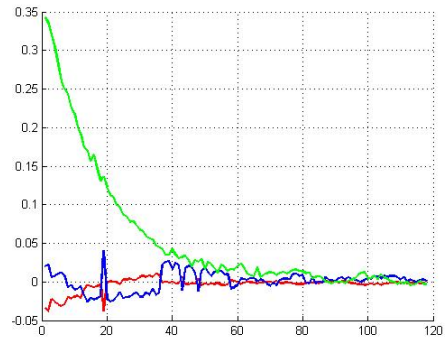
(a)



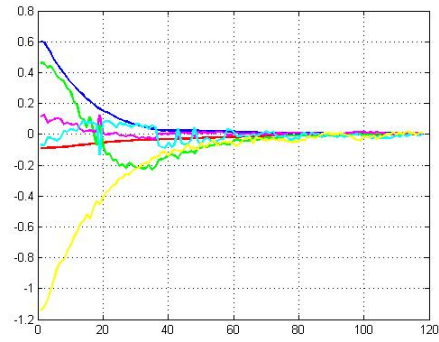
(b)



(c)

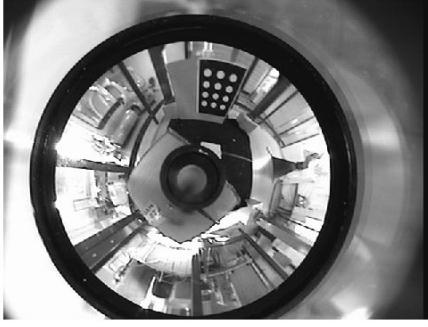


(d)

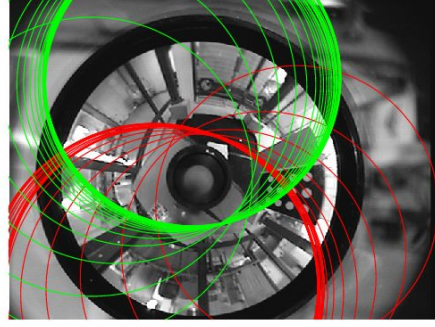


(e)

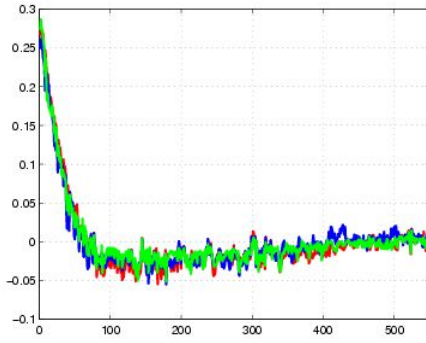
Fig. 7 Visual servoing with lines: (a) initial image, (b) desired image and trajectory of the conics in the image plane (c) translational velocities [m/s], (d) rotational velocities [rad/s], (e) $\mathbf{s}_v - \mathbf{s}_v^*$ and $u\theta$ errors.



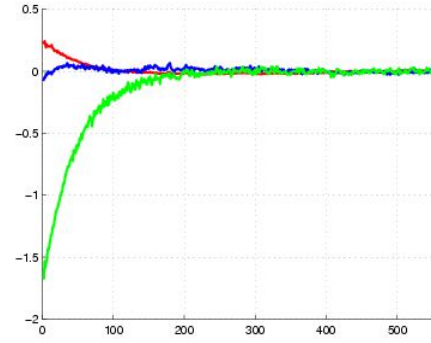
(a)



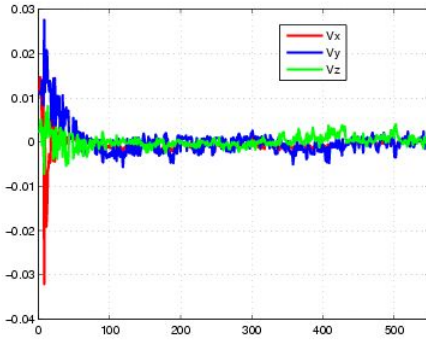
(b)



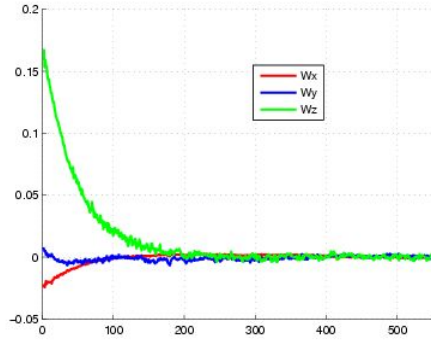
(c)



(d)



(e)



(f)

Fig. 8 Visual servoing with lines: (a) initial image, (b) desired image and trajectory of the conics in the image plane, (c) $s_v - s_v^*$, (d) $u\theta$ errors [rad] (e) translational velocities [m/s], (f) rotational velocities [rad/s]

6. Conclusion

In this paper hybrid vision-based control schemes valid for the entire class of central cameras was presented. Geometrical relationship between imaged points and lines was exploited to estimate a generic homography matrix from which partial Euclidean reconstruction can be obtained. The information extracted from the homography matrix were then used to design vision-based control laws. Results with simulated data confirmed the relevance. In future work, the robustness and stability analysis with respect to calibration errors must be studied.

7. References

- S. Baker & S. K. Nayar (1999). A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):1– 22, November 1999.
- J. Barreto and H. Araujo (2002). Geometric properties of central catadioptric line images. In 7th European Conference on Computer Vision, ECCV'02, pages 237–251, Copenhagen, Denmark, May 2002.
- R. Benosman & S. Kang (2000). *Panoramic Vision*. Springer Verlag ISBN 0387-95111-3, 2000.
- P. Blaer & P.K. Allen (2002). Topological mobile robot localization using fast vision techniques. In *IEEE International Conference on Robotics and Automation*, pages 1031–1036, Washington, USA, May 2002.
- G. Chesi, K. Hashimoto, D. Prattichizzo & A. Vicino (2003). A switching control law for keeping features in the field of view in eye-in-hand visual servoing. In *IEEE International Conference on Robotics and Automation*, pages 3929–3934, Taipei, Taiwan, September 2003.
- B. Espiau, F. Chaumette & P. Rives (1992). A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313–326, June 1992.
- S. Hutchinson, G.D. Hager, and P.I. Corke (1996). A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5):651–670, October 1996.
- E. Malis, F. Chaumette, & S. Boudet (1999). 2 1/2 d visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):238–250, April 1999.
- Y. Mezouar & F. Chaumette (2002). Path planning for robust image-based control. *IEEE Transactions on Robotics and Automation*, 18(4):534–549, August 2002.
- Y. Mezouar, H. Haj Abdelkader, P. Martinet, & F. Chaumette (2004). Central catadioptric visual servoing from 3d straight lines. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'04*, volume 1, pages 343–349, Sendai, Japan, September 2004.
- E. Malis & S. Benhimane (2003). Vision-based control with respect to planar and non-planar objects using a zooming camera. In *IEEE International Conference on Advanced Robotics*, pages 863–869, July 2003.
- R. Vidal, O. Shakernia, & S. Sastry (2003). Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation. In *IEEE International Conference on Robotics and Automation*, pages 584–589, Taipei, Taiwan, September 2003.
- N. Winter, J. Gaspar, G. Lacey, & J. Santos-Victor (2000). Omnidirectional vision for robot

navigation. In Proc. IEEE Workshop on Omnidirectional Vision, OMNIVIS, pages 21–28, South Carolina, USA, June 2000.