Proceedings of the 2007 IEEE/RSJ International
Conference on Intelligent Robots and Systems
San Diego, CA, USA, Oct 29 - Nov 2, 2007

A Generic Fisheye camera model for robotic applications

Jonathan Courbon*†, Youcef Mezouar*, Laurent Eck†, Philippe Martinet*‡

*LASMEA
24 Avenue des Landais
63177 AUBIERE - FRANCE

†CEA, List
18 route du Panorama, BP6
F- 92265 FONTENAY AUX ROSES - FRANCE
Email: laurent.eck@cea.fr

‡Sungkyunkwan University
Intelligent Systems Research Center
SUWON, SOUTH KOREA

Email: firstname.lastname@lasmea.univ-bpclermont.fr

Abstract— **Omnidirectional cameras have a wide field of view and are thus used in many robotic vision tasks. An omnidirectional view may be acquired by a fisheye camera which provides a full image compared to catadioptric visual sensors and do not increase the size and the weakness of the imaging system with respect to perspective cameras. We prove that the unified model for catadioptric systems can model fisheye cameras with distortions directly included in its parameters. This unified projection model consists on a projection onto a virtual unitary sphere, followed by a perspective projection onto an image plane. The validity of this assumption is discussed and compared with other existing models. Calibration and partial Euclidean reconstruction results help to confirm the validity of our approach. Finally, an application to the visual servoing of a mobile robot is presented and experimented.**

I. INTRODUCTION

A camera with a large field of view is usefull in many robotic applications including robot localisation, navigation and visual servoing. Conventional cameras are generally seen as perspective tools (pinhole model) which are convenient for modelisation and algorithmic design. Morever they exhibit small distorsions and thus acquired images can easily be interpreted. Unfortunately, conventional cameras suffer from a restricted field-of-view (Fig.1 (a)). When a larger field of view is required omnidirectional cameras can be employed (Fig.1 (b) and (c)). In the literature, several methods have been proposed to increase the field of view [1]. The first one consists on using a moving camera or a system made of multiple cameras. In this case, the omnidirectional view is computed from a set of images which is generally not compatible with real time applications. A second solution consists on combining conventional camera and mirrors. The obtained sensors are refered to catadioptric imaging system. Unfortunately, those cameras exhibit generally a large dead area in the center of the image (Fig.1 (c)) which can be a huge drawback. Such sensors also have the drawback of requiring a mirror which significantly increases the size and the weakness of the imaging system. Note that the resulting imaging systems have been termed central catadioptric when a single projection center describes the world-image mapping [2]. In [3], a projection model valid for the entire class of central catadioptric camera has been proposed. According to this generic model, all central catadioptric cameras can be modelled by a central projection onto a sphere followed by a central projection onto the image plane.

(a) (b) (c)

Fig. 1. Three images acquired approximately at the same position, with (a) a $1/4''$ image sensor with a perspective lens (b) the same image sensor with an Omnitech Robotics ORIFL190-3 fisheye lens with a field of view of 190° and (c) a catadioptric camera.

The last class of cameras with a wide field-of-view are the *dioptric systems* (fisheye cameras) [2]. A fisheye camera is an imaging system combining a fisheye lens and a conventional camera. Their main advantages with respect to catadioptric sensors is firstly that they do not exhibit dead area (see Fig.1 (b)) and secondly, a fisheye lens do not increase the size and the weakness of the imaging system with respect to a conventional camera. However, whereas a unified model exists for central catadioptric cameras [3], many different models are used for fisheye cameras. Distortions are generally splitted into two main components (namely tangential and radial distortions). Tangential distortions are generally negligible compared to radial distortions and thus not considered in this paper.

The radial distortions models for fisheye cameras can be classified into three main groups. The first group is based on the pinhole model. A 3D point is first perspectively projected into the image plane (pinhole model), then, a radial distortion is applied to the projected point to obtain the distorted image point. The second category is based on the captured rays. It consists on defining a mapping between the incidence ray direction and the distance between the image point and the image center. The last category is based on the unified catadioptric camera model. It is motivated by the fact that similar behaviors have been observed for those sensors (see Fig.1 (b), (c)). In [4], a 3rd order polynomial model of radial distortion is applied to the projected point obtained with the unified model. In [5], the authors propose to extend the

1-4244-0912-8/07/$25.00 ©2007 IEEE.

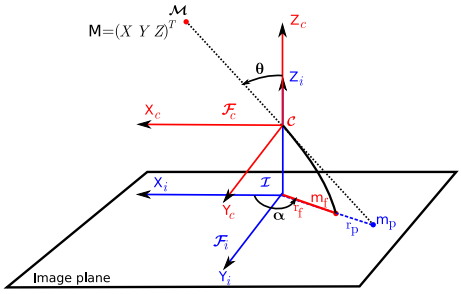

Fig. 2. Perspective and fisheye imaging process. The perspective image of a 3D point $\mathcal{M}$ is m_p and its fisheye image is m_f.

unified model to include fisheye camera by substituing the projection on a sphere by a projection on a quadric surface. In [6], the division model of radial distortions proposed in [7] is integrated in the unified model for central catadioptric cameras.

In this paper, we show that the unified model for catadioptric cameras via a spherical projection is equivalent to a pinhole-based and to a captured rays-based models with radial distortions and thus that it can be directly employed to model fisheye cameras. As a consequence, all existing algorithms for catadioptric cameras (calibration, visual servoing, ...) are also usefull for fisheye cameras. In Section II, the existing models for fisheye cameras are detailled. Then, the proposed model is presented in Section III and validated in Section IV through comparison with several existing models, experimental results of calibration, and partial Euclidean reconstruction. Finally, the Section V reports experimental results in the context of mobile robot vision-based navigation.

II. Fisheye camera models

Let $\mathcal{F}_c$ and $\mathcal{F}_i$ be the frames attached to the camera with origin $\mathcal{C}$ located at the projection center and to the image plane with origin $\mathcal{I}$ located at the principal point respectively. Let $\mathcal{M}$ be a 3D point of coordinates $\mathsf{M} = (X\,Y\,Z)^T$ with respect to the frame $\mathcal{F}_c$. The perpective image and the fisheye image coordinates with respect to $\mathcal{F}_i$ of $\mathcal{M}$ are $\mathsf{m}_p = (x_p \quad y_p)^T$ and $\mathsf{m}_f = (x_f \quad y_f)^T$ respectively. The distance between the principal axis and the image point is denoted by r, the angle between the incoming ray and the principal axis by θ and the angle between the X-axis and m by α (refer to Figure 2) . The subscripts p, f and c denote respectively perspective, fisheye and catadioptric cameras. For a perspective camera, given the focal length f, r_p satisfies:

$$r_p(\theta) = f \tan\theta \tag{1}$$

For a fisheye camera, radial distortions have to be taken into account. Many models of radial distortions exist but they have all to respect the two following fundamental constraints :

Constraint 1: the ray arriving along the principle axis is not deformed: $r_f(0) = 0$.

Constraint 2: the radius $r_f(k)$ is monotonically increasing for $k > 0$.

$$\begin{aligned}
r_f^1(r_p) &= r_p L(r_p, n)\\
r_f^2(r_p) &= \frac{r_p}{L(r_p, n)}\\
r_f^3(r_p) &= \frac{r_p}{1 + k_1 r_p^2}\\
r_f^4(r_p) &= r_p \frac{L_1(r_p, n_1)}{L_2(r_p, n_2)}\\
r_f^5(r_p) &= s \log(1 + \lambda r_p)
\end{aligned}$$

TABLE I
Pinhole-based models

$$\begin{aligned}
r_f^1(\theta) &= f\theta\\
r_f^2(\theta) &= 2f \tan\left(\frac{\theta}{2}\right)\\
r_f^3(\theta) &= f \sin\theta\\
r_f^4(\theta) &= f \sin\left(\frac{\theta}{2}\right)\\
r_f^5(\theta) &= f(k_1\theta + k_2\theta^3 + \cdots + k_n\theta^{2(n-1)+1})
\end{aligned}$$

TABLE II
Captured rays-based models

Let us now describe the two models' families classically employed.

A. Pinhole-based models

The first model is based on the pinhole projection. The coordinates of the image point m_f can be obtained thanks to a mapping T_1 linking the fisheye and the perspective radii:

$$r_p \xrightarrow[T_1]{} r_f$$

As the equation (1) is not defined for $\theta = \pi/2$, the models based on the pinhole projection are not defined too.

The mapping T_1 may be defined by several ways as detailed in Table I. In [8], radial distortions are described by a polynomial model $r_f^1(r_p)$ where: $L(r_p, n) = 1 + k_1 r_p^2 + k_2 r_p^4 + ... + k_n r_p^{2n}$. k_i, $(i = 1, 2, ..n)$ are the n parameters of the model. In practice, this model fits to lenses with small distortions but many parameters are needed for lenses with larger distortions. The division model $r_f^2(r_p)$ suggested in [7] allows to handle high distortion with few parameters. This model may be used with only one parameter $r_f^3(r_p)$ [7]. A rational model $r_f^4(r_p)$ is proposed in [9] with $n_1 + n_2$ parameters. A logarithmic mapping $r_f^5(r_p)$ is used in [10].

B. Captured rays-based models

The second group of models is based on the captured rays. It consists on defining a mapping T_2 between the fisheye radius and the incidence angle:

$$\theta \xrightarrow[T_2]{} r_f$$

For perspective cameras, Equation (1) maps the incidence angle θ to the radius r thanks to the focal length f. For fisheye cameras, this mapping is no more valid. In the sequel, we briefly review T_2-mappings proposed to take into account distortions (see Table II).
The f-theta mapping or equiangular or equidistance projection $r_f^1(\theta)$ proposed in [11] , is suitable for cameras with limited distortions. The stereographic projection $r_f^2(\theta)$ proposed in [12] preserves circularity and thus project 3D local symmetries onto 2D local symmetries. The orthogonal or sine law projection $r_f^3(\theta)$ is described in [13]. In [14], the equisolid angle projection function $r_f^4(\theta)$ is proposed. The polynomial function $r_f^5(\theta)$ proposed in [15] allow to improve the accuracy of a polynomial model at the cost of increasing the number of parameters.

III. Modeling fisheye camera via a projection on a sphere

In this section, first the unified model is detailled. Then, the equivalence of this model to distortion-based models (pinhole-based and captured rays-based models) is demonstrated.

A. World-image mapping

As outlined previously the unified projection model consists on a 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.
Let $\mathcal{F}_c$ and $\mathcal{F}_m$ be the frames attached to the conventional camera and to the unitary sphere respectively. In the sequel, we suppose that $\mathcal{F}_c$ and $\mathcal{F}_m$ are related by a simple translation along the Z-axis ($\mathcal{F}_c$ and $\mathcal{F}_m$ have the same orientation as depicted in Figure 3). The origins $\mathcal{C}$ and $\mathcal{M}$ of $\mathcal{F}_c$ and $\mathcal{F}_m$ will be termed optical center and principal projection center respectively. The optical center $\mathcal{C}$ has coordinates $[0\ 0\ -\xi]^T$ with respect to $\mathcal{F}_m$ and the image plane is orthogonal to the Z-axis and it is located at a distance $Z = f_c$ from $\mathcal{C}$.
Let $\mathcal{X}$ be a 3D point with coordinates $\mathsf{X} = [X\ Y\ Z]^T$ with respect to $\mathcal{F}_m$. The world point $\mathcal{X}$ is projected in the image plane into the point of homogeneous coordinates $\mathsf{m} = [x\ y\ 1]^T$. The image formation process can be split into three steps (Fig.3):

- First step: The 3D world point $\mathcal{X}$ is first projected on the unit sphere surface into a point of coordinates $\mathsf{X_m}$ in $\mathcal{F}_m$: $\mathsf{X_m} = \mathsf{X}/\rho$ where $\rho = \|\mathsf{X}\| = \sqrt{X^2+Y^2+Z^2}$. The projective ray $\mathsf{X_m}$ passes through the principal projection center $\mathcal{M}$ and the world point $\mathcal{X}$.

- Second step: The point $\mathcal{X}_m$ 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 $\underline{\mathsf{x}} = [\mathsf{x}^T\ 1]^T = f(\mathsf{X})$ (where $\mathsf{x} = [x\quad y]^T$):

$$\underline{\mathsf{x}} = f(\mathsf{X}) = \begin{bmatrix} \frac{X}{Z+\xi\rho} & \frac{Y}{Z+\xi\rho} & 1 \end{bmatrix}^\top \tag{2}$$

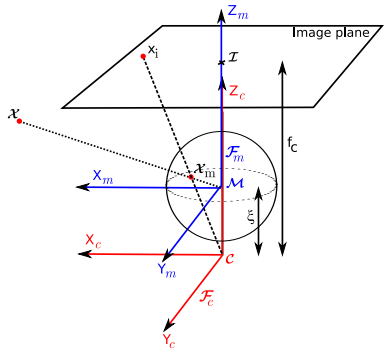

Fig. 3. Unified model for catadioptric camera.

Note that, setting $\xi = 0$, the general projection model becomes the well known pinhole model. We will see in the sequel that ξ can be seen as a parameter which allows to control the amount of radial distortions.

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

$$\mathsf{m} = \mathbf{K}\underline{\mathsf{x}} \tag{3}$$

The matrix $\mathbf{K}$ can be written as $\mathbf{K} = \mathbf{K_p}\mathbf{M}$ where the matrix $\mathbf{K_p}$ contains the perspective camera intrinsic parameters, and the diagonal matrix $\mathbf{M}$ links the frame attached to the unitary sphere to the camera frame $\mathcal{F}_m$. In the sequel, we assume that the intrinsic matrix is given by: $\mathbf{K} = \begin{pmatrix} f_c & 0 & 0 \\ 0 & f_c & 0 \\ 0 & 0 & 1 \end{pmatrix}$ where $f_c > 0$. When a pinhole camera is considered f_c will be chosen as $f_c = f_p$ and when a fisheye camera is considered as $f_c = f_f$. The perspective projection m_p of the 3D point $\mathcal{X}$ is obtained setting $\xi = 0$:

$$\begin{cases} x_p = f_p X/Z \\ y_p = f_p Y/Z \end{cases} \tag{4}$$

We will now show that the model whose final results is given by Equation (3) is equivalent to pinhole-based model or captured rays-based model.

B. Equivalence to pinhole-based model

By extracting the values of X and Y in (2) and (4), elevating squarely and then taking the square root, we obtain:

$$\frac{Z+\xi\rho}{f_f}\underbrace{\sqrt{x_f^2+y_f^2}}_{r_f} = \frac{Z}{f_p}\underbrace{\sqrt{x_p^2+y_p^2}}_{r_p} \tag{5}$$

Noticing that $Z > 0$, we can write: $\rho = Z\sqrt{\left(\frac{X}{Z}\right)^2+\left(\frac{Y}{Z}\right)^2+1}$. Thanks to (4), we thus have the relation:

$$\rho = \frac{Z}{f_p}\sqrt{r_p^2+f_p^2} \tag{6}$$

Then, from (5) we deduce:

$$\boxed{r_f = r_f(r_p) = \frac{\frac{f_f}{f_p}r_p}{1+\xi\sqrt{\frac{r_p^2}{f_p^2}+1}}} \tag{7}$$

Equation (7) is a T_1-mapping linking the perspective radius r_p and the fisheye radius r_f. It can be easily verified that Constraint 1 is respected. Furthermore, since f_f, f_p and ξ are positive, Constraint 2 is respected too.

C. Equivalence to captured rays-based model

Using the expressions of x and y given in (2) and noticing that $Z + \xi\rho > 0$, we can write:

$$r_f = \sqrt{x_f^2 + y_f^2} = \frac{f_f}{Z + \xi\rho}\sqrt{X^2 + Y^2}$$

Given that $\tan\theta = \frac{\sqrt{X^2+Y^2}}{Z}$, from the previous equation we deduce:

$$r_f = \frac{f_f}{1 + \xi\frac{\rho}{Z}}\tan\theta \tag{8}$$

Noticing that ρ can be written as a function of $\tan\theta$ and Z: $\rho = Z\sqrt{\tan^2\theta + 1}$, Equation (8) can be rewritten:

$$\boxed{r_f = r_f(\theta) = \frac{f_f \tan\theta}{1+\xi\sqrt{\tan^2\theta+1}}} \tag{9}$$

Equation (9) is clearly a T_2-mapping linking the radius r_f and the incidence angle θ. Once again, it can be easily verified that Constraints 1 and 2 are respected as f_f and ξ are positive scalars. Note that as usual with captured rays-based model, our model is not valid when $\theta = \pi/2$.
The unified model has been shown to be theoretically equivalent to the mappings T_1 and T_2 which implies that it can be used to model fisheye cameras. In the next section, the validition of the model is experimentally demonstrated.

IV. Validation

In order to validate the proposed model, firstly, the parameters f_f and ξ corresponding to the Omnitech Robotics ORIFL190-3 are computed from the captured rays model provided by the camera producer and the model (9). Secondly, we show that our model is able to fit almost all existing fisheye cameras as seen in Fig.5 and Fig.6. We then calibrate our camera using the method designed for central catadioptric cameras proposed in [4]. The obtained intrinsic parameters are finally exploited for partial Euclidean reconstruction.
Let us first define the cost functions we will use in the sequel for non-linear least square optimizations: $C(\theta) = \sum_j \|r_f^{ourmodel}(\theta_j) - r_f^i(\theta_j)\|$ and $C(r_p) = \sum_j \|r_f^{ourmodel}(r_{p_j}) - r_f^i(r_{p_j})\|$.

A. Fitting the producer data

The data of the captured ray model provided by the producer of the ORIFL190-3 of Omnitech Robotics are given as a curve representing the function $r = r(\theta)$ (see Fig.4). In order to estimate the parameters ξ and f_f of the model (9), a set of points on the producer curve is used ($r^{provider}(\theta_i)$) and $\min_{\xi,f_f} C(\theta)$ is solved. The parameters f_f and ξ have been estimated respectively at $4.823\,mm$ and 2.933 with a residual $r = 4.10^{-4}$. As it can be seen in Fig.4, the proposed model fit well with the producer's data.

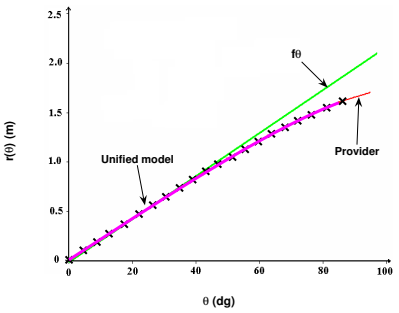

Fig. 4. The model (9) plotted in plain line fits well with the producer data (crosses).

B. Equivalence with the existing models

First, let us compare the proposed model with different T_1-mappings (pinhole-based models). Let us consider the models $r_f^1(r_p)$, $r_f^3(r_p)$ and $r_f^5(r_p)$ (refer to Table I) whose parameters have been chosen aleatory. In order to estimate the parameters f_f, f_p and ξ (model (7)), we use a set of points $r_f^i(r_{p_j})$ and we solve $\min_{\xi,f_f,f_p} C(r_p)$. As it can be

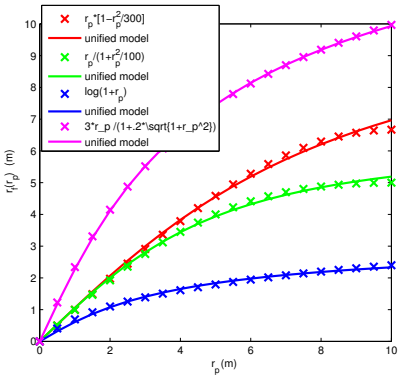

Fig. 5. Equivalence with pinhole-based models.

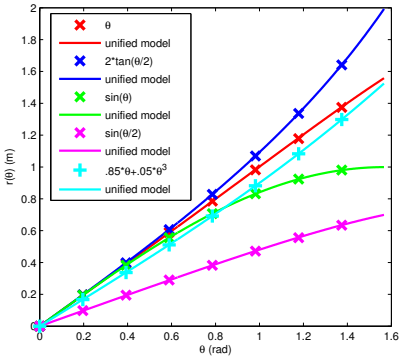

Fig. 6. Equivalence with captured rays-based models.

seen on Fig.5, almost all the tested T_1-mappings can be approximated by the proposed model.
Let us now compare the proposed model with some T_2-mappings (captured rays-based models) assuming $f = 1$:

(a) Translation

$\mathtt{t}/\|\mathtt{t}\|$ (m)	0.843	-0.334	0.422
$\widetilde{\mathtt{t}}/\|\widetilde{\mathtt{t}}\|$ (m)	0.841	-0.344	0.417
Θ_T (dg)		0.917	

(b) Rotation

$\mathtt{u}\theta$ (dg)	0	-40.00	0
$\widetilde{\mathtt{u}\theta}$ (dg)	0.206	-39.4	-3.20
Θ_R (dg)		3.20	

(c) Conjugated translation and rotation

$\mathtt{t}/\|\mathtt{t}\|$ (m)	0.688	0	0.726	$\mathtt{u}\theta$ (dg)	0	-20.00	0
$\widetilde{\mathtt{t}}/\|\widetilde{\mathtt{t}}\|$ (m)	0.673	-0.015	0.740	$\widetilde{\mathtt{u}\theta}$ (dg)	0.470	-21.77	-0.006
Θ_T (dg)		1.21		Θ_R (dg)		1.831	

TABLE III

Computation of the displacement: translation (a), rotation (b) and rotation and translation (c). The rotation is represented by $\mathtt{u}\theta$ (rotation axis $\mathtt{u}$ and angle θ expressed in degree).

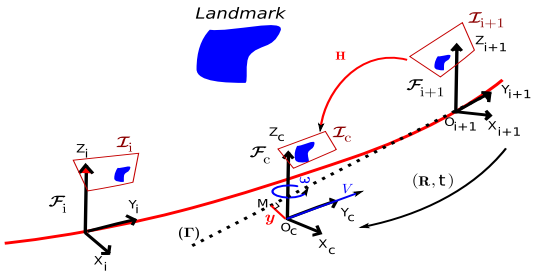

Fig. 7. Control strategy using the homography matrix: the control consists on regulating the lateral error y and the angular error θ to zero.

$r_f^i(\theta)$ ($i = 1, 2, 3, 4, 5$, refer to Table II). In order to estimate the parameters f_f and ξ of the model, we use a set of points $r_f^i(\theta_k)$ and we solve $\min_{\xi, f_f} C(\theta)$. Results are plotted in Fig.6.

Once again, the proposed model is able to approximate almost all the tested transformations T_2 and thus almost all the fisheye cameras.

C. Camera calibration

An important advantage of the proposed model is that existing algorithms designed for central catadioptric cameras can directly be used with fisheye cameras. As an example, our camera is calibrated with the Matlab toolbox provided by Mei [4]. The complete intrinsic parameters matrix is given by: $\mathbf{K} = \begin{pmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$ and the unknown parameters f_u, f_v, u_0, v_0 and ξ have been estimated to:

$$\begin{cases} f_u = 870.60 \text{ px} \\ f_v = 868.34 \text{ px} \end{cases} \quad \begin{cases} u_0 = 327.37 \text{ px} \\ v_0 = 231.87 \text{ px} \end{cases} \quad \xi = 2.916$$

The computed parameter ξ agree the value estimated from the provider's data in Section IV-A ($\xi = 2.933$).
As we will see in the sequel, the calibration results are compatible with robotic applications.

D. Scaled Euclidean reconstruction from fisheye views

In this section, the proposed model is exploited to enable a partial Euclidean reconstruction by decoupling the interaction between translation and rotation components of a homography matrix. In order to show the validity of the spherical model for fisheye cameras, the results are compared with the real displacements.
Let $\mathcal{X}$ be a 3D point with coordinates $\mathtt{X} = [X\ Y\ Z\ ]^T$ in a camera frame $\mathcal{F}$ and $\mathtt{X}^* = [X^*\ Y^*\ Z^*\ ]^T$ in a camera frame $\mathcal{F}^*$. The motion between $\mathcal{F}$ and $\mathcal{F}^*$ is composed of a rotation $\mathbf{R}$ and a translation $\mathtt{t}$ expressed in the camera frame $\mathcal{F}$. Let $\overline{\mathtt{x}}$ be: $\overline{\mathtt{x}} = \begin{bmatrix} \mathtt{x}^T & \frac{1}{1+\xi\eta} \end{bmatrix}^T$ where:

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

Then a relation between the two views (refer to [16]) exists:

$$\overline{\mathtt{x}} \propto \mathbf{H}_\pi \overline{\mathtt{x}}^*$$

$\mathbf{H}_\pi$ is the Euclidean homography matrix related to the plane (π), function of the camera displacement and of the plane coordinates in $\mathcal{F}^*$. Note that the camera has to be calibrated to estimate $\overline{\mathtt{x}}$ from the corresponding image points. From the $\mathbf{H}_\pi$-matrix, the camera motions parameters (the rotation matrix $\mathbf{R}$ and the scaled translation $\mathtt{t} = \|\mathtt{t}\|$) can be estimated.
In the reported experiments, three camera displacements have been considered: translation $\mathtt{t}$ (Tab.III (a)), rotation $\mathbf{R}$ (Tab.III (b)), rotation and translation ($\mathbf{R}$, $\mathtt{t}$) (Tab.III (c)). These displacements are estimated by using the images acquired by the ORIFL190-3 fisheye camera (calibrated as explained in Section IV-C) at the initial and final positions by decomposing the homography matrix (rotation matrix $\widetilde{\mathbf{R}}$ and normalized translation $\widetilde{\mathtt{t}}/\|\widetilde{\mathtt{t}}\|$). The following errors were then computed:

- rotational error: rotation angle Θ_R of the matrix $\mathbf{R}\widetilde{\mathbf{R}}^{-1}$
- translational error: angle Θ_T between the normalized vectors $\mathtt{t}/\|\mathtt{t}\|$ and $\widetilde{\mathtt{t}}/\|\widetilde{\mathtt{t}}\|$

As observed in Table III, displacements are well estimated. When a translation is executed, an error of less than 1° is observed. The error between the computed rotation and the executed rotation is around 3°. During a rotation and translation displacement, the obtained errors are less than 2°.

V. Application to Visual Servoing

To experiment our model in a robotic context, we extend the vision-based navigation framework for nonholonomic mobile robot presented in [17] with a perspective camera and in [18] with central cameras to fisheye cameras. Principles of the navigation framework are briefly presented in V-A. Experimental results with a catadioptric camera can be found in [18]. Experiments with the fisheye camera are presented in Section V-B.

A. Vision based navigation

The framework consists on two successive steps. During an off-line learning step, the robot performs paths which are sampled and stored as a set of ordered key images acquired by the embedded fisheye camera. The set of visual paths can be interpreted as a visual memory of the environment. In the second step, the robot is controlled by a vision-based control law along a visual route which joins the current image to the target image in the visual memory.
To design the controller, the key images of the reference

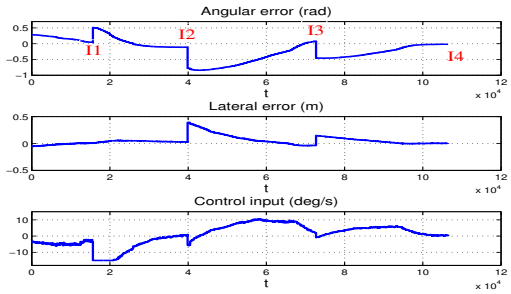

Fig. 8. Navigation task: evolution of the lateral (in m) and the angular (in rad) errors and of the control input (angular speed in deg/s). The key images $I_j^*, \{j = 1, 2, 3, 4\}$ are successively reached.

visual route are considered as consecutive checkpoint to reach in the sensor space. The robot is thus controlled from the image $\mathcal{I}_i$ to the next image of the visual route $\mathcal{I}_{i+1}$. The control problem can be formulated as a path following to guide the nonholonomic mobile robot along the visual route. Let us note $\mathcal{F}_i = (O_i, \mathbf{X_i}, \mathbf{Y_i}, \mathbf{Z_i})$ and $\mathcal{F}_{i+1} = (O_{i+1}, \mathbf{X_{i+1}}, \mathbf{Y_{i+1}}, \mathbf{Z_{i+1}})$ the frames attached to the robot when $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ were stored and $\mathcal{F}_c = (O_c, \mathbf{X_c}, \mathbf{Y_c}, \mathbf{Z_c})$ a frame attached to the robot in its current location. Figure 7 illustrates this setup. The control vector of the considered cart-like robot is $\mathbf{u} = [V, \omega]^T$ where V is the longitudinal velocity along the axle $\mathbf{Y_c}$ of $\mathcal{F}_c$, and ω is the rotational velocity around $\mathbf{Z_c}$. Consider the straight line $\Gamma = (O_{i+1}, \mathbf{Y}_{i+1})$ (see Figure 7). The control strategy consists in guiding $\mathcal{I}_c$ to $\mathcal{I}_{i+1}$ by regulating asymptotically the axle $\mathbf{Y}_c$ on Γ. The control objective is achieved if $\mathbf{Y}_c$ is regulated to Γ before the origin of $\mathcal{F}_c$ reaches the origin of $\mathcal{F}_{i+1}$. Let θ be the angle (around the axle Z) between $\mathcal{F}_c$ and $\mathcal{F}_{i+1}$ and y the distance between O_c and Γ. We use the asymptotically stable guidance control law proposed in [17]:

$$\omega(y, \theta) = -V \cos^3\theta K_p y - |V \cos^3\theta| K_d \tan\theta$$

where K_p and K_d are two positive gains which set the performances of the control law. The lateral and angular deviations (y and θ) to regulate can be obtained through partial Euclidean reconstructions as described in Section IV-D.

B. Experimentations

The proposed framework is implemented on a standard PC which controls a Pioneer 2 robot. A fisheye camera is embedded on the robot. A learning stage has been conducted off-line and images have been memorized. Four key views have been selected to drive the robot from its initial configuration to the desired one. For this experiments, the positions of five patterns are memorized and then tracked. The control is realized using the homography matrix from the projection of the patterns onto the equivalence sphere. The longitudinal velocity of the robot is set to $V = 0.1\, m/s$. The results of the experimentation (Fig.8) show that the lateral and the angular errors are regulated to zero before reaching a key image. Once again, this experimentation demonstrates that the proposed model is suitable for such a task.

VI. Conclusion

The unified model for central catadioptric systems is able to model almost all the fisheye cameras in the context of robotic applications. This result allows to directly employe all existing algorithms designed for central catadioptric cameras. Calibration and structure from motion results as well as a vision-based navigation application have confirmed the validity of our approach. Future work will be devoted to adequately combine fisheye with other cameras (perspective, catadioptric) for navigation tasks.

Acknowledgment

This work is supported by the EU-Project FP6 IST μDrones, FP6-2005-IST-6-045248.

References

[1] R. Benosman and S. Kang, *Panoramic Vision*. Springer Verlag ISBN 0-387-95111-3, 2000.

[2] S. Baker and S. Nayar, "A theory of catadioptric image formation," in *International Conference on Computer Vision*, 1998, pp. 35–42.

[3] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *European Conference on Computer Vision*, vol. 29 (3), Dublin, Ireland, May 2000, pp. 159–179.

[4] C. Mei, S. Benhimane, E. Malis, and P. Rives, "Constrained multiple planar template tracking for central catadioptric cameras," in *British Machine Vision Conference BMVC06*, 2006.

[5] X. Ying and Z. Hu, "Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model," in *European Conference on Computer Vision, ECCV'2004*, 2004, pp. Vol I: 442–455.

[6] J. Barreto, "A unifying geometric representation for central projection systems," *Computer Vision and Image Understanding*, vol. 103, pp. 208–217, 2006.

[7] A. Fitzgibbon, "Simultaneous linear estimation of multiple-view geometry and lens distortion," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2001*, vol. 1, 2001, pp. I–125–I–132.

[8] M. El-Melegy and A. Farag, "Statistically robust approach to lens distortion calibration with model selection," *Conference on Computer Vision and Pattern Recognition CVPR-03, Workshop on Intelligent Learning*, vol. 8, pp. 150–156, 2003.

[9] H. Li and R. Hartley, "An easy non-iterative method for correcting lens distortion from nine point correspondences," *OMNIVIS'05*, 2005.

[10] A. Basu and S. Licardie, "Alternative models for fish-eye lenses," *Pattern Recognition Letters*, vol. 16, no. 4, pp. 433–441, 1995.

[11] R. Kingslake, *A History of the Photographic Lens*. Academic Press, San Diego, 1989, 334 p., ISBN 0-12-408640-3.

[12] D. Stevenson and M. Fleck, "Nonparametric correction of distortion," Computer Science. University of Iowa. USA, Tech. Rep., 1995, technical report TR 95-07.

[13] S. Ray, *Applied Photographic Optics*. Focal Press, Oxford, 1994, second edition. ISBN-0240514998.

[14] W. Smith, *Modern Lens Design: A Resource Manual*. McGraw-Hill, New-York, 1992, ISBN: 0070591784.

[15] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A flexible technique for accurate omnidirectional camera calibration and structure from motion," in *Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)*, 2006.

[16] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet, "Omnidirectional visual servoing from polar lines," in *International Conference on Robotics and Automation, ICRA'06*, vol. 1, 2006, pp. 2385–2390.

[17] G. Blanc, Y. Mezouar, and P. Martinet, "Indoor navigation of a wheeled mobile robot along visual routes," in *IEEE International Conference on Robotics and Automation ICRA'05*, 2005, pp. 3365–3370.

[18] J. Courbon, G. Blanc, Y. Mezouar, and P. Martinet, "Navigation of a non-holonomic mobile robot with a memory of omnidirectional images," in *ICRA 2007 Workshop on "Planning, perception and navigation for Intelligent Vehicles"*, 2007.