

Evaluation of the Unified Model of the Sphere for Fisheye Cameras in Robotic Applications

Jonathan Courbon, Youcef Mezouar, Philippe Martinet

▶ **To cite this version:**

Jonathan Courbon, Youcef Mezouar, Philippe Martinet. Evaluation of the Unified Model of the Sphere for Fisheye Cameras in Robotic Applications. Advanced Robotics, 2012, 26 (8-9), pp.947-967. hal-02463429

HAL Id: hal-02463429
https://inria.hal.science/hal-02463429

Submitted on 31 Jan 2020

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.

Evaluation of the Unified Model of the Sphere for Fisheye Cameras in Robotic Applications

Jonathan Courbon, Youcef Mezouar * and Philippe Martinet

LASMEA, 24 Avenue des Landais, 63171 Aubiere Cedex, France

Received 3 June 2011; accepted 27 June 2011

Abstract

A wide field of view is required for many robotic vision tasks. Such an aperture may be acquired by a fisheye camera, which provides a full image compared to catadioptric visual sensors, and does not increase the size and the weakness of the imaging system with respect to perspective cameras. While a unified model exists for all central catadioptric systems, many different models, approximating the radial distortions, exist for fisheye cameras. It is shown in this paper that the unified projection model proposed for central catadioptric cameras is also valid for fisheye cameras in the context of robotic applications. This model consists of a projection onto a virtual unitary sphere followed by a perspective projection onto an image plane. This model is shown equivalent to almost all the fisheye models. Calibration with four cameras and partial Euclidean reconstruction are done using this model, and lead to persuasive results. Finally, an application to a mobile robot navigation task is proposed and correctly executed along a 200-m trajectory.

© Koninklijke Brill NV, Leiden and The Robotics Society of Japan, 2012

Keywords

Fisheye camera, projection model, robotics

1. Introduction

Conventional cameras are generally seen as perspective tools (pinhole model), which is convenient for modeling and algorithmic design. Moreover, they exhibit small distorsions and thus acquired images can easily be interpreted. Unfortunately, conventional cameras suffer from a restricted field of view. A camera with a large field of view is useful in many robotic applications including robot localization, navigation and visual servoing. In the literature, several methods have been proposed for increasing the field of view. The first method consists on using a moving camera or a system made of multiple cameras. A second solution consists of combining a conventional camera and mirrors. The obtained sensors are referred to as a

* To whom correspondence should be addressed. E-mail: youcef.mezouar@lasmea.univ-bpclermont.fr



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

catadioptric imaging system. The resulting imaging systems have been termed central catadioptric when a single projection center describes the world-image mapping [1]. In Ref. [2], a projection model valid for the entire class of central catadioptric cameras has been proposed. According to this generic model, all central catadioptric cameras can be modeled by a central projection onto a unitary sphere followed by a perspective projection onto the image plane. Such sensors have been employed in many vision-based robotic applications such as mobile robot localization, navigation [3] or visual servoing [4]. However, those cameras exhibit generally a large dead area in the center of the image (Fig. 1, right), which can be a huge drawback. Such sensors also have the drawback of requiring a mirror, which results in a more cumbersome and more fragile sensor imaging system. The last class of cameras with a wide field of view can be obtained as a dioptric system: a fisheye camera [1]. A fisheye camera is an imaging system that combines a fisheye lens and a conventional camera. Their main advantages with respect to catadioptric sensors are that (i) they do not exhibit a dead area (Fig. 1, middle) and (ii) a fisheye lens does not increase the size and the weakness of the imaging system with respect to a conventional camera. Whereas central catadioptric cameras have a single viewpoint (central point), fisheye cameras have a locus of projection centers (diacaustic). However, in many robotic applications the small locus could be approximated by a single viewpoint. As we will see in this paper, the calibration accuracy under this assumption can satisfy robotic applications.

Fisheye cameras provide an image with a large field of view (around 180° or more), but exhibit strong distortions. Those distortions are generally split into several components. The tangential distortions are generally negligible and thus are not considered here. A decentering distortion is caused by the non-orthogonality of the lens components with respect to the optical axis [5]. Distortions may also be due to the non-regular geometry of the lens. The higher distortions are the radial distortions (barrel or pincushion distortions). Those distortions only depend on the angle of incidence of the incoming ray. Considering robotic applications such as robot navigation or visual servoing, it is reasonable to take into account only radial distortions to model fisheye cameras. In this paper, we will use this assumption. While a unified model exists for central catadioptric cameras [2], many different models are used for fisheye cameras. The radial distortion models can

be classified into three main groups. The first group is based on the pinhole model. A three-dimensional (3-D) point is first perspectively projected into the image plane (pinhole model). A radial distortion is then applied to the projected point to obtain the distorted image point. Several parametrical models of distortion can be found in the literature, such as polynomial-based models [6], logarithmic mapping [7] and hyperbolic cosine mapping [8]. More recently, a parameter-free model has been proposed in Ref. [9] and used in Ref. [10] for the correction of endoscopic medical images. The second category is based on the captured rays. It consists of defining a mapping between the incidence ray direction and the distance between the image point and the image center. Some of these mappings are the f-theta mapping, the stereographic projection [11], the orthogonal projection, the equisolid angle projection and the polynomial model [12]. The last category is based on the unified catadioptric camera model [2, 4]. It is motivated by the fact that similar behaviors have been observed for those sensors. As has been noted in Ref. [2], the image of 3-D straight lines is projected onto conics in the image plane of central catadioptric cameras [4], but also in the image plane of some fisheye cameras [13]. In Ref. [14], a third-order polynomial model of radial distortion is applied to the projected point obtained with the unified model in order to model central catadioptric cameras and fisheye cameras with the division model of distortions [15]. In Ref. [16], the authors propose to extend the unified model to include fisheye cameras by substituting the projection on a sphere by a projection on a quadric surface. In Ref. [17], it is shown that the division model of radial distortions proposed in Ref. [15] can be described using a unitary paraboloid that leads to a geometry similar to paracatadioptric systems.

In this paper, the unified model for catadioptric cameras *via* a spherical projection (called the unified spherical model (USM)) will be used. This model exhibits many advantages. (i) This representation is compact in the sense that a unique parameter will be used to describe radial distortions. (ii) This model is geometrically intuitive as the sphere provides radial distortions. (iii) The transfer function has an analytic inverse. (iv) Many algorithms have been designed for central catadioptric cameras using the USM, such as for camera calibration [14], visual servoing [4] or attitude estimation [18]. The aim of this work is to show that the USM can be directly employed to model almost all fisheye cameras. As a consequence, all existing algorithms for catadioptric cameras are also useful for fisheye lenses.

This paper is organized as follows. In Section 2, the models classically employed for fisheye cameras are detailled. The USM is presented in Section 3, and validated for fisheye cameras in Section 4 through comparison with several existing models, experimental results of calibration and partial Euclidean reconstruction. Section 5 reports experimental results in the context of mobile robot navigation using a fisheye camera. Section 6 presents conclusions.

2. Fisheye Camera Models

Let $\mathcal{F}_m$ and $\mathcal{F}_i$ be the frames attached to the camera with origin $\mathcal{M}$ located at the projection center and to the image plane with origin $\mathcal{I}$ located at the principal point, respectively. Let $\mathcal{X}$ be a 3-D point of coordinates $\mathsf{X} = (X\ Y\ Z)^{\mathrm{T}}$ with respect to the frame $\mathcal{F}_m$. The subscripts p and f denote, respectively, perspective and fisheye cameras. The perspective image and the fisheye image coordinates of $\mathcal{X}$ with respect to $\mathcal{F}_i$ are $\mathsf{x}_{\mathrm{p}} = (x_{\mathrm{p}}\ y_{\mathrm{p}})^{\mathrm{T}}$ and $\mathsf{x}_{\mathrm{f}} = (x_{\mathrm{f}}\ y_{\mathrm{f}})^{\mathrm{T}}$, respectively. The distance between $\mathcal{I}$ and the image point is denoted by r, the angle between the incoming ray and the principal axis by θ (Fig. 2). For a perspective camera, given the focal length f, r_{p} satisfies:

$$r_{\mathrm{p}}(\theta) = f \tan\theta. \tag{1}$$

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

Constraint 1. *The ray arriving along the principal axis is not deformed:* $r_{\mathrm{f}}(0) = 0$.

Constraint 2. *The radius* $r_{\mathrm{f}}(k)$ *is monotonically increasing for* $k > 0$.

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

2.1. Pinhole-Based Models

The first class of models 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: $r_{\mathrm{p}} \longrightarrow_{T_1} r_{\mathrm{f}}$ or by the inverse mapping: $r_{\mathrm{f}} \longrightarrow_{T_1'} r_{\mathrm{p}}$.

As (1) is not defined for $\theta = \pi/2$, the models based on the pinhole projection are not defined too. Those mappings T_1 and T_1' can be defined by several ways as detailed in Table 1(left). In Refs [6, 19] among others, radial distortions are described

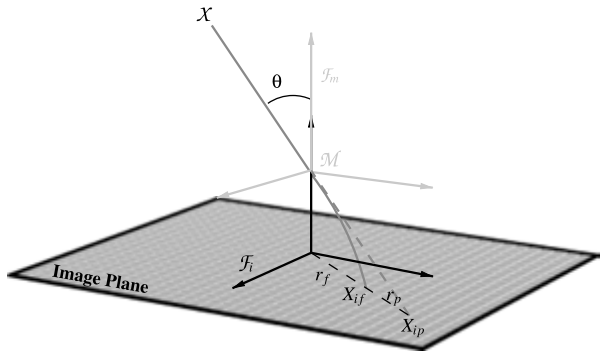


Figure 2. Perspective and fisheye imaging process. The perspective image of a 3-D point $\mathcal{X}$ is x_{p} and its fisheye image is x_{f}.

Table 1.
Models of projection classically employed

Pinhole-based models	Captured rays-based models
$r_{\mathrm{f}}^1(r_{\mathrm{p}}) = r_{\mathrm{p}} L(r_{\mathrm{p}}, n)$	$r_{\mathrm{f}}^1(\theta) = f\theta$
$r_{\mathrm{f}}^2(r_{\mathrm{p}}) = r_{\mathrm{p}} \frac{L_1(r_{\mathrm{p}}, n_1)}{L_2(r_{\mathrm{p}}, n_2)}$	$r_{\mathrm{f}}^2(\theta) = 2f \tan(\frac{\theta}{2})$
$r_{\mathrm{f}}^3(r_{\mathrm{p}}) = s \log(1 + \lambda r_{\mathrm{p}})$	$r_{\mathrm{f}}^3(\theta) = f \sin\theta$
	$r_{\mathrm{f}}^4(\theta) = f \sin(\frac{\theta}{2})$
	$r_{\mathrm{f}}^5(\theta) = f(k_1\theta + k_2\theta^3 + \cdots + k_n\theta^{2(n-1)+1})$
$r_{\mathrm{p}}^1(r_{\mathrm{f}}) = \frac{r_{\mathrm{f}}}{L(r_{\mathrm{f}}, n)}$	$r_{\mathrm{f}}^6(\theta) = \alpha \sin(\beta\theta)$
$r_{\mathrm{p}}^2(r_{\mathrm{f}}) = \frac{r_{\mathrm{f}}}{1 + k_1 r_{\mathrm{f}}^2}$	$r_{\mathrm{f}}^7(\theta) = a \tan(\theta/b) + c \sin(\theta/d)$

by a polynomial model $r_{\mathrm{f}}^1(r_{\mathrm{p}})$, where $L(r_{\mathrm{p}}, n) = 1 + k_1 r_{\mathrm{p}}^2 + k_2 r_{\mathrm{p}}^4 + \cdots + k_n r_{\mathrm{p}}^{2n}$. The n parameters of the model are k_i, $i = 1, 2, \ldots, n$. This model is classically used for perspective cameras [20]. In practice, this model fits to lenses with small distortions, but many parameters are needed for lenses with larger distortions. In Ref. [21], a piecewise polynomial model of distortion is proposed in order to obtain an inverse solution and lower-order polynomials. A rational model $r_{\mathrm{f}}^2(r_{\mathrm{p}})$ is proposed in Ref. [22] with $n_1 + n_2$ parameters. A logarithmic mapping $r_{\mathrm{f}}^3(r_{\mathrm{p}})$ with parameters s (scaling factor) and λ (gain to control the amount of distortion over the whole image) is used in Ref. [7].

The division model $r_{\mathrm{p}}^1(r_{\mathrm{f}})$ suggested in Ref. [15] allows us to handle high distortion with few parameters. This model can be used with only one parameter $r_{\mathrm{p}}^2(r_{\mathrm{f}})$ [15]. Contrary to the previous models, it is inherently an inverted model in the sense that it links the distorted radius to the perspective radius.

In Ref. [9], a generic parameter-free model has been proposed. The calibration step consists of estimating the position of the distortion center and the discrete values $\lambda_{r_{\mathrm{p}}}$ such as $r_{\mathrm{f}}(r_{\mathrm{p}}) = \lambda_{r_{\mathrm{p}}}$ where r_{f} and r_{p} are the distance to the distortion center. The 'distortion function' $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ can be then fitted by a parametric model.

2.2. Captured Rays-Based Models

The second group of models is based on captured rays. It consists of defining a mapping T_2 between the fisheye radius and the incidence angle: $\theta \longrightarrow_{T_2} r_{\mathrm{f}}$.

For perspective cameras, (1) maps the incidence angle θ to the radius r. For fisheye cameras, this mapping is no longer valid. Some T_2-mappings have been proposed to take into account distortions (refer to Table 1(right), where f is the focal length of the camera). The f-theta mapping or equiangular or equidistant projection $r_{\mathrm{f}}^1(\theta)$ proposed in Ref. [23], is suitable for cameras with limited distortions. The stereographic projection $r_{\mathrm{f}}^2(\theta)$ proposed in Ref. [11] preserves circularity and thus projects 3-D local symmetries onto 2-D local symmetries. The orthogonal or

sine law projection $r_{\mathrm{f}}^{3}(\theta)$ is described in Ref. [24]. In Ref. [25], the equisolid angle projection function $r_{\mathrm{f}}^{4}(\theta)$ is proposed. The polynomial function $r_{\mathrm{f}}^{5}(\theta)$ proposed in Refs [12, 26] allows us to improve the accuracy of a polynomial model at the cost of increasing the number of parameters. In Ref. [27], the model $r_{\mathrm{f}}^{6}(\theta)$ with two parameters α (scale factor) and β (radial mapping parameter) is proposed. The model $r_{\mathrm{f}}^{7}(\theta)$ presented in Ref. [28] is a combination of the stereographic projection (with parameters a, b) with the equisolid angle projection (with parameters c, d).

3. Modeling Fisheye Camera *via* a Projection on a Sphere

In this section, the USM first is detailed and we then show that it can be viewed as a pinhole-based model (Section 3.2) or as a captured rays-based model (Section 3.3).

3.1. World-Image Mapping

As outlined previously, the USM consists of 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 Fig. 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 $[0\ 0\ -\xi]^{\mathrm{T}}$ with respect to $\mathcal{F}_m$. The image plane is orthogonal to the Z-axis and it is located at a distance $Z = f$ from $\mathcal{C}$.

Let $\mathcal{X}$ be a 3-D point with coordinates $\mathsf{X} = [X\ Y\ Z]^{\mathrm{T}}$ with respect to $\mathcal{F}_m$. The world point $\mathcal{X}$ is projected in the image plane into the point of homogeneous coordinates $\underline{\mathsf{x}} = [x\ y\ 1]^{\mathrm{T}}$. The image formation process can be split in four steps (Fig. 3):

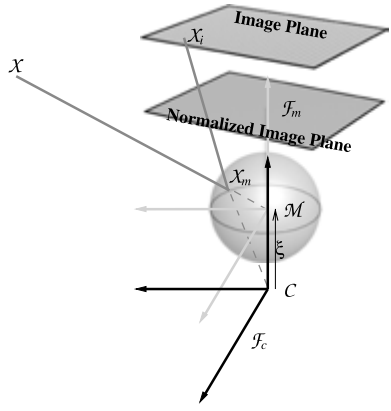


Figure 3. Unified model for a catadioptric camera.

- The 3-D 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 $\mathcal{X}$.

- The point $\mathcal{X}_m$ lying on the unitary sphere is then perspectively projected on the normalized image plane $Z = 1 - \xi$ with respect to $\mathcal{F}_m$ in a point of homogeneous coordinates $\underline{\mathsf{x}}' = f(\mathsf{X})$:

$$\underline{\mathsf{x}}' = f(\mathsf{X}) = \left[\frac{X}{\varepsilon_s Z + \xi\rho} \quad \frac{Y}{\varepsilon_s Z + \xi\rho} \quad 1 \right]^{\mathrm{T}}. \tag{2}$$

  The parameter ε_s allows us to integrate the spherical projection into this model by setting $\varepsilon_s = 0$ and $\xi = 1$. In the general case and in the sequel, this parameter is equal to 1. Note that setting $\xi = 0$ (and $\varepsilon_s = 1$), the general projection model becomes the well-known pinhole model. We will see that ξ can be seen as a parameter that allows us to control the amount of radial distortions.

- The perspective projection of the point of coordinates $\underline{\mathsf{x}}'$ onto the image plane is given by:

$$\underline{\mathsf{x}} = \mathbf{M}\underline{\mathsf{x}}', \tag{3}$$

  where $\underline{\mathsf{x}} = [x \ y \ 1]^{\mathrm{T}}$. Let us note f the focal length. The diagonal matrix $\mathbf{M}$ links the frame attached to the unitary sphere to the camera frame $\mathcal{F}_m$:

$$\mathbf{M} = \begin{pmatrix} f\kappa & 0 & 0 \\ 0 & \delta f\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

  δ is equal to ± 1 depending on the position of the image plane and $\kappa > 0$ depends on the sensor type. In the case of a perspective camera, $\kappa = 1$ and $\delta = +1$, while for catadioptric cameras, $\delta = -1$ and κ is a function of ξ and of a parameter depending on the shape of the mirror. For the sake of simplicity, we will consider that κ is included in the focal length. The radius r is computed as: $r = \sqrt{x^2 + y^2}$.

- Finally, the point of homogeneous coordinates m expressed in pixels in the image plane is obtained after a plane-to-plane collineation

$$\mathbf{K} = \begin{pmatrix} k_u & s_{uv} & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

  of the 2-D projective point of coordinates $\underline{\mathsf{x}}$:

$$\mathsf{m} = \mathbf{K}\underline{\mathsf{x}}. \tag{4}$$

  $(u_0, v_0)^{\mathrm{T}}$ is the position of the optical center in the image plane, k_u and k_v stand for the scaling along the x- and y-axes of the image plane and s_{uv} is the skew. When considering a pinhole camera, the coordinates of the projection of the

3-D point $\mathcal{X}$ in the image plane are obtained setting $\xi = 0$, $\kappa = 1$ and $\delta = +1$. Denoting the focal length f_p, we obtain:

$$\begin{cases} x_\mathrm{p} = f_\mathrm{p} X/Z \\ y_\mathrm{p} = f_\mathrm{p} Y/Z. \end{cases} \tag{5}$$

3.1.1. Inverse Mapping

An additionnal advantage of the USM is that the transfer function has an analytic inverse. Let X_m be the coordinates vector of the point $\mathcal{X}_m$ lying on the unit sphere. Knowing the camera intrinsic parameters, X_m can be computed from the image point x using the relations:

$$\mathsf{X}_m = (\eta^{-1} + \xi)\bar{\mathsf{x}} \quad \text{and} \quad \bar{\mathsf{x}} = \begin{bmatrix} \mathsf{x}^\mathrm{T} & \frac{1}{1+\xi\eta} \end{bmatrix}^\mathrm{T}, \tag{6}$$

where:

$$\begin{cases} \eta = \dfrac{-\gamma - \xi(x^2 + y^2)}{\xi^2(x^2 + y^2) - 1} \\ \gamma = \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}. \end{cases}$$

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

3.2. Equivalence to the Pinhole-Based Model

By extracting the values of X and Y in (3) and (5), elevating squarely and then taking the square root, we obtain:

$$\frac{Z + \xi\rho}{f_\mathrm{f}} \underbrace{\sqrt{x_\mathrm{f}^2 + y_\mathrm{f}^2}}_{r_\mathrm{f}} = \frac{Z}{f_\mathrm{p}} \underbrace{\sqrt{x_\mathrm{p}^2 + y_\mathrm{p}^2}}_{r_\mathrm{p}}. \tag{7}$$

Noticing that $Z > 0$, we can write: $\rho = Z\sqrt{(\frac{X}{Z})^2 + (\frac{Y}{Z})^2 + 1}$. Thanks to (5), we thus have the relation:

$$\rho = \frac{Z}{f_\mathrm{p}} \sqrt{r_\mathrm{p}^2 + f_\mathrm{p}^2}. \tag{8}$$

Then, using (7) we can write:

$$\boxed{r_\mathrm{f} = r_\mathrm{f}(r_\mathrm{p}) = \frac{(f_\mathrm{f}/f_\mathrm{p}) r_\mathrm{p}}{1 + \xi\sqrt{r_\mathrm{p}^2/f_\mathrm{p}^2 + 1}}}. \tag{9}$$

Equation (9) 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 also respected.

3.3. Equivalence to the Captured Rays-Based Model

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

$$r_{\mathrm{f}} = \sqrt{x_{\mathrm{f}}^2 + y_{\mathrm{f}}^2} = \frac{f_{\mathrm{f}}}{Z + \xi\rho}\sqrt{X^2 + Y^2}.$$

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

$$r_{\mathrm{f}} = \frac{f_{\mathrm{f}}}{1 + \xi(\rho/Z)}\tan\theta. \tag{10}$$

Noticing that ρ can be written as a function of $\tan\theta$ and Z: $\rho = Z\sqrt{\tan^2\theta + 1}$, (10) can be rewritten:

$$\boxed{r_{\mathrm{f}} = r_{\mathrm{f}}(\theta) = \frac{f_{\mathrm{f}}\tan\theta}{1 + \xi\sqrt{\tan^2\theta + 1}}}. \tag{11}$$

Equation (11) 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 the captured rays-based model, the model of the sphere presented in this form 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.

4. Validation

In order to validate the USM for fisheye camera, first, the parameters f_{f} and ξ corresponding to different cameras are computed from the captured rays model provided by the camera manufacturers and (11).

In Section 4.2, we show that this model is able to fit almost all existing fisheye cameras. We then calibrate three different visual sensors with a large field of view. Results obtained with the USM are compared to results obtained with other models. The intrinsic parameters of the ORIFL camera are exploited for partial Euclidean reconstruction.

4.1. Fitting Manufacturers Data

Let k^m be a vector containing the parameters of the projection model m. We will first fit the model m with manufacturers data (models i). The estimated parameters $\hat{\mathsf{k}}^m$ are obtained using a set of points $\{(\theta_j, r_{\mathrm{f}}^i(\theta_j)),\ j = 1, \ldots, n\}$ extracted from the manufacturer data by solving the following optimization problem:

$$\hat{\mathsf{k}}^m = \arg\min_{\mathsf{k}^m} C^m(\theta, \mathsf{k}^m),$$

where the cost function is defined as: $C^m(\theta, \mathsf{k}^m) = \sum_{j=1}^{n} \| r_{\mathrm{f}}^m(\theta_j, \mathsf{k}^m) - r_{\mathrm{f}}^i(\theta_j) \|$.

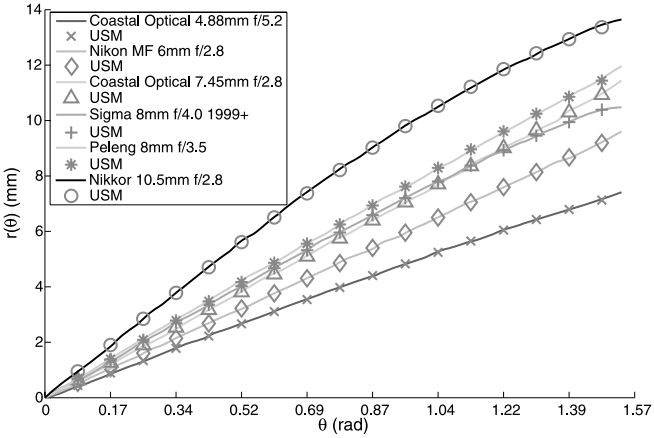


Figure 4. Plots of model (11) (crosses) fit well with the manufacturers' data (plain lines).

Table 2.
Fitting manufacturers data: mean residual on the nth data multiplied by 1000

	$r_{\mathrm{f}}^{\mathrm{pol2}}(\theta)$	$r_{\mathrm{f}}^{\mathrm{pol3}}(\theta)$	$r_{\mathrm{p}}^{\mathrm{div}}$	$r_{\mathrm{f}}^{\mathrm{USM}}(\theta)$
Coastal Optical 4.88 mm $f/5.2$	0.6385	0.6287	*0.4751*	0.6011
Nikon MF 6 mm f/2.8	0.5004	0.4286	*0.3666*	0.4225
Coastal Optical 7.45 mm $f/2.8$	1.7183	1.4344	*1.0859*	1.5150
Sigma 8 mm f/4.0 < 1998	0.9769	0.7797	1.2380	*0.7764*
Peleng 8 mm f/3.5	0.4692	0.4839	*0.2695*	0.6126
Nikkor 10.5 mm f/2.8	1.0395	0.8796	2.3702	*0.8481*
ORIFL190-3	0.1138	*0.0757*	1.4959	0.0953
Mean	0.7795	*0.6730*	1.0430	0.6958

We consider data of the captured ray models provided by the manufacturers of different lenses (Fig. 4). (Data are available on the website: http://michel.thoby.free.fr/Blur_Panorama/Nikkor10-5mm_or_Sigma8mm/Sigma_or_Nikkor/Comparison_Short_Version_Eng.html.) Table 2 summarizes the results obtained with four models of fitting m: the polynomial model $r_{\mathrm{f}}^{3}(\theta)$ with two parameters ($r_{\mathrm{f}}^{\mathrm{pol2}}(\theta) = k_1\theta + k_2\theta^3$) and with three parameters ($r_{\mathrm{f}}^{\mathrm{pol3}}(\theta) = k_1\theta + k_2\theta^3 + k_3\theta^5$), the division model ($r_{\mathrm{p}}^{\mathrm{div}}$), and the model of the unitary sphere of (11) ($r_{\mathrm{f}}^{\mathrm{USM}}$). A point was set aside during the minimization and then used for validation purposes. For each model m and different lenses, Table 2 shows the mean of the residual errors obtained with this point put aside.

The smallest residual errors are obtained with the polynomial model $r_{\mathrm{f}}^{\mathrm{pol3}}$ (using three parameters). The division model (one parameter) leads to the smallest residual errors for four lenses, but provides the highest errors for the Sigma, Nikkor and ORIFL lenses. The USM gives good results for almost all the lenses with only

Table 3.
Mean residuals (multiplied by 1000) obtained by fitting a set of points of models t by models m

t	m			
	$r_{\mathrm{f}}^{\mathrm{pol2}}(r_{\mathrm{p}})$	$r_{\mathrm{f}}^{\mathrm{pol3}}(r_{\mathrm{p}})$	$r_{\mathrm{p}}^{\mathrm{div}}(r_{\mathrm{f}})$	$r_{\mathrm{f}}^{\mathrm{USM}}(r_{\mathrm{p}})$
$r_{\mathrm{p}}(1+kr_{\mathrm{p}})$	0.1419	0.0201	0.0442	*0.0006*
$r_{\mathrm{p}}(1+kr_{\mathrm{p}}^2)$	*0.0000*	*0.0000*	0.0409	0.0029
$s\log(1+kr_{\mathrm{p}})$	1.8373	0.4455	0.0457	*0.0123*
$\frac{r_{\mathrm{p}}}{\sqrt{1+kr_{\mathrm{p}}^2}}$	0.0119	*0.0002*	0.0005	0.0003
Mean	0.4978	0.1164	0.0328	*0.0040*

one parameter. Figure 4 shows the results obtained using the USM as model m for different cameras (Nikon, Coastal, Peleng, etc.). It shows that the USM fits well with the manufacturers data.

4.2. Equivalence with the Pinhole-Based Models

In these experiments, models m are fitted with several T_1 mappings with aleatory chosen parameters (models i) by solving the previously used optimization problem. The same four models m are compared and mean residuals are reported in Table 3.

We can see that the smallest residual errors are obtained with the USM despite the fact that only one parameter is required. These results confirm that the USM is a good candidate for T_1 mapping approximation.

4.3. Equivalence with the Captured Rays-Based Models

Let us now compare the USM with some T_2 mappings (captured rays-based models) (refer to Table 1, right). The same minimization problem as in Section 4.1 is solved. Results are given in Table 4. The polynomial-based model with three parameters gives the smaller reprojection errors. The results obtained with the unified model are relatively similar. Once again, the USM is able to approximate all the tested transformations T_2 and thus almost all the fisheye cameras.

4.4. Camera Calibration

An important advantage of the USM is that existing algorithms designed for central catadioptric cameras can directly be used with fisheye cameras. As an example, our cameras have been calibrated using a planar grid with the MATLAB toolbox provided by Mei [14]. The skew is supposed to be null. The intrinsic parameters matrix is given by:

$$\mathbf{K}_{\mathrm{p}} = \begin{pmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix},$$

Table 4.
Mean residuals (multiplied by 1000) obtained by fitting a set of points of models t by models m

t	m			
	$r_{\mathrm{f}}^{\mathrm{pol2}}(\theta)$	$r_{\mathrm{f}}^{\mathrm{pol3}}(\theta)$	$r_{\mathrm{p}}^{\mathrm{div}}$	$r_{\mathrm{f}}^{\mathrm{USM}}(\theta)$
$f\theta$	*0.0000*	*0.0000*	0.0405	0.0021
$2f\tan(\frac{\theta}{2})$	0.0058	*0.0000*	*0.0000*	*0.0000*
$f\sin\theta$	0.0031	*0.0000*	1.6016	*0.0000*
$f\sin(\frac{\theta}{2})$	*0.0000*	*0.0000*	0.0105	0.0001
$k_1\theta + k_2\theta^3$	*0.0000*	*0.0000*	0.0594	0.0041
Mean	0.0018	*0.0000*	0.3424	0.0013

Table 5.
Calibration of the ORIFL190-3 of Omnitech Robotics (this camera has a field of view of 190° and a focal length of 1.24 mm; the size of the images is 640 × 480)

	Model $r_{\mathrm{f}}^{1}(r_{\mathrm{p}})$	USM with distortions	USM
Focal length: $\binom{f_u}{f_v}$	$\binom{291.5}{293.5}$	$\binom{223.3}{222.5}$	$\binom{222.9}{222.1}$
Principal point: $\binom{u_0}{v_0}$	$\binom{315.6}{255.5}$	$\binom{305.3}{266.9}$	$\binom{305.1}{266.9}$
Parameters:	$a = \begin{pmatrix} -0.427 \\ 0.142 \\ -0.017 \end{pmatrix}$	$a = \begin{pmatrix} 1.385 \\ -12.860 \\ 572.772 \end{pmatrix}$ $\xi = 4.552$	$\xi = 2.854$
Reprojection error: $\binom{err_u}{err_v}$	$\binom{2.013}{1.933}$	$\binom{0.158}{0.132}$	$\binom{0.157}{0.141}$
Reprojection error Val. $\binom{err_u}{err_v}$	$\binom{53.816}{56.798}$	$\binom{0.163}{0.136}$	$\binom{0.159}{0.142}$

where $f_u = f \cdot k_u$ and $f_v = f \cdot k_v$. $[err_u, err_v]$ indicates the reprojection error in pixels. The values prefixed by *Val* correspond to the error obtained with an image not used during the calibration step. The results of the calibration using the pinhole-based model $r_{\mathrm{f}}(r_{\mathrm{p}}) = r_{\mathrm{p}}(1 + a_1 r_{\mathrm{p}}^2 + \cdots + a_3 r_{\mathrm{p}}^6)$, the unified model for central catadioptric cameras (4) with distortions [14] (this model is only presented with fisheye lenses under the division model of distortions $r_{\mathrm{f}}^3(r_{\mathrm{p}})$ [15]) $r_{\mathrm{f}}(r_c) = r_c(1 + a_1 r_c^2 + \cdots + a_3 r_c^6)$ and the model on the sphere are provided in Table 5 for the ORIFL190-3 of Omnitech Robotics, Table 6 for the Sigma 8/3.5 EX DG of Canon and Table 7 for the Fisheye E5 of Fujinon. The same sets of data have been used to calibrate each camera model.

We note that the USM provides satisfactory reprojection errors in all cases. The difference compared to the unified model with distortion parameters is small. Using

Table 6.
Calibration of the Sigma 8/3.5 EX DG (this camera has a field of view of 180°; the size of the images is 1020 × 1020)

	Model $r_{\mathrm{f}}^{1}(r_{\mathrm{p}})$	USM with distortions	USM
Focal length: $\binom{f_u}{f_v}$	$\binom{347.5}{347.1}$	$\binom{347.7}{347.3}$	$\binom{351.0}{350.5}$
Principal point: $\binom{u_0}{v_0}$	$\binom{502.5}{525.4}$	$\binom{502.4}{525.4}$	$\binom{502.2}{525.4}$
Parameters:	$a = \begin{pmatrix} -0.329 \\ 0.132 \\ -0.027 \end{pmatrix}$	$a = \begin{pmatrix} 0.704 \\ -0.672 \\ 37.659 \end{pmatrix}$ $\xi = 3.405$	$\xi = 2.556$
Reprojection error: $\binom{err_u}{err_v}$	$\binom{0.115}{0.105}$	$\binom{0.110}{0.102}$	$\binom{0.100}{0.099}$
Reprojection error Val. $\binom{err_u}{err_v}$	$\binom{0.178}{0.187}$	$\binom{0.218}{0.205}$	$\binom{0.153}{0.168}$

Table 7.
Calibration of the Fujinon Fisheye E5 1:1.4/1.4 mm (this camera has a field of view of 185°; the size of the images is 1024 × 768)

	Model $r_{\mathrm{f}}^{1}(r_{\mathrm{p}})$	USM with distortions	USM
Focal length: $\binom{f_u}{f_v}$	$\binom{243.7}{239.5}$	$\binom{208.6}{208.7}$	$\binom{208.5}{208.5}$
Principal point: $\binom{u_0}{v_0}$	$\binom{500.0}{372.8}$	$\binom{521.9}{391.5}$	$\binom{521.7}{390.9}$
Parameters:	$a = \begin{pmatrix} -0.201 \\ 0.028 \\ -0.001 \end{pmatrix}$	$a = \begin{pmatrix} 0.066 \\ 0.173 \\ -0.324 \end{pmatrix}$ $\xi = 1.720$	$\xi = 1.586$
Reprojection error: $\binom{err_u}{err_v}$	$\binom{1.824}{2.265}$	$\binom{0.142}{0.148}$	$\binom{0.130}{0.136}$
Reprojection error Val. $\binom{err_u}{err_v}$	$\binom{2.787}{4.396}$	$\binom{0.156}{0.174}$	$\binom{0.174}{0.205}$

the unified model with three distortion parameters generally provides the smallest reprojection errors, but increases the complexity of the model for future processing.

4.5. Scaled Euclidean Reconstruction from Fisheye Views

In this section, the USM is exploited to enable a partial Euclidean reconstruction by decoupling the interaction between translation and rotation components of a homography matrix.

Let $\mathcal{X}$ be a 3-D point with coordinates $\mathsf{X} = [X\ Y\ Z]^{\mathrm{T}}$ in a camera frame $\mathcal{F}$ and $\mathsf{X}^* = [X^*\ Y^*\ Z^*]^{\mathrm{T}}$ in a camera frame $\mathcal{F}^*$. The motion $\mathbf{T}$ between $\mathcal{F}$ and $\mathcal{F}^*$ is composed of a rotation $\mathbf{R}$ and a translation t expressed in the camera frame $\mathcal{F}$: $\mathbf{T} = \begin{pmatrix} \mathbf{R} & \mathsf{t} \\ 0_{1\times 3} & 1 \end{pmatrix}$. Using the homogeneous coordinates (denoted with an underline) $\underline{\mathsf{X}} = [X\ Y\ Z\ 1]^{\mathrm{T}}$ and $\underline{\mathsf{X}}^* = [X^*\ Y^*\ Z^*\ 1]^{\mathrm{T}}$, we can write: $\underline{\mathsf{X}} = \mathbf{T}\underline{\mathsf{X}}^*$. Let (π) be a 3-D

reference plane and $d(\mathcal{X}, \pi)$ be the distance from the world point $\mathcal{X}$ to the plane (π). This plane is defined by the vector: $\pi^{*\mathrm{T}} = [\mathsf{n}^{*\mathrm{T}} \ -d^*]$, where n^* its unitary normal in $\mathcal{F}^*$ and d^* the distance from (π) to the origin of $\mathcal{F}^*$. If the point $\mathcal{X}$ belongs to the plane (π), we can deduce after some algebraic manipulation:

$$\bar{\mathsf{x}} \propto \mathbf{H}_\pi \bar{\mathsf{x}}^*, \tag{12}$$

where $\mathbf{H}_\pi = \mathbf{R} + \frac{\mathsf{t}}{d^*}\mathsf{n}^{*\mathrm{T}}$ is the Euclidean homography matrix related to the plane (π), and $\bar{\mathsf{x}}$ and $\bar{\mathsf{x}}^*$ are computed using (6). $\mathbf{H}_\pi$ is a function of the camera displacement and of the plane coordinates with respect to $\mathcal{F}^*$. Note that the camera has to be calibrated to estimate $\bar{\mathsf{x}}$ from the corresponding image points. From the $\mathbf{H}_\pi$ matrix, the camera motion parameters (the rotation matrix $\mathbf{R}$ and the scaled translation $\mathsf{t}/\|\mathsf{t}\|$) can be estimated using standard homography decomposition [29].

4.5.1. Evaluation with Synthetic Images

Synthetic images are used to evaluate the performance of the partial 3-D reconstruction using the USM (Fig. 5). The camera is placed in the virtual working space, looking to a plane π containing N points aleatory located. The 640 × 480 synthetic images are generated by projecting those points in the image plane using the projection model $r_{\mathrm{f}}^1(r_{\mathrm{p}})$ with the parameters of the ORIFL found in Section 4.4 (i.e., $r_{\mathrm{f}}^1(r_{\mathrm{p}}) = r_{\mathrm{p}}(1 - 0.238r_{\mathrm{p}}^2 + 0.040r_{\mathrm{p}}^4 + 0.002r_{\mathrm{p}}^6 - 0.001r_{\mathrm{p}}^8)$). Two-dimensional Gaussian noise with zero mean and standard deviation σ is added to each image point. For each experiment, two positions of the camera are aleatory generated with a displacement consisting of the rotation $\mathbf{R}$ and the translation t. The homography matrix $\mathbf{H}_\pi$ related to the plane π between the two images is estimated for the USM with parameters of the ORIFL lens (i.e., $\xi = 2.875$).

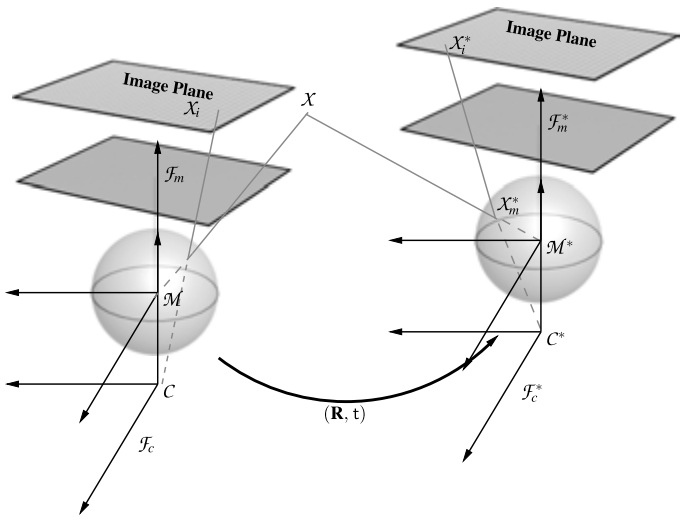


Figure 5. Geometry of two views.

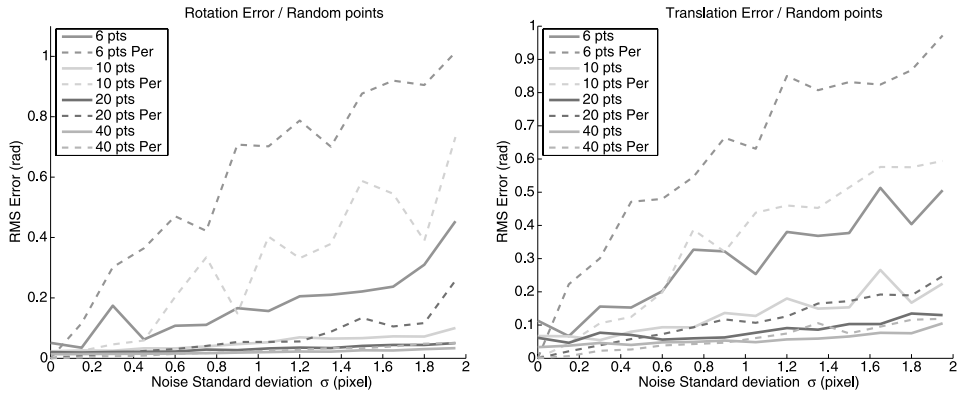


Figure 6. Partial Euclidean reconstruction with synthetic data: r.m.s. errors (in rad) using 6, 10, 20 and 40 points in correspondences versus standard deviation $\sigma = 0, \ldots, 2$ pixel(s). Plain lines, USM; dashed lines, perspective projection.

The estimated rotation $\widetilde{\mathbf{R}}$ and scaled translation $\widetilde{\mathrm{t}}$ are then extracted from $\mathbf{H}_\pi$. Two errors are computed. The rotational error is the rotation angle Θ_R of the matrix $\mathbf{R}\widetilde{\mathbf{R}}^{-1}$ and the translational error is the angle Θ_T between the normalized vectors $\mathrm{t}/\|\mathrm{t}\|$ and $\widetilde{\mathrm{t}}/\|\widetilde{\mathrm{t}}\|$. In order to compare with the well-known perspective case, the homography matrix $\mathbf{H}_\pi$ is also estimated with the pinhole projection model used to generate the images. Rotation and translation are extracted from $\mathbf{H}_\pi$, and rotational and translational errors are computed. The results are compared with the ground truth and the root mean square (r.m.s.) error is computed over 200 runs of each experiment (Fig. 6).

Figure 6 shows that the results are similar to those obtained with a perspective camera: the r.m.s. error is increasing when the standard deviation increases and is decreasing when the number of points used to compute the homography matrix increases. Moreover, for a small number of points, the error of reconstruction with the USM is better than that obtained with a pinhole model. As expected, without noise, values of the errors are zero when the same (perspective) projection model has been used for image generation and homography computation, and are not null when images are generated with the pinhole model and homography matrices are computed with the USM.

4.5.2. Experiments with Real Data

Three camera displacements have been considered: translation t, rotation $\mathbf{R}$, and rotation and translation $(\mathbf{R}, \mathrm{t})$ (Table 8). These displacements are estimated by using the images acquired by the ORIFL190-3 fisheye camera (calibrated as explained in Section 4.4) at the initial and final positions by decomposing the homography matrix (rotation matrix $\widetilde{\mathbf{R}}$ and normalized translation $\widetilde{\mathrm{t}}/\|\widetilde{\mathrm{t}}\|$).

As observed in Table 8, displacements are well estimated. When a translation is realized, an error of less than 1° is observed. The error between the computed rota-

Table 8.
Three-dimensional partial reconstruction with real data: computation of the displacement: translation, rotation, and rotation and translation (the rotation is represented by $\mathsf{u}\theta$ (rotation axis u and angle θ expressed in deg))

Translation				Rotation			
$\mathsf{t}/\|\mathsf{t}\|$ (m)	0.843	−0.334	0.422	$\mathsf{u}\theta$ (deg)	0	−40.00	0
$\widetilde{\mathsf{t}}/\|\widetilde{\mathsf{t}}\|$ (m)	0.841	−0.344	0.417	$\widetilde{\mathsf{u}\theta}$ (deg)	0.206	−39.4	−3.20
Θ_T (deg)		0.917		Θ_R (deg)		3.20	
Translation and rotation							
$\mathsf{t}/\|\mathsf{t}\|$ (m)	0.688	0	0.726	$\mathsf{u}\theta$ (deg)	0	−20.00	0
$\widetilde{\mathsf{t}}/\|\widetilde{\mathsf{t}}\|$ (m)	0.673	−0.015	0.740	$\widetilde{\mathsf{u}\theta}$ (deg)	0.470	−21.77	−0.006
Θ_T (deg)		1.21		Θ_R (deg)		1.831	

tion and the real rotation is about 3°. When a rotation and translation displacement is realized, the errors are less than 2°.

5. Navigation of a Non-Holonomic Vehicle with a Fisheye Camera

Many vision-based navigation strategies are emerging [30, 31] and may benefit for a large field-of-view camera to be more robust to occlusions. In this section, the vision-based navigation framework for a non-holonomic mobile robot presented in Ref. [32] with central cameras is extended to fisheye cameras. Principles of the navigation framework are briefly presented in Section 5.1. Experiments with the fisheye camera are presented in Section 5.2.

5.1. Vision-Based Navigation

The framework consists on two successive steps. During an off-line learning step, the robot performs paths that 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 that joins the current image to the target image in the visual memory.

To design the controller, the key images of the reference visual route are considered as consecutive checkpoints to reach in the sensor space. The robot is 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 non-holonomic 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.

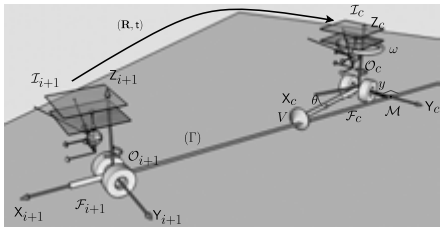


Figure 7. Control strategy using a partial Euclidean reconstruction. $\mathcal{I}_c$ is the current image (when the robot frame is $\mathcal{F}_c$) and $\mathcal{I}_{i+1}$ the desired image (when the robot frame is $\mathcal{F}_{i+1}$). The control consists of regulating the lateral error y and the angular error θ to zero.

The control vector of the considered cart-like robot is $\mathbf{u} = [V \ \ \omega]^{\mathrm{T}}$, 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 $\Gamma = (O_{i+1}, \mathbf{Y}_{i+1})$ (Fig. 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 [32]:

$$\omega(y, \theta) = -V \cos^3\theta K_{\mathrm{p}} y - |V \cos^3\theta| K_{\mathrm{d}} \tan\theta, \tag{13}$$

where K_{p} and K_{d} are two positive gains that set the performances of the control law. The lateral and angular deviations (y and θ) to regulate are obtained through partial Euclidean reconstructions using the epipolar geometry. The essential matrix $\mathbf{E}$ between two images can be estimated using five couples of matched points as proposed in Ref. [33]. Outliers are rejected using a random sample consensus (RANSAC) algorithm. From the essential matrix, the camera motion parameters (i.e., the rotation $\mathbf{R}$ and the translation t up to a scale) can be determined. Finally, the estimation of the input of the control law (13) (i.e., the angular deviation θ and the lateral deviation y) can be computed straightforwardly from $\mathbf{R}$ and t.

5.2. Experimentation

The proposed framework is implemented on a PC that controls an urban electric vehicle, named RobuCab (Fig. 8). The Fujinon fisheye lens, looking forward, is located at approximately 80 cm from the ground. The parameters of the transformation between the camera and the robot control frames are roughly estimated. Grey level images are acquired at a rate of 15 frames/s. The image resolution in the experiments was 800 $\times$ 600 pixels. The camera is modeled using the USM and it has been calibrated as proposed in Section 4.4.

In our experiment, the RobuCab is manually driven along a 200-m path. After a selection step, 82 key images are kept and form the visual memory of the vehicle. During the online step, the longitudinal velocity V has been fixed to 0.8 ms^{-1}. K_{p} and K_{d} have been set in order that the error presents a double pole located at a value of 0.3. At each frame, points are extracted from the current image and matched with



Figure 8. RobuCab vehicle with the embedded fisheye camera.

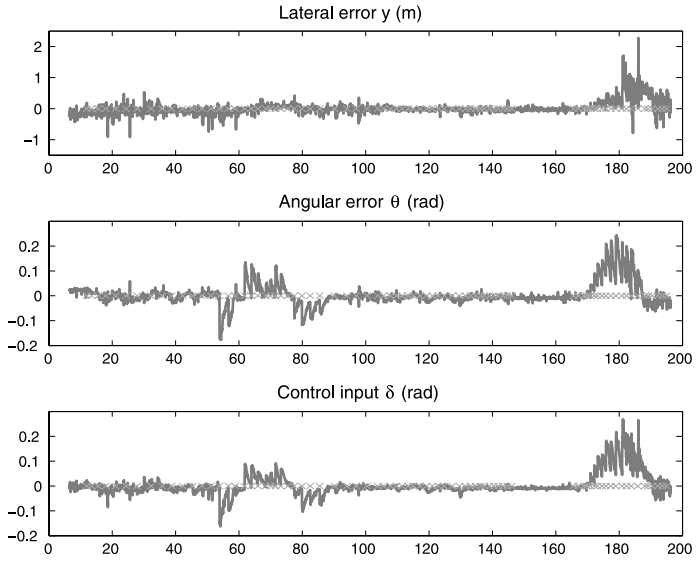


Figure 9. Lateral y (m, up to a scale factor) and angular θ errors (rad) and control input δ (rad) *versus* time (s).

the desired key image. These points are exploited to compute the rotation and the translation between the current and the desired key images. Angular and lateral errors are extracted and used as input of the control law (13).

The RobuCab successfully follows the visual path to reach its target. A mean of 105 good matches for each frame has been found. The mean computational time during the online navigation was of 72 ms by image. For the state estimation, the mean reprojection error in the three views was around 2.5 pixels. Lateral and angular errors as well as control input are represented in Fig. 9. As can be seen, those

errors are well regulated to zero for each key view. Note that the discontinuities are due to the transition between two successive key images.

6. Conclusions

In this paper, we have shown that the unified model for central catadioptric systems can be employed for fisheye cameras. This model depends on a single parameter that controls the amount of radial distortions. The equivalence with pinhole-based and captured rays-based projections have been demonstrated. The calibration of four cameras (ORIFL190-3, TS212A, SIGMA 8/3.5 EX DG and Fisheye E5) shows small reprojection errors and better results than those obtained using multi-parameters models. Three-dimensional partial Euclidean reconstruction from homography matrices shows that we can expect similar results with a fisheye camera modeled by the USM as with a classical camera modeled by the pinhole projection. The visual navigation task consisting on successively reaching key images with visual servoing algorithms was managed successfully using the USM.

Acknowledgements

This work has been supported by the ANR project — ARMEN and the Korea–France Joint Research Center on Cognitive Personal Transport Service Robot under the Gyeonggi Province International Collaborative Research Program.

References

1. S. Baker and S. K. Nayar, A theory of catadioptric image formation, in: *Proc. Int. Conf. on Computer Vision*, Bombay, pp. 35–42 (1998).
2. C. Geyer and K. Daniilidis, A unifying theory for central panoramic systems and practical implications, in: *Proc. Eur. Conf. on Computer Vision*, Dublin, pp. 159–179 (2000).
3. N. Winter, J. Gaspar, G. Lacey and J. Santos-Victor, Omnidirectional vision for robot navigation, in: *Proc. IEEE Workshop on Omnidirectional Vision*, Hilton Head, SC, pp. 21–28 (2000).
4. J. Barreto, F. Martin and R. Horaud, *Visual Servoing/Tracking Using Central Catadioptric Images*, pp. 245–254. Springer, Berlin (2003).
5. R. Swaminathan and S. K. Nayar, Non-metric calibration of wide-angle lenses, in: *Proc. DARPA Image Understanding Workshop*, Monterey, CA, pp. 1079–1084 (1998).
6. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge (2000).
7. A. Basu and S. Licardie, Alternative models for fish-eye lenses, *Pattern Recognit. Lett.* **16**, 433–441 (1995).
8. G. Klancar, M. Kristan and R. Karba, Wide-angle camera distortions and non-uniform illumination in mobile robot tracking, *Robots Autonomous Syst.* **46**, 125–133 (2004).
9. R. Hartley and S. B. Kang, Parameter-free radial distortion correction with centre of distortion estimation, *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1309–1321 (2007).
10. J. P. Barreto, R. Swaminathan and J. Roquette, Nonparametric distortion correction in endoscopic medical images, presented at: *3DTV-CON: The True Vision, Capture, Transmission and Display of 3D Video*, Kos (2007).

11. M. M. Fleck, Perspective projection: the wrong imaging model, *Technical report 95-01*, Computer Science, University of Iowa, Ames, IA (1994).
12. Y. Xiong and K. Turkowski, Creating image-based VR using a self-calibrating fisheye lens, in: *Proc. Conf. on Computer Vision and Pattern Recognition*, Puerto Rico, p. 237 (1997).
13. P. W. Smith, K. B. Johnson and M. A. Abidi, Efficient techniques for wide-angle stereo vision using surface projection models, in: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Fort Collins, CO, vol. 1, pp. 113–118 (1999).
14. C. Mei and P. Rives, Single view point omnidirectional camera calibration from planar grids, in: *Proc. IEEE Int. Conf. on Robotics and Automation*, Rome, pp. 3945–3950 (2007).
15. A. Fitzgibbon, Simultaneous linear estimation of multiple-view geometry and lens distortion, in: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Maui, HI, vol. 1, pp. I-125–I-132 (2001).
16. X. Ying and Z. Hu, Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model, in: *Proc. Eur. Conf. on Computer Vision*, Prague, vol. I, pp. 442–455 (2004).
17. J. P. Barreto, A unifying geometric representation for central projection systems, *Comp. Vis. Image Understand.* **103**, 208–217 (2006).
18. J.-C. Bazin, I. So Kweon, C. Demonceaux and P. Vasseur, UAV attitude estimation by combining horizon-based and homography-based approaches for catadiopric images, in: *Proc. 6th IFAC Symposium on Intelligent Autonomous Vehicles*, Toulouse (2007).
19. T. Pajdla, T. Werner and V. Hlaváč, Correcting radial lens distortion without knowledge of 3-D structure, *Technical Report K335-CMP-1997-138*, Czech Technical University (1997).
20. Z. Zhang, A flexible new technique for camera calibration, *Technical Report MSR-TR-98-71*, Microsoft Research (1998).
21. L. Ma, Y. Chen and K. L. Moore, Analytical piecewise radial distortion model for precision camera calibration, *Vision Image Signal Process.* **153**, 468–474 (2006).
22. H. Li and R. Hartley, An easy non-iterative method for correcting lens distortion from nine point correspondences, presented at: *Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras*, Beijing (2005).
23. R. Kingslake, *A History of the Photographic Lens*. Academic Press, San Diego, CA (1989).
24. S. F. Ray, *Applied Photographic Optics*, 2nd edn. Focal Press, Oxford (1994).
25. W. J. Smith, *Modern Lens Design: A Resource Manual*. McGraw-Hill, New York, NY (1992).
26. E. Schwalbe, Geometric modelling and calibration of fisheye lens camera systems, presented at: *ISPRS Panoramic Photogrammetry Workshop*, Berlin (2005).
27. J. Kumler and M. Bauer, Fisheye lens designs and their relative performance, *SPIE* **4093**, 360–369 (2000).
28. H. Bakstein and T. Pajdla, Panoramic mosaicing with a 180° field of view lens, in: *Proc. IEEE Workshop on Omnidirectional Vision*, Los Alamitos, CA, pp. 60–67 (2002).
29. O. Faugeras and F. Lustman, Motion and structure from motion in a piecewise planar environment, *Int. J. Pattern Recognit. Artif. Intell.* **2**, 485–508 (1988).
30. T. Goedemé, T. Tuytelaars, G. Vanacker, M. Nuttin, L. Van Gool and L. Van Gool, Feature based omnidirectional sparse visual path following, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Edmonton, pp. 1806–1811 (2005).
31. A. Diosi, A. Remazeilles, S. Segvic and F. Chaumette, Outdoor visual path following experiments, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, San Diego, CA, pp. 4265–4270 (2007).

32. J. Courbon, G. Blanc, Y. Mezouar and P. Martinet, Navigation of a non-holonomic mobile robot with a memory of omnidirectional images, presented at: *ICRA 2007 Workshop on 'Planning, Perception and Navigation for Intelligent Vehicles'*, Rome (2007).
33. D. Nistér, An efficient solution to the five-point relative pose problem, *Trans. Pattern Anal. Mach. Intell.* **26**, 756–770 (2004).

About the Authors



Jonathan Courbon graduated from the Institut Français de Mécanique Avancée (IFMA), Clermont-Ferrand, France, in 2005. He received the PhD degree in Vision and Robotics from the Université Blaise Pascal, Clermont-Ferrand, in 2009. He is now pursuing his Postdoc at the Robotics and Vision Group of LASMEA-CNRS, Clermont-Ferrand. His research focuses on the use of visual informations for navigation and control of mobile robots.



Youcef Mezouar received the PhD degree in Computer Science from the Université de Rennes 1, Rennes, France, in 2001, and the 'Habilitation à Diriger les Recherches' degree from the Université Blaise Pascal, Clermont-Ferrand, France, in 2009. For 1 year, he was a Postdoctoral Associate with the Robotics Laboratory, Computer Science Department, Columbia University, New York. In 2002, he joined the Robotics and Vision Group, LASMEA-CNRS, Blaise Pascal University, where he is currently a Colead of the GRAVIR Group (http://www.lasmea.univ-bpclermont.fr) and the ROSACE Team (http://www.lasmea.univ-bpclermont.fr/rosace). His research interests include automatics, robotics and computer vision, especially visual servoing and mobile robot navigation.



Philippe Martinet graduated from the CUST, Clermont-Ferrand, France, in 1985, and he received the PhD degree in Electronics Science from the Blaise Pascal University, Clermont-Ferrand, in 1987. From 1990 to 2000, he was Assistant Professor with CUST in the Electrical Engineering Department, Clermont-Ferrand. From 2000 to 2011, he was Professor with Institut Français de Mécanique Avancée (IFMA), Clermont-Ferrand. He was performing research at the Robotics and Vision Group of LASMEA-CNRS, Clermont-Ferrand. In 2006, he spent 1 year as a Visiting Professor in ISRC at the Sungkyunkwan University in Suwon, South Korea. In 2011, he moved to Ecole Centrale de Nantes and he is doing research at IRCCYN. He was the Leader of the group GRAVIR (over 74 persons) from 2001 to 2006. In 1999, he set up a team in Robotics and Autonomous Complex Systems (over 20 persons). Between 2008 and 2011, he was Colead of a Joint Unit of Technology called 'Robotization Meets Industry'. Between 2009 and 2011, he was Colead of a Korea–France Joint Research Center on Cognitive Personal Transport Service Robots. His research interests include visual servoing of robots (position, image and hybrid based, omnidirectional), multi-sensor-based control, force–vision coupling), autonoumous guided vehicles (control, enhanced mobility, uncertain dynamics, monitoring), modeling, identification and control of complex machines (vision-based kinematic/dynamic modeling identification and control). From 1990, he is Author and Coauthor of more than two hundred thirty references.