Models and Calibration Methods

Guillaume Caron

Chapter 3

Models and Calibration Methods

Author: Guillaume CARON
MIS Laboratory
University of Picardie Jules Verne Amiens
France
and
CNRS-AIST Joint Robotics Laboratory (JRL)
Tsukuba
Japan

3.1 Abstract

This chapter reviews the state-of-the-art projection models for omnidirectional cameras and the calibration methods to compute their parameters. These models range from the expression of explicit shape of mirrors and lenses to unified and generic models. The calibration methods estimate the parameters of these models thanks to partially or totally known scene geometry. Various geometric features are considered for camera calibration such as points, lines circles used as input of linear estimators or non-linear optimization schemes.

Several examples of cameras and images they capture are provided together with sketches illustrating the path of light rays entering cameras as well as the main equations using common notations to ease their understanding and comparison.

3.2 Introduction

This chapter deals with the geometrical modeling of omnidirectional image formation and the methods to compute its parameters for each omnidirectional

camera. The models are formalized by relying on the geometry of the optics of each omnidirectional camera, exploiting the explicit shape of the mirrors and lenses, *i.e.* the ad hoc models (Section 3.3.2), or by more abstract relations valid for several types of omnidirectional cameras, such as the unified central projection model (Section 3.3.3). On the other hand, the generic models (Section 3.3.4) make it possible to compensate for the impossibility of characterizing the omnidirectional camera by the other models or the great difficulty in calibrating it because of a too large number of parameters.

Finally, the calibration methods (Section 3.4) estimate the parameters of the above-mentioned models from correspondences between the content of the omnidirectional image and the observed scene, which is known either totally by the use of calibration patterns or partially, whether the scene is structured or not.

3.3 Projection models

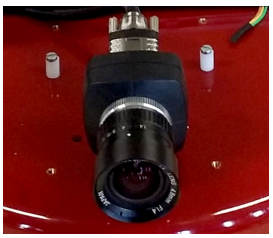
3.3.1 Perspective Projection Recalls

Describing the geometrical formation of an image (Figure 3.1c), observation of a scene by a camera (Fig. 3.1a) assimilable to a pinhole or a camera lens with infinite depth of field, the perspective projection model, very well known, involves a projection center $\mathbf{C} \in \mathbb{R}^3$ and the image plane $\boldsymbol{\pi}$ (Figure 3.1b). The image being a finite rectangle of $\boldsymbol{\pi}$, it defines a section of the pyramid of vertex $\mathbf{C}$ and infinite height characterizing the field of view of the camera. The greater the focal length $f \in \mathbb{R}_+^*$ between the optical center and the image plane, the smaller the field of view, and conversely.

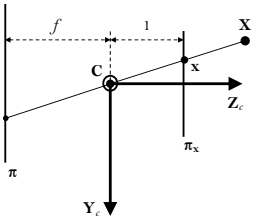
We define the camera frame $\mathcal{F}_c$ with origin $\mathbf{C}$ and axes $\mathbf{X}_c \in \mathbb{R}^3$, such that $||\mathbf{X}_c|| = 1$, and $\mathbf{Y}_c \in \mathbb{R}^3$, such that $||\mathbf{Y}_c|| = 1$, parallel to the horizontal and, respectively, vertical edges of the image rectangle. By usual convention, the direction of these axes follows that of the organization of the pixels of a digital image acquired by a camera, *i.e.* "to the right" for $\mathbf{X}_c$ and "downward" for $\mathbf{Y}_c, \mathbf{Z}_c \in \mathbb{R}^3$, such that $||\mathbf{Z}_c|| = 1$, is orthogonal to the previous axes making $\mathcal{F}_c$ is a direct reference frame, so $\mathbf{Z}_c$ points to the front of the camera. Thus, the three-dimensional (3D) points ${}^c\mathbf{X} = \begin{pmatrix} {}^cX & {}^cY & {}^cZ \end{pmatrix}^\mathsf{T} \in \mathbb{R}^3$ of the scene observed by the camera, expressed in the camera reference frame $\mathcal{F}_c$, have their third coordinate positive.

The perspective projection model expresses an image point by the intersection of the line of sight (${}^c\mathbf{C}{}^c\mathbf{X}$), representing the path followed by a light ray, with the image plane. If, physically, the image plane is beyond the optical center with respect to the 3D point, any plane in space parallel to it can form a virtual image plane, identical to the real image plane, with a similarity transformation. The normalized image plane $\boldsymbol{\pi}_\mathbf{x}$, *i.e.* distant of one unit from ${}^c\mathbf{C}$, on the side of ${}^c\mathbf{X}$, is generally chosen in computer vision.

Therefore, the perspective projection model expresses the projection of a 3D point ${}^c\mathbf{X}$ in $\boldsymbol{\pi}_\mathbf{x}$ as $\mathbf{x} = \begin{pmatrix} x & y \end{pmatrix}^\mathsf{T} \in \mathbb{R}^2$ by:



(a)



(b)

(c)

Figure 3.1: Illustration of the perspective projection. (a) Conventional camera. (b) Diagram of the perspective projection of a point $\mathbf{X}$ of the 3D world in the normalized image plane $\boldsymbol{\pi}_{\mathbf{x}}$. (c) Example of an image acquired by a camera that can be characterized by the perspective projection model from the triforium of the Amiens cathedral: despite the exceptional regularity of the building, the width of the central vessel in the image decreases from the bottom of the image towards its center while it is approximately the same along the entire length of the building.

$$\begin{cases} x = \dfrac{{}^cX}{{}^cZ}, \\ y = \dfrac{{}^cY}{{}^cZ}. \end{cases} \tag{3.1}$$

$\mathbf{x}$ is defined in the normalized image plane whose origin is its own intersection with the $\mathbf{Z}_c$ axis. It is to be noted that

$${}^c\mathbf{x} = \begin{pmatrix} x & y & 1 \end{pmatrix}^\mathsf{T} \in \mathbb{R}^3, \tag{3.2}$$

gives the direction of the line of sight, expressed in $\mathcal{F}_c$ and passing through its origin. It is distinguished from $\tilde{\mathbf{x}} = \begin{pmatrix} x & y & 1 \end{pmatrix}^\mathsf{T} \in \mathbb{P}^2$, the homogeneous representation of $\mathbf{x}$ in $\boldsymbol{\pi}_{\mathbf{x}}$. We then write the perspective projection function $pr()$:

$$\tilde{\mathbf{x}} = pr({}^c\mathbf{X}), \tag{3.3}$$

with x and y expressed as in (3.1).

The origin of the digital image acquired by a camera being at the top left and its sampling being in pixels, the perspective projection model considers

an additional affine transformation $\mathbf{K} \in \mathrm{Aff}(2)$, to transform $\boldsymbol{\pi}_{\mathbf{x}}$ to the digital image plane $\boldsymbol{\pi}_{\mathbf{u}}$. This transformation involves, generally, four parameters $\gamma_p = \{\alpha_u, \alpha_v, u_0, v_0\}$ of which $\alpha_u \in \mathbb{R}^*$ and $\alpha_v \in \mathbb{R}^*$ are the horizontal and, respectively, vertical scale factors and $(u_0, v_0) \in \mathbb{R}^2$ are the coordinates of the principal point, *i.e.* the intersection of $\boldsymbol{\pi}_{\mathbf{x}}$ and $\mathbf{Z}_c$, expressed in the digital image. These parameters, called intrinsic, characterize the optics of the camera, according to the perspective projection model, and are linked to the physical realization of an image, in particular $\alpha_u = f/k_u$ and $\alpha_v = f/k_v$, with $k_u \in \mathbb{R}^*_+$ and $k_v \in \mathbb{R}^*_+$ the dimensions of a photodiode giving a pixel in the digital image. Thus, the point $\tilde{\mathbf{u}} = \begin{pmatrix} u & v & 1 \end{pmatrix}^\mathsf{T} \in \mathbb{P}^2$ of the digital image is obtained from $\tilde{\mathbf{x}}$ by:

$$\tilde{\mathbf{u}} = \mathbf{K}\tilde{\mathbf{x}} \text{ with } \mathbf{K} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.4}$$

By putting together the two steps (3.3) and (3.4), we obtain the projection function of a 3D point ${}^c\mathbf{X}$ in the digital image plane:

$$\tilde{\mathbf{u}} = pr_{\gamma_p}({}^c\mathbf{X}) = \mathbf{K}pr({}^c\mathbf{X}). \tag{3.5}$$

Remark 1 (Enrichment of the perspective projection model) *The above projection model takes into account the fact that pixels may not be perfectly square but rectangular. It can be simplified to a minimum with a single scale factor. This model can also be enriched to take into account: the fact that the pixels are parallelograms, radial or tangential distortions caused by the optics used, a misalignment of the optics and the photodiode array of the camera, etc.*

The perspective projection model is suitable for cameras with a limited field of view, around 122 degrees without distortion[1]. Indeed, the division by ${}^cZ$ in the perspective projection equations (3.1) raises three problems when the field of view approaches or exceeds 180 degrees, for example when using a curved mirror catadioptric lens (Fig. 3.2a) or a fisheye lens (up to[2] 280 degrees fisheye lens). The first problem, when ${}^cZ$ tends to 0, generates the need for a very large image rectangle on $\boldsymbol{\pi}_{\mathbf{x}}$, either because of a very large photosensitive matrix, or at the expense of the resolution in the center of the image. Then, when ${}^cZ = 0$, x and y are not defined in (3.1), not to mention the numerical instability of the calculations when ${}^cZ$ is very close to 0. Finally, when the field of view is greater than 180 degrees, two 3D points ${}^c\mathbf{X}$ and ${}^c\mathbf{X}'$ such that ${}^cX' = -{}^cX$, ${}^cY' = -{}^cY$ and ${}^cZ' = -{}^cZ$ project to the same coordinates in the image plane. This is why the projection function $pr_{\gamma_p}()$ (3.5) must be adapted for omnidirectional cameras by taking into account either explicitly the optical geometry of the lenses (Section **??**), or an abstraction replacing the plane surface of the perspective projection model by a sphere (Section 3.3.3), or by relaxing the idea of the projection surface by reasoning on a vector field (Section 3.3.4).

[1] Distortions less than 1% according to `www.dxomark.com` as of March 1, 2017 for the Sigma 12-24 mm camera lens

[2] Entaniya M12 280.

3.3.2 Ad hoc models

Ad hoc projection models explicitly exploit the optical geometry of panoramic and omnidirectional cameras, whether they involve one or more mirrors (Section 3.3.2), a fisheye lens (Section 3.3.2), or a combination of cameras (Section 3.3.2).

Catadioptric cameras

A catadioptric camera combines curved lenses (diopter) and mirrors (catoptric) to acquire a panoramic image (Fig. 3.2c).

Central catadioptric cameras The panoramic catadioptric camera is single view point, or central, when we associate a hyperbolic mirror of revolution to a perspective camera (hypercatadioptric camera, Fig. 3.2a), a parabolic mirror of revolution to an orthographic camera (paracatadioptric camera), for convex mirrors, or a concave elliptic mirror of revolution to a perspective camera [2]. The single view point is also almost respected by combining a convex paraboloid mirror and a concave spherical mirror with a perspective camera (Fig. 3.2b). For all these configurations, ensuring the single view point requires a precise relative placement of the perspective or orthographic camera and the mirror(s): the principal axes must be aligned and the optical center of the perspective camera must be coincident with the focus of the conical surface of the mirror (of the spherical mirror in the case of two mirrors mentioned above). Other combinations of mirrors, although more complex to implement, also allow the realization of a single view point panoramic camera [8, 17].

The ad hoc modeling of central catadioptric cameras is based on the equations of the surfaces of each mirror and those of the associated cameras. Thus,

(a) (b) (c)

Figure 3.2: Panoramic catadioptric vision. (a) Single catadioptric camera (V-Stone VS-C450MR-TK lens). (b) Dual catadioptric camera (RemoteReality lens) acquiring panoramic (c) image (Place Saint-Michel, Amiens, 2015).

when the placement of the mirror with respect to the camera respects the geometrical constraints to ensure the uniqueness of the view point, we express the coordinates of $\mathbf{x}$ the point of the normalized image plane from the coordinates of the corresponding 3D point ${}^c\mathbf{X}$ and the parameters of the mirror equation [15] (Tab. 3.1) with:

$$\rho = \sqrt{{}^cX^2 + {}^cY^2 + {}^cZ^2}. \tag{3.6}$$

Non-central catadioptric cameras Even with a camera-mirror pair that can lead to the single view point, when their relative placement is imprecise, the catadioptric camera is non-central [23]. This is also the case for any combination of mirror and camera shape, other than those discussed in Section 3.3.2.

In this case, the light rays passing through the camera to reach the image plane do not intersect at a single point [48] but form a caustic surface [19].

The explicit formulation of the caustics associated with the shape of a mirror is based on the flux flow model from geometrical optics. The equation of the caustic surface associated to a mirror is obtained by solving the differential equation canceling the Jacobian of a point $\mathbf{F} \in \mathbb{R}^3$ of the caustic with respect to the height ${}^cZ_m$ of the mirror surface and the distance between $\mathbf{F}$ and the mirror

Table 3.1: Ad hoc projection models of central catadioptric cameras (summary of [15] and of the approximation for the two-mirror projection model [17]).

mirror	**parameters**	**camera**	**projection equations**
parabolic convex	h (semi-latus rectum)	orthographic	$\begin{cases} x = \frac{h\,{}^cX}{\rho - {}^cZ} \\ y = \frac{h\,{}^cY}{\rho - {}^cZ} \end{cases}$ (3.7)
hyperbolic convex	h et e (excentricity)	perspective	$\begin{cases} x = \frac{2eh\,{}^cX/\sqrt{4e^2+h^2}}{\frac{2e}{\sqrt{4e^2+h^2}}\rho - {}^cZ} \\ y = \frac{2eh\,{}^cY/\sqrt{4e^2+h^2}}{\frac{2e}{\sqrt{4e^2+h^2}}\rho - {}^cZ} \end{cases}$ (3.8)
ellipsoidal concave	h et e (excentricity)	perspective	$\begin{cases} x = \frac{2eh\,{}^cX}{2e\rho + {}^cZ\sqrt{4e^2+h^2}} \\ y = \frac{2eh\,{}^cY}{2e\rho + {}^cZ\sqrt{4e^2+h^2}} \end{cases}$ (3.9)
parabolic convex and spherical concave	h et r (radius of the spherical mirror)	perspective	$\begin{cases} x = \frac{2h\,{}^cX}{r/(\rho - {}^cZ) - h^2(\rho + {}^cZ)/r} \\ y = \frac{2h\,{}^cY}{r/(\rho - {}^cZ) - h^2(\rho + {}^cZ)/r} \end{cases}$ (3.10)

along the reflected ray. Using the parameters of excentricity e and semi-latus rectum h of the mirror surface introduced in Section 3.3.2, the mirror surface is then written parametrically [48]:

$$g({}^cZ_m) = \sqrt{\frac{h^2}{4} - h\,{}^cZ_m - (e^2-1)\,{}^cZ_m^2}, \tag{3.11}$$

leading to the implicit expression of the caustic:

$$f({}^cZ_m, g) = (e^2-1)\,{}^cZ_m^2 + g({}^cZ_m)^2 + h\,{}^cZ_m - \frac{h^2}{4} = 0. \tag{3.12}$$

If the shape of the mirror is not known, the factors of the parameters of (3.12) are replaced by unknowns to be estimated numerically.

Fisheye cameras

A fisheye camera has a hemispherical field of view (180 degrees) or more. As for the panoramic catadioptric camera, the radial distortions encountered in the image acquired by the fisheye camera are not aberrations but the result of the projection of a sphere on a plane. When the radial distortions $r \in \mathbb{R}$ are symmetrical with respect to the principal point, they are expressed from the polar angle $\theta \in \mathbb{R}$, formed by the line of sight of the point ${}^c\mathbf{X}$ and the optical axis $\mathbf{Z}_c$, *i.e.* $\theta = \arccos({}^cZ/||{}^c\mathbf{X}||)$. Several fisheye projection models exist, mainly the equidistant [35] projections:

$$r = \theta, \tag{3.13}$$

and equisolid [35]:

$$r = 2\sin\frac{\theta}{2}, \tag{3.14}$$

to compare with the equivalent relationship for a perspective camera (Sec. 3.3.1):

$$r = \tan\theta. \tag{3.15}$$

The equidistant projection has a regular radial resolution while the equisolid projection has a better resolution at the center than at the edges of the image.

Whatever the model considered among the three previous ones, we express the coordinates of the normalized image point $\mathbf{x}$ from r and the second angle $\phi \in \mathbb{R}$ defining the direction of the line of sight, *i.e.* the azimuthal angle $\phi = \arctan({}^cY/{}^cX)$, by:

$$\mathbf{x} = r\begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}. \tag{3.16}$$

However, real fisheye lenses rarely follow perfectly the ideal models of the equations (3.13) and (3.14). In this case, a more general projection model is to be considered, such as the angular polynomial projection model [27]. This model expresses the radial distortions r from a polynomial function of the polar angle

θ, denoted $r(\theta)$, considering only the odd powers of θ, weighted by coefficients k_1, k_2, ...:

$$r(\theta) = k_1\theta + k_2\theta^3 + k_3\theta^5 + k_4\theta^7 + k_5\theta^9 + \dots, \tag{3.17}$$

without loss of generality. Thus this model has as many additional intrinsic parameters to the perpective projection model as terms considered in the polynomial (3.17), *i.e.* $\gamma_{pa} = \{\alpha_u, \alpha_v, u_0, v_0, k_1, k_2, \dots\}$.

Alternatively, the Cartesian polynomial projection model [43] reasons by expressing the radial distortions $r(\rho'_{\mathbf{u}})$ in the digital image plane, centered at the principal point $\mathbf{u}_0$ in which the coordinates of a point are expressed as $u' = u - u_0$ and $v' = v - v_0$. This model defines $r(\rho_{\mathbf{u}'})$ such that the vector $[u', v', r(\rho_{\mathbf{u}'})]$ is collinear with the line of sight associated with the 3D point ${}^c\mathbf{X}$, *i.e.*, by posing $\alpha = \alpha_u = \alpha_v$:

$$\frac{\rho}{\alpha}[u', v', r(\rho_{\mathbf{u}'})] = {}^c\mathbf{X}, \tag{3.18}$$

where ρ is the norm of ${}^c\mathbf{X}$ (3.6), and:

$$r(\rho_{\mathbf{u}'}) = a_0 + a_1\rho_{\mathbf{u}'} + a_2\rho^2_{\mathbf{u}'} + \dots + a_N\rho^N_{\mathbf{u}'}, \tag{3.19}$$

such that:

$$\rho_{\mathbf{u}'} = \sqrt{u'^2 + v'^2}. \tag{3.20}$$

In summary, similar to the angular polynomial projection model, the number of intrinsic parameters of the Cartesian polynomial projection model depends on the number of coefficients considered in the polynomial (3.19). They are noted as $\gamma_{pc} = \{\alpha, u_0, v_0, a_1, a_2, \dots\}$.

Multi-camera systems

The multi-camera systems combine several cameras, identical or not, in a rig with mainly complementary fields of view in order to reach, together, an omnidirectional field of view up to 360 degrees. This type of polydioptric visual sensor goes from the combination of several perspective cameras distributed on the surface of a sphere [49] (polycamera) to the combination of two back-to-back fisheye lenses, for the most compact [28]. Some products are available on the market for professionals of photography and panoramic and 360 movies (or *virtual reality*), for the first (*e.g.* LadyBug, Dodeca 2, Insta360Pro), and many consumer products (*e.g.* Ricoh Theta, Insta360One, Samsung Gear 360, Garmin Virb 360, etc.), for the second.

Generally, a projection model characterizes each camera, so there are as many sets of intrinsic parameters $\gamma_{m,j}$ as there are cameras, m designating the projection model considered and j, the camera index. To these we add the extrinsic parameters which express the pose $\mathbf{p}_{s,j} \in \mathbb{R}^6$ of each camera j, of frame $\mathcal{F}_{c_j}$, in a common coordinates system associated with the polydioptric system, $\mathcal{F}_s$.

Considering $\mathbf{p}_{s,j} = \begin{bmatrix} {}^{c_j}\mathbf{t}_s^\mathsf{T} & | & \mathbf{r}_{s,j}^\mathsf{T} \end{bmatrix}^\mathsf{T}$, such that ${}^{c_j}\mathbf{t}_s \in \mathbb{R}^3$ and $\mathbf{r}_{s,j} \in \mathbb{R}^3$, the axis-angle representation of the rotation matrix ${}^{c_j}\mathbf{R}_s \in SO(3)$ [31], we express

${}^{c_j}\mathbf{M}_s \in SE(3)$, the rigid transformation matrix from frame $\mathcal{F}_s$ to frame $\mathcal{F}_{c_j}$, by:

$$ {}^{c_j}\mathbf{M}_s = \begin{bmatrix} {}^{c_j}\mathbf{R}_s & {}^{c_j}\mathbf{t}_s \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix}. \tag{3.21} $$

This modeling of polydioptric system is very similar to that of stereoscopic vision systems.

3.3.3 Unified central projection and its extensions

Unlike ad hoc models, unified central projection is based on a mathematical abstraction of the geometry of central omnidirectional camera image formation. It introduces a virtual sphere as an intermediate projection surface between the scene and the image plane [15, 3]. This sphere is also another image representation space, common to all central cameras.

Unified central projection model

The unified central projection model (Fig. 3.3) can be seen as a generalization of the perspective projection model consisting in adding a preliminary step which first projects the 3D point ${}^c\mathbf{X}$ onto a unit sphere [3] of center $\mathbf{C}$ in $\mathbf{X}_\mathcal{S} \in \mathbb{R}^3$, such that $||\mathbf{X}_\mathcal{S}|| = 1$:

$$ \mathbf{X}_\mathcal{S} = \begin{pmatrix} X_\mathcal{S} \\ Y_\mathcal{S} \\ Z_\mathcal{S} \end{pmatrix} = pr_\mathcal{S}({}^c\mathbf{X}) \quad \text{with} \quad \begin{cases} X_\mathcal{S} = \frac{{}^cX}{\rho} \\ Y_\mathcal{S} = \frac{{}^cY}{\rho} \\ Z_\mathcal{S} = \frac{{}^cZ}{\rho} \end{cases}, \ \rho = ||{}^c\mathbf{X}||, \tag{3.22} $$

before projecting it on the normalized image plane in $\mathbf{x}$, using a second center of projection $\mathbf{C} \in \mathbb{R}^3$, distant of $\xi \in \mathbb{R}_+$ from the first one, along the axis $\mathbf{Z}_c$ [3]:

$$ \mathbf{x} = pr\left(\mathbf{X}_\mathcal{S} + \begin{pmatrix} 0 & 0 & \xi \end{pmatrix}^\mathsf{T}\right) \quad \text{with} \quad \begin{cases} x = \frac{X_\mathcal{S}}{Z_\mathcal{S}+\xi} \\ y = \frac{Y_\mathcal{S}}{Z_\mathcal{S}+\xi} \end{cases}. \tag{3.23} $$

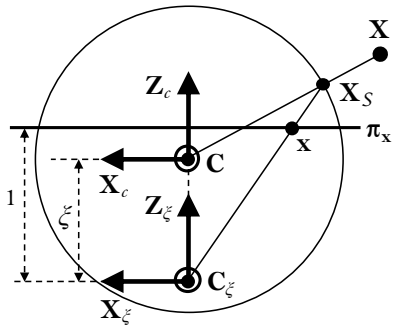


Figure 3.3: Diagram of the unified central projection model.

$pr_{\mathcal{S}}()$ and $pr()$ combine into a single projection function pr_ξ of the 3D point ${}^c\mathbf{X}$ into $\tilde{\mathbf{x}}$:

$$\tilde{\mathbf{x}} = pr_\xi({}^c\mathbf{X}) \quad \text{with} \quad \begin{cases} x = \frac{{}^cX}{{}^cZ+\xi\rho} \\ y = \frac{{}^cY}{{}^cZ+\xi\rho} \end{cases} . \tag{3.24}$$

To finalize the unified central projection model, the transformation from the normalized image plane to the digital image plane is done in the same way as with the perspective (3.4) projection model. ξ joins the set of intrinsic parameters of the unified central projection model $\gamma_u = \{\alpha_u, \alpha_v, u_0, v_0, \xi\}$. Thus, the projection function of a 3D point ${}^c\mathbf{X}$ in the digital image plane is written:

$$\tilde{\mathbf{u}} = pr_{\gamma_u}({}^c\mathbf{X}) = \mathbf{K} pr_\xi({}^c\mathbf{X}). \tag{3.25}$$

The projection $pr_\xi()$ of the sphere to the image plane is invertible, which makes it possible to express a spherical point, and thus the associated line of sight, from an image point:

$$\mathbf{X}_{\mathcal{S}} = pr_\xi^{-1}(\mathbf{x}) = \begin{pmatrix} \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} x \\ \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} y \\ \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} - \xi \end{pmatrix} . \tag{3.26}$$

The equation (3.25) of projection to the digital image plane is equivalently rewritten by introducing the intrinsic parameters $\tau \in [0, 1[$, $\alpha'_u \in \mathbb{R}^*$ and $\alpha'_v \in \mathbb{R}^*$ and by posing $xi = \tau/(1-\tau)$, $\alpha_u = \alpha'_u/(1-\tau)$ and $\alpha_v = \alpha'_v/(1-\tau)$ [51]. We then express the equivalent unified central projection function $pr_{\gamma'_u}$:

$$\tilde{\mathbf{u}} = pr_{\gamma'_u}({}^c\mathbf{X}) = \mathbf{K}' pr_\tau({}^c\mathbf{X}) \quad \text{with} \quad \mathbf{K}' = \begin{pmatrix} \alpha'_u & 0 & u_0 \\ 0 & \alpha'_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{3.27}$$

and $pr_\tau({}^c\mathbf{X})$ giving:

$$\begin{cases} x = \frac{{}^cX}{(1-\tau){}^cZ+\tau\rho} \\ y = \frac{{}^cY}{(1-\tau){}^cZ+\tau\rho} \end{cases} . \tag{3.28}$$

This rewriting of the unified central projection model with $\gamma'_u = \{\alpha'_u, \alpha'_v, u_0, v_0, \tau\}$, inverts and has better numerical properties for the calibration [51].

Finally, representing the coordinates of points on the sphere by Cartesian coordinates is redundant because, since the sphere is unitary, $X_{\mathcal{S}}$, $Y_{\mathcal{S}}$ and $Z_{\mathcal{S}}$ are not independent ($||\mathbf{X}_{\mathcal{S}}|| = 1$). The minimal representation of a point on the sphere is done by the spherical coordinates of azimuthal ϕ and polar θ angles[3] ($\boldsymbol{\theta} = [\theta, \phi]^\mathsf{T} \in \mathbb{R}^2$), which are expressed from $\mathbf{X}_{\mathcal{S}}$ by the function $c2s()$ (Cartesian to spherical):

$$\boldsymbol{\theta} = \begin{pmatrix} \theta \\ \phi \end{pmatrix} = c2s(\mathbf{X}_{\mathcal{S}}) = \begin{pmatrix} \arccos(Z_{\mathcal{S}}) \\ \arctan(Y_{\mathcal{S}}/X_{\mathcal{S}}) \end{pmatrix} . \tag{3.29}$$

[3]This chapter considers the standard ISO 80000-2:2019 to note spherical coordinates.

Remark 2 (Enrichment of the unified central projection model) *Similar to the perspective projection model, the unified central projection model can be extended by taking into account additional intrinsic parameters, e.g. for radial distortions (Remark 1).* ◇

Remark 3 (Generalization of the perspective projection model) *The unified central projection model is valid for any single point of view camera, including perspective cameras. Indeed, it is enough to cancel ξ (or τ) to find back the perspective projection.* ◇

Extensions for fisheyes cameras

The unified central projection model (Sec. 3.3.3) is also equivalent to the ad hoc projection models of fisheye lens cameras [13] seen in Section 3.3.2. In practice, this model needs to be completed by a distortion parameter to better approximate the projection of most fisheye lenses [53]. However, this model poorly characterizes fisheye lenses whose field of view is greater than 180 degrees [51].

In this case, the unified central projection model can be extended by adding, after the spherical projection (3.22), a second spherical projection, of center $\mathbf{C}_\xi$, before the perspective projection to the image plane of center $\mathbf{C}_\tau = (0, 0, \tau/(1-\tau) - \xi)^\mathsf{T}$, which thus becomes the third center of projection. The equation of this double-sphere projection model to the digital image plane is then written from the two versions (3.25) and (3.27) of the unified central projection [51]:

$$\tilde{\mathbf{u}} = pr_{\gamma_d}({}^c\mathbf{X}) = \mathbf{K}' pr_\tau(pr_\xi({}^c\mathbf{X})), \tag{3.30}$$

such that the composition of projection functions $pr_\tau(pr_\xi({}^c\mathbf{X}))$ gives:

$$\begin{cases} x = \frac{{}^cX}{(1-\tau)({}^cZ+\xi\rho)+\tau\rho_2} \\ y = \frac{{}^cY}{(1-\tau)({}^cZ+\xi\rho)+\tau\rho_2} \end{cases}, \tag{3.31}$$

where ρ is defined in Equation (3.6), and with:

$$\rho_2 = \sqrt{{}^cX^2 + {}^cY^2 + ({}^cZ + \xi\rho)^2}. \tag{3.32}$$

This last model has therefore 6 intrinsic parameters $\gamma_d = \{\alpha'_u, \alpha'_v, u_0, v_0, \xi, \tau\}$.

Extension for 360° cameras

The most compact spherical polydioptric systems have the specificity of being designed so that their lenses have complementary fields of view, thus reducing their number to a minimum. Two fisheye lenses placed back-to-back are enough to cover the full 360 degrees of the spherical field of view thanks to two flat mirrors judiciously positioned between the lenses, thus reflecting the light to a single photosensitive matrix [28]. For even more compactness, the camera manufacturer Ricoh has replaced, in its Theta[4] brand, the mirrors by two

[4] `http://theta360.com`

prisms redirecting the light rays passing through the fisheye lenses towards two photosensitive matrices.

Beyond the interest of the very small volume of this spherical polydioptric camera (Fig. 3.4a), for the general public, as well as in robotics, the proximity of the two fisheye lenses also makes their optical centers very close. We can then make the approximation that they are at the same location, in particular when the elements of the observed scene are far enough from the spherical camera[5] (minimum working distance of polycameras [49]).

The projection model of a compact polydioptric spherical camera is thus restricted to a single sphere, each hemisphere of which is associated with one of the two fisheye images (Fig. 3.4b), thus considering two image planes associated with the same sphere [11]. As the lenses, the photosensitive matrices and their alignment can be slightly different from one fisheye camera to the other, two sets of intrinsic parameters $\gamma_{u,j}$, $j \in \{1, 2\}$ are considered. However, since both fisheye cameras are assumed to share the same origin, we can set $\mathcal{F}_s = \mathcal{F}_{c_1}$ (Fig. 3.4c) and the extrinsic parameters, *i.e.* the pose of the second fisheye camera, relative to the first, thus, are reduced to the orientation $\mathbf{r}_{s,2}$ (or $\mathbf{r}_{1,2}$) ${}^{c_2}\mathbf{R}_s = {}^{c_2}\mathbf{R}_{c_1} \in SO(3)$. We then re-express the projection of a spherical point ${}^{s}\mathbf{X}_{\mathcal{S}}$ of the hemisphere associated to the camera j in the normalized image plane of the latter (Eq. (3.23)) by:

$$\mathbf{x} = pr_j \left({}^{c_j}\mathbf{R}_s \, {}^{s}\mathbf{X}_{\mathcal{S}} + \begin{pmatrix} 0 & 0 & \xi_j \end{pmatrix}^{\mathsf{T}} \right), \tag{3.33}$$

with ${}^{c_1}\mathbf{R}_s = {}^{c_1}\mathbf{R}_{c_1} = \mathbf{I}_{3\times 3}$. In practice, the sign of the third coordinate of ${}^{c_j}\mathbf{X}_s = {}^{c_j}\mathbf{R}_s \, {}^{s}\mathbf{X}_S$ is sufficient to determine which of the two cameras perceives it.

3.3.4 Generic Models

Generic models, also called discrete models, associate to each pixel a line of sight, also called light ray [46, 41] or raxel [18]. These rays intersect at a single point, the optical center, for central cameras. They intersect at the same line, the camera axis, for axial cameras. Finally, the rays are not constrained for the other non-central cameras [41].

For some axial cameras, the generic model can be quasi-central when the intersection points of the rays and the camera axis form a short segment [7]. The projection of a 3D point ${}^{c}\mathbf{X}$ is then a ray $\boldsymbol{\psi} = [\theta, \phi, \delta_Z]^{\mathsf{T}} \in \mathbb{R}^3$ where θ and ϕ are the polar and azimuthal angles (see (3.29)) of the line of sight formed by ${}^{c}\mathbf{X}$ and the point ${}^{c}[0, 0, \delta_Z]^{\mathsf{T}}$ [7]:

$$\boldsymbol{\psi} = \begin{bmatrix} \theta \\ \phi \\ \delta_Z \end{bmatrix} = \begin{bmatrix} c2s \left(\dfrac{{}^{c}\mathbf{X} - {}^{c}[0, 0, \delta_Z]^{\mathsf{T}}}{||{}^{c}\mathbf{X} - {}^{c}[0, 0, \delta_Z]^{\mathsf{T}}||} \right) \\ \delta_Z \end{bmatrix}. \tag{3.34}$$

[5] The Theta brand of Ricoh is mentioned here as an example because it has one of the smallest distances on the market between its fisheyes lenses. It is thus a spherical camera for which the approximation of uniqueness of optical center for the two fisheye lenses is among the most tolerable.

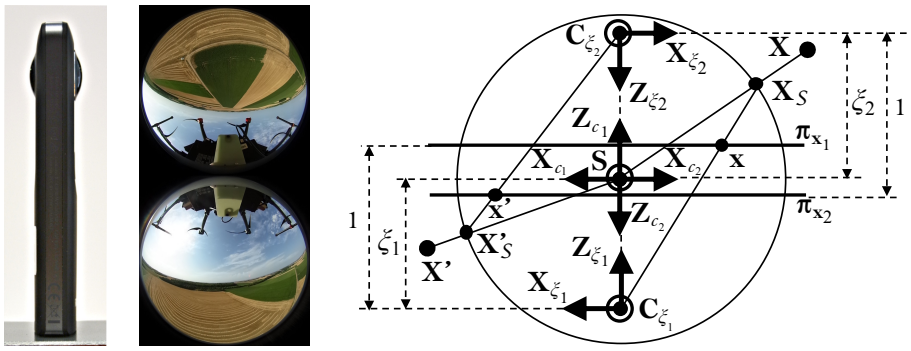


Figure 3.4: Spherical vision: (a) Compact polydioptric spherical camera (Ricoh Theta S, side view with the two fisheye lenses in its upper part) acquiring (b) double-fisheye images (Le champ à cailloux, Vaux-en-Amiénois, 2019) and whose projection can be represented, under hypotheses, by (c) an extension of the unified central projection model to two image planes.

When the camera is central, $\delta_Z = 0$.

δ_Z is possibly unique for each pixel $\mathbf{u}$ and the projection model is a table associating to each pixel $\mathbf{u}$ a ray $\boldsymbol{\psi}$.

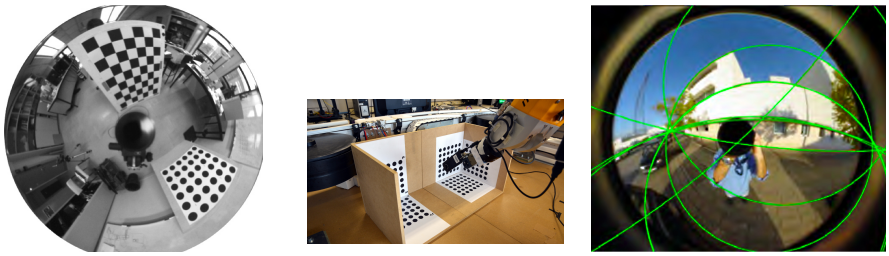
3.4 Calibration methods

Calibration methods exploit points (most methods), straight lines [16, 4, 6], circles [27] or spheres [52]. These *primitives* observed by the camera on calibration patterns (Fig. 3.5a, 3.5b, 3.5c) or in the scene (Fig. 3.5d) are the data used to estimate the intrinsic parameters of the projection model characterizing the camera. The calibration patterns can be two-dimensional (2D) or 3D, the former requiring to be observed at several distinct exposures (to be added in the parameters to be estimated) to perform the calibration, when a single observation of a 3D calibration pattern may suffice [40]. Without requiring a calibration pattern or the presence of a particular structure in the scene, self-calibration relies on the detection and matching of points of interest in the scene in several images acquired at different camera exposures [26, 34, 36].

Most of the available methods and software estimate the intrinsic parameters minimizing the reprojection error (Sum of Squared Differences, SSD) of reference calibration patterns ${}^o\mathbf{X}_i$ with respect to their detection in the image $\mathbf{u}_i$:

$$\hat{\gamma}_\cdot = \underset{\gamma_\cdot}{\operatorname{argmin}} \frac{1}{2} \sum_i ||pr_{\gamma_\cdot}({}^c\mathbf{M}_o \, {}^o\mathbf{X}_i) - \mathbf{u}_i||^2, \tag{3.35}$$

$\hat{\gamma}_\cdot$ being the estimate of the intrinsic parameters according to the considered projection model: $\gamma_\cdot \in \{\gamma_p, \gamma_u, ...\}$ (Sec. 3.3). In practice, as many ${}^c\mathbf{M}_o$ as



(a) (b) (c) (d)

Figure 3.5: Examples of measurements in the image for calibration. (a) checkerboard and disk plane patterns in a paracatadioptric [9] image: primitives are points (corners of squares or centers of disks). (b-c) 3D multi-planar pattern observed by a 360 polydioptric camera [11]. (d) straight lines extracted from the environment structure (green lines) [6].

there are calibration patterns must be computed when solving the minimization problem (3.35), even if they are not used afterwards.

As for the solution of the minimization problem (3.35), it is usually iterative by nonlinear optimization [27, 51, 9, 33, 45, 11] by a Newton, Gauss-Newton or Levenberg-Marquardt type method. Indeed, the formation of omnidirectional images is nonlinear. On the other hand, a nonlinear optimization ensures a better management of erroneous or inaccurate measurements [51] than a linear resolution method, being able nevertheless, classically, to initialize the intrinsic parameters to ensure the convergence of the optimization method [39].

Some variants are based on a collinearity criterion, not only when the considered primitives are straight lines [4], but also for the lines of sight associated to the pixels of the omnidirectional image, whether they form the projection model [43, 7] or not [39].

Table 3.2 gathers the most used and among the most recent omnidirectional camera calibration methods (see [40] to complete the oldest references). Most of them are associated with free software [12, 14, 25, 5, 22, 38, 29, 37] and some of them are even integrated to other libraries like OpenCV[6]. In addition to the attributes already mentioned (use of calibration pattern, type of primitive, criterion considered, solver) to classify the calibration methods, the maximum field of view of the camera calibrated with each method is reported, when known. At most, a monocular camera with a field of view of 280 degrees and a polydioptric

[6] `https://opencv.org`

camera with a field of view of 360 degrees have been calibrated[7].

Finally, the calibration methods reported in Table 3.2 consider various projection models:

- [24, 42] use perspective projection (Sec. 3.3.1) with distortions for multi-camera systems (Sec. 3.3.2)
- [28] uses the equidistant fisheye projection (Sec. 3.3.2, Eq. (3.13))
- [27, 45] are based on an angular polynomial projection model (Sec. 3.3.2, Eq. (3.17)), combined with a quadric-based non-central projection model [1] for [45]
- [44] is based on the Cartesian polynomial projection model (Sec. 3.3.2, Eq. (3.19))
- [4, 39, 9, 33, 11] use the unified central projection model (Sec. 3.3.3)
- [51] is based on the double sphere projection model (Sec. 3.3.3)
- [7] is based on a generic projection model (Sec. 3.3.4)

3.5 Conclusion of the chapter

This chapter has presented the most frequently encountered projection models in omnidirectional vision. Faced with the variety of these models, the choice depends both on the camera used but also on the intended application, particularly according to criteria of accuracy, efficiency and image representation. The calibration method, its ease of implementation, and even the availability of software, are all practical elements to take into account. Nevertheless, all calibration methods require a distribution of the considered primitives in the whole field of view of the camera, particularly at its periphery, to ensure a correct estimation of the intrinsic parameters of the camera [40].

Once the camera is calibrated, the representation of the omnidirectional image can be transformed, for example from the acquired image plane to rectified planes so that the straight lines of the scene become straight and not curved in the image [4], or to a spherical image [32]. This issue of omnidirectional image representation depends on the algorithms that exploit it, whether they are video surveillance applications [50], 3D reconstruction [47], virtual reality [21], robotics [10] or autonomous vehicles [20].

[7] Only polydioptric camera calibration methods leading to an omnidirectional image like monocular are mentioned. For the others, see for example [30] and references.

Table 3.2: Main calibration methods and their characteristics: use of pattern or not, primitive and criterion considered, solver, maximum calibrated field of view (in degrees) and available software.
Abbreviations: *lin* for *linear*, *reproj* for *reprojection*, *SSD* for *Sum of Squared of Differences*, *rSSD* for *robust SSD*, *opt* for *optimization*, *LM* for *Levenberg-Marquardt*, *GN* for *Gauss-Newton*, *alt* for *alternate*.
Notes: † lacks specification; * maximum assumed by following that of [9] because the methods are similar or the same cameras are calibrated.

method	**pattern**	**primitive**	**criterion**	**solver**	**max field**	**software**
[4]	no	lines	colinearity	linear	180†	[12]
[39]	yes	points	colinearity and reproj/SSD	lin and opt/LM	180†	[14]
[27]	yes	points/ circles	reproj/SCE	opt/LM	190	[25]
[51]	yes	points	reproj/SrCE	opt/GN	195	[5]
[9]	yes	points	reproj/SCE	opt/LM	210	[22]
[33]	yes	points	reproj/SCE	opt/LM	210*	[38]
[45]	yes	points	reproj/SCE	opt	210*	[29]
[43]	yes	points	colinearity and reproj/SCE	lin/alt and opt/LM	220	[37]
[7]	yes	points (denses)	colinearity and distance	opt	280	no
[28, 11] [24, 42]	yes	points	reproj/SCE	opt/LM	360	no

Bibliography

[1] A. Agrawal, Y. Taguchi, and S. Ramalingam. Beyond alhazen's problem: Analytical projection model for non-central catadioptric cameras with quadric mirrors. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 2993–3000, 2011.

[2] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *Int. Journal on Computer Vision*, 35(2):175–196, 1999.

[3] J. P. Barreto and H. Araujo. Issues on the geometry of central catadioptric imaging. In *IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, USA, Dec. 2001.

[4] J. P. Barreto and H. Araujo. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(8):1327–1333, 2005.

[5] Basalt. The double sphere camera model calibration software (c++). `https://gitlab.com/VladyslavUsenko/basalt`, 2018. Author: V. Usenko.

[6] Jesus Bermudez-Cameo, Gonzalo López-Nicolás, and Josechu J. Guerrero. Automatic line extraction in uncalibrated omnidirectional cameras with revolution symmetry. *Int. J. Comput. Vis.*, 114(1):16–37, 2015.

[7] Pierre-Andre Brousseau and Sebastien Roy. Calibration of axial fisheye cameras through generic virtual central models. In *IEEE/CVF International Conference on Computer Vision*, Seoul, Corée du Sud, Oct. 2019.

[8] A. M. Bruckstein and T. J. Richardson. Omniview cameras with curved surface mirrors. In *IEEE Workshop on Omnidirectional Vision*, pages 79–84, Hilton Head Island, USA, Jui. 2000.

[9] Guillaume Caron and Damien Eynard. Multiple camera types simultaneous stereo calibration. In *IEEE Int. Conf. on Robotics and Automation*, pages 2933–2938, Shanghai, China, May 2011.

[10] Guillaume Caron, Eric Marchand, and El Mouaddib. Photometric visual servoing for omnidirectional cameras. *Autonomous Robots*, 35(2-3):177–193, 2013.

[11] Guillaume Caron and Fabio Morbidi. Spherical Visual Gyroscope for Autonomous Robots using the Mixture of Photometric Potentials. In *IEEE Int. Conf. on Robotics and Automation*, pages 820–827, Brisbane, Australie, May 2018.

[12] CatPack. Catpack toolbox: Matlab software package for the calibration of central catadioptric cameras using line images. `https://home.deec.uc.pt/~jpbar/CatPack/main.htm`, 2002. Authors: J. P. Barreto, H. Araujo.

[13] Jonathan Courbon, Youcef Mezouar, and Philippe Martinet. Evaluation of the Unified Model of the Sphere for Fisheye Cameras in Robotic Applications. *Advanced Robotics*, 26(8-9):947–967, 2012.

[14] DLTOmniCalibration. Direct linear transform-like omnidirectional camera calibration. `http://webdiis.unizar.es/~lpuig/DLTOmniCalibration`, 2011. Authors: L. Puig, Y. Bastanlar, P. Sturm, J. J. Guerrero and J. Barreto.

[15] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical applications. In *European Conf. on Computer Vision*, Dublin, Irelande, Jui. 2000.

[16] C. Geyer and K. Daniilidis. Paracatadioptric camera calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(5):687–695, 2002.

[17] J. J. Gonzalez-Barbosa. *Vision panoramique pour la robotique mobile : stéréovision et localisation par indexation d'images*. PhD thesis, Université Toulouse III, Ecole doctorale en informatique et télécomunications, 2004.

[18] M. D. Grossberg and S. K. Nayar. A general imaging model and a method for finding its parameters. In *IEEE Int. Conf. on Computer Vision*, volume 2, pages 108–115, 2001.

[19] W.R. Hamilton. Theory of systems of rays. *Trans. of the Royal Irish Academy*, 15:69–174, 1828.

[20] L. Heng, B. Choi, Z. Cui, M. Geppert, S. Hu, B. Kuan, P. Liu, R. Nguyen, Y. C. Yeo, A. Geiger, G. H. Lee, M. Pollefeys, and T. Sattler. Project autovision: Localization and 3d scene perception for an autonomous vehicle with a multi-camera system. In *Int. Conf. on Robotics and Automation*, pages 4695–4702, 2019.

[21] J. Huang, Z. Chen, D. Ceylan, and H. Jin. 6-dof vr videos with a single 360-camera. In *IEEE Virtual Reality*, pages 37–44, 2017.

[22] Hyscas. Hyscas: Hybrid stereoscopic calibration software. `https://mis.u-picardie.fr/~g-caron/software`, 2011. Authors: G. Caron, D. Eynard.

[23] S.-H. Ieng and R. Benosman. *Geometric Construction of the Caustic Surface of Catadioptric Non-Central Sensors*, pages 39–54. Springer Netherlands, 2006.

[24] Sei Ikeda, Tomokazu Sato, and Naokazu Yokoya. Calibration method for an omnidirectional multicamera system. In *Stereoscopic Displays and Virtual Reality Systems X*, volume 5006, pages 499 – 507. SPIE, 2003.

[25] Kalibr. Kalibr: C++ calibration toolbox. `https://github.com/ethz-asl/kalibr`, 2014. Authors: P. Furgale, H. Sommer, J. Maye, J. Rehder, T. Schneider, L. Oth.

[26] Kang, S. B. Catadioptric self-calibration. In *IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 201–207, Hilton Head Island, USA, Jui. 2000.

[27] Juho Kannala and Sami S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(8):1335–1340, 2006.

[28] S. Li. Full-View Spherical Image Camera. In *IEEE Int. Conf. Pattern Recogn.*, volume 4, pages 386–390, 2006.

[29] Libomnical. Libomnical: Omnidirectional camera calibration. `http://www.cvlibs.net/projects/omnicam/`, 2014. Authors: M. Schonbein, T. Strauss, A. Geiger.

[30] Albert Liu, Steve Marschner, and Noah Snavely. Caliber: Camera localization and calibration using rigidity constraints. *Int. Journal of Computer Vision*, 118(1):1–21, 2016.

[31] Y. Ma, S. Soatto, J. Košecká, and S. Sastry. *An invitation to 3D vision.* Springer, 2004.

[32] A. Makadia and K. Daniilidis. Rotation recovery from spherical images without correspondences. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(7):1170–1175, 2006.

[33] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In *IEEE Int. Conf. on Robotics and Automation*, pages 3945–3950, Rome, Italie, Avr. 2007.

[34] B. Micusik and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(7):1135–1149, 2006.

[35] K. Miyamoto. Fish eye lens. *Journal of the Optical Society of America*, 54(8):1060–1061, 1964.

[36] Thanh-Tin Nguyen and Maxime Lhuillier. Self-calibration of omnidirectional multi-cameras including synchronization and rolling shutter. *Computer Vision and Image Understanding*, 162:166 – 184, 2017.

[37] OCamCalib. Ocamcalib: Omnidirectional camera calibration toolbox for matlab. `https://sites.google.com/site/scarabotix/ocamcalib-toolbox`, 2013. Author: D. Scaramuzza.

[38] OmniCalib. Improved omnidirectional calibration toolbox. `http://www.robots.ox.ac.uk/~cmei/Toolbox.html`, 2007. Author: C. Mei.

[39] Luis Puig, Yalin Bastanlar, Peter Sturm, Josechu Guerrero, and Joao Barreto. Calibration of Central Catadioptric Cameras Using a DLT-Like Approach. *Int. Journal of Computer Vision*, 93(1):101–114, 2011.

[40] Luis Puig, J. Bermúdez, Peter Sturm, and J.J. Guerrero. Calibration of omnidirectional cameras in practice: A comparison of methods. *Computer Vision and Image Understanding*, 116(1):120 – 137, 2012.

[41] Srikumar Ramalingam and Peter Sturm. A Unifying Model for Camera Calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 39(7):1309–1319, 2017.

[42] J. Y. Rau, B. W. Su, K. W. Hsiao, and J. P. Jhan. Systematic calibration for a backpacked spherical photogrammetry imaging system. *ISPRS - Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1:695–702, 2016.

[43] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *IEEE Int. Conf. on Computer Vision Systems*, pages 45–53, New York, USA, Mar. 2006.

[44] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A toolbox for easily calibrating omnidirectional cameras. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 5695–5701, Pékin, Chine, Oct. 2006.

[45] Miriam Schönbein, Tobias Strauss, and Andreas Geiger. Calibrating and centering quasi-central catadioptric cameras. In *IEEE Int. Conf. on Robotics and Automation*, pages 4443–4450, Hong Kong, Jui. 2014.

[46] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, Simone Gasparini, and Joao Barreto. Camera Models and Fundamental Concepts Used in Geometric Computer Vision. *Foundations and Trends in Computer Graphics and Vision*, 6(1-2):1–183, 2011.

[47] Shinya Sumikura, Mikiya Shibuya, and Ken Sakurada. OpenVSLAM: A Versatile Visual SLAM Framework. In *ACM Int. Conf. on Multimedia*, pages 2292–2295, Nice, France, Oct. 2019.

[48] R. Swaminathan, M. D. Grossberg, and S. K. Nayar. Caustics of catadioptric cameras. In *IEEE Int. Conf. on Computer Vision*, volume 2, pages 2–9, Vancouver, Canada, Jui. 2001.

[49] R. Swaminathan and S. K. Nayar. Polycameras: camera clusters for wide angle imaging. Technical Report CUCS-013-99, Columbia university, Computer science, Avr. 1999.

[50] Yazhe Tang, Zhi Gao, Feng Lin, Y.F. Li, and Fei Wen. Visual adaptive tracking for monocular omnidirectional camera. *Journal of Visual Communication and Image Representation*, 55:253 – 262, 2018.

[51] V. Usenko, N. Demmel, and D. Cremers. The double sphere camera model. In *Int. Conf. on 3D Vision*, Vérone, Italie, Sep. 2018.

[52] Xianghua Ying and Zhanyi Hu. Catadioptric camera calibration using geometric invariants. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(10):1260–1271, 2004.

[53] X. Ying and Z. Hu. Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model. In *European Conf. on Computer Vision*, pages 442–455, Prague, République Tchèque, May 2004.