HAL
open science

Calibration of Cameras with Radially Symmetric Distortion

Jean-Philippe Tardif, Peter Sturm, Martin Trudeau, Sébastien Roy

To cite this version:

Jean-Philippe Tardif, Peter Sturm, Martin Trudeau, Sébastien Roy. Calibration of Cameras with Radially Symmetric Distortion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31 (9), pp.1552-1566. 10.1109/TPAMI.2008.202 . inria-00384203

HAL Id: inria-00384203

https://inria.hal.science/inria-00384203

Submitted on 14 May 2009

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.

Calibration of Cameras with Radially Symmetric Distortion

Jean-Philippe Tardif, Peter Sturm, Martin Trudeau, and Sébastien Roy

***Abstract*— We present algorithms for plane-based calibration of general radially distorted cameras. By this we understand cameras that have a distortion center and an optical axis such that the projection rays of pixels lying on a circle centered on the distortion center, form a right viewing cone centered on the optical axis. The camera is said to have a single viewpoint (SVP) if all such viewing cones have the same apex (the optical center), otherwise we speak of NSVP cases. This model encompasses the classical radial distortion model [4], fisheyes and most central or non-central catadioptric cameras.**

Calibration consists in the estimation of the distortion center, the opening angles of all viewing cones and their optical centers. We present two approaches of computing a full calibration from dense correspondences of a single or multiple planes with known Euclidean structure. The first one is based on a geometric constraint linking viewing cones and their intersections with the calibration plane (conic sections). The second approach is an homography-based method. Experiments using simulated and a broad variety of real cameras show great stability. Furthermore, we provide a comparison with Hartley-Kang's algorithm [14], which however can not handle such a broad variety of camera configurations, showing similar performance.

***Index Terms*— Calibration, omnidirectional vision, fisheye, catadioptric camera.**

I. Introduction

In the last few years, we have seen an increasing interest in non-conventional cameras and projection models, going beyond affine or perspective projection. There exists a large diversity of camera models, many of which specific to certain types of projections [17], [20], others applicable to families of cameras such as central catadioptric systems [1], [2], [7], [9]. All these models are described by a few intrinsic parameters, much like the classical pinhole model, possibly enhanced with radial or decentering distortion coefficients. Calibration methods exist for all these models, and they are usually tailor-made for them, *i.e.* can not be used for any other projection model [18]. Several works address the calibration problem from an opposite point of view, by adopting a very generic imaging model that incorporates most commonly used cameras [5], [10], [11], [23], [25].

In the most general case, cameras are modeled by attributing an individual projection ray to each pixel. Such a model is highly expressive, but it is difficult to obtain a stable calibration of cameras with it, at least with few input images taken in an uncontrolled environment. Finally, several researchers have proposed a compromise between parametric and such generic models. They assume radial symmetry around the distortion center (often considered as coinciding with the principal point), but with a general distortion function [14], [27]–[29], [31], [32].

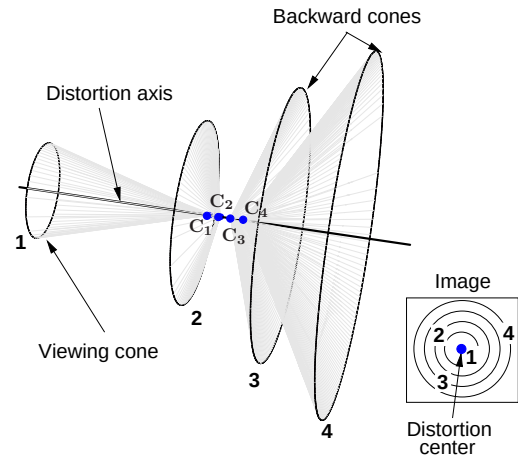


Fig. 1. Our camera model in the case of a non-central projection (see text for explanations). The inlayed illustrations show the distortion center (in blue) and four distortion circles and their corresponding viewing cone and optical center.

In this paper, we use a model of this category which we find sufficiently general to model many common types of cameras. By having fewer parameters than the fully generic model, calibration remains easy and stable. It encompasses many common camera models, such as pinhole (modulo aspect ratio and skew), the classical polynomial radial distortion model, fisheyes, or any catadioptric system whose mirror is a surface of revolution, and for which the optical axis of the perspective (or affine) camera looking at it is aligned with the mirror's revolution axis. We describe our model in the following and refer to Fig. 1 for an illustration. Cameras are modeled using the notion of a **distortion center** in the image whose back-projection yields the **optical axis** in 3D. For cameras with **radially symmetric distortion**, the projection rays associated with pixels lying on a same **distortion circle** centered on the distortion center, lie on a right **viewing cone** centered on the optical axis. In the following, we denote by d the radius of a distortion circle; for ease of expression, we also speak of distortion circle d, meaning the distortion circle of radius d.

For a perspective camera with focal length f, the opening angle of a distortion circle's viewing cone, is given by $2\arctan\frac{d}{f}$. Hence, the opening angles of all viewing cones are "dictated" by the camera's focal length. In our model, this is generalized: in the most general case, one may not assume any relation between opening angles of different viewing cones. In other words, we may have one individual focal length f_d per distortion circle d. We may define a **focal length function** $f_d = f(d)$. In practice, we handle such a general model in two

different ways: by assuming a polynomial form for the focal length function, or by subsampling it and estimating/using one focal length per discrete sample of d. We also consider its reciprocal, the **distortion function** that maps the opening angle of a viewing cone to the radius of the associated distortion circle.

Let us introduce a few additional notations. A line spanned by a pixel and the distortion center is called a **radial line** and the plane spanned by the projection ray associated with the pixel and the optical axis, a **radial plane**. In our camera model, we specify that angles between radial lines and associated radial planes are identical (otherwise, it would be a very uncommon camera). Our model comprises **central** cameras (SVP), where all viewing cones have the same apex (the **optical center**), but also **non-central** ones (NSVP), for which the viewing cones' apexes lie anywhere on the optical axis. In the latter case, we may speak of one optical center per viewing cone. For convenience, these terms are summarized in Table I.

Problem statement. We want to calibrate cameras based on the above model from one or several images of a calibration plane in unknown positions. Calibration consists in estimating the camera's focal length function or, equivalently, its distortion function. Further, for the NSVP case, we have to estimate the optical centers associated with all distortion circles.

The problem of calibration from a planar scene of known geometry has been studied intensively in computer vision. It is widely accepted that one seeks a solution that minimizes the sum of reprojection errors of the points on the calibration planes. This being a non-convex function, iterative methods must be used. Besides the problem of convergence to a local minimum, they require some estimate for the internal parameters of the camera, generally provided by the camera manufacturer [15], [16]. Both these difficulties can be overcome by non-iterative approaches relying on algebraic error functions as long as they provide good accuracy. Then, these results can be refined iteratively using a geometric error.

Contributions. Our contributions take the form of two non-iterative calibration algorithms. Their input is the Euclidean structure of the plane(s) and a dense or sparse matching between the plane(s) and the camera image. The first approach is based on algebraic constraints relating the projection of the points on the calibration plane on the image. It is based on the observation that each distortion circle and the associated viewing cone can be considered as an individual perspective camera. Hence, the projection of a calibration plane to the image, when reduced to a single distortion circle, can be expressed by an homography and we can readily apply plane-based calibration algorithms designed for perspective cameras. Our second approach is based on a detailed geometric analysis. Consider one image of a calibration plane; each viewing cone of the camera intersects the calibration plane in a conic, which we call **calibration conic**. We describe geometric constraints relating calibration conics to the orientation and position of the camera as well as its intrinsic parameters; these are at the basis of our second calibration approach. Our first approach usually performs best in practice. Furthermore, it can accurately estimate the projection model for a non-central camera. Based on our experimental results, we argue that in the case of a general model such as the one we propose, an approach directly estimating a non-central projection gives better accuracy than an approach based on: 1) estimating an approximated central projection, 2) iterative refinement using a non-central model [18].

Organization. A geometric study of our model is presented in Section II. Our first calibration approach, akin to plane-based calibration of perspective cameras, is described in Section III followed by our second, more geometric, approach in Section IV. Our algorithms assume a known position of the distortion center, but we also show how to estimate it, using repeated calibrations for different candidates, cf. Section V. Several practical issues and experimental results, including a comparison with Hartley-Kang's methods [14], are presented in Section VI and Section VII, respectively. Finally, we conclude in Section VIII.

II. Geometry

In this section, we present geometric constraints for our camera model. Readers interested in implementing the homography-based approach may wish to skip to Sections II-C and III.

A. One Distortion Circle in One Image of a Calibration Plane

Let us consider one distortion circle in the image plane. Its associated viewing cone cuts the calibration plane in a **calibration conic**. From dense matches between image and calibration plane, we can compute this calibration conic (see Section VI for more on the dense matching). This conic can be either an ellipse or a hyperbola (the parabola is only a theoretically interesting case). If we knew the position of the camera's optical center relative to the calibration plane, then we could directly compute the viewing cone of our distortion circle, i.e. the cone that has the optical center as apex and that contains the above calibration conic. As mentioned, that cone has several useful properties: its axis is the camera's optical axis and it is a right cone, *i.e.* rotationally symmetric with respect to its axis. From the cone, the focal length of the considered distortion circle can be easily computed: the cone's opening angle equals the field of view.

In practice, we do not know the optical center's position relative to the calibration plane. In the following, we show geometrical relations between the calibration conic, the optical center and the optical axis of the camera, that allow to compute the optical center. Recall that when talking about optical center, we mean the optical center per distortion circle; they all lie on the optical axis and in the SVP case, they are identical.

Without loss of generality, we assume that the calibration plane is the plane $Z = 0$, and that the calibration conic ε is given by the matrix $\varepsilon \propto \mathrm{diag}(a, b, -1)$ with $|b| \geq a > 0$ ($\propto$ represents equality up to scale), *i.e.* the X-axis is the conic's major axis. The type of the calibration conic ε depends on a and b as follows:

$$\begin{cases} b \geq a > 0 & \text{ellipse} \\ b < 0 \ \text{ and } \ |b| > a > 0 & \text{hyperbola.} \end{cases} \tag{1}$$

TABLE I

SUMMARY OF THE TERMS RELATED TO OUR CAMERA MODEL.

Terms	Description
Optical axis	A line perpendicular to the image plane
Distortion center	The intersection between the image plane and the optical axis
Distortion circle	A circle in the image centered in the distortion center
Viewing cone	Back-projection of a distortion circle; a right cone centered in the optical axis
Calibration conic	Intersection of a viewing cone and a calibration plane
Radial line	A line passing through the distortion center
Radial plane	Back-projection of a radial line
Central projection	All viewing cones have identical apex, the optical center
Non-central projection	The viewing cones only share the same (optical) axis

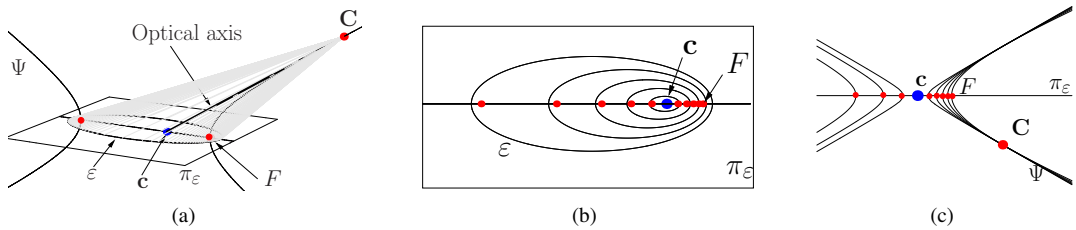


Fig. 2. Illustrations of the geometry of viewing cones, calibration conics (here ellipses) and location of optical center in the SVP case. **(a)** Complete illustration for one viewing cone. **(b)** View of the calibration plane, showing many cones' calibration ellipses. Note that their major axes are collinear. **(c)** Side view of the viewpoint conics (hyperbolas in this case) associated with many calibration conics.

Our aim is to provide constraints on the position of the optical center, as well as on the orientation of the optical axis, from this conic. We first do this for the case of a calibration ellipse, then for the case of a hyperbola.

The case of calibration ellipses. Let us first state a well-known result. Consider a right cone whose apex is a point with real coordinates, and its intersection with a plane. For now, we assume that the intersection is an ellipse (the case of the hyperbola will be discussed later). It is easy to prove that the orthogonal projection of the cone's apex onto the plane, lies on the ellipse's major axis (*cf.* Fig. 2(a)). This implies that the cone's apex lies in the plane that is orthogonal to the ellipse's supporting plane and that contains its major axis.

For our problem, this means that the optical center must lie in the plane $Y = 0$ (since the ellipse lies in plane $Z = 0$ and has the X-axis as major axis). We may further constrain its position $\mathbf{C} = (X, 0, Z, 1)^\mathsf{T}$, as follows [3]. The cone with $\mathbf{C}$ as apex and that contains the calibration ellipse ε, is given by

$$\Lambda \propto \begin{bmatrix} a\,Z^2 & 0 & -a\,X\,Z & 0 \\ 0 & b\,Z^2 & 0 & 0 \\ -a\,X\,Z & 0 & a\,X^2 - 1 & Z \\ 0 & 0 & Z & -Z^2 \end{bmatrix}.$$

For this cone to be a right one, its upper left 3×3 matrix $\bar{\Lambda}$ must have a double eigenvalue. The three eigenvalues are:

$$b\,Z^2\,,\ \frac{a\,(X^2{+}Z^2) - 1 \pm \sqrt{4a\,Z^2 + (1{-}a\,(X^2{+}Z^2))^2}}{2}. \quad (2)$$

The second and third eigenvalues can not be equal for real values of X and Z (besides in the trivial case $X = Z = 0$ which is excluded since it would correspond to the optical center lying in the calibration plane). The first eigenvalue is equal to the third one if $Z = 0$ (excluded for the same reason as above) and to the second one if:

$$a\,b\,X^2 + b(a-b)\,Z^2 + (a-b) = 0. \quad (3)$$

This equation tells us that the optical center lies on a **viewpoint conic** given by the following matrix and the associated equation

$$\Psi \propto \begin{bmatrix} a\,b & & \\ & b(a-b) & \\ & & a-b \end{bmatrix}, \qquad \begin{pmatrix} X & Z & 1 \end{pmatrix} \Psi \begin{pmatrix} X \\ Z \\ 1 \end{pmatrix} = 0. \quad (4)$$

This is a hyperbola, due to $b \geq a > 0$ (cf. (1)); it is sketched in Fig. 2(a). Furthermore, its asymptotes correspond to the directions of the two circular cylinders that contain the calibration ellipse.

Let us now consider the orientation of the optical axis. Due to (3), we have an optical center given by:

$$Z = \pm\sqrt{\frac{a\,b\,X^2 + a - b}{b(b-a)}}. \quad (5)$$

Here, we exclude the case $a = b$, which would correspond to the camera looking straight at the calibration plane.

The direction of the cone's axis is given by the eigenvector associated with the single eigenvalue of $\bar{\Lambda}$, *i.e.* the third one, augmented with a homogeneous coordinate 0:

$$\begin{pmatrix} \pm\sqrt{b(b-a)(a\,b\,X^2 + a - b)} & 0 & a\,b\,X & 0 \end{pmatrix}^\mathsf{T}. \quad (6)$$

We now show that the cone's axis is identical to the tangent of the hyperbola Ψ in the optical center $\mathbf{C}$, which is given by the line (in the plane $Y = 0$)

$$\Psi \begin{pmatrix} X \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} a\,b\,X \\ \mp\sqrt{b(b-a)(a\,b\,X^2 + a - b)} \\ a - b \end{pmatrix}.$$

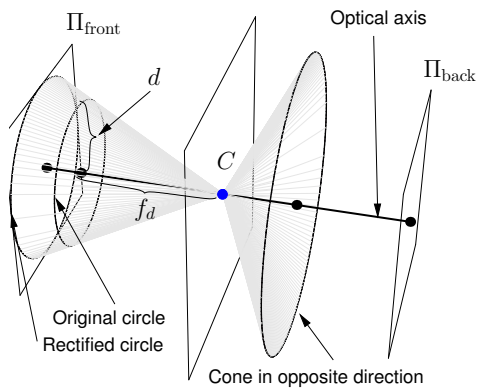


Fig. 3. Viewing cones can also be seen as individual perspective cameras with different focal length. A rectified image can be obtained by projecting the distortion circles (which lie in different planes) on one plane Π_{front} (or Π_{back} for a field of view larger than 180°).

Its point at infinity is (still in the plane $Y = 0$)

$$\begin{pmatrix} \pm\sqrt{b(b-a)(a\,b\,X^2+a-b)} & a\,b\,X & 0 \end{pmatrix}^{\mathsf{T}},$$

i.e. it is identical with the point given in (6). Hence, for an optical center on Ψ, the optical axis is directly given by the associated tangent.

Coming back to (2), the eigenvalues of the cone can be used to compute the focal length f_d associated to a distortion circle of radius d (see Fig. 3). The value f_d/d is the tangent of half the opening angle of the viewing cone. Furthermore, it can be shown that $(f_d/d)^2$ equals the negative of the ratio of the double and the single eigenvalue of $\bar{\Lambda}$, *i.e.* the negative of the ratio of the first and the third eigenvalues given in (2). Using (5), we get

$$\left(\frac{f_d}{d}\right)^2 = -\frac{b\left(abX^2+a-b\right)}{a(a-b)}. \tag{7}$$

This relation will prove useful in Section IV-B.

The case of calibration hyperbolas. The case where the intersection between a cone and the calibration plane yields a hyperbola is accounted for automatically by the previous formulation. All above findings hold here, with the exception that the viewpoint conic Ψ is an ellipse. This case typically occurs with very wide angle cameras or when the angle between the camera's optical axis and the calibration plane is large.

B. Multiple Distortion Circles

So far, we have shown that for an individual distortion circle, the associated viewing cone can be determined from the associated calibration conic, up to 1 degree of freedom (location on the viewpoint conic Ψ and associated orientation of the optical axis). We now show how to get a unique solution, when considering several distortion circles simultaneously. Let us first note that calibration conics corresponding to different distortion circles are not independent: their major axes are collinear (*cf.* Fig. 2(b)), even in the NSVP case. Their centers are not identical however, unless they are all circles, which can only happen when the camera looks straight at the calibration plane.

Let Ψ_d be the viewpoint conic associated with distortion circle (of radius) d, all Ψ_d being given in the same coordinate frame. In the case of a single viewpoint camera, the optical center must lie on all these conics. Furthermore, the optical axis is tangent to all of them. This implies that all conics touch each other (have a double intersection point) in the optical center. This is illustrated in Fig. 2(c). A naïve algorithm would compute the viewpoint conic for all calibration conics and seek their single intersection/contact point. However, very little noise can cause two viewpoint conics to have no real intersection point at all, instead of a double one.

Interestingly, this constraint gives a geometric explanation of the ambiguity, or correlation, between camera position and focal length that often occurs when calibrating a camera from a single view. Consider perfectly recovered viewpoint conics (such as in Fig. 2 (c)), the ambiguity is observed as the area where the viewpoint conics are "very close" to each other. Clearly, a very low perturbation of the curves can result in significant but correlated errors on the optical center and the focal length.

In the NSVP case, a separate optical center corresponds to each distortion circle and viewing cone. Hence, the viewpoint conics will not have a single contact point anymore. However, the optical axis is shared by all viewing cones. Hence, it is the single (in general) line that is tangent to all viewpoint conics. Furthermore, each contact point with the optical axis is the associated optical center.

C. Model parameterization

It is possible to consider the distortion circles and associated viewing cones as individual *perspective cameras*, with different focal lengths but identical principal points [27]–[29]. In the SVP case, extrinsic parameters of all these cameras are identical, whereas in the NSVP case, they all share the same orientation and the optical centers are merely displaced along the optical axis.

Let us consider the distortion circle of radius d and one image of a calibration plane. From point correspondences between pixels on this circle and points on the calibration plane (on the calibration conic), we can compute a plane-to-image homography H_d. For simplicity, let us assume that image coordinates have been translated to put the distortion center at the origin. The homography can then be decomposed such that

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathsf{H}_d \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathsf{K}_d\,\mathsf{R} \begin{bmatrix} 1 & 0 & \\ 0 & 1 & -\mathbf{t}+t_d\mathbf{r}_3 \\ 0 & 0 & \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{8}$$

where $\propto$ means equality up to scale, (x, y) is a calibration point, (u, v) a pixel on the distortion circle, and R and $\mathbf{t}$, a rotation matrix and translation vector representing camera pose (same for all d). The scalar t_d allows to model translational displacement of individual viewing cones along the optical axis (given by $\mathbf{r}_3^{\mathsf{T}}$, the third row of R), which is needed for NSVP cameras. For SVP cameras, t_d is set to 0 for all d.

As for K_d, it is a calibration matrix[1] $\mathrm{diag}(f_d, f_d, 1)$, where f_d is the focal length associated with the considered distortion circle. We may interpret the relation between d and f_d as a distortion function applied to a perspective camera whose undistorted image plane is π_{front} (*cf.* Fig. 3).

Note that this parameterization only accounts for viewing cones with field of view smaller than 180°. Larger fields of view can be modeled by adding a reflection to the rotational component of the pose, $\mathsf{R}' = \mathrm{diag}(1, 1, -1)\mathsf{R}$, and a corresponding image plane π_{back}. Equivalently, one may describe these cones as cameras with negative focal length as presented in [28].

D. Overview of the approaches

We briefly describe our two calibration approaches in terms of the concepts previously introduced in this section. Both proceed in two steps. In the homography based algorithm (next section), the optical axis for each view of the calibration plane is first estimated, each one up to four possibilities. The correct choice is made at the second step, which consists of simultaneously estimating the position of the camera on the optical axis for all views and the opening angle (equivalently the focal length) of all the cones. The second approach is based on above introduced geometric constraints, although ultimately some algebraic manipulations are required to avoid an iterative solution. We discuss two solutions for estimating the pose of the camera for each view, using the calibration conics. Then, the opening angles of the cones are estimated by re-fitting the calibration conics using a parameterization enforcing the fit to be consistent with the pose of the camera.

III. Homography-based Calibration

A. Unconstrained approach

In [27], an approach based on recovering the *Image of the Absolute Conic*, associated to each distortion circle, from one or many images of a calibration plane, is introduced. Once the internal parameters are recovered, the camera external parameters are estimated using the approach presented in [22], [24], [34]. The weakness of this approach is that since the calibration is individually performed on each distortion circle, it does not enforce SVP or NSVP constraints. However, it is useful for estimating the distortion center as described in Section V. This algorithm will be referred to "`IAC`".

B. Constrained approach

We present another homography-based approach that can enforce the constraints directly. It is related to Hartley-Kang's approach, but it also handles non-central omnidirectional cameras.

[1]As mentioned in the introduction, this model does not include a skew between pixel axes or an aspect ratio different from 1. Also, it assumes that the distortion center is at the principal point.

SVP case, one image. In the SVP case, t_d is zero for all d. It follows that

$$\mathsf{K}_d^{-1}\mathbf{q} \propto \begin{bmatrix} 1 & & \\ & 1 & \\ & & f_d \end{bmatrix} \mathbf{q} \propto \mathsf{R} \underbrace{\begin{bmatrix} 1 & 0 & \\ 0 & 1 & -\mathbf{t} \\ 0 & 0 & \end{bmatrix}}_{\mathsf{M}} \mathbf{Q} \tag{9}$$

holds up to scale. If we divide the first by the second coordinates of both sides, we obtain an equation which is independent from the distortion circle d

$$\frac{q_1}{q_2} = \frac{(\mathsf{M}\mathbf{Q})_1}{(\mathsf{M}\mathbf{Q})_2}. \tag{10}$$

We finally obtain a linear equation on the six coefficients in the first two rows of the pose matrix M

$$\begin{aligned} q_1\left(M_{21}Q_1 + M_{22}Q_2 + M_{23}Q_3\right) - \\ q_2\left(M_{11}Q_1 + M_{12}Q_2 + M_{13}Q_3\right) \quad = \quad 0. \end{aligned} \tag{11}$$

The above equation being linear homogeneous, we may use all point correspondences between the calibration plane and the image to estimate the first two rows of M, up to scale. Until here, our algorithm is very similar to [14], but the following differs significantly. The pose, *i.e.* R and $\mathbf{t}$, can be partially estimated from M. There exists a scalar λ such that

$$\begin{bmatrix} R_{11} & R_{12} & -\mathbf{r}_1^\mathsf{T}\mathbf{t} \\ R_{21} & R_{22} & -\mathbf{r}_2^\mathsf{T}\mathbf{t} \end{bmatrix} = \lambda\bar{\mathsf{M}}, \tag{12}$$

where $\mathbf{r}_i^\mathsf{T}$ is the ith row of R and $\bar{\mathsf{M}}$ the upper 2×3 part of M. This leads to the following observations:

- the rotation matrix R can be estimated up to 4 solutions, from the left 2×2 part of $\bar{\mathsf{M}}$. They differ by a choice of sign for columns and rows of the solution: if R is one correct solution, then the other three are given by DR, RD, DRD, with $\mathsf{D} = \mathrm{diag}(-1, -1, 1)$. Distinguishing the valid rotation matrix will be done later when estimating the full camera position and internal parameters;
- the intersection $\mathbf{t}_0$ between the optical axis of the camera and the calibration plane can be recovered as the right nullspace of $\bar{\mathsf{M}}$, since it is the point on the plane projecting onto the distortion center (the origin). Hence, the translation vector $\mathbf{t}$ can be estimated up to one degree of freedom, a displacement μ along the optical axis

$$\mathbf{t} = \mathbf{t}_0 + \mu\mathbf{r}_3. \tag{13}$$

Estimating μ can be done simultaneously with estimating the focal lengths of all distortion circles. By inserting the solution for $\mathbf{t}$ into (9), we obtain

$$\mathbf{q} \propto \begin{bmatrix} f_d & & \\ & f_d & \\ & & 1 \end{bmatrix} \mathsf{R} \begin{bmatrix} 1 & 0 & \\ 0 & 1 & -\mathbf{t}_0 - \mu\mathbf{r}_3 \\ 0 & 0 & \end{bmatrix} \mathbf{Q}.$$

Due to the orthonormality of R, we get

$$\mathbf{q} \propto \begin{bmatrix} f_d & & \\ & f_d & \\ & & 1 \end{bmatrix} \left(\mathsf{R} \underbrace{\begin{bmatrix} 1 & 0 & \\ 0 & 1 & -\mathbf{t}_0 \\ 0 & 0 & \end{bmatrix}}_{\mathsf{T}_0} - \begin{bmatrix} 0 & & \\ & 0 & \\ & & \mu \end{bmatrix} \right) \mathbf{Q}$$

Let us denote: $\mathbf{S} = \mathsf{R}\mathsf{T}_0\mathbf{Q}$ which is a known vector. We thus can write the above equation as

$$\mathbf{q} \propto \begin{bmatrix} f_d S_1 \\ f_d S_2 \\ S_3 - \mu Q_3 \end{bmatrix} = \begin{bmatrix} S_1 & & \\ & S_2 & \\ & & Q_3 \end{bmatrix} \begin{pmatrix} f_d \\ f_d \\ S_3/Q_3 - \mu \end{pmatrix} \quad (14a)$$

and obtain the following set of linear equations on the unknowns f_d and μ

$$[\mathbf{q}]_\times \begin{bmatrix} S_1 & & \\ & S_2 & \\ & & -Q_3 \end{bmatrix} \begin{pmatrix} f_d \\ f_d \\ \mu - S_3/Q_3 \end{pmatrix} = \mathbf{0} \quad (14b)$$

or

$$\begin{bmatrix} q_3 S_2 & q_2 Q_3 \\ q_3 S_1 & q_1 Q_3 \\ q_1 S_2 - q_2 S_1 & 0 \end{bmatrix} \begin{pmatrix} f_d \\ \mu \end{pmatrix} = \begin{pmatrix} q_2 S_3 \\ q_1 S_3 \\ 0 \end{pmatrix}. \quad (14c)$$

Note that the equation system is non-homogeneous, *i.e.* the solution for the f_d and μ is computed exactly, not only up to scale. Also note that the third equation is useless: in the absence of noise, the term $q_1 S_2 - q_2 S_1$ is zero and gives no constraint on f_d. With noise however, the third equation admits random coefficients and using it was found to bias f_d towards small values. We thus only use the first two equations of (14c).

All point correspondences, from all distortion circles, can be used simultaneously: each correspondence contributes to estimating μ and to the focal length of the distortion circle it belongs to. Hence, overall we have a linear system of size $2n \times (D+1)$, where n is the number of point correspondences and D the number of distortion circles.

This assumed a known rotation. All rotation matrices among the four possibilities shown above, give the same solution for f_d and μ, up to different signs. Let d_0 be distortion circle of smallest radius. Then the correction rotation is the one with positive f_{d_0} and μ.

Using many images of a calibration plane. The above equations can be used in a slightly modified form to simultaneously use many images. Let v be the index for each view out of V. Each associated 2×3 partial homography $\bar{\mathsf{M}}_v$ can be computed using (11). Then $\mathbf{t}_{v0}$ and R_v can be estimated individually from each of them. Finally, (14c) is extended naturally to account for many views by simultaneously solving for all displacements μ_v and focal lengths f_d. The resulting equation system is of size $2n \times (D + V)$.

NSVP case, many views. In the NSVP case, the previous algorithm can be applied nearly without modification. Let us consider the form of the plane-to-image homography for an individual distortion circle, given in equation (8), for the NSVP case. In comparison to the SVP case (equation (9)), there is an additional term t_d for the position of the displaced optical center on the optical axis:

$$\begin{aligned} \mathbf{q} &\propto \mathsf{H}_{vd}\mathbf{Q} = \mathsf{K}_d \mathsf{R}_v \begin{bmatrix} 1 & 0 & \\ 0 & 1 & -(\mathbf{t}_{v0} + (\mu_v + t_d)\mathbf{r}_{v3}) \\ 0 & 0 & \end{bmatrix} \mathbf{Q} \\ &= \mathsf{K}_d \left(\mathsf{R}_v \mathsf{T}_{v0} - \mathrm{diag}(0, 0, \mu_v + t_d)\right) \mathbf{Q}. \end{aligned} \quad (15)$$

Note that t_d is the same for all views. Since K_d is a diagonal matrix, the first two coordinates of the right hand side do not depend on μ_v and t_d. Hence, like in the SVP case, (11) can be used to estimate linearly the first two rows of the pose matrix M_v, using all point correspondences, for all distortion circles. The pose parameters R_v and $\mathbf{t}_v$ can also be extracted in the same manner, with $\mathbf{t}_v$ being obtained up to a displacement along the optical axis $\mathbf{r}_{v3}$.

INPUT: Dense or Sparse plane–image correspondences

ALGORITHM:

For each calibration plane v:

- Estimate $\bar{\mathsf{M}}_v$ using (11) and all the correspondences.
- Obtain the rotation R_v up to 4 solutions using (12).
- Obtain $\mathbf{t}_{0v}$ as the right nullspace of $\bar{\mathsf{M}}_v$, The four possible optical axes are given by (13).
- Disambiguate R_v and optical axis by estimating $\forall d$: f_d and μ_v using (14c)

DENSE CORRESPONDENCES:

- Simultaneously estimate $\forall d : f_d$ and $\forall v : \mu_v$ combining (14c) for each view.

SPARSE CORRESPONDENCES:

- Apply the previous step replacing f_d by a polynomial.

Perform iterative refinement of the reprojection error.

Fig. 4. Complete algorithm for the HB calibration approach for an SVP camera.

We now consider how to estimate the focal lengths f_d and the optical center positions $\mu_v + t_d$. The equations are identical to those in the SVP case, with the difference that μ_v has to be replaced by $\mu_v + t_d$, *i.e.* is not the same for all distortion circles. The set of μ_v and t_d is an overparameterization, since subtracting a value from all μ_v and adding it to all t_d leaves (15) unchanged. Hence, we may set one of them to any fixed value. The equation system is thus of size $2n \times (2D - 1 + V)$.

Polynomial model. Calibrating a general model requires many data points to obtain precise and accurate results, i.e. to avoid over-fitting. Otherwise, relying on a more restricted model may give better results. In [28], [29], [31], a polynomial model was used to represent the relationship between focal length and radius d. Our formulation can also be trivially modified to use polynomials for f_d as well as t_d. For NSVP cameras, it turns out that in practice, estimating the fully general model can be unstable because of the focal length–displacement ambiguity mentioned in Section II-B. The result is that both f_d and t_d are not smooth functions for small radius values d. Instead of relying on smoothing terms (whose weights are difficult to set) to circumvents this, we prefer using polynomials.

The above method is summarized in Fig. 4 and referred to as "HB" in the following. In practice, this is the most accurate and flexible of the approaches presented here (see Section VII).

IV. Geometric approach

A. Solving the Calibration and the Pose

In the following, we present a second approach based on the geometric constraints discussed in Section II. We first propose a naïve calibration approach; it is not optimal though and a better one will be discussed in the next section. Recall the

constraints relating calibration conics, viewpoint conics and position of the camera illustrated in Fig. 2. In the SVP case, one easily deduces a calibration algorithm consisting in first estimating the optical center (relative to the calibration plane) as the 3D point which is closest on average to all viewpoint conics (see below). Then, (7) can be used to compute the focal lengths for all distortion circles. In the NSVP case, a possibility would be to compute the optical axis: the line $\mathbf{L}$ that minimizes the sum of squared distances to the viewpoint conics. However, this is not very accurate as discussed below.

SVP case: Computing the closest point to the viewpoint conics. This is the original approach presented in [27]. Computing the orthogonal distance of a point to a general conic requires solving a fourth degree polynomial [33]. Using this to compute the closest point to our set of viewpoint conics is not very practical. Instead, we iteratively minimize a cost function subject to constraints. The closest point $\mathbf{q}$ is found by solving

$$\min_{\mathbf{q},\mathbf{q}_d} \sum_d dist(\mathbf{q},\mathbf{q}_d)^2, \text{ subject to } \mathbf{q}_d^\mathsf{T}\,\Psi_d\,\mathbf{q}_d = 0,$$

i.e. we also estimate one point per conic Ψ_d that will, after convergence, be the orthogonal projection of $\mathbf{q}$ on Ψ_d. Since the function and constraints are polynomial, this problem can be optimized using an algorithm relying on Cylindrical Algebraic Decomposition (CAD) which guarantees a global minimum [21]. Such an algorithm is available through the *Minimize* function of *Mathematica*. In this work, we present another solution that gives much better result. Instead of proceeding into two steps, we directly use the fitted calibration conics to estimate the camera position, as described below.

B. A formulation enforcing the constraints directly

In our first experiments, we found the previous approach to be unstable, even though good results could be obtained in some cases [27]. This is because the formulation has several drawbacks. First, although CAD optimization is algorithmic, it becomes computationally intractable when the number of viewpoint conics increases. Secondly, it appears that finding the closest point to the set of the recovered viewpoint conics is not the optimal criterion. As discussed above and shown in Fig. 5, when the noise is large, the shape and position of the viewpoint conics may become very perturbed.

For an SVP camera, a better formulation would directly enforce that the viewpoint conics all touch in one location and have the same tangent at this point. Then, CAD optimization could be avoided. Before giving our solution, we come back to our calibration conics. As mentioned above, they share an axis (in the absence of noise). We assume that their common axis can be estimated with high accuracy despite noise in the data, due to using many calibration conics. Once the conics' axis is estimated, it is convenient to change the coordinate system of the calibration plane such that the conics are aligned with the X-axis. Then, each one of them can be parameterized by its major and minor axis lengths b_d and a_d, as well as by a displacement k_d along the X-axis:

$$\varepsilon_d \propto \begin{bmatrix} a_d & 0 & -a_d\,k_d \\ 0 & b_d & 0 \\ -a_d k_d & 0 & a_d\,k_d^2 - 1 \end{bmatrix}. \tag{16}$$

The parameters a_d, b_d and k_d can be estimated much like in a classical conic fitting algorithm. We perform all subsequent computations with those axis-aligned conics. This parameterization guarantees that all viewpoint conics lie in the same plane ($Y = 0$), in which they can be expressed as

$$\Psi_d \propto \begin{bmatrix} a_d\,b_d & 0 & -a_d\,b_d\,k_d \\ 0 & b_d\,(a_d - b_d) & 0 \\ -a_d\,b_d k_d & 0 & a_d\,b_d\,k_d^2 + a_d - b_d \end{bmatrix}. \tag{17}$$

However, this parameterization does not guarantee other properties given in Section II-B, especially that these viewpoint conics all touch in a single point. To achieve this, we state a new result. For a central radially symmetric camera, the intersections of all its viewing cones with the calibration plane are calibration conics $\hat{\varepsilon}_d$ given by

$$\hat{\varepsilon}_d \propto \begin{bmatrix} \gamma^2\rho_d & 0 & \epsilon\rho_d\gamma + \gamma \\ 0 & \delta^2\,(2\rho_d + 1) & 0 \\ \epsilon\rho_d\gamma + \gamma & 0 & \rho_d\epsilon^2 + 2\epsilon - 2 \end{bmatrix}, \tag{18}$$

where γ, ϵ and δ encode the external parameters of the camera, and ρ_d is a parameter for the distortion circle of radius d. For each calibration conic $\hat{\varepsilon}_d$, the corresponding viewpoint conic is given by

$$\hat{\Psi}_d \propto \begin{bmatrix} \gamma & 0 & \epsilon + \frac{1}{\rho_d} \\ 0 & \frac{\gamma^2 - 2\delta^2 - \frac{\delta^2}{\rho_d}}{\gamma} & 0 \\ \epsilon + \frac{1}{\rho_d} & 0 & \frac{\gamma^2 + 2\delta^2(\epsilon - 1) + \delta^2\epsilon^2\rho_d}{\gamma\delta^2\rho_d} \end{bmatrix}. \tag{19}$$

The derivation and the proof of this formulation are available in [30]. Note that in practice, we are only interested in the estimation of $\hat{\varepsilon}_d$. Much like (7) can be obtained for Ψ_d, we can compute $(f_d/d)^2$ for $\hat{\Psi}_d$. After some algebraic manipulation using that result (not shown here) we substitute ρ_d in (18) with

$$\rho_d = -\frac{\phi_d\gamma^2}{2\,(\gamma^2 - 2\delta^2)} - \frac{1}{2}, \tag{20}$$

where ϕ_d is chosen so to be equal to $(f_d/d)^2$.

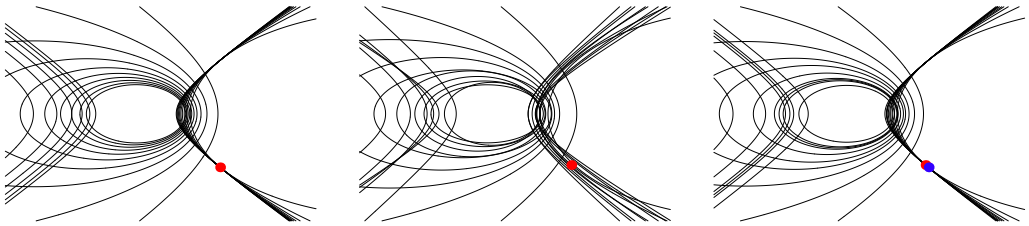
The formulas for estimating the external parameters of the camera are summarized in Table II and briefly derived below. The position of the camera is obtained by solving

$$(X, Z, 1)\,\Psi_i\,(X, Z, 1)^\top = (X, Z, 1)\,\Psi_j\,(X, Z, 1)^\top$$

for $(i \neq j)$. The value $-\frac{\epsilon}{\gamma}$ gives the X-coordinate of the intersection point of the optical axis with the calibration plane. Indeed, we can verify that

$$\forall d : (-\epsilon/\gamma, 0, 1)\,\Psi_d\,(X, Z, 1)^\top = 0,$$

where $\Psi_d\,(X, Z, 1)^\top$ is the tangent to the viewpoint conic at the optical center, *i.e.* the optical axis. One may wonder why the extrinsics were not directly "encoded" in the parameterization. Indeed, such parameterizations exist and were investigated. However, they were abandoned because the resulting calibration conics $\hat{\varepsilon}_d$ were not as "simple" as the one



(a) (b) (c)

Fig. 5. The effect on viewpoint conics of errors on the calibration conics (shown in the same plane for visualization). **(a)** Original configuration with the optical center. **(b)** Added noise to the calibration conics and resulting viewpoint conics. Obviously, finding the closest point to these curves will yield a large error. **(c)** Recovered optical center in dark/blue and corrected conics given by our approach, *cf.* Section IV-B.

TABLE II

COMPUTING THE EXTERNAL PARAMETERS OF THE CAMERA FROM THE PARAMETERS

Information	Formula
X-coordinate of the optical center	$-\frac{\gamma}{\delta^2} - \frac{\epsilon}{\gamma} + \frac{2}{\gamma}$
Z-coordinate of the optical center	$-\sqrt{-\frac{\gamma^2-2\delta^2}{\delta^4}}$
Intersection of optical axis and calibration plane	$-\epsilon/\gamma$
Intersection of calibration plane and principal plane[1] (setting $\phi_d = 0$)	$y = (2-\epsilon)/\gamma$

[1] The principal plane refers to the plane passing through the optical center and being parallel to the image plane.

we propose. When a calibration algorithm could be deduced from any of these formulations, it was less stable than the one we show next.

With this global formulation, all calibration conics $\hat{\varepsilon}_d$ should be fitted a the same time. However, this task is difficult since the function is non-linear and requires an initial estimate of the parameters.

Perhaps surprisingly, there exists an analytic solution to computing the extrinsics parameters once the values of a_d, b_d and k_d for all calibration conics are estimated. The position of the camera can be estimated, while ignoring the focal length at each circle. We exploit the fact that without noise $\forall d : \varepsilon_d = \hat{\varepsilon}_d$, *i.e.* that the unconstrained and constrained formulation should give identical results[2]. We solve this equation for each parameter ϵ, γ and ϕ_d and obtain

$$\phi_d = 2\delta^{-1/2} a_d^{-1} b_d \left(b_d - a_d\right) - 1 \tag{21}$$

$$\tilde{\epsilon}_d = -a_d^{-1/2}\sqrt{\gamma^2 + a_d} - \gamma k_d + 1 \tag{22}$$

$$\tilde{\gamma}_d = \pm\delta b_d^{-1}\sqrt{a_d}\sqrt{\delta^2 + 2b_d}, \tag{23}$$

for each calibration conic. Note that the two last equations involve only the extrinsics parameters. We rename them with a $\tilde{}$ and a d subscript because, in the presence of noise, they are different for each calibration conic. This is because each calibration conic is fitted individually. Therefore, an estimate for δ can be obtained by minimizing the variance of the $\tilde{\gamma}_d$. In our implementation, a minimization over $\tilde{\gamma}_d^2$ is prefered, since it eliminates the square root in the expression. This results in solving

$$\delta^2 = \arg\min_{\delta} \sigma^2, \quad \text{and} \quad \sigma^2 = \sum_d (\mu - \tilde{\gamma}_d^2)^2 \tag{24}$$

[2]We scale ε_d and $\hat{\varepsilon}_d$ such that they have unity at the upper-left coordinate.

where μ is the mean. The expression σ^2 is a fourth degree polynomial in δ^2, *i.e.* it admits only three extrema in δ^2. One of them corresponds to $\delta = 0$ and is a maximum, and the two others are minima with identical absolute value but opposite sign. The one larger than zero is our solution for δ^2 and δ can always be taken positive. Given this value, we can estimate the other parameters with $\epsilon = \text{Mean}(\tilde{\epsilon}_d)$ and $\gamma = \pm\text{Mean}(\tilde{\gamma}_d)$ using (22) and (23). This sign ambiguity will be resolved later. In the presence of noise, we perform a final optimization:

$$\arg\min_{\delta,\epsilon,\gamma} \sum_d (\epsilon - \tilde{\epsilon}_d)^2 + (\gamma - \tilde{\gamma}_d)^2 \tag{25}$$

using a standard Gauss-Newton algorithm.

Once the external parameters of the camera are known, that is, the value of δ, ϵ and γ, we are seeking the intrinsics ϕ_d. Using (21) did not give satisfying results in practice. Instead, we find $\hat{\varepsilon}_d$ that best fits the original data points. Given the external parameters, each calibration conic can be fitted using a least square algebraic error function resulting in a second degree polynomial involving the variable ϕ_d. We try both signs for γ and keep the one that best fits the calibration conics.

Many calibration planes. If many calibration planes are available, the focal length of each circle must be consistent over all the views. This is accomplished by simultaneously fitting the calibration conics $\hat{\epsilon}_d$ for different planes. This global Linear Least Square problem is a second degree polynomial where ϕ_d is the only variable.

The steps of this algorithm are summarized in Fig. 6. In the following, it is referred to as the Right Cone Constraint method, "RCC".

C. NSVP extension

Our previous SVP model can be generalized allowing an NSVP. Due to lack of space, it is not described here. Our tests demonstrated that such a parameterization is not that useful in practice. Indeed, the parameters can only be recovered by means of optimization that would require initial estimates of the parameters, which are difficult to obtain.

V. COMPUTING THE DISTORTION CENTER

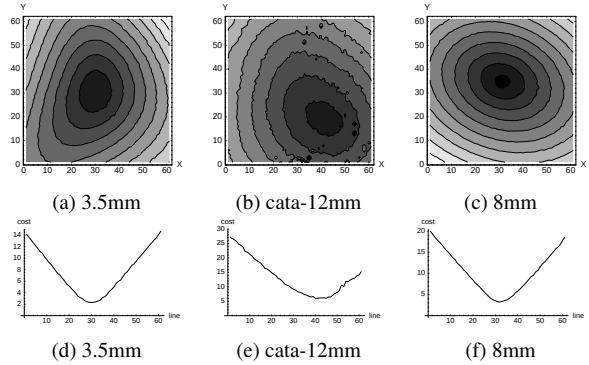
Until now, we have assumed, for both algorithms, that the distortion center was known; this information was used to select the distortion circles. Recall that the distortion center is also the principal point of the camera in our model. Tests with

INPUT: Dense plane–image correspondences
ALGORITHM:
For each calibration plane v:

- Estimate individually every calibration conic ε_d.
- Compute the average major axis $\mathbf{m}$ and select a point, denoted $\mathbf{g}$, on the axis.
- Change the coordinate system of the original data to set $\mathbf{m}$ on the X-axis and $\mathbf{g}$ at the origin.
- Estimate the new ε_d of the form: $\propto \begin{bmatrix} a & 0 & c \\ 0 & b & 0 \\ c & 0 & d \end{bmatrix}$.
- Solve (24) to obtain δ_v.
- Minimize (25) and obtain ϵ_v, γ_v.

Compute, for all distortion circles d, ϕ_d using (18).
Perform iterative refinement of the reprojection error.

Fig. 6. Complete algorithm for the RCC calibration approach.



(a) 3.5mm (b) cata-12mm (c) 8mm

(d) 3.5mm (e) cata-12mm (f) 8mm

Fig. 7. Plots of the goodness measure for the distortion center, obtained for three tested lenses (*cf.* Section VII). **(a,c,e)** 60×60 grid around the image center (yellow/dark meaning smaller). **(b,d,f)** One slice per plot, through the respective minimum.

noiseless simulated data showed that the calibration may be quite sensitive to a bad choice of distortion center; indeed, like for real cameras, using the image center as an approximation was not satisfying in general. Hence, the distortion center must be estimated as part of the calibration process. Note that using Hartley-Kang's algorithm is not satisfactory in general since the criteria for choosing this point is in terms of image rectification. Besides not being applicable to non-central cameras, these criteria do not correspond to our model. Below, we propose one where the recovered distortion center is identical to the principal point. The sensitivity of calibration we observed in simulations suggests that it should be possible to estimate the distortion center rather reliably, which was confirmed in practice.

Algorithm. We used the following heuristics to define an optimization criterion for the distortion center. Let us apply the IAC approach of Section III-A with several images as input. The plane-based calibration for each distortion circle is then capable of estimating a principal point, besides the focal length f_d. It seems plausible that the better the assumed distortion center was, the closer the estimated principal points will be to it. Since plane-based calibration is applied on images centered on the assumed distortion center, we can consider the average distance of the estimated principal points (one per distortion circle) as a measure for the goodness of the center.

Figure 7 shows the values of this measure, computed for distortion center positions on a 60×60 grid around the image center, for real cameras. The shape of the cost surface indicates that we can find the optimum distortion center using a simple steepest descent type method. We implemented such an approach that accurately finds the distortion center within a couple of minutes of computation. Note that the second column of Fig. 7 shows that, although the principal points used to plot it were computed individually per distortion circle, they are very densely clustered (average distance to assumed distortion center of less than 3 pixels). This suggests a high stability of the calibration.

Discussion. Compared to other approaches, our optimization criterion is chosen to find the best optical axis. Thus, the distortion center is identical to a principal point in a pinhole camera. This formulation is also used in [9], [31], [32]. Our criterion is different from the one in image-based distortion functions, where the distortion center (together with the distortion function) is chosen to maximize the linearity of rectified line images [6], [14], [28], [34]. However, the latter can result in instability for the estimation under very low distortion and, more importantly, is not optimal when the camera is NSVP since image rectification is not possible. In theory, ours is not subject to these problems even when no distortion is present. This is confirmed by the nice shape of the error function for the 8.0mm camera as show in Fig. 7. For these reasons, the distortion center computation will not be included in the comparison with Hartley-Kang's approach.

VI. Practical Issues

A. Dense Plane–Image Correspondences

The easiest approach we found to get dense correspondences between the calibration plane and the camera is to use a flat screen. We used a simple coded structured light algorithm [19], which consists in displaying a sequence of patterns of horizontal and vertical black and white stripes of varying thickness on the screen to encode the position of each screen pixel (*cf.* Fig. 8). Then, for each camera pixel, we identify the corresponding position on the calibration plane by decoding the observed intensities in each pattern. We found that, when performed in a controlled environment (low-constant ambient lighting, screen of high contrast and resolution), the accuracy of such a method is good enough for calibration. Indeed, we only used correspondences of pixel precision (see [26] for details). Since the points located on the distortion circles are given in floating point coordinates, we compute their correspondences by a weighted sum of the correspondences recovered for the four closest image pixels. Besides allowing dense correspondences with the calibration plane, these approaches render trivial the problem of recovering the structure of the calibration plane. This is as opposed to using calibration grid images where grid points must be automatically extracted and identified. This is especially difficult when the distortion is very large.

(a) (b) (c)

Fig. 8. Projected patterns for correspondences are horizontal and vertical black and white stripes. Images taken with **(a)** the Goyo 3.5mm, **(b)** catadioptric, and **(c)** paracatadioptric camera (cata-12mm).

B. Omnidirectional Cameras

There are several issues worth mentioning for omnidirectional cameras. If the field of view is larger than $180°$, then some distortion circles will have viewing cones that actually approach planes. For the RCC approach, this means that fitting the calibration conic may become unstable. These cases can be detected as the ones whose correspondences on the calibration plane are close to collinear. In practice, they are discarded from the actual calibration procedure. In the case of the homography-based algorithm that uses all matches simultaneously, no special attention is needed.

C. Non-linear optimization

Because f_d is a function of the radius in the distorted image, it is not straightforward to perform the projection of a 3D point into the image (as opposed to the backprojection of image points to 3D). This means that (8) can not be used directly to perform a non-linear optimization of the calibration unknowns. Since the focal length function can not be inverted when it crosses zero, it is preferable to define the distortion in terms of view angle θ_d. As seen in Fig. 14(b), this function is generally simple, so easily invertible. Given this function, the radius of the image of a 3D point is computed from the angle between the optical axis and the line spanned by the optical center and the 3D point. The non-linear optimization is then performed with θ_d instead of f_d. The NSVP case can not be handled as simply, since there is no single optical center relative to which to compute the angle θ. In this case, we use (8) and estimate a radius d_i associated to each 3D–2D correspondence, along with the other parameters. This yields a sparse non-linear problem. We also enforce monotonicity on f_d and θ_d. For example, this can be done approximatively by adding terms to the cost function, of the form: $\left(|f_d - f_{d+s}| - (f_d - f_{d+s})\right)^2$, $s > 0$, which is a quadratic penalty if the constraints are not enforced, but gives 0 otherwise.

VII. Experiments

We tested our approaches using different types of cameras with simulated and real data. They were also compared to Hartley-Kang's algorithm [14] (referred to as "HK"). Since the view angle of some of our camera is larger than $180°$, we implement the algorithm for a spherical retina.

A. Simulation

SVP cameras. We simulated two types of cameras: wide-angle with small radial distortion and fisheyes with large distortion[3]. The focal length (and distortion) function of each type of camera was built randomly via monotonically decreasing polynomials of fifth degree. We used an image size comparable to our own camera: 1 Mega pixel (see real data). The shapes of focal length functions of the wide-angle cameras were similar to our 15mm (Fig. 14) and we used f_d with $f_0 = 1000 \pm 200$ pixels. The simulated fisheyes were analogous to our 3.5mm, so we used $f_0 = 400 \pm 100$ pixels and the shape of the focal length function yielded a field of view close to $180°$. In these tests, we assumed a known distortion center and also made sure the camera was never placed in a near fronto-parallel position w.r.t. the calibration plane.

We compared the reprojection error, the error on the pose and on the calibration. We define the latter as the average difference between the recovered focal length function f_d and the ground truth. Our tests showed that all of them are highly related, so we only show results for the reprojection error. The reprojection errors for the three algorithms w.r.t. noise and the number of used cameras are shown in Fig. 9. We added Gaussian noise of standard deviation up to 4 pixels to the original data, which consisted of 50 points per distortion circle (which is rather small compared to the several hundred usually available from a structured light dense mapping). In all our tests, data on the distortion circles are not necessarily evenly distributed. Especially in the case of very large field of view, only a portion of the image effectively sees the calibration plane. We made sure to properly simulate this effect. Finally, our tests were performed using 3, 7 and 10 views.

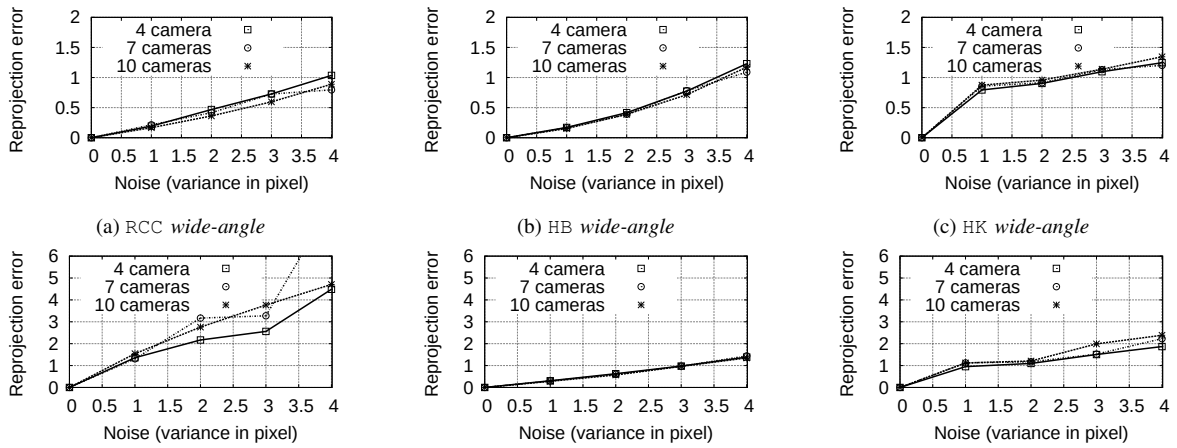
In general, all three algorithms performed similarly. The RCC obtains results for the wide-angle camera similar to the other two. However, this is not the case for the fisheye camera because the data points were not uniformly distributed around the distortion circles. Hence, the pose estimation was not as stable. This exposes a weakness of this approach: if the pose for one view is badly estimated, it can potentially destroy the estimation of the focal length function even using many views. This effect is not as important for the HB since the focal length function is estimated with the position of the camera on the optical axis using all the views.

NSVP cameras. We performed an in-depth analysis of the performance of our algorithms for NSVP cameras. Three aspects were considered. First, how well the displacement along the optical axis can be recovered under noise. Second, is the linear NSVP algorithm useful in practice? Third, does the NSVP model overfit when the camera is actually SVP? We performed our tests using the homography based approach because it is the only one that naturally enforces both SVP and NSVP constraints. Two approaches were tested:

1) Initialization using linear calibration based on an SVP assumption, followed by non-linear optimization of the NSVP model.
2) Initialization as well as optimization using the NSVP model.

The tests were performed on simulated SVP and NSVP catadioptric cameras with viewpoints moving along the optical axis

[3]We did not compare catadioptric cameras with the HK approach because the comparison would have been unfair.



(a) RCC *wide-angle* (b) HB *wide-angle* (c) HK *wide-angle*

(d) RCC *fisheye* (e) HB *fisheye* (f) HK *fisheye*

Fig. 9. Reprojection errors (in pixel) for the simulated *wide-angle* and *fisheye* cameras.

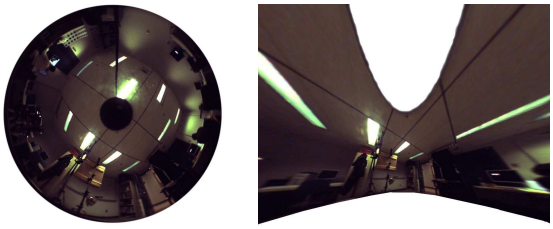
(Fig. 10(a,d)). We used 10 different positions for the plane and added Gaussian noise of standard deviation 1 pixel to the data. Beyond this noise level, we found that our approach could not accurately estimate the viewpoint displacement. Three hundred points per image were used and polynomial models of degree 5 for the calibration were used. A typical behavior of the two approaches is shown in Fig. 10. This leads to the following observations:

- For an SVP camera, both optimizations should lead to similar results (negligible NSVP);
- For an NSVP camera, enforcing an SVP yields a biased focal length function (*cf.* Fig. 10(f)). In some cases, this can be satisfying in terms of reprojection error, like for one of our real cameras (*cf.* Fig. 13);
- When the model is refined to include t_d, the optimization might not converge to a satisfying minimum;
- If the camera is NSVP, the second approach should perform better, but only if the noise is sufficiently low (*cf.* Fig. 10(d)). Otherwise it can result in worse results because solving with (15) is not as stable as with (9).

TABLE III

COMPARISON OF THE AVERAGE REPROJECTION ERRORS. NL REFERS TO NON-LINEAR OPTIMIZATION.

	Algorithms					
Cameras	RCC	RCC+NL	HB	HB+NL	HK	HK+NL
cata-12mm	4.60	1.21	1.28	1.14	1.02	0.96
cata-30mm	—	—	1.42	1.33	—	—
3.5mm	6.89	1.15	1.46	1.06	2.53	1.16
8mm	10.74	3.01	6.64	3.08	10.92	3.10
Canon 15mm	0.56	0.48	0.51	0.48	0.51	0.48



(a) (b)

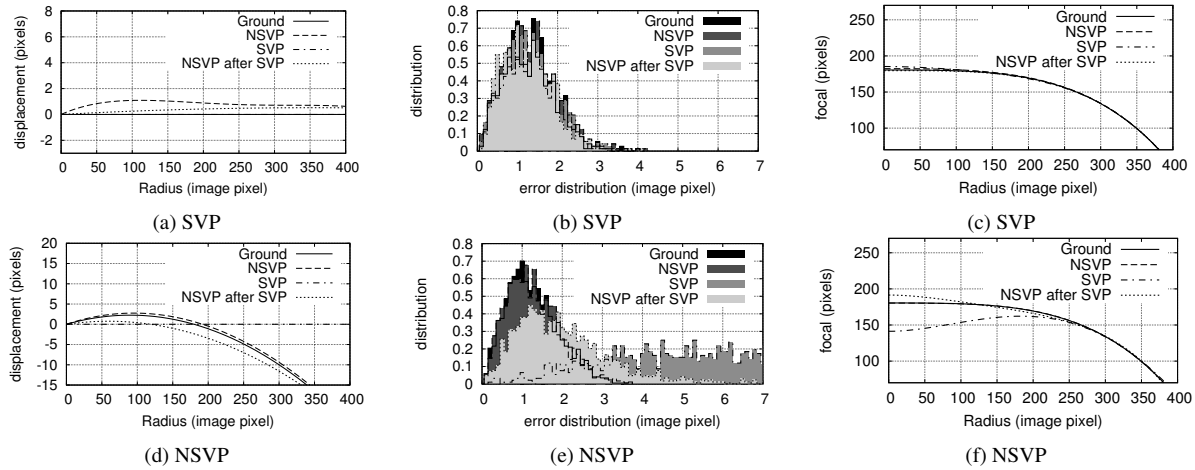
Fig. 11. Image rectification for the Basler camera with the 12.5mm lens and the RemoteReality lens (cata-12mm). **(a)** Original image. **(b)** Rectified image for a rotated view.

B. Real data

Several camera configurations were tested. First, a CCTV Basler A201bc was combined with a fisheye Goyo 3.5mm lens, to an 8mm Cosmicar lens with small distortion and to a RemoteReality catadioptric lens combined with a 12.5mm Cosmicar lens (referred to as "cata-12mm"). Secondly, a Canon SLR was combined with a fisheye 15mm lens and to a 0-360 catadioptric lens combined with a 30mm lens (referred to as "cata-30mm"). In all cases, the calibration plane of known Euclidean structure was a 20 inch LCD screen. The number of calibration views was between 8 and 10 for the different experiments.

Figure 14(a) gives the computed focal length of the 15mm, 3.5mm, cata-12mm and cata-30mm, w.r.t. the distance d to the distortion center, using all methods. These are the functions that were recovered without further optimization based on the reprojection error. Table III shows the average reprojection errors of the three algorithms for all cameras. In many cases, the cameras could be calibrated from a single image of the screen (*cf.* Fig. 12 for the RCC), although in general, we recommend using at least five images for good stability. Recall that our structured light based matching provides a large number of correspondences. The cata-12mm was calibrated with all the approaches (except for Hartley-Kang's approach where only the portion of the image corresponding to forward cones was used) with very similar results (*cf.* Fig. 14). As for the cata-30mm, since it was found to be non-central, as described later in this section. The RCC gave very accurate results, however only with a limited number of planes. The difficulties arose when not enough data were available to allow a good fitting of the calibration conics. We handled this by dropping these planes and use only the other ones.

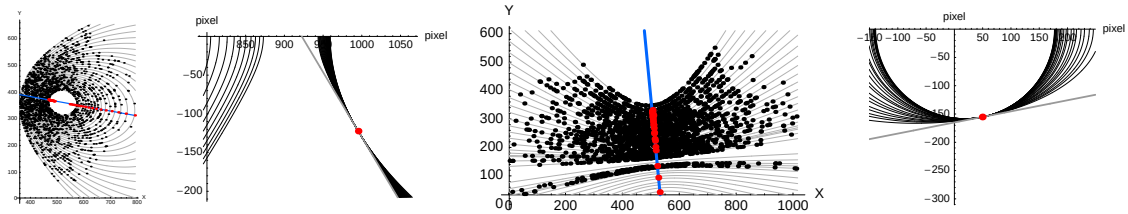
If discrete values for f_d are computed, instead of e.g. a polynomial function, then only a subset of distortion circles are used for calibration; others can then be extrapolated or



(a) SVP (b) SVP (c) SVP

(d) NSVP (e) NSVP (f) NSVP

Fig. 10. The two optimization algorithms (see text for details) with simulated data. **(a)+(d)** The recovered displacement t_d. **(b)+(e)** The error distribution. **(c)+(f)** The focal length functions compared to the ground truth.



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

Fig. 12. Calibration with the RCC approach. **(a)** Fitted ellipses for the Goyo 3.5mm lens and **(b)** Corresponding hyperbolas, computed intersection and optical axis (gray line). **(c)** For the cata-12mm camera, the intersection between the calibration plane and the cones yielded ellipses and hyperbolas, constraining the viewpoint to lie respectively on hyperbolas and ellipses. **(d)** Intersection of the viewpoint conics.

interpolated from a polynomial fitting of the data. Let us define this polynomial p; from the camera model, it is best to ensure that its derivative at 0 (corresponding to the distortion center) is 0. This constraint is due to the symmetry of the distortion model. Another criterion is that the function should be monotonically decreasing. This last constraint is not directly enforced in our algorithms. However, this did not seem to be an issue in our tests. In practice, polynomials of degree 5 appeared to be sufficient. To handle the case of omnidirectional cameras more appropriately, the interpolation is carried out with the view angle instead of the focal length. In this case, a monotonically increasing polynomial passing through 0 can also be fitted (see Fig. 14(b)).

Both catadioptric cameras cata-12mm and cata-30mm are typical examples of configurations yielding multiple viewpoints. Indeed, both mirrors are parabolic and the mounted lenses are perspective [25]. However, only the second one was found to be NSVP (*cf.* Fig. 13 and Table IV). We conjecture that although our 12.5mm camera is not orthographic, it has a field of view sufficiently small to provide a locus of viewpoints very close to a single effective viewpoint. To verify our hypothesis, it would be useful to perform the test with more specialized equipment like in [25].

Evaluating the results based on the reprojection error can lead to biased conclusions in the case of a generic model. Indeed, the model offers more freedom which allows to fit the data better. Meaningful quantitative results were obtained for the Goyo 3.5mm lens, using a pose estimation procedure.

TABLE IV

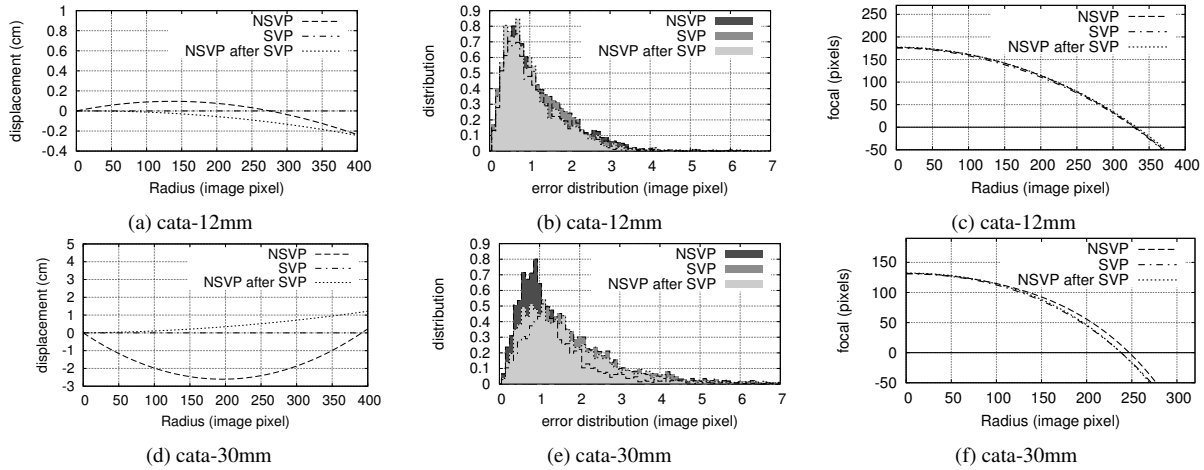
COMPARISON OF THE AVERAGE REPROJECTION ERROR FOR DIFFERENT CONSTRAINTS ON THE VIEWPOINT. 'L' REFERS TO THE LINEAR ALGORITHMS AND 'NL' TO NON-LINEAR OPTIMIZATION OF THE PARAMETERS.

	Algorithms-Constraints				
Cameras	SVP-L	SVP-NL	NSVP after SVP-NL	NSVP-L	NSVP-NL
cata-12mm	1.65	1.08	1.06	1.58	1.12
cata-30mm	2.22	1.93	1.73	1.42	1.33

Using a translation stage, the camera was moved to three positions with known relative motion (no rotation, known translation). Using the calibration information (obtained using other images), the pose of the camera relative to the calibration plane was computed for all three positions. From this, the relative motions were computed and compared to the ground truth. The results presented in Table V show a good stability for all methods.

Images from three panoramic cameras were rectified based on the calibration results (*cf.* Fig. 16 and 11(a,b)). For the wide-angle Goyo lens and the cata-12mm, the results seem to be very good, even towards the image borders (*cf.* fig. 16(b) and the inset images in 11(b)).

Finally, a home-made catadioptric device built from a Fujinon 12.5mm lens pointed at a roughly spherical mirror was tested (*cf.* Fig. 15). Although its radial configuration was not



(a) cata-12mm (b) cata-12mm (c) cata-12mm

(d) cata-30mm (e) cata-30mm (f) cata-30mm

Fig. 13. The two optimization algorithms (see text for details) with our catadioptric cameras. **(a)+(d)** The recovered displacement t_d, **(b)+(e)** The distribution of reprojection errors. **(c)+(f)** The recovered focal length functions.

(a) (b) (c)

Fig. 16. Image rectification. **(a)** Original images. **(b)** Rectified image for the Goyo 3.5mm. **(c)** Rectified image for the home-made catadioptric camera. Small inset images show rectification of the border regions.

TABLE V

RESULT FOR POSE ESTIMATION. THE CAMERA WAS MOVED TO THREE POSITIONS WITH KNOWN RELATIVE MOTION. COEFFICIENTS p_{ij} AND a_{ij} DENOTE THE DISTANCE (IN CENTIMETERS) AND RELATIVE ANGLE (IN DEGREES) BETWEEN CAMERA POSITIONS i AND j.

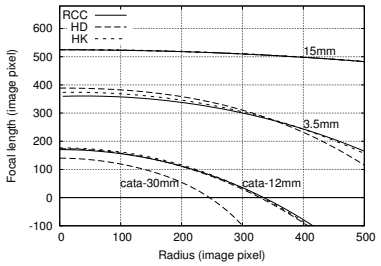
	Position			Angle		
Algorithms	p_{01}	p_{12}	p_{02}	a_{01}	a_{12}	a_{02}
Ground truth	5	5	10	0°	0°	0°
RCC	4.94	5.00	9.93	0.83°	0.1°	0.92°
HB	4.90	4.94	9.85	0.79°	0.79°	1.6°
HK	4.90	4.96	9.86	0.84°	0.68°	1.51°

perfect, the distortion center could be found and a satisfying calibration could be obtained with our methods. The "HB" approach gave the best results because it could take advantage of up to eight images, which is more robust to the imperfect configuration of the camera. The rectification is surprisingly good for a large part of the image, especially around the borders (*cf.* Fig. 16(a,c)). The remaining distortions in the center were found to be caused by a small bump on the "mirror's" surface.

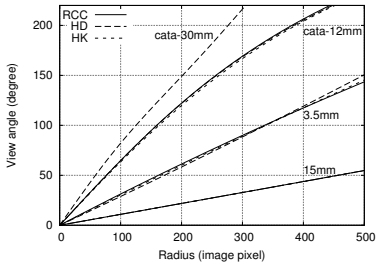
VIII. SUMMARY AND CONCLUSION

We have proposed new calibration approaches for a camera model that may be a good compromise between flexibility and stability for many camera types, especially wide-angle ones. Previous work showed that the RCC approach might have a limited practical usability because of stability issues [27]. This was because only one calibration plane could be used directly and because camera position was recovered in two steps: conic fitting and finding the closest point to a set of viewpoint conics. Both issues were addressed in this paper and we showed that the use of the RCC approach is very well suited when performed with only few camera poses. However, our homography-based approach introduced in this paper is preferable. It can be adapted to using a polynomial distortion model and extended to NSVP configurations. This allows to perform calibration without a dense plane-to-image matching, unlike the previous approach. Those reasons, and the fact it was the most reliable in our experiments, lead us to recommend the HB approach over RCC.

Hartley-Kang's approach gives very good results with the benefit that it has an elegant solution for the distortion center estimation. Our approach gives a similar accuracy, but can also



(a)



(b)

Fig. 14. - **(a)** Recovered focal length (in pixels) for the three algorithms (after polynomial fitting of the data). **(b)** Recovered view angle (in degrees). These calibration curves were obtained without optimization based on the reprojection error. Performing such an optimization lead to very similar results in general. Observe that for the catadioptric cameras there are negative focal lengths, meaning that their view angle is larger than 180°.

Fig. 15. Home-made catadioptric camera built from a Basler A201bc camera with a Fujinon 12.5mm lens pointed at a Christmas ornament representing a roughly spherical mirror.

deal with NSVP cameras, which is one of the main goals of this work.

ACKNOWLEDGMENT

Jean-Philippe Tardif and Peter Sturm gratefully acknowledge support by the French ANR project CAVIAR. Funding was also provided to Jean-Philippe Tardif by the Natural Sciences and Engineering Research Council of Canada. The authors thank the reviewers for useful comments.

REFERENCES

[1] S. Baker, S.K. Nayar. A Theory of Single-Viewpoint Catadioptric Image Formation. International Journal of Computer Vision, 35(2), 1-22, 1999.
[2] J.P. Barreto, K. Daniilidis. Unifying image plane lifting for central catadioptric and dioptric cameras. Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, 151-162, 2004.
[3] W. Boehm, H. Prautzsch. *Geometric Concepts for Geometric Design*. A.K. Peters, 1994.
[4] D.C. Brown. Close-Range Camera Calibration. Photogrammetric Engineering, 37(8), 855-866, 1971.
[5] G. Champleboux, S. Lavallée, P. Sautot, P. Cinquin. Accurate Calibration of Cameras and Range Imaging Sensors: the NPBS Method. International Conference on Robotics and Automation, 1552-1558, 1992.
[6] F. Devernay, O.D. Faugeras. Straight lines have to be straight. Machine Vision and Applications, 13(1), 14-24, 2001.
[7] D. Claus, A.W. Fitzgibbon. A Rational Function Model for Fish-eye Lens Distortion. International Conference on Computer Vision and Pattern Recognition, 213-219, 2005.
[8] W. Gander, G.H. Golub, R. Strebel. Fitting of Circles and Ellipses. *BIT*, 34, 556-577, 1994.
[9] C. Geyer, K. Daniilidis. A Unifying Theory for Central Panoramic Systems and Practical Applications. European Conference on Computer Vision, 445-461, 2000.
[10] K.D. Gremban, C.E. Thorpe, T. Kanade. Geometric Camera Calibration using Systems of Linear Equations. International Conference on Robotics and Automation, 562-567, 1988.
[11] M.D. Grossberg, S.K. Nayar. The Raxel Imaging Model and Ray-Based Calibration. International Journal of Computer Vision, 61(2), 119-137, 2005.
[12] P. Gurdjos, R. Payrissat. Plane-based Calibration of a Camera with Varying Focal Length: the Centre Line Constraint. British Machine Vision Conference, 623-632, 2001.
[13] P. Gurdjos, A. Crouzil, R. Payrissat. Another Way of Looking at Plane-Based Calibration: the Centre Circle Constraint. European Conference on Computer Vision, 252-266, 2002.
[14] R.I. Hartley, S.B. Kang. Parameter-free Radial Distortion Correction with Centre of Distortion Estimation. International Conference on Computer Vision, 1834-1841, 2005. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(7), 1209-1321, Aug. 2007.
[15] Intel Open Source Computer Vision Library. `http://www.intel.com/research/mrl/research/opencv/`.
[16] J. Kannala, S.S. Brandt. A Generic Camera Model and Calibration Method for Conventional, Wide-Angle, and Fish-Eye Lenses. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(8), 1335-1340, 2006.
[17] S.-S. Lin, R. Bajcsy. True single view point cone mirror omni-directional catadioptric system. International Conference on Computer Vision, 102-107, 2001.
[18] B. Micusik, T. Pajdla. Structure from Motion with Wide Circular Field of View Cameras. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(7), 1135-1149, Jul. 2006.
[19] J. Salvi, J. Pagès, J. Batlle. Pattern codification strategies in structured light systems. Pattern Recognition, 37(4), 827-849, 2004.
[20] D.E. Stevenson, M.M. Fleck. Nonparametric correction of distortion. TR 95-07, University of Iowa, 1995.
[21] A. Strzebonski. Cylindrical Algebraic Decomposition. From MathWorld – A Wolfram Web Resource, created by E.W. Weisstein. `http://mathworld.wolfram.com/CylindricalAlgebraicDecomposition.html`.
[22] P. Sturm. Algorithms for plane-based pose estimation. International Conference on Computer Vision and Pattern Recognition, 1010-1017, 2000.
[23] P. Sturm, S. Ramalingam. A generic concept for camera calibration. European Conference on Computer Vision, 1-13, 2004.
[24] P. Sturm, S. Maybank. On Plane-Based Camera Calibration. International Conference on Computer Vision and Pattern Recognition, 432-437, 1999.
[25] R. Swaminathan, M. Grossberg, S. Nayar. Caustics of catadioptric cameras. International Conference on Computer Vision, 2-9, 2001.
[26] J.-P. Tardif, S. Roy, A MRF formulation for coded structured light. Internation Conference on 3D Modeling and Imaging, 22-29, 2005.
[27] J.-P. Tardif, P. Sturm. Calibration of Cameras with Radially Symmetric Distortion. Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, 44-51, 2005.
[28] J.-P. Tardif, P. Sturm, S. Roy. Self-calibration of a general radially symmetric distortion model. European Conference on Computer Vision, 186-199, 2006.
[29] J.-P. Tardif, P. Sturm, S. Roy. Plane-based Self-calibration of Radial Distortion. International Conference on Computer Vision, 2007.
[30] J.-P. Tardif, P. Sturm., M. Trudeau, S. Roy. Calibration of Cameras with Radially Symmetric Distortion. Technical Report, INRIA, 2008 (to appear).
[31] S. Thirthala, M. Pollefeys. The Radial Trifocal Tensor. A tool for calibrating the radial distortion of wide-angle cameras. International Conference on Computer Vision and Pattern Recognition, 321-328, 2005.
[32] S. Thirthala, M. Pollefeys. Multi-View Geometry of 1D Radial Cameras and its Application to Omnidirectional Camera Calibration. International Conference on Computer Vision, 1539-1546, 2005.
[33] Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. Rapport de Recherche 2676, INRIA, 1995.
[34] Z. Zhang. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330-1334, 2000.