On the Calibration of Non Single Viewpoint Catadioptric Sensors

Alberto Colombo[1], Matteo Matteucci[2], and Domenico G. Sorrenti[1]

[1] Università degli Studi di Milano Bicocca,
Dipartimento di Informatica, Sistemistica e Comunicazione
[2] Politecnico di Milano, Dipartimento di Elettronica e Informazione

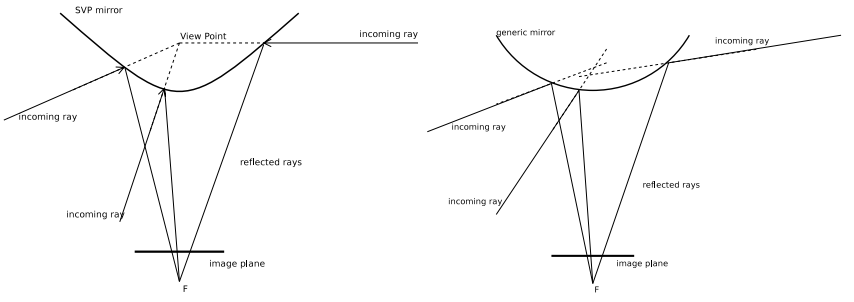
Abstract. A method is proposed for calibrating Catadioptric Omnidirectional Vision Systems. This method, similarly to classic camera calibration, makes use of a set of fiducial points to find the parameters of the geometric image formation model. The method makes no particular assumption regarding the shape of the mirror or its position with respect to the camera. Given the camera intrinsic parameters and the mirror profile, the mirror pose is computed using the projection of the mirror border and, eventually, the extrinsic parameters are computed by minimizing distance between fiducial points and back-projected images.

1 Introduction

A catadioptric imaging system is built combining reflective (catoptric) and refractive (dioptric) elements. They allowed building omni-directional sensors [1] at an affordable price, which made catadioptric systems become of widespread use also in fields like video-surveillance and teleconferencing; in contrast omnidirectional dioptric-based systems usually cost much more.

The choice of a specific mirror shape is based mainly on required resolution, manufacturing cost, assembly cost and geometric properties. With respect to these properties, mirror shapes can be classified in two categories: *Single Viewpoint* (SVP, also called *central*) imaging systems, and *Non-SVP* (*non-central*) imaging systems [2]. In SVP mirror systems, the lines of those incoming rays, which are reflected to the image plane, intersect at a single point, called *Effective Viewpoint* (see Fig. 1). It has been proved that the only useful mirror shapes with this property are hyperboloid mirrors (coupled with a perspective camera) and paraboloid mirrors (coupled with an orthographic camera). The uniqueness of the viewpoint allows a mapping of every scene point on the image plane, similarly to what happens with perspective cameras. The result of the catadioptric projection is equivalent to a panoramic image taken with a camera rotating around its viewpoint, and the resulting catadioptric image can be rectified to obtain an omni-directional undistorted perspective image called *panorama* [3,4]. On the other hand, SVP imaging systems suffer from a number of problems. Firstly, the SVP constraint leaves little room for changing other optical parameters that could be useful to tune because the relative pose of the mirror and

G. Lakemeyer et al. (Eds.): RoboCup 2006, LNAI 4434, pp. 194–205, 2007.



(a) central sensor (b) non-central sensor

Fig. 1. Comparison between a central catadioptric system (featuring an hyperbolic mirror) and a non-central one. Note that in *a* all rays intersect at the hyperboloid focus, while in *b* there is no unique intersection for the rays.

the camera are strictly bound. Due to the curved vertical section of the mirrors, spatial resolution is badly distributed: typically the ground and the camera fill the central part of the omni-directional image, leaving little room, on the peripheral part of the image, for the (interesting) surrounding environment. Regarding assembly costs, cameras based on hyperboloid and paraboloid mirrors are more expensive needing very accurate assembly, for the viewpoint to be unique.

The work on non-SVP sensors aimed at overcoming these limitations. Early non-SVP systems were based on a spheric mirror [5] or on a conical mirror [1] due to the low cost of such mirrors and the mathematical simplicity of the projection. In general, however, non-SVP systems pose no constraints on mirror shape and position. Catadioptric images taken with such devices cannot be rectified though, i.e., there is no function that takes a non-SVP omni-directional image and unwraps it into a perspective image. The freedom in the mirror shape, however, can be exploited for designing optimal mirrors for each specific application. Examples include single-camera stereo-vision [6], multi-part mirrors [7] and many kind of non-distorting mirrors [8].

This paper contribution is an algorithm to calibrate non-SVP systems using the projection of the mirror border and a set of points whose scene and image coordinates are known. By *calibration* we mean to estimate the mirror pose (with respect to the camera), as well as the camera pose (with respect to a *world reference frame*). We will be assuming that the camera intrinsics are known. We will therefore call *intrinsic parameters* of the catadioptric system the mirror pose; *extrinsic parameters* of the catadioptric system the camera pose in the world reference frame. By knowing the calibration parameters it will be possible to map every pixel onto its interpretation line. These lines can then be exploited for metrology. The assumption we make is that the mirror shape is known.

Most of the scientific literature regarding catadioptric system calibration deals with SVP systems, probably because of their availability off-the-shelf. Since SVP systems assume that the mirror pose is known, their calibration consists of finding 1) the intrinsic parameters of the mirror-and-camera pair (seen as a single

indivisible object) and 2) the extrinsic parameters, similarly to what happens with traditional cameras. There are even some ready-to-use calibration tools available on the Internet, e.g. [9], [10], [3].

The calibration of non-SVP systems, on the other hand, is a surprisingly unstudied problem. The only exception is the particular case of non-SVP systems featuring a conic mirror, like the COPIS [1]. In this case, if the cone axis matches the camera optical axis, it can be proved that there is one viewpoint for each direction the optical rays come from. The set of these viewpoints, sometimes referred to as *view-circle*, allows the conic projection to be described through relatively simple equations; it also allows to calibrate the camera intrinsic parameters together with the rest of the catadioptric system, like one would do with a SVP system. In our experiments, for the sake of ease, we used a conic mirror, but this will not necessarily be coaxial to the camera. In fact, the algorithm we develope (see Sect. 2.2) is equally fit to any mirror shape.

Strelow's method [11] for calibrating generic catadioptric systems separates the perspective projection from the catadioptric part and deals only with the latter. The mirror can be of any shape as long as the surface equation is known. The method consists of calculating the mirror position w.r.t. the camera (i.e., the catadioptric system intrinsic parameters) using the image of one or more targets whose three-dimensional coordinates (w.r.t. the camera) are known, and then minimizing the difference between the observed position on the image and their respective predicted back-projections, where these depend on the roto-translation describing the mirror pose. This requires solving a non-linear system of equations. This method obtains a precision of about 1%, but it suffers from two problems. Firstly, the calibration points coordinates must be known w.r.t. the camera reference frame, i.e., the camera extrinsic parameters must be known *before* the calibration. The second problem is tied to the intrinsic complexity of a non-SVP projection model. In order to back-project a scene point onto the image one has to find the line, exiting from the scene point, which, once reflected by the mirror, passes through the camera pin-hole. To minimize the distance between image points and back-projected scene points this method requires for each minimization iteration, and for each point, to numerically resolve a non linear system.

2 Our Method

In the first place, we shall start by looking at the geometric model adopted to describe catadioptric imaging systems; we shall then examine the two phases of our method: mirror localization and extrinsic parameters computation. In Fig. 2 we see a schematic drawing of a typical catadioptric system. The cube in the middle, indicated with C, represents the camera (*camera reference frame* $x_C y_C z_C$). Note that this reference frame has a precise physical meaning, as its origin is the center of projection. It is not defined arbitrarily, but depends on the camera intrinsic parameters. In front of the camera there is a solid of revolution (SOR) mirror. Solids of revolution feature an *axis of symmetry* and a *profile*. We

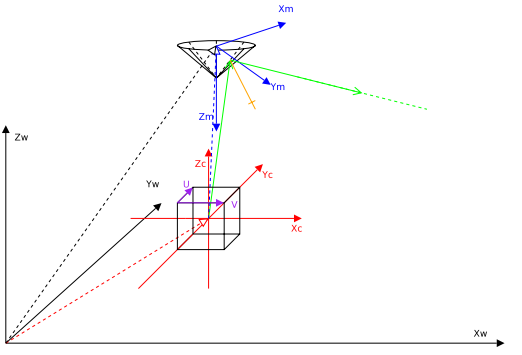


Fig. 2. Geometric model of a catadioptric imaging system

can define a reference frame local to the mirror whose z-axis corresponds to the SOR axis and whose origin lies in the center of the mirror base (*mirror reference frame* $x_M y_M z_M$). The mirror is in a generic position with respect to the camera: its axis is not parallel to the optical axis and it does not pass through the pinhole. The unknown roto-translation from the camera reference frame to the mirror is $\mathcal{T}_{C\to M}$: we will see how to compute it in Sect. 2.1. In Fig. 2 we see a third reference frame, $x_W y_W z_W$, called *world reference frame*; all distances measured by the user are defined in this frame. Its z-axis is vertical upward and the origin is usually located in a easily recognizable point. In Sect. 2.2 we will see how to compute the roto-translation $\mathcal{T}_{W\to C}$ and, as a consequence, the roto-translation $\mathcal{T}_{W\to M}$. In analogy with the terminology of traditional camera calibration, we will call the coefficients of $\mathcal{T}_{W\to C}$ *extrinsic parameters* of the catadioptric system. Because all rigid transformations are invertible, we will refer to the inverse of the above mentioned roto-translations, i.e., $\mathcal{T}_{M\to C}$, $\mathcal{T}_{M\to W}$ e $\mathcal{T}_{W\to C}$, whenever this would help readability.

2.1 Mirror Localization

The method exposed here is based on the analysis of the conic on the image plane corresponding to the projection of a circle. *The perspective projection of a circle is an ellipse.* The mirror pose can be estimated by this circle-to-ellipse correspondence. The idea of using the projected ellipse to estimate the circle pose is not new and it has found various applications (e.g. [12]); an introduction to this problem can be found in [13]. From the observed perspective projection of a circle with known radius, the circle support plane and the position of its center can be inferred analytically. Let the real, symmetric, non-zero matrix Q be the ellipse corresponding to the projection of a circle of radius r in the normalized camera reference frame (i.e., where the camera pinhole is in the origin and the image plane has equation $z = 1$). Because the matrix representation of quadrics is unique, up to an arbitrary scale factor, we choose a scale κ such that $\det \kappa Q = 1$. From matrix theory, we know $\kappa = \sqrt[3]{-\frac{1}{\det Q}}$. In the rest of this section, Q will

refer to the normalized matrix. Let λ_1, λ_2 e λ_3 be the eigenvalues of Q, where $\lambda_3 < 0 < \lambda_1 \leq \lambda_2$, and let $\hat{u}_1$, $\hat{u}_2$ and $\hat{u}_3$ be the (orthonormal) eigenvectors corresponding to λ_1, λ_2 and λ_3 respectively. The normal $\hat{n}$ to the circle support plane is given by:

$$\hat{n} = \hat{u}_2 \sin\theta + \hat{u}_3 \cos\theta \tag{1}$$

where $\sin\theta = \sqrt{\frac{\lambda_2 - \lambda_1}{\lambda_2 - \lambda_3}}$, $\cos\theta = \sqrt{\frac{\lambda_1 - \lambda_3}{\lambda_2 - \lambda_3}}$. The distance of the support plane from the pinhole is

$$d = \lambda_1^{\frac{3}{2}} r\,. \tag{2}$$

To compute the position of the circle centre one only needs to multiply the inverse of the ellipse with the normal to the support plane:

$$\boldsymbol{c} = Q^{-1} \cdot \hat{n}\,. \tag{3}$$

Note that the sign of eigenvectors $\hat{u}_{1,2,3}$ are arbitrary. Since both $\hat{n}$ and $-\hat{n}$ represent the same plane orientation, from (1) we deduce that

1. If $\lambda_1 = \lambda_2$, then $\hat{n} = \sqrt{\frac{\lambda_1 - \lambda_3}{\lambda_2 - \lambda_3}}\hat{u}_3$, therefore there exists only one solution regardless of the sign of $\hat{u}_3$.
2. If $\lambda_1 \neq \lambda_2$, then there exist two interpretations.

In the second case, both interpretations are physically plausible, but of course only one is real. Unfortunately, a single image is not enough to solve this ambiguity, therefore an implementation of this algorithm will need more information in order to choose one of the two solutions. We will bring both hypothesis foreward up to the solution of the extrinsic parameters step. Only at that point we will have the ambiguity solved, by choosing the solution that gives the best estimate (with the smallest residual). In other applications, where two concentric circles are used, the solution can be directly disambiguated [14]; this could be useful for multi-part mirrors, like the one in [7].

2.2 Calibration of the Extrinsic Parameters

The extrinsic parameters are the coefficients of the transformation $\mathcal{T}_{C \to W}$ that converts a point coordinates from the camera reference frame $x_C y_C z_C$ to the world reference frame $x_W y_W z_W$. To compute these parameters we take a set of N *calibration points* or *fiducial points* whose scene coordinates (w.r.t. the world reference frame) and image coordinates are known. If we know the camera intrinsic parameters we can use the fiducial points image coordinates to compute their interpretation lines, with respect to the camera system. If we know the pose of the mirror, with respect to the camera system, we can compute the reflected interpretation lines. We also know that these lines come from their corresponding scene points. We know the reflected interpretation lines with respect to the camera system, while the fiducial points scene coordinates are known with respect to the world system, but we do not know the transformation between the two. We

first see how to compute the reflection onto a surface of revolution, then we use the reflected interpretation lines for the estimation of the extrinsic parameters.

Reflection of the Interpretation Lines. If the mirror profile can be represented by a second degree equation, then the mirror shape is a quadric (a cone, a paraboloid or a hyperboloid) and the reflected line can be determined analytically. The same holds if the profile is a spline[1] of second degree or less, as long as the reflection is computed for each segment of the spline. On the other hand, when the profile is described by a third degree equation, the mirror shape is represented by a sixth degree equation that cannot be solved analytically [15]. In this work, without loss of generality, we implemented the reflection for quadric-shaped mirrors and we only applied it to conic mirrors. As explained in the following section, the same theory can be applied to mirrors of any shape using a local spline approximation of the profile.

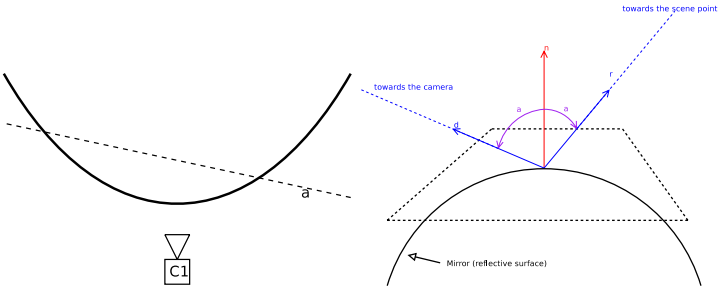
Let $Q \in \Re^{4\times4}$ be a symmetric, non-zero matrix representing the quadric. The locus of points belonging to the quadric surface is given by $F = \boldsymbol{x}^{\mathrm{T}} Q \boldsymbol{x} = 0$ where $\boldsymbol{x} = [x,\, y,\, z,\, 1]^{\mathrm{T}}$. Given a point $\boldsymbol{p} = [p_x, p_y, p_z]^{\mathrm{T}}$ and a direction $\hat{d} = [d_z, d_y, d_z]^{\mathrm{T}}$, $(||\hat{d}|| = 1)$, let $\boldsymbol{x}(t) = \boldsymbol{p} + \hat{d} \cdot t$ be the parametric equation of the line passing through $\boldsymbol{p}$ with direction $\hat{d}$, whose intersection with the quadric we want to compute. By substituting the line into the quadric, grouping and rewriting as a function of t we obtain $A \cdot t^2 + B \cdot t + C = 0$ which is a second degree polynomial in t. We solve it to find t_1, t_2 and substitute them into the line to find the intersection points $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ of the line with the quadric. Only one of the two intersections has a physical meaning, while the other is a point either behind the camera or outside the mirror or also occluded by the mirror itself (Fig. 4a). Once the intersection point $\boldsymbol{P}_i$ is known, we compute the normal to the quadric in $\boldsymbol{P}_i$ as

$$\hat{n}_i = \left[\left.\frac{\partial F}{\partial x_i}\right|_{\boldsymbol{P}_i} \left.\frac{\partial F}{\partial y_i}\right|_{\boldsymbol{P}_i} \left.\frac{\partial F}{\partial z_i}\right|_{\boldsymbol{P}_i} \right]^{\mathrm{T}}, \tag{4}$$

and then we apply the law of reflection to the line $\hat{d}$: $\hat{r} = 2\left(\hat{n}^{\mathrm{T}}\hat{d}\right)^{\mathrm{T}}\hat{n} - \hat{d}$ (Fig. 3).

Reflection on Mirrors of any Shape. In genral, mirrors are defined by the points of their profile, which can therefore approximated by a linear spline. This means that the mirror surface is a stack of conic trunks. To intersect a line with a conic trunk, one must first intersect the line with the full cone, and then check whether this intersection falls within the trunk. As we have a stack of trunks, the intersection must be computed separately for each trunk. Typically, only one intersection will survive the interval test; if more than one does, the one closest to the origin of the ray is chosen. In Fig. 4 ray r intersects all cones $(m_{1..3})$, but only the one with m_2 falls within the validity interval of the trunk.

[1] A *spline* is a curved line formed by two or more *vertices*, or *control points*, and a mathematical formula describing the curve(s) between them; if this formula is an n-th degree polynomial, then we get an n-th degree spline.



(a) Intersection of incoming ray with the mirror surface

(b) Reflection of a ray on a curved surface

Fig. 3. The reflection model for the interpretation lines

Computing the Roto-translation. The reflected interpretation lines pass through their corresponding scene points, i.e., the distance between a line and its point is zero. We look for the roto-translation $\mathcal{T}_{C\to W}$ that zeroes the distances between the fiducials and their reflected interpretation lines. Actually, because of noise, the lines will only pass *close* to the fiducials, i.e., we can only find a roto-translation M that minimizes the sum of the squares of the fiducial distances to their interpretation lines. If we call $\boldsymbol{p}_i$ and $\boldsymbol{q}_i(t)$, $(i = 1...N)$, respectively the fiducials and the parametric equations of the reflected interpretation lines. We look for a M that minimizes

$$\mathcal{J} = \sum_{i=1}^{N} \|\boldsymbol{p}_i - M\boldsymbol{q}_i(t_i)\| \tag{5}$$

where t_i is the value of the parameter t that minimize the distances between the lines $\boldsymbol{q}_i\,(t)$ and the fiducials $\boldsymbol{p}_i$, for a given M. M is computed by solving a nonlinear system of equations. The non-linearity comes from the orthonormality constraint on the rotational part of M.

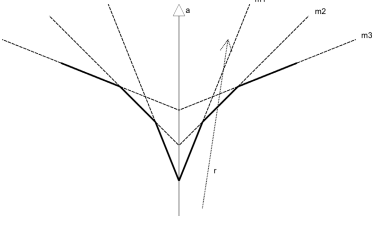


Fig. 4. Intersection of a line with a mirror profile approximated by a linear spline

3 Experimental Results

In order to test the correctness of the algorithm we tried it on synthetic images, so as to compare the results computed by our method against a ground truth. We performed the intrinsic and extrinsic calibrations separately, to avoid errors in the first phase to corrupt the results in the second phase. At the end of this section, we provide a complete calibration for a COPIS system.

3.1 Intrinsic Parameters

For space limits we present only two tests, using the same conic mirror, tilted by $\theta = 0°, 10°$ respectively. For each test we present the expected value (ground truth, GT) and the values obtained for the mirror position ($\boldsymbol{c}$, $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ respectively), for the projection of the mirror centre ($\boldsymbol{p}$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, in pixels) and for the normal to the support plane ($\hat{n}$, $\hat{n}_1$ and $\hat{n}_2$, where $\hat{n} = [\sin\theta, 0, \cos\theta]^{\mathrm{T}}$). We also computed the angles ϕ, ϕ_1, ϕ_2 respectively formed by $\hat{n}$, $\hat{n}_1$ and $\hat{n}_2$ with the z-axis of the camera reference frame. We present the differences $\Delta_1 = \phi - \phi_1$ and $\Delta_2 = \phi - \phi_2$ (expressed in degrees) between the expected angles and the computed ones. This difference is a good estimate for the mirror orientation error.

1. $\theta = 0°$

$$\hat{c} = \begin{bmatrix}0\ 0\ 100\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{c}_1 = \begin{bmatrix}0.066\ -0.146\ 100.003\end{bmatrix}^{\mathrm{T}}\ \boldsymbol{c}_2 = \begin{bmatrix}0.01\ -0.312\ 100.003\end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{p} = \begin{bmatrix}319.5\ 239.5\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_1 = \begin{bmatrix}319.681\ 240.703\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_2 = \begin{bmatrix}319.521\ 238.499\end{bmatrix}^{\mathrm{T}}$$

$$\hat{n} = \begin{bmatrix}0\ 0\ 1\end{bmatrix}^{\mathrm{T}},\ \hat{n}_1 = \begin{bmatrix}0.001\ 0.010\ 0.999\end{bmatrix}^{\mathrm{T}},\ \hat{n}_2 = \begin{bmatrix}0.000\ -0.008\ 0.999\end{bmatrix}^{\mathrm{T}}$$

$$\Delta_1 = 0.632,\ \Delta_2 = 0.547$$

1. $\theta = 5°$

$$\hat{c} = \begin{bmatrix}0\ 0\ 100\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{c}_1 = \begin{bmatrix}3.492\ 0.097\ 100.22\end{bmatrix}^{\mathrm{T}}\ \boldsymbol{c}_2 = \begin{bmatrix}0.174\ 0.069\ 100.281\end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{p} = \begin{bmatrix}319.5\ 239.5\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_1 = \begin{bmatrix}303.092\ 239.389\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_2 = \begin{bmatrix}319.169\ 239.66\end{bmatrix}^{\mathrm{T}}$$

$$\hat{n} = \begin{bmatrix}0.087\ 0\ 0.996\end{bmatrix}^{\mathrm{T}},\ \hat{n}_1 = \begin{bmatrix}0.048\ 0.000\ 0.998\end{bmatrix}^{\mathrm{T}},\ \hat{n}_2 = \begin{bmatrix}0.085\ 0.001\ 0.996\end{bmatrix}^{\mathrm{T}}$$

$$\Delta_1 = 2.222,\ \Delta_2 = 0.127$$

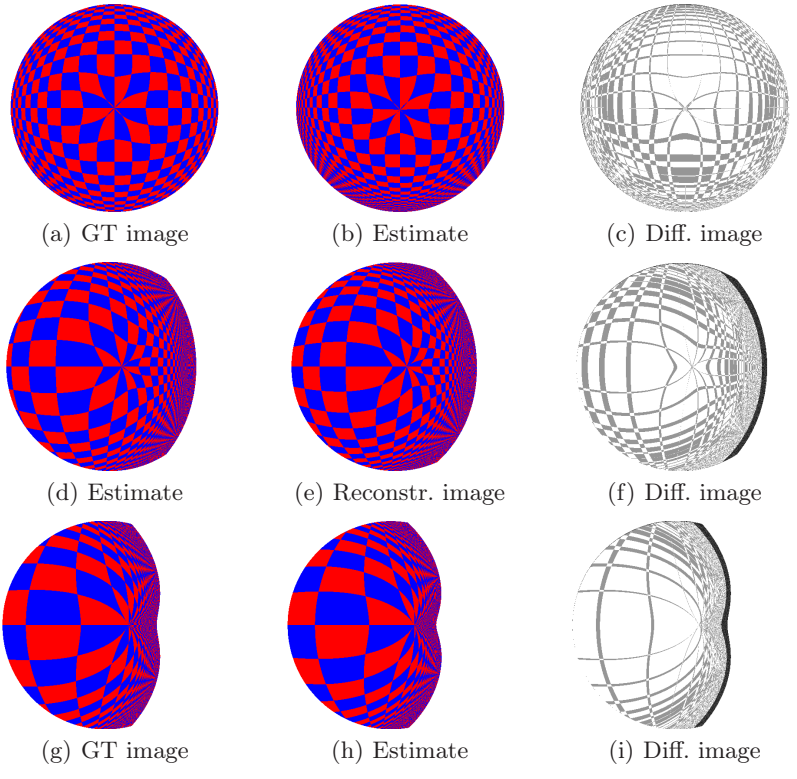
2. $\theta = 10°$

$$\hat{c} = \begin{bmatrix}0\ 0\ 100\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{c}_1 = \begin{bmatrix}7.003\ 0.092\ 100.029\end{bmatrix}^{\mathrm{T}}\ \boldsymbol{c}_2 = \begin{bmatrix}0.102\ 0.074\ 100.274\end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{p} = \begin{bmatrix}319.5\ 239.5\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_1 = \begin{bmatrix}285.694\ 238.972\end{bmatrix}^{\mathrm{T}},\ \boldsymbol{p}_2 = \begin{bmatrix}319.51\ 239.645\end{bmatrix}^{\mathrm{T}}$$

$$\hat{n} = \begin{bmatrix}0.173\ 0\ 0.984\end{bmatrix}^{\mathrm{T}},\ \hat{n}_1 = \begin{bmatrix}0.105\ 0.000\ 0.994\end{bmatrix}^{\mathrm{T}},\ \hat{n}_2 = \begin{bmatrix}0.175\ 0.001\ 0.984\end{bmatrix}^{\mathrm{T}}$$

$$\Delta_1 = 3.997,\ \Delta_2 = 0.067$$



(a) GT image (b) Estimate (c) Diff. image

(d) Estimate (e) Reconstr. image (f) Diff. image

(g) GT image (h) Estimate (i) Diff. image

Fig. 5. Mirror not tilted (a,b,c); tilted 5° (d,e,f); tilted 10° (g,h,i)

Note how an increase in the mirror tilt improves the precision in the (correct) pose estimate and increases also the difference between the two solutions.

3.2 Extrinsic Parameters

We generated a set of images where the world reference frame matched the camera reference frame, in order to have a trivial ground truth. We then applied the calibration algorithm keeping the fiducials set fixed but varying the initial estimate. This test allowed us to verify when, and how quickly, the algorithm converged. As a measure of the quality of the result we used (5), normalized with respect to the number N of fiducials (ξ_N). For each test, we present the initial estimate for rotation angle $\tilde{\theta}$ (in degrees, around the world z-axis), the initial estimate for translation $\tilde{T}$, the goodness of the initial estimate $\tilde{\xi}_N$ (before a solution is computed) and the goodness of the calibration ξ_N (after a solution is computed), and the number n of iterations performed by the algorithm. The results are reported in Tab. 1.

3.3 Calibration of a Real COPIS System

Before the catadioptric calibration can begin, the *intrinsic* parameters of the camera must be known. We used *camcal* [16] to calibrate our 8mm Sony XC711, which yielded the following intrinsic parameters matrix:

$$K = \begin{bmatrix} -626.438 & 0 & 317.218 \\ 0 & -604.007 & 231.424 \\ 0 & 0 & 1 \end{bmatrix}.$$

The mirror base is well visible and, once the ellipse equation is extracted, the following values are found for the normal $\hat{n}_{1,2}$, the centre position $\boldsymbol{c}_{1,2}$, and its projection $\boldsymbol{p}_{1,2}$:

$$\hat{n}_1 = \begin{bmatrix} 0.133 & 0.0059 & 0.991 \end{bmatrix}^{\mathrm{T}}, \quad \hat{n}_2 = \begin{bmatrix} -0.0072 & -0.0265 & 0.999 \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{c}_1 = \begin{bmatrix} 6.388 & -1.603 & 123.501 \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{c}_2 = \begin{bmatrix} 9.216 & -0.951 & 123.329 \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{p}_1 = \begin{bmatrix} 284.815 & 239.268 \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{p}_2 = \begin{bmatrix} 270.028 & 235.754 \end{bmatrix}^{\mathrm{T}}$$

Because the two solutions are so similar, we cannot solve the ambiguity yet and so we estimate the extrinsic parameters for both. We will eventually choose the solution with the smallest residual. The next step is the collection of fiducial points. We are currently developing a new version of the tool, whose calibration pattern is a checkerboard to ease fiducials collection. That not available, we choose 40 among the most visible dots of the grid pattern (the ones in red in Fig. 6*a*). Having collected the fiducials, an initial estimate is needed to compute the extrinsic parameters. This initial estimate consists of a rotation matrix $\tilde{R}$ and a translation vector $\tilde{T}$. The latter can be measured by hand, while the rotation may be input as the angle $\tilde{\theta}$ around the world z-axis, assuming that the camera is approximately pointed vertically upward. In this case, the initial estimate was

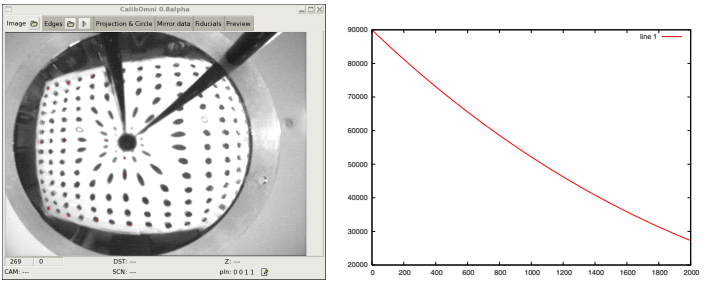
$$\tilde{R} = I, \quad \tilde{T} = \begin{bmatrix} -1105 & -780 & -20 \end{bmatrix}^{\mathrm{T}}$$

which yielded the following residuals (for the two possible mirror poses): $\tilde{\xi}_{N_1} = 80924$, $\tilde{\xi}_{N_2} = 2306.78$. After the extrinsic parameters are computed, we get the following results (the subscripts indicate either the first or the second pose):

$$R_1 = I, \quad T_1 = \begin{bmatrix} -1105 & -780 & -20 \end{bmatrix}^{\mathrm{T}}, \quad \xi_{N_1} = 80924$$

Table 1. Convergence speed and quality of the results w.r.t. the initial estimate

$\tilde{\theta}$	0	0	0	0	0	0	1	2	10	1	1
$\tilde{T}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 10 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$
$\tilde{\xi}_N$	0	1.232	4.926	123.143	33.222	470.88	48.321	1753.97	5109.54	205.589	808.635
ξ_N	0	0	0	0	0	0	0	0	112.241	0	0
n	1	7	14	68	35	116	82	161	490	107	225



(a) Calibration pattern (b) Convergence speed

Fig. 6. Calibration of a COPIS system

$$R_2 = \begin{bmatrix} 0.999 & 0.019 & -0.02 \\ -0.019 & 0.999 & -0.021 \\ 0.019 & 0.022 & 0.999 \end{bmatrix}, \quad T_2 = \begin{bmatrix} -1106.29 \\ -760.09 \\ -87.88 \end{bmatrix}, \quad \xi_{N_2} = 699.438\,.$$

Note how the residual of option 1 is much greater than the residual of option 2. Also note that a solution could not be computed for option 1. This means that the correct mirror position is option 2. The convergence graph can be seen in Fig. 6*b*. Using these 40 fiducials, an AMD 2400+ can compute approximately 100 iterations per second.

A comparison with our method is unfortunately unfeasable. The only other tool that matches ours in flexibility is Strelow's [11], but it is based on a different setup and no software is available for download. All other tools are not generic w.r.t. mirror shape or pose.

4 Conclusions

We described a mathematical model for catadioptric image formation which is generic with respect to mirror shape and pose and to the camera optics. Then, we introduced a calibration method that allows to estimate the model parameters. The experimental session proves the method to be valid and gives a quantitative measure of the precision The main flaw of decomposing the calibration process in mirror localization and extrinsic parameters estimate is that an error in the first phase propagates to the second phase. A way to overcome this problem is to add a third phase where a global post-optimization is performed. This phase consists of solving a non-linear system where all parameters (both intrinsic and extrinsic) are unknown, and where the results of the previous two phases are used as an initial estimate; this is what we are currently developing.

References

1. Yagi, Y., Kawato, S.: A panorama scene analysis with conic projection. In: Proc. IROS90, vol. 1, pp. 181–187 (1990)
2. Baker, S., Nayar, S.: A theory of catadioptric image formation. In: Proc. ICCV98
3. Geyer, C., Daniilidis, K.: Paracatadioptric Camera Calibration. IEEE T-PAMI 25(5), 687–695 (2002)
4. Kaidan Incorporated: `http://www.kaidan.com/`
5. Charles, J.: How to build and use an all-sky camera. Astronomy Journal (1987)
6. Conroy, T.L., Moore, J.B.: Resolution invariant surfaces for panoramic vision systems. In: Proc. ICCV99
7. Blind for the review: 36(2-3), 87–102
8. Gaspar, J., Decco, C., Okamoto, J.J., Santos-Victor, J.: Constant resolution omnidirectional cameras. In: Proc. OmniVis02, pp. 27–34 (2002)
9. Mei, C.: Omnidir. Calibr. Toolbox Extension (2005), `http://www-sop.inria.fr/icare/personnel/Christopher.Mei/ChristopherMeiPhDStudentToolbox.html`
10. Scaramuzza, D., Martinelli, A., Siegwart, R.: A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion. In: Proc. of ICVS06 (2006)
11. Strelow, D., Mishler, J., Koes, D., Singh, S.: Precise Omnidirectional Camera Calibration. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01) 1, 689 (2001)
12. Wang, J.G., Sung, E.: Eye gaze estimation from a single image of one eye. Ninth IEEE International Conference on Computer Vision 1, 136–143 (2003)
13. Kanatani, K.: Geometric Computer Vision. Oxford Science Publications, Oxford (1993)
14. Fremont, V., Ryad, C.: Direct camera calibration using two concentric circles from a single view. In: Proc. ICAT02 (2002)
15. Van Wijk, J.: Ray Tracing Objects Defined by Sweeping Planar Cubic Splines. In: ACM Transaction on Graphics, ACM Press, New York (1984)
16. Tarel, J.P., Vezien, J.M.: CamCal v.10 Manual - A Complete Software Solution for Camera Calibration. Technical report, INRIA (1996)