

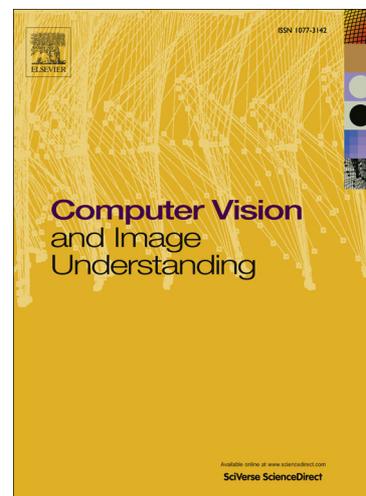
## Accepted Manuscript

Calibration of mirror position and extrinsic parameters in axial non-central catadioptric systems

Luis Perdigoto, Helder Araujo

PII: S1077-3142(13)00063-5  
DOI: <http://dx.doi.org/10.1016/j.cviu.2013.04.001>  
Reference: YCVIU 1975

To appear in: *Computer Vision and Image Understanding*



Please cite this article as: L. Perdigoto, H. Araujo, Calibration of mirror position and extrinsic parameters in axial non-central catadioptric systems, *Computer Vision and Image Understanding* (2013), doi: <http://dx.doi.org/10.1016/j.cviu.2013.04.001>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Calibration of mirror position and extrinsic parameters in axial non-central catadioptric systems

Luis Perdigoto<sup>a,b,\*</sup>, Helder Araujo<sup>a</sup>

<sup>a</sup>*Institute for Systems and Robotics, Dept. of Electrical and Computer Engineering,  
University of Coimbra, 3030 Coimbra, Portugal*

<sup>b</sup>*ESTG, Polytechnic Institute of Leiria, 2411-901 Leiria, Portugal*

---

## Abstract

We propose a novel calibration method for catadioptric systems made up of an axial symmetrical mirror and a pinhole camera with its optical center located at the mirror axis. The calibration estimates the relative camera/mirror position and the extrinsic rotation and translation w.r.t. the world frame. The procedure requires a single image of a (possibly planar) calibration object. We show how most of the calibration parameters can be estimated using linear methods (Direct-Linear-Transformation algorithm) and cross-ratio. Two remaining parameters are obtained by using non-linear optimization. We present experimental results on simulated and real images.

*Keywords:* Non-central catadioptric vision systems, Calibration

---

## 1. Introduction

Catadioptric vision systems use a combination of cameras and mirrors to acquire images. They can provide some advantages over more traditional camera systems, namely in terms of increased field-of-view (usually through reflection off curved mirrors) and/or single image multi-view geometry (with the use of multiple mirrors). Several configurations have been proposed and studied, alongside with tailor-made or more generic calibration methods.

---

\*Corresponding author. Tel.: +351 244820300; fax: +351 244820310  
Email addresses: [luis.perdigoto@ipleiria.pt](mailto:luis.perdigoto@ipleiria.pt) (Luis Perdigoto), [helder@isr.uc.pt](mailto:helder@isr.uc.pt) (Helder Araujo)

8 *1.1. Previous work on catadioptric calibration*

9 Central catadioptric systems [1] allow for a single-viewpoint projection model,  
10 by the use of particular mirror shapes restrictively aligned with an orthographic  
11 or perspective camera. Calibration methods for these systems include using the  
12 image of lines in the scene [2, 3, 4], self-calibration using tracked point on sev-  
13 eral images [5, 6] or using a calibration pattern [3, 7, 8, 9]. A recent review and  
14 comparison of calibration techniques focusing on central systems can be found  
15 in [10].

16 Some calibration methods propose a general, un-parameterized, camera [11,  
17 12, 13]. These can model central and non-central catadioptric systems, as well  
18 as more unconventional camera designs. The intrinsic calibration of the camera  
19 consists on associating a 3D direction with each pixel in image.

20 Most non-central catadioptric systems are modeled as a perspective camera  
21 and an axial symmetric mirror of conical section (sphere, paraboloid, ellipsoid  
22 and hyperboloid). The geometry of image formation is dependent on the in-  
23 trinsic parameters of the camera, on the particular shape and relative position  
24 of the mirror. Some calibration methods assume an independent calibration  
25 of the perspective camera [14, 15, 16], which can be robustly achieved using  
26 well-established techniques for conventional cameras. Many consider the mirror  
27 shape to be accurately known from the manufacturing process [17, 18, 19, 20].  
28 The mirror pose w.r.t. the camera is often estimated by identifying the mirror  
29 boundaries in the image (usually a conic) [17, 14, 21, 6, 20].

30 Self-calibration approaches use point correspondences on several (at least  
31 two) images [22, 17, 23]. Caglioti et al. [24] used the reflected image of lines  
32 on axial-symmetric mirrors as the calibration object. Sagawa et al. [25] applied  
33 projected parallel light to estimate mirror location. Grossberg and Nayar [26]  
34 and Tardif and Sturm [27] used a computer screen and a projector to generate  
35 coded structured-light calibration patterns and achieve a dense mapping of the  
36 image pixels. Morel and Fofi [28] used polarized light.

37 A comprehensive survey of camera models used in panoramic image acqui-

38 sition devices, as well as calibration approaches, is presented in [29].

### 39 1.2. Proposed method

40 Our paper presents a novel calibration method for axial catadioptric systems.  
41 By “axial catadioptric” [21] we mean a vision system made up of a pinhole  
42 camera and a mirror, such that

- 43 • The mirror is rotationally symmetric around an axis;
- 44 • The camera’s optical center is placed on the mirror’s axis.

45 There are no additional constraints on the relative position of the camera and  
46 mirror. The camera’s principal axis is not necessarily aligned with the axis of  
47 the mirror.

48 The constraint of placing the projective camera’s optical center on the mir-  
49 ror axis is acceptable for most systems. When using spherical mirrors, this  
50 constraint becomes irrelevant, as a symmetry axis passing through the camera  
51 always exists. We note that, although we focus on non-central systems, this  
52 model also includes central cameras. Our calibration technique can be easily  
53 applied to dioptric systems of similar characteristics, like fisheye lenses.

54 Our method is capable of calibrating

- 55 • The mirror position w.r.t. the pinhole camera;
- 56 • The extrinsic parameters of the camera, i.e., pose w.r.t. world coordinates.

57 It uses a single image of a known point pattern, i.e., a calibration object. This  
58 calibration object can be planar, although, as we will show, additional processing  
59 is required in that case.

60 The method is divided in 3 steps, executed in sequence. The first step esti-  
61 mates the intersection point between the mirror axis and the image plane, which  
62 we will call the *vertex point*. In systems where the camera is aligned with the  
63 mirror, this point coincides with the image center. The calibration is achieved  
64 by using the cross-ratio as an invariant in our axial-symmetric projection model.

65 This property was first noted by Wu and Hu in [30]. Although their paper was  
66 focused on central systems, the underlying principle is the same. We provide,  
67 however, proof of its applicability to our model, geometrical insight about the  
68 procedure (the solution is derived from the intersection of conical loci in the  
69 image) and additional techniques to deal with noise.

70 The second step estimates the extrinsic rotation and translation of the cam-  
71 era coordinate frame w.r.t. the world reference frame. The rotation matrix  
72 is completely determined while the translation vector is estimated up to one  
73 unknown component (the  $Z$ -component).

74 The method relies on establishing a linear projection from 3D world points  
75 to a 1D image feature, which is possible given the axial catadioptric geometry.  
76 A similar 3D-1D linear mapping was used by Thirthala and Pollefeys [31] in a  
77 self-calibration framework. Although it does not rely on knowledge of the scene  
78 structure, their method requires at least 15 point correspondences in 4 views  
79 (for non-central cameras).

80 We show how the Direct-Linear-Transformation (DLT) algorithm [32] can  
81 be used to recover the extrinsic parameters from a set of world-to-image point  
82 correspondences. No knowledge of the mirror shape (besides the axial symme-  
83 try) is needed at this stage. We assume, however, that the pinhole camera is  
84 internally calibrated (a common assumption, e.g. [33, 14, 18, 34]).

85 The third and final step estimates the remaining calibration parameters:  
86 the distance from camera to mirror along the symmetry axis and the undeter-  
87 mined component of the extrinsic translation. It takes into account the com-  
88 plete (non-linear) projection geometry of the system and depends on the mirror  
89 shape, which is assumed to be known a priori. The procedure relies on non-  
90 linear optimization methods (e.g. Levenberg-Marquardt algorithm). Non-linear  
91 optimization and bundle-adjustment are recurring techniques in these types of  
92 systems (e.g. [23, 22, 17, 21]), but we perform the optimization in a single pa-  
93 rameter and show that a precise initial estimate is not required for convergence.

94 *1.3. Notation*

95 Some background concepts regarding cross-ratio and vector representation of  
96 conic curves are briefly reviewed in Appendix A. The notation used throughout  
97 the paper is now introduced.

98 Vectors are denoted by bold symbols. Homogeneous coordinates of points  
99 in  $\mathbb{P}^3$  are represented in upper-case bold symbols (e.g.  $\mathbf{X}$ ), points in  $\mathbb{P}^2$  are in  
100 lower-case (e.g.  $\mathbf{x}$ ) and points in  $\mathbb{P}^1$  are represent in lower-case with an overbar  
101 (e.g.  $\bar{\mathbf{x}}$ ). A tilded symbol denotes an inhomogeneous vector (e.g.  $\tilde{\mathbf{X}}$ ).

102 Matrices are represented by symbols in sans serif font (e.g.  $\mathbf{R}$ ). The super-  
103 script “ $r_i$ ” denotes the  $i$ -th row of a matrix, as in  $\mathbf{R}^{r_i}$ .

104 Equality of matrices or vectors up to a scalar factor is written as “ $\sim$ ”.

105 *1.4. Paper structure*

106 The following sections are organized as follows. Section 2 discusses the  
107 system geometry assumed by our method and deduces the linear projection  
108 equation that can be established from 3D world points to 1D image features.  
109 Section 3 describes the estimation of the *vertex point* (the intersection point  
110 between the image plane and the the mirror axis), which is the first calibra-  
111 tion parameter to be obtained. Section 4 shows how to estimate the extrinsic  
112 parameters, up to one unknown translation component, using a linear method  
113 based on the DLT algorithm. The estimation of the two remaining parameters,  
114 the distance between camera and mirror and the unknown translation compo-  
115 nent, is addressed in section 5. Experimental results are presented in section 6.  
116 Finally, section 7 presents the conclusions.

117 **2. System geometry**

118 We will now discuss the axial catadioptric geometry, and show how a linear  
119 projection equation can be established.

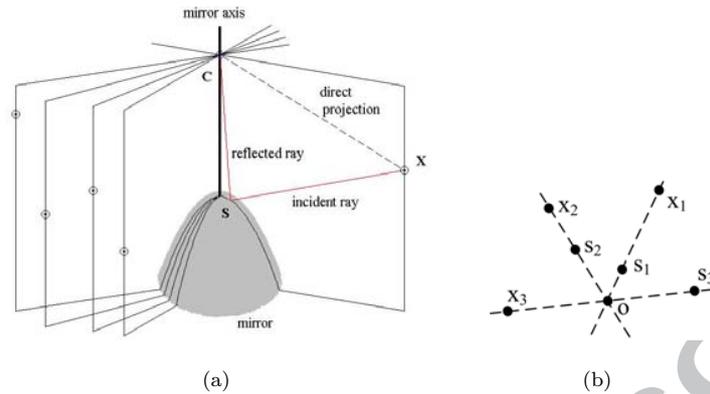


Figure 1: The axial catadioptric geometry. Fig.(a): The pencil of *projection planes* has the mirror axis as the common intersection line.  $\mathbf{C}$  is the camera's optical center.  $\mathbf{X}$  is a world point.  $\mathbf{S}$  is the reflection point on the surface of the mirror. Fig.(b): The pencil of *projection planes* is imaged as a pencil of lines, with  $\mathbf{o}$ , the image of the mirror axis, as the common point. For a given world point  $\mathbf{X}_i$ , there is a line in the image passing through its reflected image  $\mathbf{s}_i$ , its direct image  $\mathbf{x}_i$ , and the image of the axis  $\mathbf{o}$ . Note that the direct image of a point,  $\mathbf{x}_i$ , may not be available in practice. As shown for point  $\mathbf{x}_3$ , if the world point is behind the camera (i.e., negative coordinate on the camera's principal axis), its direct and reflected images have opposite directions w.r.t. the central point  $\mathbf{o}$ .

### 120 2.1. The pencil of projection planes and its image

121 Consider Fig. 1(a). Let  $\mathbf{C}$  be the camera's optical center and  $\mathbf{X}$  a point in  
 122 the world. An incident ray from  $\mathbf{X}$  intersects the mirror's surface at point  $\mathbf{S}$   
 123 and is reflected to the camera, forming the reflected image of the world point,  
 124 denoted by  $\mathbf{s}$ .

125 From the laws of reflection, we know that the incident ray, the reflected ray  
 126 and the surface's normal at point  $\mathbf{S}$  must belong to the same plane. Also in  
 127 this plane is the direct projection ray, i.e., the projective line, from  $\mathbf{X}$  to  $\mathbf{C}$ ,  
 128 that forms the real (not reflected) image of  $\mathbf{X}$ , denoted by  $\mathbf{x}$ . We refer to this  
 129 plane as a *projection plane*, in the sense that it contains the direct and reflected  
 130 projection rays of a given point in space.

131 As a consequence of the previous assumptions made on system geometry,  
 132 every *projection plane* is part of a pencil of planes, with the mirror axis as the

133 common intersection line. Furthermore, this pencil of planes is projected in the  
 134 image plane as a pencil of lines, where the common point,  $\mathbf{o}$ , is the image of  
 135 the axis (see Fig. 1(b)). For every world point  $\mathbf{X}$ , there is a line in the image  
 136 passing through its reflected image  $\mathbf{s}$ , its direct image  $\mathbf{x}$ , and the image of the  
 137 axis  $\mathbf{o}$  (the *vertex point*).

138 Changes in camera orientation (i.e., rotation around the optical center) in-  
 139 duce homographic transformations in the image (c.f. [32]) and, of course, do  
 140 not affect the collinearity between  $\mathbf{s}$ ,  $\mathbf{x}$  and  $\mathbf{o}$ . Thus, the pencil of *projection*  
 141 *planes* are always imaged as a pencil of lines, as long as the center of the camera  
 142 is placed on the mirror axis. In the particular configuration where the camera's  
 143 principal axis is coincident with the mirror axis (which is of great practical in-  
 144 terest, e.g., in central catadioptric systems), point  $\mathbf{o}$  becomes the principal point  
 145 of the image.

146 It should be noted that the direct image of a point,  $\mathbf{x}$ , is in most practical  
 147 situations not visible in the image, because it is out of the field-of-view or behind  
 148 the camera. This fact does not change, obviously, the validity of the discussion.  
 149 In the algorithms we present in this paper, the position of  $\mathbf{x}$  is always assumed  
 150 to be unknown.

## 151 2.2. Parameterizing the line pencil

152 Now, let  $\mathbf{x} \sim [x \ y \ 1]^T$  and  $\mathbf{s} \sim [s_x \ s_y \ 1]^T$  be the direct and reflected  
 153 image of  $\mathbf{X}$ , respectively, and  $\mathbf{o} \sim [o_x \ o_y \ 1]^T$  be the vertex of pencil. Each  
 154 line on the pencil can be specified by a single parameter, that we will define to  
 155 be the line slope. Thus, the line containing point  $\mathbf{x}$  and passing through the  
 156 vertex  $\mathbf{o}$ , is specified by the slope  $\frac{x-o_x}{y-o_y}$ .

We define the 1D homogenous vector

$$\bar{\mathbf{x}} \sim \begin{bmatrix} x - o_x \\ y - o_y \end{bmatrix} \sim \begin{bmatrix} \frac{x-o_x}{y-o_y} \\ 1 \end{bmatrix}.$$

157 as the *reduced coordinates* of point  $\mathbf{x}$ . Vector  $\bar{\mathbf{x}}$  uniquely specifies the line in the  
 158 pencil that  $\mathbf{x}$  belongs to. Note that, because  $\bar{\mathbf{x}}$  is an homogenous vector, infinite  
 159 slopes can be handled seamlessly.

Since  $\mathbf{s}$  and  $\mathbf{x}$  belong to the same line of the pencil, we have

$$\bar{\mathbf{x}} \sim \bar{\mathbf{s}} \sim \begin{bmatrix} s_x - o_x \\ s_y - o_y \end{bmatrix}. \quad (1)$$

160 *2.3. Linear mapping between  $\mathbf{X}$  and  $\bar{\mathbf{s}}$*

The direct image of world point  $\mathbf{X}$  is given by the projection equation

$$\mathbf{x} \sim \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \mathbf{X},$$

161 where  $\mathbf{K}$  is the intrinsic parameter matrix, and  $\mathbf{R}$  and  $\mathbf{T}$  are the extrinsic rotation  
162 and translation relating the world reference frame with the camera frame.

Using equation 1 we can rewrite the projection equation as

$$\bar{\mathbf{s}} \sim \underbrace{\begin{bmatrix} 1 & 0 & -o_x \\ 0 & 1 & -o_y \end{bmatrix} \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \mathbf{X}}_{\sim \mathbf{P}}. \quad (2)$$

163 The  $2 \times 4$  matrix  $\mathbf{P}$  establishes a linear mapping between points in the world  
164 reference frame and a 1D image parameter computed from the image position  
165 of the reflected points.

166 Given enough known correspondences between  $\mathbf{X}$  and  $\mathbf{s}$ , matrix  $\mathbf{P}$  can be  
167 obtained up to scale, from equation 2, by using the DLT algorithm (Direct  
168 Linear Transform) [32]. We note that in the case that all world points  $\mathbf{X}$  lie  
169 in a single plane, the size of the recovered matrix  $\mathbf{P}$  is reduced to  $2 \times 3$ . This  
170 particular case will be addressed in Section 4.4.

171 **3. Finding the *vertex point***

172 In this section we show how the cross-ratio can be used as an invariant under  
173 the axial catadioptric geometry to obtain the image of the mirror axis, the *vertex*  
174 *point*  $\mathbf{o}$ . By determining its location, the axis direction w.r.t. the camera frame  
175 is immediately defined (assuming an internally calibrated camera).

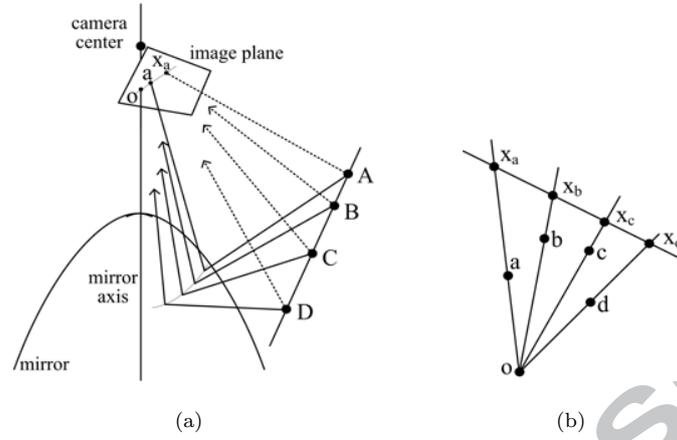


Figure 2: The cross-ratio as an invariant under the axial catadioptric geometry. Fig.(a):  $A$ ,  $B$ ,  $C$  and  $D$  are four collinear 3D points.  $a$ ,  $b$ ,  $c$  and  $d$  are their images after reflection from the mirror.  $x_a$ ,  $x_b$ ,  $x_c$  and  $x_d$  are their direct images, i.e., the direct projection in image. Fig.(b): The cross-ratio relation between image points. Point  $o$  is the image of the mirror axis.

### 176 3.1. Cross-ratio as an invariant

177 Consider Fig. 2(a). Let  $A$ ,  $B$ ,  $C$  and  $D$  be four collinear 3D points. Consider  
 178  $a$ ,  $b$ ,  $c$  and  $d$  to be their reflected images and  $x_a$ ,  $x_b$ ,  $x_c$  and  $x_d$  their direct  
 179 images (i.e., the direct projection in image, not reflected through the mirror).

Fig. 2(b) shows points in the image plane. Being the projection of collinear 3D points,  $x_a$ ,  $x_b$ ,  $x_c$  and  $x_d$  are also collinear. Since the cross-ratio is invariant under a projective transformation,

$$\{x_a x_b x_c x_d\} = \{ABCD\} .$$

Each pair of reflected and direct images of a point (e.g.,  $a$  and  $x_a$ ) is on a line that passes through the image of the mirror axis,  $o$ , so we can write

$$\{o; abcd\} = \{x_a x_b x_c x_d\} = \{ABCD\} . \quad (3)$$

180 We see, thus, that the cross-ratio of four collinear space points is the same as  
 181 the cross-ratio of the lines through their reflected images and the common point  
 182  $o$ , which is the image of the mirror axis.

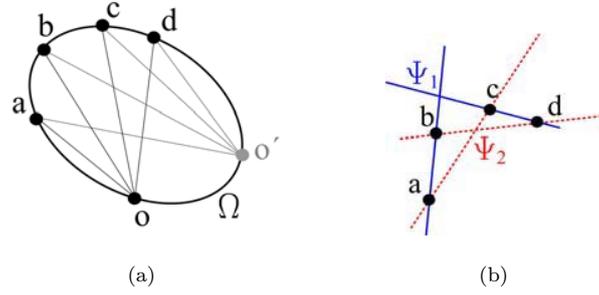


Figure 3: The conic locus of possible solutions for point  $o$ . Fig.(a): The conic  $\Omega$  is completely defined by four image points ( $a, b, c, d$ ) and the value of the cross-ratio,  $k$ . Chasles' theorem states that  $\{o; abcd\} = \{o'; abcd\} = k$ . Fig.(b): The degenerate conics  $\Psi_1$  and  $\Psi_2$  are defined by line-pairs passing through the image points  $a, b, c$  and  $d$ . The solid blue lines define  $\Psi_1$ , while the dashed red lines define  $\Psi_2$ . The conic locus  $\Omega$  is a function of  $\Psi_1, \Psi_2$  and  $k$ .

### 183 3.2. Conic locus for point $o$

Assume that the cross-ratio of a 4-tuple of collinear world points is known,  $k = \{ABCD\}$ . Given the reflected images of these points,  $a, b, c$  and  $d$ , the location of point  $o$  is restricted by (review equation 3):

$$\{o; abcd\} = k. \quad (4)$$

184 We can see that, as a direct application of Chasles' theorem [35], equation 4  
 185 defines a conic locus of possible solutions for  $o$  (see Fig. 3(a)). It should be  
 186 noted that the conic is completely defined by the four points,  $a, b, c$  and  $d$   
 187 (belonging to the conic), and the value of the cross-ratio,  $k$ .

We now show how to obtain the expression of the conic. Consider Fig. 3(b). Let  $\Psi_1$  and  $\Psi_2$  be degenerate conics, defined by the line pairs  $(l_1, m_1)$  and  $(l_2, m_2)$ , respectively, where

$$l_1 = a \times c, \quad m_1 = d \times b, \quad l_2 = a \times b, \quad m_2 = c \times d$$

and with the conics given (in matrix form) by

$$\Psi_i = l_i m_i^T + m_i l_i^T, \quad i = 1, 2$$

It can be verified that the conic locus of point  $\mathbf{o}$  can be obtained from these degenerate conics and the cross-ratio by the expression<sup>1</sup>:

$$\Omega \sim k \Psi_1 - \Psi_2 \quad (5)$$

188 As an additional insight, the conic  $\Omega$  in equation 5 can be viewed as a 1-  
189 parameter family of conics (passing through 4 fixed points,  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$  and  $\mathbf{d}$ ), with  
190 that parameter being  $k$ , the desired value for the cross-ratio.

### 191 3.3. Obtaining a solution

192 Given enough 4-tuples of points in the scene with known cross-ratio, a unique  
193 solution for  $\mathbf{o}$  can be found, corresponding to the common intersection point of  
194 all the conic loci. The minimum number of sets of points required to obtain  
195 a single solution depends on their location and on the number of intersection  
196 points between the conics (as two conics can intersect in up to 4 points). Assum-  
197 ing general position, three sets of points will normally be sufficient to produce  
198 a single solution.

In the presence of noise, however, a common intersection point for the conics may not exist. We can, thus, obtain an estimate for  $\mathbf{o}$  using the following procedure: Let  $\omega_i$  be the vector representation (review equation A.2) of conic  $\Omega_i$ , corresponding to the  $i$ -th 4-tuple of image points with known cross-ratio. Construct a matrix  $\mathbf{Q}$  by stacking the conics  $\omega_i$  for all  $N$  sets of tuples:

$$\mathbf{Q} = \begin{bmatrix} \omega_1^T \\ \vdots \\ \omega_N^T \end{bmatrix}$$

199 Without noise, the right null space of  $\mathbf{Q}$  is the solution for  $\mathbf{o}$ , i.e.,  $\mathbf{Q}\hat{\mathbf{o}} = 0$ . The  
200 estimate for  $\mathbf{o}$  can, thus, be obtained by picking the eigenvector corresponding

---

<sup>1</sup>This expression is valid for a cross-ratio calculated using the formula in A.1. Alternative formulas for the cross-ratio produce different combinations of points in the expressions of  $l_i$  and  $m_i$



Figure 4: Finding the *vertex point*  $\mathbf{o}$ . Fig. (a) and (b) show test images of grid patterns reflected on a spherical mirror. Several 4-tuples of image points and their corresponding conics  $\omega$  are marked in the images. The tuples of image points correspond to equally-spaced collinear world points (cross-ratio=1/4). The intersection point of all conics is the *vertex point*  $\mathbf{o}$ , indicated by with a red arrow. Note that point  $\mathbf{o}$  corresponds to the reflection of the camera's optical center because its projection ray coincides with the mirror axis.

201 to the smallest singular value associated with matrix  $\mathbf{Q}$ . At least  $N = 6$  tuples  
 202 are required for building  $\mathbf{Q}$ .

203 Fig. 4 shows examples, using real images, of conics generated from 4-tuples  
 204 of image points and how the common intersection point is the *vertex point*  $\mathbf{o}$ .

### 205 3.4. Refining the estimate

206 If an intersection point does not exist due to noise, the estimate for vector  
 207  $\hat{\mathbf{o}}$  will not belong to the subspace of lifted coordinates (equation A.3) and the  
 208 extracted vertex  $\mathbf{o}$  will be only an approximation. Furthermore, we have found  
 209 that the cross-ratio conics  $\omega$  show a relatively high sensitivity to noise, which  
 210 degrades the accuracy of the estimate of the *vertex point*. Fig. 5 quantifies this  
 211 sensitivity. It plots the distance between the cross-ratio conic  $\omega$ , obtained from  
 212 image points corrupted with noise, and the ground truth point  $\mathbf{o}$ . Since point  
 213  $\mathbf{o}$  should belong to the conic, the distance provides an error measurement.

214 To improve the accuracy of the estimation of the *vertex point*, we propose an  
 215 additional refinement procedure using a non-linear optimization method. The  
 216 computation of the *reduced coordinates*  $\bar{\mathbf{s}}$  of a given image point is a function of

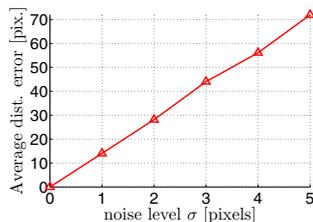


Figure 5: Sensitivity to noise of the cross-ratio conics  $\omega$ . Plot of the distance between the conic curve  $\omega$  and the ground truth point  $\mathbf{o}$ , as a function of noise level  $\sigma$  ( $\sigma$  is the standard deviation of the gaussian noise added to the position of image points). Point  $\mathbf{o}$  should belong to the conic, so the distance to the conic curve provides an error measurement. The figure was obtained by simulation (we used the same simulation setups that are described in detail in subsection 6.1). The results were obtained by averaging repeated simulations, using sets of points in different positions. In total, the result for each noise level was obtained from 2000 simulations.

217 point  $\mathbf{o}$ , i.e.,  $\bar{\mathbf{s}}(\mathbf{o})$ . Let  $\{\bar{\mathbf{s}}_i(\mathbf{o}); \mathbf{X}_i\}$ ,  $i = 1, \dots, N$ , denote the set of  $N$  world-to-  
 218 image point correspondences. As stated in the previous section (review equa-  
 219 tion 2), a linear mapping can be estimated from the set of correspondences using  
 220 the DLT algorithm.

221 Consider a function  $SSV_{DLT}(\{\bar{\mathbf{s}}_i(\mathbf{o}); \mathbf{X}_i\})$  that returns the smallest singu-  
 222 lar value obtained during the Singular-Value-Decomposition factorization of the  
 223 DLT procedure. The closer to zero the value is, the better the linear mapping  
 224 fits the set of points. Thus, function  $SSV_{DLT}$  can be used to evaluate a candi-  
 225 date point  $\mathbf{o}$ , quantifying how the estimates for the coordinates of that point fit  
 226 into the linear projection model.

Starting at the initial solution obtain in the previous subsection, we can refine the estimate for point  $\mathbf{o}$  by apply non-linear optimization to

$$\min_{\mathbf{o}} SSV_{DLT}(\{\bar{\mathbf{s}}_i(\mathbf{o}); \mathbf{X}_i\}) . \quad (6)$$

227 In our implementation we used the Levenberg-Marquardt method. Addi-  
 228 tionally, we used the RANSAC algorithm [32] to handle outliers on the set of  
 229 point correspondences  $\{\bar{\mathbf{s}}_i(\mathbf{o}); \mathbf{X}_i\}$ .

230 Fig. 6 shows an example, with a real image, of the output of function

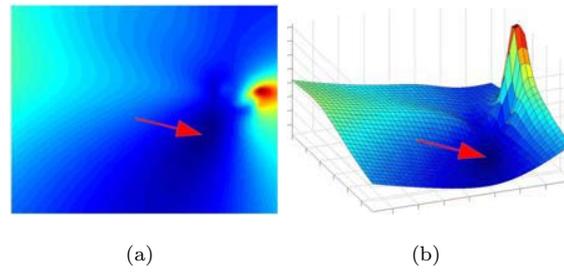


Figure 6: The function  $SSV_{DLT}$ . Fig. (a) shows the output of function  $SSV_{DLT}$  evaluated at every pixel of the test image shown in Fig. 4(a). Fig. (b) shows the same surface but in a 3D perspective. A blue hue represents lower values on the surface, while a red hue represents higher values. Point  $\mathbf{o}$  is located at the global minimum of the surface, indicated by the red arrow.

231  $SSV_{DLT}$ .

#### 232 4. Estimating the extrinsic parameters using linear methods

233 In this section we show how the extrinsic parameters can be obtained, up to  
 234 one undetermined component of the translation vector, from a linear method  
 235 and using a single image of a calibration object. We first consider a generic 3D  
 236 calibration object, but then adapt the algorithm to handle the case, of practical  
 237 interest, when all the calibration points belong to a single plane.

238 We assume that the position of the *vertex point*  $\mathbf{o}$  (discussed in the last  
 239 section) has already been determined, and that the pinhole camera is internally  
 240 calibrated. In most cases, the camera can be previously calibrated (internally),  
 241 without the mirror, using standard methods [36, 37].

##### 242 4.1. Pre-alignment of the camera frame

243 To derive the method to estimate the extrinsic parameters we assume that  
 244 the camera is aligned with the mirror, i.e., the camera's principal axis coincides  
 245 with the symmetry axis of the mirror, with the camera pointing towards the  
 246 mirror. This assumption does not imply a loss of generality since a pre-rotation  
 247 can always be performed to align the camera axis.

248 Given an internally calibrated camera, the knowledge of point  $\mathbf{o}$  provides,  
 249 implicitly, the direction of the mirror axis in the camera reference frame. A  
 250 rotation can then be calculated that would align the principal axis with that  
 251 direction.

The alignment rotation is implemented by an homographic transformation  
 in the image. This homography is called a *conjugate rotation* (c.f. [32], section  
 8 and appendix A7) and is given by

$$\mathbf{H} = \mathbf{K}\mathbf{R}\mathbf{K}^{-1},$$

252 where  $\mathbf{K}$  is the intrinsic parameters matrix and  $\mathbf{R}$  is the rotation matrix. All im-  
 253 age points are transformed from their original positions into the aligned camera  
 254 frame using the homography  $\mathbf{H}$ . It should be noted that after the alignment the  
 255 *vertex point*  $\mathbf{o}$  is moved to the image center, i.e.,  $\mathbf{o} \sim \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ . In subsequent  
 256 sections, any reference to an image point ( $\mathbf{s}$ ) assumes an aligned camera.

257 In many applications (e.g., central systems) the camera is in fact aligned  
 258 with the mirror, and this initial step is unnecessary.

#### 259 4.2. The projection matrix $\mathbf{P}$

260 Please recall that a point in the world reference frame is denoted by  $\mathbf{X}$ .  
 261 Point  $\mathbf{X}$  has known position (belongs to the calibration object). Its projection  
 262 in the image after reflection from the mirror is denoted by point  $\mathbf{s}$ . Consider  
 263  $\mathbf{T} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T$  to be the extrinsic translation vector and let  $R^{r_i}$  denote the  
 264  $i$ -th row of the extrinsic rotation matrix  $\mathbf{R}$ .

Assuming that the camera is internally calibrated ( $\mathbf{K} = \mathbf{I}$ ) and that the  
 camera frame is aligned with the mirror axis ( $\mathbf{o} \sim \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ ), the  $2 \times 4$   
 projection matrix of equation 2 is simplified to

$$\bar{\mathbf{s}} \sim \underbrace{\begin{bmatrix} R^{r_1} & t_x \\ R^{r_2} & t_y \end{bmatrix}}_{\sim \mathbf{P}} \mathbf{X}. \quad (7)$$

265 4.3. Retrieving  $R$ ,  $t_x$  and  $t_y$

266 As previously discussed, given enough known correspondences between  $\mathbf{X}$   
 267 and  $\mathbf{s}$ , matrix  $\mathbf{P}$  can be obtain up to scale, from equation 7, by using the DLT  
 268 algorithm. It should be noted that, for the moment, we are considering a gen-  
 269 eral non-planar calibration object. The case of a planar calibration pattern is  
 270 analyzed in the next subsection.

Let  $p_{ij}$  denote the element of  $\mathbf{P}$  at row  $i$  and column  $j$ . Noting that  $\mathbf{P}$  is determined only up to a scale factor  $\lambda$ , the extrinsic parameters, with the exception of  $t_z$ , can be recovered from

$$\mathbf{R}^{r1} = \lambda \begin{bmatrix} p_{11} & p_{12} & p_{13} \end{bmatrix} \quad (8)$$

$$\mathbf{R}^{r2} = \lambda \begin{bmatrix} p_{21} & p_{22} & p_{23} \end{bmatrix}$$

$$\mathbf{R}^{r3} = \mathbf{R}^{r1} \times \mathbf{R}^{r2}$$

$$t_x = \lambda p_{14}$$

$$t_y = \lambda p_{24}$$

As  $\mathbf{R}^{r1}$  and  $\mathbf{R}^{r2}$  are normal vectors, the value of  $\lambda$  is subjected to the constraint

$$\|\lambda \begin{bmatrix} p_{11} & p_{12} & p_{13} \end{bmatrix}\| = \|\lambda \begin{bmatrix} p_{21} & p_{22} & p_{23} \end{bmatrix}\| = 1,$$

which yields

$$\lambda = \pm \frac{1}{\|\begin{bmatrix} p_{11} & p_{12} & p_{13} \end{bmatrix}\|} = \pm \frac{1}{\|\begin{bmatrix} p_{21} & p_{22} & p_{23} \end{bmatrix}\|}. \quad (9)$$

The signal ambiguity of  $\lambda$  can be solved by means of a simple procedure, taking into consideration the geometric properties of image formation. Consider  $(X_c, Y_c, Z_c)$  as the coordinates of  $\mathbf{X}$  in the camera frame. We have that

$$\begin{bmatrix} X_c \\ Y_c \end{bmatrix} = \begin{bmatrix} \mathbf{R}^{r1} & t_x \\ \mathbf{R}^{r2} & t_y \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (10)$$

271 where point  $(X_c, Y_c)$  can be seen as the orthogonal projection of  $(X_c, Y_c, Z_c)$  in  
 272 the image plane.

273 Since we are considering an aligned camera frame, the image plane is per-  
 274 pendicular to the *projection planes*, and point  $(X_c, Y_c)$  and the corresponding  
 275 reflected image point  $(s_x, s_y)$  are on a line that passes through the image ori-  
 276 gin (see section 2). More so, in the presence of a convex mirror<sup>2</sup>,  $(X_c, Y_c)$  and  
 277  $(s_x, s_y)$  have the same direction w.r.t. the image origin. In other words, vectors  
 278  $(X_c, Y_c)$  and  $(s_x, s_y)$  must have the same orientation and direction.

279 The correct value for  $\lambda$  can, thus, be obtained using the following procedure:

- 280 1. Choose one known pair of correspondences  $\mathbf{X}$  and  $\mathbf{s}$ ;
- 281 2. For both solutions of equation 9,  $+\lambda$  and  $-\lambda$ :
  - 282 • Compute  $R$ ,  $t_x$  and  $t_y$  using (8);
  - 283 • Compute  $(X_c, Y_c)$  using (10);
- 284 3. From the two opposing vectors resulting from step 2,  $(\pm X_c, \pm Y_c)$ , choose  
 285 the one pointing in the same direction as  $(s_x, s_y)$  (in the presence of noise,  
 286 choose the closest direction). The value of  $\lambda$  that corresponds to the  
 287 correct vector is the solution.

288 It should be noted that, in the presence of noise, the recovered matrix  $R$  may  
 289 not be a true rotation matrix. Using Singular-Value-Decomposition,  $R = U\Sigma V^T$ ,  
 290  $R$  can be projected to a matrix  $R'$  in orthonormal space by substituting all the  
 291 singular values by 1, i.e.,  $R' = UV^T$ . Matrix  $R'$  is the closest orthonormal matrix  
 292 to  $R$  in the sense that it minimizes the Frobenius norm  $\|R - R'\|_F$ .

#### 293 4.4. Planar calibration pattern

294 We now show how the algorithm can be changed in order to allow for a  
 295 planar calibration object.

---

<sup>2</sup>In the axial geometry we are considering, when the camera is pointing at a convex mirror, the reflection is seen on the same direction (or “side”) as the object is in the world. For a concave mirror, the opposite is true. In this algorithm we assume the convex case because of its far greater practical interest.

We will assume, without loss of generality, that the calibration points belong to plane  $Z = 0$  w.r.t. the world frame (in a similar manner as in [36]). Equation 7 becomes

$$\bar{\mathbf{s}} \sim \underbrace{\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \end{bmatrix}}_{\sim \mathbf{P}} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad (11)$$

where  $r_{ij}$  denotes the element of matrix  $\mathbf{R}$  at row  $i$  and column  $j$ . With some abuse of notation, let us redefine  $\mathbf{P}$  to be the  $2 \times 3$  matrix mapping the planar world points to the 1D image feature.

Matrix  $\mathbf{P}$  is, again, obtained up to a scale factor  $\lambda$  using the DLT algorithm. Similarly to equation 8, we have that

$$\begin{aligned} \mathbf{R}^{r_1} &= \lambda \begin{bmatrix} p_{11} & p_{12} & a \end{bmatrix} \\ \mathbf{R}^{r_2} &= \lambda \begin{bmatrix} p_{21} & p_{22} & b \end{bmatrix} \end{aligned} \quad (12)$$

with  $\lambda$ ,  $a$  and  $b$  to be determined.

Since  $\mathbf{R}^{r_1}$  and  $\mathbf{R}^{r_2}$  are orthonormal we can write

$$\begin{bmatrix} p_{11} & p_{12} & a \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & a \end{bmatrix}^T = \begin{bmatrix} p_{21} & p_{22} & b \end{bmatrix} \begin{bmatrix} p_{21} & p_{22} & b \end{bmatrix}^T$$

and

$$\begin{bmatrix} p_{11} & p_{12} & a \end{bmatrix} \begin{bmatrix} p_{21} & p_{22} & b \end{bmatrix}^T = 0.$$

It can be shown that these constraints generate 2 real solutions for  $a$  and  $b$ . The solutions are symmetric and will be denoted as  $\{a_+; b_+\}$  and  $\{a_-; b_-\}$ , where

$$a_{\pm} = \pm \frac{(k_{\alpha} - k_{\gamma})}{2k_{\beta}} \sqrt{\frac{k_{\alpha} + k_{\gamma}}{2}}; \quad b_{\pm} = \pm \sqrt{\frac{k_{\alpha} + k_{\gamma}}{2}}$$

with

$$k_{\alpha} = r_{11}^2 + r_{12}^2 - r_{21}^2 - r_{22}^2$$

$$k_{\beta} = r_{11}r_{21} + r_{12}r_{22}$$

$$k_{\gamma} = \sqrt{k_{\alpha}^2 + 4k_{\beta}^2}$$

300 The unknown scale factor  $\lambda$  is determined using equation 9, where variables  
 301  $p_{13}$  and  $p_{23}$  are substituted, respectively, by  $a_+$  and  $b_+$  (or by  $a_-$  and  $b_-$ ,  
 302 yielding the same result). The signal ambiguity of  $\lambda$  can, again, be solved with  
 303 the procedure described in the previous subsection. It should be noted that in  
 304 equation 10 we now have  $Z = 0$ , which causes the equation to be independent  
 305 of the values of  $a$  and  $b$ , and so  $\lambda$  is still uniquely determined.

306 Two solutions are, thus, possible for the extrinsic rotation matrix  $\mathbf{R}$ , obtained  
 307 by substituting the values  $\{\lambda; a_+; b_+\}$  and  $\{\lambda; a_-; b_-\}$  in equation 12 (the  
 308 procedure to determine the correct solution is discussed in the next subsection).  
 309 The 3rd row of  $\mathbf{R}$  is given by  $\mathbf{R}^{r_3} = \mathbf{R}^{r_1} \times \mathbf{R}^{r_2}$ .

The first two components of the extrinsic translation are determined without  
 ambiguity and can be obtained from

$$t_x = \lambda p_{13}$$

$$t_y = \lambda p_{23}.$$

#### 310 4.5. Discussion and summary

311 Using a 3D (non-planar) calibration object produces an unique solution for  
 312 the extrinsic rotation matrix  $\mathbf{R}$ . Regarding the minimum number of world-to-  
 313 image point correspondences required to apply the DLT algorithm to equation 7,  
 314 it can be seen that each  $\bar{\mathbf{s}} \leftrightarrow \mathbf{X}$  pair establishes two equations up to scale.  
 315 Eliminating the unknown scale factor between them results in one constraint on  
 316 the variables of  $\mathbf{P}$  for every point correspondence. Since the  $2 \times 4$  matrix  $\mathbf{P}$  is  
 317 recovered only up to scale, 7 independent variables need to be determined, which  
 318 means that at least 7 world-to-image correspondences are required. The world  
 319 points can not be located on a single plane (i.e. the calibration object must be  
 320 non-planar), or else one column of matrix  $\mathbf{P}$  is left undetermined (equation 11).  
 321 Furthermore, to determine that column of  $\mathbf{P}$ , at least two off-plane world points  
 322 are needed to constrain the two variables in the column.

323 For a simpler experimental setup, the use of a planar calibration pattern  
 324 is possible. A minimum of 5 point correspondences is needed in this case (a

325 similar reasoning as in the previous case, now with a  $2 \times 3$  matrix  $\mathbf{P}$ ). In this  
 326 situation, two possible solutions are obtained for matrix  $\mathbf{R}$ . This ambiguity can,  
 327 however, be solved by carrying both solutions to the next step in the calibration  
 328 procedure and performing a complete reprojection of the world object into the  
 329 image. The correct solution is the one that produces the image closest to the  
 330 original.

331 The  $t_x$  and  $t_y$  components of the extrinsic translation are unambiguously  
 332 recovered, regardless of the use of a non-planar or planar calibration object.  
 333 The  $t_z$  component is undetermined at this stage. The value of the translation  
 334 vector  $\mathbf{T}$  is, thus, restricted to a line space.

### 335 **5. Estimating remaining parameters using non-linear optimization** 336 **methods**

337 The previous sections described how to obtain most of the parameters related  
 338 to the mirror position and to the extrinsic calibration: Section 3 showed how  
 339 to determine the mirror axis direction w.r.t. the camera while section 4 showed  
 340 how to calculate the complete extrinsic rotation, and the extrinsic translation  
 341 up to one component.

342 In this section we estimate the remaining parameters: the distance  $d$  between  
 343 camera and mirror along the symmetry axis, and the last component of the  
 344 extrinsic translation,  $t_z$ .

345 Previously, we have taken advantage of the axial geometry of the system and  
 346 avoided the use of the non-linear reflections associated with a (possibly) non-  
 347 central catadioptric system. From now on, we take into consideration mirror  
 348 shape and reflection geometry in order to estimate  $d$  and  $t_z$ , using non-linear  
 349 optimization methods. We show, given the previously calculated parameters,  
 350 that the optimization is performed on a single variable.

351 Our method requires the computation of back-projection rays from the cam-  
 352 era and mirror geometry. In Appendix B we briefly outline the procedure for  
 353 a mirror with a conic section. The derivation is based on [21]. We note, how-

354 ever, that any mirror profile is admissible as long as it is known a priori so that  
 355 back-projection rays can be calculated.

356 *5.1. 3D reconstruction from back-projection and partial extrinsics*

Let  $\tilde{\mathbf{X}}_c = [X_c \ Y_c \ Z_c]^\top$  be the inhomogeneous coordinates, in the aligned camera frame, of a known world point  $\tilde{\mathbf{X}}$  belonging to the calibration object. Point  $\tilde{\mathbf{X}}_c$  is obtained from the extrinsic parameters  $\mathbf{R}$  and  $\mathbf{T}$  by

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}} \\ 1 \end{bmatrix} = \begin{bmatrix} R^{r1}\tilde{\mathbf{X}} + t_x \\ R^{r2}\tilde{\mathbf{X}} + t_y \\ R^{r3}\tilde{\mathbf{X}} + t_z \end{bmatrix}. \quad (13)$$

357 Since the parameter  $t_z$  is not yet determined, the position of point  $\tilde{\mathbf{X}}_c$  is  
 358 defined only up to a linear locus in space, which we denote as line  $\mathbf{L}_z$ . The  
 359 line is orthogonal to the image plane and intersects this plane at coordinates  
 360  $(X_c, Y_c)$ .

361 On the other hand, point  $\tilde{\mathbf{X}}_c$  must also belong to the back-projected ray  
 362 obtained from its reflected image  $\mathbf{s}$ . We denote that back-projected ray, after  
 363 reflection on the mirror surface, as space line  $\mathbf{L}_{BP}$ .

364 Consequently, space point  $\tilde{\mathbf{X}}_c$  can be reconstructed by intersecting both 3D  
 365 lines,  $\mathbf{L}_z$  and  $\mathbf{L}_{BP}$ . While line  $\mathbf{L}_z$  is fully defined (it is a function of the already  
 366 estimated  $\mathbf{R}$ ,  $t_x$  and  $t_y$ ), line  $\mathbf{L}_{BP}$  depends on the yet undetermined distance  
 367  $d$  (see Appendix B). It should be noted, however, that despite the fact that  
 368 different values of  $d$  produce distinct back-projection rays, an intersection point  
 369 between  $\mathbf{L}_z$  and  $\mathbf{L}_{BP}$  always exists, as both lines belong to the same *projection*  
 370 *plane* (see section 2).

371 *5.2. Estimating distance to mirror  $d$  and the extrinsic translation parameter  $t_z$*

372 Let  $\{\mathbf{X}_i\}$  and  $\{\mathbf{s}_i\}$ , with  $i = 1..N$ , denote the set of points from the calibra-  
 373 tion object, expressed in the world frame, and their reflected images. Consider,  
 374 also,  $\{\mathbf{X}_{c_i}\}$  to be the set of points from the calibration object expressed in the  
 375 camera frame coordinates. The problem of determining  $d$  can be stated in the

376 following manner: *Given a set of correspondences between world points  $\{\mathbf{X}_i\}$*   
 377 *and image points  $\{\mathbf{s}_i\}$ , and the knowledge of the extrinsic parameters  $\mathbf{R}$  and  $\mathbf{T}$ ,*  
 378 *with the exception of  $t_z$ , find the value of  $d$  that reconstructs the set of points*  
 379  *$\{\mathbf{X}_{c_i}\}$  in such a way that they “fit” the original pattern  $\{\mathbf{X}_i\}$  from the cali-*  
 380 *bration object. The evaluation function is, in general, a measure of how “well”*  
 381  *$\{\mathbf{X}_i\}$  and  $\{\mathbf{X}_{c_i}\}$  can be related by a rigid transformation, as both sets should*  
 382 *represent the same object. Alternatively, other metric characteristics regarding*  
 383 *shape, distances, angles, etc., can be used, depending on the specific geometric*  
 384 *properties of the calibration object.*

385 The well known Iterative Closest Point (ICP) [38] algorithm can be used to  
 386 obtain the rotation and translation that registers the calibration object  $\{\mathbf{X}_i\}$   
 387 to its reconstruction  $\{\mathbf{X}_{c_i}\}$ . The values of this rotation and translation will, of  
 388 course, dependent on  $d$ , and we denote them as  $\mathbf{R}_d$  and  $\mathbf{T}_d$ , respectively.

The distance  $d$  can be obtained by minimizing

$$\min_d \sum_i \|\tilde{\mathbf{X}}_{c_i} - (\mathbf{R}_d \tilde{\mathbf{X}}_i + \mathbf{T}_d)\|^2. \quad (14)$$

389 Once the value of  $d$  that achieves the minimization is found, the last unknown  
 390 parameter  $t_z$  is obtained from the z-component of  $\mathbf{T}_d$ .

391 Since the estimation of  $\mathbf{R}_d$  and  $\mathbf{T}_d$  relies on the ICP procedure, a closed-form  
 392 solution for equation 14 can not be easily obtained. However, standard non-  
 393 linear optimization methods can be used (e.g. Levenberg-Marquardt algorithm).  
 394 We have found that, even in the presence of noise, the minimization achieves  
 395 convergence to the global minimum without an accurate initial estimate of  $d$ . In  
 396 our experiments we considered  $d = \text{focal length}$  (camera touching the mirror)  
 397 as the initial estimate.

398 To provide intuition, Fig. 7 illustrates the idea behind the procedure by  
 399 showing the effect that errors in  $d$  have on the shape of a reconstructed planar  
 400 calibration pattern.

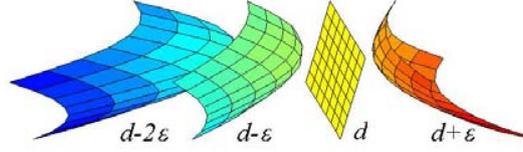


Figure 7: Estimating of the distance to mirror,  $d$ . The goal is to find the value of  $d$  that reconstructs the original calibration object that, in this example, consists on a planar grid. The figure, obtained from simulation, exemplifies how a reconstructed object deviates from the original shape as an error  $\epsilon$  is added to the true value of  $d$ .

## 401 6. Experimental Results

402 We now present experimental results obtained with the proposed method.  
 403 First we show tests with simulated data and then results from real images. We  
 404 also include a comparison with methods designed for central systems.

405 To provide an intuitive representation to the reader, rotation matrices are  
 406 presented as a 3 element vector containing the corresponding Euler angles,  
 407 in degrees. Rotation matrix  $R = R_z(\theta_z)R_y(\theta_y)R_x(\theta_x)$  is represented by  $r =$   
 408  $(\theta_x, \theta_y, \theta_z)$ , where  $R_a(\theta)$  denotes a rotation of angle  $\theta$  along axis  $a = x, y, z$ .

409 We refer to the rotation error in the following terms: given a ground truth  
 410 rotation matrix  $R_{GT}$  and the corresponding noise affected estimate  $R_{est}$ , the  
 411 rotation error matrix  $R_{err}$  is defined as:  $R_{est} = R_{err}R_{GT}$ .

412 Translation errors are quantified in two distinct values: an angle error, cor-  
 413 responding to the angle between the estimated and ground truth vectors, and as  
 414 a length percentage error, given by the ratio  $\|\mathbf{T}_{est} - \mathbf{T}_{GT}\|/\|\mathbf{T}_{GT}\|$ , where  $\mathbf{T}_{est}$   
 415 and  $\mathbf{T}_{GT}$  are the estimated and ground truth translation vectors, respectively,  
 416 and  $\|\cdot\|$  denotes the L2-norm.

### 417 6.1. Simulated data

418 The simulations were run on three distinct setups. Each setup had different  
 419 parameters regarding the mirror shape, mirror position, and pose of the cali-  
 420 bration object. Table 1 summarizes the values of the parameters in each setup.  
 421 The image size of the simulated camera was  $1500 \times 1500$  pixels, with a focal

422 length of 1200 pixels. The calibration pattern consisted of a planar square grid,  
 423 with  $8 \times 8$  points. The distance between adjacent points on the grid was 2 world  
 424 metric units.

Setup	mirror (A,B,C) [w.m.u]	$d$ [w.m.u]	$\mathbf{o}$ [pixels]	$\mathbf{R}$ [Euler ang.]	$\mathbf{T}$ [w.m.u]
#1	spherical (1, 0, 4)	3	(100, 150)	(40°, 100°, 45°)	(-4, 5, -6)
#2	parabolic (0, 1, 1)	4	(100, 150)	(30°, 100°, 0°)	(4, -5, -2)
#3	hyperbolic (-1, 4, -1)	5	(100, 150)	(0°, 60°, 0°)	(6, -5, 2)

Table 1: Simulation setups. The proposed methods were simulated in different setups, each with distinct mirror shape, mirror position ( $d$  and  $\mathbf{o}$ ), and pose of the calibration object ( $\mathbf{R}$  and  $\mathbf{T}$ ). Mirror parameters are defined in equation B.1 of Appendix B. “w.m.u” stands for “world metric units”.

425 Gaussian noise of zero mean and  $\sigma$  standard deviation was added to the  
 426 position of the image points before running the calibration procedure. For a  
 427 given  $\sigma$  value, each of the setups was repeated 100 times and the data compiled  
 428 from the 3 setups, to provide a statistical analysis on the estimation error.  
 429 Fig. 8(a)-(d) shows the root mean square (RMS) error, as a function of the noise  
 430 level  $\sigma$ , in the extrinsic parameters  $\mathbf{R}$  and  $\mathbf{T}$ , and in mirror position parameters  
 431  $d$  and  $\mathbf{o}$ .

432 Fig. 8(e) plots the reprojection error as a function of the noise level. Since  
 433 our method does not rely on direct minimization of the reprojection error (like  
 434 bundle adjustment techniques), this error can be considered as a measure of  
 435 the overall quality of the calibration. Also shown in Fig. 8(e) is the result from  
 436 repeating the simulations assuming that point  $\mathbf{o}$  is known a priori (without  
 437 noise), and estimating only the remaining parameters. This situation is relevant  
 438 in systems where the camera is aligned with the mirror axis, and  $\mathbf{o}$  corresponds  
 439 to (or approximates) the image center. When using spherical mirrors, the *vertex*  
 440 *point* can be estimated from the reflected image of the camera itself (if visible)  
 441 as point  $\mathbf{o}$  corresponds to the reflection of the optical center.

442 In additional simulations we studied the effect of using more than one image  
 443 in the calibration procedure. For each simulation setup, the calibration pattern

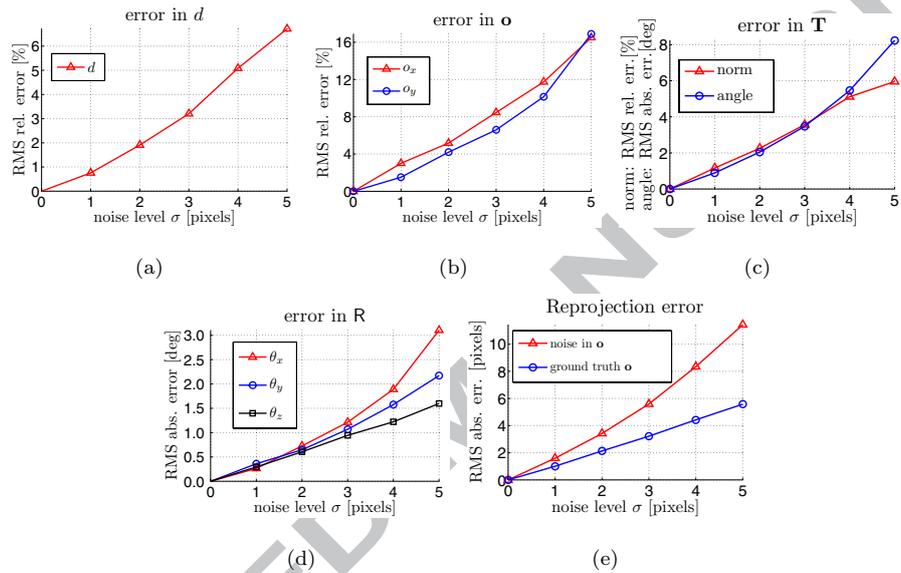


Figure 8: Simulation results. Fig.(a) to (d) show the root mean square (RMS) error in the estimation of the calibration parameters, as a function of noise level  $\sigma$  ( $\sigma$  is the standard deviation of the gaussian noise added to the position of image points).  $\mathbf{o}$  is the *vertex point*;  $d$  is the distance between camera and mirror;  $\mathbf{R}$  and  $\mathbf{T}$  are the extrinsic rotation and translation, respectively. Fig.(e) shows the RMS error in image position obtained from reprojecting the calibration points using the estimated calibration parameters.

444 was rotated around the mirror axis, producing images with different extrinsic  
 445 parameters, but with the camera/mirror relative position kept constant. The  
 446 estimates for the *vertex point*  $\mathbf{o}$  and for the mirror distance  $d$  were computed by  
 447 the minimization of expression 6 (for  $\mathbf{o}$ ) and 14 (for  $d$ ) taking into account all  
 448 images simultaneously. Fig. 9 shows the estimation error as a function of the  
 449 number of images used, for a fixed noise level of  $\sigma = 4$  pixels. It is seen that  
 450 the using multiple images (with different positions of the calibration pattern)  
 451 can help reduce the effect of noise and increase the accuracy in the estimation  
 452 of the mirror relative position.

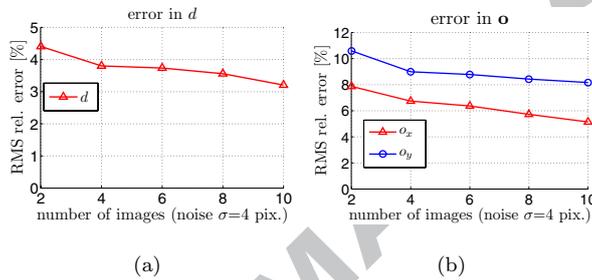


Figure 9: Reduction of the estimation error of the *vertex point*  $\mathbf{o}$  and mirror distance  $d$  by using more than one image of the calibration pattern. The extrinsic parameters change from image to image, but the camera/mirror relative position ( $\mathbf{o}$  and  $d$ ) was kept constant. The results were compiled from repeated simulations with different mirror types, as before. The image noise standard deviation was fixed at  $\sigma = 4$  pixels.

## 453 6.2. Comparison with methods designed for central systems

454 As previously stated, although we focus on non-central catadioptric systems,  
 455 our method can be applied to central systems. Using a simulated setup, we  
 456 applied our technique to a central system and compared its performance with  
 457 two widely used methods from Sacaramuzza et al. [8, 39], and Mei and Rives [9],  
 458 both available as OpenSource toolboxes [40, 41]. The two methods use images  
 459 of a planar calibration object.

460 We simulated a central system with an hyperbolic mirror (parameters [mm]:  
 461  $A = -0.76$ ;  $B = 0$ ;  $C = -600$ ) and a pinhole camera (resolution of  $1000 \times 1000$

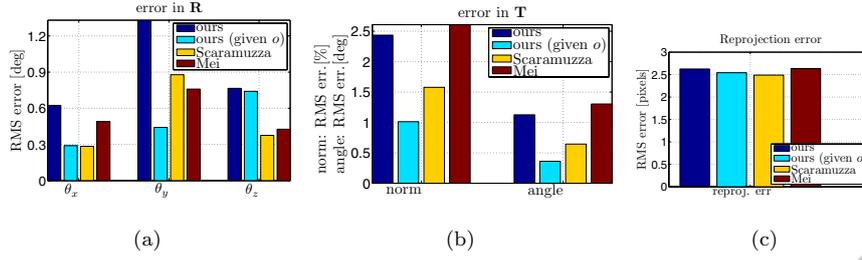


Figure 10: Comparison with methods for central systems. Performance comparison between our method (full calibration and partial calibration assuming known *vertex point*  $\mathbf{o}$ ) and the methods of Sacaramuzza et al., and Mei and Rives, designed for central systems. Fig. (a) and (b) show the error in the estimation of the extrinsic rotation and translation, respectively. Note that  $\theta_x$ ,  $\theta_y$  and  $\theta_z$  are the euler angles of the rotation error (please review the beginning of the section for details). Fig.(c) shows the reprojection error. Results obtained from a simulated setup of a central hypercatadioptric system with 10 calibration images (with added noise). The error values shown were computed from all the images.

462 pixels) placed at the focus of the hyperbola. A  $9 \times 10$  point grid was placed  
 463 in 10 positions around the mirror, generating 10 different calibration images.  
 464 Gaussian noise of zero mean and 2 pixels standard deviation was added to  
 465 the image position of each point. The toolboxes were modified to bypass any  
 466 imaging processing and to use the simulated image points instead.

467 We applied our method in two distinct conditions. First with a complete  
 468 calibration, and then assuming that the *vertex point*  $\mathbf{o}$  was known a priori, and  
 469 only estimating the remaining parameters. In a central system the camera is  
 470 aligned with the mirror and point  $\mathbf{o}$  corresponds to the image center.

471 The results are presented in Fig. 10. Since our methods assumes a calibrated  
 472 pinhole camera, we only compare the estimation of the extrinsic parameters,  $\mathbf{R}$   
 473 and  $\mathbf{T}$ , and the reprojection error. The values presented are the RMS errors  
 474 obtained from the set of the 10 images.

### 475 6.3. Experiments with real images

476 We now present results obtained with real images. The experiments were  
 477 setup as follows. The projective camera was previously (internally) calibrated  
 478 using standard methods [37]. Two different mirrors were used, one spherical and

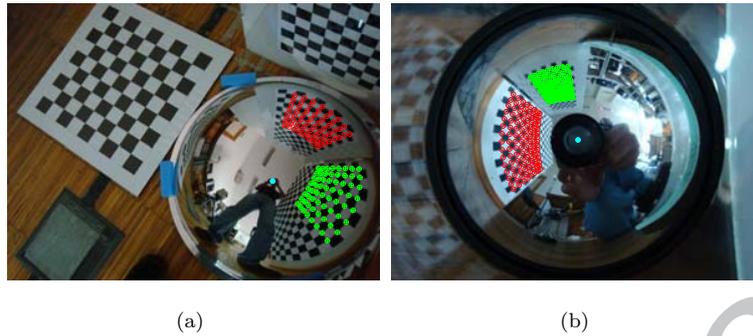


Figure 11: Test images obtained with a spherical and an hyperbolic mirror, shown in Fig.(a) and Fig.(b), respectively. Two separate planar calibration grids are seen reflected in each mirror. The calibration points used in the experiments are highlighted in the images: points in grid 1 are marked with a red “ $\odot$ ”; points in grid 2 are marked with a green “ $\oplus$ ”. In Fig.(a) (spherical mirror),  $8 \times 8$  points were used in each grid. In Fig.(b) (hyperbolic mirror),  $15 \times 8$  points were used in grid 1 and  $9 \times 10$  points were used in grid 2. The *vertex point*  $\circ$  is marked in each image with a cyan “\*”.

479 one hyperbolic. An image containing two distinct planar calibration patterns  
 480 was acquired for each mirror. We applied our method to each pattern separately,  
 481 obtaining two independent results for each setup. Fig. 11 shows the test images  
 482 acquired with both mirrors, and the calibration points used in each grid pattern.  
 483 Each image has a resolution of  $1600 \times 1200$  pixels.

484 To compare and evaluate the output of our algorithm, reference values for the  
 485 calibration parameters were obtained independently, from direct measurement  
 486 and from image analysis, using Bouguet’s camera calibration toolbox [37]. Each  
 487 mirror was aligned with a third, auxiliary, grid pattern. The relative pose  
 488 between the auxiliary patterns and the mirrors was calculated from the grid  
 489 alignment and by direct measurement. Then, from an external projective image  
 490 (capturing all the grids) the transformations between the mirror frame and the  
 491 calibration grids were extracted using the toolbox.

492 In the spherical mirror setup, the camera was placed so that the auxiliary  
 493 mirror grid was directly visible in the test image (alongside the mirror itself),  
 494 and the camera/mirror pose was computed, again using [37]. In the hyperbolic

495 mirror setup, due to the small mirror size and camera alignment, the auxil-  
 496 iary mirror grid was not directly visible in the image. In this case, we relied  
 497 on careful camera placement and measurement to estimate the camera/mirror  
 498 transformation.

499 Combining the camera/mirror relative pose with the information from the  
 500 external image, the geometry of the scene was fully reconstructed for each setup,  
 501 and reference values for the extrinsic parameters ( $\mathbf{R}$  and  $\mathbf{T}$ ) and mirror position  
 502 parameters ( $d$  and  $\mathbf{o}$ ) were obtained.

503 Table 2 summarizes the reference values and the estimation error (with re-  
 504 spect to the reference values) obtained for each experiment.

Setup	mirror (A,B,C) [mm]	calib. grid	reference values				estimation error				
			$d$ [mm]	$\mathbf{o}$ [pix]	$\mathbf{R}$ [Euler ang]	$\mathbf{T}$ [mm]	$d$ [%]	$\mathbf{o}$ [pix]	$\mathbf{R}$ [Euler ang]	$\mathbf{T}$ norm[%]; ang	reproj. RMS [pix]
#1	sphere:	#1	1164	$\begin{bmatrix} 571 \\ 386 \end{bmatrix}$	$51^\circ$	$-502$	0.8	$\begin{bmatrix} -4.9 \\ 3.0 \end{bmatrix}$	$2.3^\circ$	1.3; 1.9°	0.7
					$71^\circ$	$84$					
#2	$\begin{bmatrix} 1 \\ 0 \\ 300^2 \end{bmatrix}$	#2			$109^\circ$	$936$			$-0.2^\circ$		
					$-56^\circ$	$-680$	3.0	$\begin{bmatrix} -3.3 \\ 1.0 \end{bmatrix}$	$-1.2^\circ$	1.4; 0.5°	1.1
					$5^\circ$	$-11$					
					$-73^\circ$	$871$			$0.4^\circ$		
#3	hyperb.:	#1	45	$\begin{bmatrix} 401 \\ 296 \end{bmatrix}$	$90^\circ$	$620$	0.8	$\begin{bmatrix} -4.4 \\ -2.7 \end{bmatrix}$	$1.9^\circ$	0.4; 2.9°	0.4
					$0^\circ$	$-398$					
					$-90^\circ$	$-24$			$-0.5^\circ$		
#4	$\begin{bmatrix} -0.76 \\ 0 \\ -600 \end{bmatrix}$	#2			$178^\circ$	$438$	8.4	$\begin{bmatrix} 6.6 \\ 7.3 \end{bmatrix}$	$-3.2^\circ$	16.0; 9.6°	1.5
					$-1.4^\circ$	$536$					
					$-90^\circ$	$8$			$-3.3^\circ$		

Table 2: Experimental results with real images. For each mirror type two independent calibration grids were used. Mirror parameters are defined in equation B.1 of Appendix B. The reference values for the calibration parameters were obtained using direct measurement and Bouguet’s camera calibration toolbox (see text for details).  $\mathbf{o}$  is the *vertex point*;  $d$  is the distance between camera and mirror;  $\mathbf{R}$  and  $\mathbf{T}$  are the extrinsic rotation and translation, respectively.

#### 505 6.4. Discussion

506 The simulation results show that the method described in this paper allows  
 507 the estimation of the calibration parameters with good accuracy. The values  
 508 of the estimated parameters remain stable even in the presence of considerable  
 509 noise (i.e., when  $\sigma = 5$  pixels). At first sight, the value of the coordinates of  
 510 image point  $\mathbf{o}$  appears to be the most affected parameter, but the error loses

511 relevance when compared to the full image resolution (for  $\sigma = 5$ , the position  
512 error in  $\mathbf{o}$  is less than 2% of the image dimension).

513 Regarding the comparison with methods designed for central systems, we  
514 focused the analysis on the estimation of the extrinsic parameters as the re-  
515 maining parameters differ from our model. Besides the extrinsic parameters,  
516 we aim at the reconstruction of the mirror/camera position while the method  
517 of Mei and Rives uses the spherical camera model [42, 43] and the method of  
518 Scaramuzza et. al uses a distortion model for the image. Our method had a  
519 performance similar to the other techniques, especially when assuming that the  
520 *vertex point* was given a priori. The reprojection error was also presented to  
521 provide an overall evaluation, and all methods provided very similar results.

522 The experiments with real images demonstrate how a good estimation of  
523 the calibration parameters can be achieved from a very simple and practical  
524 setup, even with the highly non-linear image formation geometry of non-central  
525 catadioptric systems. We note, however, that in setup #4 (hyperbolic mirror,  
526 grid pattern 2) the estimation of  $d$  and  $\mathbf{T}$  presented larger errors, which can be  
527 explained by the fact that the reflection of the grid pattern occupied a relatively  
528 small area of the mirror surface (see topmost pattern in Fig. 11(b)), making the  
529 calibration points more sensitive to noise.

## 530 7. Conclusions

531 We presented a method for the estimation of the mirror position and extrinsic  
532 parameters in axial non-central catadioptric systems, i.e., systems made up  
533 of an axial symmetric mirror and a projective camera with its optical center  
534 located along the symmetry axis (but not necessarily orientated with the axis).  
535 We assume an internally calibrated pinhole camera and require the use of a  
536 calibration object, that can be planar. A single image is sufficient to perform  
537 the calibration procedure.

538 The camera/mirror relative position is determined with two parameters: the  
539 image position of the intersection of the symmetry axis with the image plane

540 and the distance from the camera center to the mirror. The extrinsic parameters  
541 are fully determined through a rotation matrix and a translation vector.

542 A linear projection equation is established between 3D points and 1D image  
543 features, which enables the use of the DLT algorithm in the estimation of the  
544 extrinsic rotation and translation, the latter up to one undetermined component.  
545 The cross-ratio is used as an invariant under the axial-symmetric geometry to  
546 determine the image of the axis. Non-linear optimization methods are applied  
547 in the estimation of the remaining parameters.

548 Regarding the estimation of mirror/camera relative position, our approach  
549 provides a significant alternative to methods that require the identification of  
550 the mirror boundary in the image (e.g. [17, 14, 21, 6, 34]). The calibration  
551 procedure is accurate and much easier to automate. Since the calibration object  
552 can be planar, the setup is easy to implement.

553 The estimation of the extrinsic parameters, up to one translation parameter,  
554 is achieved with a simple and linear procedure, even in the presence of a highly  
555 non-linear image formation geometry. In applications that do not require the  
556 z-component of the extrinsic translation to be determined, the extrinsic para-  
557 meters are obtained without full knowledge of the vision system characteristics  
558 (unknown mirror shape and distance to mirror). As an example of one such  
559 application, consider a robot navigating on a plane, equipped with an omni-  
560 directional vision system. If known landmarks (calibration points) are visible in  
561 the image, the robot's pose (extrinsic parameters) can be fully retrieved using  
562 the method of section 4, as the z-component of the translation is constrained  
563 by the plane on which the robot moves.

## 564 **Appendix A. Notation and Background**

565 This appendix briefly reviews some background concepts used in the paper.

566 *Appendix A.1. Cross-ratio*

Consider four collinear points  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{D}$ . Their cross-ratio  $\{\mathbf{ABCD}\}$  is defined as

$$\{\mathbf{ABCD}\} = \frac{|\mathbf{AB}||\mathbf{CD}|}{|\mathbf{AC}||\mathbf{BD}|}, \quad (\text{A.1})$$

567 where  $|\mathbf{XY}|$  denotes the (signed) distance between points  $\mathbf{X}$  and  $\mathbf{Y}$ .

Let  $\mathbf{O}$  be the intersection point of four concurrent lines, with each line passing through  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{D}$ , respectively. The cross-ratio of the four lines is given by

$$\{\mathbf{O}; \mathbf{ABCD}\} = \frac{\sin(\mathbf{AOB}) \sin(\mathbf{COD})}{\sin(\mathbf{AOC}) \sin(\mathbf{BOD})},$$

and we have that (c.f. [35], chapter 2)

$$\{\mathbf{ABCD}\} = \{\mathbf{O}; \mathbf{ABCD}\}$$

568 *Appendix A.2. Vector representation of conic curves*

Consider a 2D point, with homogeneous coordinates

$$\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix}^T,$$

and a conic curve represented by the symmetric matrix

$$\Omega \sim \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}.$$

Point  $\mathbf{x}$  is on the conic curve *iff*

$$\mathbf{x}^T \Omega \mathbf{x} = 0.$$

This second order polynomial can be re-written in the following form

$$\boldsymbol{\omega}^T \hat{\mathbf{x}} = 0, \quad (\text{A.2})$$

with  $\hat{\mathbf{x}}$  being the *lifted point coordinates* of  $\mathbf{x}$

$$\hat{\mathbf{x}} = \begin{bmatrix} x^2 & xy & y^2 & xz & yz & z^2 \end{bmatrix}^T, \quad (\text{A.3})$$

and  $\boldsymbol{\omega}$  a vector representation of the conic curve

$$\boldsymbol{\omega} = \begin{bmatrix} a & b & c & d & e & f \end{bmatrix}^T.$$

569 **Appendix B. Back-projection with conic section mirror**

570 In this appendix we show how to obtain the back-projection ray described  
 571 in section 5.1. The derivation is based on the work of Agrawal, Taguchi and  
 572 Ramalingam in [21]. That paper addressed the forward projection equations  
 573 in axial catadiotric systems with conic section mirrors, but concerning back-  
 574 projection, only the case with a spherical mirror was explicitly derived. We  
 575 present the back-projection equations for a generic conic section mirror.

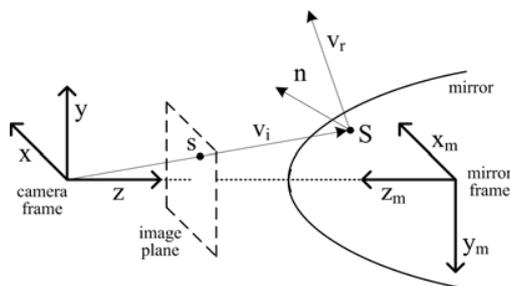


Figure B.12: Back-projection of an image point after reflection on a conic section mirror. See text for details.

576 Consider Fig. B.12. The camera principal axis ( $z_c$ ) is aligned with the mirror  
 577 symmetry axis ( $z_m$ ). The distance between the camera frame origin and the  
 578 mirror frame origin is given by  $d$ . Vector  $\mathbf{v}_i$  is the incident ray and  $\mathbf{v}_r$  is the  
 579 reflected ray.  $\mathbf{S}$  is the reflection point on the surface of the mirror.  $\mathbf{n}$  is the  
 580 surface normal vector at point  $\mathbf{S}$ .

The mirror is specified by three parameters,  $A$ ,  $B$  and  $C$ , that define its conic section in the  $x_m z_m$  plane:

$$Az_m^2 + x_m^2 + Bz_m = C. \quad (\text{B.1})$$

The incident ray direction for a image point  $\mathbf{q}$  (in pixels) is given, in the camera reference frame, by  $\mathbf{s} = \mathbf{K}^{-1}\mathbf{q}$ , where  $\mathbf{K}$  is the camera intrinsic calibration matrix. Let  $\mathbf{s} = [s_1 \ s_2 \ s_3]^T$ . The inhomogeneous coordinates of the reflection point are given, in the camera reference frame, by  $\tilde{\mathbf{S}} = \beta\mathbf{s}$ , with  $\beta$

obtained from

$$\beta = \frac{s_3(B+2Ad) \pm \sqrt{4(s_1^2+s_2^2)(-Bd-Ad^2+C)+s_3^2(B^2+4AC)}}{2(s_1^2+s_2^2+As_3^2)} \quad (\text{B.2})$$

As can be seen from equation B.2,  $\beta$  has, in general, two solutions, corresponding to two intersection points between the incident ray and the mirror surface. The smallest value of  $\beta$  that verifies  $\beta s_3 > 0$  is the one that corresponds to the reflection point closest to, and in front of, the camera.

Finally, using the laws of reflection, the direction of the reflected ray is obtained from

$$\mathbf{v}_r = \mathbf{v}_i - \frac{2\mathbf{nn}^\top}{\mathbf{n}^\top\mathbf{n}}\mathbf{v}_i,$$

with the incident ray given by  $v_i = \tilde{\mathbf{S}}$  and the normal vector at point  $\tilde{\mathbf{S}} = [S_x \ S_y \ S_z]^\top$  given by  $\mathbf{n} = [S_x \ S_y \ AS_z - Ad - B/2]^\top$ .

#### Acknowledgement

Helder Araujo would like to thank the support of project FCT/PTDC/EIA-EIA/122454/2010, funded by the Portuguese Science Foundation (FCT) by means of national funds (PIDDAC) and co-funded by the European Fund for Regional Development (FEDER) through COMPETE Operational Programme Competitive Factors (POFC). Luis Perdigoto acknowledges the support of the Portuguese Science Foundation through grant SFRH/BD/50281/2009.

#### References

- [1] S. Baker, S. K. Nayar, A theory of single-viewpoint catadioptric image formation, *International Journal of Computer Vision* 35 (1999) 175–196.
- [2] J. P. Barreto, H. Araujo, Paracatadioptric camera calibration using lines, in: *International Conference on Computer Vision*, pp. 1359–1365.
- [3] C. Geyer, K. Daniilidis, Paracatadioptric camera calibration, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24 (2002) 687–695.

- 602 [4] X. Ying, Z. Hu, Catadioptric camera calibration using geometric invariants,  
603 IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2004)  
604 1260–1271.
- 605 [5] B. Micusik, T. Pajdla, Para-catadioptric camera auto-calibration from  
606 epipolar geometry, in: Asian Conference on Computer Vision (ACCV),  
607 Korea.
- 608 [6] S. B. Kang, Catadioptric self-calibration, in: Computer Vision and Pattern  
609 Recognition, 2000. Proceedings. IEEE Conference on, volume 1, pp. 201 –  
610 207 vol.1.
- 611 [7] L. Zhang, X. Du, Y. Zhu, J. Liu, Central catadioptric camera calibration  
612 with single image, in: Acoustics, Speech and Signal Processing, 2009.  
613 ICASSP 2009. IEEE International Conference on, pp. 1253 –1256.
- 614 [8] D. Scaramuzza, A. Martinelli, R. Siegwart, A flexible technique for ac-  
615 curate omnidirectional camera calibration and structure from motion, in:  
616 Computer Vision Systems, 2006 ICVS '06. IEEE International Conference  
617 on, p. 45.
- 618 [9] C. Mei, P. Rives, Single view point omnidirectional camera calibration  
619 from planar grids, in: Robotics and Automation, 2007 IEEE International  
620 Conference on, pp. 3945 –3950.
- 621 [10] L. Puig, J. Bermúdez, P. Sturm, J. J. Guerrero, Calibration of omnidirec-  
622 tional cameras in practice: A comparison of methods, Comput. Vis. Image  
623 Underst. 116 (2012) 120–137.
- 624 [11] M. D. Grossberg, S. K. Nayar, A general imaging model and a method for  
625 finding its parameters, in: In Proc. ICCV, pp. 108–115.
- 626 [12] M. Grossberg, S. Nayar, The Raxel Imaging Model and Ray-Based Cali-  
627 bration, International Journal on Computer Vision 61 (2005) 119–137.

- 628 [13] P. Sturm, S. Ramalingam, A generic concept for camera calibration, in:  
629 Proceedings of the European Conference on Computer Vision, Prague,  
630 Czech Republic, volume 2, Springer, 2004, pp. 1–13.
- 631 [14] Y. Taguchi, A. Agrawal, A. Veeraraghavan, S. Ramalingam, R. Raskar,  
632 Axial-cones: Modeling spherical catadioptric cameras for wide-angle light  
633 field rendering, ACM Transactions on Graphics (Proceedings of SIG-  
634 GRAPH Asia 2010) 29 (2010) 172:1–172:8.
- 635 [15] D. Lanman, D. Crispell, M. Wachs, G. Taubin, Spherical catadioptric ar-  
636 rays: Construction, multi-view geometry, and calibration, in: 3D Data  
637 Processing, Visualization, and Transmission, Third International Symposi-  
638 um on, pp. 81 –88.
- 639 [16] C.-S. Chen, W.-Y. Chang, On pose recovery for generalized visual sensors,  
640 Pattern Analysis and Machine Intelligence, IEEE Transactions on 26 (2004)  
641 848 –861.
- 642 [17] B. Micusik, T. Pajdla, Autocalibration and 3d reconstruction with non-  
643 central catadioptric cameras, in: Computer Vision and Pattern Recogni-  
644 tion, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society  
645 Conference on, volume 1, pp. I–58 – I–65 Vol.1.
- 646 [18] N. Goncalves, H. Araujo, Linear solution for the pose estimation of noncen-  
647 tral catadioptric systems, in: Computer Vision, 2007. ICCV 2007. IEEE  
648 11th International Conference on, pp. 1 –7.
- 649 [19] R. Swaminathan, M. D. Grossberg, S. K. Nayar, Non-Single Viewpoint  
650 Catadioptric Cameras: Geometry and Analysis, Technical Report, Inter-  
651 national Journal of Computer Vision, 2001.
- 652 [20] J. Fabrizio, J.-P. Tarel, R. Benosman, Calibration of panoramic catadiop-  
653 tric sensors made easier, in: Omnidirectional Vision, 2002. Proceedings.  
654 Third Workshop on, pp. 45 – 52.

- 655 [21] A. Agrawal, Y. Taguchi, S. Ramalingam, Analytical forward projection  
656 for axial non-central dioptric and catadioptric cameras, in: Proceedings of  
657 the 11th European conference on computer vision conference on Computer  
658 vision: Part III, ECCV'10, pp. 129–143.
- 659 [22] S. Ramalingam, P. Sturm, E. Boyer, A factorization based self-calibration  
660 for radially symmetric cameras, in: 3D Data Processing, Visualization, and  
661 Transmission, Third International Symposium on, 2006, pp. 480–487.
- 662 [23] C. Geyer, K. Daniilidis, Structure and motion from uncalibrated catadiop-  
663 tric views, in: Computer Vision and Pattern Recognition, 2001. CVPR  
664 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol-  
665 ume 1, pp. I-279 – I-286 vol.1.
- 666 [24] V. Caglioti, P. Taddei, G. Boracchi, S. Gasparini, A. Giusti, Single-image  
667 calibration of off-axis catadioptric cameras using lines, in: Computer Vi-  
668 sion, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1–6.
- 669 [25] R. Sagawa, N. Aoki, Y. Mukaigawa, T. Echigo, Y. Yagi, Mirror localization  
670 for a catadioptric imaging system by projecting parallel lights, in: Robotics  
671 and Automation, 2007 IEEE International Conference on, pp. 3957–3962.
- 672 [26] M. D. Grossberg, S. K. Nayar, The raxel imaging model and ray-based  
673 calibration, *Int. J. Comput. Vision* 61 (2005) 119–137.
- 674 [27] J.-P. Tardif, P. Sturm, M. Trudeau, S. Roy, Calibration of cameras with  
675 radially symmetric distortion, *Pattern Analysis and Machine Intelligence,*  
676 *IEEE Transactions on* 31 (2009) 1552–1566.
- 677 [28] O. Morel, D. Fofi, Calibration of catadioptric sensors by polarization imag-  
678 ing, in: Robotics and Automation, 2007 IEEE International Conference  
679 on, pp. 3939–3944.
- 680 [29] P. Sturm, S. Ramalingam, J.-P. Tardif, S. Gasparini, J. Barreto, Cam-  
681 era models and fundamental concepts used in geometric computer vision,  
682 *Found. Trends. Comput. Graph. Vis.* 6 (2011) 1–183.

- 683 [30] Y. Wu, Z. Hu, Geometric invariants and applications under catadioptric  
684 camera model, in: *Computer Vision, 2005. ICCV 2005. Tenth IEEE Inter-*  
685 *national Conference on*, volume 2, pp. 1547 –1554 Vol. 2.
- 686 [31] S. Thirthala, M. Pollefeys, Multi-view geometry of 1d radial cameras and  
687 its application to omnidirectional camera calibration, in: *Computer Vision,*  
688 *2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pp.  
689 1539 –1546 Vol. 2.
- 690 [32] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision.*  
691 2nd edn., Cambridge University Press, 2004.
- 692 [33] D. Lanman, D. Crispell, M. Wachs, G. Taubin, Spherical catadioptric ar-  
693 rays: Construction, multi-view geometry, and calibration, in: *3D Data*  
694 *Processing, Visualization, and Transmission, Third International Sympo-*  
695 *sium on*, pp. 81 –88.
- 696 [34] T. Mashita, Y. Iwai, M. Yachida, Calibration method for misaligned cata-  
697 dioptric camera, in: *In Proceedings of the 6th Workshop on Omnidirec-*  
698 *tional Vision (OMNIVIS 2005).*
- 699 [35] J. Semple, G. Kneebone, *Algebraic Projective Geometry*, Clarendon Press,  
700 1998.
- 701 [36] Z. Zhang, A flexible new technique for camera calibration, *Pattern Analysis*  
702 *and Machine Intelligence, IEEE Transactions on* 22 (2000) 1330 – 1334.
- 703 [37] J.-Y. Bouguet, Camera calibration toolbox for matlab,  
704 [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html), 2012.
- 705 [38] P. Besl, H. McKay, A method for registration of 3-d shapes, *Pattern*  
706 *Analysis and Machine Intelligence, IEEE Transactions on* 14 (1992) 239  
707 –256.

- 708 [39] D. Scaramuzza, A. Martinelli, R. Siegwart, A toolbox for easily calibrat-  
 709 ing omnidirectional cameras, in: Intelligent Robots and Systems, 2006  
 710 IEEE/RSJ International Conference on, pp. 5695–5701.
- 711 [40] D. Scaramuzza, Ocamcalib: Omnidirectional camera calibration  
 712 toolbox for matlab, [https://sites.google.com/site/scarobotix/ocamcalib-](https://sites.google.com/site/scarobotix/ocamcalib-toolbox/ocamcalib-toolbox-download-page)  
 713 [toolbox/ocamcalib-toolbox-download-page](https://sites.google.com/site/scarobotix/ocamcalib-toolbox/ocamcalib-toolbox-download-page), 2012.
- 714 [41] C. Mei, Omnidirectional calibration toolbox, [http://www.robots.ox.ac.uk/](http://www.robots.ox.ac.uk/~cmei/Toolbox.html)  
 715 [~cmei/Toolbox.html](http://www.robots.ox.ac.uk/~cmei/Toolbox.html), 2012.
- 716 [42] C. Geyer, K. Daniilidis, A unifying theory for central panoramic systems  
 717 and practical applications, in: Proceedings of the 6th European Conference  
 718 on Computer Vision-Part II, ECCV '00, Springer-Verlag, London, UK, UK,  
 719 2000, pp. 445–461.
- 720 [43] J. Barreto, H. Araujo, Issues on the geometry of central catadioptric image  
 721 formation, in: Computer Vision and Pattern Recognition, 2001. CVPR  
 722 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol-  
 723 ume 2, pp. II-422 – II-427 vol.2.



**Luis Perdigoto** received the M.S. degree in Electrical and Computer Engineering from University of Coimbra (UC), Portugal, in 2007. He is currently a Ph.D. student at the Institute for Systems and Robotics (UC) and an Assistant in the Dept. of Electrical Engineering, ESTG - Polytechnic Institute of Leiria, Portugal.



**Helder Araujo** is a Professor in the Department of Electrical and Computer Engineering of the University of Coimbra, Portugal. He is also a researcher at the Institute for Systems and Robotics-Coimbra. His research interests include computer and robot vision, robot navigation and localization, and sensor modeling for robot navigation and localization.

We propose a novel calibration method for non-central catadioptric systems.

We assume an axial symmetrical mirror and a pinhole camera placed on the mirror axis.

The calibration estimates the camera/mirror position and the extrinsic parameters.

The procedure requires a single image of a (possibly planar) calibration object.

The Direct-Linear-Transformation algorithm and cross-ratio are used.

ACCEPTED MANUSCRIPT