HAL
open science

General Imaging Geometry for Central Catadioptric Cameras

Peter Sturm, Joao Barreto

To cite this version:

Peter Sturm, Joao Barreto. General Imaging Geometry for Central Catadioptric Cameras. ECCV 2008 - 10th European Conference on Computer Vision, Oct 2008, Marseille, France. pp.609-622, 10.1007/978-3-540-88693-8_45 . inria-00384256

HAL Id: inria-00384256

https://inria.hal.science/inria-00384256

Submitted on 14 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

General Imaging Geometry for Central Catadioptric Cameras

Peter Sturm[1] and João P. Barreto[2]

[1] INRIA Rhône-Alpes and Laboratoire Jean Kuntzmann, Grenoble, France
[2] ISR/DEEC, University of Coimbra, 3030 Coimbra, Portugal

Abstract. Catadioptric cameras are a popular type of omnidirectional imaging system. Their imaging and multi-view geometry has been extensively studied; epipolar geometry for instance, is geometrically speaking, well understood. However, the existence of a bilinear matching constraint and an associated fundamental matrix, has so far only been shown for the special case of para-catadioptric cameras (consisting of a paraboloidal mirror and an orthographic camera). The main goal of this work is to obtain such results for all central catadioptric cameras. Our main result is to show the existence of a general 15×15 fundamental matrix. This is based on and completed by a number of other results, e.g. the formulation of general catadioptric projection matrices and plane homographies.

1 Introduction and Previous Work

The geometry of single and multiple images has been extensively studied in computer vision and photogrammetry [1]. The picture is rather complete for perspective cameras and many results have been obtained for other camera models too, e.g. catadioptric [2–5], fisheyes [6, 7], pushbroom [8], x-slit [9], oblique [10, 11], non-central mosaics [12, 11]; this list is not intended to be exhaustive.

Besides perspective cameras, the most studied case is probably that of catadioptric ones. Baker and Nayar have shown which catadioptric devices have a single effective viewpoint, i.e. are central cameras [13]. Among those, the most useful ones are the para-catadioptric and the hyper-catadioptric models, using a mirror of paraboloidal/hyperboloidal shape, coupled with an orthographic/perspective camera. The epipolar geometry of these devices has been studied by Svoboda and Pajdla [2] who showed the existence of epipolar *conics*. Geyer and Daniilidis have shown the existence of a fundamental matrix for para-catadioptric cameras [4, 14]; Sturm has extended this to fundamental matrices and trifocal tensors for mixtures of para-catadioptric and perspective images [5]. Barreto showed that the framework can also be extended to cameras with lens distortion due to the similarities between the para-catadioptric and division models [15, 16]

However, no such results have so far been obtained for the general catadioptric camera model, i.e. including hyper-catadioptric cameras. In this paper, we present a number of novel results concerning (bi-) linear formulations for single and two-view geometry, valid for all central catadioptric cameras. First, we show

that the projection of a 3D point can be modeled using a projection matrix of size 6×10 and how this may be used for calibrating catadioptric cameras with a straightforward DLT approach, something which has not been possible up to now. We then give analogous results for the backprojection of image points and the projection of quadrics and conics. These are the basis for our main result, the general fundamental matrix for catadioptric cameras. It is of size 15×15 and an explicit compact expression is provided. Finally, we also show the existence of plane homographies, again of size 15×15, that relate sets of matching image points that are the projections of coplanar scene points.

Our results, like those cited above for para-catadioptric cameras, are based on the use of so-called *lifted coordinates* to represent geometric objects. For example, 2D points are usually represented by 3-vectors of homogeneous coordinates; their lifted coordinates are 6-vectors containing all degree-2 monomials of the original coordinates. Lifted coordinates have also been used to model linear pushbroom cameras [8] and to perform multi-body structure from motion [17, 18].

Organization. We describe our results in two ways, geometrically and algebraically. In the next section, we immediately describe all main results in a purely geometrical way, which we find rather intuitive and which hopefully guides the reader through the more technical later sections. In section 3, we introduce notations and background, mostly associated to Veronese maps which are extensively used in this work. In sections 4 to 8, we develop algebraic formulations for the projection of 3D points, backprojection of image points, projection of quadrics and conics, epipolar geometry, and plane homographies.

2 Geometrical Description of Our Results

We use the sphere based model for catadioptric projection introduced by Geyer and Daniilidis [19]. All central catadioptric cameras can be modeled by a unit sphere and a perspective camera, such that the projection of 3D points can be performed as follows in two steps. First, one projects the point onto the sphere, to an intersection of the sphere and the line joining its center and the 3D point. There are two such intersection points; it is usually assumed that the only physically feasible one can be singled out. That point is then projected into the perspective camera. This model covers all central catadioptric cameras, the type of which is encoded by the distance between the perspective camera and the center of the sphere, e.g. 0 for perspective, 1 for para-catadioptric, $0 < \xi < 1$ for hyper-catadioptric. Note that even for the para-catadioptric case, where the true camera is an affine one, the camera in the sphere based model is still perspective. An algebraic formulation of the model is given in section 3. We now give geometrical and intuitive descriptions of our main results.

Projection of a 3D point. In the first step of the projection as described above, the intersections of the sphere with the line spanned by its center and the 3D point, are computed. There are two mathematical solutions which when projected to the perspective camera, give the two mathematical image points, cf. figure 1. For most real catadioptric cameras, only one of them can be observed;

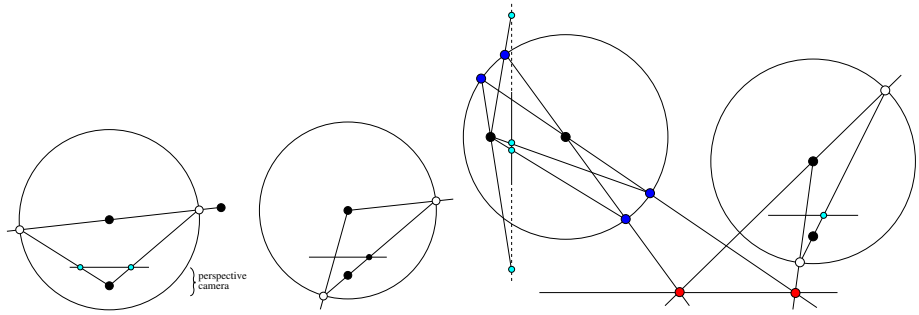

Fig. 1. *Left:* Projection of a 3D point to two image points (in cyan). *Middle:* Backprojection of an image point to two 3D lines. *Right:* Illustration of plane homography. The image point in cyan in the right camera is backprojected to the scene plane, giving the two red points. Each one of them is projected to the unit sphere of the second camera on two blue points. The four image points in the second camera are shown in cyan.

the second intersection point with the sphere is hidden from the 3D point by the mirror. An exception is the case of an elliptical mirror where a 3D point may actually be seen twice in the image. Although for the most useful catadioptric cameras, a 3D point is in reality visible in only one image point, it turns out that in order to obtain multi-linear expressions for epipolar geometry, plane homographies etc., both mathematical image points have to be considered, see later. So, from now on we consider two image points per 3D point and similar considerations will be done for backprojection etc. in the following. The algebraic formulation of the projection of 3D points, in the form of a 6×10 general *catadioptric projection matrix*, is given in section 4.

Projection of a 3D line. It is well known that a central catadioptric image of a 3D line, is a conic [19, 20]. In this work, we do not explicitly require the projection of 3D lines, but we still mention this result since it is helpful to understand the epipolar geometry, see below.

Backprojection of an image point. First, the image point is backprojected relative to the perspective camera, giving rise to a 3D line. Then, its two intersection points with the sphere are computed. Finally, the two lines spanned by the sphere center and these intersection points, are generated. These are the backprojection lines of the image point, cf. figure 1 (middle). In section 5 we show how to represent the union of two 3D lines by a single algebraic object, leading to the formulation of a 6×6 *backprojection matrix*.

Two images of a plane. Consider $\mathbf{q}_1$, the image of a 3D point on a scene plane Π, in the first image. What are the possible matching points in the second image? These can be determined as follows. Let us first backproject $\mathbf{q}_1$ to 3D; as we have seen just before, this gives two 3D lines. Their intersections with Π are the two points on that plane that may be observed in $\mathbf{q}_1$. Let us project both of them into the second image. This gives a total of four points in the second image, all of which are mathematically plausible matches of $\mathbf{q}_1$, cf. figure 1 (right). In

section 8 we give an algebraic representation of a *catadioptric plane homography*, that maps one image point onto an entity representing its four possible matches.

Epipolar geometry. The basic question of epipolar geometry is: what is the locus of points in the second image, that may be matches of a point $\mathbf{q}_1$ in the first image? The answer follows from the insights explained so far. Let us first backproject $\mathbf{q}_1$. The two 3D lines we get can then be projected into the second image, giving one conic each. Hence, the locus of matching points is the union of two conics. This can be represented by a single geometric entity, a quartic curve (note of course that not every quartic curve is the union of two conics).

Hence, if a multi-linear fundamental matrix exists that represents this epipolar geometry, it must map an image point into some representation of a quartic curve. The equation of a planar quartic curve depends on 15 coefficients (defined up to scale), one per 4-th order monomial of a 2D point's homogeneous coordinates. Hence, we may expect the fundamental matrix to be of size $15 \times \cdots$. Like for perspective images, we may expect that the transpose of the fundamental matrix gives the fundamental matrix going from the second to the first image. The fundamental matrix for catadioptric images should thus intuitively be of size 15×15. This is indeed the case, as is shown in section 7.

3 Background

Notations. We do not distinguish between a projective transformation and the matrix representing it. Matrices are represented by symbols in sans serif font, e.g. M and vectors by bold symbols, e.g. $\mathbf{Q}$. Equality of matrices or vectors up to a scalar factor is written as $\sim$. $[\mathbf{a}]_\times$ denotes the skew-symmetric matrix associated with the cross product of a 3-vector $\mathbf{a}$.

Camera model. As mentioned before, we use the sphere based model [19]. Without loss of generality, let the unit sphere be located at the origin and the optical center of the perspective camera, at the point $\mathbf{C}_p = (0, 0, -\xi)^\mathsf{T}$. The perspective camera is modeled by the projection matrix $\mathsf{P} \sim \mathsf{A}_p\mathsf{R}_p \left(\mathtt{I} - \mathbf{C}_p\right)$. For full generality, we include a rotation R_p; this may encode an actual rotation of the true camera looking at the mirror, but may also simply be a projective change of coordinates in the image plane, like for para-catadioptric cameras, where the true camera's rotation is fixed, modulo rotation about the mirror axis. Note that all parameters of the perspective camera, i.e. both its intrinsic and extrinsic parameter sets, are intrinsic parameters for the catadioptric camera. Hence, we replace $\mathsf{A}_p\mathsf{R}_p$ by a generic projective transformation K from now on. The **intrinsic parameters of the catadioptric camera** are thus ξ and K.

The *projection* of a 3D point $\mathbf{Q}$ goes as follows (cf. section 2). The two intersection points of the sphere and the line joining its center and $\mathbf{Q}$, are $\left(Q_1, Q_2, Q_3, \pm\sqrt{Q_1^2 + Q_2^2 + Q_3^2}\right)^\mathsf{T}$. Their images in the perspective camera are

$$\mathbf{q}_\pm \sim \mathsf{K}\mathbf{r}_\pm \sim \mathsf{K} \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \pm \xi\sqrt{Q_1^2 + Q_2^2 + Q_3^2} \end{pmatrix}$$

In the following, we usually first work with the intermediate image points $\mathbf{r}_\pm \sim \mathsf{K}^{-1}\mathbf{q}_\pm$, before giving final results for the actual image points $\mathbf{q}_\pm$.

Plücker line coordinates. 3D lines may be represented by 6-vectors of so-called Plücker coordinates. Let $\mathbf{A}$ and $\mathbf{B}$ be the non-homogeneous coordinates of two generic 3D points. Let us define the line's Plücker coordinates as the 6-vector $\mathbf{L} = \left(\mathbf{A}^\mathsf{T} - \mathbf{B}^\mathsf{T}, (\mathbf{A} \times \mathbf{B})^\mathsf{T}\right)^\mathsf{T}$.

All lines satisfy the Plücker constraint $\mathbf{L}^\mathsf{T}\mathsf{W}\mathbf{L} = 0$ where W is

$$\mathsf{W} = \begin{pmatrix} 0 & \mathtt{I} \\ \mathtt{I} & 0 \end{pmatrix}$$

Two lines $\mathbf{L}$ and $\mathbf{L}'$ cut one another if and only if $\mathbf{L}^\mathsf{T}\mathsf{W}\mathbf{L}' = 0$. Consider a rigid transformation for points

$$\begin{pmatrix} \mathsf{R} & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix}$$

Lines are mapped accordingly using the transformation

$$\mathsf{T} = \begin{pmatrix} \mathsf{R} & 0_{3\times 3} \\ [\mathbf{t}]_\times \mathsf{R} & \mathsf{R} \end{pmatrix}$$

Second order line complexes. A second order line complex is a set of 3D lines that satisfy a quadratic equation in the Plücker coordinates [21]. It can be represented by a symmetric 6×6 matrix C such that exactly the lines on the complex satisfy $\mathbf{L}^\mathsf{T}\mathsf{C}\mathbf{L} = 0$. Note that C is only defined up to adding multiples of W. Henceforth we call second order line complexes shortly *line complexes*. In this paper, we use line complexes to represent the union of two 3D lines. Rigid displacements of line complexes are carried out as

$$\mathsf{T}^{-\mathsf{T}}\mathsf{C}\mathsf{T}^{-1} \quad \text{with} \quad \mathsf{T}^{-1} = \begin{pmatrix} \mathsf{R}^\mathsf{T} & 0_{3\times 3} \\ -\mathsf{R}^\mathsf{T}[\mathbf{t}]_\times & \mathsf{R}^\mathsf{T} \end{pmatrix} \tag{1}$$

Lifted coordinates from symmetric matrix equations. The derivation of (multi-) linear relations for catadioptric imagery requires the use of lifted coordinates. The Veronese map $V_{n,d}$ of degree d maps points of $\mathcal{P}^n$ into points of an m dimensional projective space $\mathcal{P}^m$, with $m = \binom{n+d}{d} - 1$.

Consider the second order Veronese map $V_{2,2}$, that embeds the projective plane into the 5D projective space, by lifting the coordinates of point $\mathbf{q}$ to

$$\hat{\mathbf{q}} = \left(q_1^2 \; q_1 q_2 \; q_2^2 \; q_1 q_3 \; q_2 q_3 \; q_3^2\right)^\mathsf{T}$$

Vector $\hat{\mathbf{q}}$ and matrix $\mathbf{q}\mathbf{q}^\mathsf{T}$ are composed by the same elements. The former can be derived from the latter through a suitable re-arrangement of parameters. Define $\mathbf{v}(\mathsf{U})$ as the vector obtained by stacking the columns of a generic matrix U [22]. For the case of $\mathbf{q}\mathbf{q}^\mathsf{T}$, $\mathbf{v}(\mathbf{q}\mathbf{q}^\mathsf{T})$ has several repeated elements because of

matrix symmetry. By left multiplication with a suitable permutation matrix S that adds the repeated elements, it follows that

$$\hat{\mathbf{q}} = \mathsf{D}^{-1} \underbrace{\begin{pmatrix} 1&0&0&0&0&0&0&0&0 \\ 0&1&0&1&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&1&0&0 \\ 0&0&0&0&0&1&0&1&0 \\ 1&0&0&0&0&0&0&0&1 \end{pmatrix}}_{\mathsf{S}} \mathbf{v}(\mathbf{q}\mathbf{q}^{\mathsf{T}}), \tag{2}$$

with D a diagonal matrix, $D_{ii} = \sum_{j=1}^{9} S_{ij}$. This process of computing the lifted representation of a point $\mathbf{q}$ can be extended to any second order Veronese map $V_{n,2}$ independently of the dimensionality of the original space. It is also a mechanism that provides a compact representation for square symmetric matrices. If U is symmetric, then it is uniquely represented by $\mathbf{v}_{sym}(\mathsf{U})$, the column-wise vectorization of its upper right triangular part:

$$\mathbf{v}_{sym}(\mathsf{U}) = \mathsf{D}^{-1}\mathsf{S}\mathsf{U} = (U_{11}, U_{12}, U_{22}, U_{13}, \cdots, U_{nn})^{\mathsf{T}}$$

Let us now discuss the lifting of linear transformations. Consider A such that $\mathbf{r} = \mathsf{A}\mathbf{q}$. The relation $\mathbf{r}\mathbf{r}^{\mathsf{T}} = \mathsf{A}(\mathbf{q}\mathbf{q}^{\mathsf{T}})\mathsf{A}^{\mathsf{T}}$ can be written as a vector mapping

$$(\mathbf{r}\mathbf{r}^{\mathsf{T}}) = (\mathsf{A} \otimes \mathsf{A})(\mathbf{q}\mathbf{q}^{\mathsf{T}}),$$

with $\otimes$ denoting the Kronecker product [22]. Using the symmetric vectorization, we have $\hat{\mathbf{q}} = \mathbf{v}_{sym}(\mathbf{q}\mathbf{q}^{\mathsf{T}})$ and $\hat{\mathbf{r}} = \mathbf{v}_{sym}(\mathbf{r}\mathbf{r}^{\mathsf{T}})$, thus:

$$\hat{\mathbf{r}} = \underbrace{\mathsf{D}^{-1}\mathsf{S}(\mathsf{A} \otimes \mathsf{A})\mathsf{S}^{\mathsf{T}}}_{\hat{\mathsf{A}}} \hat{\mathbf{v}}$$

We have just derived the expression for lifting linear transformations. A has a lifted counterpart $\hat{\mathsf{A}}$ such that $\mathbf{r} = \mathsf{A}\,\mathbf{q}$ *iff* $\hat{\mathbf{r}} = \hat{\mathsf{A}}\,\hat{\mathbf{q}}$. For the case of a second order Veronese map, the lifting of a 2D projective transformation A is $\hat{\mathsf{A}}$ of size 6×6. This lifting generalizes to any projective transformation, independently of the dimensions of its original and target spaces, i.e. it is also applicable to rectangular matrices. We summarize a few useful properties [22].

$$\widehat{\mathsf{A}\mathsf{B}} = \hat{\mathsf{A}}\hat{\mathsf{B}} \qquad \widehat{\mathsf{A}^{-1}} = \hat{\mathsf{A}}^{-1} \qquad \widehat{\mathsf{A}^{\mathsf{T}}} = \mathsf{D}^{-1}\hat{\mathsf{A}}^{\mathsf{T}}\mathsf{D} \tag{3}$$

Also, for symmetric matrices U and M, we have the following property:

$$\mathsf{U} = \mathsf{A}\mathsf{M}\mathsf{A}^{\mathsf{T}} \Rightarrow \mathbf{v}_{sym}(\mathsf{U}) = \hat{\mathsf{A}}\ \mathbf{v}_{sym}(\mathsf{M}) \tag{4}$$

Also, note that $\hat{\mathsf{A}}$ is non-singular *iff* A is non-singular. In this paper, we use the following liftings: 3-vectors $\mathbf{q}$ to 6-vectors $\hat{\mathbf{q}}$, 4-vectors $\mathbf{Q}$ to 10-vectors $\hat{\mathbf{Q}}$, and 6-vectors $\mathbf{u}$ to 21-vectors $\hat{\mathbf{u}}$. Analogously, 3×3 matrices W are lifted to 6×6 ones $\hat{\mathsf{W}}$ and 3×4 matrices to 6×10 ones. We also use the fourth order Veronese map of $\mathcal{P}^2$, mapping 3-vectors $\mathbf{q}$ to 15-vectors $\hat{\hat{\mathbf{q}}}$ containing the quartic monomials of $\mathbf{q}$. We call this *double lifting*; it applies analogously to 3×3 matrices W, which are doubly lifted to 15×15 matrices $\hat{\hat{\mathsf{W}}}$. Finally, note that applying two second

order Veronese maps in succession, is **not** equivalent to applying one fourth order Veronese map: for a 3-vector $\mathbf{q}$, $\hat{\hat{\mathbf{q}}}$ is a 15-vector, whereas $\hat{\mathbf{w}}$ where $\mathbf{w}$ is the 6-vector $\mathbf{w} = \hat{\mathbf{q}}$, is of length 21. We thus denote successive application of two second order liftings by $\mathring{}$, e.g. for a 3×3 matrix E, we get a 21×21 matrix $\mathring{\mathsf{E}}$.

4 Projection of 3D Points

As explained in section 2, a 3D point is mathematically projected to two image points. How to represent two 2D points via a single geometric entity? One way is to compute the degenerate dual conic generated by them, i.e. the dual conic containing exactly the lines going through at least one of the two points. Let the two image points be $\mathbf{q}_+$ and $\mathbf{q}_-$ (see section 3). The dual conic is given by

$$\Omega \sim \mathbf{q}_+\mathbf{q}_-^\mathsf{T} + \mathbf{q}_-\mathbf{q}_+^\mathsf{T} \sim \mathsf{K} \begin{pmatrix} Q_1^2 & Q_1Q_2 & Q_1Q_3 \\ Q_1Q_2 & Q_2^2 & Q_2Q_3 \\ Q_1Q_3 & Q_2Q_3 & Q_3^2 - \xi^2(Q_1^2 + Q_2^2 + Q_3^2) \end{pmatrix} \mathsf{K}^\mathsf{T}$$

This can be written as a linear mapping of the 3D point's lifted coordinates, onto the vectorized matrix of the conic:

$$\mathbf{v}_{sym}(\Omega) \sim \widehat{\mathsf{K}}_{6\times 6} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -\xi^2 & 0 & -\xi^2 & 0 & 0 & 1-\xi^2 \end{pmatrix}}_{\mathsf{X}_\xi} \begin{pmatrix} \mathtt{I}_6 & \mathtt{0}_{6\times 4} \end{pmatrix} \hat{\mathbf{Q}}$$

So far, we have projected a 3D point given in the catadioptric camera's local coordinate system. If we introduce extrinsic parameters of the camera, i.e. a pose matrix $\mathsf{T} = \mathsf{R}\begin{pmatrix} \mathtt{I} & -\mathbf{t} \end{pmatrix}$ then we can write the projection operation as

$$\mathbf{v}_{sym}(\Omega) \sim \widehat{\mathsf{K}}_{6\times 6}\mathsf{X}_{\xi,6\times 6}\widehat{\mathsf{R}}_{6\times 6} \begin{pmatrix} \mathtt{I}_6 & \mathsf{T}'_{4\times 6} \end{pmatrix} \hat{\mathbf{Q}}_{10}$$

where T' depends only on $\mathbf{t}$.

We have thus derived a 6×10 **catadioptric projection matrix** P_{cata}. Furthermore, it can be decomposed, like the projection matrix of a perspective camera, into two matrices containing either intrinsic or extrinsic parameters:

$$\mathsf{P}_{cata} = \underbrace{\widehat{\mathsf{K}}\mathsf{X}_\xi}_{\mathsf{A}_{cata}} \underbrace{\widehat{\mathsf{R}}_{6\times 6} \begin{pmatrix} \mathtt{I}_6 & \mathsf{T}'_{4\times 6} \end{pmatrix}}_{\mathsf{T}_{cata}}$$

We find here a 6×6 **catadioptric calibration matrix** A_{cata}. Note that the restriction of the projection to points lying in a plane in 3D, leads straightforwardly to a 6×6 homography, analogous to the plane-to-image homographies used e.g. for perspective camera calibration.

Camera Calibration. Let us consider how to set up equations for calibrating a camera. In the perspective case, a 2D–3D point correspondence allows to write $\mathbf{q} \sim \mathsf{P}\mathbf{Q}$. One way to set up linear equations on P is to write $[\mathbf{q}]_\times \mathsf{P}\mathbf{Q} = \mathbf{0}_3$. What is the analogous expression in the catadioptric case? Let again $\mathbf{q}$ and $\mathbf{Q}$ be a 2D–3D point correspondence. Since each 3D point is projected to two 2D points, one may not directly be able to compare $\mathbf{q}$ to the image of $\mathbf{Q}$, unlike in the perspective case. Instead, as mentioned above, the projection matrix maps $\mathbf{Q}$ (rather, its lifted version) onto the coefficients of a degenerate dual conic Ω. The point $\mathbf{q}$ must be one of the two generators of Ω. This implies that all lines through $\mathbf{q}$ must lie on Ω. Hence: $\forall \mathbf{p} : \mathbf{p} \times \mathbf{q} \in \Omega$, which gives $\forall \mathbf{p} : \mathbf{p}^\mathsf{T}[\mathbf{q}]_\times \Omega [\mathbf{q}]_\times \mathbf{p} = 0$. Thus, $[\mathbf{q}]_\times \Omega [\mathbf{q}]_\times = \mathsf{0}_{3\times 3}$. This gives 6 constraints that can be written as

$$\left(\widehat{[\mathbf{q}]_\times}\right)_{6\times 6} \mathbf{v}_{sym}(\Omega) = \widehat{[\mathbf{q}]_\times}\ \mathsf{P}_{cata}\ \hat{\mathbf{Q}} = \mathbf{0}_6$$

We thus find an expression that is very similar to that for perspective cameras and that may be directly used for calibrating catadioptric cameras using e.g. a standard DLT like approach. While a 3×3 skew symmetric matrix has rank 2, its lifted counterpart is rank 3. Therefore, each 3D-to-2D match provides 3 linear constraints on the 59 parameters of P_{cata}, and DLT calibration can be done with a minimum of 20 matches.

5 Backprojection of Image Points

This is essential for deriving the proposed expression of the fundamental matrix. Similarly to the case of projection, we want to express the backprojection function of a catadioptric camera as a linear mapping. Recall from section 2, that the backprojection of an image point gives **two** 3D lines. How to represent two 3D lines via a single geometric entity? Several possibilities may exist; the one that seems appropriate is to use a second order line complex: consider two 3D lines $\mathbf{L}_+$ and $\mathbf{L}_-$. All lines that cut at least one of them, form a second order line complex, represented by a 6×6 matrix C such that lines on the complex satisfy $\mathbf{L}^\mathsf{T}\mathsf{C}\mathbf{L} = 0$. The matrix C is given as (with W as defined in section 3)

$$\mathsf{C} \sim \mathsf{W}\left(\mathbf{L}_+\mathbf{L}_-^\mathsf{T} + \mathbf{L}_-\mathbf{L}_+^\mathsf{T}\right)\mathsf{W}$$

The backprojection lines $\mathbf{L}_\pm$ of an image point $\mathbf{q}$ are spanned by the origin (center of the sphere) and points $(\mathbf{b}_\pm^\mathsf{T}, 1)^\mathsf{T}$, thus $\mathbf{L}_\pm \sim \left(\mathbf{b}_\pm^\mathsf{T}\ \mathbf{0}^\mathsf{T}\right)^\mathsf{T}$. Here,

$$\mathbf{b}_\pm = (\mathbf{r}^\mathsf{T}\mathbf{r})\mathbf{C}_p + \left(\xi r_3 \pm \sqrt{\xi^2 r_3^2 - (\mathbf{r}^\mathsf{T}\mathbf{r})(\xi^2 - 1)}\right)\mathbf{r}$$

with $\mathbf{r} \sim \mathsf{K}^{-1}\mathbf{q}$ and $\mathbf{C}_p$ the center of the perspective camera (cf. section 3). The line complex C generated by the two lines, is

$$\mathsf{C} \sim \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{b}_+\mathbf{b}_-^\mathsf{T} + \mathbf{b}_-\mathbf{b}_+^\mathsf{T} \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \\ 0 & \xi^2(\mathbf{r}^\mathsf{T}\mathbf{r})\mathbf{e}_3\mathbf{e}_3^\mathsf{T} - \xi^2 r_3(\mathbf{e}_3\mathbf{r}^\mathsf{T} + \mathbf{r}\mathbf{e}_3^\mathsf{T}) + (\xi^2 - 1)\mathbf{r}\mathbf{r}^\mathsf{T} \end{pmatrix}$$

where $\mathbf{e}_3 = (0,0,1)^\mathsf{T}$. C is by construction symmetric and of rank 2 and it has 9 non-zero coefficients. Let M be the lower right 3×3 submatrix of C and the 6-vector $\mathbf{m}$ its vectorized version: $\mathbf{m} = \mathbf{v}_{sym}(\mathsf{M})$. We have the following linear backprojection equation:

$$\mathbf{m} \sim \mathsf{B}_\xi \hat{\mathbf{r}} = \mathsf{B}_\xi \hat{\mathsf{K}}^{-1} \hat{\mathbf{q}} \tag{5}$$

with

$$\mathsf{B}_\xi = \begin{pmatrix} \xi^2-1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi^2-1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \xi^2-1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ \xi^2 & 0 & \xi^2 & 0 & 0 & -1 \end{pmatrix}$$

We call $\mathsf{B}_{cata} = \mathsf{B}_\xi \hat{\mathsf{K}}^{-1}$ the **backprojection matrix**.

6 Projection of Line Complexes and Quadrics

To the best of our knowledge the projection of general quadric surfaces in catadioptric cameras has never been studied. The existing literature concerns only the projection of spheres, for calibration purposes [23]. The problem can be conveniently addressed by considering a line-based representation of quadrics, via line complexes. The set of 3D lines tangent to a quadric form a line complex [21]. A conic on a 3D scene plane can also be represented by a line complex [24, 25], and the following results apply thus to the projection of both quadrics and conics. As discussed in the previous section, a line complex can be represented by a 6×6 symmetric matrix C. Let C be split in 3×3 blocks:

$$\mathsf{C} \sim \begin{pmatrix} \mathsf{U} & \mathsf{N} \\ \mathsf{N}^\mathsf{T} & \mathsf{M} \end{pmatrix}$$

The image of C consists of all points $\mathbf{q}$ such that at least one of their backprojection rays lies on C. Let $\mathbf{L}_\pm \sim (\mathbf{b}_\pm^\mathsf{T}\ \mathbf{0}^\mathsf{T})^\mathsf{T}$ be the Plücker coordinates of the two backprojections of $\mathbf{q}$ (cf. section 5). Hence, $\mathbf{q}$ lies on the image of C *iff*

$$(\mathbf{L}_+^\mathsf{T} \mathsf{C} \mathbf{L}_+)(\mathbf{L}_-^\mathsf{T} \mathsf{C} \mathbf{L}_-) = (\mathbf{b}_+^\mathsf{T} \mathsf{U} \mathbf{b}_+)(\mathbf{b}_-^\mathsf{T} \mathsf{U} \mathbf{b}_-) = 0$$

By replacing $\mathbf{b}_\pm$ with its definition (section 5) and developing this equation, we get the following constraint on the doubly lifted coordinates of $\mathbf{q}$:

$$\hat{\hat{\mathbf{q}}}^\mathsf{T} \underbrace{\hat{\hat{\mathsf{K}}}^{-\mathsf{T}} \mathsf{X}_{lc}}_{\mathsf{P}_{lc}} \hat{\mathbf{u}} = 0$$

where $\mathbf{u} = \mathbf{v}_{sym}(\mathsf{U})$, $\hat{\mathbf{u}}$ is a 21-vector and the 15×21 matrix X_{lc} only depends on the intrinsic parameter ξ and is highly sparse (not shown due to lack of space). Since the coefficients of $\hat{\hat{\mathbf{q}}}$ are 4th order monomials of $\mathbf{q}$, we conclude that a central catadioptric image of any line complex, and thus of any quadric or conic, is a quartic curve. We may call the 15×21 matrix P_{lc} the **line complex projection matrix** for catadioptric cameras. It maps the lifted coefficients of the line complex to the 15 coefficients of the quartic curve in the image.

7 The General Catadioptric Fundamental Matrix

We are now ready to derive an analytical expression for the fundamental matrix for any pair of catadioptric images. As suggested in section 2, we perform the following steps: *(i)* Backproject a point $\mathbf{q}$ from the first image to its two backprojection rays, represented by a line complex. *(ii)* Map the line complex from the coordinate system of the first camera, to that of the second one, via a rigid transformation (rotation and translation). *(iii)* Project the transformed line complex into the second camera.

We already know from the previous sections that the result has to be a *quartic epipolar curve* since it is the image of a line complex. In our case, the line complex is degenerate (the "envelope" of just two lines – the backprojection rays – not a full quadric). Hence, and as described in section 2, the quartic epipolar curve is indeed the union of two conics.

Let us now derive the full expression of the catadioptric fundamental matrix. The only remaining missing piece is that the backprojection of an image point (step *(i)*) gives the coefficients of a line complex, but that the projection of the line complex (step *(iii)*) requires its *lifted* coefficients. Hence, we need to insert that lifting between steps *(ii)* and *(iii)*.

Recall from section 5 that the backprojection line complex is obtained as:

$$\mathsf{C} \sim \begin{pmatrix} 0 & 0 \\ 0 & \mathsf{M} \end{pmatrix} \quad \text{with} \quad \mathbf{v}_{sym}(\mathsf{M}) = \mathsf{B}_{cata}\hat{\mathbf{q}}$$

The rigid transformation of step *(ii)* gives (cf. equation (1) in section 3)

$$\mathsf{C}' \sim \mathsf{T}^{-\mathsf{T}}\mathsf{C}\mathsf{T}^{-1} \sim \begin{pmatrix} [\mathbf{t}]_\times \mathsf{R}\mathsf{M}\mathsf{R}^\mathsf{T}[\mathbf{t}]_\times^\mathsf{T} & -[\mathbf{t}]_\times^\mathsf{T}\mathsf{R}\mathsf{M}\mathsf{R}^\mathsf{T} \\ -\mathsf{R}\mathsf{M}\mathsf{R}^\mathsf{T}[\mathbf{t}]_\times & \mathsf{R}\mathsf{M}\mathsf{R}^\mathsf{T} \end{pmatrix}$$

Recall from section 6 that the projection of a line complex, when expressed in the local camera coordinate system, only involves its upper left 3×3 submatrix:

$$\mathsf{U} = [\mathbf{t}]_\times \mathsf{R}\mathsf{M}\mathsf{R}^\mathsf{T}[\mathbf{t}]_\times^\mathsf{T} = \mathsf{E}\mathsf{M}\mathsf{E}^\mathsf{T}$$

where we encounter the well-known essential matrix $\mathsf{E} = [\mathbf{t}]_\times \mathsf{R}$.

Since U and M are symmetric, property (4) from section 3 allows to write the folllwing relation between their vectorized versions $\mathbf{u} = \mathbf{v}_{sym}(\mathsf{U}), \mathbf{m} = \mathbf{v}_{sym}(\mathsf{M})$:

$$\mathbf{u} = \hat{\mathsf{E}}_{6\times 6}\mathbf{m}$$

Finally, the required lifted coefficients of the line complex are obtained as:

$$\hat{\mathbf{u}} = \mathring{\mathsf{E}}_{21\times 21}\hat{\mathbf{m}}$$

The last remaining detail is to express $\hat{\mathbf{m}}$ in terms of the image point $\mathbf{q}$. From $\mathbf{m} = \mathsf{B}_{cata}\ \hat{\mathbf{q}}$, we deduce

$$\hat{\mathbf{m}} = \hat{\mathsf{B}}_{cata,21\times 15}\ \hat{\hat{\mathbf{q}}}$$

We can now introduce the **catadioptric fundamental matrix**:

$$\mathsf{F}_{cata,15\times 15} \sim \mathsf{P}_{lc}\ \mathring{\mathsf{E}}\ \hat{\mathsf{B}}_{cata} \tag{6}$$

We have already explained that $\mathsf{F}_{cata}\ \hat{\hat{\mathbf{q}}}$ gives a quartic epipolar curve. The **epipolar constraint** can thus be written as

$$\hat{\hat{\mathbf{q}}}_2^{\mathsf{T}}\ \mathsf{F}_{cata}\ \hat{\hat{\mathbf{q}}}_1 = 0$$

which has the familiar form known for perspective cameras. F_{cata} has rank 6 and its left/right null space has dimension 9. While a perspective view has a single epipole, in an omnidirectional view there are a pair of epipoles, $\mathbf{e}_+$ and $\mathbf{e}_-$, corresponding to the two antipodal intersections of the baseline with the sphere, cf. section 4. The nullspace of F_{cata} comprises the doubly lifted coordinates vectors of both epipoles. We conjecture that they are the only doubly lifted 3-vectors in the nullspace of F_{cata}, but this has to be proven in future work.

We have no space to discuss it, but for mixtures involving one hyper-catadioptric and one other camera, the size of F_{cata} is smaller (15×6 for para–hyper and 6×6 for hyper–perspective). Other special cases are already known [14, 5].

8 The General Catadioptric Plane Homography

We give an algebraic formulation of the different steps involved in the plane homography operation, as described in section 2. This section omits many details due to lack of space. Let $\Pi = (\mathbf{n}^{\mathsf{T}}, d)^{\mathsf{T}}$ be the plane and $\mathbf{q}_1$ a point in the first image. We start by backprojecting the point to the plane, giving two 3D points

$$\mathbf{Q}_\pm \sim \underbrace{\begin{pmatrix} d\mathtt{I}_3 \\ -\mathbf{n}^{\mathsf{T}} \end{pmatrix}}_{\mathsf{Y}_{4\times 3}} \mathbf{b}_\pm$$

with $\mathbf{b}_\pm$ as in section 5. We project them to the second image using the projection matrix given in section 4. To do so, we first have to lift the coordinates of these 3D points: $\hat{\mathbf{Q}}_\pm \sim \hat{\mathsf{Y}}_{10\times 6}\ \hat{\mathbf{b}}_\pm$. The projection then gives two dual conics in the second image (cf. section 4), represented by 6-vectors $\omega_\pm \sim \mathsf{P}_{cata}\ \hat{\mathsf{Y}}\ \hat{\mathbf{b}}_\pm$.

Let us compute the following symmetric 6×6 matrix:

$$\Gamma \sim \omega_+\omega_-^{\mathsf{T}} + \omega_-\omega_+^{\mathsf{T}} \sim \mathsf{P}_{cata}\hat{\mathsf{Y}} \underbrace{\left(\hat{\mathbf{b}}_+\hat{\mathbf{b}}_-^{\mathsf{T}} + \hat{\mathbf{b}}_-\hat{\mathbf{b}}_+^{\mathsf{T}}\right)}_{\mathsf{Z}} \hat{\mathsf{Y}}^{\mathsf{T}}\mathsf{P}_{cata}^{\mathsf{T}} \tag{7}$$

What does Γ represent? It is the "union" of two degenerate dual conics, ω_+ and ω_-. Hence, it represents a dual quartic curve (this can also be proven more formally). Further, each of the two dual conics represents two image points; it contains exactly the lines going through at least one of them. Hence, Γ contains exactly the lines going through at least one of the total of four considered image

points. These four points are nothing else than the four points explained in section 2, i.e. the possible four matches of $\mathbf{q}_1$.

The expression of the dual quartic curve represented by Γ can be written as $\hat{\hat{\mathbf{l}}}^{\mathsf{T}}\ \gamma = 0$, where $\hat{\hat{\mathbf{l}}}$ denotes the fourth order Veronese map of a generic line $\mathbf{l}$, and γ is a 15-vector containing sums of the 21 coefficients of Γ (this is analogous to the reduction explained in equation (2)).

When developing the expression for Z in (7), it can be seen that its coefficients are linear in $\hat{\hat{\mathbf{q}}}_1$. The coefficients of γ may thus be computed via a linear mapping of $\hat{\hat{\mathbf{q}}}_1$; that mapping is given by a 15×15 matrix:

$$\gamma \sim \mathsf{H}_{cata,15\times 15}\ \hat{\hat{\mathbf{q}}}_1$$

The matrix H_{cata} is the **catadioptric plane homography**. Its explicit form is omitted due to lack of space.

By the same approach as in section 4, we can derive the following constraint equation:

$$\widehat{\widehat{[\mathbf{q}_2]_\times}}\ \mathsf{H}_{cata}\ \hat{\hat{\mathbf{q}}}_1 = \mathbf{0}_{15}$$

Of the 15 constraints contained in the above equation, only five are linearly independent. Hence, in order to estimate the $15^2 = 225$ coefficients of H_{cata}, we need at least 45 matches.

In the special case of para-catadioptric cameras, the homography is of size 6×6 and each match gives 6 equations, 3 of which are linearly independent. Hence, 12 matches are needed to estimate the 36 coefficients of that plane homography.

9 Conclusions, Discussion, and Perspectives

Our motivation for this work is to get a complete picture of the imaging and multi-view geometry of catadioptric (and other) cameras. We have shown that the basic concepts – projection, backprojection, epipolar geometry, or plane homography – can all be written as (multi-) linear mappings. These results are first of all of conceptual value, and we consider them as theoretical contributions.

Concerning potential practical applications, we note that a linear estimation of the catadioptric fundamental matrix requires 224 matches... We thus do not currently believe that F_{cata} will be of practical use.

However, the catadioptric projection matrix and the plane-to-image homography described in section 4 may indeed prove useful for calibrating catadioptric and possibly other omnidirectional cameras. We show this by an illustrative experiment, cf. Figure 2. Corner extraction for calibration grids, despite being trivial for perspective cameras, is still problematic for images with strong non-linear distortions [26]. In the perspective case we typically indicate the area of interest by manually clicking 4 corners; they enable the estimation of an homography and the projection of the grid into the image. The final position of all corners is accurately determined by refining the initial estimate using image processing techniques. Such a procedure has not been possible until now for non-conventional imagery with non-perspective distortions.

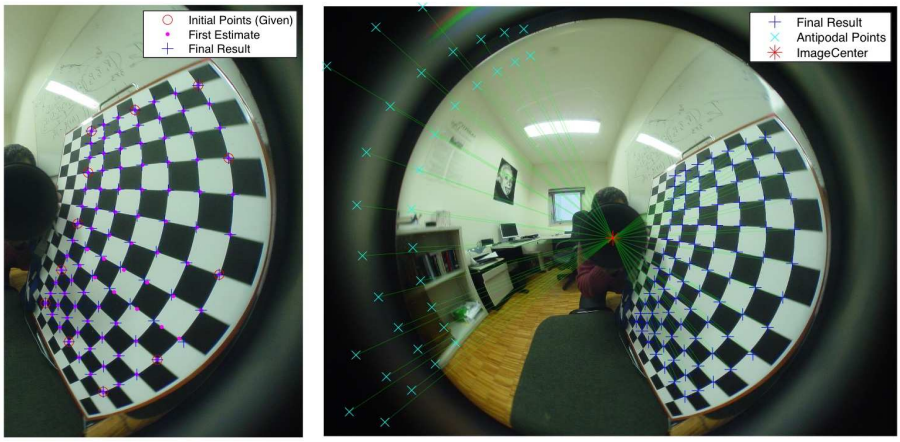

Fig. 2. Estimation of the homography mapping a planar grid into a catadioptric image. It was determined from 12 clicked points (left side). Each corner is mapped into a pair of image projections. The lines joining corresponding pairs form a pencil going through the principal point which confirms the correctness of the estimation.

From section 4 it follows that the homography from a plane to any catadioptric image is represented by a 6×6 matrix H. We estimated it from the required minimum of 12 manually selected matches by the DLT procedure suggested in section 4 (left side of Fig. 2). All corners were then projected into the image and refined using a corner detector. From this initial step only 7 out of 91 points were missed. The procedure was repeated a second time using all the good points and 6 more points were correctly detected. The estimated homography maps each plane point to a pair of antipodal image points (right side of Fig. 2). The shown result suggests that the plane-to-image homography can be well estimated and that it is useful for extracting and matching corners of a planar grid. Current work deals with calibrating the camera from such homographies, from multiple images of the grid.

There are several perspectives for our work. The shown results can be specialized to e.g. para-catadioptric cameras, leading to simpler expressions. It may also be possible that due to the coordinate liftings used, some of the results hold not only for catadioptric cameras, but also for other models, e.g. classical radial distortion; this will be investigated. Current work is concerned with developing practical DLT like calibration approaches for catadioptric cameras, using 3D or planar calibration grids. Promising results have already been obtained.

Acknowledgements. João P. Barreto is grateful to the Portuguese Science Foundation by generous funding through grant PTDC/EEA-ACR/68887/2006. Peter Sturm acknowledges support by the French ANR project CAVIAR.

References

1. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. 2nd edn. Cambridge University Press (2004)
2. Svoboda, T., Pajdla, T.: Epipolar geometry for central catadioptric cameras. IJCV **49** (2002) 23–37
3. Kang, S.: Catadioptric self-calibration. CVPR (2000) 1201–1207
4. Geyer, C., Daniilidis, K.: Structure and motion from uncalibrated catadioptric views. CVPR (2001) 279–286
5. Sturm, P.: Mixing catadioptric and perspective cameras. OMNIVIS (2002) 37–44
6. Mičušík, B., Pajdla, T.: Structure from motion with wide circular field of view cameras. PAMI. **28**(7) (2006) 1135-1149
7. Claus, D., Fitzgibbon, A.: A rational function for fish-eye lens distortion. CVPR (2005) 213–219
8. Gupta, R., Hartley, R.: Linear pushbroom cameras. PAMI **19** (1997) 963–975
9. Feldman, D., Pajdla, T., Weinshall, D.: On the epipolar geometry of the crossed-slits projection. ICCV (2003) 988–995
10. Pajdla, T.: Stereo with oblique cameras. IJCV **47** (2002) 161–170
11. Seitz, S., Kim, J.: The space of all stereo images. IJCV **48** (2002) 21–38
12. Menem, M., Pajdla, T.: Constraints on perspective images and circular panoramas. BMVC (2004)
13. Baker, S., Nayar, S.: A theory of catadioptric image formation. ICCV (1998) 35–42
14. Geyer, C., Daniilidis, K.: Properties of the catadioptric fundamental matrix. ECCV (2002) 140–154
15. Barreto, J.: A unifying geometric representation for central projection systems. CVIU **103** (2006)
16. Barreto, J.P., Daniilidis, K.: Epipolar geometry of central projection systems using veronese maps. CVPR (2006) 1258–1265
17. Wolf, L., Shashua, A.: Two-body segmentation from two perspective views. CVPR (2001) 263–270
18. Vidal, R., Ma, Y., Soatto, S., Sastry, S.: Two-view multibody structure from motion. IJCV **68** (2006) 7–25
19. Geyer, C., Daniilidis, K.: A unifying theory for central panoramic systems. ECCV (2000) 445–461
20. Barreto, J., Araujo, H.: Geometric properties of central catadioptric line images and its application in calibration. PAMI **27** (2005) 1237–1333
21. Semple, J., Kneebone, G.: Algebraic Projective Geometry. Claredon Press (1998)
22. Horn, R., Johnson, C.: Topics in Matrix Analysis. Cambridge Univ. Press (1991)
23. Ying, X., Zha, H.: Using sphere images for calibrating fisheye cameras under the unified imaging model of the central catadioptric and fisheye cameras. ICPR (2006) 539–542
24. Ponce, J., McHenry, K., Papadopoulo, T., Teillaud, M., Triggs, B.: On the absolute quadratic complex and its application to autocalibration. CVPR (2005) 780–787
25. Valdés, A., Ronda, J., Gallego, G.: The absolute line quadric and camera autocalibration. IJCV **66** (2006) 283–303
26. Mei, C., Rives, P.: Single view point omnidirectional camera calibration from planar grids. ICRA (2007) 3945–3950