Structure and Motion from Uncalibrated Catadioptric Views

Christopher Geyer and Kostas Daniilidis*
GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104
{cgeyer,kostas}@seas.upenn.edu

Abstract

In this paper we present a new algorithm for structure from motion from point correspondences in images taken from uncalibrated catadioptric cameras with parabolic mirrors. We assume that the unknown intrinsic parameters are three: the combined focal length of the mirror and lens and the intersection of the optical axis with the image. We introduce a new representation for images of points and lines in catadioptric images which we call the circle space. This circle space includes imaginary circles, one of which is the image of the absolute conic. We formulate the epipolar constraint in this space and establish a new 4×4 catadioptric fundamental matrix. We show that the image of the absolute conic belongs to the kernel of this matrix. This enables us to prove that Euclidean reconstruction is feasible from two views with constant parameters and from three views with varying parameters. In both cases, it is one less than the number of views necessary with perspective cameras.

1. Introduction

During the last 10 years there has been a considerable effort in studying the reconstruction of scenes from uncalibrated perspective views given point correspondences. This is considered now a thoroughly understood problem. Solutions and insights gained from these studies boosted applications in video processing and image based rendering. Two books [10] and [5] contain comprehensive treatments of the subject.

In the meantime, computer vision researchers realized that perspective cameras are just one modality among many. Motivated by the need for a panoramic field of view, catadioptric cameras have been designed and can be already purchased off-the-shelf. For an extensive coverage the reader is referred to the recent book by Benosman and Kang [2] and the proceedings of the Workshop for Omnidirectional Vision [4]. Among several designs, the catadioptric systems with a single effective viewpoint, called central catadioptric, attracted special attention due to their elegant and useful geometric properties. Several authors have studied the properties of central catadioptric cameras and the image formation in them [15, 20, 3, 22, 12, 8]. Kang [12] proposed a single view approach from the image of the circular mirror boundary of a paraboloid mirror. Geyer and Daniilidis showed [7, 8] how calibration of a parabolic catadioptric system can be achieved from a single view of three lines in space or from a single view of two sets of parallel lines.

In this paper, we study the recovery of motion and scene structure from *multiple* parabolic catadioptric views. Such views can be obtained from a reflective surface of revolution of parabolic profile and an orthographic lens. We assume that the optical axes of the lens and the mirror are parallel. They do not have to coincide but to avoid aberrations and enable maximal coverage of the CCD-chip they should be close to each other. We assume, thus, that the catadioptric system is correctly aligned. We further assume that the aspect ratio and skew parameter are known leaving only the focal length (combined scaling factor of mirror, lens, and CCD-chip) and the image center (intersection of the optical axis with the image plane) as unknown.

It is already known that in such parabolic catadioptric systems lines project onto circles. We introduce a new representation for circles in the image plane: the circle space of three dimensions. This space is divided into two parts by an abstract paraboloid. The exterior of the paraboloid represents all circles with real radius and the interior all circles with imaginary radius. The space does not contain circles with complex radii but the paraboloid itself represents all circles with zero radius which are just points on the plane. By lifting each image point to a point of the paraboloid and each image circle to a point outside the paraboloid we have one space for both points and circles.

The fact that we can represent imaginary circles enables us to represent the image of the absolute conic. In the calibrated case, the image of the absolute conic is the focus of the abstract paraboloid in the circle space. In the non-calibrated case, the imaginary image of the absolute conic is a point inside the abstract paraboloid that is vertically symmetric to the point representing the real image of the fronto-parallel horizon.

We formulate the calibration problem as the question for a linear transformation that will map uncalibrated points on the abstract paraboloid to "calibrated" points on a paraboloid and the image of the absolute conic to its focus. Indeed, such a linear transformation K exists and encodes all three intrinsic parameters (focal length and image center). The question is now to find this mapping from multiple views.

It turns out that we can formulate the epipolar constraint using projective coordinates of the circle space we have been working on. A new 4×4 "catadioptric" fundamental matrix is composed from the essential matrix E and an induced projection following the mapping K above. We prove that the circle representation of the images of the absolute conic in the left and the right view respectively lie in the left and right nullspaces of the catadioptric fundamental matrix. Because the catadioptric fundamental matrix is rank 2, the image of the absolute conic is in the intersection of the left and right nullspace if the intrinsic parameters are constant and rotation does not vanish and is not about the translation direction. For three views, it is even possible to determine the image of the three different absolute conics in the case of varying intrinsics.

Thus, the main result of this paper is that, with unknown focal length and image center, Euclidean reconstruction from parabolic catadioptric views is feasible:

1. From two views with the same camera parameters.

2. From three views with varying camera parameters.

In both cases, it is one view less, than in the case of perspective views with the same unknowns (focal length and image center): Three views are necessary for constant parameters [14, 13] and four views are necessary for varying parameters [11]. In particular the fundamental matrix has seven degrees of freedom whereas the intrinsics have three and the motion has five for a total of eight. In the three view case the trifocal tensor has 19 free parameters whereas the three intrinsincs have nine plus 11 for the motion, yielding a total of 20.

We are not going to review here the vast amount of literature on uncalibrated Euclidean reconstruction which has been comprehensively summarized in the two recent books [10, 5]. The main result [14] is that three views suffice for Euclidean reconstruction with all intrinsics unknown but constant. The results still hold for known aspect ratio and skew. Hartley [10] showed that a varying focal length can be recovered from two views with all other intrinsic parameters fixed. Sturm [18] studied the degenerate configurations for the same assumption. Heyden and Astrom [11] proved that four views suffice for unknown varying focal length and image center but known aspect ratio and skew. Pollefeys et al. [17] studied several configurations of unknown and varying parameters.

In the omnidirectional vision literature, there are very few approaches dealing with structure from motion. Gluckman and Nayar [9] studied ego-motion estimation by mapping the catadioptric image to the sphere. Svoboda et al [20] first established the epipolar geometry for all central catadioptric systems. Kang [12] proposed a direct self-calibration by minimizing the epipolar constraint. Fermueller and Aloimonos [6] proved the superiority of the sphere over the plane regarding stability. Teller [1] showed how to compute ego-motion from spherical mosaics. Multiple view algorithms for the perspective case which assume piecewise planar environments are simpler when modified for catadioptric imagery. [21, 19].

In the next section we mention introductory facts about catadioptric geometry. We introduce the notion of circle space and we find the image of the absolute conic on that space. We finish the second section with the recovery of the image of the absolute conic from the catadioptric fundamental matrix. In the third section we present reconstruction algorithms for two and three views. In the fourth section a real experiment is described.

2. Preparations

2.1. Known Facts

We recall from [7] some facts about the projection induced by parabolic mirror.

Fact 1. In a coordinate system whose origin is the focus of the paraboloid and axis of symmetry coincides with the z-axis, the projection of a space point $(x, y, z, 1)$ is in image coordinates:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} c_x + \frac{2fx}{-z + \sqrt{x^2 + y^2 + z^2}} \\ c_y + \frac{2fy}{-z + \sqrt{x^2 + y^2 + z^2}} \\ 1 \end{pmatrix}, \qquad (1)$$

where f is the combined focal length of the mirror and camera, and (c_x, c_y) is the image center, the intersection of the axis of the parabola with the image plane. We assume that the aspect ratio is 1 and that there is no skew. The image point is obtained by intersecting the ray through the focus and the space point with the parabola, then orthographically projecting the intersection to a plane perpendicular to the axis of the paraboloid.

Fact 2. The horizon of the fronto-parallel plane, the plane perpendicular to the axis of the paraboloidal mirror, is the circle

$$(c_x - u)^2 + (c_y - v)^2 = 4f^2. \qquad (2)$$

This circle of radius $2f$ centered about the image center is the equivalent of the calibrating conic which we call ω' since we call the image of the absolute conic ω.
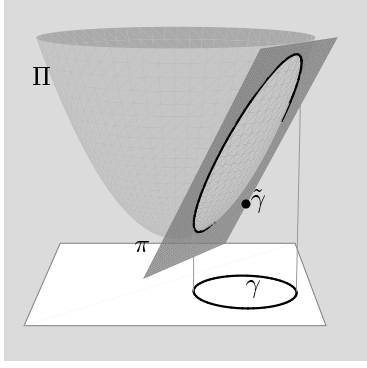
Figure 1. A circle γ is represented by the point $\tilde{\gamma}$. The plane π is the polar plane of $\tilde{\gamma}$ with respect to Π. γ is obtained by projecting the intersection of π with Π to the plane.

Fact 3. The projection of a line is an arc of a circle. If ξ is the center and R the radius of the circle and if $d^2 = (c_x - \xi_x)^2 + (c_y - \xi_y)^2$ then

$$4f^2 + d^2 = R^2. \tag{3}$$

This condition is equivalent to the condition that the circle intersect ω' antipodally.

Fact 4. The image ω of the absolute conic Ω_∞ is the circle

$$(c_x - u)^2 + (c_y - v)^2 = -4f^2, \tag{4}$$

centered at the image center with radius $2if$. This can be derived by solving for x and y in the projection formula (1) after substituting $x^2 + y^2 + z^2 = 0$,

$$x = \frac{(u - c_x)z}{2f} \qquad y = \frac{(v - c_y)z}{2f}.$$

Substitute the right hand sides into $x^2 + y^2 + z^2 = 0$, obtaining

$$\frac{z^2}{4f^2}\left(4f^2 + (c_x - u)^2 + (c_y - v)^2\right) = 0.$$

Dividing by $z^2/4f^2$ leaves (4). Thus, knowledge of either the absolute conic or the calibrating conic yields the intrinsic parameters.

2.2. Parabolic Circle Space

In the next few paragraphs we consider an abstract paraboloid which is different from the physical paraboloid of the mirror. Following Pedoe [16], we use this surface to describe a correspondence between points in space and circles in the plane. Lines in this circle space correspond to one parameter systems of coaxial circles. Planes in the space correspond to two parameter systems of circles which intersect a single circle antipodally. See Figure 1 in which a circle is obtained from a point in space by taking the polar of the point with respect to the paraboloid, and projecting to a plane the intersection of the polar plane with the paraboloid; this projection will be a circle.

We call the paraboloid Π; it is given by the quadratic form

$$C_\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{4} \end{pmatrix}. \tag{5}$$

Its focus is at the origin and has a focal length equal to $\frac{1}{4}$. So,

$$\Pi = \left\{p : p^T C_\Pi\, p = 0\right\} = \left\{\left(x, y, x^2 + y^2 - 1/4, 1\right)^T\right\}.$$

Definition. *Suppose γ is the circle centered at (p, q) with radius R:*

$$(p - x)^2 + (q - y)^2 = R^2, \tag{6}$$

where R is possibly zero or imaginary, but never complex. Let the point representation *of γ be the the projective point*

$$\tilde{\gamma} = \left(p, q, p^2 + q^2 - R^2 - \frac{1}{4}, 1\right)^T. \tag{7}$$

Note that the circle's radius is real iff it lies outside of Π. Its radius is imaginary iff it lies inside of (above) Π. If $R = 0$ then γ is a single point and $\tilde{\gamma}$ lies on Π. The set of points $\{\tilde{\gamma}\}$ is the parabolic circle space.

When γ is a point, because $\tilde{\gamma}$ has the same x and y coordinates as γ but lying on Π, we say that $\tilde{\gamma}$ is the lifting of γ to Π.

Proposition. *If π is the polar plane of the point $\tilde{\gamma}$ with respect to the paraboloid Π, the orthographic projection in the direction of the z-axis of the intersection of π with Π is the circle γ.*

Proof: The implicit equation of the polar plane π of $\tilde{\gamma}$ is

$$\begin{aligned} 0 &= \tilde{\gamma}^T C_\Pi \begin{pmatrix} x & y & z & 1 \end{pmatrix}^T \\ &= \frac{1}{2}\left(-\frac{1}{4} - p^2 - q^2 + 2px + 2q - z\right). \end{aligned}$$

Substitute $z = x^2 + y^2 - 1/4$, yielding (6). $\square$

Therefore the point $(p, q, r, 1)$ represents the circle

$$(p - x)^2 + (q - y)^2 = p^2 + q^2 - r - \frac{1}{4}. \tag{8}$$

We can extend the definition to encompass lines as well; they are represented by points on the plane at ∞. The polar plane of a point $(p, q, r, 0)$ at infinity is the plane

$$0 = -\frac{r}{2} + px + qy \, ,$$

which is independent of z and so the line in the plane has the same equation.

2.3. Application of Circle Representation

First, note the point representations of the calibrating conic,

$$\tilde{\omega}' = \left(c_x, c_y, c_x^2 + c_y^2 - 4f^2 - \frac{1}{4}, 1 \right)^T , \qquad (9)$$

which, because it has a real radius, lies outside of Π; and the absolute conic,

$$\tilde{\omega} = \left(c_x, c_y, c_x^2 + c_y^2 + 4f^2 - \frac{1}{4}, 1 \right)^T , \qquad (10)$$

which, because it has an imaginary radius, lies inside of Π. The points $\tilde{\omega}$ and $\tilde{\omega}'$ lie the same vertical distance, $4f^2$, away from Π.

Proposition. *The point representations of circles which are images of lines in a parabolic projection lie in a plane whose pole with respect to Π is $\tilde{\omega}$.*

Proof: If $(p, q, r, 1)$ is a circle which is the parabolic projection of a line it must satisfy (3). Using (8),

$$4f^2 + (c_x - p)^2 + (c_y - q)^2 = p^2 + q^2 - r - \frac{1}{4}$$

$$4f^2 + c_x^2 + c_y^2 + \frac{1}{4} = 2pc_x + 2qc_y - r, \quad (11)$$

which, in the variables p, q, and r, is the equation of a plane. This plane is represented by the row vector

$$\pi = \left(c_x, c_y, -1/2, -2f^2 - c_x^2/2 - c_y^2/2 - 1/8 \right) .$$

The point

$$C_\Pi^{-1} \pi^T = \left(c_x, c_y, c_x^2 + c_y^2 + 4f^2 - 1/4, 1 \right) = \tilde{\omega} \, ,$$

is the pole of the plane π. $\qquad \square$

The paraboloid Π was defined so that its focus is the origin. The point $\tilde{\omega}$ is located at the origin when $c_x, c_y = 0$ and $f = \frac{1}{4}$. The polar plane of this point (11) reduces to $r = -\frac{1}{2}$. In this case, image points lifted to the parabola exactly correspond to calibrated rays. When these intrinsics hold, the lifting of a space point projected by formula (1) is a point on the parabola which is collinear with the focus and

the point in space. In particular, the projection of the point $(x, y, z, 1)^T$ in space is

$$\left(\frac{\frac{1}{2}x}{-z + \sqrt{x^2 + y^2 + z^2}}, \frac{\frac{1}{2}y}{-z + \sqrt{x^2 + y^2 + z^2}}, 1 \right)^T ,$$

according to (1). The lifting of this point is,

$$\begin{pmatrix} \frac{\frac{1}{2}x}{-z+\sqrt{x^2+y^2+z^2}} \\ \frac{\frac{1}{2}y}{-z+\sqrt{x^2+y^2+z^2}} \\ \frac{\frac{1}{4}x^2+\frac{1}{4}y^2}{(-z+\sqrt{x^2+y^2+z^2})^2} - \frac{1}{4} \\ 1 \end{pmatrix} \propto \begin{pmatrix} x \\ y \\ z \\ -2z + 2\sqrt{x^2 + y^2 + z^2} \end{pmatrix}$$

which lies on the line through the focus and the point $(x, y, z, 1)^T$.

Is there a linear transformation which transforms point representations of uncalibrated image points, in which $\tilde{\omega}$ is in general position, to calibrated rays, in which $\tilde{\omega}$ is the origin? In the next section we show that this is indeed the case.

2.4. Transformations Fixing Π

In this section we find linear transformations under which Π is invariant. The four transformations,

$$R_\theta = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; S_\alpha = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha^2 & \frac{\alpha^2-1}{4} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$T_\tau = \begin{pmatrix} 1 & 0 & 0 & \tau_x \\ 0 & 1 & 0 & \tau_y \\ 2\tau_x & 2\tau_y & 1 & \tau_x^2 + \tau_y^2 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \; H = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ,$$

are such that for any choice of θ, α, and vectors τ,

$$R_\theta^T C_\Pi R_\theta \propto C_\Pi \, , \qquad S_\alpha^T C_\Pi S_\alpha \propto C_\Pi \, ,$$
$$T_\tau^T C_\Pi S_\tau \propto C_\Pi \, , \qquad H^T C_\Pi H \propto C_\Pi \, ,$$

where C_Π was previously defined in (5) and is the quadratic form of Π. Therefore these transformations affect the parabolic circle space such that they take points to points, as opposed to say points to circles. The transformations have the following effect on points in the image plane: R_θ induces a rotation of θ about the origin; S_α induces a scale of α also about the origin; T_τ translates points by τ; and H reflects about the line $x = 0$.

Any composition of these transformations will also leave Π invariant. Note that these transformations also leave π_∞ invariant. They are therefore affine transformations, and also they send lines to lines.

These transformations act as similarity transformations on the points. Do they change the image of the absolute conic and the line image plane so as to correctly reflect the transformation induced on the points? In other words, say

c_x, c_y, and f are fixed, applying T_τ would induce a translation of τ on points; it should therefore transform $\tilde{\omega}$ into

$$\left(c_x + \tau_x, c_y + \tau_y, (c_x + \tau_x)^2 + (c_y + \tau_y)^2 + 4f^2 - \tfrac{1}{4}, 1\right)^T ,$$

and the line image plane (11) to

$$2p(c_x + \tau_x) + 2q(c_y + \tau_y) - r = (c_x + \tau_x)^2 + (c_y + \tau_y)^2 + \frac{1}{4} ,$$

so that the new image center is $(c_x + \tau_x, c_y + \tau_y)$ as desired; any rotation or scaling should act similarly. One can verify that all four transformations transform $\tilde{\omega}$ and the line image plane in a manner consistent with the way in which the transformations affect points.

Thus, there is a linear transformation taking point representations of image points obtained from a camera with intrinsic parameters c_x, c_y, and f, to calibrated rays. This transformation is the 4×4 matrix,

$$K = S_{\frac{1}{4f}} T_{(-c_x, -c_y)} \qquad (12)$$

$$= \begin{pmatrix} 4f & 0 & 0 & -4c_x f \\ 0 & 4f & 0 & -4c_y f \\ -2c_x & -2c_y & 1 & \frac{1}{4} + c_x^2 + c_y^2 - 4f^2 \\ 0 & 0 & 0 & 16f^2 \end{pmatrix} . \qquad (13)$$

This is an important point, for if $q = (u, v, 1)^T$ is the parabolic projection (with intrinsics c_x, c_y, f) of the space point $p = (x, y, z, 1)^T$ then for some scalar λ,

$$p = \lambda K \left(u, v, u^2 + v^2 - \frac{1}{4}, 1 \right)^T .$$

Implying that if

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} ,$$

then

$$PK \left(u, v, u^2 + v^2 - \frac{1}{4}, 1 \right)^T \propto \left(\frac{x}{z}, \frac{y}{z}, 1 \right) \qquad (14)$$

which is the **perspective projection of** $(x, y, z, 1)$ with image center $(0, 0, 1)$ and focal length $f = 1$. Note that K is different from the usual camera matrix: it is not actually a projection; P induces the projection. Leaving K non-singular (i.e. not incorporating P) will make it easier to prove that a matrix, a fundamental matrix, created with it has a certain rank.

2.5. The Catadioptric Fundamental Matrix

Let m and n be calibrated rays pointing to the same point $(x, y, z, 1)$ in space taken from two views related by a rotation R and translation t. The points m and n must satisfy the epipolar constraint which is specified by

$$n^T [t]_\times R m = n^T E m = 0 , \qquad (15)$$

where $E = [t]_\times R$ is called the essential matrix. Say $p = (u_1, v_1, 1)^T$ and $q = (u_2, v_2, 1)^T$ are two parabolic catadioptric projections of the space point, and say the camera matrices are K and K', with $\tilde{\omega}$ and $\tilde{\omega}'$ the point representations of the image of the absolute conic. If $\tilde{p}$ and $\tilde{q}$ are their liftings to Π, then using equation (14), so that $m = PK\tilde{p}$ and $n = PK'\tilde{q}$, the epipolar constraint (15) becomes,

$$\tilde{q}^T K'^T P^T E P K \tilde{p} = 0 . \qquad (16)$$

Let the 4×4 matrix

$$F = K'^T P^T E P K \qquad (17)$$

be called the *catadioptric fundamental matrix*. Then the epipolar constraint for parabolic catadioptric cameras is

$$\tilde{q}^T F \tilde{p} = 0 . \qquad (18)$$

Theorem. *The catadioptric fundamental matrix defined in (17) has rank* 2. *Let* $\tilde{\omega}_1$ *be the point representation of the image of the absolute conic in the first image, corresponding to* K, *and similarly for* $\tilde{\omega}_2$ *corresponding to* K' *in the second image. Then,*

$$\tilde{\omega}_2 F = 0 \quad and \quad F \tilde{\omega}_1 = 0 . \qquad (19)$$

Proof: The essential matrix E is known to be of rank 2, thus $P^T E P = \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}$ has rank 2. Since K and K' are non-singular then F must also have rank 2. Let us calculate the left and right null vectors of F. First, let t and t' be the images of the viewpoints from each camera,

$$t'^T E = 0 , \text{ and } E t = 0 .$$

Then by inspection, linearly independent left and right null vectors of $P^T E P$ are

$$f_1 = \left(t'^T \quad 0 \right) , \quad f_2 = \left(0 \quad 0 \quad 0 \quad 1 \right) \quad \text{and}$$

$$f_1' = \left(t \quad 0 \right)^T , \quad f_2' = \left(0 \quad 0 \quad 0 \quad 1 \right)^T .$$

Hence $g_{i=1,2} = K^{-1} f_i$ are vectors spanning the right nullspace of F and $g_{i=1,2}' = f_i'^T K'^{-T}$ are vectors spanning the left nullspace. Note that $g_2 = \tilde{\omega}_1$ and $g_2' = \tilde{\omega}_2^T$. Therefore,

$$\tilde{\omega}_2^T F = 0 \quad \text{and} \quad F \tilde{\omega}_1 = 0 .$$

Corollary. *If* $K = K'$ *and* $t \neq t'$ *then,*

$$\ker F \cap \ker F^T = \{\lambda \tilde{\omega}\} .$$

The condition $t \neq t'$ is true when the rotation is not trivial and when the axis of rotation is not the translation vector.

3. Algorithm

The algorithm proceeds in three steps. First estimate the fundamental matrix, from the fundamental matrix extract the intrinsic parameters via the image of the absolute conic, and reconstruct using well known perspective methods.

3.1. Estimating F

We use a non-linear method to estimate F. An algorithm based on singular value decomposition which is similar to the the 8-point algorithm for the perspective case exists for parabolic catadioptric projections but is equally sensitive.

1. Obtain images $p_{i,j} = (u_{i,j}, v_{i,j}, 1)^T$ of the same point $q_{j=1,\dots,n}$ in space in two catadioptric views $i = 1, 2$. Let

$$\tilde{p}_{i,j} = \left(u_{i,j}, v_{i,j}, u_{i,j}^2 + v_{i,j}^2 - \frac{1}{4}, 1 \right)^T .$$

2. Minimize the sum of first-order geometric errors,

$$\sum_j \frac{(\tilde{p}_{2,j} F \tilde{p}_{1,j})^2}{(F\tilde{p}_{2,j}) \circ (F\tilde{p}_{2,j}) + (F^T\tilde{p}_{1,j}) \circ (F^T\tilde{p}_{1,j})} ,$$

where the minimization is over F and using the notation $p \circ q = p \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} q$. F is parameterized as in:

$$\begin{pmatrix} a & b & \alpha a + \beta b & \gamma a + \delta b \\ c & d & \alpha c + \beta d & \gamma c + \delta d \\ e & f & \alpha e + \beta f & \gamma e + \delta f \\ g & h & \alpha g + \beta h & \gamma g + \delta h \end{pmatrix} ,$$

where one of $a, \dots, f$ is held constant at 1. This ensures that F has rank 2. Initial estimates for F can be obtained using the singular value decomposition method since the components of F are linear in coordinates of the lifted image points.

3.2. Estimating ω

In the case where $K = K'$ the left and right nullspaces of F contain the point representation of the image of the absolute conic. In the presence of noise the nullspaces will not intersect. Once we have calculated the two-dimensional nullspaces, we choose the point equidistant to the two lines as the estimate of $\tilde{\omega}$.

When the intrinsics vary and we have images from three views, with three matrices $K_{i=1,2,3}$ and point representations $\tilde{\omega}_{i=1,2,3}$, we then have

$$\begin{aligned} F_{12} &= K_2^T P^T E_{12} P K_1 , \\ F_{23} &= K_3^T P^T E_{23} P K_2 , \\ F_{31} &= K_1^T P^T E_{31} P K_3 . \end{aligned}$$

Then once we have estimated the three fundamental matrices we calculate say $\tilde{\omega}_1$ from the fact that,

$$\ker F_{12} \cap \ker F_{31}^T = \{\tilde{\omega}_1\} .$$

Again, the estimate of $\tilde{\omega}_1$ is the point equidistant to the two nullspaces.

3.3. Reconstruction

Reconstruction proceeds as in the calibrated perspective case. Once we have determined $\tilde{\omega}$ and consequently ω, we can transform the image points into calibrated rays with which we determine the essential matrix E using a non-linear optimization and then back-project the rays into space using a linear algorithm, both algorithms described in [10].

4. Experiments

We use the algorithm to perform a reconstruction of a scene from two views. The two pictures in Figure 2 are of a building on the campus of our institution and are assumed to have the same intrinsic parameters. First we manually choose and correspond points in the two images. We calculate the fundamental matrix F between the two views from the point correspondences using the algorithm described in the previous section. We estimate the point representation of the image of the absolute conic by finding the left and right nullspaces of F and finding the point equidistant to each. Using the intrinsic parameters we back-project the image points to calibrated rays. Using the calibrated rays we estimate the essential matrix E, decompose E into translation and rotation, and determine the perspective camera projection matrices P_1 and P_2. We then back-project the rays and use homogeneous linear triangulation to estimate scene points.

The reconstruction is shown in the top and bottom of Figure 3. In the reconstruction we have fitted a plane to the points on the front facade of the building and to points on the ground plane, these are highlighted in Figure 2 and shaded differently in Figure 3. The viewpoints and poses are also displayed in the figures. The triangulation is manually added and shown for visualization purposes only. The ground plane and front facade were reconstructed to almost planar surfaces and are close to perpendicular. The other facade of the building, on the left in the images, did not reconstruct true to the scene, this is because this plane is perpendicular to the axis of motion which makes estimating depth more error-prone. In two views with such small motion, the reconstruction performs remarkably well.
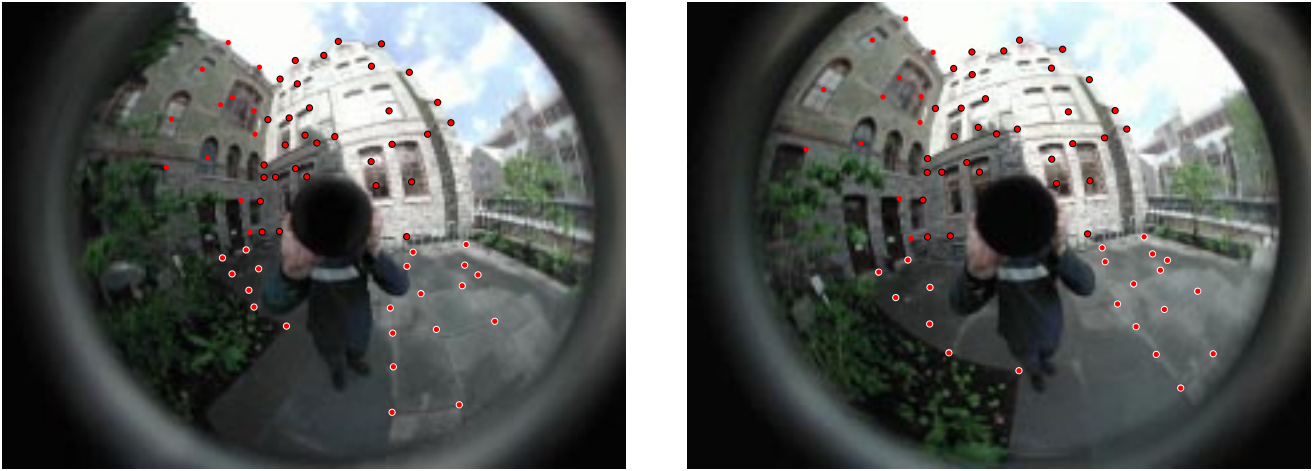
Figure 2. Two images taken with the same parabolic catadioptric camera. Points are those used for correspondence. Points highlighted in white are on the ground plane; points highlighted in black are on one side of the building facade.
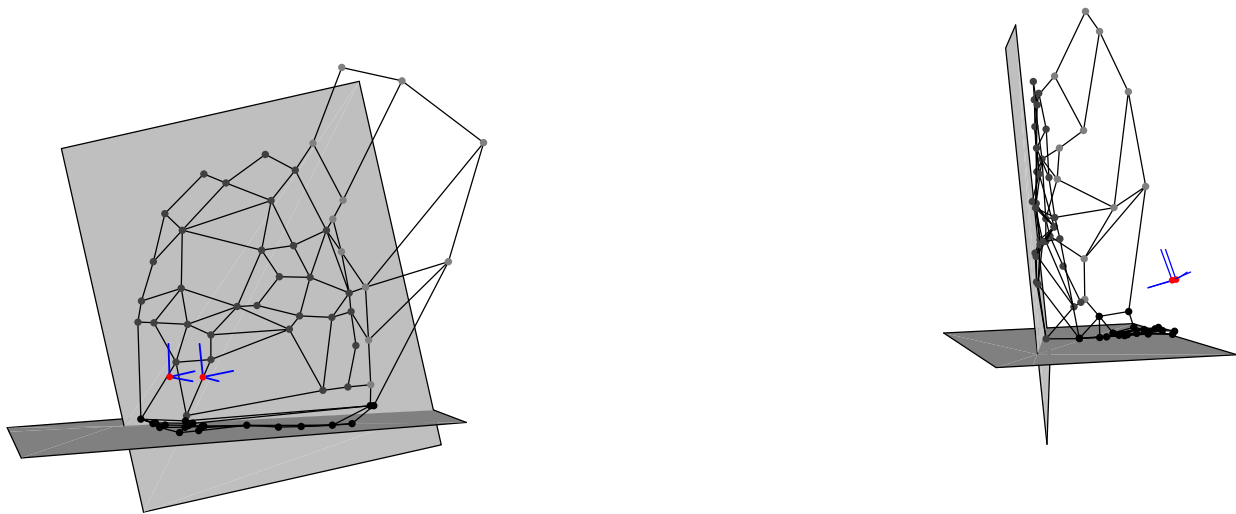

Figure 3. Reconstruction from two images. Black points are in the ground plane. Darkly shaded points are on the front facade of the building; lightly shaded points are on the other facade (which is on the left in the images). Planes are fitted to the facade and ground plane (and translated slightly so points are made visible). The coordinate systems at the points are the pose estimates. Tilt of the fitted plane is irrelevant to the results of the reconstruction. The top view is taken looking straight at the front facade; the bottom view is from the side. Note that the mirror reverses the orientation; this has been accounted for in the reconstruction.

5. Conclusion

We have established a new representation for images of lines and points in parabolic catadioptric cameras. Based on this representation we found a natural representation for the image of the absolute conic if aspect ratio and skew are assumed known. Writing the epipolar constraint in this new space yields a new catadioptric fundamental matrix. It turns out that the image of the absolute conic belongs to the two-dimensional kernel of this matrix. Applying thus only subspace recovery and intersection we can obtain Euclidean reconstructions:

- from two views with the same camera

- from three views with three different cameras.

The corresponding minimal views for the perspective case are three and four, respectively. This approach opened new questions which we address in our current work: What is the number of independent conditions on F to be decomposable? What is the degree of the manifold of all catadioptric fundamental matrices? Which point configurations make the recovery of the fundamental matrix degenerate? What is the minimal number of points for directly computing motion and the intrinsics?

Sensor resolution of commercial catadioptric cameras is increasing every year. We believe that geometrically intuitive algorithms working directly on catadioptric images can provide flexible solutions for panoramic image-based rendering and visualization.

References

[1] M. Antone and S. Teller. Automatic recovery of relative camera rotations in urban scenes. In *IEEE Conf. Computer Vision and Pattern Recognition*, Hilton Head Island, SC, June 13-15, 2000.

[2] R. Benosman and S.B. Kang. *Panoramic Vision*. Springer-Verlag, 2000.

[3] A. Bruckstein and T. Richardson. Omniview cameras with curved surface mirrors. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 79–86, 2000. originally published as Bell Labs Technical Memo, 1996.

[4] K. Daniilidis, editor. *IEEE Workshop on Omnidirectional Vision*, Hilton Head Island, SC, June 12, 2000.

[5] O. Faugeras, Q.-T. Luong, and T. Papadopoulo. *The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. MIT Press, 2001.

[6] C. Fermüller and Y. Aloimonos. Ambiguity in structure from motion: Sphere vs. pla. *International Journal of Computer Vision*, 28:137–154, 1998.

[7] C. Geyer and K. Daniilidis. Catadioptric camera calibration. In *Proc. Int. Conf. on Computer Vision*, pages 398–404, Kerkyra, Greece, Sep. 20-23, 1999.

[8] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems. In *Proc. Sixth European Conference on Computer Vision*, pages 445–462, Dublin, Ireland, 2000.

[9] J. Gluckman and S.K. Nayar. Ego-motion and omnidirectional cameras. In *Proc. Int. Conf. on Computer Vision*, pages 999–1005, Bombay, India, Jan. 3-5, 1998.

[10] R. Hartley and A. Zisserman. *Multiple View Geometry*. Cambridge Univ. Press, 2000.

[11] A. Heyden and K. Aström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 438–443, 1997.

[12] S.B. Kang. Catadioptric self-calibration. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages I–201–207, Hilton Head Island, SC, June 13-15, 2000.

[13] Y. Ma, S. Soatto, J. Kosecka, and S.S. Sastry. Euclidean reconstruction and reprojection up to subgroups. *International Journal of Computer Vision*, 38:217–227, 2000.

[14] S.J. Maybank and O.D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8:123–151, 1992.

[15] S. Nayar. Catadioptric omnidirectional camera. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 482–488, Puerto Rico, June 17-19, 1997.

[16] D. Pedoe. *Geometry: A comprehensive course*. Dover Publications, New York, NY, 1970.

[17] M. Pollyfeys, R. Koch, and L. van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. Int. Conf. on Computer Vision*, pages 90–95, Bombay, India, Jan. 3-5, 1998.

[18] P. Sturm. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. In *BMVC*, 1999.

[19] P. Sturm. A method for 3d-reconstruction of piecewise planar objects from single panoramic images. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 119–126, 2000.

[20] T. Svoboda, T. Pajdla, and V. Hlavac. Epipolar geometry for panoramic cameras. In *Proc. 5th European Conference on Computer Vision*, pages 218–231, 1998.

[21] C.J. Taylor. Video plus. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 3–10, 2000.

[22] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *Trans. on Robotics and Automation*, 10:11–22, 1994.