Autocalibration & 3D Reconstruction with Non-central Catadioptric Cameras*

Branislav Mičušík Tomáš Pajdla
Center for Machine Perception, Dept. of Cybernetics, Faculty of Elec. Eng.
Czech Technical University in Prague, Karlovo nám. 13, 121 35 Prague, Czech Rep.
{micusb1,pajdla}@cmp.felk.cvut.cz

Abstract

We present a technique for modeling non-central catadioptric cameras consisting of perspective cameras and curved mirrors. The real catadioptric cameras have to be treated as non-central cameras, since they do not possess a single viewpoint. We present a method for solving the correspondence problem, auto-calibrating cameras, and computing a 3D metric reconstruction automatically from two uncalibrated non-central catadioptric images. The method is demonstrated on spherical, parabolic, and hyperbolic mirrors.

We observed that the reconstruction & auto-calibration with non-central catadioptric cameras is as easy (or as difficult) as with central catadioptric cameras, provided that the correspondence problem can be solved with a suitable approximate central model. It turns out that it is the number of parameters of the camera model that matters rather than the exact centrality of the projection. Our technique allows to autocalibrate catadioptric cameras even with genuinely non-central mirrors such as spheres (simple model, low blur, easy to manufacture) or uniform resolution mirrors (optimized projection).

1. Introduction

Omnidirectional cameras have large field of view, provide stable ego-motion estimation [5] and very complete 3D reconstruction from few (usually 2) images. There exist omnidirectional catadioptric cameras exploiting mirrors [3] or purely dioptric fish-eye lenses [20].

Some catadioptric cameras, consisting, e.g., of a parabolic, a hyperbolic, or an elliptical mirror, can be aligned with a central camera in such a way that they possess a single projection center [2]. Such catadioptric cameras can be treated as central cameras and the theory of central catadioptric cameras [2, 3, 9, 10, 21] can be employed.

In practice, however, catadioptric cameras are not central. The most common reasons are that *i)* a non-telecentric lens is used for a parabolic mirror or the lens and the parabolic mirror axes are not aligned properly, *ii)* a perspective camera is not placed in one of the focal length points of hyperbolic or elliptical mirror, or *iii)* the mirror shape, e.g. a sphere or uniform resolution mirrors [13, 8], are designed so that they do not posses a single viewpoint property. All the above may cause that the catadioptric camera becomes non-central and there is no single viewpoint from which all rays would emanate.

*This research was supported by the following grants: MSM 212300013, MŠMT Kontakt 22-2003-04, BeNoGo–IST–2001–39184.

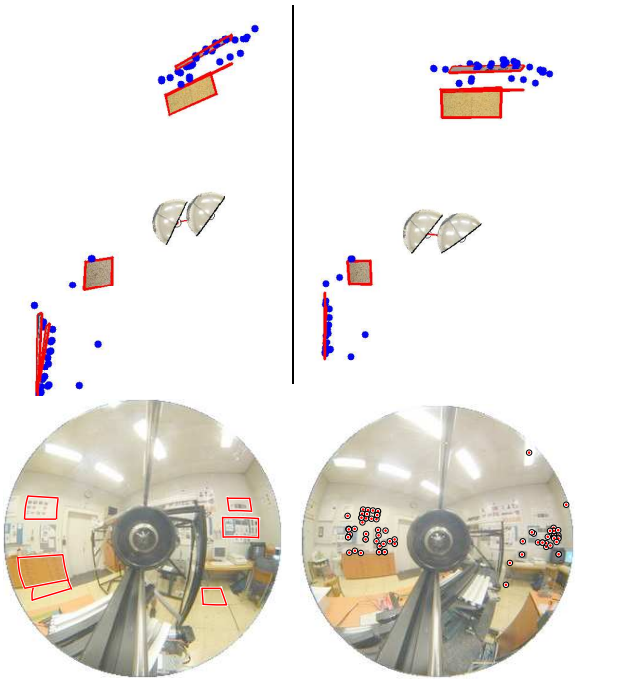

Figure 1: The bird's eye views of the 3D reconstructions of a square room. Top Left: A skewed reconstruction. A central model for a real para-catadioptric (PCD) camera was used. Top Right: The correct reconstruction. A non-central PCD camera model was used. Bottom: Input image pair showing manually marked rectangular regions and point matches established automatically by [14].

Using the central camera model for a non-central camera leads to an inaccurate determination of 3D rays corresponding to image points and consequently to a skewed 3D reconstruction as Fig. 1 (top left) shows. A remedy is to de-

rive the correct non-central camera model capturing the true distribution of reflected rays.

To build a 3D reconstruction from automatically established point correspondences contaminated by mismatches, we need to be able to incorporate the autocalibration procedure into some robust estimation technique, e.g. RANSAC [12]. However, there is a price to pay for the accuracy. A non-central model is often very complex with large number of parameters and it is often impossible (or very computationally complex) to use the non-central model in a RANSAC based estimation of multiple-view geometry. Thus, we are facing a problem: "To be accurate, we have to use a complex non-central model. To be automatic, we cannot use a complex model."

We show in this work how to solve the above problem by constructing a hierarchy of camera models. The hierarchy starts with a simplified model that is accurate enough to distinguish between correct (inliers) and incorrect (outliers) correspondences and simple enough to provide a tractable optimization problem when using the RANSAC. In all cases presented here, the simplified model is central. The simplified model allows to recover inliers automatically. Secondly, an accurate and complex non-central model is fitted to the correspondences validated in the first step. The non-central model allows to reconstruct scene geometry accurately. The art is to find *i)* a simplified model that is accurate enough to reject outliers but simple enough to provide a tractable RANSAC estimation problem and *ii)* a sufficiently accurate non-central model providing an accurate 3D reconstruction from correct image correspondences.

When an imaging system does not maintain a single viewpoint, a caustic, a locus of viewpoints in three dimensions is formed and the system has to be treated as non-central. Previous work related to the non-central cameras can be found in [4, 6, 11, 17, 19, 22]. In most relevant papers, [11, 22], the non-central devices like curved mirrors, meniscus lenses, camera clusters, or compound cameras were represented by points on the caustic and directional vectors towards scene points. These models capture geometrical, radiometrical, and optical properties of the omnidirectional camera. Our approach to modeling the optics differs in that we represent the catadioptric camera by points on the mirror surface (not on the caustic). Our model captures the geometric (not radiometric and optical) properties of the camera what is sufficient for doing 3D reconstruction. Since the caustic does not have to be computed, the derivation of our model is simpler and the computation of the Jacobian, as in [22], is not necessary. Moreover, our model can be estimate from point correspondences only.

In [6], the non-central spherical catadioptric camera is approximated by a camera with a single viewpoint (the top of the mirror) to construct perspective images. We also use a central approximate model of the spherical catadioptric camera. However, we chose the position of the center, which minimizes the sum of squares of angular differences between approximated rays and the correct ones.

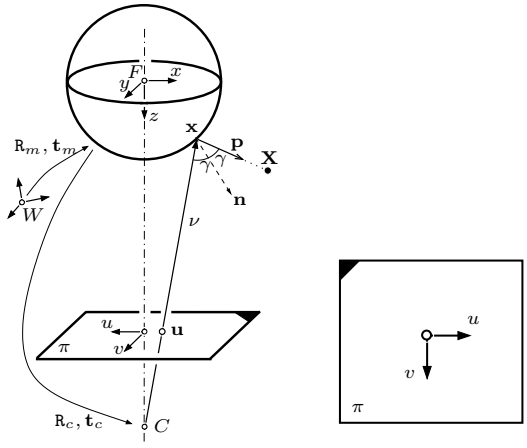

Figure 2: Left: Coordinate system of the spherical catadioptric camera. The origin is located in F. Right: A coordinate system in an acquired image.

The structure of the paper is the following. The non-central and the approximate central model, the autocalibration and the 3D reconstruction method are derived in full detail for the spherical mirror. For parabolic and a hyperbolic mirrors, the procedure is analogical, therefore only final equations, results, and observations are given.

2. Non-central model

The advantage of the spherical mirror is that it is readily available, provides images with little blur, has large field of view ($\sim 250°$), and the equation of the surface is simple (one-parametric). However, the spherical mirror does not possess a single projection center when observed by any central (perspective or orthographic) camera. To derive an appropriate central model, an approximation has to be used.

Let us derive a general imaging model for a non-central spherical catadioptric camera. The technique of the derivation is in some steps similar to the derivation of the central catadioptric camera models [21]. A sphere in the origin with center F and radius a can be expressed as

$$(\mathbf{x}^\top \quad 1)\, \mathtt{Q} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = 0, \tag{1}$$

where

$$\mathtt{Q} = \mathrm{diag}(1, 1, 1, -a^2) \tag{2}$$

and $\mathbf{x} = (x,\, y,\, z)^\top$ is a point on the sphere.

Suppose a Cartesian world coordinate system W. A Cartesian coordinate system of a mirror, placed in the center of the sphere F, is rotated and translated by $\mathtt{R}_m \in \mathbb{R}^{3\times3}$ and $\mathbf{t}_m \in \mathbb{R}^{3\times1}$ w.r.t. W. A Cartesian coordinate system

of a perspective camera, placed in the optical center C, is related to the mirror coordinate system by $\mathtt{R}_c$ and $\mathbf{t}_c$, see Fig. 2.

2.1. From image point to its projection ray

A conventional perspective camera projects a 3D point $\mathbf{x}$ on the mirror surface into the image plane by the standard projection equation [12]

$$\mathbf{u} = \frac{1}{\lambda}\mathtt{K}\mathtt{R}_c[\mathtt{I}|-\mathbf{t}_c]\begin{pmatrix}\mathbf{x}\\1\end{pmatrix}, \quad (3)$$

where $\mathtt{K}$ is a 3×3 upper triangular camera calibration matrix [12], $\mathbf{u}\in\mathbb{R}^3$ and $\lambda\in\mathbb{R}\backslash\{0\}$.

The ray ν emanating from camera center C in the direction $\mathbf{u}$ can be parametrized by λ and written, using (3), as

$$\nu = \lambda\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u} + \mathbf{t}_c = \lambda\mathbf{v} + \mathbf{t}_c, \quad (4)$$

where $\mathbf{v} = (r,s,t)^\top$ and $\mathbf{t}_c = (k,l,m)^\top$. Substituting (4) and (2) into the mirror equation (1) yields

$$(r^2{+}s^2{+}t^2)\lambda^2{+}2(ls{+}kr{+}mt)\lambda{+}(k^2{-}a^2{+}m^2{+}l^2) = 0. \quad (5)$$

Solving quadratic equation (5) gives two solutions for λ. Going from C in direction $\mathbf{u}$, we are interested in the intersection which is closer to the point C. Therefore the smaller λ is the correct solution.

Substituting λ from Eq. (5) into (4) yields the point on the mirror surface. This point $\mathbf{x} = \lambda\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u} + \mathbf{t}_c$ is expressed in the world coordinate system W as

$$\mathbf{x}_w = \mathtt{R}_m^{-1}\mathbf{x} + \mathbf{t}_m = \mathtt{R}_m^{-1}\left(\lambda\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u} + \mathbf{t}_c\right) + \mathbf{t}_m. \quad (6)$$

The ray ν is reflected at the point $\mathbf{x}$ by the mirror such that the angles γ of incident and coincident rays to the surface normal are equal. The normal to the mirror surface in mirror point $\mathbf{x}$, i.e. $\mathbf{n} = (\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, -1)^\top$, is given by

$$\mathbf{n} = \left(\frac{-x}{\sqrt{a^2-x^2-y^2}}, \frac{-y}{\sqrt{a^2-x^2-y^2}}, -1\right)^\top. \quad (7)$$

If the normal vector $\mathbf{n}$ is normalized to $\hat{\mathbf{n}} = \frac{\mathbf{n}}{\|\mathbf{n}\|}$ of unit length, the directional vector of the reflected ray $\mathbf{p}$ can be computed as

$$\begin{aligned}\mathbf{p} &= \mathbf{v} - 2(\mathbf{v}\cdot\hat{\mathbf{n}})\cdot\hat{\mathbf{n}}\\ &= \mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u} - 2\left(\left(\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u}\right)\cdot\hat{\mathbf{n}}\right)\cdot\hat{\mathbf{n}}. \quad (8)\end{aligned}$$

Vector $\mathbf{p}$, expressed in the world coordinate system W, i.e. $\mathbf{p}_w = \mathtt{R}_m^{-1}\mathbf{p}$, is given by

$$\mathbf{p}_w = \mathtt{R}_m^{-1}\left(\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u} - 2\left(\left(\mathtt{R}_c^\top\mathtt{K}^{-1}\mathbf{u}\right)\cdot\frac{\mathbf{n}}{\|\mathbf{n}\|}\right)\cdot\frac{\mathbf{n}}{\|\mathbf{n}\|}\right). \quad (9)$$

Eq. (6) and (9) represent a complete mapping from an image point $\mathbf{u}$ in the camera coordinate system to a pair $(\mathbf{x}_w;\mathbf{p}_w)$ in a world coordinate system. The pair $(\mathbf{x}_w;\mathbf{p}_w)$ consists of the point $\mathbf{x}_w$ on the mirror surface and the directional vector $\mathbf{p}_w$ pointing towards a scene point $\mathbf{X}$. A simulation of the derived model shows the corresponding caustic, see Fig. 3.

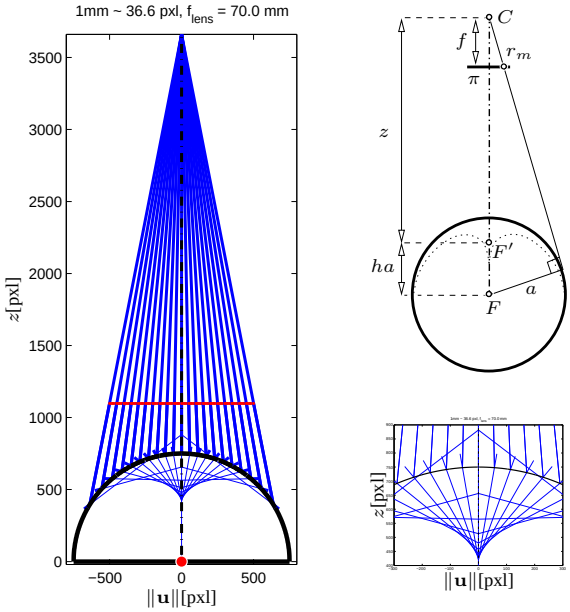

Figure 3: Rays reflected by the mirror are tangent to a caustic surface. Left: A perspective camera looking at a spherical mirror. Right top: An approximation of a non-central spherical model by a central one with a fictive viewpoint F'. Right bottom: A detail view of the caustic.

2.2. Projection of a scene point to an image

Given a scene point $\mathbf{X}$, there is no closed-form solution for its image projection $\mathbf{u}$. It can be easily designed as an iterative method incorporating the mapping from an image point to its projection ray (derived in subsection 2.1).

First, the initial image $\mathbf{u}$ of $\mathbf{X}$ is obtained using a central model (derived later), for which a direct solution exists. Second, the iterative method is used. The iterative method minimizes (over coordinates of $\mathbf{u}$) the distance between a ray, computed from $\mathbf{u}$ by the non-central model from Sec. 2.1, and the 3D point $\mathbf{X}$. The method converges very quickly thanks to a good initial estimate of $\mathbf{u}$ provided by the approximate central model.

3. Approximation by a central model

A non-central model can be approximated by a central one in order to find a sufficiently simple model which is accurate enough to distinguish between correct (inliers) and incorrect (outliers) point matches and which can be used to obtain an initial estimate of camera motion.

Suppose that the optical axis of a perspective camera goes through the center of the spherical mirror, i.e. $\mathtt{R}_c = \mathrm{diag}(-1,1,-1)$ (image is mirrored), $\mathbf{t}_c = (0,0,z)^\top$, and the whole spherical mirror is visible in the image.

Since the rim of the mirror is visible in the image, the principal point [12] can easily be estimated by fitting an ellipse to the image of the mirror rim. The center of the

ellipse can be regarded as the principal point. The ellipse can be transformed to a circle to obtain radially symmetric image appropriate for the approximate *precalibrated* central model. In this *precalibrated* image, the calibration matrix becomes $\mathtt{K} = \text{diag}\,(f, f, 1)$. In all central models in the rest of the paper, an image point $\mathbf{u}$ stands for the image point in the precalibrated image.

3.1. From image point to its projection ray

Suppose a fictive optical center F' moved by $h\,a$ from the center of a sphere, see Fig. 3. Unknown parameters are: f (focal length of the perspective camera), z (distance camera – mirror), a (mirror parameter, radius of the sphere), h (position of the fictive center).

The matrix $\mathtt{Q}$ of a sphere with the origin in the fictive point F' in Eq. (1) becomes

$$\mathtt{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & ha \\ 0 & 0 & ha & (ha)^2 - a^2 \end{bmatrix}. \tag{10}$$

It follows from perspective projection, see Fig. 3 (right top), that the focal length

$$f = r_m \sqrt{\left(\frac{z}{a} + h\right)^2 - 1} \tag{11}$$

is a function of z, a, h. It holds in the precalibrated image that

$$\mathbf{v} = \mathtt{R}_c^\top \mathtt{K}^{-1} \mathbf{u} = (-u/f, v/f, -1)^\top. \tag{12}$$

Substituting (10) and (4) with $\mathbf{v}$ from (12) and λ from (5) where $k, l = 0$ (aligned optical axes), into the mirror equation (1) yields the vector $\mathbf{p}$ going towards a scene point

$$\mathbf{p} \simeq \lambda \begin{pmatrix} -\frac{u}{f} \\ \frac{v}{f} \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}, \tag{13}$$

with

$$\lambda = \frac{b + h - \sqrt{(b+h)^2 - \left(\frac{r^2}{f^2} + 1\right)(b^2 + 2hb + h^2 - 1)}}{\frac{r^2}{f^2} + 1},$$

where $r = \sqrt{u^2 + v^2}$ is the radius of an image point, f is from (11), and $b = \frac{z}{a}$ is a new parameter absorbing z and a. Unknown parameters are reduced to the ratio $\frac{z}{a}$ and h.

Let us investigate the relationship of the distance h of the fictive center F' and the ratio $\frac{z}{a}$. We minimized (over h) the sum of squares of angular differences between real 3D rays reflected by a mirror surface to scene points and the approximated rays emanating from the fictive center F', see Fig. 4. It can be seen that the distance of the fictive center is changing moderately. Small change of h does not affect significantly the error arised by approximating the non-central model by the central one. Therefore we fixed $h = 0.64$. By that we reduced the number of parameter to one, $b = \frac{z}{a}$.

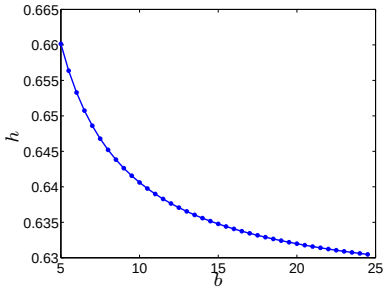

Figure 4: The position h of the fictive center F' as a function of the ratio $b = \frac{z}{a}$.

3.2. Projection of a scene point to an image

Let $\mathbf{X} \in \mathbb{R}^3$ be a scene point, and $\mathtt{R}_m$, $\mathbf{t}_m$, $\mathtt{R}_c$, $\mathbf{t}_c$, $\mathtt{K}$ be as before. The intersection of a ray going from $\mathbf{X}$ towards a fictive center F' with the mirror surface, Eq. (10), can be computed from Eq. (1)

$$\begin{bmatrix} \lambda[\mathtt{R}_m, -\mathtt{R}_m\mathbf{t}_m]\mathbf{X} \\ 1 \end{bmatrix}^\top \mathtt{Q} \begin{bmatrix} \lambda[\mathtt{R}_m, -\mathtt{R}_m\mathbf{t}_m]\mathbf{X} \\ 1 \end{bmatrix} = 0$$

as the smaller λ from two solutions of the quadratic equation

$$(r^2 + s^2 + t^2)\lambda^2 + 2\,t\,h\,a\lambda + h^2a^2 - a^2 = 0, \tag{14}$$

where $(r, s, t)^\top = [\mathtt{R}_m, -\mathtt{R}_m\mathbf{t}_m]\mathbf{X}$. A point on the mirror surface, written in the camera coordinate system,

$$\mathbf{x}_c = (x_c, y_c, z_c)^\top = \mathtt{R}_c\,(\lambda[\mathtt{R}_m, -\mathtt{R}_m\mathbf{t}_m]\mathbf{X} - \mathbf{t}_c) \tag{15}$$

with λ from (14) is projected to the image as $\mathbf{u} = \mathtt{K}\frac{1}{z_c}\mathbf{x}_c$.

4. Autocalibration from EG

Suppose that we have two images of a rigid scene acquired by the same spherical catadioptric camera from two different positions. Assume that automatically established tentative point correspondences contaminated by outliers are available. We know neither the internal calibration parameters of the catadioptric camera, nor the position of cameras.

We want to validate the tentative correspondences (mark inliers/outliers), estimate parameter b in Eq. (13), and find the relative position of the second camera w.r.t. the first one, i.e. $\mathtt{R}_m$, $\mathbf{t}_m$. To do so, we will use the approximated central model derived in Sec. 3.

The basic property of central cameras is that they possess a single viewpoint and therefore the epipolar geometry can be formulated for vectors $\mathbf{p}$ in the 1st camera and $\check{\mathbf{p}}$ in the 2nd one

$$\check{\mathbf{p}}^\top \mathtt{F}\,\mathbf{p} = 0. \tag{16}$$

The autocalibration method based on epipolar constraint is motivated by papers [7, 15]. In order to obtain a direct solution of $\mathtt{F}$ and b, a linearization of the vector $\mathbf{p}$ in b_0 in Eq. (13) is performed using the Taylor series

$$\begin{aligned}\mathbf{p} &\approx \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}_{b=b_0} + (b-b_0)\begin{pmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{p}_z \end{pmatrix}_{b=b_0} = \\ &= \begin{pmatrix} p_x - \dot{p}_x b_0 \\ p_y - \dot{p}_y b_0 \\ p_z - \dot{p}_z b_0 \end{pmatrix}_{b=b_0} + b\begin{pmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{p}_z \end{pmatrix}_{b=b_0} = \mathbf{q} + b\,\mathbf{s}, \quad (17)\end{aligned}$$

where vectors $\mathbf{q} = (p, q, r)^\top$ and $\mathbf{s} = (s, t, u)^\top$ are computed directly from known image coordinates, and $\dot{p}_i$, $i = \{x, y, z\}$, stand for $\frac{\partial p_i}{\partial b}$. The epipolar constraint, Eq. (16), can be written as

$$(\check{\mathbf{q}} + b\check{\mathbf{s}})^\top \mathtt{F}\,(\mathbf{q} + b\mathbf{s}) = 0.$$

Gathering the point coordinates into three design matrices, we obtain an equation that is quadratic in parameter b and linear in $\mathbf{f}$,

$$\left(\mathtt{D}_1 + b\mathtt{D}_2 + b^2\mathtt{D}_3\right)\mathbf{f} = 0, \quad (18)$$

known as the Quadratic Eigenvalue Problem (QEP) [1]. Efficient algorithms are available, e.g. MATLAB solves QEP by the function `polyeig`.

Every point correspondence adds one row into the matrices $\mathtt{D}_i$. The unknown vector $\mathbf{f}$ (created from the matrix $\mathtt{F}$) and one row of three design matrices $\mathtt{D}_i$ composed from image points coordinates have the following form:

$$\begin{aligned}\mathbf{f} &= [\,F_{11},\, F_{12} \quad \dots \quad F_{33}\,]^\top \\ \mathtt{D}_1 &= [\;\check{p}p \quad \check{p}q \quad \check{p}r \quad \check{q}p \quad \check{q}q \quad \check{q}r \quad \check{r}p \quad \check{r}q \quad \check{r}r\;] \\ \mathtt{D}_2 &= [\;\check{p}s + \check{s}p \quad \check{p}t + \check{s}q \quad \check{p}u + \check{s}r \quad \check{q}s + \check{t}p \quad \check{q}t + \check{t}q \\ &\qquad \check{q}u + \check{t}r \quad \check{r}s + \check{u}p \quad \check{r}t + \check{u}q \quad \check{r}u + \check{u}r\;] \\ \mathtt{D}_3 &= [\;\check{s}s \quad \check{s}t \quad \check{s}u \quad \check{t}s \quad \check{t}t \quad \check{t}u \quad \check{u}s \quad \check{u}t \quad \check{u}u\;]. \quad (19)\end{aligned}$$

The matrices $\mathtt{D}_i$ have to be square. Therefore, the minimum number of correspondences is 9. Usage for 9 matches: `polyeig`$(\mathtt{D}_3, \mathtt{D}_2, \mathtt{D}_1)$. If more than 9 points are to be used, the least-squares solution can be obtained by premultiplication by $\mathtt{D}_i^\top$. It does not change the solution, but allows square solvers to be used. Usage for more than 9 matches: `polyeig`$(\mathtt{D}_1^\top\mathtt{D}_3, \mathtt{D}_1^\top\mathtt{D}_2, \mathtt{D}_1^\top\mathtt{D}_1)$.

As a result of solving Eq. (18), there are 18 solutions for b and their corresponding $\mathtt{F}$'s. Many of the solutions are zero, infinite or complex. In practice, no more than one solution is considerable. If there are more then one solution, the one with the minimal angular error (the sum of angles between an epipolar plane and 3D vectors $\mathbf{p}$ [18]) is chosen.

The most important property of this autocalibration method is that the direct solution of Eq. (18) allows to incorporate the autocalibration method into a RANSAC robust estimation technique and thus to handle outliers. As the result of the autocalibration, the essential matrix (stronger than the fundamental one, allowing a metric reconstruction), the camera parameter b, and the inliers, see Fig. 5, are obtained.

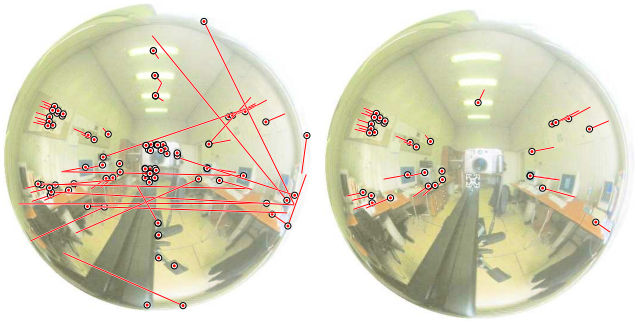

Figure 5: Outlier detection during the calibration process. Left: An acquired image from a stereo pair with all tentative correspondences found by [14]. Circles mark points in the image, lines join them to their matches in the next one. Right: The same image with the validated inliers only and their matches in the next image.

The reconstruction error (explained later), computed by using estimated b, $\mathtt{F}$ $(\mathtt{R}_m, \mathbf{t}_m)$ and the non-central camera model, derived in Sec. 2.1, is used as the criterion in RANSAC to classify a point match to be an inlier or an outlier.

The linearization was done in the point $b_0 = 10$, what is a typical ratio of camera distance to the radius of a sphere in which it is possible to focus whole mirror surface. A rough initial estimate of b_0 is sufficient.

5. 3D reconstruction

Every i^{th} point correspondence can be represented by $(\mathbf{x}_w^i; \mathbf{p}_w^i)$ in the first camera and by $(\check{\mathbf{x}}_w^i; \check{\mathbf{p}}_w^i)$ in the second one, computed from Eq. (6) and (9). The reconstructed 3D point $\mathbf{X}$ for one point match $(\mathbf{x}_w^i; \mathbf{p}_w^i) \leftrightarrow (\check{\mathbf{x}}_w^i; \check{\mathbf{p}}_w^i)$ is obtained as the point in the center of the shortest transversal of the respective projection rays

$$d_i = \frac{|\left(\mathbf{x}_w^i - \check{\mathbf{x}}_w^i\right) \cdot \left(\mathbf{p}_w^i \times \check{\mathbf{p}}_w^i\right)|}{|\mathbf{p}_w^i \times \check{\mathbf{p}}_w^i|}. \quad (20)$$

The final *metric* reconstruction $\mathcal{R}_M$ is obtained by minimizing the sum of squared lengths of shortest transversals

$$\mathcal{R}_M = \underset{b, \mathtt{R}_c, \mathbf{t}_c, \mathtt{R}_m, \mathbf{t}_m, \mathtt{K}}{\operatorname{argmin}} \sum_{i=1}^{N} d_i^2, \quad (21)$$

where N is the number of point matches. This is almost equivalent to minimizing the image reprojection errors since our cameras are calibrated.

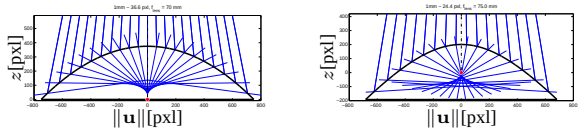

Figure 6: Reflected rays by a parabolic (left) and a hyperbolic (right) non-central catadioptric camera create caustics.

6. Other mirrors

6.1. Parabolic mirror

If a standard lens with a large focal length is used instead of an ideal telecentric lens or if the orthographic projection is not exactly parallel with the axis of symmetry of the parabolic mirror, the para-catadioptric camera becomes non-central, see Fig. 6 (left).

Suppose a paraboloid of revolution with equation $z = \frac{a^2 - \|\mathbf{u}\|}{2a}$ with the origin located at the focal point of the paraboloid, where a is the parameter of the paraboloid and $\mathbf{u}$ is an image point. Analogically to Eq. (4) we have

$$(-r^2 - s^2)\lambda^2 + 2(-kr - at - ls)\lambda + (a^2 - k^2 - l^2 - 2am) = 0,$$

what gives the solution for λ in (6). Substituting λ and the normal vector to the parabolic mirror surface at the mirror point $\mathbf{x}$, i.e. $\mathbf{n} = \left(\frac{-x}{a}, \frac{-y}{a}, -1\right)^\top$, into (6) and (9) gives the *non-central* para-catadioptric camera model.

The autocalibration method from automatically established point matches for *central* para-catadioptric cameras, analogical to the method described in Sec. 4, can be used to solve the correspondence problem and to obtain an initial estimate of camera model and motion [16]. The method leads to the Polynomial Eigenvalue Problem of degree 5 [16] .

The parabolic mirror has an advantage over a spherical one since the auto-calibration method does not require a linearization of the camera model and thus an initial estimate of a point of the linearization is not needed. However, the parabolic mirror provides more blurred images than the spherical one.

6.2. Hyperbolic mirror

If the optical center of a perspective camera is not placed in the second focal length of the hyperbola precisely the camera becomes non-central, see Fig. 6 (right).

Suppose a hyperboloid of revolution with equation $\frac{(z-e)^2}{a^2} - \frac{x^2+y^2}{b^2} = 1$ with the origin located in the first focal point of the hypeboloid, where a, b are parameters of the hyperboloid and $e = \sqrt{a^2 + b^2}$ is the eccentricity. Analogically to Eq. (4) we have

$$(t^2b^2 - s^2a^2 - r^2a^2)\lambda^2 + 2(mtb^2 - etb^2 - lsa^2 - 2kra^2)\lambda + \\ + (m^2b^2 - k^2a^2 - 2meb^2 - l^2a^2 + b^4) = 0, \qquad (22)$$

what gives the solution for λ in (6). Substituting λ and the normal vector to the mirror surface at the mirror point $\mathbf{x}$, i.e. $\mathbf{n} \simeq \left(-\frac{a}{b}x, -\frac{a}{b}y, -\sqrt{b^2 + x^2 + y^2}\right)^\top$, into (6) and (9) gives the *non-central* hyperbolic catadioptric camera model.

We designed an autocalibration method analogical to the methods for spherical and parabolic mirrors. The method assumes the central model with the center of perspective camera placed into the second focal length of the hyperboloid.

Substituting λ, computed from Eq. (22) for $k, l = 0$ (aligned optical axes), $m = 2e$ (central camera), and $\mathtt{R}_c^\top \mathtt{K}^{-1}\mathbf{u}$ from Eq. (12) to Eq. (4), yields the vector $\mathbf{p}$

$$\mathbf{p} \simeq \frac{c^2\left(\sqrt{1+c^2} + \sqrt{\frac{r^2}{f^2}+1}\right)}{c^2 - \frac{r^2}{f^2}} \begin{pmatrix} -\frac{u}{f} \\ \frac{v}{f} \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2\sqrt{1+c^2} \end{pmatrix}, \qquad (23)$$

where $r = \sqrt{u^2 + v^2}$ is the radius of the image point $\mathbf{u}$ w.r.t. the principal point, $c = \frac{b}{a}$ is unknown parameter, and f is the focal length of the perspective camera. The focal length f can be expressed as a function of c, r_m (radius of a view field circle), and g (the ratio height/width of a mirror) as

$$f = -\frac{r_m(g^2c^2 - 1)}{2gc^2}\left(\sqrt{1+c^2} - \frac{2g^2c^2}{g^2c^2 - 1} + 1\right). \qquad (24)$$

The vector $\mathbf{p}$ in Eq. (23) can thus be expressed as a one-parametric function of parameter c and represents the *central model* of a hyperbolic catadioptric camera.

To derive the autocalibration method, the procedure from Sec. 4 can be directly applied. The difference from the spherical mirror is that the linearization, Eq. (17), of $\mathbf{p}$ in Eq. (23) with substituted f, Eq. (24), is done at the point c_0 (instead of b_0). The method leads to the Quadratic Eigenvalue Problem and a 9-point RANSAC.

There is a disadvantage in comparison to the spherical or parabolic mirror: the ratio height/width of the mirror needs to be known to arrive at a 9-point RANSAC. If the ratio is unknown, the vector $\mathbf{p}$ in Eq. (23) is a two-parametric function of c and g. The method would then lead to the QEP again, but to a 15-point RANSAC, what is computationally much more demanding.

7. Experiments

We acquired two images by a catadioptric camera consisting of a perspective camera (Canon PowerShot G2) and spherical, parabolic and hyperbolic mirrors, respectively. The size in pixels of the minimal image containing a view field circle was $1042 \times 1042 \sim 252°$ for the spherical, $1474 \times 1474 \sim 210°$ for the parabolic, and $1560 \times 1560 \sim 186°$ for the hyperbolic mirror.

The tentative correspondences between the gravity centers of feature regions were obtained by [14]. As a result

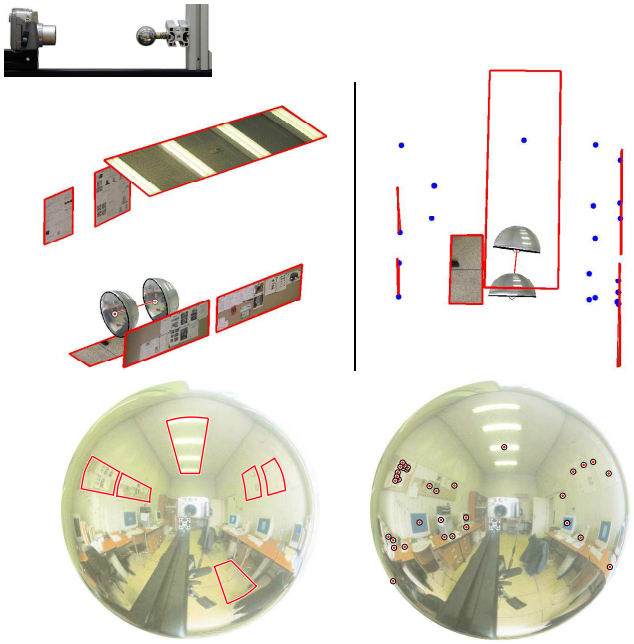

Figure 7: Middle left: 3D metric reconstruction from two uncalibrated spherical catadioptric images. Top right: Bird's eye view of the room. Bottom: Input image pair together with manually marked rectangular regions and automatically detected point correspondences[14].

of applying the calibration method based on central models (described in detail for spherical mirror in Sec. 4), initial camera models and the corresponding essential matrices were obtained, and outliers were rejected, see Fig. 5 and 8 (2nd row).

The camera model parameters and relative camera positions obtained from the estimated essential matrix were used to compute an initial 3D reconstruction as Sec. 5 describes. The 3D reconstruction was improved by the non-linear bundle adjustment, Eq. (21), assuming the non-central model and tuning the second camera matrix $\mathtt{P}_2 = [\mathtt{R}_m \,|\, \mathtt{t}_m]$ (the first camera was fixed to $\mathtt{P}_1 = [\mathtt{I} \,|\, \mathbf{0}]$) and internal camera model parameters, same for both cameras. The internal camera model parameters are *i)* the center of radial symmetry in the image $(u_0, v_0)^\top$, i.e. the intersection of the mirror axis with the image plane, *ii)* mirror shape parameters a, b, *iii)* perspective camera orientation and position w.r.t. the mirror coordinate system $\mathtt{R}_c$, $\mathbf{t}_c$, and *iv)* the focal length of the perspective camera f. The square pixel and zero skew of the perspective camera were assumed.

To show the quality of the 3D reconstruction, some rectangular regions have been marked manually. The estimated parameters of the non-linear camera model were used for the reconstruction these regions.

Results in Fig. 7 and 8 show that the method leads to an accurate 3D metric reconstruction (for a very small resolution) what can be easily verified on bird's eye views that the reconstructed rooms is indeed rectangular. Notice thus for a spherical mirror, 1024pxl for the FOV 252° corresponds to 182pxl for the FOV 45° of a typical perspective camera.

8. Conclusion

We derived accurate non-central and suitable approximate central models for spherical, parabolic, and hyperbolic mirrors. We designed a method allowing to build a 3D metric reconstruction from two uncalibrated non-central catadioptric images. One can observe that an accurate and quite complete 3D reconstruction of surrounding scene can be obtained already from two views.

The proposed autocalibration and 3D reconstruction technique is not restricted to the three mirrors presented in the paper. The same procedure can be applied for non-central catadioptric cameras composed of uniform resolution mirrors especially designed to fulfil a given type of projection [13, 8].

Two important conclusions follow from our work. First, the correspondences between catadioptric images can (and need to) be validated using an approximation of the non-central camera model by a suitable central one. Second, the non-central camera model has to be used to get a geometrically correct 3D metric reconstruction.

References

[1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems : A Practical Guide*. SIAM, Philadelphia, 2000.

[2] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *Proc. IJCV*, 35(2):175–196, 1999.

[3] R. Benosman and S. B. Kang. *Panoramic Vision : Sensors, Theory, and Applications*. Springer Verlag, Germany, 2001.

[4] T. Bonfort and P. Sturm. Voxel carving for specular surfaces. In *Proc. ICCV*, pages 591–596, 2003.

[5] T. Brodsky, C. Fermuller, and Y. Aloimonos. Directions of motion fi elds are hardly ever ambiguous. *IJCV*, 26(1):5–24, 1998.

[6] S. Derrien and K. Konolige. Approximating a single viewpoint in panoramic imaging devices. In *Proc. of the IEEE Workshop on Omnidirectional Vision*, pages 85–90, 2000.

[7] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proc. CVPR*, pages I: 125–132, 2001.

[8] S. Gächter, T. Pajdla, and B. Micusik. Mirror design for an omnidirectional camera with a space variant imager. In *Proc. of the IEEE Workshop on Omnidir. Vision*, pages 99–105, 2001.

[9] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. In *Proc. CVPR*, pages I: 279–286, 2001.

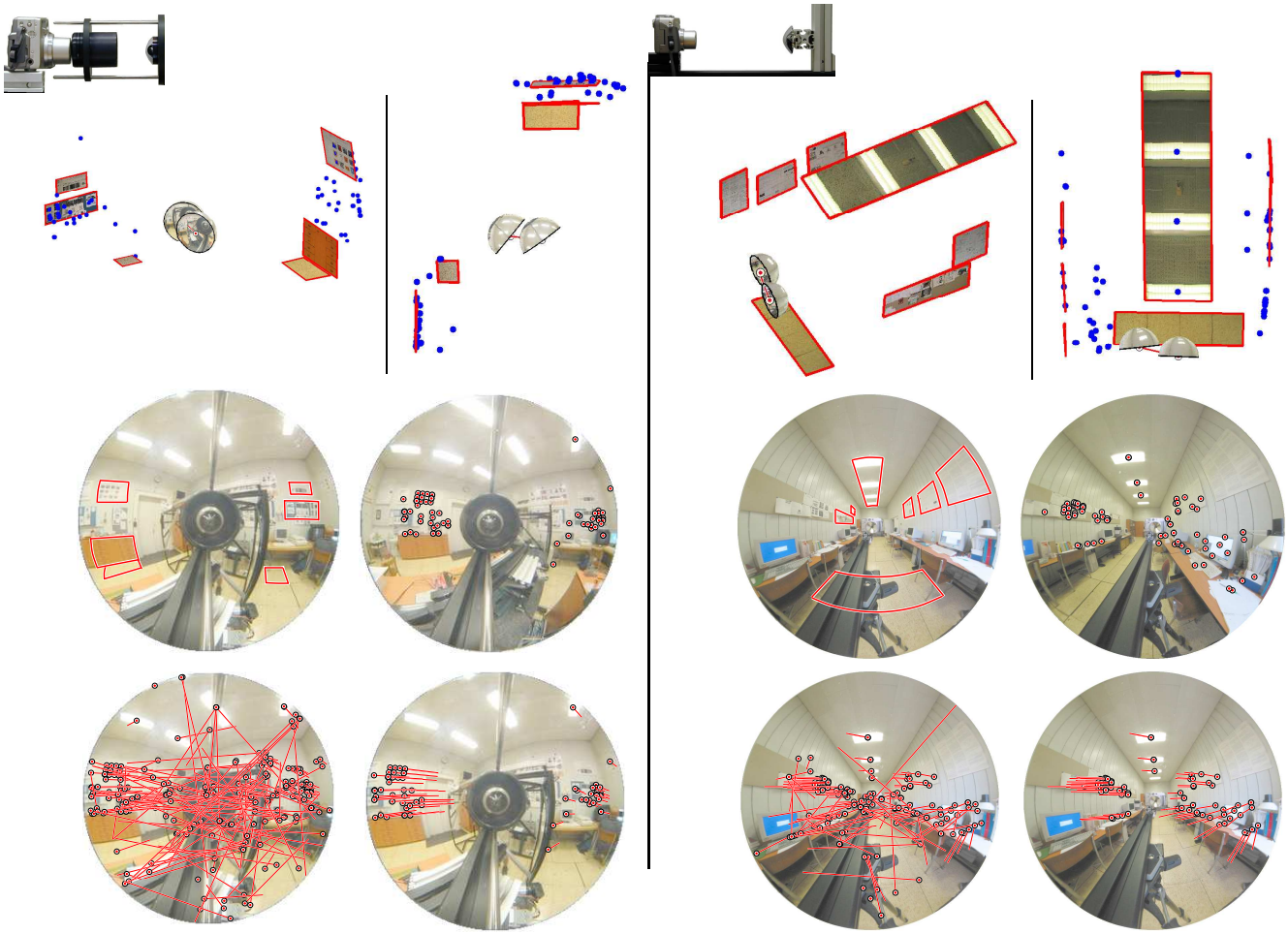

Figure 8: 3D metric reconstructions from two uncalibrated parabolic (left) hyperbolic (right) catadioptric images. 1st row: Some views of 3D reconstructions together with both cameras. 2nd row: Input image pair together with manually marked rectangular regions and automatically detected point correspondences by [14]. 3rd row: Inliers/outliers detection, see Fig. 5 for more explanation.

[10] C. Geyer and K. Daniilidis. Mirrors in motion: Epipolar geometry and motion estimation. In *Proc. ICCV*, pages 766–773, Nice, France, 2003.

[11] M. D. Grossberg and S. K. Nayar. A general imaging model and a method for finding its parameters. In *Proc. ICCV*, pages 108–115, 2001.

[12] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

[13] A. Hicks and R. Bajcsy. Catadioptric sensors that approximate wide-angle perspective projections. In *Proc. of the IEEE Workshop on Omnidir. Vision*, pages 97–103, 2000.

[14] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc. BMVC*, pages I: 384–393, UK, 2002.

[15] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *Proc. CVPR*, pages I: 485–490, 2003.

[16] B. Micusik and T. Pajdla. Para-catadioptric camera auto-calibration from epipolar geometry. In *Proc. ACCV*, volume II, pages 748–753, 2004.

[17] J. Neumann, C. Fermüller, and Y. Aloimonos. Eye design in the plenoptic space of light rays. In *Proc. ICCV*, pages 1160–1167, 2003.

[18] J. Oliensis. Exact two–image structure from motion. *PAMI*, 24(12):1618–1633, 2002.

[19] R. Pless. Using many cameras as one. In *Proc. CVPR*, pages II: 587–593, 2003.

[20] S. F. Ray. *Applied photographic optics : Lenses and optical systems for photography, film, video, electronic and digital imaging*. Focal Press, Oxford, 2002.

[21] T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *IJCV*, 49(1):23–37, 2002.

[22] R. Swaminathan, M. Grossberg, and S. Nayar. Caustics of catadioptric cameras. In *Proc. ICCV*, pages II: 2–9, 2001.