

Structure from Motion with Wide Circular Field of View Cameras

Branislav Mičušík and Tomáš Pajdla

Abstract—This paper presents a method for fully automatic and robust estimation of two-view geometry, autocalibration, and 3D metric reconstruction from point correspondences in images taken by cameras with wide circular field of view. We focus on cameras which have more than 180° field of view and for which the standard perspective camera model is not sufficient, e.g., the cameras equipped with circular fish-eye lenses Nikon FC-E8 (183°), Sigma 8mm-f4-EX (180°), or with curved conical mirrors. We assume a circular field of view and axially symmetric image projection to autocalibrate the cameras. Many wide field of view cameras can still be modeled by the central projection followed by a nonlinear image mapping. Examples are the above-mentioned fish-eye lenses and properly assembled catadioptric cameras with conical mirrors. We show that epipolar geometry of these cameras can be estimated from a small number of correspondences by solving a polynomial eigenvalue problem. This allows the use of efficient RANSAC robust estimation to find the image projection model, the epipolar geometry, and the selection of true point correspondences from tentative correspondences contaminated by mismatches. Real catadioptric cameras are often slightly noncentral. We show that the proposed autocalibration with approximate central models is usually good enough to get correct point correspondences which can be used with accurate noncentral models in a bundle adjustment to obtain accurate 3D scene reconstruction. Noncentral camera models are dealt with and results are shown for catadioptric cameras with parabolic and spherical mirrors.

Index Terms—Omnidirectional vision, fish-eye lens, catadioptric camera, autocalibration.

1 INTRODUCTION

STABLE ego-motion estimation and a 3D metric reconstruction from a small number of images can be achieved by using cameras with a large angle of view [8], [11], [45]. There exists a variety of omnidirectional cameras consisting of mirrors or wide-angle (fish-eye) lenses with an angle of view larger than 180° [6]. In this work, we show how to do autocalibration and structure from motion with such cameras.

We are interested in central and slightly noncentral *omnidirectional* cameras with a circular field of view which cannot be described by the standard perspective model [10], [22]. See Fig. 1 for examples of central dioptric (only lenses) and slightly noncentral catadioptric (combination of lenses and mirrors) omnidirectional cameras. The reason why the center of projection is so desirable is that the models and algorithms are simpler and it is possible to generate pure perspective images by transforming cutouts from the acquired images. The advantage of omnidirectional cameras is primarily that they have a large field of view and, thus, the ability to see a large part of the surroundings in one image. The large angle of view often helps to establish extended point correspondences which provide stable and very complete 3D reconstructions from few images. Degenerate

cases (e.g., when only a single plane is observed in the image) are less likely and therefore more stable ego-motion estimation is often achieved [8], [11] compared to standard cameras.

This paper studies omnidirectional camera geometry, properties, image formation, and the possibility of their autocalibration from point correspondences in order to compute the Structure from Motion from two views, see Fig. 2. Often, no information about mirror or lens parameters and no calibration objects is available. Nevertheless, it still is often possible to acquire two or more images of the surroundings and establish point correspondences manually or automatically.

The main contribution of this work is the formulation of the autocalibration methods for omnidirectional cameras with a circular field of view as the Polynomial Eigenvalue Problem, which allows one to robustly estimate the two-view geometry just from image correspondences even if they are contaminated by mismatches.

We show that the Fitzgibbon's technique [12] can be generalized to a wider class of cameras with radially symmetric image formation. Interestingly, there is a direct generalization for the para-catadioptric cameras, leading to the quartic polynomial eigenvalue problem. More importantly, a very large class of (probably all practically useful) radially symmetric mappings can be approximated in such a way that the model estimation can be formulated as a polynomial eigenvalue problem. For one-parametric models, we get a direct generalization of [12] and need to solve a quadratic eigenvalue problem. For more-parametric models, a coordinate lifting can be used to attain the quadratic eigenvalue problem on the cost of using extra correspondences. This technique allows one to estimate epipolar geometry and one-parametric models from nine and two-parametric models from 15 correspondences in two views.

• B. Mičušík is with the Pattern Recognition and Image Processing Group, Institute of Computer Aided Automation, Vienna University of Technology, Favoritenstr. 9/1832, A-1040 Vienna, Austria. E-mail: micusik@prip.tuwien.ac.at.

• T. Pajdla is with the Center for Machine Perception, Department of Cybernetics, Czech Technical University in Prague, Karlovo nám. 13, 121 35 Prague, Czech Republic. E-mail: pajdla@cmp.felk.cvut.cz.

Manuscript received 4 Jan. 2005; revised 6 Oct. 2005; accepted 19 Oct. 2005; published online 11 May 2006.

Recommended for acceptance by C. Taylor.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0006-0105.



Fig. 1. Omnidirectional cameras and acquired images. From the left: Nikon Coolpix digital camera and Nikon FC-E8 fish-eye lens; Canon EOS-1Ds and Sigma 8mm-f4-EX fish-eye lens; perspective camera and hyperbolic mirror; orthographic camera and parabolic mirror.



Fig. 2. The 3D metric reconstruction built from *two* uncalibrated omnidirectional images. Input images with drawn epipolar curves are shown.

We present experimental results for some single view-point catadioptric cameras as well as for two wide field of view fish eye lenses. We demonstrate that the technique can be integrated into a RANSAC estimation paradigm and used in a hierarchical manner to cope with a large number of mismatches. The simple approximate image formation model allows us to detect and remove mismatches efficiently. Moreover, the approximate model computed from correct matches can be used to initiate bundle adjustment with very accurate but complex models. We demonstrate this on slightly noncentral catadioptric cameras which need a complicated model to get accurate reconstructions. We present experiments with short as well as long baseline camera setups and automatically established tentative correspondences where contamination by mismatches reached 80 percent.

To enable the formulation of a polynomial eigenvalue problem, the center of radial symmetry of the image formation has to be known and the image has to have “square pixels.” In general, these conditions are not satisfied. The technique [12] has been used only for cameras with a standard angle of view. In such situations, it is safe to assume that the center of radial symmetry is in the center of the image. Moreover, work [12] assumed that the camera had square pixels. We demonstrate that, if cameras have a circular field of view, which is common for very large angle of view imaging, the origin can be put into the symmetry center and the image can be transformed to get squared pixels by mapping the view field to a circle. In a way, the circular view field becomes a calibration object. Even though this precalibration may not be accurate enough for metrology, it was always accurate enough to get a working

polynomial eigenvalue problem for rejecting mismatches. Our experiments demonstrate a hierarchy of the two-view geometry estimation based on a coarse-to-fine strategy for some types of central and slightly noncentral omnidirectional cameras. As a side effect, we also show that accurate 3D metric reconstructions can be obtained from two uncalibrated images with very wide circular field of view.

The aim of the proposed calibration method is not to be more accurate than offline calibration methods using calibration targets and many-parametric models. The aim of the method is to produce 3D reconstruction *fully automatically* when only uncalibrated omnidirectional images are available.

2 PREVIOUS WORK

Essential concepts describing the geometry of omnidirectional cameras were studied and some of the models of catadioptric cameras were proposed in [44], [3], [15], [5], [6], [41], [18], [52]. Structure from motion with the omniscameras are discussed in [8], [11], [45], [20], [19], [6], [1], [38], [48]. We review the most relevant work.

The idea of autocalibrating catadioptric omnidirectional cameras from point correspondences has appeared in [16], [26]. In [16], the para-catadioptric camera calibration was done using an image of the absolute conic. It was shown that Euclidean reconstruction is feasible from two views with constant parameters and from three views with varying parameters. The paper [26] introduced a calibration technique for para-catadioptric cameras from epipolar geometry by doing bundle adjustment on the distance of point

correspondences to the epipolar curves without any calibration object, knowledge of camera motion, or knowledge of scene geometry. To start the bundle adjustment, the calibration had to be a priori known accurately enough. Moreover, it was assumed that all image correspondences, were correct. Our method also uses point correspondences but we provide a technique applicable to a much wider class of cameras in the form of a direct solution which allows the use of the autocalibration process in the RANSAC robust estimation technique. Compared to [16], [26], our method also allows one to use automatically established point correspondences possibly infected by mismatches.

Our method leads to a Polynomial Eigenvalue Problem, which was first used in camera calibration from point matches in [12] for standard pinhole narrow-angle view cameras. We extend the method to omnidirectional cameras [30], [33] by a linearization of projection models. Our method is not limited to a particular projection model, but can be used for all radially symmetric models.

In [40], a method for interactive 3D reconstruction of piecewise planar objects from a single panoramic view was proposed. The 3D reconstruction was done using geometrical constraints (coplanarity, perpendicularity, and parallelism) provided by the user.

A method for calibrating general central or noncentral cameras proposed in [42] is based on several views of objects with known structure but acquired from unknown viewpoints. Our method does not need any knowledge about the scene since we can model our noncentral cameras from first principles using a small number of parameters compared to [42].

Other calibration methods for omnicaleras have appeared in literature. They used calibration patterns [4], [37], plumb lines [7], [17], [47], or special camera motions, e.g., pure rotation [26], [51]. Our method does not need any calibration objects, no assumptions about the scene are made, and the camera motions may be general except for some degenerate motions.

When an imaging system does not maintain a single viewpoint, a caustic, i.e., a locus of viewpoints in three dimensions, is formed and the system has to be treated as noncentral [9], [21], [46], [39], [42]. Accurate models of noncentral cameras become very complex and may have tens of parameters. Such models are difficult to estimate since many matches are necessary to fix a model and, thus, a workable RANSAC becomes too expensive. We demonstrate that simplified central models can be estimated, mismatches removed, and the simplified models used to initialize a bundle adjustment with a complex but accurate model.

A method for modeling slightly noncentral catadioptric cameras, which is similar to our idea [32], was suggested in [39]. However, they calibrate the camera from known 3D positions of target points, which are not necessary in our case.

A generalization of the “homography-based calibration of radial distortion” [49] also estimates camera models for very wide view angles. The technique needs to observe correspondences on a plane in at least three images (or general points and rotation of the camera). We can calibrate from only two images of a general scene that are related by a general motion.

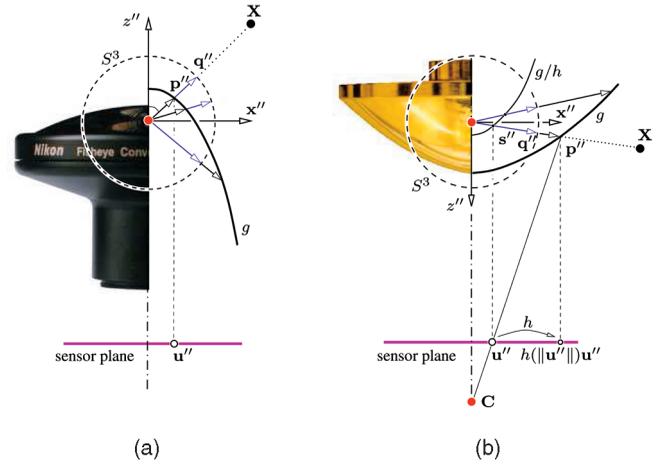


Fig. 3. The mapping of a scene point X into a sensor plane to a point u'' for a fish-eye lens and a hyperbolic mirror.

3 OMNIDIRECTIONAL CAMERA GEOMETRY

Omnidirectional cameras with a large field of view have to be modeled differently than standard perspective cameras with a narrow angle of view. We introduce a new model to represent omnidirectional cameras with an angle of view larger than 180° .

3.1 Image Formation

We assume that the lenses and mirrors are 1) symmetric with regard to an axis and 2) that the axis is perpendicular to the sensor plane.

In practical situations, the axial symmetry is guaranteed by manufacturing. The perpendicularity of the lens axis and the sensor plane is also often guaranteed by camera construction since lenses are mounted directly on cameras. For catadioptric cameras, mirrors have to be carefully placed in front of the cameras to make their symmetry axis perpendicular to the image sensor. The perpendicularity allows the recovery of the affine transformation caused by the digitization process, as described later.

In the following, a fish-eye lens is shown in illustrative images, but the equations hold for both dioptric and catadioptric cameras. By fish-eye lens, we mean a “true fish-eye” lens with a circular field of view (FOV) following the terminology of Fig. 34.5 in [36].

Suppose we are observing a scene point X with an omnidirectional camera, see Fig. 3. The projection of the scene point X onto the unit sphere around the projection center C is represented by unit vector $q'' \in S^3 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ expressed in a sensor Cartesian coordinate system (measured in some world units, e.g., in millimeters). There is always a (possibly nonunit) \mathbb{R}^3 vector $p'' = (x''^T, z'')^T$ linearly dependent with q'' , which maps to the *sensor plane* point $u'' \in \mathbb{R}^2$ so that u'' is *linearly dependent* on x'' as

$$p'' = \begin{pmatrix} h(\|u''\|, a'') u'' \\ g(\|u''\|, a'') \end{pmatrix}, \quad (1)$$

where g, h are functions $\mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$, which depend on the radius $\|u''\|$ of the sensor point with regard to the image of the optical axis (the center of symmetry) and on some parameters $a'' \in \mathbb{R}^N$, where N is the number of parameters. In the following, we will, for the sake of simplicity, write $h(\|u''\|)$

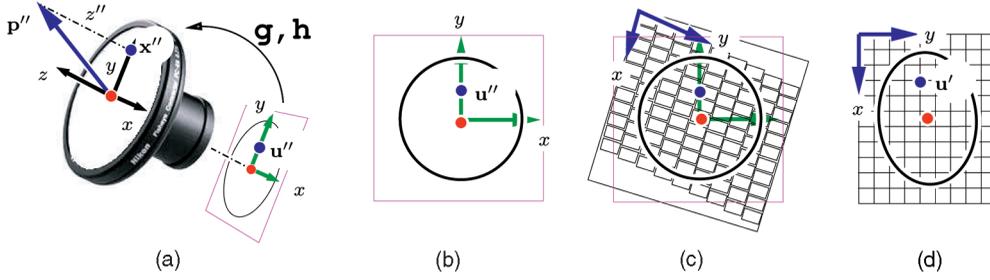


Fig. 4. Omnidirectional image formation. (a) A lens projects scene points to a sensor plane. (b) The sensor plane with a view field circle. (c) The digitization process. (d) An acquired image related to the image on the sensor plane by an affine transformation. The view field circle is transformed to an ellipse.

instead of $h(\|\mathbf{u}''\|, \mathbf{a}'')$. The linear dependence of \mathbf{u}'' on \mathbf{x}'' , i.e., $\mathbf{x}'' = \gamma \mathbf{u}''$ is clear from Fig. 3b, where $\gamma = h(\|\mathbf{u}''\|)$ and, thus, $\mathbf{x}'' = h(\|\mathbf{u}''\|) \mathbf{u}''$.

The functions g, h differ according to the type of lenses and mirrors. For lenses, the functions depend on the type of the lens projection (equisolid, equiangular, etc.) and for the mirrors they depend on the shape of the mirror (parabolic, hyperbolic, elliptical, etc.). The mapping of the vector \mathbf{p}'' to the sensor plane point \mathbf{u}'' through the functions g, h is shown in Fig. 3 for a fish-eye lens and a hyperbolic mirror. For all fish-eye lenses in this work (and for a parabolic mirror as well), $h = 1$ and the vector \mathbf{p}'' is mapped orthographically to the sensor plane, Figs. 4a and 4b.

Suppose a general digital sensor composed of nonsquare pixels is aligned in a “linear” but nonrectangular grid (nonzero skew in the standard terminology of [22]), see Fig. 4c. The digitization process transforms the circular field of view on the sensor into an elliptical one in the digital image, see Fig. 4d. The digitization can be represented by an affine transformation

$$\mathbf{u}'' = \mathbf{A}' \mathbf{u}' + \mathbf{t}', \quad (2)$$

where $\mathbf{u}' \in \mathbb{R}^2$ is a point in the digital image, $\mathbf{A}' \in \mathbb{R}^{2 \times 2}$ is a regular matrix and $\mathbf{t}' \in \mathbb{R}^2$ is a translational vector. The point \mathbf{u}' represents the sensor point \mathbf{u}'' in a nonorthogonal coordinate system with the origin usually placed at the top left corner of the image, see Fig. 4c.

The complete image formation can be written as

$$\begin{aligned} \frac{1}{\alpha''} \mathbf{P}'' \mathbf{X} = \mathbf{p}'' &= \begin{pmatrix} \mathbf{x}'' \\ z'' \end{pmatrix} = \begin{pmatrix} h(\|\mathbf{u}''\|) \mathbf{u}'' \\ g(\|\mathbf{u}''\|) \end{pmatrix} \\ &= \begin{pmatrix} h(\|\mathbf{A}' \mathbf{u}' + \mathbf{t}'\|) (\mathbf{A}' \mathbf{u}' + \mathbf{t}') \\ g(\|\mathbf{A}' \mathbf{u}' + \mathbf{t}'\|) \end{pmatrix}, \end{aligned} \quad (3)$$

where $\mathbf{P}_{3 \times 4}'' = [\mathbf{R}_c'' \mid -\mathbf{R}_c'' \mathbf{t}_c'']$ is a camera projection matrix [22] with regard to a world coordinate system.

3.2 Camera Calibration

The aim of the calibration is to find the mapping from a digital image point \mathbf{u}' to a corresponding 3D ray \mathbf{p}'' . The calibration can be divided into two steps. First, the mapping from the digital image to the sensor plane caused by digitization has to be recovered. Second, the mapping from the sensor plane to rays, caused by the optics (or mirror reflection), has to be found.

The advantage of the circular field of view cameras over the standard ones is that the *full* view field circle is projected into the sensor plane and is thus observable in the digital

image as a view field ellipse, see Fig. 6b. Let us emphasize that seeing the view field ellipse in the image is essential for recovering digitization parameters [22]. Usually, the ellipse is very close to a circle since the pixel skew is negligible and the pixels are nearly square.

We can transform the view field ellipse, which can often be easily obtained by fitting an ellipse to the boundary of the view field in the digital image, to a circle. The transformation can be written again as

$$\mathbf{u} = \mathbf{A} \mathbf{u}' + \mathbf{t}, \quad (4)$$

where $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ and $\mathbf{t} \in \mathbb{R}^2$. The point $\mathbf{u} \in \mathbb{R}^2$ is expressed in a camera Cartesian coordinate system with the origin placed at the center of radial symmetry. The columns of the matrix \mathbf{A} represent the coordinates of the basis vectors of the image coordinate system expressed in the basis of the camera Cartesian coordinate system. The angle between the vectors is the skew and the ratio of the vector lengths is the aspect ratio. The vector \mathbf{t} gives the position of the symmetry center.

There is an ambiguity with two degrees of freedom left in the affine transform generated by mapping an ellipse to a circle. We do not know the radius of the circle and the rotation $\mathbf{R} \in \mathbb{R}^{2 \times 2}$ around the circle center is not fixed. The following relationship between \mathbf{A}', \mathbf{t}' and \mathbf{A}, \mathbf{t} holds

$$\rho \mathbf{t} = \mathbf{R}^{-1} \mathbf{t}', \quad \rho \mathbf{A} = \mathbf{R}^{-1} \mathbf{A}', \quad (5)$$

with the unknown rotation \mathbf{R} and unknown scale ρ .

The process of determining matrix \mathbf{A} and the vector \mathbf{t} is called the *precalibration* step in the rest of the paper. The precalibrated image corresponds to the image acquired by a camera with square pixels, see Fig. 5c. After the precalibration step, the image mapping can be modeled by a radially symmetric function, i.e., in such a precalibrated image, the nonlinear mapping of an image point to its 3D ray can be described by a function of the point distance from the origin.

Using (5) gives the relationship between the point \mathbf{u}'' in the sensor plane and the point \mathbf{u} in the precalibrated image

$$\mathbf{u}'' = \mathbf{A}' \mathbf{u}' + \mathbf{t}', \quad \mathbf{u} = \mathbf{A} \mathbf{u}' + \mathbf{t} \Rightarrow \mathbf{u}'' = \rho \mathbf{R} \mathbf{u}, \quad (6)$$

which means that we are able to compute the point \mathbf{u} in a camera Cartesian coordinate system up to a scale and a rotation with regard to the point \mathbf{u}'' in a sensor Cartesian coordinate system.

Let us further focus on doing autocalibration avoiding the need for a calibration target. Instead, the epipolar constraints will be used. The functions g, h should have a special property with regard to rotation and scaling, which are left

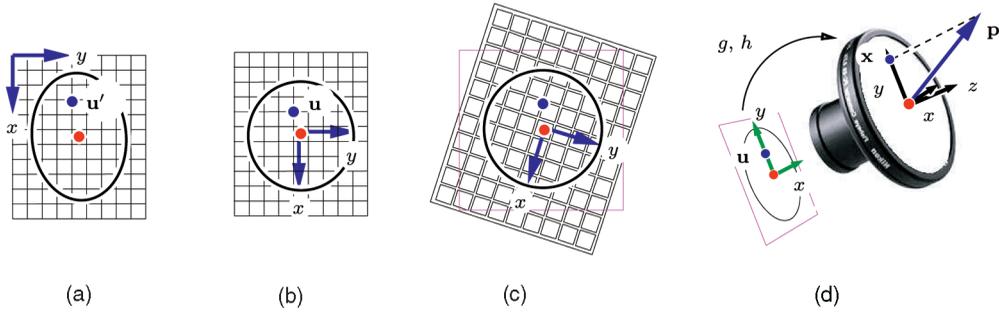


Fig. 5. Omnidirectional camera calibration process. (a) An acquired digital image with an elliptical field of view. (b) A transformed image with a circular field of view. (c) Simulated sensor with zero skew parameter and square pixels. (d) Back-projection of the sensor plane to 3D rays.

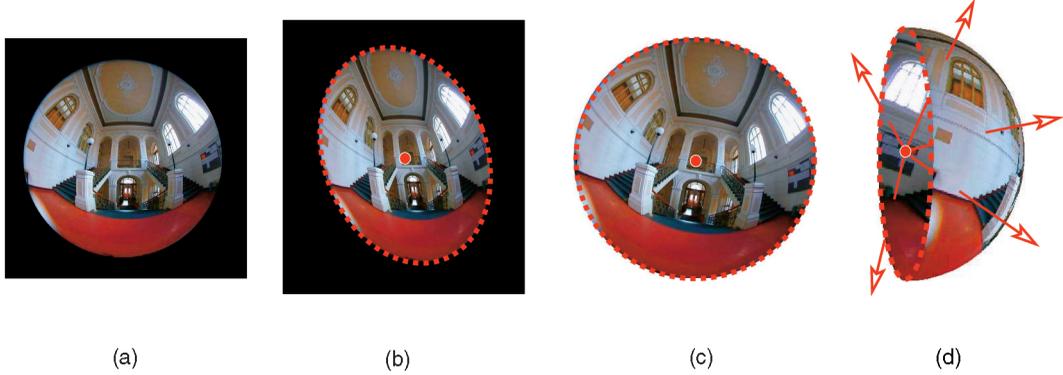


Fig. 6. The calibration process shown on a real image. (a) An image in a sensor plane. (b) An acquired digital image with an ellipse fitted on the view field. (c) A transformed circular precalibrated image. (d) Representation of an image on a sphere.

undetermined in the precalibration step. The rotation is no problem since the functions are radially (i.e., rotationally) symmetric. The scaling, however, might introduce an extra parameter into g and h and would increase the number of parameters by one. This would be possible, in general, but it was not necessary in our case since we were able to design g and h for which the scaling was “absorbed” into the change of other parameters of g and h and, thus, did not introduce any extra parameter. Every extra parameter doubles the RAN-SAC estimation computational cost. Therefore, it is important to keep the number of parameters as low as possible.

The following theorem shows that if functions g and h behave with regard to their parameters, as in (7), the scaling of their argument does not change the resulting ray direction. This is important since the absolute size of image pixels is usually unknown.

Theorem 1. Assume that functions $g: \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$ and $h: \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$ model a concrete projection. If there are parameters $\mathbf{a} \in \mathbb{R}^N$ such that, for all image points $\mathbf{u} \in \mathbb{R}^2$, all $\rho > 0$, and all $\mathbf{a}'' \in \mathbb{R}^N$ holds

$$\begin{aligned} h(\|\rho\mathbf{u}\|, \mathbf{a}'') &= \alpha h(\|\mathbf{u}\|, \mathbf{a}), \\ g(\|\rho\mathbf{u}\|, \mathbf{a}'') &= \alpha \rho g(\|\mathbf{u}\|, \mathbf{a}), \end{aligned} \quad (7)$$

for some $\alpha \in \mathbb{R}$,

$$\mathbf{p}'' \simeq \begin{pmatrix} \mathbb{R} \\ 1 \end{pmatrix} \mathbf{p}. \quad (8)$$

It is important that the functions g , h , which map the point \mathbf{u}'' in the sensor plane to vector \mathbf{p}'' through parameters \mathbf{a}'' , i.e.,

$$\mathbf{u}'' \xrightarrow{g, h(\|\mathbf{u}''\|, \mathbf{a}'')} \mathbf{p}'',$$

can be used to map the point \mathbf{u} in the precalibrated image to the vector \mathbf{p} , i.e.,

$$\mathbf{u} \xrightarrow{g, h(\|\mathbf{u}\|, \mathbf{a})} \mathbf{p},$$

with only the values of parameters being changed. The unknown scale ρ is absorbed into the parameters (for each omnidirectional camera differently) and does not have to be estimated separately. The vectors \mathbf{p}'' and \mathbf{p} differ in their lengths and they are mutually rotated around the optical axis. The change of their lengths and the rotation does not affect the angles between the vectors. Therefore, the camera becomes *metrically* calibrated. The full calibration process is illustrated on a real image in Fig. 6.

All our camera models use functions f and g that satisfy Theorem 1. We were able to design the functions for all common omnidirectional cameras, i.e., for catadioptric cameras with parabolic (one-parametric model), hyperbolic (two-parametric), and spherical mirror (one-parametric when assuming central projection), and for dioptric cameras with Nikon FC-E8 and Sigma 8mm-f4-EX fish-eye converter (one and two-parametric models). See [29] for a detailed description of all the models. Here, we present only those that are necessary to demonstrate our autocalibration method.

3.3 Epipolar Geometry

The epipolar geometry can be formulated for central omnidirectional central cameras, i.e., for catadioptric [43] and for dioptric (with fish-eye lenses) omnidirectional cameras. The epipolar constraint for vectors \mathbf{p}_1'' and \mathbf{p}_2'' reads as

$$\mathbf{p}_2^{\prime\prime\top} \mathbf{E}'' \mathbf{p}_1'' = 0, \quad (9)$$

where \mathbf{E}'' is an essential matrix [22] and \mathbf{p}_1'' and \mathbf{p}_2'' are ray direction vectors of the first and the second camera, respectively.

We have shown in Section 3.2 that we are able to obtain vectors \mathbf{p}_1 , \mathbf{p}_2 that are related to \mathbf{p}_1'' , \mathbf{p}_2'' by (8). Substituting (8) into (9) leads to the constraint

$$\mathbf{p}_2^\top \begin{pmatrix} \mathbf{R}_2 & \\ & \mathbf{1} \end{pmatrix}^\top \mathbf{E}'' \begin{pmatrix} \mathbf{R}_1 & \\ & \mathbf{1} \end{pmatrix} \mathbf{p}_1 = \mathbf{p}_2^\top \mathbf{E} \mathbf{p}_1 = 0. \quad (10)$$

We would like to estimate the matrix \mathbf{E} from a small number of point correspondences, similar to what is done by the 7 or 8-point algorithm [22]. However, the vector \mathbf{p} , (10), cannot be computed since the parameters \mathbf{a} of the functions g , h are unknown. Equation (10) can be rearranged to the following ‘‘homogeneous’’ form:

$$(u_2, v_2, 1) \text{diag}(h_2(\|\mathbf{u}\|), h_2(\|\mathbf{u}\|), g_2(\|\mathbf{u}\|)) \\ \mathbf{E} \text{diag}(h_1(\|\mathbf{u}\|), h_1(\|\mathbf{u}\|), g_1(\|\mathbf{u}\|)) (u_1, v_1, 1)^\top = 0.$$

The unknowns $g_{1,2}$, $h_{1,2}$ cannot be merged with \mathbf{E} into a single matrix \mathbf{F} as can be done with the unknown calibration matrix \mathbf{K} for standard cameras [22], i.e., $\mathbf{u}_2^\top \mathbf{K}_2^\top \mathbf{E} \mathbf{K}_1^{-1} \mathbf{u}_1 = \mathbf{u}_2^\top \mathbf{F} \mathbf{u}_1$, since $g_{1,2}$, $h_{1,2}$ are different for image points with different radii.

In the next section, we propose a method for simultaneous estimation of \mathbf{E} and \mathbf{a} . Correctly estimated parameters \mathbf{a} lead to correctly recovered angles between vectors \mathbf{p} . It allows us to evaluate the quality of the estimated matrix \mathbf{E} and the camera model parameters \mathbf{a} using an angular error measure. The angular error [35] is defined as the minimum (over the normals of epipolar planes) of the sum of squared sines of angles ϕ_1 and ϕ_2 between the two corresponding rays and the epipolar plane, i.e.,

$$\epsilon(\mathbf{q}_1, \mathbf{q}_2, \mathbf{E}) = \min_{\mathbf{n}} (\sin^2 \phi_1 + \sin^2 \phi_2) \\ = \min_{\mathbf{n}} (|\mathbf{n} \cdot \mathbf{q}_1|^2 + |\mathbf{n} \cdot \mathbf{q}_2|^2), \quad (11)$$

where \mathbf{q} stands for \mathbf{p} normalized to unit length and \mathbf{E} stands for the essential matrix. Direct solution of this error for one corresponding pair \mathbf{q}_1 , \mathbf{q}_2 (see [35]) is

$$\epsilon(\mathbf{q}_1, \mathbf{q}_2, \mathbf{E}) = \frac{A}{2} - \sqrt{\frac{A^2}{4} - B}, \quad \text{where} \\ A = \mathbf{q}_1^\top \mathbf{E}^\top \mathbf{E} \mathbf{q}_1 + \mathbf{q}_2^\top \mathbf{E} \mathbf{E}^\top \mathbf{q}_2, B = (\mathbf{q}_2^\top \mathbf{E} \mathbf{q}_1)^2.$$

It could be argued that image reprojection errors would be more appropriate than the angular errors, (11). The image reprojection errors could be evaluated provided that a procedure for evaluating the distance of a point from an epipolar curve was available. Such a procedure can always be given, but it depends on the camera projection model and is often iterative, since for many projection models, there is no direct solution for the distance. For calibrated cameras, the angular error is equivalent to the reprojection error provided that the distribution of noise across the image is correctly taken into account. Suppose there is isotropic Gaussian noise in the image. Then, points on circles centered in image points have equal probability to be generated from those points. If the projection was perspective, then to each such circle would correspond an almost circular cone of rays backprojected into

space. For wide field of view lenses and mirrors, the cone is circular in the middle of the image, but becomes elliptical toward the image boundary. In our experiments with fish-eye lenses, the maximum change between the circle and the ellipse axes was 20 percent. For a catadioptric camera with a spherical mirror with FOV 220°, the change was 40 percent. Thus, to be absolutely correct, the covariance matrices characterizing the noise distribution on angles should be used when evaluating the angular error. However, in all our experiments, we have neglected this when rejecting mismatches and still obtained enough correct matches.

4 AUTOCALIBRATION FROM EPIPOLAR GEOMETRY

We adapt the Fitzgibbon’s technique [12] for the epipolar geometry estimation for standard cameras with radial distortion by solving the Polynomial Eigenvalue Problem (PEP) to omnidirectional cameras with a circular field of view. We introduce new models that relate the radius of a point in the image to the angle contained between the corresponding ray and the optical axis. In [12], the ‘‘division model’’ has been used to get a Quadratic Eigenvalue Problem. We show that the para-catadioptric camera model leads directly to a Quartic Eigenvalue Problem. Furthermore, we show that other algebraic and even nonalgebraic models of catadioptric cameras and fish-eye lenses can be suitably linearized to get a Quadratic Eigenvalue Problem.

In the following, we explain the idea of the autocalibration method and we demonstrate it on a para-catadioptric camera and on an omnidirectional camera with a fish-eye lens. The methods for other omnidirectional cameras (with a hyperbolic or spherical mirror or other lenses) are analogous and some of them can be found in [32], [29].

4.1 Para-Catadioptric Camera

The para-catadioptric camera (PCD), see Fig. 7, is composed of a convex parabolic mirror and an orthographic camera [6]. The orthographic camera is assembled with the parabolic mirror so that the rays of the orthographic camera are parallel to the mirror symmetry axis, see Fig. 7. An image point \mathbf{u} is orthographically projected onto the mirror and reflected such that the ray \mathbf{p} passes through the focal point \mathbf{F} . All rays intersect at \mathbf{F} and, therefore, the PCD camera possesses a central projection.

4.1.1 Camera Model

All coordinates will be expressed in the mirror Cartesian coordinate system placed at \mathbf{F} , see Fig. 7, with the z axis aligned with the axis of the mirror. Let us consider a paraboloid of revolution

$$z = \frac{a''^2 - \|\mathbf{u}''\|^2}{2a''},$$

where $\mathbf{u}'' = (-u'', v'')^\top$ is a point in the sensor plane ($-u''$ since the image is mirrored), a'' is the parameter of the mirror. The model of the PCD camera has the form of 1, where $h(\|\mathbf{u}''\|) = 1$ and

$$g(\|\mathbf{u}''\|) = \frac{a''^2 - \|\mathbf{u}''\|^2}{2a''}.$$

Therefore, when setting $a = a''/\rho$ and taking all $\alpha = 1$, we get from Theorem 1 that vectors

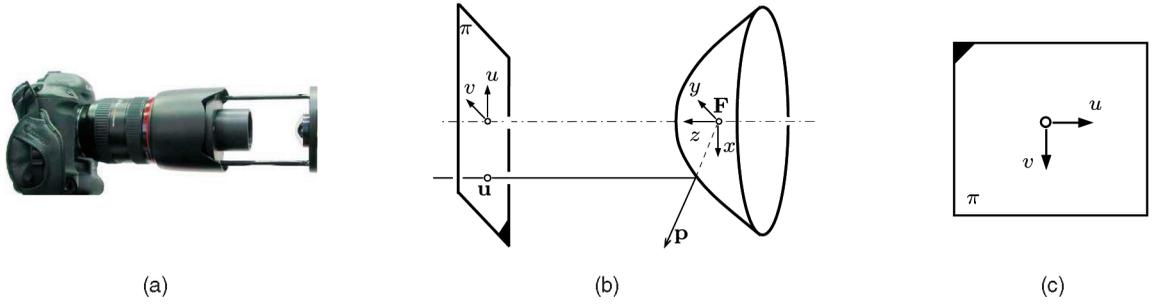


Fig. 7. (a) A para-catadioptric camera consisting of an orthographic camera and a parabolic mirror. (b) Coordinate system of the para-catadioptric camera. The origin is located at F . (c) A coordinate system in the digital image.

$$\mathbf{p} = \begin{pmatrix} \mathbf{u} \\ \frac{a^2 - \|\mathbf{u}\|^2}{2a} \end{pmatrix}$$

are related to vectors \mathbf{p}'' by a common rotation and individual scalings.

4.1.2 Camera Autocalibration

By the calibration of the PCD camera we understand the determination of the matrix \mathbf{A} and the vector \mathbf{t} , (4), and the parameter a of the nonlinear function g so that all vectors \mathbf{p} fulfill the epipolar geometry, (10). The vector \mathbf{p}_1 in the left image and \mathbf{p}_2 in the right image can be substituted into the epipolar constraint $\mathbf{p}_2^T \mathbf{E} \mathbf{p}_1 = 0$, (10), as follows:

$$\begin{pmatrix} -u_2 & v_2 & \frac{a^2 - \|\mathbf{u}_2\|^2}{2a} \end{pmatrix} \mathbf{E} \begin{pmatrix} -u_1 & v_1 & \frac{a^2 - \|\mathbf{u}_1\|^2}{2a} \end{pmatrix}^T = 0. \quad (12)$$

Arranging and gathering the point coordinates and radii into five design matrices, we obtain the equation

$$(\mathbf{D}_1 + a \mathbf{D}_2 + a^2 \mathbf{D}_3 + a^3 \mathbf{D}_4 + a^4 \mathbf{D}_5) \mathbf{f} = \mathbf{0}, \quad (13)$$

which is quartic (degree 4) in parameter a and linear in elements of \mathbf{f} , known as the Polynomial Eigenvalue Problem (PEP) [2]. Efficient algorithms for solving the PEP are available [2], e.g., MATLAB solves the PEP by the function `polyeig`.

Every image point correspondence adds one row to the matrices \mathbf{D}_i . The vector \mathbf{f} of unknowns (created from the matrix \mathbf{E}) and one row from each of the five design matrices \mathbf{D}_i have the following form:

$$\begin{aligned} \mathbf{f} &= [E_{11}, E_{12}, E_{13}, E_{23} \quad \dots \quad E_{33}]^T, \\ \mathbf{D}_1 &= [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ r_2^2 \ r_1], \\ \mathbf{D}_2 &= [0 \ 0 \ 2u_2 \ r_1^2 \ 0 \ 0 \ -2v_2 \ r_1^2 \ 2u_1 \ r_2^2 \ -2v_1 \ r_2^2 \ 0], \\ \mathbf{D}_3 &= [4u_2 \ u_1 \ -4u_2 \ v_1 \ 0 \ 0 \ -4v_2 \ u_1 \ 4v_2 \ v_1 \ 0 \ 0 \ 0 \ -r_2^2 - r_1], \\ \mathbf{D}_4 &= [0 \ 0 \ -2u_2 \ 0 \ 0 \ 0 \ 2v_2 \ -2u_1 \ 2v_1 \ 0], \\ \mathbf{D}_5 &= [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1], \end{aligned} \quad (14)$$

where $r_1 = \|\mathbf{u}_1\|$ and $r_2 = \|\mathbf{u}_2\|$.

Automatically established point correspondences are usually contaminated by mismatches. Since the autocalibration method needs nine points, it can be easily incorporated in a 9-point RANSAC estimation technique to handle outliers in point matches. The computational complexity increases by a factor of 2 for 50 percent contamination by outliers in comparison to the standard linear 8-point RANSAC algorithm [22] for the fundamental matrix computation. The

convergence of the RANSAC algorithm can be speeded up by using a bucketing technique [31].

4.1.3 Experiment

We rotated an off-centered PCD camera, see Fig. 7, on a turntable such that the trajectory of its optical center (the mirror focal point F) was a circle. The PCD camera was set up from a Canon EOS-1Ds digital camera with a view field circle of diameter 2500 pixels and a commercial parabolic mirror [25]. One image was acquired every 9° , in total 40 images. The correspondences for every consecutive pair of images were automatically established and used in the autocalibration process based on the RANSAC with bucketing [31]. In the rest of the paper, if not otherwise specified, the correspondences were found by the wide baseline stereo technique [28].

An essential matrix \mathbf{E} , the camera parameter a , and correct point matches, see Fig. 8, were obtained between every consecutive pair as a result of the autocalibration method. From the essential matrix \mathbf{E} , the relative camera rotation and the camera translation direction were computed [22]. For obtaining the magnitudes of the translational vectors, we would need to reconstruct the observed scene. Instead, we normalized the translational vectors to have unit length since we made the motion such that all translation vectors had the same length. The final trajectory of the optical center can be seen in Fig. 9a.

After calibrating all consecutive image pairs, the parameter \hat{a} was computed as the mean over all estimated a s. We used the estimated \hat{a} to create 3D rays corresponding to image points and applied the standard linear 8-point algorithm [22] to recompute the essential matrices.

It can be seen in Fig. 9 that the estimation of the camera parameter \hat{a} from more pairs of images leads to a “more circular” trajectory. However, the end position does not approach the starting one exactly as it should, neither using the various a s nor using the one \hat{a} for the entire sequence. The error accumulated over the sequence reaches 18° at the end, which is 5 percent relative error. There is approximately 0.5° error in the angle estimation on each pair. The camera orientations, however, were correctly estimated for all images of the sequence, see Fig. 9c.

The error in the trajectory estimation was caused by using a nonideal orthographic camera. It is difficult to make an ideal orthographic camera, i.e., to have all the image rays reflected from the mirror parallel to each other. Often, the PCD camera becomes slightly noncentral and a more complicated imaging model has to be used to obtain a more accurate camera trajectory. The real noncentral catadioptric cameras will be

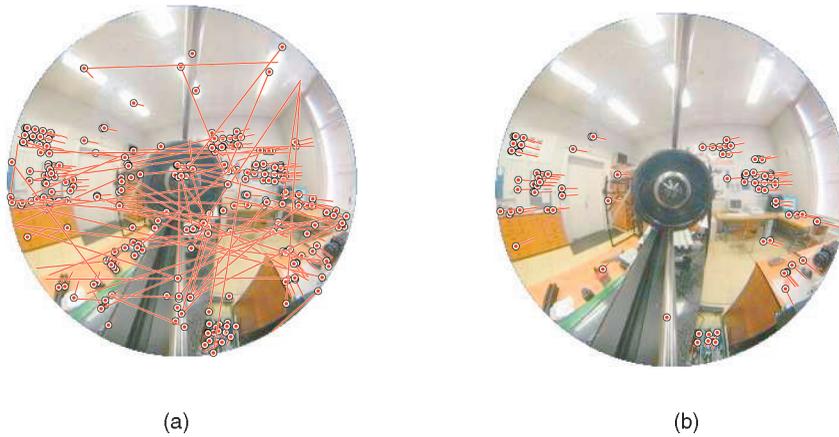


Fig. 8. Outliers removal during the autocalibration process for a PCD camera. (a) An acquired digital image from a stereo pair with all tentative correspondences. Circles mark points in the first image, lines join their matches in the next one. (b) The same image with the validated inliers only.

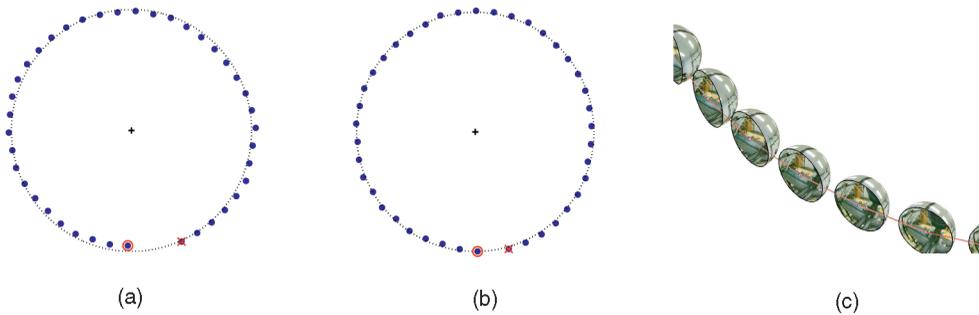


Fig. 9. Estimated trajectories from the circle sequence. Circle \circ depicts the starting position, \times depicts the end position. (a) The trajectory estimated during the autocalibration. (b) The trajectory estimated after calibration using the same \hat{a} for each pair. (c) A detail view of the estimated camera orientations in the circle sequence.

described in Section 5 and it will be shown how the accuracy of camera trajectory estimation increases.

The correspondences, however, were always correctly recovered by using the central model, see Fig. 8. It shows the strength of the method. It is possible, and advisable, to validate the tentative matches using the central model before starting a nonlinear bundle adjustment with a more accurate and complicated model.

4.2 Fish-Eye Lens

Recently, a number of high quality and widely available lenses with angle of view larger than 180° appeared, e.g., the Nikon FC-E8 fish-eye converter for the Nikon COOLPIX digital camera or the Sigma 8mm-f4-EX fish-eye lens for cameras with 35mm film format. The fish-eye lenses with so large a field of view can be regarded as dioptric central omnidirectional cameras, see Fig. 10a.

Depending on the desired accuracy, lens models with various numbers of parameters can be used. Here, we show how to simplify two fish-eye lens models and create a hierarchy of models from the simplest (but not very accurate) one to a more complicated (accurate) one. Even though the simplest model is not very accurate, it is accurate enough to reject many outliers. The more complicated models are more accurate but more points have to be used to estimate them with much bigger computational effort if outliers are present.

The main message of this section is in showing that various lens models can be uniquely treated by suitable linearizations to obtain the PEP even if the functions g, h in (1) are not rational polynomial functions. If they are (such as

for PCD cameras), no linearization is needed to get the PEP formulation of the calibration.

In what follows, we used concrete fish-eye lens models to simplify the explanation, however, the procedure can be repeated for any other model, e.g., central hyperbolic or approximated spherical catadioptric camera [29]. We chose the fish-eye lens model to show a two-parametric model and the treatment of the second parameter. Every two-parametric model can be handled in exactly the same way.

4.2.1 Fish-Eye Lens Models

For the Nikon and the Sigma fish-eye lenses, respectively, we used the following *two-parametric nonlinear models*

$$\theta = \frac{a'' \|\mathbf{u}''\|}{1 + b'' \|\mathbf{u}''\|^2} \quad \text{and} \quad \theta = \frac{1}{b''} \arcsin\left(\frac{b'' \|\mathbf{u}''\|}{a''}\right), \quad (15)$$

where both a'' and b'' are the parameters of the models.

Both models in (15) were obtained experimentally in the following manner: The Nikon and the Sigma lenses approximately implement the equi-angular projection (given by the manufacturer), therefore, we took the equi-angular model with parameter a'' as the base and added an extra parameter b'' to model the lenses more accurately. For $b'' \rightarrow 0$, both models, (15), become the following *one-parametric linear models*

$$\theta = a'' \|\mathbf{u}''\| \quad \text{and} \quad \theta = \|\mathbf{u}''\|/a''. \quad (16)$$

Adding an extra parameter b'' is often tricky. We tried out some combinations of known models [13], [27], [4], [36] and, finally, we ended up with the two models given by (15) as a

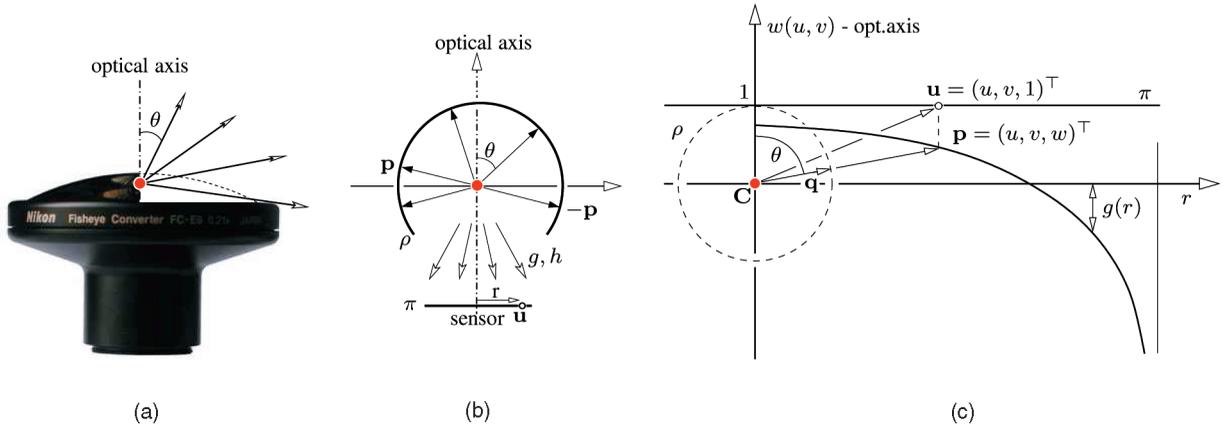


Fig. 10. (a) The Nikon FC-E8 fish-eye converter. All rays emanate from the projection center shown as a red bold dot. (b) The image point on the planar sensor π can be represented by intersecting a spherical retina ρ with camera half-rays. (c) The projection of the vector \mathbf{q} into a sensor plane point \mathbf{u} through a function $g(r)$.

compromise between the accuracy and the number of parameters.

Suppose that the maximal view angle θ_{max} of the fish-eye lens is known (given by a manufacturer). The maximal radius r''_{max} corresponding to the θ_{max} can be directly measured on the view field circle in the precalibrated image. Using (15) allows one to express the parameter a'' as a function of b'' , the radius r''_{max} and the angle θ_{max} . Substituting it into (15) yields the following *one-parametric nonlinear models*

$$\begin{aligned} \theta &= \frac{(1 + b'' r''_{max}{}^2) \theta_{max} \|\mathbf{u}''\|}{r''_{max} (1 + b'' \|\mathbf{u}''\|^2)} \quad \text{and} \\ \theta &= \frac{1}{b''} \arcsin \frac{\|\mathbf{u}''\| \sin(b'' \theta_{max})}{r''_{max}}. \end{aligned} \quad (17)$$

The relationship between the 3D vector \mathbf{p} emanating from the optical center \mathbf{C} toward a scene point and the corresponding sensor plane point \mathbf{u} , see Fig. 10, for the Nikon and the Sigma fish-eye lens, respectively, can be expressed in the camera Cartesian coordinate system, according to (1), as

$$\mathbf{p} = \begin{pmatrix} \mathbf{u} \\ g(\|\mathbf{u}\|, a, b) \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \frac{\|\mathbf{u}\|}{\tan \theta} \end{pmatrix}. \quad (18)$$

Depending on the model used, the θ is given by formulae in (15), (16), or (17). Similarly to the PCD camera, the relation between \mathbf{a} and \mathbf{a}'' can be found without increasing the number of parameters [29] and, therefore, Theorem 1 holds and double primes are further omitted.

When taking the simplifications to the limit, we can set $b'' := 0$ and $a'' := \theta_{max}/r_{max}$ to arrive at a “zero-parametric” model and, thus, convert the problem into the standard problem of rejecting mismatches by estimating the standard epipolar geometry. This may be practical in some situations, especially with equiangular lenses and mirrors.

4.2.2 Camera Autocalibration

The above three models can be used to stratify the autocalibration in three steps: 1) estimate the one-parametric linear model, 2) estimate the one-parametric nonlinear model, and 3) estimate the two-parametric nonlinear

model. Optionally, it would be possible to start with the zero-parametric model to finish earlier.

The estimation hierarchy allows the gradual rejection of outliers depending on their error. The assumption is that the mismatches with large error fit neither the most accurate model nor a less accurate one. The worst outliers can thus be detected by a simpler model and a more complicated two-parametric model can be estimated from data with a lower fraction of outliers. Using the more accurate (but more computationally expensive) model on the inliers found by the simpler models speeds up the convergence of the RANSAC-based estimation and helps to avoid overfitting.

We follow the framework of Section 4.1.2, therefore we stress only the different parts in the derivation process compared to the para-catadioptric (PCD) camera.

Autocalibration for the Two-Parametric Nonlinear Model. The fish-eye model (18) involves $\tan \theta$ and, thus, does not lead directly to a PEP. However, a PEP can be obtained by linearizing the function $g(\cdot)$ in (18) with regard to a and b in a_0 and b_0 ,

$$\begin{aligned} \tilde{g}(r, a, b) &= g(r, a_0, b_0) + g_a(r, a_0, b_0)(a - a_0) \\ &\quad + g_b(r, a_0, b_0)(b - b_0), \end{aligned} \quad (19)$$

where the functions $g_a(\cdot)$, respectively, $g_b(\cdot)$, are the partial derivatives of $g(r, a, b)$ with regard to a , respectively, b , evaluated at the point a_0 and b_0 [29]. Parameter b often represents “distortion” correction of the optics with regard to intended design and then it is possible to linearize in $b_0 = 0$. For a circular field of view cameras, moreover, it is also straightforward to get $a_0 = \theta_{max}/r_{max}$. The vector $\tilde{\mathbf{p}}$ can be written using (19) as follows:

$$\begin{aligned} \tilde{\mathbf{p}} &= \begin{pmatrix} \mathbf{u} \\ g(\cdot) - a_0 g_a(\cdot) - b_0 g_b(\cdot) \end{pmatrix} + a \begin{pmatrix} \mathbf{0} \\ g_a(\cdot) \end{pmatrix} + b \begin{pmatrix} \mathbf{0} \\ g_b(\cdot) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{u} \\ w \end{pmatrix} + a \begin{pmatrix} \mathbf{0} \\ s \end{pmatrix} + b \begin{pmatrix} \mathbf{0} \\ t \end{pmatrix} = \mathbf{x} + a \mathbf{s} + b \mathbf{t}, \end{aligned}$$

where \mathbf{x} , \mathbf{s} , and \mathbf{t} are vectors composed of image coordinates of corresponding points. The epipolar constraint, (10), holds for the linearized vectors, i.e., $\tilde{\mathbf{p}}_2^T \mathbf{E} \tilde{\mathbf{p}}_1 = 0$. After rearranging, we get the following Quadratic Eigenvalue Problem (QEP):

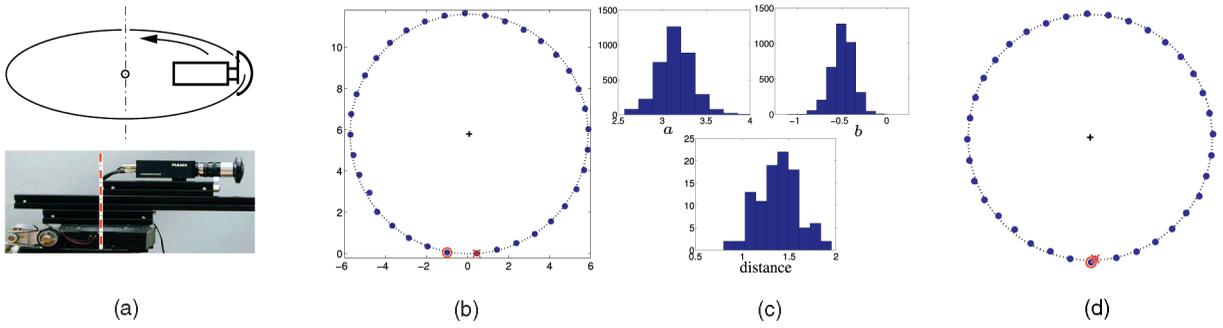


Fig. 11. Estimated trajectories for the circular sequence. Circle \circ depicts the start position, \times depicts the end position. (a) The Nikon fish-eye converter mounted on the digital camera is rotated along a circle. (b) The trajectory estimated during the calibration. (c) Histograms of the parameters a , b and the Euclidean distance of the end point with regard to the starting one computed from 100 runs. (d) The trajectory recomputed when using averaged \hat{a} , \hat{b} .

$$(D_1 + a D_2 + a^2 D_3) \mathbf{l} = \mathbf{0}. \quad (20)$$

Rows of the matrices D_i and the vector \mathbf{l} are as follows:

$$\begin{aligned} D_1 &= \begin{bmatrix} u_1 u_2 & v_1 v_2 & w_1 w_2 & u_1 v_2 & v_1 v_2 & w_1 v_2 & u_1 w_2 & v_1 w_2 & w_1 w_2 \\ t_1 u_2 & t_1 v_2 & t_1 w_2 & u_1 t_2 & v_1 t_2 & w_1 t_2 & u_1 t_2 + v_1 t_2 & t_1 t_2 \end{bmatrix}, \\ D_2 &= \begin{bmatrix} 0 & 0 & s_1 u_2 & 0 & 0 & s_1 v_2 & u_1 s_2 & v_1 s_2 & s_1 w_2 + w_1 s_2 & 0 & 0 & 0 & 0 \\ t_1 s_2 + s_1 t_2 & 0 \end{bmatrix}, \\ D_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s_1 s_2 & 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{l} &= (E_{11} \ E_{12} \ E_{13} \ E_{21} \ \dots \ E_{33} \ b E_{13} \ b E_{23} \ b E_{31} \ b E_{32} \\ &\quad b E_{33} \ b^2 E_{33})^\top. \end{aligned} \quad (21)$$

The vector \mathbf{l} contains elements of the essential matrix and six additional (dependent) products $b E_{ij}$. The parameter b can be determined from any of them. The matrices D_i should be square and, therefore, at least 15 point matches are needed for solving the PEP given by (20). There are 30 possible solutions of a . For every a , six possible solutions of b exist. Since there is noise in data, we choose the solution which best fits our model in terms of the smallest error in (11). With more than 15 points the least-square solution by left-multiplication of (20) by D_1^\top can be used.

The QEP autocalibration method for the one-parametric fish-eye model is straightforward. The vector \mathbf{l} in (20) then contains only elements of the essential matrix, similar to (13). See [29] for more details.

4.2.3 Experiments

We carried out the same circular motion experiment as for the PCD described in Section 4.1.3. A PULNIX digital camera with a Nikon FC-E8 fish-eye lens was used. The diameter of the view field circle was 870 pixels, 36 images in total were acquired with one every 10° . Correspondences were obtained by feature tracking using *boujou* [24] to show that the method works for another matching technique and is not restricted to [28] only. However, a similar result was achieved also for matches found by [28].

The difference from the experiment with the PCD camera in Section 4.1.3 is that two parameters a and b are estimated instead of one and the calibration procedure was done in three steps. First, a 9-point RANSAC with the linear model, (16), and then a 9-point RANSAC with the nonlinear model, (17), as a pretest to detect most of outliers were used. Finally, a 15-point RANSAC with the nonlinear model, (15), was run to

compute parameters a , b , and E . See Fig. 11 for results. To show the robustness, the autocalibration was run 100 times on the same sequence, see histograms in Fig. 11c. The two top histograms show that the distribution of the recovered camera parameters is Gaussian. It therefore makes sense to take mean values as estimates of the correct (unknown) parameter values. The bottom histogram shows the final error in the trajectory end point estimation for 100 runs. None of the runs correctly approaches the starting position. However, taking mean values of the parameters significantly improves the result as shown in Fig. 11d. Ideally, the end position should coincide with the starting position, which approximately holds for the trajectory in Fig. 11d.

In the next experiment, one image pair was selected from the Venice Yard QY data set, acquired by a Sigma 8mm-f4-EX fish-eye lens with view angle 180° and mounted on a Canon EOS-1Ds digital camera with the diameter of the view field circle of 2,528 pixels. The same hierarchical approach as in the previous experiment (the 9-points RANSAC followed by the 15-points one) was applied. The obtained calibrated cameras, essential matrix, and validated point matches were used to reconstruct the scene by a linear technique [22].

The final 3D reconstruction was improved by a nonlinear bundle adjustment minimizing the reprojection error and enforcing the same internal parameters for both cameras. To show the quality of the 3D reconstruction, some correspondences, such as corners on the walls, were picked manually. The estimated camera matrices were used for reconstructing these points. See Fig. 2 for the result. Notice how accurate and complete a 3D reconstruction obtained from only *two* omnidirectional images can be. The RMS of the reprojection error was 0.25 pxl.

4.3 Algorithm

All autocalibration methods presented here are described by the following algorithm. Suppose we have two omnidirectional images of the same rigid scene taken from two different positions by the same camera.

1. Find the ellipse corresponding to the view field of the camera. Transform the images so that the ellipse becomes a circle with radius r_{max} , determine \mathbf{A} and \mathbf{t} . Establish 9 (or 15 for TPM¹ point) correspondences between two images $\{\mathbf{u}_1 \leftrightarrow \mathbf{u}_2\}$.

1. TPM = two-parametric model.

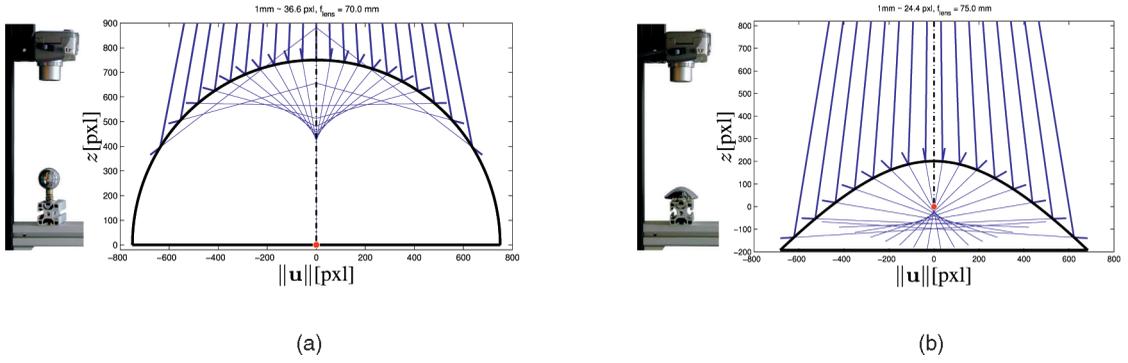


Fig. 12. Reflected rays by (a) a spherical and (b) a hyperbolic noncentral catadioptric camera create caustics.

2. Scale the image points $\mathbf{u} := \mathbf{u}/r_{max}$ to get a better conditioned problem.
3. Create matrices D_i from $\{\mathbf{u}_1 \leftrightarrow \mathbf{u}_2\}$ according to the type of mirror or lens and solve the PEP. Use, e.g., MATLAB: $[\mathbf{H} \mathbf{a}] = \text{polyeig}(D_1, D_2, \dots)$, where \mathbf{H} is the matrix with columns $\mathbf{f} \in \mathbb{R}^{9 \times 1}$ (or $\mathbf{h} \in \mathbb{R}^{15 \times 1}$ for TPM) and \mathbf{a} is a vector with solutions of \mathbf{a} .
4. Choose only real positive “finite” \mathbf{a} s. Typically, one to four solutions remain. For every \mathbf{a} , there is the corresponding essential matrix \mathbf{E} reshaped from vector \mathbf{f} (for TPM the first nine elements of \mathbf{h} give \mathbf{E} , the last six elements give six possible \mathbf{b} s).
5. Compute 3D rays using the \mathbf{a} (and the \mathbf{b} for TPM) for all point correspondences. Compute the error, (11), for all pairs $\{\mathbf{a} \leftrightarrow \mathbf{E}(\leftrightarrow \mathbf{b})\}$ as the sum of errors for all correspondences. The pair with the minimal error is the solution and \mathbf{a} , \mathbf{b} , and the essential matrix \mathbf{E} are obtained.

For integrating the algorithm into RANSAC, 9 (or 15) points are “randomly” selected from the whole set of automatically detected correspondences and Steps 1-4 are repeated until the model fitting the highest number of matches is found.

To speed up the convergence of RANSAC and to avoid a biased estimate, the bucketing technique [31] should be applied. The main idea is to sample points from three concentric zones with equal areas. Moreover, the correspondences in the central part of the image are not used for sampling to estimate the camera model, but they are included in the evaluation of the model fitting error.

Once the mismatches are removed, the Maximum Likelihood estimate known as the Gold Standard method [22] can be used to obtain the optimal estimate of an essential matrix from its initial estimate \mathbf{E} provided by the auto-calibration. In our experiments, we skipped this step. Instead, we directly initiated the bundle adjustment of the final 3D metric reconstruction with \mathbf{E} and always reached an appealing reconstruction.

4.4 Degenerate Configurations and Motions

Suppose that two 3×1 vectors

$$\mathbf{p} = \begin{pmatrix} \mathbf{u} \\ g(\|\mathbf{u}\|, \mathbf{a}) \end{pmatrix} \text{ and } \bar{\mathbf{p}} = \begin{pmatrix} \mathbf{u} \\ g(\|\mathbf{u}\|, \mathbf{a}) + d \end{pmatrix}, \quad d \neq 0, \quad (22)$$

are constructed from the same precalibrated image point \mathbf{u} . Assume that \mathbf{p} , $\bar{\mathbf{p}}$ are such that the epipolar geometry $\mathbf{p}_2^\top \mathbf{E} \mathbf{p}_1 = 0$ and $\bar{\mathbf{p}}_2^\top \mathbf{E} \bar{\mathbf{p}}_1 = 0$ holds for them. To simplify the

situation, assume further that \mathbf{E} represents a pure translation, i.e.,

$$\mathbf{E} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & t_x \\ -t_y & -t_x & 0 \end{pmatrix}.$$

Let $w = g(\|\mathbf{u}\|, \mathbf{a})$, then, using $\bar{\mathbf{p}}_2^\top \mathbf{E} \bar{\mathbf{p}}_1 = 0$,

$$[\mathbf{p}_2^\top + (0, 0, d)] \mathbf{E} [\mathbf{p}_1 + (0, 0, d)^\top] = 0$$

$$d \mathbf{p}_2^\top (t_y, t_x, 0)^\top + d (-t_y, -t_x, 0) \mathbf{p}_1 = 0,$$

which is for $d \neq 0$ equivalent to $u_2 t_y + v_2 t_x - u_1 t_y - v_1 t_x = 0$, where $\mathbf{p}_1 = (u_1, v_1, w_1)^\top$ and $\mathbf{p}_2 = (u_2, v_2, w_2)^\top$.

Theorem 2. Let the motion between two cameras be a pure translation $\mathbf{T} = (t_x, t_y, t_z)^\top$. If, for all correspondences $(u_1, v_1) \leftrightarrow (u_2, v_2)$,

$$u_2 t_y + v_2 t_x - u_1 t_y - v_1 t_x = 0, \quad (23)$$

then the function $g(\cdot)$ can be recovered only up to an additive constant.

If (23) is fulfilled, the ambiguity in the third coordinate of the vector \mathbf{p} for pure translation causes the vector of parameters $\bar{\mathbf{a}}$ instead of \mathbf{a} to be estimated in such a way that the parameters $\bar{\mathbf{a}}$ move the function $g(\cdot)$ along the z axis, i.e., $g(\|\mathbf{u}\|, \bar{\mathbf{a}}) = g(\|\mathbf{u}\|, \mathbf{a}) + d$. However, the essential matrix \mathbf{E} is estimated correctly.

Assume a forward motion of the camera, i.e., motion along the optical axis. Then, $t_x = t_y = 0$ and (23) holds. It means that there remains an ambiguity in the computed parameters of the camera model and the camera cannot be completely calibrated from forward translations.

Assume a sideways motion of the camera, i.e., motion perpendicular to the optical axis. Let $t_x \neq 0, t_y = 0$, then (23) becomes $(v_2 - v_1) t_x = 0$, which is fulfilled iff $v_2 = v_1$. By the sideways motion, the correspondences go from one epipole to the other, see center image in Fig. 2. In the central part of the image, the condition $v_2 = v_1$ approximately holds. It means that, for a sideways motion, correspondences near the view field center should be avoided for point sampling in RANSAC (but still used to evaluate residuals) in the proposed calibration procedure since they always fulfill (23).

5 REAL NONCENTRAL CATADIOPTIC CAMERAS

In practice, many catadioptric cameras are often slightly noncentral. The most common reasons are that 1) a

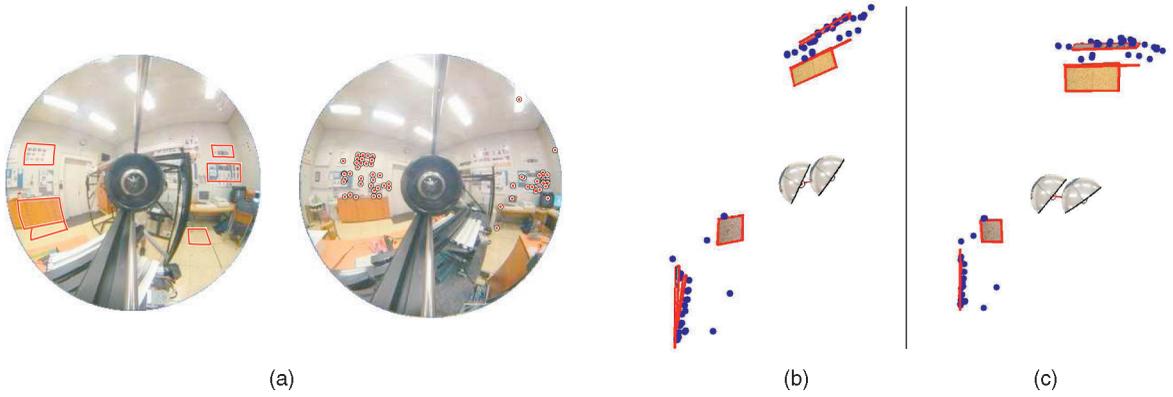


Fig. 13. The top views of the 3D reconstructions of a square room. (a) Input image pair showing manually marked rectangular regions and point matches established automatically. (b) A skewed reconstruction with a central model for a real para-catadioptric (PCD) camera. RMS of reprojection error 2.5 pxl. (c) The correct reconstruction with a noncentral PCD camera model. RMS of reprojection error 0.2 pxl.

nontelecentric lens is used for a parabolic mirror (Fig. 14), or that the lens and the parabolic mirror axes are not properly aligned, 2) a perspective camera is not placed in one of the focal points of a hyperbolic (Fig. 12a) or an elliptical mirror, or 3) the mirror shapes, e.g., the spherical (Fig. 12b) or uniform resolution mirrors [23], [14] are used, which do not possess a single viewpoint property.

Noncentral models of catadioptric cameras may have many parameters. For instance, the models obtained by using the principles of ray optics for a slightly misplaced central camera and a hyperbolic mirror would have 19 parameters. We could hope to estimate it from 19 correspondences if we could compute its minimal solution [34]. It would, however, be overkill to estimate the full model to reject mismatches. We demonstrate that, for common catadioptric cameras, an approximate central model can be designed, calibrated by our method, and used to detect and reject mismatches. On the other hand, using the central camera model for a noncentral camera leads to an inaccurate determination of 3D rays corresponding to image points and, consequently, to a skewed 3D reconstruction, as Fig. 13c shows. A remedy is to estimate the accurate noncentral model from the remaining correct matches by bundle adjustment. We demonstrate that this approach gives good reconstructions even from very low resolution omnidirectional images. See Fig. 13 to compare how the 3D reconstruction was improved by using a noncentral camera model. This result also indicates that formulations leading to generalized eigenvalue problems are stable and work even if the true image projection does not exactly follow the mathematical model used.

We show in this section how to build a 3D reconstruction 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 RANSAC. In the case presented here, the simplified models are central, allowing the mismatches to be rejected automatically through the autocalibration method derived in Section 4. Second, an accurate and complex noncentral model allowing accurate reconstruction is fitted to the correspondences validated in the first step. In general, the art is to find 1) a simplified model that is accurate enough to reject outliers but simple enough to provide a tractable RANSAC estimation problem and 2) a sufficiently accurate

noncentral model providing accurate 3D reconstruction from correct image correspondences.

5.1 Approximate Central Model

For central catadioptric cameras (i.e., parabolic, hyperbolic, elliptical mirrors) [43], it is straightforward and advisable to use their central models to approximate the slightly noncentral projections caused by a misalignment.

For other mirrors (e.g., spherical, mirrors guaranteeing uniform resolution) not possessing center of projection for any camera-mirror setting, a fictive center of projection has to be introduced [32].

5.2 Noncentral Model

Assume a Cartesian world coordinate system W . The Cartesian coordinate system of a mirror, placed at \mathbf{F} , is rotated and translated by $\mathbf{R}_m \in \mathbb{R}^{3 \times 3}$ and $\mathbf{t}_m \in \mathbb{R}^{3 \times 1}$ with regard to W . The Cartesian coordinate system of a perspective camera, placed at the optical center \mathbf{C} , is related to the mirror coordinate system by \mathbf{R}_c and \mathbf{t}_c , see Fig. 14.

From image point to its projection ray. Following the derivation of a central catadioptric camera model [43] and using that rays are reflected by the mirror such that the angles of incident and coincident rays to the surface normal are equal, we can easily derive the full noncentral model [32]

$$\begin{aligned} \mathbf{x}_w &= \mathbf{R}_m^{-1} \mathbf{x} + \mathbf{t}_m = \mathbf{R}_m^{-1} (\lambda \mathbf{R}_c^\top \mathbf{K}^{-1} \mathbf{u} + \mathbf{t}_c) + \mathbf{t}_m, \\ \mathbf{p}_w &= \mathbf{R}_m^{-1} \left(\mathbf{R}_c^\top \mathbf{K}^{-1} \mathbf{u} - 2 \left((\mathbf{R}_c^\top \mathbf{K}^{-1} \mathbf{u}) \cdot \frac{\mathbf{n}}{\|\mathbf{n}\|} \right) \cdot \frac{\mathbf{n}}{\|\mathbf{n}\|} \right), \end{aligned} \quad (24)$$

where \mathbf{R}_i , \mathbf{t}_i are as stated before, \mathbf{K} is a 3×3 upper triangular camera calibration, and \mathbf{n} is a normal vector (different for each type of mirror) to the mirror surface at the point corresponding to the image point \mathbf{u} , see Fig. 14. The number λ comes from parametrization of the ray ν giving the point of intersection with the mirror [32].

Equation (24) represents the complete mapping from an image point \mathbf{u} in the camera coordinate system to a pair $(\mathbf{x}_w; \mathbf{p}_w)$ in the 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 toward a scene point \mathbf{X} . A simulation of the derived model shows the corresponding caustic for a para-catadioptric camera, see Fig. 14.

Projection of a scene point to the image. Given a scene point \mathbf{X} , there is no direct solution for its image projection \mathbf{u} . It

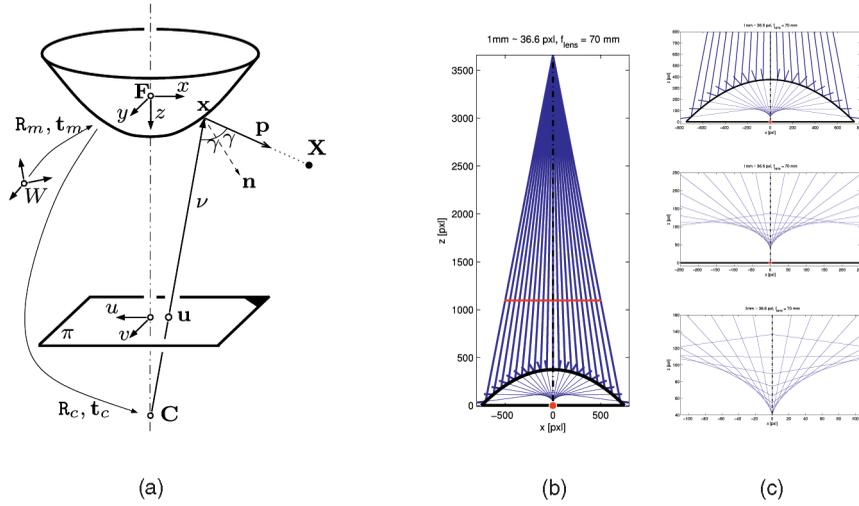


Fig. 14. Noncentral para-catadioptric camera. (a) Coordinate system of the camera. (b) The caustic. Rays reflected by the mirror are tangent to a curve, called the caustic. (c) Detailed views of the caustic.

can be easily designed as an iterative method incorporating the mapping from an image point to its projection ray (derived in the previous section).

First, the initial image \mathbf{u} of the point \mathbf{X} is obtained using a central model for which a direct solution exists [43]. 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 noncentral model, and the 3D point \mathbf{X} . The method converges very quickly because of a good initial estimate of \mathbf{u} provided by the approximate central model.

5.3 3D Reconstruction

Every i th point correspondence can be represented by $(\mathbf{x}_{w1}^i; \mathbf{p}_{w1}^i)$ in the first camera and by $(\mathbf{x}_{w2}^i; \mathbf{p}_{w2}^i)$ in the second one, computed from (24). The reconstructed 3D point \mathbf{X} for one point match $(\mathbf{x}_{w1}^i; \mathbf{p}_{w1}^i) \leftrightarrow (\mathbf{x}_{w2}^i; \mathbf{p}_{w2}^i)$ is obtained as the point in the center of the shortest transversal of the respective projection rays

$$d_i = \frac{|(\mathbf{x}_{w1}^i - \mathbf{x}_{w2}^i) \cdot (\mathbf{p}_{w1}^i \times \mathbf{p}_{w2}^i)|}{|\mathbf{p}_{w1}^i \times \mathbf{p}_{w2}^i|}.$$

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

$$\mathcal{R}_M = \underset{\mathbf{a}, \mathbf{R}_c, \mathbf{t}_c, \mathbf{R}_m, \mathbf{t}_m, \mathbf{K}}{\operatorname{argmin}} \sum_{i=1}^N d_i^2, \quad (25)$$

where N is the number of point matches.

5.4 Experiments

We present two experiments, 1) a parabolic mirror theoretically having the central model and 2) a spherical mirror having only the noncentral model. For both cases, the approximated central models were accurate enough to provide initial estimates for further nonlinear minimization.

5.4.1 Noncentral Para-Catadioptric Camera

We acquired two images using a real para-catadioptric camera (the same parabolic mirror as in the experiment in Section 4.1.3 and a Canon PowerShot G2 digital camera with

the diameter of the view field circle of 1,474 pixels $\sim 210^\circ$, which is equivalent to 351 pixels for the common 50° view angle).

The camera model parameter and relative camera position encoded in the essential matrix (obtained by the autocalibration method described in Section 4.1.2) were used to perform an initial 3D reconstruction using the triangulation method [22] with an RMS of reprojection error of 2.5 pxl. The 3D reconstruction was improved by the Levenberg-Marquardt bundle adjustment using the full noncentral model in (24) minimizing the error in (25). The square pixel and zero skew of the perspective camera were assumed.

The final RMS of the reprojection error was 0.19 pxl. Minimization of the reprojection error in the images instead of shortest transversals in 3D lead to the almost equal RMS error of 0.21 pxl. However, the time increased eight times due to the iterative backprojection model. Thus, the first error using the shortest transversal is more practical.

To show the quality of the 3D reconstruction, some correspondences, such as corners on the walls, have been established manually. The estimated camera matrices were used to reconstruct these points. See Fig. 13c for the result.

5.4.2 Noncentral Spherical-Catadioptric Camera

The same experiment with the same digital camera as was done for para-catadioptric camera, described in the previous section, was carried out for a spherical mirror (Fig. 12a). See Fig. 15 and notice the accurate 3D reconstruction that can be obtained even for very low resolution images with 1,042 pixels for 252° , which corresponds to 208 pixels for the common 50° view angle.

6 CONCLUSION

The paper presented the theory and practice of two-view geometry and structure from motion for all common central omnidirectional cameras with a circular field of view. An algorithm for camera autocalibration and 3D metric reconstruction was presented and demonstrated. The paper extended a nonlinear camera model estimation using epipolar geometry, introduced by Fitzgibbon, to omnidirectional

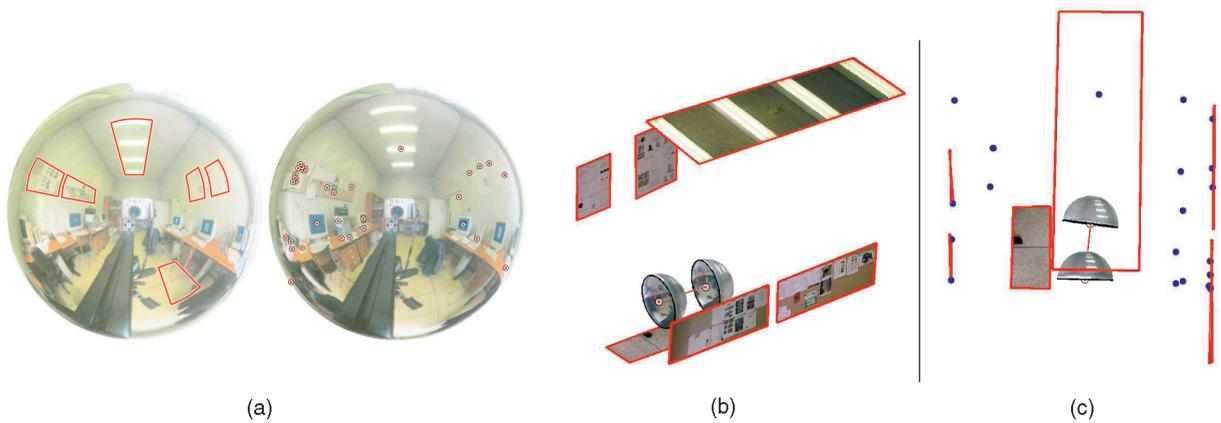


Fig. 15. 3D reconstruction with a spherical mirror. (a) Input image pair together with manually marked rectangular regions and automatically detected and validated point correspondences. (b) 3D metric reconstruction from two uncalibrated spherical catadioptric images. (c) Top view of the room. Notice that the reconstructed plan is really rectangular.

cameras with an angle of view larger than 180° . It was shown that omnidirectional cameras with the circular FOV can be, in practical situations, fully autocalibrated from point correspondences by solving Polynomial Eigenvalue Problems. In particular, we have 1) designed a robust method allowing a 3D metric reconstruction to be built from two uncalibrated omnidirectional images automatically from point correspondences, 2) formulated a theory of image formation and autocalibration of central omnidirectional cameras with a circular field of view, 3) demonstrated that the proposed autocalibration method can be applied to real (often non-central) catadioptric cameras to obtain initial values for an optimization process with a noncentral model giving accurate 3D reconstruction.

APPENDIX

STABILITY OF THE POLYNOMIAL EIGENVALUE PROBLEM

Stability of the Quadratic Eigenvalue Problems is a well-understood mathematical problem. We refer the reader to [2], [50], especially to Section 4 (Perturbation Analysis) in [50] where the stability of the Quadratic Eigenvalue Problem is studied. The accuracy of a numerical solver can be measured by the backward error [50]. See [50, p. 267, Table 5.1] that shows how the backward error decreases with the size of an eigenvalue. In all our cases, estimated parameters corresponding to the eigenvalues are larger than 0.8 and, thus, the most sensitive area around 0 was avoided. The magnitude of the estimated parameter can be, to some extent, controlled by normalizing the image coordinates in the second step of the algorithm described in Section 4.3.

In general, a polynomial eigenvalue problem of order l is well-posed if at least one of the two matrices D_1 and D_l is regular because then it can be "linearized" and transformed to a generalized eigenvalue problem [2]. If they are both singular, the problem is potentially ill-posed. The solutions might not exist or might not be unique.

For the quadratic eigenvalue problem in (20), the 15×15 matrix D_1 in (21) is regular, therefore it guarantees a well-posed problem. Except for the para-catadioptric camera case, all other cameras presented here lead to a

quadratic eigenvalue problem. In case of the quartic eigenvalue problem for the para-catadioptric camera, (13), both 9×9 matrices D_1 , D_5 , (14), are singular. Thus, the problem could be ill-posed and there could be a problem with stability of solutions. We, nevertheless, always obtained results useful for further optimization.

Only real solutions of the PEP are useful and zero solutions can be directly discarded. The number of solutions does not play a role since we choose the one which provides the smallest error in (11). If the PEP does not have a real solution, it indicates that there is large noise or an outlier in point correspondences used to build the matrices D_i in (13) or (20). Since we use RANSAC, we simply skip such cases and use another 9 or 15 points to build the matrices.

ACKNOWLEDGMENTS

This research was supported by MSM 212300013, EU Project Dirac FP6-IST-027787, and IS-3D 1ET101210406 grants.

REFERENCES

- [1] D.G. Aliaga, "Accurate Catadioptric Calibration for Real-Time Pose Estimation in Room-Size Environments," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 127-134, 2001.
- [2] *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst eds., Philadelphia: SIAM, 2000.
- [3] S. Baker and S.K. Nayar, "A Theory of Single-Viewpoint Catadioptric Image Formation," *Int'l J. Computer Vision*, vol. 35, no. 2, pp. 175-196, 1999.
- [4] H. Bakstein and T. Pajdla, "Panoramic Mosaicing with a 180° Field of View Lens," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 60-67, 2002.
- [5] J.P. Barreto and H. Araujo, "Issues on the Geometry of Central Catadioptric Image Formation," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. II:422-427, 2001.
- [6] *Panoramic Vision: Sensors, Theory, and Applications*, R. Benosman and S.B. Kang, eds., New York: Springer Verlag, 2001.
- [7] C. Bräuer-Burchardt and K. Voss, "A New Algorithm to Correct Fish-Eye- and Strong Wide-Angle-Lens-Distortion from Single Images," *Proc. IEEE Conf. Image Processing*, pp. 225-228, 2001.
- [8] T. Brodsky, C. Fermuller, and Y. Aloimonos, "Directions of Motion Fields Are Hardly Ever Ambiguous," *Proc. European Conf. Computer Vision*, 1996.
- [9] S. Derrien and K. Konolige, "Approximating a Single Viewpoint in Panoramic Imaging Devices," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 85-90, 2000.

- [10] O. Faugeras, Q. Tuan Luong, and T. Papadopoulos, *The Geometry of Multiple Images: The Laws that Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. Cambridge, Mass.: MIT Press, 2001.
- [11] C. Fermuller and Y. Aloimonos, "Ambiguity in Structure from Motion: Sphere vs. Plane," *Int'l J. Computer Vision*, vol. 28, pp. 137-154, 1998.
- [12] A. Fitzgibbon, "Simultaneous Linear Estimation of Multiple View Geometry and Lens Distortion," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 125-132, 2001.
- [13] M.M. Fleck, "Perspective Projection: The Wrong Imaging Model," Technical Report TR 95-01, Dept. of Computer Science, Univ. of Iowa, 1995.
- [14] S. Gächter, T. Pajdla, and B. Mičušík, "Mirror Design for an Omnidirectional Camera with a Space Variant Imager," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 99-105, 2001.
- [15] C. Geyer and K. Daniilidis, "A Unifying Theory for Central Panoramic Systems and Practical Applications," *Proc. European Conf. Computer Vision*, 2000.
- [16] C. Geyer and K. Daniilidis, "Structure and Motion from Uncalibrated Catadioptric Views," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 279-286, 2001.
- [17] C. Geyer and K. Daniilidis, "Para-Catadioptric Camera Calibration," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 687-695, May 2002.
- [18] C. Geyer and K. Daniilidis, "Mirrors in Motion: Epipolar Geometry and Motion Estimation," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 766-773, 2003.
- [19] C. Geyer and K. Daniilidis, "Omnidirectional Video," *The Visual Computer*, vol. 19, no. 6, pp. 405-416, 2003.
- [20] J. Gluckman and S.K. Nayar, "Ego-Motion and Omnidirectional Cameras," *Proc. Int'l Conf. Computer Vision*, pp. 999-1005, 1998.
- [21] M.D. Grossberg and S.K. Nayar, "A General Imaging Model and a Method for Finding Its Parameters," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 108-115, 2001.
- [22] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [23] A. Hicks and R. Bajcsy, "Catadioptric Sensors that Approximate Wide-Angle Perspective Projections," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 97-103, 2000.
- [24] <http://www.2d3.com>, 2006.
- [25] <http://www.remotereality.com>, 2006.
- [26] S.B. Kang, "Catadioptric Self-Calibration," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 201-207, 2000.
- [27] J. Kumlér and M. Bauer, "Fish-Eye Lens Designs and Their Relative Performance," (*SPIE Proc.—Current Developments in Lens Design and Optical Systems Eng.*, vol. 4093, pp. 360-369, 2000).
- [28] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions," *Proc. British Machine Vision Conf.*, vol. 1, pp. 384-393, 2002.
- [29] B. Mičušík, "Two View Geometry of Omnidirectional Cameras," PhD thesis, Dept. of Cybernetics, Faculty of Electrical Eng., Czech Technical Univ., Prague, Czech Republic, 2004.
- [30] B. Mičušík and T. Pajdla, "Estimation of Omnidirectional Camera Model from Epipolar Geometry," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 485-490, 2003.
- [31] B. Mičušík and T. Pajdla, "Omnidirectional Camera Model and Epipolar Geometry Estimation by RANSAC with Bucketing," *Proc. Scandinavian Conf. Image Analysis*, 2003.
- [32] B. Mičušík and T. Pajdla, "Autocalibration and 3D Reconstruction with Noncentral Catadioptric Cameras," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2004.
- [33] B. Mičušík and T. Pajdla, "Para-Catadioptric Camera Autocalibration from Epipolar Geometry," *Proc. Asian Conf. Computer Vision*, vol. 2, pp. 748-753, 2004.
- [34] D. Nister, "A Minimal Solution to the Generalised 3-Point Pose Problem," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1 pp. 560-567, 2004.
- [35] J. Oliensis, "Exact Two-Image Structure from Motion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1618-1633, Dec. 2002.
- [36] S.F. Ray, *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging*, third ed. Oxford, U.K.: Focal Press, 2002.
- [37] S. Shah and J.K. Aggarwal, "Intrinsic Parameter Calibration Procedure for a (High Distortion) Fish-Eye Lens Camera with Distortion Model and Accuracy Estimation," *Pattern Recognition*, vol. 29, no. 11, pp. 1775-1788, 1996.
- [38] O. Shakernia, R. Vidal, and S. Sastry, "Omnidirectional Vision-Based Egomotion Estimation from Backprojection Flows," *Proc. IEEE Workshop Omnidirectional Vision*, 2003.
- [39] D. Strelow, J. Mishler, D. Koes, and S. Singh, "Precise Omnidirectional Camera Calibration," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 689-694, 2001.
- [40] P. Sturm, "A Method for 3D Reconstruction of Piecewise Planar Objects from Single Panoramic Images," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 119-126, 2000.
- [41] P. Sturm, "Mixing Catadioptric and Perspective Cameras," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 60-67, 2002.
- [42] P. Sturm and S. Ramalingam, "A Generic Concept for Camera Calibration," *Proc. European Conf. Computer Vision*, pp. 1-13, 2004.
- [43] T. Svoboda and T. Pajdla, "Epipolar Geometry for Central Catadioptric Cameras," *Int'l J. Computer Vision*, vol. 49, no. 1, pp. 23-37, 2002.
- [44] T. Svoboda, T. Pajdla, and V. Hlaváč, "Epipolar Geometry for Panoramic Cameras," *Proc. European Conf. Computer Vision*, pp. 218-232, 1998.
- [45] T. Svoboda, T. Pajdla, and V. Hlaváč, "Motion Estimation Using Central Panoramic Cameras," *Proc. IEEE Int'l Conf. Intelligent Vehicles*, pp. 335-340, 1998.
- [46] R. Swaminathan, M.D. Grossberg, and S.K. Nayar, "Caustics of Catadioptric Cameras," *Proc. IEEE Int'l Conf. Computer Vision*, vol. 2, pp. 2-9, 2001.
- [47] R. Swaminathan and S.K. Nayar, "Nonmetric Calibration of Wide-Angle Lenses and Polycameras," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1172-1178, Oct. 2000.
- [48] C.J. Taylor, "Videoplus: A Method for Capturing the Structure and Appearance of Immersive Environments," *IEEE Trans. Visualization and Computer Graphics*, vol. 8, no. 2, pp. 171-182, Apr.-June 2003.
- [49] S. Thirithala and M. Pollefeys, "The Radial Trifocal Tensor: A Tool for Calibrating the Radial Distortion of Wide-Angle Cameras," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 321-328, 2005.
- [50] F. Tisseur and K. Meerbergen, "The Quadratic Eigenvalue Problem," *SIAM Rev.*, vol. 43, no. 2, pp. 235-286, 2001.
- [51] Y. Xiong and K. Turkowski, "Creating Image-Based VR Using a Self-Calibrating Fisheye Lens," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 237-243, 1997.
- [52] X. Ying and Z. Hu, "We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model?" *Proc. European Conf. Computer Vision*, 2004.



Branislav Mičušík received the MSc degree (with honors) in electrical engineering from the Slovak University of Technology in Bratislava in 2001 and the PhD degree from the Czech Technical University in Prague in 2004. This was followed by a postdoctoral fellowship at the Vienna University of Technology. His research interests include omnidirectional vision, self-calibration, and epipolar geometry estimation of panoramic cameras, image segmentation, texture detection, and image understanding.



Tomáš Pajdla received the MSc and PhD degrees from the Czech Technical University in Prague. He coauthored works that introduced epipolar geometry of panoramic cameras, investigated the use of panoramic images for robot localization, contributed to studies of panoramic mosaics, and studied noncentral cameras and generalized epipolar geometries. He is a member of the IEEE, ACM, and the Czech Pattern Recognition Society. He coauthored works awarded the best paper prize at OAGM '98 and BMVC '02.