International Journal of Computer Vision manuscript No.
(will be inserted by the editor)

Radial Multi-focal Tensors

Applications to Omnidirectional Camera Calibration

SriRam Thirthala · Marc Pollefeys

Received: date / Accepted: date

Abstract The 1D radial camera maps all points on a plane, containing the principal axis, onto the radial line which is the intersection of that plane and the image plane. It is a sufficiently general model to express both central and non-central cameras, since the only assumption it makes is of known center of distortion. In this paper, we study the multi-focal tensors arising out of 1D radial cameras. There exist no two-view constraints (like the fundamental matrix) for 1D radial cameras. However, the 3-view and 4-view cases are interesting. For the 4-view case we have the radial quadrifocal tensor, which has 15 d.o.f and 2 internal constraints. For the 3-view case, we have the radial trifocal tensor, which has 7 d.o.f and **no internal constraints**. Under the assumption of a purely rotating central camera, this can be used to do a **non-parametric estimation** of the radial distortion of a 1D camera. Even in the case of a non-rotating camera it can be used to do parametric estimation, assuming a planar scene. Finally we examine the mixed trifocal tensor, which models the case of two 1D radial cameras and one standard pin-hole camera. Of the above radial multifocal tensors, only the radial trifocal tensor is useful practically, since it doesn't require any knowledge of the scene and is extremely robust. We demonstrate results based on real-images for this.

For the quadrifocal tensor, too, we present a way to do a metric reconstruction of the scene and to undistort the image (given a sufficiently dense set of point-correspondences). We also show results on synthetic images. However, it must be noted that currently the quadrifocal and mixed trifocal tensors are useful *only* from a theoretical stand-point.

SriRam Thirthala
Dept of Computer Science
UNC-Chapel Hill
E-mail: tvnsriram@google.com
Present address: Google Engineering, Bangalore

Marc Pollefeys
Dept of Computer Science
UNC-Chapel Hill,
Institute of Computational Sciences
ETH-Zurich
E-mail: marc.pollefeys@inf.ethz.ch

Keywords Calibration · Non-Standard Cameras · Auto-Calibration · Radial Distortion · Multi-view tensors · Quadrifocal Tensor · Trifocal Tensor · Fish-eye lenses · Catadioptric Cameras

1 Introduction

Omnidirectional cameras have found increasing use in medical imaging [35], surveillance and robot navigation [3]. Wider field-of-view is obtained using fish-eye lenses or by using the camera in combination with a mirror. There has been previous work [1] which tries to identify the conditions under which the resulting camera would be a central camera.

The work described in this paper is useful for both central omnidirectional cameras and perspective cameras (with severe radial distortion). Hence here we briefly note work done in camera calibration in both the fields.

Calibration techniques, which assume a generalized model for the camera, have been discussed in literature [13,39,33,32,14,29,8,7]. However these require knowledge about the scene and/or camera motion. Further there have been approaches [15] that think of the cameras as a bunch of photocells, and try to find the angualr separation of these photocells, based on the (dis)similarity of the signals (received by each of the cells). There has been other prior work that tries to solve the problem only for a particular type of camera. For example, there have been approaches tailored to calibrating fish-eye lenses [37,40,6,2,24]. Some of these additionally require knowledge of the scene [51]. Geyer and Danalidis [11] came up with approaches for parabolic catadioptric systems. An angular-error minimizatio approach, based on a similar camera-model, was used in [25]. Pajdla et. al. [28] introduced a stratified approach that reconstructs the scene and estimates the camera parameters in the spirit of Structure-from-Motion.

There have been approaches that try to estimate radial distortion based on knowledge of the scene [49,12,50]. Recently there has been work [18,26,22,21] that tries to estimate radial distortion *and* center of distortion using a planar calibration grid. Some of these [18] have been extended to do auto-calibration, but the methods are too sensitive to noise. Tardif et. al., in [41], introduced two plane-based approaches. A plumb-line type one and another based on the assumption that one can generate dense correspondences across two views. Both approaches assume a known distortion center, though a minimization-based method is given to estimate it.

Many methods have been proposed that try to exploit the following property of the pin-hole model: Straight lines in the scene must project onto straight lines in the image [4,40]. Kang [23] used snakes to represent distortion curves. Devernay and Faugeras [6] proposed an approach in which the system does edge-detection, followed by polygonal approximation, to group edgels which could possibly have come from an edge segment. The system then tries to minimize the distortion error by optimizing over the distortion parameters. This is done iteratively till the relative change in error is below a threshold.

Requiring knowledge of the scene is a serious limitation, for the first catgeory of methods, because it makes them unsuitable for situations where the camera lens geometry might change (variable zoom etc.). For the second category of methods, one requires straight-lines in the scene. Further, differentiating straight-lines which are curved, due to distortion, and real-world curves is a non-trivial task.

The third category of methods does so by using point correspondences across multiple views. Stein [38] proposed a method based on epipolar and trifocal constraints.

The objective here is to minimize the reprojection error over distortion parameters. Fitzgibbon [10] proposed a technique, for estimating small distortions, that does a simultaneous estimation of the fundamental matrix and a single distortion parameter by formulating a QEP [46]. Micusik and Pajdla [28] extended this method to cameras that produce large distortions. All these approaches assume a specific (parametric) model for radial distortion. In fact, a lot of the analysis goes into investigating the interaction of the assumed model with the multi-view constraints. Recently, Tardif et. al. [42] have come up with an approach that does non-parametric radial calibration using two images of a planar scene. This is achieved by formulating the problem as a convex optimization, by introducing an approximation.

One of the central motivations of the radial 1D camera is to achieve a method based on point correspondences that does reconstruction of the scene *without having to make any assumption about the kind of radial distortion* (other than the sufficiently general one of radial symmetry). Thus both the methods, based on the radial trifocal and quadrifocal tensor, follow the 2-step process. One, to estimate the scene structure independent of unknown distortion. Two, to use the reconstructed scene as a calibration object to estimate a non-parametric model of radial distortion.

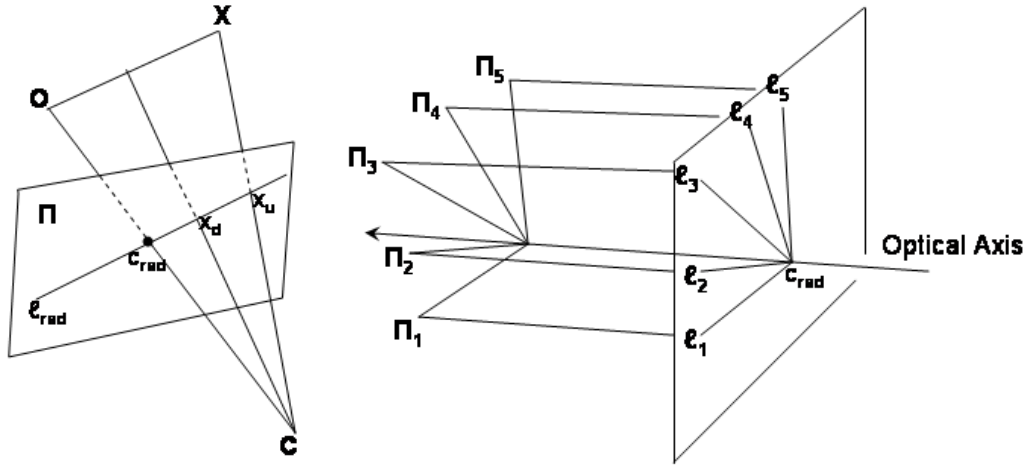
We have tabulated the results, discussed in this paper, in Table 1 on page number 28. The reader is encouraged to use it, as a reference, as we study the radial 1D camera and the different multi-focal constraints. We would also like to note that the work presented in this paper was previously introduced in [43], [44] and [45].

2 Radial 1D Camera

Suppose that the center of radial distortion is known. In the absence of any information, the image center is a good approximation for the center of distortion. However, if we have more information (for example, if the rim of the mirror/fish-eye is visible), we can use that information too. The image can then be transformed such that the center of radial distortion is the origin. Consider a point in the world $\mathbf{X}$ that projects onto $\mathbf{x}_d = (x_d, y_d, 1)^T$ in the distorted (input) image. Further let $\mathbf{C}$ be the camera center. Because of large unknown and possibly varying distortion, the point $\mathbf{X}$ does not lie on the ray passing through $\mathbf{C}$ and $\mathbf{x}_d$ (see Figure 1(a)).

However, consider the line passing through the center of radial distortion and $\mathbf{x}_d$ in the image ($\mathbf{l_{rad}} = \mathbf{x}_d \times \mathbf{c_{rad}}$). The undistorted image point (one that would have been obtained if the camera had followed a pin-hole projection model) $\mathbf{x}_u$, would lie on this line. This is because though the distance of an image point from the center of radial distortion is not preserved by radial distortion, the direction (which is what the radial line $\mathbf{l_{rad}}$ encodes) is. If instead of back-projecting a ray, we back-project the line $\mathbf{l_{rad}}$ using the camera center, it would contain the ray passing through $\mathbf{C}$ and $\mathbf{x}_u$, and thus would contain $\mathbf{X}$. Thus, by representing the distorted image as a 1D image of radial lines passing through the center of radial distortion, we can factor out the *unknown deviation* from the pinhole model (which is along the radial line), but preserve the *known direction* of radial line. The insight of separating the unknown deviation but preserving the known direction, has been used earlier by Tsai [49] in the context of camera calibration using a grid. Tsai referred to this as the *radial alignment constraint*.

The radial 1D camera can be thought of as projecting the pencil of planes containing the optical axis onto the pencil of lines passing through the $\mathbf{c_{rad}}$ (Figure 1(b)). A radial line can be represented as $\mathbf{l} = (y, x)^T$ if $\mathbf{c}_{rad}$ has been mapped to the origin. Note that a



(a) (b)

Fig. 1: *Radial 1D Camera*

radial 1D camera can be obtained for most single effective viewpoint cameras (standard pin-hole cameras, low radial distortion cameras, fish-eye lenses, catadioptric cameras [1]). In fact we can deal with non-central cameras also. The only requirement is that all points that lie in one plane, of the pencil around the optical axis, project onto the same radial line (passing through $\mathbf{c}_{rad}$). For catadioptric systems, this corresponds to the requirement that all the normals on the mirror have to be contained within radial planes. This constraint is automatically satisfied for mirror shapes that are symmetric around the optical axis.

Definition: The radial 1D camera represents the mapping of a point in $\mathbb{P}^3$ onto a radial line in the image. Since it is a $\mathbb{P}^3 \rightarrow \mathbb{P}^1$, it can be represented by a 2×4 matrix and has 7 degrees of freedom.

The projection of a 3D point $\mathbf{X}$ on a radial line $\mathbf{l}$ using radial camera $\mathbf{P}$ is then given by:

$$\lambda \mathbf{l} = \mathbf{P}\mathbf{X} \tag{1}$$

with λ a non-zero scale factor. Note that the image of a point $\mathbf{U}$ on the optical axis does not have a proper image in $\mathbb{P}^1$ as we obtain $\mathbf{P}\mathbf{U} = (0,0)^T$. Since $\mathbf{l}$ $(= [l_1\ l_2]^T)$ and $\mathbf{P}$ $(= [\mathbf{p}_1\ \mathbf{p}_2]^T)$ are equal up to scale, we get:

$$(l_2\mathbf{p}_1 - l_1\mathbf{p}_2)^T\mathbf{X} = 0 \tag{2}$$

Comparing this with the equation of a point $\mathbf{X}$ lying on a plane $\mathbf{\Pi}$, i.e. $\mathbf{\Pi}^T\mathbf{X} = 0$, we get that the plane back-projected by the radial line is

$$\mathbf{\Pi} = (l_2\mathbf{p}_1 - l_1\mathbf{p}_2) \tag{3}$$

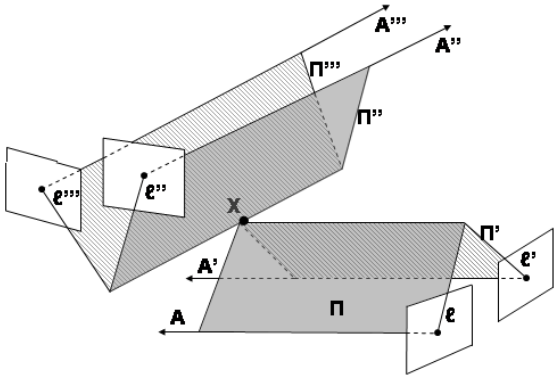


Fig. 2: *The Quadrifocal Constraint*

3 Radial Quadrifocal Tensors

Let us examine the possible multi-view constraints using these 1D radial cameras in general configuration. Note that we only have back-projected planes and no back-projected rays (as the distance information from the $\mathbf{c}_{rad}$ is unknown, only the radial line is preserved as a set). Three planes in 3D-space in general position intersect at a point and hence 3 views give us no constraints. However, four planes intersecting at a common point yields a non-trivial constraint (Figure 2). Thus we have multi-view constraints among four 1D radial cameras.

Consider a point $\mathbf{X}$ in $\mathbf{P}^3$ that projects onto the radial lines, $\mathbf{l}, \mathbf{l}', \mathbf{l}'', \mathbf{l}'''$. Then the radial projection equations (Eq. 1), can be collected in the following matrix equation:

$$\underbrace{\begin{bmatrix} \mathbf{P} & \mathbf{l} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}' & \mathbf{0} & \mathbf{l}' & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'' & \mathbf{0} & \mathbf{0} & \mathbf{l}'' & \mathbf{0} \\ \mathbf{P}''' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{l}''' \end{bmatrix}}_{\mathbf{M}_{8\times 8}} \begin{bmatrix} \mathbf{X} \\ -\lambda \\ -\lambda' \\ -\lambda'' \\ -\lambda''' \end{bmatrix} = \mathbf{0} \tag{4}$$

Since we know that a solution exists, the right null-space of $\mathbf{M}$ should have non-zero dimension, which implies that the determinant of the matrix has to be zero.

Following the approach of [47], expansion of the determinant yields the quadrilinear constraint for 1D radial cameras:

$$\mathbf{Q}^{ijkl}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k\mathbf{l}'''_l = 0 \tag{5}$$

$\mathbf{Q}^{ijkl}$ is the $2 \times 2 \times 2 \times 2$ homogenous quadrifocal tensor (introduced in [43]) of four 1D cameras. We use the Einstein summation convention in which indices repeated in covariant and contravariant positions denote implicit summations.

A general $2 \times 2 \times 2 \times 2$ tensor (up to scale) has 15 degrees of freedom. However, to describe four uncalibrated 1D radial cameras, up to a projectivity, we need only $(4 \times (2 \times 4 - 1) - (4 \times 4 - 1) = 13)$ degrees of freedom. Thus the radial quadrifocal tensor has *exactly 2* internal constraints. Compare this to $(80 - 29)$ internal constraints

for the quadrifocal tensor of 4 perspective views [1]. The radial quadrifocal tensor can thus be linearly estimated given 15 corresponding quadruplets. Given more than 15 corresponding quadruplets, a linear least squares solution can be obtained.

Lemma: There exist no higher-order tensors for 1D cameras.

Proof: Each camera projection matrix is 2×4. Thus if there are n (> 4) cameras, $\mathbf{M}$ (similar to Eq. (4)) will be of dimensions $2n \times (4+n)$.Following a similar argument, since the right null-space is of non-zero dimension, the column rank of $\mathbf{M}$ is $< (4+n)$. Hence any $(4+n)\times(4+n)$ sub-matrix (say, $\mathbf{M_s}$) has determinant zero. Note that all the columns of $\mathbf{M}$ are selected in $\mathbf{M_s}$ but only $(4+n)$ rows are selected. If we omit picking a row from the projection matrix of the i^{th} camera, then $(i+4)^{th}$ column (whose only non-zero elements are $\mathbf{l}_{i1}$ and $\mathbf{l}_{i2}$) , of $\mathbf{M_s}$ is all zero. Thus the determinant is trivially zero and doesn't give us a *constraint*.

To get a non-trivial constraint we need to select *at least* one row from each camera, *and* choose 4 specific cameras, to select the 2^{nd} row from. Let these cameras be (p, q, r, s). Expansion of this $(4 + n) \times (4 + n)$ matrix, gives us a constraint which is of the form, $\mathbf{l}_{1w_1} \ldots \mathbf{l}_{1w_n}\mathbf{D}^2$ where $\mathbf{D} = (\mathbf{l}_{p1}\mathbf{l}_{q1}\mathbf{l}_{r1}\mathbf{l}_{s1}\det(\begin{bmatrix}\mathbf{P}_{p2}^T & \mathbf{P}_{q2}^T & \mathbf{P}_{r2}^T & \mathbf{P}_{s2}^T\end{bmatrix}^T) + \ldots)$ Note that $\mathbf{D}$ is nothing but the quadrifocal constraint (Eq. (5)) of the 4 selected views (p, q, r, s).

3.1 Nature of the Internal Constraints

We will now give a geometric interpretation of the *two* internal constraints of the radial quadrifocal tensor. It depends on the following observation: Given four lines (the optical axes of the 4 cameras), in general configuration, in 3D space, there exists two lines that intersect all of them [17]. We will call these special lines, that intersect **all** four optical axes, **quadrifocal lines**.

Let us denote one quadrifocal line as $\mathbf{L}$. Further, let it project onto $\mathbf{l}, \mathbf{l}', \mathbf{l}'', \mathbf{l}'''$ in the four images. The planes back-projected from the radial lines $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ will contain $\mathbf{L}$. So irrespective of which plane is chosen, among the pencil back-projected by the 4th camera (Figure 1(b)), we will have a point of intersection for the four planes (since a line and a plane always intersect in 3D space). This means that the quadrifocal constraint (Eq. (5)) is satisfied *for all* $\mathbf{l}'''$. This means that,

$$(\mathbf{Q}^{ijk1}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k)l'''_1 + (\mathbf{Q}^{ijk1}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k)l'''_2 = 0 \tag{6}$$

Since the above is satisfied for all values of l'''_1 and l'''_2, it implies that the respective coefficients should be zero.

$$\begin{aligned}\mathbf{Q}^{ijk1}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0\\ \mathbf{Q}^{ijk2}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0\end{aligned} \tag{7}$$

Since the above argument will be true even if we choose any of the other cameras instead of the 4th camera (for example, we can back-project $\mathbf{l}', \mathbf{l}'', \mathbf{l}'''$, which would allow us to choose any arbitrary line, in the 1st image, as $\mathbf{l}$), we have the following 8 equations:

$$\mathbf{Q}^{1jkl}\mathbf{l}'_j\mathbf{l}''_k\mathbf{l}'''_l = 0 \tag{8a}$$

[1] In fact, for perspective cameras, the radial quadrifocal tensor corresponds to the upper $2 \times 2 \times 2 \times 2$ part of the full quadrifocal tensor.

[2] $w_1 \ldots w_n$ are the rows selected from all the cameras other than (p, q, r, s)

$$\mathbf{Q}^{2jkl}\mathbf{l}'_j\mathbf{l}''_k\mathbf{l}'''_l = 0 \tag{8b}$$

$$\mathbf{Q}^{i1kl}\mathbf{l}_i\mathbf{l}''_k\mathbf{l}'''_l = 0 \tag{8c}$$

$$\mathbf{Q}^{i2kl}\mathbf{l}_i\mathbf{l}''_k\mathbf{l}'''_l = 0 \tag{8d}$$

$$\mathbf{Q}^{ij1l}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}'''_l = 0 \tag{8e}$$

$$\mathbf{Q}^{ij2l}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}'''_l = 0 \tag{8f}$$

$$\mathbf{Q}^{ijk1}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0 \tag{8g}$$

$$\mathbf{Q}^{ijk2}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0 \tag{8h}$$

Notice that filling in Eq. (8a,8c,8e,8g) in the quadrifocal constraint (Eq. (5)), yields the other four equations (Eq. (8b,8d,8f,8h)). Hence the above 8 equations yield only 5 independent constraints, Eq. (8a,8c,8e,8g) as well as the quadrifocal constraint (Eq. (5)) itself.

Let us denote the tuple of 4 lines as $\mathbf{r} = \left[\mathbf{l}_{2\times1}, \mathbf{l}'_{2\times1}, \mathbf{l}''_{2\times1}, \mathbf{l}'''_{2\times1}\right]$[3] and equations Eq. (8a,8c,8e,8g) as $f_1(\mathbf{Q},\mathbf{r}) = 0, f_2(\mathbf{Q},\mathbf{r}) = 0, f_3(\mathbf{Q},\mathbf{r}) = 0, f_4(\mathbf{Q},\mathbf{r}) = 0$. Finally, let us denote by $f_5(\mathbf{Q},\mathbf{r}) = 0$, the quadrifocal constraint (Eq. (5)).

Thus we have the following condition: If $\mathbf{Q}$ is a *valid* radial quadrifocal tensor, then the set of equations:

$$f_{1\ldots5}(\mathbf{Q},\mathbf{r}) = 0 \tag{9}$$

has at least *two* solutions (in general configuration).

Assume that $\mathbf{Q}$ is a *valid* radial quadrifocal tensor. If $\mathbf{r}'$ is a solution of Eq. (9) we can solve for $\mathbf{r}'$ from $f_{1\ldots4}(\mathbf{Q},\mathbf{r}) = 0$ in terms of f_i and $\mathbf{Q}$. Since there are two solutions, we have $\mathbf{r}_1$, $\mathbf{r}_2$ such that $\mathbf{r}_i = g_i(\mathbf{Q})$, $i = 1, 2$. Note that we *are not* solving f_i to compute g_i. We are just claiming the *existence* of such g_is. If we substitute each solution $\mathbf{r}_i$ in $f_5(\mathbf{Q},\mathbf{r}) = 0$, we will get two equations in $\mathbf{Q}$. Thus, if $\mathbf{Q}$ is a valid radial quadrifocal tensor, then,

$$\begin{aligned} f_5(\mathbf{Q}, g_1(\mathbf{Q})) &= 0 \\ f_5(\mathbf{Q}, g_2(\mathbf{Q})) &= 0 \end{aligned} \tag{10}$$

This proves that we have the two internal constraints sitting within the set of 5 equations. In other words, the two internal constraints that a valid radial quadrifocal tensor has to satisfy corresponds to the need for two quadrifocal lines to exist. This is very comparable to the rank-2 constraint of the fundamental matrix which implies the existence of both epipoles.

[3] Even though these are vectors, they are projective entities and hence $\mathbf{l}_{2\times1} \sim \lambda$ [u v] has only 1 variable. Thus $\mathbf{r}$ has 4 variables and *not* 8.

3.2 3D Reconstruction

3.2.1 Projective Reconstruction

We now consider the problem of 3D reconstruction of points whose correspondences have been specified across the input images. Given a radial quadrifocal tensor, we can easily compute the four uncalibrated camera matrices [19]. For every valid radial quadrifocal tensor, two non-equivalent projective reconstructions are obtained. As we can not disambiguate between them at this stage we will carry them through to the metric reconstruction stage and potentially the radial calibration where in general only a single solution will yield consistent results. Once the projection matrices have been recovered, points in 3D can be reconstructed by back-projecting planes. This corresponds to computing the right nullspace of the following matrix:

$$\mathbf{R}_{8\times 8} = \begin{bmatrix} \mathbf{P} & \mathbf{l} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}' & \mathbf{0} & \mathbf{l}' & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'' & \mathbf{0} & \mathbf{0} & \mathbf{l}'' & \mathbf{0} \\ \mathbf{P}''' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{l}''' \end{bmatrix} \tag{11}$$

Since only three planes are required to define a point uniquely in 3D space, we can in fact reconstruct all points seen in at least three views.

3.2.2 Metric Reconstruction

The dual absolute quadric, $\mathbf{\Omega}^*_\infty$ encodes both the absolute conic and the plane at infinity. To upgrade our reconstruction to metric, we need to estimate this degenerate quadric in the projective frame in which the cameras and the points have been determined [48,30]. $\mathbf{\Omega}^*_\infty$ projects into the radial 1D image as,

$$\tilde{\mathbf{K}}\tilde{\mathbf{K}}^T = \tilde{\omega}^* = \mathbf{P}\mathbf{\Omega}^*_\infty\mathbf{P}^T \tag{12}$$

with $\tilde{\mathbf{K}} = \begin{bmatrix} f_x & s \\ 0 & f_y \end{bmatrix}$ the upper 2×2 part of the calibration matrix. Using the assumptions of (i) known principal point (and it being at the origin) (ii) *zero skew* ($s = 0$) (iii) *known aspect ratio* ($f_y = af_x$), we obtain 8 linear constraints on $\mathbf{\Omega}^*_\infty$, from the 4 views. Since $\mathbf{\Omega}^*_\infty$ is a 4×4 homogenous symmetric matrix it has 9 d.o.f (10 up to scale). Using the additional rank-3 constraint we obtain a fourth-degree equation ($det\mathbf{\Omega}^*_\infty = 0$) and thus obtain up to 4 solutions. Only positive semi-definite solutions for the absolute quadric have to be considered. If more than one solution persists, we can generate multiple alternative metric reconstructions and disambiguate them later by verifying the radial symmetry in the next section. If $\mathbf{\Omega}^*_\infty$ is decomposed as $\mathbf{\Omega}^*_\infty = \mathbf{HIH}^T$, then $\mathbf{H}^{-1}$ is the point homography that takes the projective frame to the metric frame [20].

3.3 Radial calibration

Once a metric reconstruction has been obtained using the 1D radial property of the camera, it can be used to calibrate the remaining unknowns of the projection. In this section we will present a non-parametric approach to calibrate central and non-central radially symmetric cameras. This process can be done independently for each image

and it is thus possible to calibrate four different cameras -or a camera with different settings- using a single quadrifocal tensor.

For each cameras, all reconstructed feature points can be represented in a cylindrical coordinate system relative to the optical axis of the camera, i.e. (ρ, ϕ, z). The origin along the z-axis can in a first phase be chosen arbitrarily. Because we assume radial symmetry, the ϕ coordinate is irrelevant for us. The goal of our calibration procedure is to obtain an expression for rays, $\mathbf{r}$, in the ρz-plane as a function of the radius r, i.e. $\mathbf{r}(r) : a_1(r)\rho + a_2(r)z + a_3(r) = 0$. This can be done by fitting lines to all the points that have (almost) the same r value.

4 Radial Trifocal Tensor

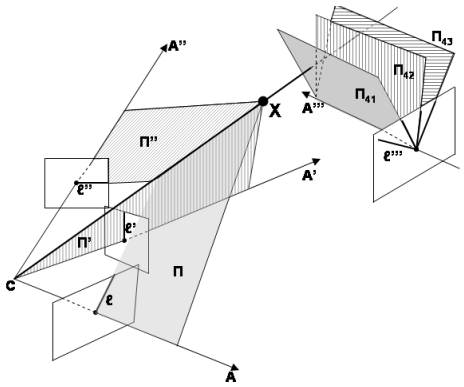


Fig. 3: *The Trifocal Constraint*

Suppose that three optical axes, $\mathbf{A}, \mathbf{A}'$ and $\mathbf{A}''$ intersect at some point $\mathbf{C}$. Also, suppose that a 3D point $\mathbf{X}$ projects onto the lines $\mathbf{l}, \mathbf{l}'$and $\mathbf{l}''$ in the three views (Figure 3) . Consider the plane, $\mathbf{\Pi}$ containing $\mathbf{C}$ and the line $\mathbf{l}$ (corresponding to the back-projection of the radial line $\mathbf{l}$). Similarly, one has the planes $\mathbf{\Pi}^{'}$ and $\mathbf{\Pi}^{''}$. Note that for every 3D point, $\mathbf{X}$, the corresponding planes back-projected from the 3 views intersect in the line passing through $\mathbf{C}$ and $\mathbf{X}$. Three planes in 3D space intersecting in a line is a non-trivial constraint. This non-trivial constraint between the three 1D radial views is encoded by the radial trifocal tensor (introduced in [44]). We can now formulate this constraint mathematically.

Without loss of generality we can assume that the three optical axes intersect in the origin $(0, 0, 0, 1)^T$. Since $\mathbf{PC} = (0, 0)^T$, the 1D radial cameras whose optical axes contain the origin must have the following form $\mathbf{P} = [\tilde{\mathbf{P}}|\mathbf{0}]$. Let $\tilde{\mathbf{X}}^T$ correspond to the first three coefficients of $\mathbf{X}$. In this case, the first 6 rows of Eq. (4) can be rewritten as:

$$\begin{bmatrix} \tilde{\mathbf{P}} & \mathbf{l} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{P}}' & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \tilde{\mathbf{P}}'' & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}} \\ -\lambda \\ -\lambda' \\ -\lambda'' \end{bmatrix} = \mathbf{0} \tag{13}$$

The non-zero dimension of the right null-space implies that the 6×6 measurement matrix must have a zero determinant.

Expansion of the determinant produces the unique trilinear constraint for 1D views yields

$$\mathbf{T}^{ijk}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0 \tag{14}$$

$\mathbf{T}^{ijk}$ is the $2 \times 2 \times 2$ homogeneous radial trifocal tensor of the three 1D radial cameras. The expression for the coefficients of the trifocal tensor is

$$\mathbf{T}^{ijk} = \det\begin{bmatrix}\tilde{\mathbf{P}}_i^T & \tilde{\mathbf{P}}_j'^T & \tilde{\mathbf{P}}_k''^T\end{bmatrix} \tag{15}$$

The radial trifocal tensor is a minimal parameterization of the three $\mathbb{P}^2 \rightarrow \mathbb{P}^1$ mapping cameras as the d.o.f can be shown to match, $2\times2\times2-1 = 7 = 3\times(2\times3-1)-(3\times3-1)$ (with the LHS being the d.o.f of $\mathbf{T}$ and the RHS being the d.o.f of the three uncalibrated views up to a projectivity) and has no internal constraints.

The radial trifocal tensor can be linearly estimated given seven corresponding triplets (where every triplet gives a linear constraint on the parameters of the radial trifocal tensor using Eq. (14)) Given more than seven correspondences, we can obtain the linear least squares solution.

It is interesting to verify the relation between the radial trifocal constraint and the radial quadrifocal constraint. When three optical axes intersect, adding a fourth view doesn't yield any additional constraint and the quadrifocal constraint becomes degenerate. Since in $\mathbf{P}^3$ a line and a plane always intersect, we no longer need the precise plane $\mathbf{\Pi}'''$, back-projected from $\mathbf{l}'''$. Instead we could choose any of the planes among the pencil back-projected by the fourth camera. Let us examine the radial quadrifocal constraint, Eq. (5), in this scenario:

$$(\mathbf{Q}^{ijk1}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k)l'''_1 + (\mathbf{Q}^{ijk2}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k)l'''_2 = 0 \tag{16}$$

Choosing an arbitrary back-projected plane from the fourth camera corresponds to arbitrary values for l'''_1 and l'''_2. Since Eq. (16) is valid for arbitrary values of l'''_1 and l'''_2, it implies that the coefficients are zero. Further, the above condition is valid for any 3D point $\mathbf{X}$. Comparing this to the trifocal constraint, we see that in this case the quadrifocal tensor must be related to the trifocal tensor as follows

$$\mathbf{Q}^{ijkl} = (\lambda_1\mathbf{T}^{ijk}, \lambda_2\mathbf{T}^{ijk}) \tag{17}$$

and can only be determined up to one degree of freedom, i.e. $\frac{\lambda_1}{\lambda_2}$.

The trifocal tensor for 1D cameras and its properties were first studied by Quan and Kanade [31] in the context of structure and motion using line correspondences under affine cameras. They showed that by neglecting the position of the lines and considering only their direction, this problem was equivalent to the structure and motion problem for points in one lower dimension. Faugeras et. al. [9] studied the 1D trifocal tensor in the context of planar motion recovery and self-calibration.

4.1 Reconstruction

4.1.1 Projective Reconstruction of $\mathbf{\Pi}_\infty$

Let us now consider the problem of reconstructing directions from $\mathbf{C}$. Directions correspond to points on $\mathbf{\Pi}_\infty$. Given the radial trifocal tensor, $\mathbf{T}$, we can estimate the three uncalibrated camera matrices, $\tilde{\mathbf{P}}, \tilde{\mathbf{P}}'$ and $\tilde{\mathbf{P}}''$ [44]. These projection matrices can be thought of projecting points on $\mathbf{\Pi}_\infty$ to radial lines in the corresponding views.

Suppose a point on $\mathbf{\Pi}_\infty$, $\mathbf{X}$ (or direction $\mathbf{X}$ in 3D space) projects onto the radial lines $\mathbf{l}$, $\mathbf{l}'$ and $\mathbf{l}''$. The point then satisfies Eq. (13). Thus it can be computed as the right null-space of

$$\mathbf{R}_{6\times 6} = \begin{bmatrix} \tilde{\mathbf{P}} & \mathbf{l} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{P}}' & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \tilde{\mathbf{P}}'' & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} \tag{18}$$

Note that for reconstruction is possible even if the point is visible in only two views. In this case, $\mathbf{R}$ reduces to the following 4×5 matrix, which always has a solution:

$$\begin{bmatrix} \tilde{\mathbf{P}} & \mathbf{l} & \mathbf{0} \\ \tilde{\mathbf{P}}' & \mathbf{0} & \mathbf{l}' \end{bmatrix} \tag{19}$$

4.1.2 Metric Reconstruction

Let $\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2$ be the two rows of the projection matrix, $\tilde{\mathbf{P}}_{2\times 3}$. Similarly let $\tilde{\mathbf{p}}'_1, \tilde{\mathbf{p}}'_2$ be the rows of $\tilde{\mathbf{P}}'$ and $\tilde{\mathbf{p}}''_1, \tilde{\mathbf{p}}''_2$ be the rows of $\tilde{\mathbf{P}}''$. Let ω^*_∞ be dual of the absolute conic in the projective frame in which we have reconstructed the points on $\mathbf{\Pi}_\infty$ (directions). It is a 3×3 homogenous symmetric matrix and hence has 5 degrees of freedom (6 up to scale). To upgrade the projective reconstruction to metric it is sufficient to estimate ω^*_∞ [20].

We have the assumptions of (i) known principal point (and it being at the origin) (ii) *zero skew* and (iii) *constant (but possibly unknown) aspect ratio*. Note that this is equivalent to assuming rectangular pixels. It can be shown the assumption of zero skew in the three views gives us the following set of equations linear in the parameters of ω^*_∞ :

$$\begin{aligned} \tilde{\mathbf{p}}_1 \omega^*_\infty \tilde{\mathbf{p}}_2^T &= 0 \\ \tilde{\mathbf{p}}'_1 \omega^*_\infty \tilde{\mathbf{p}}_2'^T &= 0 \\ \tilde{\mathbf{p}}''_1 \omega^*_\infty \tilde{\mathbf{p}}_2''^T &= 0 \end{aligned} \tag{20}$$

Further, the assumption of constant aspect ratio gives us the following equations:

$$\frac{\tilde{\mathbf{p}}_1 \omega^*_\infty \tilde{\mathbf{p}}_1^T}{\tilde{\mathbf{p}}_2 \omega^*_\infty \tilde{\mathbf{p}}_2^T} = \frac{\tilde{\mathbf{p}}'_1 \omega^*_\infty \tilde{\mathbf{p}}_1'^T}{\tilde{\mathbf{p}}'_2 \omega^*_\infty \tilde{\mathbf{p}}_2'^T} = \frac{\tilde{\mathbf{p}}''_1 \omega^*_\infty \tilde{\mathbf{p}}_1''^T}{\tilde{\mathbf{p}}''_2 \omega^*_\infty \tilde{\mathbf{p}}_2''^T} = A \tag{21}$$

If the aspect ratio is known (A is known in Eq. (21)), we have 3 more equations, linear in the parameters of ω^*_∞. If the aspect ratio is unknown, then we have two equations, quadratic in the parameters of ω^*_∞. Using the linear equations in Eq. (20), we can reduce the equations in Eq. (21) to two quadratic equations in two variables. And these can be solved analytically.

4.2 Non-Parametric Radial Calibration

Once a metric reconstruction of $\mathbf{\Pi}_\infty$ has been obtained using the 1D radial property of the camera, it can be used to calibrate the remaining unknowns of the projection. In this section we present a non-parametric approach to calibrate purely rotating central radially symmetric cameras. The objective thus is to compute a function that maps the radius in the distorted image (r) to an angle with the optical axis (θ): $f : r \rightarrow \theta$

Note that the procedure described here can be done independently for each image and it is thus possible to calibrate three different cameras using a single trifocal tensor.

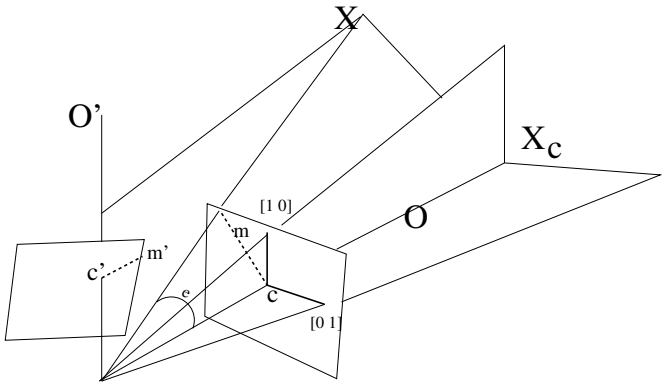


Fig. 4: *Shows two radial cameras with optical axes* $\mathbf{O}$ *and* $\mathbf{O}'$ *and centers of distortion* $\mathbf{c}$, $\mathbf{c}'$ *respectively.* $\mathbf{X}$ *is reconstructed as the point on* $\mathbf{\Pi}_\infty$ *that projects onto radial lines* $\mathbf{m}$ *and* $\mathbf{m}'$ *(dotted lines) in the images.* $\mathbf{X}_c$ *is reconstructed as the point that projects onto [1 0] and [0 1] (indicated by bold lines) in the first radial image. If we have a metric reconstruction of* $\mathbf{\Pi}_\infty$*, we can compute the angle e*

Since we have upgraded the reconstruction to metric, we can compute the angle between rays corresponding to $\mathbf{X}$ and $\mathbf{X}_c$ (defined in figure above) as [20]:

$$\cos(\theta) = \frac{\mathbf{X}^T \mathbf{X}_c}{\sqrt{\mathbf{X}^T \mathbf{X}} \sqrt{\mathbf{X}_c^T \mathbf{X}_c}} \tag{22}$$

For each camera, the rays corresponding to reconstructed feature points can be represented in a coordinate system relative to the optical axis of the camera, i.e. (θ, ϕ). Because we assume radial symmetry, the ϕ coordinate is irrelevant for us. Thus for each camera and each reconstructed point visible in it, we will get a point on the function f. We can use the map we get, i.e. (r, θ) to then do undistortion (by fitting a function, use as samples in a lookup table etc:.).

5 Radial Trifocal Tensor for Cameras in General Position

In some situations it might not be possible to ensure a rotating camera Hence in this section we give an approach that allows reconstruction even when the camera-center doesn't remain stationary across the three views. However, we will need to make two allowances here:

- **All the scene-points should come from a planar scene**. This is not really difficult to achieve because we can image a building facade, checkerboard images etc:. Further, the trifocal tensor equation will not be satisfied for scene-points in non-planar position and hence can form the basis of a robust sieve.
- We will have to use a parametric model of distortion in the last stage of estimation. Note that this isn't too stringent a constraint because the the reconstruction still is independent of the model of distortion.

It should be noted, here, that for the rotating camera case (discussed in the previous section), all scene points being on a real-world plane is *not a degenerate case* for reconstruction.

Consider a point $\mathbf{X}$, on the real-world plane $\mathbf{\Pi}$ that projects onto the radial lines $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ in the three cameras. Then each projection matrix is a $\mathbf{P}^2 \rightarrow \mathbf{P}^1$ mapping and we get equations of the form Equation (13). Since the right null space exists, the determinant of 6×6 matrix is 0 which implies Equation (14). Evaluation of the projection matrices from the trifocal tensor and reconstruction follow similar lines.

In this case, however instead of $\mathbf{\Pi}_\infty$ we have $\mathbf{\Pi}$ which is real-world plane from which the scene-points come. Let us denote the undistorted first image as I_u^1 and the distorted (which is what we have) image as I_d^1. We consider the formation of the distorted image to be a two-step procedure. A homography, denoted by $\mathbf{H}$, that does $\mathbf{\Pi} \rightarrow I_u^1$. Followed by distortion for which we will assume a parametric model.

5.1 Estimating the homography from $\mathbf{\Pi}$ to the undistorted images

Consider the projection matrix of the first radial camera, $\mathbf{P}_{3\times 2}^T = \begin{bmatrix} \mathbf{p}_1^\top & \mathbf{p}_2^\top \end{bmatrix}$, where $\mathbf{p}_1$ and $\mathbf{p}_2$ are the rows of 2×3 matrix, $\mathbf{P}$.

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{bmatrix} \mathbf{X} \tag{23}$$

Suppose $\mathbf{X}$ projects onto $\mathbf{x}_u$ in the first image (I_u^1, conforming to the pin-hole model). Also, suppose that $\mathbf{X}$ projects onto the line $\mathbf{l} = [l_1\, l_2]^T$, in the first distorted image(I_d^1). Then, $\mathbf{x}_u$ is of the form $\lambda \begin{bmatrix} -l_2 \\ l_1 \end{bmatrix}$ (since the center of distortion is $(0,0)^T$, and deviation is only along the radial line).

The homography $\mathbf{H}$ from $\mathbf{\Pi}$ to I_u^1, would map $\mathbf{X}$ to $\mathbf{x}_u$. From the observation made above, we can estimate the first two rows of $\mathbf{H}$ as

$$\mathbf{H} = \begin{bmatrix} -\mathbf{p}_2 \\ \mathbf{p}_1 \\ \mathbf{h}_3 \end{bmatrix} \tag{24}$$

where $\mathbf{h}_3 = (h_{31}, h_{32}, h_{33})^T$ is unknown.

Let

$$S_u = \{ \begin{bmatrix} x_u^i \\ y_u^i \\ 1 \end{bmatrix} \mid i = 1 \ldots n \}$$

be the set of coordinates of the feature points in the undistorted image, I_u^1.

Then by estimating the homography, $\mathbf{H}$, up to three unknown parameters, as we have done above, we are able to express the set, S_u, as

$$S_u(h_{31}, h_{32}, h_{33}) = \{ \begin{bmatrix} -\mathbf{p}_2 \cdot \mathbf{X}^i \\ \mathbf{p}_1 \cdot \mathbf{X}^i \\ [h_{31}\, h_{32}\, h_{33}] \cdot \mathbf{X}^i \end{bmatrix} \mid i = 1 \dots n \} \tag{25}$$

The undistorted coordinates ($\mathbf{x}_u$) of all the feature points, together, are thus now known up to only three parameters (of $\mathbf{h}_3$) in total.

5.2 Computing the distortion parameters

We will now estimate the distortion parameters of the division model [4]. We will assume that the transformation from I_d^1 to I_u^1, follows the division model [10] i.e, induced by the distortion parameters,is

$$\mathbf{x}_u = \frac{\mathbf{x}_d}{(1 + K_1 {r_d}^2 + K_2 {r_d}^4 + K_3 {r_d}^6 + \dots)} \tag{26}$$

The transformation from $\mathbf{\Pi}$ to I_u^1, induced by $\mathbf{H}$, is

$$\rho \begin{bmatrix} \mathbf{x}_u \\ 1 \end{bmatrix} = \begin{bmatrix} -\mathbf{p}_2\mathbf{X} \\ \mathbf{p}_1\mathbf{X} \\ \mathbf{h}_3\mathbf{X} \end{bmatrix} \tag{27}$$

with ρ an unknown scale factor. Since the two points are the same, the vectors representing them should be parallel. Thus their cross-product should be equal to zero [20].

$$\begin{bmatrix} -\mathbf{p}_2\mathbf{X} \\ \mathbf{p}_1\mathbf{X} \\ \mathbf{h}_3\mathbf{X} \end{bmatrix} \times \begin{bmatrix} x_d \\ y_d \\ (1 + K_1 {r_d}^2 + \dots) \end{bmatrix} = \mathbf{0} \tag{28}$$

Thus every point gives us two equations,

$$\begin{bmatrix} x_d(\mathbf{h}_3\mathbf{X}) + \mathbf{p}_2\mathbf{X}(K_1 r_d^2 + \dots) \\ y_d(\mathbf{h}_3\mathbf{X}) - \mathbf{p}_1\mathbf{X}(K_1 r_d^2 + \dots) \end{bmatrix} = \begin{bmatrix} (-\mathbf{p}_2\mathbf{X}) \\ (\mathbf{p}_1\mathbf{X}) \end{bmatrix} \tag{29}$$

which can be rewritten as,

$$\begin{bmatrix} x_d\mathbf{X} & (\mathbf{p}_2\mathbf{X})[r_d^2 & r_d^4 \dots] \\ y_d\mathbf{X} & (-\mathbf{p}_1\mathbf{X})[r_d^2 & r_d^4 \dots] \end{bmatrix} \begin{bmatrix} \mathbf{h}_3^\top \\ K_1 \\ K_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} (-\mathbf{p}_2\mathbf{X}) \\ (\mathbf{p}_1\mathbf{X}) \end{bmatrix} \tag{30}$$

These two equations are dependent, but it is best to use them both to avoid degenerate cases and deal with orientation ambiguities.

[4] Note that everything up to this stage was independent of any assumption on the form of the radial distortion. Therefore, we could also use a different distortion model. Depending on the type/parameters of distortion, we may or may not be able to estimate the last row of the homography and the distortion parameters linearly. However, the relations that we will derive are valid irrespective of the model used.

Given more than $3+n$ feature points (where n is the number of distortion parameters), we can solve the system of equations we would get, in a least-squares sense.

Using the above set of equations directly, we minimize an algebraic error. A better solution would be to minimize the geometric error in the distorted image, I_d^1 (since that is the input image). For that we need to divide each of the equations given in Eq (30), by $\frac{1}{\mathbf{h}_3\mathbf{X}}$. This would then minimize the sum (over all the feature points) of the following squared-error.

$$|| \begin{bmatrix} x_d - \frac{-\mathbf{p}_2\mathbf{X}}{\mathbf{h}_3\mathbf{X}}(1 + K_1 r_d^2 + \dots) \\ y_d - \frac{\mathbf{p}_1\mathbf{X}}{\mathbf{h}_3\mathbf{X}}(1 + K_1 r_d^2 + \dots) \end{bmatrix} ||^2 \tag{31}$$

which is distance, in I_d^1, from $(x_d, y_d)^T$ to $\begin{bmatrix} \frac{-\mathbf{p}_2\mathbf{X}}{\mathbf{h}_3\mathbf{X}} & \frac{\mathbf{p}_1\mathbf{X}}{\mathbf{h}_3\mathbf{X}} \end{bmatrix}^T (1 + K_1 r_d^2 + \dots)$ i.e, the pixel corresponding to the feature point in I_u^1, warped by the distortion parameters $((1 + K_1 r_d^2 + \dots))$. However, we don't have $\frac{1}{\mathbf{h}_3\mathbf{X}}$, since $\mathbf{h}_3$ is unknown, but by scaling with $\frac{||(x_d, y_d)^T||}{||(-\mathbf{p}_2\mathbf{X}, \mathbf{p}_1\mathbf{X})^T||}$ we can at least normalize for the arbitrary scale of $\mathbf{X}$. We scale both of the equations, generated by each feature point, before stacking them in the matrix to obtain the least-squares solution.

This system of equations could be refined iteratively using the previous approximation of $\mathbf{h}_3$ to normalize the equations or alternatively a non-linear minimization of Eq. (31) could be used to refine our linear solution. The results described in the experimental section are obtained using the linear method only.

6 Radial Tensor for Heterogeneous Cameras

The next natural question to ask is: how does the radial 1D camera model interact with the more standard pin-hole camera model. This is particularly useful when we consider today's camera networks which contain cameras of different types, some of which will need the pin-hole model and others the radial 1D model. In particular we will examine the multi-view relationship among *two 1D cameras and one pin-hole camera*.

Consider the point $\mathbf{X}$ that projects onto the point $\mathbf{x}_{3\times1}$ in the pin-hole camera and the radial lines $\mathbf{l}'_{2\times1}$ and $\mathbf{l}''$ in the radial cameras. Then it projects by the following set of equations,

$$\begin{aligned} \lambda\mathbf{x} &= \mathbf{P}_{3\times4}\mathbf{X} \\ \lambda'\mathbf{l}' &= \mathbf{P}'_{2\times4}\mathbf{X} \\ \lambda''\mathbf{l}'' &= \mathbf{P}''_{2\times4}\mathbf{X} \end{aligned} \tag{32}$$

These equations can be rewritten in matrix format as,

$$\begin{bmatrix} \mathbf{P}_{3\times4} & \mathbf{x} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'_{2\times4} & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \mathbf{P}''_{2\times4} & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ -\lambda \\ -\lambda' \\ -\lambda'' \end{bmatrix} = \mathbf{0} \tag{33}$$

Since we know that a solution exists, the right null-space of the 7×7 measurement matrix should have non-zero dimension, which implies that

$$\det \begin{bmatrix} \mathbf{P}_{3\times4} & \mathbf{x} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'_{2\times4} & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \mathbf{P}''_{2\times4} & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} = 0 \tag{34}$$

Following the approach of [47], expansion of the determinant produces the unique trilinear constraint for the pin-hole view and the two 1D views,

$$\mathbf{T}_i^{jk}\mathbf{x}^i\mathbf{l}'_j\mathbf{l}''_k = 0 \tag{35}$$

$\mathbf{T}_i^{jk}$ is the $3 \times 2 \times 2$ mixed trifocal tensor (introduced in [45]) for the pin-hole camera and two 1D radial cameras. Elements of $\mathbf{T}$ can be written as 4×4 minors of the joint projection matrix $\begin{bmatrix}\mathbf{P}^T \mathbf{P}'^T \mathbf{P}''^T\end{bmatrix}^T$ where two rows (of the minor) come from the pin-hole projection matrix, $\mathbf{P}$, and a row each is contributed by $\mathbf{P}'$ and $\mathbf{P}''$.

One useful way of understanding the mixed trifocal tensor is by considering the intersection of the ray back-projected from the pin-hole camera and the planes back-projected from the two radial cameras, in 3D space. The back-projected ray intersects the plane back-projected from the first radial camera at a point, say $\mathbf{X}$. This constrains the radial line in the second radial camera to be such that the plane back-projected from it contains $\mathbf{X}$. The mixed trifocal tensor captures this incidence relationship.

A general $3\times2\times2$ tensor has $12-1 = 11$ degrees of freedom up to scale. Subtracting from these the degrees of freedom required to describe one pin-hole camera and two radial cameras up to a 3D projectivity, $(3 \times 4 - 1) + 2 \times (2 \times 4 - 1) - (4 \times 4 - 1) = 10$, we observe that the mixed trifocal tensor has *only one* internal constraint.

6.1 Nature of the Internal Constraint

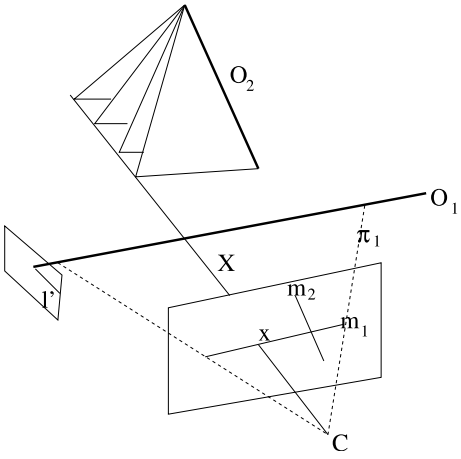


Fig. 5: $\mathbf{O}_1$ *and* $\mathbf{O}_2$ *are the two optical axes.* $\mathbf{C}$ *is the center of the pin-hole camera.* $\mathbf{\Pi}_1$ *is the plane defined by* $\mathbf{C}$ *and* $\mathbf{O}_1$*. Planes* $\mathbf{\Pi}_i$ *project onto* $\mathbf{m}_i$ *in the pin-hole image.* $\mathbf{X}$ *(lying on* $\mathbf{\Pi}_1$*) projects onto* $\mathbf{x}$ *in the pin-hole image and* $\mathbf{l}'$ *in the first radial image. The intersection of* $\mathbf{\Pi}_1$ *(the plane back-projected from* $\mathbf{l}'$*) and the ray back-projected from* $\mathbf{x}$ *is the ray itself. Any plane back-projected from the second radial camera will intersect this ray.*

We will now characterise the internal constraint of the mixed trifocal tensor. Consider Figure 2. Let $\mathbf{C}$ be the camera center of the pin-hole camera and $\mathbf{O}_1$ and $\mathbf{O}_2$ be the optical axes of the two radial cameras. Let us denote the plane defined by $\mathbf{C}$ and $\mathbf{O}_i$ by $\mathbf{\Pi}_i$.

Consider a point $\mathbf{X}$, lying in $\mathbf{\Pi}_1$, that projects onto $\mathbf{x}$ in the pin-hole image and $\mathbf{l}'$ in the first radial image. Note that the plane back-projected from the first radial image is $\mathbf{\Pi}_1$. Further, the ray back-projected from the pin-hole image lies in $\mathbf{\Pi}_1$. Thus the intersection of these two is the ray itself. Since in 3D space a ray and a plane always intersect, we can choose any arbitrary plane from the pencil back-projected by the second radial camera. This is equivalent to selecting an arbitrary $\mathbf{l}'' = \begin{bmatrix} l''_1, l''_2 \end{bmatrix}$. This implies that their coefficients in Eq. (35) are zero.

$$\mathbf{T}_i^{j1}\mathbf{x}^i\mathbf{l}'_j = 0 \quad \mathbf{T}_i^{j2}\mathbf{x}^i\mathbf{l}'_j = 0 \tag{36}$$

Eq. (36) can be interpreted to imply that *there exists* $\mathbf{l}' = \begin{bmatrix} l'_1, l'_2 \end{bmatrix}$ such that

$$\underbrace{\begin{bmatrix} \mathbf{T}_i^{11}\mathbf{x}^i & \mathbf{T}_i^{21}\mathbf{x}^i \\ \mathbf{T}_i^{12}\mathbf{x}^i & \mathbf{T}_i^{22}\mathbf{x}^i \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} l'_1 \\ l'_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{37}$$

Since the right null-space of the $\mathbf{M}$ is of non-zero dimension, it implies that $\det(\mathbf{M}) = 0$. Thus,

$$\mathbf{T}_p^{11}\mathbf{x}^p\mathbf{T}_q^{22}\mathbf{x}^q - \mathbf{T}_p^{12}\mathbf{x}^p\mathbf{T}_q^{21}\mathbf{x}^q = 0 \tag{38}$$

Note that the above equation is a conic in $\mathbf{x}$ of the form $\mathbf{x}^T\mathbf{W}\mathbf{x} = 0$ where the $(i,j)^{th}$ entry of $\mathbf{W}$ is of the form

$$\mathbf{W}_{ij} = \frac{\mathbf{T}_i^{11} * \mathbf{T}_j^{22} + \mathbf{T}_j^{11} * \mathbf{T}_i^{22} - \mathbf{T}_i^{12} * \mathbf{T}_j^{21} - \mathbf{T}_j^{12} * \mathbf{T}_i^{21}}{2} \tag{39}$$

Thus for a point $\mathbf{x}$ lying on $\mathbf{W}$ there exists a radial line, $\mathbf{l}'$ in the first radial image such that for any arbitrary $\mathbf{l}''$, Eq. (35) is satisfied (or there exists a radial line in the second image such that for any arbitrary radial line in the first image, Eq. (35) is satisfied). Note that the above condition holds only for points $\mathbf{x}$ which come from either line $\mathbf{m}_1$, which is the image of plane $\mathbf{\Pi}_1$ or $\mathbf{m}_2$ which is the image of the plane $\mathbf{\Pi}_2$. This implies that $\mathbf{W}$ is a degenerate conic (pair of lines). Thus

$$\det(\mathbf{W}) = 0 \tag{40}$$

Eq. (40) is the degree-six internal constraint on the entries of $\mathbf{T}_i^{jk}$.

6.2 Computation of the Mixed Trifocal Tensor

The mixed trifocal tensor can be linearly estimated given at least 11 corresponding triplets of features, with each triplet giving a linear constraint on the parameters of the tensor using Eq. (35). However this method would not impose the internal constraint we have discussed above. We will now describe a technique by which the degree-six internal constraint discussed above can be imposed.

Given 10 corresponding triplets of features, we obtain a 10×12 measurement matrix and thus a two-dimensional right null-space. The trifocal tensor, $\mathbf{T}$ can then expressed as

$$\mathbf{T} = \mathbf{S} + \lambda\mathbf{R} \tag{41}$$

where λ needs to be determined.

Given two $3 \times 2 \times 2$ tensors $\mathbf{U}$ and $\mathbf{V}$, define an operator $\mathbf{Y}(\mathbf{U}, \mathbf{V})$, which gives as output a 3×3 matrix, as following:

$$\mathbf{Y}_{ij} = \frac{\mathbf{U}_i^{11} * \mathbf{V}_j^{22} + \mathbf{V}_j^{11} * \mathbf{U}_i^{22} - \mathbf{U}_i^{12} * \mathbf{V}_j^{21} - \mathbf{V}_j^{12} * \mathbf{U}_i^{21}}{2} \tag{42}$$

The degree-six internal constraint expressed by Eq. (40) can be rewritten in the above terms (using Eq. (39)) as:

$$\det(\mathbf{Y}(\mathbf{T}, \mathbf{T})) = 0 \tag{43}$$

Using Eq. (41), we can reduce Eq. (43) as following:

$$\begin{array}{l} \det(\mathbf{Y}(\mathbf{T}, \mathbf{T})) = 0 \\ \det(\mathbf{Y}(\mathbf{S} + \lambda\mathbf{R}, \mathbf{S} + \lambda\mathbf{R})) = 0 \\ \det(\underbrace{\mathbf{Y}(\mathbf{S}, \mathbf{S})}_{\mathbf{D}_1} + \lambda(\underbrace{\mathbf{Y}(\mathbf{S}, \mathbf{R}) + \mathbf{Y}(\mathbf{R}, \mathbf{S})}_{\mathbf{D}_2}) + \lambda^2 \underbrace{\mathbf{Y}(\mathbf{R}, \mathbf{R})}_{\mathbf{D}_3}) = 0 \\ \qquad\qquad \underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\mathbf{D}} \end{array} \tag{44}$$

Note that in Eq. (44), $\mathbf{D}_1, \mathbf{D}_2$ and $\mathbf{D}_3$ are known. Further, since the value of λ would be such that the determinant of $\mathbf{D}$ is zero, there will exist $\mathbf{z}$ such that:

$$(\mathbf{D}_1 + \lambda\mathbf{D}_2 + \lambda^2\mathbf{D}_3)\mathbf{z} = 0 \tag{45}$$

Computing the tuple $(\lambda, \mathbf{z})$, such that Eq. (45) is satisfied, is the standard Quadratic Eigenvalue Problem (QEP) [46] and can be efficiently solved. For example, in MATLAB, one can use the `polyeig` function.

Since the size of the minimal hypothesis is 10 and the kernel can be efficiently implemented ($\mathbf{S}$ and $\mathbf{R}$ are estimated linearly and Eq. (45) can be solved efficiently), we can use a robust sieve, like RANSAC, to estimate the mixed trifocal tensor.

7 Experiments

7.1 Radial Quadrifocal Tensor

7.1.1 Using Synthetic Data

We will now describe simulations that we carried out to test the validity and robustness of reconstruction using the quadrifocal tensor. The following 4 cameras were chosen: a pin-hole camera looking at a spherical mirror, a pin-hole camera looking at a hyperbolic mirror (satisfying the single effective view-point condition [1]), a perspective camera and a fish-eye camera. Only the points which were imaged in all the 4 views were considered (a total of 2300 points were imaged, see Figure 7). To every point in every image, Gaussian noise with $\sigma = 1$ pixel (in an image of 2000x2000 pixels) was added. The final 4 images are shown in Figure 6. The 4 cameras were modeled to have zero skew and unit aspect ratio.

In Figure 8 the results of the metric reconstruction are shown. Only the difference vectors, between the ground truth and the reconstruction obtained, are plotted. The ratio of the RMS reconstruction error and the standard deviation of the ground truth point set is less than 1 percent. This ratio grew to around 3-5 percent when noise of $\sigma = 2$ pixel was introduced.

In the second phase, we perform radial calibration. For each camera, the first phase would have given us a precise optical axis in metric space. We select an arbitrary point on the optical axis and compute (ρ, z) pairs for each reconstructed point . In this 2D coordinate system, all points which project onto the same radial circle, in the input image, should lie on the same line. Given a sufficiently dense set of points, we can estimate these incoming rays (see Figure 6). Note that all the incoming rays for the three central cameras (views 2,3 and 4) pass through a point on the optical axis, as expected. This happens without enforcing any explicit constraint. For a non-central camera (view 1), the envelope of rays corresponds to the caustic of a spherical mirror as expected.

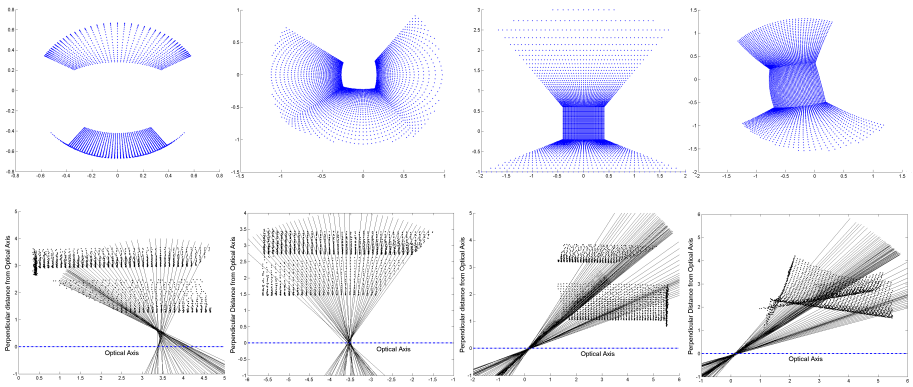


Fig. 6: *(L to R) Spherical mirror, Hyperbolic mirror, Pin-Hole Camera, Fish-eye Lens.* ***Top****:Images obtained by the four cameras* ***Bottom****:* (ρ, z) *plots for reconstructed features and estimated incoming rays. Notice the caustic of the spherical-mirror camera (extreme left)*

7.2 Radial Trifocal Tensor

7.2.1 Non-parametric Radial Calibration

In our first experiment, a triplet of images obtained using a rotating fish-eye camera was fed as input to the system. The images were acquired using a Nikon 8mm FC-E8 fish-eye converter mounted on a Nikon Coolpix 8400 camera. An online implementation of Lowe's feature matcher [27] was used to obtain triplets of corresponding points in the three images. Note that despite the severe non-perspective distortion, most automatic feature matching techniques work well because the views were obtained using a purely rotating camera.

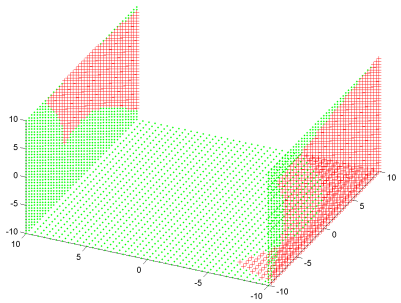


Fig. 7: *The scene that is imaged by the 4 cameras. Plus (+) signs mark the points imaged in all cameras and dot (.) mark points which werent.*

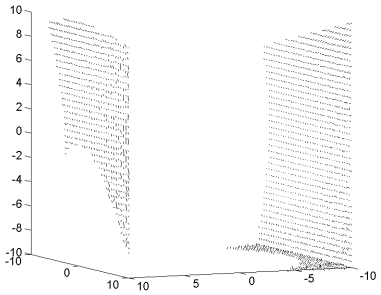


Fig. 8: *These line segments connect the ground-truth points to the metric reconstruction obtained. Compare the length of the segments to the extent of the scene to get an idea of the error in reconstruction.*

The image resolution was 1024x768 pixels. Approximately 560 triplets were returned by the feature matcher. Next, RANSAC based on the radial trifocal tensor identified about 220 inliers (the threshold was set to 3 pixels). The input images, with the triplets of corresponding points (those that were identified as inliers after RANSAC) marked are shown in Figure 9. A projective reconstruction was obtained and upgraded to metric based on the assumptions of zero skew and known aspect ratio of unity. For every 3D point that has correspondences across at least two images, we obtain the angle of the ray, passing through that point and the camera center, and the optical axis of the corresponding view. This gives us a point on the *angle vs. distorted radius* curve. Figure 10 shows the plots for the three views. We see that the angle of a ray with the optical axis is related to the distorted radius almost linearly. This is expected as a fish-eye camera roughly follows the equidistant model.

Note that at no point during the whole procedure did we make any assumptions about the type/amount of radial distortion. Further, an automatic feature matcher has been able to give us features that span the whole range of distorted radii. Finally, note that no additional constraint (smoothness etc) was enforced across the three views. Finally in Figure 11 we show a cubemap of the undistorted left view. Note that straight lines in the world are indeed mapped to straight lines in the image. The unwarping was carried out by computing the distorted radius for a given undistorted radius using

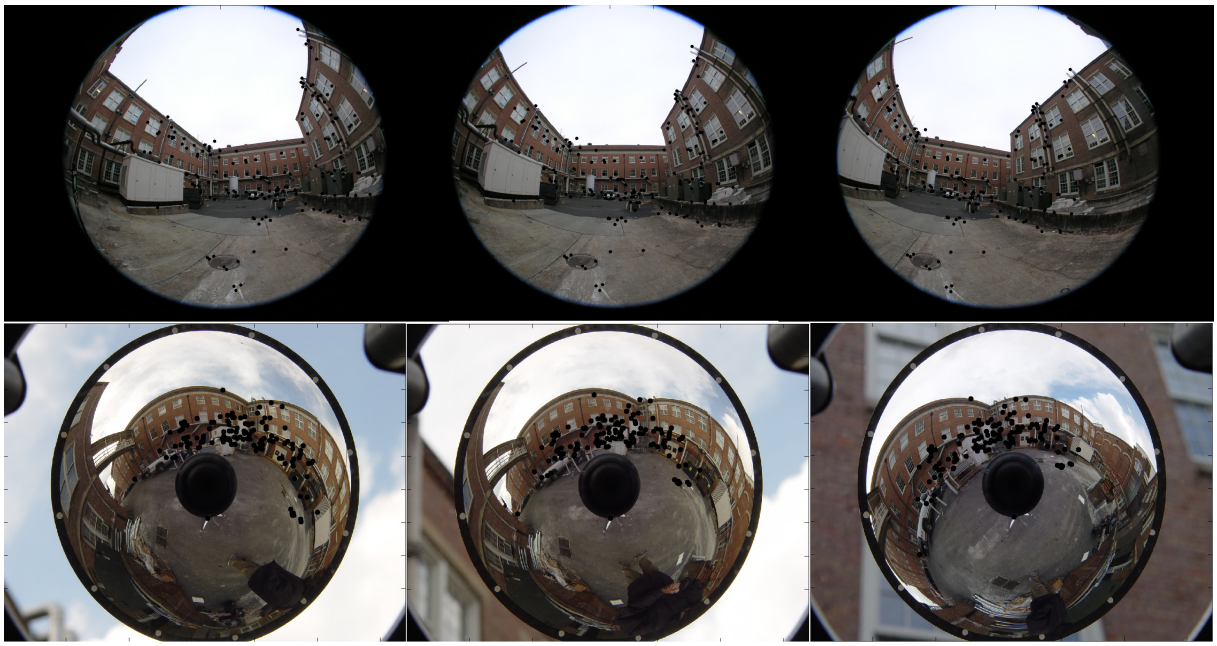


Fig. 9: *The triplet of images input to the system with features that were automatically matched overlaid.* ***Top****: Fish-eye images* ***Bottom****: Catadioptric images*

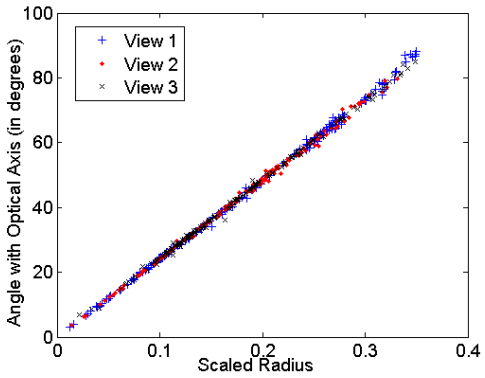


Fig. 10: *Plot for fish-eye camera. Radius in Distorted Image (r)* vs *Angle with Optical-Axis of ray passing through a pixel at that radius (θ, in degrees)*

a simple line interpolation on the plot in Figure 10. More complex models could also be used .

In our second experiment, three images obtained from a purely rotating single viewpoint catadioptric camera were used. The image resolution was 1280x960 pixels. Lowe's feature matcher [27] produced approximately 220 matching triplets across the three views. An in the previous experiment, RANSAC based on the radial trifocal tensor produced around 130 inlier triplets. Figure 9 (lower row) shows the input images

with the inlier triplets marked. We compute a projective reconstruction and upgrade it to metric based on assumptions of zero skew and unit aspect ratio. Note that since our method handles all types of radial distortion uniformly, the complete procedure in this experiment in exactly the same as in the previous experiment. Figure 12 shows the plots of the *Angle with Optical Axis vs. Distorted Radius* for each of the views. Finally, Figure 13 shows a cubemap of the undistorted left view. One can refine the estimates produced by our method using techniques like bundle adjustment. In the first experiment, it reduced the RMS reprojection error from 1.13 pixels to 0.43 pixels.

Fig. 11: *Cubemap of undistorted left image (unwarping done using a simple line interpolation between θ and r, the distorted radius)*

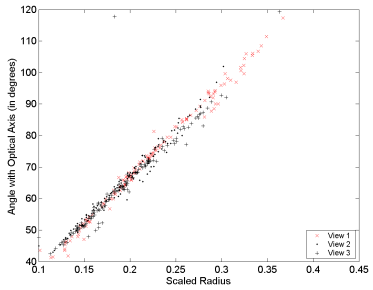


Fig. 12: *Plot for catadioptric camera. Radius in Distorted Image (r)* vs *Angle with Optical-Axis of ray passing through a pixel at that radius (θ, in degrees). Instead of a single curve, we have a band, probably because of the non-central nature of the camera.*

Fig. 13: *Cubemap of undistorted left image (unwarping done using a simple line interpolation between angle and r, the distorted radius)*

7.2.2 Radial Calibration from Camera in General Position

In this experiment, 3 images of a courtyard, acquired by a Sigma 8mm-f4-EX fish-eye lens with view angle 180^o mounted on a Canon EOS-1Ds digital camera were used. The image resolution was 2560x2560 pixels. Since the 3 views weren't obtained with a purely rotating camera, we input 44 corresponding triplets, that lie on a real-world plane (see Figure 14). We observed that the average clicking error was 1-3 pixels. As in the previous experiment, RANSAC, based on the radial trifocal tensor, was used, resulting in 30 inlier triplets. A second RANSAC based on reprojection error, was used to estimate the distortion parameters. Figure 15 plots the distortion curves when different number of parameters $(4-8)$ were used in the distortion model. A distortion model with 5 parameters was used to compute a undistorted image, for one of the views, using a cubemap projection (see Figure 16). Note that we are able to accurately undistort, not only regions in the center of the image, but also the periphery of the image. Since the images were acquired using a full 180^o fish-eye lens, it shows that the model is robust for wide-angle lenses with *very high* degree of distortion. In this case, the RMS reprojection error was around 2-3 pixels.

7.3 Comparison with Previous work

Among our approaches we selected the approach developed in Section 5 (i.e., using the radial trifocal tensor to calibrate a camera in general position). And we selected the OCamCalib toolbox (Omnidirectional Camera Calibration Toolbox [34,36]) as the reference implementation. We used average reprojection error (in the distorted input images), as a measure of the accuracy of the approach. Both the methods use the fact that all the grid-points lie on a plane. However, the toolbox has more input information (the metric coordinates of the grid-points) than us (we just use the fact that the grid-points in the three images are corresponding, but do not know the coordinates of the points on the plane). The results reported here are by using the direct linear technique for estimation (results will improve further on doing bundle adjustment). Also, the toolbox recommends the use of a routine that iteratively tries to find the best center-

Fig. 14: *Three images, taken with different camera centers, input to the system (matching points input to the system are marked). Images courtesy Tomas Pajdla.*

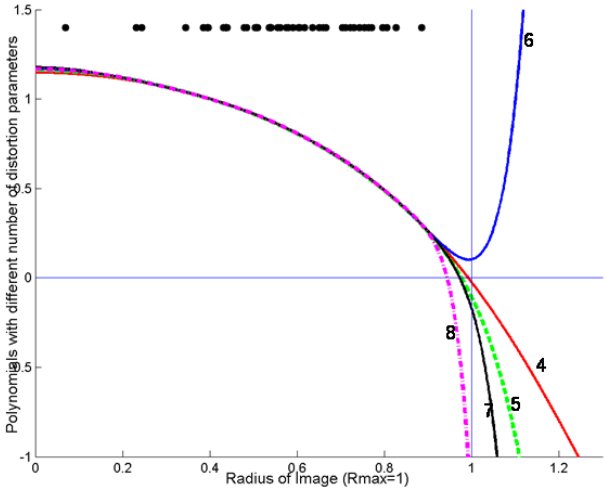


Fig. 15: *Distortion Curves* $1 + K_1 r_d^2 + \cdots + K_n r_d^{2n}$ *when different number parameters* ($n = 4 - 8$, *marked next to corresponding curve) are used. Note that most of the curves are well-behaved even at* $r = R_{max}$.

of-distortion. The results discussed here, for the toolbox, are after having done that step.

Three images of a checkerboard taken from a fish-eye camera were used in the first experiment (these involved relative motion of the camera with respect to the world-plane, in this case the checkerboard). 25 points across 3 views were manually extracted. These were fed to the toolbox and 4 coefficients (as directed by the toolbox tutorial) were used for parameterizing the distortion polynomial. The toolbox gave an average error of 0.25 pixels. In our approach we used the division-model of radial distortion with 3 parameters. This gave us an average reprojection error of 0.45 pixels.

Fig. 16: *Cubemap of undistorted left image (warping done using 5 distortion parameters)*

In our second experiment, we selected three images of the checkboard taken with a Ladybug2 camera. Once again 25 correspondences across 3 views were used. With the same parameters as the first experiment, the toolbox gave an average reprojection error of 0.24 pixels. With our approach we got a reprojection error of 0.7 pixels.

Using the average reprojection-error as a cost function, we developed a gradient-descent approach to estimate the center-of-distortion. It used widely spaced positions as initial guesses to the optimization. Preliminary results have been encouraging, which we will pursue in the future.

8 Conclusion

The 1D radial camera maps 3D points on radial lines. This allows us to derive multilinear constraint between three and four views recorded with central or non-central omnidirectional cameras. One of the main contributions of this work has been to come up with an approach for reconstruction, while still making only very minimal assumptions about the kind of distotion. This allows us to use this reconstruction as an accurate calibration object to estimate the distortion in a second step.

Given 15 or more correspondences across four views taken with a moving camera (or multiple cameras), the radial quadrifocal tensor allows us to compute the corresponding 1D radial cameras and a metric reconstruction of 3D points. The reconstruction is then used to estimate a non-parametric camera model for different cameras, including a non-central cameras. Although the approach has been demonstrated to work well on synthetic data, the required number of point correspondences makes it hard to develop a robust automatic approach for real images.

For a purely rotating camera a simpler constraint (radial trifocal tensor) is obtained requiring 7 point correspondences across 3 views. As with perspective cameras [16, 5], pure rotation turns out to be particularly well suited for self-calibration of central omnidirectional cameras. In particular, we present an automatic approach that recovers the accurate non-parametric distortion curve relating image radius to angle of incoming rays. Finally we introduce an approach that does not need the assumption of pure rotation, but does parametric distortion estimation.

In the future, we plan to study various mixed camera tensors (for example, with two pin-hole and a single 1D camera). This will be very useful as heterogeneous camera networks become more and more commonplace.

References

1. S. Baker and S. Nayar. A theory of single-viewpoint catadioptric image formation. *IJCV*, 35(2):1 – 22, 1999.
2. H. Bakstein and T. Pajdla. Panoramic mosaicing with a field of view lens. In *Proc. IEEE Workshop on Omnidirectional Vision*, pages 60–67, 2002.
3. R. Benosman and S.B. Kang. *Panoramic Vision: Sensors, Theory and Applications*. 2001.
4. D.C. Brown. Close-range camera calibration. *PhEng*, 37(8):855–866, August 1971.
5. L. d. Agapito, R. Hartley, and E. Hayman. Linear selfcalibration of a rotating and zooming camera. In *CVPR*, 1999.
6. F. Devernay and O.D. Faugeras. Straight lines have to be straight. *MVA*, 13(1):14–24, 2001.
7. F. Espuny and J.I. Burgos Gil. Generic self-calibration of central cameras from two real rotational flows. 2008.
8. Ferran Espuny. A closed-form solution for the generic self-calibration of central cameras from two rotational flows. In *VISAPP (1)*, pages 26–31, 2007.
9. Olivier D. Faugeras, Long Quan, and Peter Sturm II. Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1179–1185, 2000.
10. A.W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR01*, pages I:125–132, 2001.
11. C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. In *CVPR*, 2001.
12. A. Goshtasby. Correction of image deformation from lens distortion using bezier patches. *CVGIP*, 47(3):385–394, September 1989.
13. M. Grossberg and S. Nayar. A general imaging model and a method for finding its parameters. In *ICCV*, 2001.
14. E. Grossmann, E.J. Lee, P. Hislop, D. Nister, and H. Stewenius. Are two rotational flows sufficient to calibrate a smooth non-parametric sensor? In *CVPR06*, pages 1222–1229, 2006.
15. Etienne Grossmann, Jos Antnio Gaspar, and Francesco Orabona. Discrete camera calibration from pixel streams. *Computer Vision and Image Understanding*, 114:198–209, 2010.
16. R. Hartley. Self-calibration from multiple views with a rotating camera. In *ECCV*, pages 471–478, 1994.
17. R.I. Hartley. Invariants of lines in space. In *DARPA93*, pages 737–744, 1993.
18. R.I. Hartley and S.B. Kang. Parameter-free radial distortion correction with center of distortion estimation. 29(8):1309–1321, August 2007.
19. R.I. Hartley and F. Schaffalitzky. Reconstruction from projections using grassmann tensors. In *ECCV04*, pages Vol I: 363–375, 2004.
20. R.I. Hartley and A. Zisserman. Multiple view geometry in computer vision. In *Cambridge*, 2000.
21. C. Hughes, P. Denny, M. Glavin, and E. Jones. Equidistant fish-eye calibration and rectification by vanishing point extraction. 32(12):2289–2296, December 2010.
22. C. Hughes, R. McFeely, P. Denny, M. Glavin, and E. Jones. Equidistant f-theta fish-eye perspective with application in distortion centre estimation. 28(3):538–551, March 2010.

23. S.B. Kang. Radial distortion snakes. In *IAPR Workshop on Machine Vision Applications (MVA2000), Tokyo, Japan*, pages 603–606, 2000.
24. J. Kannala and S. Brandt. A generic camera calibration method for fish-eye lenses. In *ICPR*, 2004.
25. Juho Kannala, Sami S. Brandt, and Janne Heikkilä. Self-calibration of central cameras by minimizing angular error. In *VISAPP (1)*, pages 28–35, 2008.
26. H.D. Li and R. Hartley. Plane-based calibration and auto-calibration of a fish-eye camera. In *ACCV06*, pages I:21–30, 2006.
27. D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
28. B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *CVPR03*, pages I: 485–490, 2003.
29. D. Nister, H. Stewenius, and E. Grossmann. Non-parametric self-calibration. pages I: 120–127, 2005.
30. M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *IJCV*, 1999.
31. L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *PAMI*, 1997.
32. Srikumar Ramalingam, Peter Sturm, and Suresh Lodha. Generic self-calibration of central cameras. *Computer Vision and Image Understanding*, 114(2):210–219, feb 2010.
33. Srikumar Ramalingam, Peter Sturm, and Suresh K. Lodha. Towards generic self-calibration of central cameras. In *In proc. ICCV workshop on Omnidirectional Vision, Camera Networks and Nonclassical cameras OMNIVIS*, 2005.
34. M. Rufli, D. Scaramuzza, and R. Siegwart. Automatic detection of checkerboards on blurred and distorted images. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3121 –3126, sept. 2008.
35. Ryusuke Sagawa, Takurou Sakai, Tomio Echigo, Keiko Yagi, Masatsugu Shiba, Kazuhide Higuchi, Tetsuo Arakawa, and Yasushi Yagi. Omnidirectional vision attachment for medical endoscopes. In *OMNIVIS08*, 2008.
36. Davide Scaramuzza and Agostino Martinelli. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *Proceedings of IEEE International Conference on Computer Vision Systems*, 2006.
37. S. Shah and JK Aggarwal. Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, 29(11):1175–1788, 1996.
38. G.P. Stein. Lens distortion calibration using point correspondences. In *CVPR97*, pages 602–608, 1997.
39. P. Sturm and S. Ramalingam. A generic concept for camera calibration. In *ECCV*, volume 2, pages 1–13, 2004.
40. R. Swaminathan and S.K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *PAMI*, 22(10):1172–1178, October 2000.
41. J.P. Tardif, P.F. Sturm, and S. Roy. Self-calibration of a general radially symmetric distortion model. In *ECCV06*, pages IV: 186–199, 2006.
42. J.P. Tardif, P.F. Sturm, and S. Roy. Plane-based self-calibration of radial distortion. In *ICCV07*, pages 1–8, 2007.
43. S. Thirthala and M. Pollefeys. Multi-view geometry of 1d radial cameras and its application to omnidirectional camera calibration. In *ICCV05*, pages II: 1539–1546, 2005.
44. S. Thirthala and M. Pollefeys. The radial trifocal tensor: A tool for calibrating the radial distortion of wide-angle cameras. In *CVPR*, 2005.
45. S. Thirthala and M. Pollefeys. Trifocal tensor for heterogenous cameras. In *OMNIVIS05*, 2005.
46. Françoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. Technical Report No. 370, Manchester, England, 2000.
47. B. Triggs. Matching constraints and the joint image. In *ICCV*, pages 338–343, 1995.
48. B. Triggs. Autocalibration and the absolute quadric. In *CVPR*, pages 609–614, 1997.
49. R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *RA*, 3(4):323–344, 1987.
50. J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *PAMI*, 14(10):965–980, October 1992.
51. Y. Xiong and K. Turkowski. Creating image-based vr using a self-calibrating fisheye lens. In *CVPR*, pages 237–243, 1997.

Table 1: Summary of Results

Radial Tensor	Results
Quadrifocal	• **Camera Type**: Both central and non-central • **Practical**: No. Too sensitive to noise. More for theoretical interest. • **Reconstruction**: Metric[a] • **Radial Distortion Calibration**: Non-parametric but requires very dense correspondences. • **Comments**: Has 13 d.o.f. The nature of two constraints explained geometrically.

Trifocal	• **Comments**: Has 7 d.o.f. And **no** internal constraints. Hence easy to estimate.
	• *Rotating Camera Triplet*
	  ◦ **Camera Type**: Required to be central for radial calibration to work.
	  ◦ **Practical**: Yes. Very robust to noise and easy to estimate.
	  ◦ **Reconstruction**: Metric. Reconstruct Π_∞ (plane of all directions). And use that a calibration device.
	  ◦ **Radial Distortion Calibration**: Non-parametric. Using Π_∞ compute the angle between a reconstructed ray and the optical axis.
	• *Cameras in general position but scene is a plane*
	  ◦ **Camera Type**: Central only.
	  ◦ **Practical**: Yes. Very robust to noise and easy to estimate.
	  ◦ **Reconstruction**: Projective reconstruction of scene-plane.
	  ◦ **Radial Distortion Calibration**: Only parametric. Simultaneous estimation of distortion parameters *and* last row of homography, $\mathbf{H}$: scene-plane $\rightarrow$ undistorted-image.

Trifocal Tensor for Heterogeneous Cameras[b]	• **Camera Type**: Non-central handled too. • **Practical**: Haven't completely explored sensitivity to noise. • **Reconstruction**: Projective. • **Radial Distortion Calibration**: Not discussed. • **Comments**: Has 10 d.o.f and 1 internal constraint. This constraint is derived geometrically.

[a] based on zero skew, known principal point and aspect ratio.

[b] Lists results as discussed in this paper. For example, radial distortion calibration might still be feasible on further investigation.