submitted to CVPR 2005

The Radial Trifocal Tensor: A tool for calibrating the radial distortion of wide-angle cameras

SriRam Thirthala [*] Marc Pollefeys [†]

Abstract

We present a technique to linearly estimate the radial distortion of a wide-angle lens given three views of a real-world plane. The approach can also be used with pure rotation as in this case all points appear as lying on a plane. The three views can even be recorded using three different cameras as long as the deviation from the pin-hole model for each camera is distortion along radial lines. We introduce the 1D radial camera which projects scene points onto radial lines and the radial trifocal tensor which encodes the multi-view relations between radial lines. Given at least seven triplets of corresponding points the radial trifocal tensor can be computed linearly. This allows recovery of the radial cameras and the projective reconstruction of the plane up to a two-fold ambiguity. This 2D reconstruction is unaffected by radial distortion and can be used in different ways to compute the radial distortion parameters. We propose to use the division model as in this case we obtain a linear algorithm that computes the radial distortion coefficients and the 3 remaining degrees of freedom of the homography relating the reconstructed 2D plane to the undistorted image. Each feature point that has at least one corresponding point yields one linear constraint on those unknowns. Our method is validated on real-world images. We successfully calibrate several wide-angle cameras.

1. Introduction

For many vision applications, cameras with large field of view are required. Wide-angle lenses or curved mirrors obtain a large field of view, by severely bending the rays, but the corresponding camera projection model is far more complicated than the traditional pin-hole model. A major problem in the calibration procedure is the non-linear relation between the image and space coordinates.

This paper deals with the problem of estimating the coefficients of the non-linear transformation (henceforth, called the distortion parameters), that maps points in the distorted (input) image to points in the undistorted image (i.e, one that would conform to the pin-hole model). Computing the distortion parameters would allow us to use images having large radial distortion, for most applications in 3D computer vision (which make the pin-hole assumption), by using the *transformed image coordinates* rather than the input image coordinates.

We now present a short overview of methods used for recovering distortion parameters.

The first class of methods do so with the aid of features whose coordinates in the 3D space are known (for example [14]). In [5], Goshtasby uses Bezier patches to model the distortions and uses a uniform grid as a calibration object. Weng et al. [15] also uses calibration objects to extract distortion parameters.

The second category of methods do not rely on known scene points but use the property of the pin-hole model, that straight lines in space must project onto straight lines in the image. Brown [1] used the above technique but had noiseless image data by imaging plumb-lines. The method proposed in [13] also falls into this category. In their approach, the user clicks points on image curves that (s)he knows are straight lines in the scene and an objective function is constructed that tries to minimize the deviation of these curves from straight lines. The parameters, over which this function is minimized, are the distortion parameters. In [8], Kang used snakes to represent the distortion curves. Devernay and Faugeras [2] 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.

The first category of methods suffer from the requirement of known calibration objects. This requirement makes them unsuitable to use with variable lens geometries (for eg., with variable zoom), because of the strong coupling that exists among the estimates of the parameters of a camera. The second category of methods require the presence of straight lines in the scene (which might not always be the case). Further, it requires that that image curves which could have possibly come from lines in the scenes, be robustly detected. This is non-trivial in general, since an automatic system can confuse real-world curves with straight lines and thus may require manual input of points or supervision of the system.

[*] tvn@cs.unc.edu, Dept of Computer Science, UNC-Chapel Hill
[†] marc@cs.unc.edu, Dept of Computer Science, UNC-Chapel Hill

The third category of methods does so, by using point correspondences. Stein [12] proposes a method in which he uses epipolar and trifocal constraints. That is, given corresponding points in three (distorted) images, he computes the parameters of the trilinear equations. These parameters are then used to reproject points into the third image, given corresponding points in the first two images. The cost function is defined as the RMS reprojection error and minimized over the distortion parameters. In [4], Fitzgibbon proposes a technique for simultaneous estimation of the fundamental matrix and the lens distortion parameters, by formulating the problem as a quadratic-eigenvalue problem (QEP). However, his approach concentrates on being able to "allow matching of image pairs via interest-point correspondences, when lens distortion would otherwise hinder the process" and *does not* yield an accurate estimation of the distortion parameters themselves. While the above method may be applicable for small lens distortions, it is not be suitable for large distortions, such as those produced by curved mirrors/fish-eye lenses etc. Micusik and Pajdla ([9]), also formulate the estimation of the fundamental matrix and the distortion coefficients, as a QEP.

The method that is proposed in this paper also requires corresponding points (which come from any plane in the scene) across three views. However, we consider the distorted input image as a 1D image of radial lines generated by a radial camera. Thus, only the *directions* of the feature points in the image (from the center of radial distortion), which are known precisely, are used. This allows us to factor out the radial component of the projection model (where, by the projection model, we mean a conversion of 3D space coordinates to undistorted image coordinates, i.e. adhering to the pin-hole model, followed by some distortion along the radial line). Thus whatever be the deviation from the pin-hole model along radial lines (i.e points being pulled towards/away from the center of radial distortion), it does not affect the estimation of the parameters of the trifocal tensor or the projective reconstruction of the scene plane that we obtain.

Since we have obtained a projective reconstruction of the plane (which is equivalent to one obtained from 3 pin-hole camera images), we have upto a homography what the plane looks like in the undistorted image. Further the parameters of the 1D radial camera that we can estimate from the trifocal tensor, fix 5 parameters of the homography. Thus we have the undistorted positions of the feature points upto the three unknown parameters of the homography. This allows us to linearly estimate the distortion parameters.

Therefore, the contribution of our paper is two-fold. First, by introducing the radial trifocal tensor, we are able to linearly estimate the structure of the observed plane independent of arbitrary radial distortion. Once this structure has been computed the plane can be used as a (projective) calibration object and several approaches can be used to recover a model for radial distortion. Secondly, we propose a linear method to compute the radial distortion of wide-angle cameras. The main contribution of our approach is to separate the estimation of the multi-view relation and the estimation of the distortion coefficients in two linear steps.

Notation: Vectors will be denoted in bold, for example $\mathbf{x}$ while scalars will be in normal, like x. For the camera matrices, the letters, $\mathbf{P}$, $\mathbf{P}^{'}$, $\mathbf{P}^{''}$ will be used. Whether the coordinates are distorted or undistorted, will be made clear by the subscripts (such as $\mathbf{x}_d$ and $\mathbf{x}_u$ respectively). The scene plane ,which contains the points whose images are matched in the three images, is denoted by Π. The distorted (input) images are denoted by I_d^i where $i = 1, 2, 3$. The undistorted images that conform to the pin-hole model are denoted by I_u^i where $i = 1, 2, 3$. If the size of a matrix is not clear, it will be pointed out in the subscript (such as $\mathbf{P}_{2\times3}$ and so on).

2. Radial Distortion Models

Let the center of radial distortion be $\mathbf{c_{rad}} = (c_{xr}, c_{yr})$. The standard model ([11]) for lens distortions gives the mapping from the distorted image coordinates, $\mathbf{x_d} = (x_d, y_d)$, that are observable to the undistorted coordinates $\mathbf{x_u} = (x_u, y_u)$, by the equation

$$\mathbf{x_u} = \mathbf{x_d} + \mathbf{x'_d}(K_1 {r'_d}^2 + K_2 {r'_d}^4 + K_3 {r'_d}^6 + \ldots) \qquad (1)$$

where $\mathbf{x'_d} = (\mathbf{x_d} - \mathbf{c_{rad}})$ and $r_d = || \mathbf{x'_d} ||$

Other models for radial distortion have been proposed. Fitzgibbon [4] proposed the *division model* where,

$$\mathbf{x_u} = \frac{\mathbf{x_d}}{(1 + K_1 {r_d}^2 + K_2 {r_d}^4 + K_3 {r_d}^6 + \ldots)} \qquad (2)$$

The above equation assumes that the center of radial distortion is given and the distorted images, I_d^is are transformed so that the center of radial distortion is the origin.

Among the other models proposed, an important one is the fish-eye or the equidistant model. The model proposes that the distance between the an image point and the center of radial distortion is proportional to the angle between the corresponding 3D point, the optical center and the optical axis.

In this paper, we *assume that the center of radial distortion is known* and the distorted images are transformed so that the center of radial distortion is the origin. Typically, we assume that the center of radial distortion coincides with the center of the image. We have experimentally verified that this is a good approximation and that including parameters for the center of radial distortion in the estimation does not significantly improve the results. For the purpose of estimating the radial distortion parameters, we show results by using the division model. However, we are free to choose the type/parameters of the radial distortion model, that we deem fit, because we have been able to separate the the estimation of the multi-view relation (the radial trifocal tensor) and the estimation of the parameters of radial distortion into two different stages (in contrast to [4], [9], where the estimation is done simultaneously and the problem formulation is dependent on the radial distortion model used).

3. Radial 1D Camera

Let the center of radial distortion be the origin. In the presence of large, unknown radial distortion, only the direction in the image is precisely known. Consider the image point, $\mathbf{x_d} = (x_d, y_d, 1)^T$. The direction to this point from the center of radial distortion can be represented by the 1D homogenous vector $\mathbf{d} = (y, -x)^T$. A line passing through $\mathbf{x_d} = (x, y, 1)^T$ and the $\mathbf{c}_{rad}$ (which is equal to the origin) is given by $\hat{\mathbf{l}}_{rad}$=$\mathbf{x_d} \times \mathbf{c}_{rad}$ =$(y, -x, 0)^T$. Since all radial lines, $\hat{l}_{rad}$, have their last component equal to zero, we can represent the space of radial lines, using 1D homogenous vectors. Thus, we will denote the radial lines, by $\mathbf{l}_{rad} = (y, -x)^T$. Note that the undistorted image point corresponding to $\mathbf{x_d}$ lies on $\hat{\mathbf{l}}_{rad}$. Thus by representing the distorted image as a 1D image consisting of radial lines, we factor out the *unknown deviation* from the pin-hole model (which is along the radial line), but preserve the *precise information* (which is the direction of the radial line).

Definition: The radial 1D camera represents the mapping of a point on the scene-plane, Π, to a radial line in the image (i.e., the line passing through the the center of radial distortion). Since it is a mapping from $\mathbf{P}^2$ to $\mathbf{P}^1$, it can be represented by a 2×3 matrix and has 5 degrees of freedom.

4. Radial Trifocal Tensor

Consider the point X, lying on Π, that projects onto the lines $\mathbf{l}_{2\times1}$, $\mathbf{l}'$ and $\mathbf{l}''$. Then it projects by the following set of equations,

$$\begin{array}{rcl} \lambda\mathbf{l} & = & \mathbf{P}_{2\times3}\mathbf{X} \\ \lambda'\mathbf{l}' & = & \mathbf{P}'_{2\times3}\mathbf{X} \\ \lambda''\mathbf{l}'' & = & \mathbf{P}''_{2\times3}\mathbf{X} \end{array} \tag{3}$$

These equations can be rewritten in matrix format as,

$$\begin{bmatrix} \mathbf{P}_{2\times3} & \mathbf{l} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'_{2\times3} & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \mathbf{P}''_{2\times3} & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ -\lambda \\ -\lambda' \\ -\lambda'' \end{bmatrix} = \mathbf{0} \tag{4}$$

Since we know that a solution exists, the right null-space of the 6×6 measurement matrix should have non-zero dimension, which implies that

$$\det \begin{bmatrix} \mathbf{P}_{2\times3} & \mathbf{l} & \mathbf{0} & \mathbf{0} \\ \mathbf{P}'_{2\times3} & \mathbf{0} & \mathbf{l}' & \mathbf{0} \\ \mathbf{P}''_{2\times3} & \mathbf{0} & \mathbf{0} & \mathbf{l}'' \end{bmatrix} = 0 \tag{5}$$

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

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \mathbf{T}_{ijk}\mathbf{l}_i\mathbf{l}'_j\mathbf{l}''_k = 0 \tag{6}$$

$\mathbf{T}_{ijk}$ is the $2 \times 2 \times 2$ homogeneous radial trifocal tensor of the three 1D radial cameras. Elements of $\mathbf{T}$ can be written as 3×3 minors of the joint projection matrix $\left[\mathbf{P}^T\,\mathbf{P}'^T\,\mathbf{P}''^T\right]^T$ with each row (of the minor) coming from a different camera matrix.

It can be shown that for 1D cameras observing a plane, we can obtain no higher-order constraints (i.e, from 4 or more views). Further, the radial trifocal tensor is a minimal parameterization of the three 1D cameras as the d.o.f can be shown to match, $2 \times 2 \times 2 - 1 = 7 = 3 \times (2 \times 3 - 1) - (3 \times 3 - 1)$ (with the LHS being the d.o.f of $\mathbf{T}$ and the RHS being the d.o.f of the three uncalibrated views upto 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 equation [6]) Given more than seven correspondences, we can obtain the linear least squares solution. Since the size of the minimal hypothesis, for the radial trifocal tensor, is 7 and it can be estimated linearly, we can use a robust sieve, like RANSAC, to estimate it.

The trifocal tensor for 1D cameras and its properties were first studied in [3] in the context of planar motion recovery.

5. Reconstruction of the Plane

We now consider the problem of reconstructing points (on the plane, Π) whose corresponding image triplets have been identified in the three views. Note that the input points are in the distorted images and hence only the direction information from these points is precise, but the distance from the center of radial distortion is unknown. However, by considering the 2D distorted images as 1D images consisting of radial lines, and then computing the corresponding *radial* trifocal tensor, we have been able to glean only the information that is conforming to the pin-hole model and now we will do a reconstruction based on it. Also note that we have not made any assumption about the type/parameters of the distortion model during the estimation of the radial trifocal tensor (which was dealt with in the previous section) nor shall any assumption be made about the distortion model during the projective reconstruction of the plane, Π.

Given the radial trifocal tensor, we can estimate the three uncalibrated camera matrices (see Appendix for the details). However, for every valid radial trifocal tensor, we will have two possible triplets of camera matrices that generate the same radial trifocal tensor. This inherent two-way ambiguity was studied by [10] and also in [6]. We obtain two possible (projective) reconstructions of the plane Π, from the two sets of camera matrices. This ambiguity will be resolved once we include additional constraints by fitting our radial distortion model (see next section).

Suppose we have calibrated the three radial cameras upto a projectivity. Points on the real-world plane, Π , can then be reconstructed by back-projecting the corresponding radial lines [7].

$$\begin{aligned}\mathbf{L} &= \mathbf{P}^T\mathbf{l}\\ \mathbf{L}' &= \mathbf{P}'^T\mathbf{l}'\\ \mathbf{L}'' &= \mathbf{P}''^T\mathbf{l}''\end{aligned} \tag{7}$$

Since we are reconstructing points on a plane, Π, *only two* lines are required to obtain a unique point. With three lines, we can find a least-square solution as the right singular vector of the line matrix, $[\mathbf{L}\,\mathbf{L}'\,\mathbf{L}'']^T$. Note that one can find more matching features in between two views as compared to across three views. Thus, once we have estimated the radial trifocal tensor using corresponding triplets of points, we can reconstruct any feature on Π that can be matched in two views. These additional points, thus would give us more data to estimate the radial distortion parameters for a particular view. To avoid the inclusion of outliers, a robust procedure is also used when computing the distortion parameters.

6. Estimating Distortion Parameters

We have only used the direction of the triplets in their corresponding distorted images, to compute the reconstruction. Thus, we now have a projective reconstruction of points on the real-world plane, Π, as if we had started from three images conforming to the pin-hole model. We now wish to estimate the distortion parameters that would take points from I_d^1 to I_u^1.

6.1. Estimating the homography from Π to the undistorted images

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

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{bmatrix}\mathbf{X} \tag{8}$$

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 Π 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

$$\mathbf{H} = \begin{bmatrix} -\mathbf{p}_2 \\ \mathbf{p}_1 \\ \mathbf{h}_3 \end{bmatrix} \tag{9}$$

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, $\mathbf{H}$, upto three unknown parameters, as we have done above, we are able to express the set, S_u, as

$$S_u(h_{31}, h_{32}, h_{33}) = \{ \begin{bmatrix} -\mathbf{p}_2\cdot\mathbf{X}^i \\ \mathbf{p}_1\cdot\mathbf{X}^i \\ [h_{31}\,h_{32}\,h_{33}]\cdot\mathbf{X}^i \end{bmatrix} \mid i = 1\ldots n\} \tag{10}$$

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

6.2. Computing the distortion parameters

We will now estimate the distortion parameters of the division model [1]. The transformation from I_d^1 to I_u^1, induced by the distortion parameters,is

$$\mathbf{x}_u = \frac{\mathbf{x}_d}{(1 + K_1 {r_d}^2 + K_2 {r_d}^4 + K_3 {r_d}^6 + \ldots)} \tag{11}$$

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

$$\lambda\begin{bmatrix} \mathbf{x}_u \\ 1 \end{bmatrix} = \begin{bmatrix} -\mathbf{p}_2\mathbf{X} \\ \mathbf{p}_1\mathbf{X} \\ \mathbf{h}_3\mathbf{X} \end{bmatrix} \tag{12}$$

with λ 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 [7].

$$\begin{bmatrix} -\mathbf{p}_2\mathbf{X} \\ \mathbf{p}_1\mathbf{X} \\ \mathbf{h}_3\mathbf{X} \end{bmatrix} \times \begin{bmatrix} x_d \\ y_d \\ (1 + K_1 {r_d}^2 + \ldots) \end{bmatrix} = \mathbf{0} \tag{13}$$

[1] 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.

Thus every point gives us two equations,

$$\begin{bmatrix} x_d(\mathbf{h}_3\mathbf{X}) + \mathbf{p}_2\mathbf{X}(K_1 r_d^2 + \ldots) \\ y_d(\mathbf{h}_3\mathbf{X}) - \mathbf{p}_1\mathbf{X}(K_1 r_d^2 + \ldots) \end{bmatrix} = \begin{bmatrix} (-\mathbf{p}_2\mathbf{X}) \\ (\mathbf{p}_1\mathbf{X}) \end{bmatrix} \tag{14}$$

which can be rewritten as,

$$\begin{bmatrix} x_d\mathbf{X} & (\mathbf{p}_2\mathbf{X})[r_d^2 \quad r_d^4 \ldots] \\ y_d\mathbf{X} & (-\mathbf{p}_1\mathbf{X})[r_d^2 \quad r_d^4 \ldots] \end{bmatrix} \begin{bmatrix} \mathbf{h}_3^\top \\ K_1 \\ K_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} (-\mathbf{p}_2\mathbf{X}) \\ (\mathbf{p}_1\mathbf{X}) \end{bmatrix} \tag{15}$$

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

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(15), by $\frac{1}{\mathbf{h}_3\mathbf{X}}$. This would then minimize the sum (over all the feature points) of the following squared-error.

$$\left|\left| \begin{bmatrix} x_d - \frac{-\mathbf{p}_2\mathbf{X}}{\mathbf{h}_3\mathbf{X}}(1 + K_1 r_d^2 + \ldots) \\ y_d - \frac{\mathbf{p}_1\mathbf{X}}{\mathbf{h}_3\mathbf{X}}(1 + K_1 r_d^2 + \ldots) \end{bmatrix} \right|\right|^2 \tag{16}$$

which is distance, in I_d^1, from $(x_d, y_d)^T$ to $\begin{bmatrix} \frac{-\mathbf{p}_2\mathbf{X}}{\mathbf{h}_3\mathbf{X}} & \frac{\mathbf{p}_1\mathbf{X}}{\mathbf{h}_3\mathbf{X}} \end{bmatrix}^T (1 + K_1 r_d^2 + \ldots)$ i.e, the pixel corresponding to the feature point in I_u^1, warped by the distortion parameters ($(1 + K_1 r_d^2 + \ldots)$). 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. (16) could be used to refine our linear solution. The results described in the experimental section are obtained using the linear method only.

7. Experiments

In our first experiment, the input image-set was a triplet obtained by a rotating camera. The images were acquired using a Nikon 16mm fish-eye lens mounted on a Kodak-DCS760 camera. The image resolution was 3032x2008 pixels. 40 triplets were hand-clicked and fed as input to the system. We estimate the input error was 2-3 pixels/point (see Figure 3). RANSAC based on the radial trifocal tensor, produced 36 inliers (when the threshold was set to 3-4 pixels). A second RANSAC based on reprojection error, produced 29 inliers (threshold being 2 pixels). Figure 1 plots $(1 + K_1 r_d^2 + \cdots K_4 r_d^8)$ vs. r_d, obtained when we consider points from only one view (for each view) and when points from all the three views are combined. Note that the curve for view 2, deviates from the others around radius of 1.2-1.4 . This is expected as view 2 (see Figure 3) has most of the input points concentrated at the center of the image. We

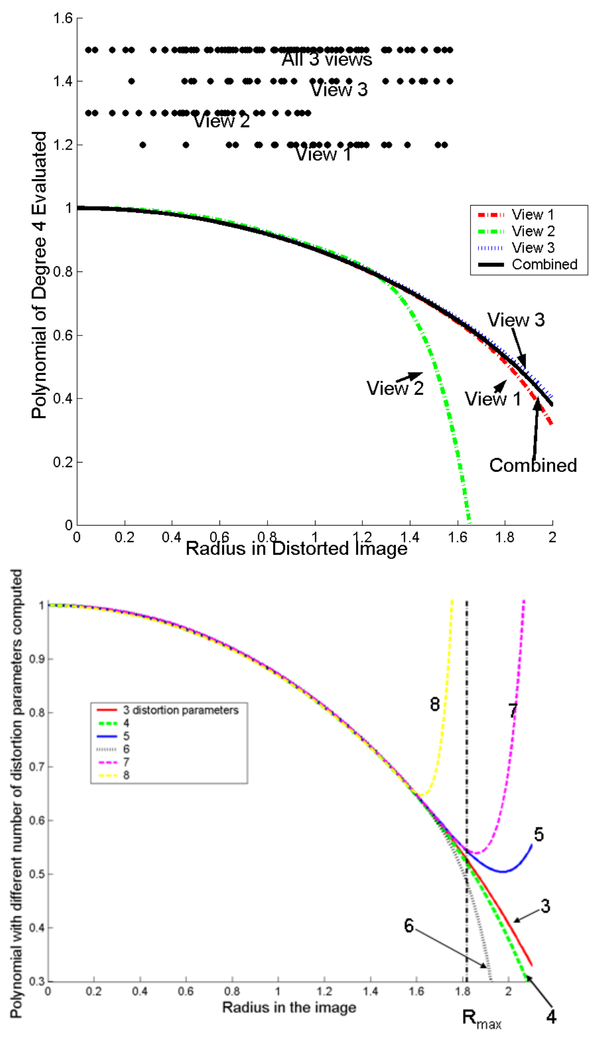

Figure 1: *Top*:Distortion curves, when coefficients are computed by using feature points from individual/all views. Radius of points used shown at the top. *Bottom*: Linearly estimated distortion curves for varying number of coefficients (3-8)

also examined the performance of our procedure with different number of distortion parameters (3-8). Figure 1 plots

Figure 3: Three images taken with a rotating camera (with selected feature triplets marked)

the corresponding distortion curves. Note that for most of the image the curves are very close. The deviation occurs only in the periphery of the image (the radius of the image is shown by the vertical line at R_{max}). Finally in Figure 2 we display the undistorted image obtained by warping the input images with the computed distortion parameters. Note how straight lines in the scene appear as straight lines in the images, even at the periphery of the image. The RMS reprojection error is less than a pixel.

Figure 2: Left Image unwarped to conform to pin-hole model, using 4 distortion coefficients

In our second 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 camera center in the 3 views is not the same, we input 44 corresponding triplets, that lie on a real-world plane (see Figure 4). 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

Figure 5: Cubemap of undistorted left image (warping done, per pixel, onto a 2000x2000 image, using 5 distortion parameters)

error, was used to estimate the distortion parameters. A distortion model with 5 parameters was estimated and used to compute a undistorted image, for one of the views, using a cubemap projection (see Figure 5). 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.

8. Conclusion and Future Work

In this paper we have presented a stratified approach to recover the radial distortion of a camera observing a plane or undergoing pure rotation. In a first step, we linearly estimate the *radial trifocal tensor* from a minimum of seven correspondences across three views. This allows us to recover the projective structure of the plane and the radial camera matrices. From this point on several approaches could be used to recover the radial distortion. We propose a linear approach that can estimate any number of radial dis-

Figure 4: Three images, taken with different camera centers, input to the system (matching points input to the system are marked)

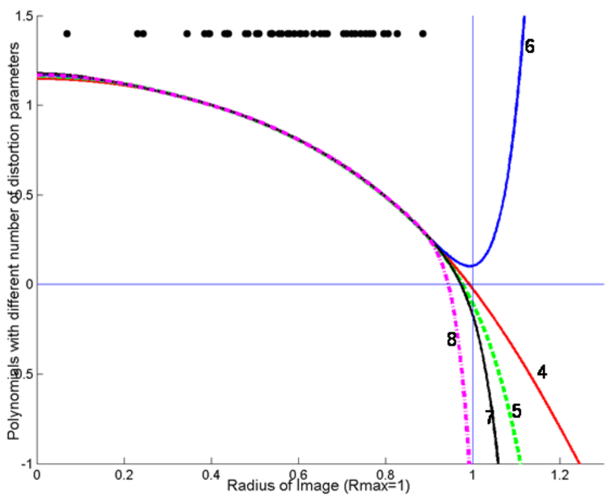

Figure 6: Distortion Curves $1 + K_1 r_d^2 + \cdots + K_n r_d^{2n}$ when different number parameters ($n = 4 - 8$, marked next to corresponding curve) are used. Note that most of the curves are well-behaved even at $r = R_{max}$.

tortion parameters of the division model. We have validated our approach using two real-world datasets. One of a fish-eye lens observing a plane and one of a wide-field of view camera undergoing pure rotation. We show that the results of our linear approach are very good.

In the future, we intend to investigate the possibility of using a similar multiple view relation between four views, i.e. the radial quadrifocal tensor, to calibrate omnidirectional cameras from images of a general 3D scene. We also intend to investigate more in depth the possibilities offered by the radial trifocal tensor for pure rotation as we believe a direct non-parametric estimation of the radial distortion should be possible.

Appendix

Consider Eq(5) in Section 4. Let the camera matrices be $\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \end{bmatrix}$ and so on (i.e, $\{\mathbf{P}_i\}_{i=1}^2$ is the i^{th} row of the corresponding 2×3 camera matrix). It can be written as,

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} det \begin{bmatrix} \mathbf{P}_{\sim i} \\ \mathbf{P}^{'}_{\sim j} \\ \mathbf{P}^{''}_{\sim k} \end{bmatrix} (-1)^{i+j+k+1} l_i l^{'}_j l^{''}_k = 0 \quad (17)$$

where $(\sim i)_{i=1} = 2$ and vice-versa.

Once we have evaluated $\mathbf{T}$, we can compute $\mathbf{S}$, which is a $2 \times 2 \times 2$ homogenous tensor, such that, $\mathbf{S}_{\sim i \sim j \sim k}(-1)^{i+j+k+1} = \mathbf{T}_{ijk}$. We then have $\mathbf{S}_{ijk} = det \begin{bmatrix} \mathbf{P}_i \\ \mathbf{P}^{'}_j \\ \mathbf{P}^{''}_k \end{bmatrix}$. We can then set up a projective basis, by choosing

$$\begin{bmatrix} \mathbf{P} \\ \mathbf{P}' \\ \mathbf{P}'' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ p & p & p \\ 0 & 1 & 0 \\ p_{21} & p_{22} & p_{23} \\ 0 & 0 & 1 \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{18}$$

Then, if we normalize $\mathbf{S}$ such that $\mathbf{S}_{111} = 1$, we can obtain $p = \mathbf{S}_{211}$, $p_{22} = \mathbf{S}_{121}$, $p_{33} = \mathbf{S}_{112}$, $\mathbf{S}_{221} = p(p_{22} - p_{21})$, $\mathbf{S}_{212} = (-p)(p_{33} - p_{31})$. We then have to evaluate p_{23} and p_{32} and have two equations, $\mathbf{S}_{122} = p_{22}p_{33} - p_{23}p_{32}$ and $\mathbf{S}_{222} = p(\mathbf{S}_{122} - (p_{21}p_{33} - p_{31}p_{23}) + p_{21}p_{32} - p_{31}p_{22})$. This allows us to solve for $\{p_{23}, p_{32}\}$ by solving a quadratic equation. And when we get two real unequal roots, we have a two-way ambiguity in the projective structure of the cameras.

References

[1] D.C. Brown. Close-range camera calibration. *PhEng*, 37(8):855–866, August 1971.

[2] F. Devernay and O.D. Faugeras. Straight lines have to be straight. *MVA*, 13(1):14–24, 2001.

[3] 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.

[4] A.W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR01*, pages I:125–132, 2001.

[5] A. Goshtasby. Correction of image deformation from lens distortion using bezier patches. *CVGIP*, 47(3):385–394, September 1989.

[6] R.I. Hartley and F. Schaffalitzky. Reconstruction from projections using grassmann tensors. In *ECCV04*, pages Vol I: 363–375, 2004.

[7] R.I. Hartley and A. Zisserman. Multiple view geometry in computer vision. In *Cambridge*, 2000.

[8] S.B. Kang. Radial distortion snakes. In *IAPR Workshop on Machine Vision Applications (MVA2000), Tokyo, Japan*, pages 603–606, 2000.

[9] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *CVPR03*, pages I: 485–490, 2003.

[10] L. Quan. Two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. *PAMI*, 23(2):212–216, February 2001.

[11] C. Slama, editor. *Manual of Photogrammetry.* American Society of Photogrammetry, Falls Church, VA, 4th edition edition, 1980.

[12] G.P. Stein. Lens distortion calibration using point correspondences. In *CVPR97*, pages 602–608, 1997.

[13] R. Swaminathan and S.K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *PAMI*, 22(10):1172–1178, October 2000.

[14] 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.

[15] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *PAMI*, 14(10):965–980, October 1992.