Fundamental Matrix for Cameras with Radial Distortion

Joao P. Barreto
Institute of Systems and Robotics
University of Coimbra
3030 Coimbra, Portugal
jpbar@deec.uc.pt

Kostas Daniilidis
GRASP Laboratory
University of Pennsylvania
Philadelphia, PA 19104
kostas@grasp.cis.upenn.edu

Abstract

When deploying a heterogeneous camera network or when we use cheap zoom cameras like in cell-phones, it is not practical, if not impossible to off-line calibrate the radial distortion of each camera using reference objects. It is rather desirable to have an automatic procedure without strong assumptions about the scene. In this paper, we present a new algorithm for estimating the epipolar geometry of two views where the two views can be radially distorted with different distortion factors. It is the first algorithm in the literature solving the case of different distortion in the left and right view linearly and without assuming the existence of lines in the scene. Points in the projective plane are lifted to a quadric in three-dimensional projective space. A radial distortion of the projective plane results to a matrix transformation in the space of lifted coordinates. The new epipolar constraint depends linearly on a 4x4 radial fundamental matrix which has 9 degrees of freedom. A complete algorithm is presented and tested on real imagery.

1 Introduction

The multiple view geometry on the context of camera self- calibration and structure from motion has been deeply studied and nowadays is subject for textbooks [5, 8]. However most of the work on the topic assumes that the camera follows the pin-hole model. This is not the case of many vision sensors that are broadly used in everyday applications. Some examples are the cameras with wide-angle lenses or motorized zoom where the lens radial distortion, caused by the bending of the light rays when crossing the optics, introduces a non-linearity in the image formation model.

Since these cameras deviate from the pin-hole model, the application of uncalibrated stereo usually requires a partial calibration to determine the non-linear relation between the image and the space coordinates. There are several techniques to recover the lens distortion. The first class of methods rely on calibration targets (e.g. [13]). The second type of algorithms use images of lines in order to explore the property that in a pin-hole lines in the scene must project onto straight lines in the image (e.g. [3]). Both categories require imagery with specific features (lines or calibration targets) and are not suitable to deal with generic images and video sequences of random scenes.

In an effort to extend the applicability of uncalibrated vision to cameras with radial distortion, some authors proposed methods that just rely on the rigidity assumptions that allow the computation of the fundamental matrix. In [15, 11] the rigidity constraint is extended to include the distortion parameters. The radial distortion is estimated using nothing more than image correspondences. The major drawback of these techniques is that the estimation relies on iterative non-linear methods. Such techniques require good initial estimates in order to guarantee a correct convergence. Moreover they are computationally too expensive to be included in a RANSAC strategy that automatically detects outliers in the correspondences.

Closer to the approach herein presented are the works of Fitzgibbon [6] and Micusik et al. [10]. Fitzgibbon proposes a linear method that simultaneously estimates the fundamental matrix and the radial distortion from corresponding points between two views. This is achieved by formulating the estimation problem as a quadratic-eigenvalue problem (QEP). In [10] the framework is extended for cameras equipped with fish-eye lenses. *The problem of these algorithms is that they require the distortion to be the same on both views*. Such requirement considerably limits the range of applications that can benefit from the approach. One of the examples that is presented in [6] is the 3D reconstruction from a video sequence acquired by a commercial handheld camera. In general these cameras are equipped with a motorized lens and, during normal operation, the amount of radial distortion varies. This is a situation for which the method proposed in [6] is not applicable. Another example is the automatic calibration of wide area camera networks

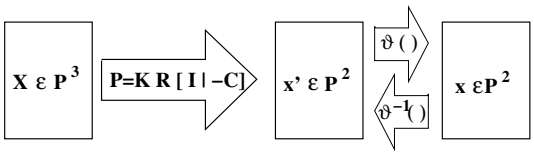

Figure 1. Image formation model on dioptric cameras with radial distortion

from natural features. Typically the cameras employed in this infrastructures are equipped with wide-angle lenses. The presence of significant radial distortion avoids the use of conventional techniques to estimate the fundamental matrix. Since the radial distortion is different for each camera, the QEP approach does not solve the problem either.

The present work overcomes these limitations by introducing the radial fundamental matrix (RFM). We establish for the first time a bilinear relation between two views acquired by cameras with different lens distortion. This is achieved by embedding the projective plane into the 3D projective space. The RFM is a 4×4 matrix that relates the lifted coordinates of corresponding image points and encodes both the standard fundamental matrix (SFM) and the distortion parameters of the two views. Like the SFM, the RFM is a projective correlation that maps lifted points in one view into epipolar curves in the other view. The main contributions can be summarized as follows:

- The Radial Fundamental Matrix that relates views with equal or different radial distortion.
- A linear technique to estimate the RFM using a minimum of 15 image correspondences.
- A RANSAC strategy, based on the linear estimator, that is able to automatically detect outliers on the correspondences between views.

2 Modeling Lens Radial Distortion

Fig. 1 is a scheme of the assumed image formation model. In the pin-hole model to each 3D point $\mathbf{X}$ corresponds an image point $\mathbf{x}' = \mathbf{PX}$. The projection matrix $\mathbf{P}$ encodes the relative pose between the world and camera coordinate systems as well as the matrix of intrinsic parameters. Points in the scene are mapped into points in the image by a linear transformation. In perspective cameras with lens distortion this mapping can no longer be described in a linear way. In this paper the radial distortion is modeled using the so called division model. Point $\mathbf{X}$ is projected into $\mathbf{x}$ in the image plane, which is related with $\mathbf{x}'$ by a non-linear transformation modeling the radial distortion. This transformation, originally introduced in [2, 6], is provided in equation 1 where ξ is a parameter quantifying the amount of distortion. If $\xi = 0$ then $\mathbf{x} = \mathbf{x}'$ and the camera follows the conventional pin-hole model.

$$\eth^{-1}(\mathbf{x}) = (xz, yz, z^2 + \xi(x^2 + y^2))^t \tag{1}$$

The model of equation 1 requires that points $\mathbf{x}'$ and $\mathbf{x}''$ are referenced in a coordinate system with origin in the image distortion center. If the distortion center is not known in advance, we can place it at the image center without significantly affecting the correction [14].

2.1 Lines Projected into Circles

Consider a 3D line lying on a plane $\mathbf{\Pi}$ that goes through the camera center $\mathbf{O}$. According to the pin-hole model the 3D line is projected into a line $\mathbf{n}' = (n'_x, n'_y, n'_z)^t$ such that $\mathbf{\Pi} = \mathbf{P}^t\mathbf{n}'$ [8]. In the presence of lens distortion the points $\mathbf{x}'$ on $\mathbf{n}'$ are mapped into points $\mathbf{x} = \eth(\mathbf{x}')$ (Fig. 1), and the straight line $\mathbf{n}'$ is imaged into a curve. Since $\mathbf{n'}^t\mathbf{x}' = 0$ and $\mathbf{x}' = \eth^{-1}(\mathbf{x})$, then the curve equation is $\mathbf{n'}^t\eth^{-1}(\mathbf{x}) = 0$. Replacing $\eth^{-1}(\mathbf{x})$ by the result of equation 1 yields

$$\xi n'_z(x^2 + y^2) + n'_x xz + n'_y yz + n'_z z^2 = 0. \tag{2}$$

The curve of equation 2 is a circle. Thus, we can state that in general a line in the scene is imaged into a circle when the radial distortion is modeled by the transformation of equation 1. The circle degenerates into a line whenever $\xi = 0$ or $n'_z = 0$. In the former situation the camera follows the pin-hole model. In the latter the line is not bent by the lens distortion because it goes through the image center. The formulas for the circle center and radius are provided in equation 3.

$$\begin{cases} \mathbf{C_n} = (-\frac{n'_x}{2\xi n'_z}, -\frac{n'_y}{2\xi n'_z}) \\ r_n = \frac{1}{2}\sqrt{\frac{{n'_y}^2 + {n'_x}^2 - 4\xi {n'_z}^2}{\xi^2 {n'_z}^2}} \end{cases} \tag{3}$$

The general equation of a circle is $a(x^2 + y^2) + dxz + eyz + fz^2 = 0$ where a, d, e and f are 4 homogeneous parameters. The circle is uniquely represented, up to a scale factor, by a point $\tilde{\boldsymbol{\omega}} = (a, d, e, f)^t$ in the dual 3D projective space $\wp^{\mathbf{3*}}$. From equation 2 follows that if $\tilde{\boldsymbol{\omega}}$ is the distorted image of a line $\mathbf{n}'$ then

$$\mathbf{n}' = \begin{bmatrix} n'_x \\ n'_y \\ n'_z \end{bmatrix} \longrightarrow \tilde{\boldsymbol{\omega}} = \begin{bmatrix} \xi n'_z \\ n'_x \\ n'_y \\ n'_z \end{bmatrix} \tag{4}$$

The representation of circles by points in $\wp^{\mathbf{3*}}$ allows us to establish a linear correspondence between the line images before distortion ($\mathbf{n}'$) and after distortion ($\tilde{\boldsymbol{\omega}}$)

$$\widetilde{\boldsymbol{\omega}} = \underbrace{\begin{bmatrix} 0 & 0 & \xi \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{\Delta_\xi}} \mathbf{n}' \tag{5}$$

We proved that in a camera with radial distortion $\xi \neq 0$ the image of a line is in general a circle. However not all circles are the projection of a line. The circle $\tilde{\boldsymbol{\omega}} = (a, d, e, f)^t$ is a line image *iff* $a = \xi f$. Considering the representation in $\boldsymbol{\wp^{3*}}$, the points $\widetilde{\omega}$ which correspond to line projections must lie on a 2 dimensional linear subspace.

2.2 Embedding the Projective Plane into $\boldsymbol{\wp^3}$

The goal of this work is to establish a bilinear relation between corresponding points on two views captured by cameras with lens distortion. If both cameras follow the pinhole model then the 3D point $\mathbf{X}$ is respectively projected on points $\mathbf{x}'$ and $\mathbf{y}'$. The corresponding image points satisfy the relation $\mathbf{y'}^t\mathbf{F'x'} = 0$ where $\mathbf{F}'$ is the fundamental matrix between the two views. However, in the presence of radial distortion, the image points become $\mathbf{x}$ and $\mathbf{y}$ (see Fig. 1), and the previous bilinear equation gives place to the non-linear relation $(\eth^{-1}(\mathbf{y}))^t\mathbf{F}'(\eth^{-1}(\mathbf{x})) = 0$.

A standard technique used in algebra to render a non-linear problem into a linear one is to find an embedding that lifts the problem into a higher dimensional space. Such technique is applied in [1] to derive the necessary and sufficient conditions for a conic to be the central catadioptric projection of a line, and in [7, 12] to establish fundamental matrices between views taken by any mixture of paracatadioptric, perspective or affine cameras. In order to establish a radial fundamental matrix we propose the embedding of the projective plane $\boldsymbol{\wp^2}$ into the 3D projective space $\boldsymbol{\wp^3}$. Our polynomial embedding preserves homogeneity and creates a natural duality between the lifted image points and the circles where the lines are projected.

Consider the image point $\mathbf{x}$ which is the distorted projection of $\mathbf{x}'$ ($\mathbf{x} = \eth(\mathbf{x}')$. Each point $\mathbf{x}$ in $\boldsymbol{\wp^2}$ is mapped into a point $\widetilde{\mathbf{x}} = (x_0, x_1, x_2, x_3)^t$ in $\boldsymbol{\wp^3}$ such that

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \longrightarrow \tilde{\mathbf{x}} = \begin{bmatrix} x^2 + y^2 \\ xz \\ yz \\ z^2 \end{bmatrix}. \tag{6}$$

Remark that the projective plane is mapped into a second order surface in $\boldsymbol{\wp^3}$. This surface is defined by equation $x_0x_3 - x_1^2 - x_2^2 = 0$. Therefore not all points in $\boldsymbol{\wp^3}$ have a correspondence in the projective plane.

If point $\mathbf{x}$ lies on circle $\widetilde{\boldsymbol{\omega}}$ then its lifted representation $\widetilde{\mathbf{x}}$ satisfies the relation $\widetilde{\boldsymbol{\omega}}^t\widetilde{\mathbf{x}} = 0$. The embedding of equation 6 creates an algebraic duality between points and lines in the distorted image plane. Moreover since circle $\widetilde{\boldsymbol{\omega}}$ is the distorted projection of line $\mathbf{n}'$, then $\widetilde{\boldsymbol{\omega}} = \boldsymbol{\Delta_\xi}\mathbf{n}'$ (equation 5) and $\widetilde{\boldsymbol{\omega}}^t\widetilde{\mathbf{x}} = 0$ gives rise to

$$\mathbf{n'}^t \underbrace{\boldsymbol{\Delta}_{\boldsymbol{\xi}}^{\boldsymbol{t}}\widetilde{\mathbf{x}}}_{\mathbf{x}'} = 0.$$

Taking into account that in the pin-hole image $\mathbf{n'}^t\mathbf{x}' = 0$ it follows that

$$\mathbf{x}' = \boldsymbol{\Delta}_{\boldsymbol{\xi}}^{\boldsymbol{t}}\widetilde{\mathbf{x}} \tag{7}$$

The proposed embedding linearizes the division model for radial distortion (equation 1). The non-linear mapping $\mathbf{x}' = \eth^{-1}(\mathbf{x})$ in $\boldsymbol{\wp^2}$ gives place to the linear transformation of equation 7.

3 Radial Fundamental Matrix

Consider two images acquired by two cameras with projection centers $\mathbf{O_x}$ and $\mathbf{O_y}$. If $\mathbf{x}' \leftrightarrow \mathbf{y}'$ are corresponding points on the two views then, in the absence of radial distortion, there is a 3×3 fundamental matrix $\mathbf{F}'$ such that

$$\mathbf{y'}^t\mathbf{F'x'} = 0.$$

Assume that both lenses have significant radial distortion which is described by parameters ξ_x and ξ_y. In this case the pair of corresponding points becomes $\mathbf{x} \leftrightarrow \mathbf{y}$ (equation 1), and is represented in the lifted space by $\widetilde{\mathbf{x}} \leftrightarrow \widetilde{\mathbf{y}}$ (equation 6). Since $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{y}}$ are linearly related with $\mathbf{x}'$ and $\mathbf{y}'$ (equation 7) then the fundamental relation can be rewritten as follows

$$\widetilde{\mathbf{y}}^t \underbrace{\boldsymbol{\Delta_y} \overbrace{\begin{bmatrix} f'_{11} & f'_{12} & f'_{13} \\ f'_{21} & f'_{22} & f'_{23} \\ f'_{31} & f'_{32} & f'_{33} \end{bmatrix}}^{\mathbf{F}'} \boldsymbol{\Delta}_{\boldsymbol{x}}^{\boldsymbol{t}}}_{\mathbf{F}} \widetilde{\mathbf{x}} = 0 \tag{8}$$

By considering the lifted representation of section 2.2 we established a bilinear relation between corresponding points on two distorted views. This bilinear relation is described by a 4×4 matrix which we will refer as the radial fundamental matrix (RFM). Matrix $\mathbf{F}$ encodes the conventional fundamental matrix $\mathbf{F}'$ and the distortion parameters ξ_x and ξ_y.

$$\mathbf{F} = \begin{bmatrix} \xi_x\xi_y f'_{33} & \xi_y f'_{31} & \xi_y f'_{32} & \xi_y f'_{33} \\ \xi_x f'_{13} & f'_{11} & f'_{12} & f'_{13} \\ \xi_x f'_{23} & f'_{21} & f'_{22} & f'_{23} \\ \xi_x f'_{33} & f'_{31} & f'_{32} & f'_{33} \end{bmatrix} \tag{9}$$

Matrix $\mathbf{F}$ has similarities with the conventional fundamental matrix $\mathbf{F}'$. While $\mathbf{F}'$ is a correlation that maps points

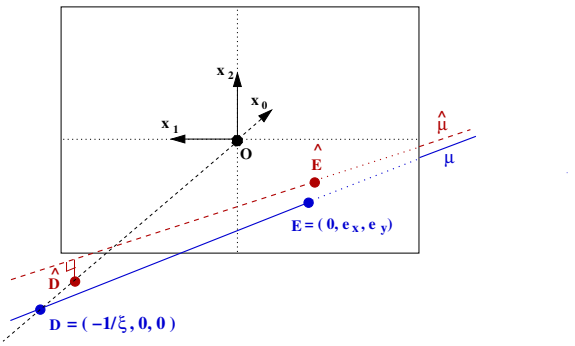

Figure 2. Line μ represents the left/right null space of $\mathbf{F}$. The line goes through points $\mathbf{E}$ and $\mathbf{D}$. The former encodes the epipole on the undistorted image plane and the latter encodes the distortion parameter

on one view into epipolar lines on the other view, the radial fundamental matrix transforms lifted points into epipolar circles ($\widetilde{\boldsymbol{\omega_y}} = \mathbf{F}\widetilde{\mathbf{x}}$ and $\widetilde{\boldsymbol{\omega_x}} = \mathbf{F^t}\widetilde{\mathbf{y}}$). Moreover matrix $\mathbf{F}$ is also rank deficient. Since $\mathbf{F}'$ has rank 2 and matrices $\boldsymbol{\Delta_x}$ and $\boldsymbol{\Delta_y}$ have rank 3, then the rank of the RFM is always 2.

Equation 9 shows the structure of the 4×4 matrix $\mathbf{F}$. The RFM has 9 degrees of freedom (DOF): the 7 DOF of $\mathbf{F}'$ plus the 2 DOF associated with the distortion parameters ξ_x and ξ_y. However matrix $\mathbf{F}$ has a total of 16 parameters f_{ij} with $i = 1,\dots 4$ and $j = 1,\dots 4$. Therefore it must satisfy the 6 independent constraints of equation 10 (the scale factor is ignored). The first constraint is the well known rank constraint of $\mathbf{F}'$. The remaining 5 bilinear constraints are due to the linear dependence between the firsts and fourths lines and columns. If the radial distortion on both views is the same then $\xi_x = \xi_y$ and matrix $\mathbf{F}$ must verify the additional constraint $f_{14} - f_{41} = 0$.

$$\begin{cases} det(\mathbf{F_{2\dots4,2\dots4}}) = 0 \\ f_{11}f_{i4} - f_{14}f_{i1} = 0, \quad i = 2,3,4 \\ f_{11}f_{4j} - f_{41}f_{1j} = 0, \quad j = 2,3 \end{cases} \tag{10}$$

Henceforth we will be considering the right null space $\aleph(\mathbf{F})$ of the RFM. However the reasoning and statements are equally valid for the left null space. Since $\mathbf{F}$ is a rank 2 matrix then $\aleph(\mathbf{F})$ is bi-dimensional and can be represented by a line $\boldsymbol{\mu}$ on the lifted space $\wp^{\mathbf{3}}$. As shown on Fig. 2, line $\boldsymbol{\mu}$ intersects the plane X_1OX_2 on a point $\mathbf{E} = (0, \mathbf{e}'^t)^t$. From equation 9 follows that $\mathbf{e}' = (e'_x, e'_y, 1)^t$ is the null space of $\mathbf{F}'$ and represents the epipole on the undistorted image plane. Moreover if ξ is the distortion parameter then $\mathbf{D} = (-\xi^{-1}, 0, 0, 1)^t$ satisfies $\mathbf{F}.\mathbf{D} = 0$. Point $\mathbf{D}$ also lies on line $\boldsymbol{\mu}$ and is the intersection of $\boldsymbol{\mu}$ with the OX_0 axis. These observations will prove to be useful.

4 Estimating the Fundamental Matrix from Image Correspondences

Consider a set of K corresponding points $\mathbf{x_k} \leftrightarrow \mathbf{y_k}$ on two views with radial distortion parameters ξ_x and ξ_y ($K \geq 15$). The radial fundamental matrix $\mathbf{F}$ can be directly estimated from the lifted correspondences $\widetilde{\mathbf{x_k}} \leftrightarrow \widetilde{\mathbf{y_k}}$ ($k = 1,2\dots K$). The bilinear relation of equation 8 gives rise to the set of equations of the form

$$\underbrace{\begin{bmatrix} \widetilde{\mathbf{x_1}}^t \otimes \widetilde{\mathbf{y_1}}^t \\ \widetilde{\mathbf{x_2}}^t \otimes \widetilde{\mathbf{y_2}}^t \\ \vdots \\ \widetilde{\mathbf{x_K}}^t \otimes \widetilde{\mathbf{y_K}}^t \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{44} \end{bmatrix}}_{\mathbf{f}} = 0 \tag{11}$$

If the correspondences are noiseless then $\mathbf{A}$ has rank 15 and it is possible to solve for $\mathbf{F}$ up to a scale factor. In the presence of noise the radial fundamental matrix can be estimated by minimizing the objective function $\hbar(\mathbf{f}) = \mathbf{f}^t\mathbf{A}^t\mathbf{A}\mathbf{f}$. However, and since $\mathbf{F}$ has to satisfy the constraints of equation 10, the minimum can only be found using non-linear optimization methods. This is similar to what happens with the conventional fundamental matrix and the rank constraint.

In this section we propose a linear method to compute an initial estimate for the radial fundamental matrix. We will show that the estimated $\mathbf{F}$ is accurate enough for many applications. The method can also be used on a RANSAC based approach in order to detect outliers in the correspondences. Our 15-point algorithm is for the radial fundamental matrix as the 8-point algorithm is for the conventional fundamental matrix.

4.1 The 15-Point Algorithm

Tab. 1 summarizes the 15-point algorithm. The 3 first steps resemble the steps of the well known 8-point algorithm. In order to avoid statistical bias a proper careful normalization of the lifted correspondences is required [8, 5]. Then we find a least square solution for $\mathbf{Af} = 0$ and enforce the rank constraint using singular value decomposition. The problem is that in general the obtained 4×4 matrix $\hat{\mathbf{F}}$ is not a radial fundamental matrix. Despite of having rank 2, matrix $\hat{\mathbf{F}}$ does not satisfy the bilinear constraints shown in equation 10. The goal of steps 4 to 8 is to find a matrix with the correct structure (equation 9).

As discussed on the previous section the right null space of $\hat{\mathbf{F}}$ has dimension 2 and corresponds to a line $\hat{\boldsymbol{\mu}}$ in $\wp^{\mathbf{3}}$ (see Fig. 2). In an ideal situation of noise free correspondences the lines $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}$ would be coincident. Unfortunately this is not the case when dealing with real images. However, for moderate levels of noise, we expect that $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}$ are close

Step 1	Build matrix $\mathbf{A}$ (equation 11) using the normalized lifted correspondences $\widetilde{\mathbf{x_k}} \leftrightarrow \widetilde{\mathbf{y_k}}$. The normalizing transformations are $\mathbf{T_x}$ and $\mathbf{T_y}$ consisting of a translation and scaling.
Step 2	Determine the 4×4 matrix $\dot{\mathbf{F}}$ from the singular vector corresponding to the smallest singular value of $\mathbf{A}$ (total least squares)
Step 3	Replace $\dot{\mathbf{F}}$ by $\hat{\mathbf{F}}$, the closest rank 2 matrix to $\dot{\mathbf{F}}$ under a Frobenius norm. The correction is performed using SVD
Step 4	Compute the line $\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}(\hat{\boldsymbol{\mu}}_{\boldsymbol{y}})$ corresponding to the left(right) null space of $\hat{\mathbf{F}}$.
Step 5	Estimate the undistorted epipole $\mathbf{e}'_{\mathbf{x}}(\mathbf{e}'_{\mathbf{y}})$ by intersecting line $\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}(\hat{\boldsymbol{\mu}}_{\boldsymbol{y}})$ with plane X_1OX_2 (Fig. 2)
Step 6	Estimate the distortion parameter $\xi_x(\xi_y)$ by minimizing the geometric distance between $\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}(\hat{\boldsymbol{\mu}}_{\boldsymbol{y}})$ and OX_0 (Fig. 2 and equation 12).
Step 7	Parameterize the radial fundamental matrix in terms of $\mathbf{e}'_{\mathbf{x}}$ and $\mathbf{e}'_{\mathbf{y}}$ and compute the least square solution for the epipolar collineation [8, 5] (ξ_x, ξ_y, $\mathbf{e}'_{\mathbf{x}}$ and $\mathbf{e}'_{\mathbf{y}}$ are kept constant)
Step 8	Knowing the distortion parameters, the epipoles and the epipolar collineation build matrix $\ddot{\mathbf{F}}$ which will satisfy the constraints of equation 10.
Step 9	Denormalize the result in order to obtain the radial fundamental matrix ($\mathbf{F} = \mathbf{T}_{\mathbf{y}}^{\mathbf{t}}\ddot{\mathbf{F}}\mathbf{T}_{\mathbf{x}}$)

Table 1. The 15-point algorithm

enough, and that both the epipole and the distortion parameter can be estimated with reasonable accuracy. Thus, the epipole $\mathbf{e}'_{\mathbf{x}}$ before distortion is determined by intersecting $\hat{\boldsymbol{\mu}}$ with X_1OX_2 (point $\hat{\mathbf{E}}$ in Fig. 2). Since in general lines $\hat{\boldsymbol{\mu}}$ and OX_0 are not concurrent, the distortion parameter ξ_x is computed using point $\hat{\mathbf{D}}$. This is the point on the X_0 axis which is geometrically closer to $\hat{\boldsymbol{\mu}}$. Point $\hat{\mathbf{D}}$ can be easily determined by applying the closed form method presented in [4]. We have used this method to derive equation 12 which provides an analytical solution for the distortion parameter. Vector $\mathbf{u} = (u_0, u_1, u_2)^t$ is any 3D vector parallel to $\hat{\boldsymbol{\mu}}$ and $\mathbf{Q} = (q_0, q_1, q_2)^t$ is any point lying in the line. This same procedure is repeated for the left null space in order to estimate ξ_y and $\mathbf{e}'_{\mathbf{y}}$.

$$\xi = \left(\frac{u_0u_1q_1 + u_0u_2q_2}{u_1^2 + u_2^2} - q_0\right)^{-1} \tag{12}$$

Consider the epipolar parameterization of the fundamental matrix as described in [8, 5]. The radial fundamental matrix is written as a function of the distortion parameters (ξ_x and ξ_y), the epipoles on the undistorted views ($\mathbf{e}'_{\mathbf{x}}$ and $\mathbf{e}'_{\mathbf{y}}$) and the epipolar collineation. Since ξ_x, ξ_y, $\mathbf{e}'_{\mathbf{x}}$ and $\mathbf{e}'_{\mathbf{y}}$ have already been determined we just need to estimate the epipolar collineation that better fits the correspondences. We can find a least square solution for the problem by applying SVD. The resultant matrix satisfies the constraints of equation 10 and has the right structure to be a RFM.

4.2 Performance Evaluation

In this section synthetic images are used to evaluate the performance of the 15-point algorithm. Camera X and Y, with different radial distortion parameters, are placed in the virtual working space. The 480×640 synthetic images are generated by projecting in both views a set of NC points. These 3D points are randomly selected by uniformly sampling the volume. Two-dimensional gaussian noise with zero mean and standard deviation σ is added to each image point. The radial fundamental matrix is estimated from the noisy image correspondences using the 15-point algorithm summarized in Tab. 1. The estimation results are compared with the ground truth and the root mean square (RMS) error is computed over 200 runs of each experiment.

Before discussing the results exhibited in Fig. 3 we would like to make a couple of remarks. The first one is that we are showing the RMS error which is more penalizing that the mean error. The difference between these two metrics is specially noticeable when the results present significant standard deviation. The second remark is that the case that we are analyzing is particularly hard because the amounts of radial distortion on the two views are unbalanced. The distortion in camera X causes a deviation of 85 pixels at the image corner while in camera Y the shift along the radial direction is only 15 pixels.

The experiment of Fig. 3 evaluates the performance for an increasing number of correspondences (NC= 50, 150 and 300). For comparison purposes we estimate both the radial and the conventional fundamental matrices. The former is determined with the 15-point algorithm. The latter is computed with the 8-point algorithm after correcting the image distortion (which is assumed to be known). The graphics show the geometric distance between the data points and the estimated epipolar circles, the error in determining the epipole, and the relative error on the distortion estimation.

The 15-point algorithm performs poorly when the number of correspondences is small (NC=50). However when NC increases the performance rapidly converges to the one of the 8-point algorithm. This is due to the fact that the number of correspondences is directly proportional to the probability of having points in the image periphery. Without points in the periphery the algorithm is unable to determine the distortion and correctly estimate the RFM. The estimation of radial distortion also performs reasonably well. According to the graphic of Fig. 3 for $\sigma = 2$ and NC=150

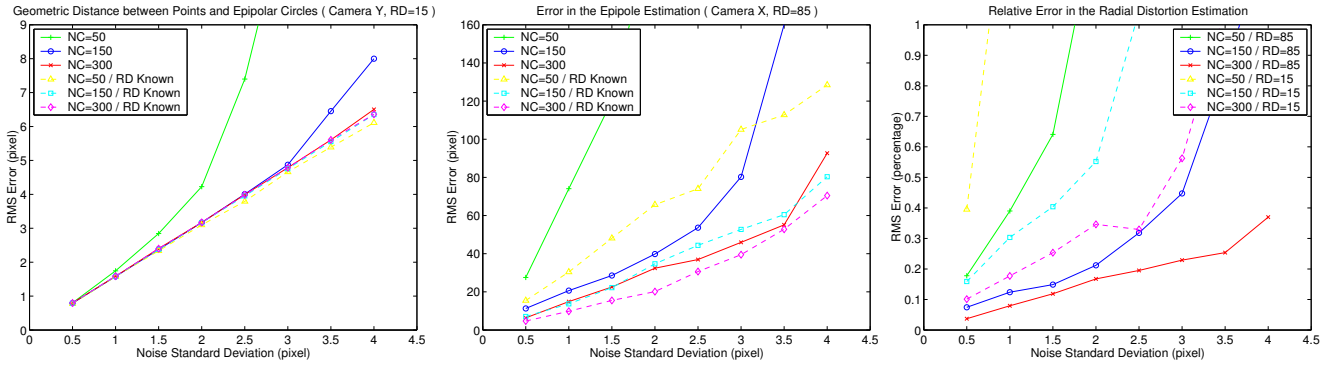

Figure 3. Performance evaluation using synthetic generated views with different amounts of distortion. The distortion of view X is high (85 pixels of deviation at the image corner) while the radial distortion of view Y is almost negligible (15 pixels). The graphics show the performance of the 15-point algorithm in estimating the radial fundamental matrix from 50, 150 and 300 correspondences.

the relative errors for the distortion parameters ξ_x and ξ_y are 18% and 55%. These values correspond to an error of 15.3 and 8.25 pixels on the deviation at the image corner. If we take into account that in a 480×640 image the deviations up to 15 pixels are barely noticeable, we can state that the estimation is accurate enough for many applications.

4.3 Automatic Computation of F

Until now, we assumed that the point correspondences were given. The matches between two captured images can be obtained using different techniques [9, 5, 8]. Typically interest points are computed on both views, and each point on the first view is matched to each point on the second view accordingly to a certain correlation metric. The pairs with a correlation score above a certain threshold are considered valid image correspondences. However, some of these pairs are incorrect, and including them on the fundamental matrix estimation severely degrades the accuracy.

In general the input to the estimation method is a set of noisy image correspondences $\mathbf{x} \leftrightarrow \mathbf{y}$, some of which are incorrect. An automatic algorithm must be able to detect these outliers and exclude them from the estimation process. To achieve this goal we propose a RANSAC strategy similar to the one used for the conventional fundamental matrix. At each iteration 15 correspondences are randomly selected and $\mathbf{F}$ is estimated following the linear method outlined in Tab. 1. A test is used to mark each correspondence as inlier or outlier. The procedure is repeated for a certain number of times and the final estimation of $\mathbf{F}$ is obtained from the largest set of inliers.

The most important component of such an algorithm is the test used to mark each correspondence as inlier or outlier. In the pin-hole case the classification is typically performed using the distance from points to epipolar lines. However, in the case of cameras with radial distortion, this distance becomes the distance from points to epipolar curves. In order to correctly perform the computation Fitzgibbon suggests the correction of the radial distortion of the noisy image points [6]. In section 2.1 we proved that lines are projected into circles. This means that the epipolar curves are in fact epipolar circles which greatly simplifies the problem. The geometric distance from points to circles can be easily determined by subtracting the circle radius (r_n) to the distance from the point to the circle center ($d(\mathbf{x}, \mathbf{C_n})$). Therefore we will classify the correspondences as inliers or outliers using the geometric distances ϵ from points to epipolar circles. The centers ($\mathbf{C_x}$, $\mathbf{C_y}$) and radius (r_x, r_y) are calculated using the formulas of equation 3.

$$\epsilon = \| \, d(\mathbf{x}, \mathbf{C_x}) - r_x \, \|^2 + \| \, d(\mathbf{y}, \mathbf{C_y}) - r_y \, \|^2 \qquad (13)$$

5 Experiments

Fig. 4 shows an experiment where the 15-point algorithm is used to estimate the radial fundamental matrix between two images acquired by different cameras. The radial distortion in the first view is high, while in the second view is just slightly noticeable. The correspondences between images were automatically established using SIFT features [9] and the outliers were detected after 500 iterations of RANSAC. The Lowe detector found 408 correspondences from which 269 were marked as inliers (34% of rejection). Since we model both the epipolar geometry and the lenses radial distortion, the inliers are spread all over the image in-

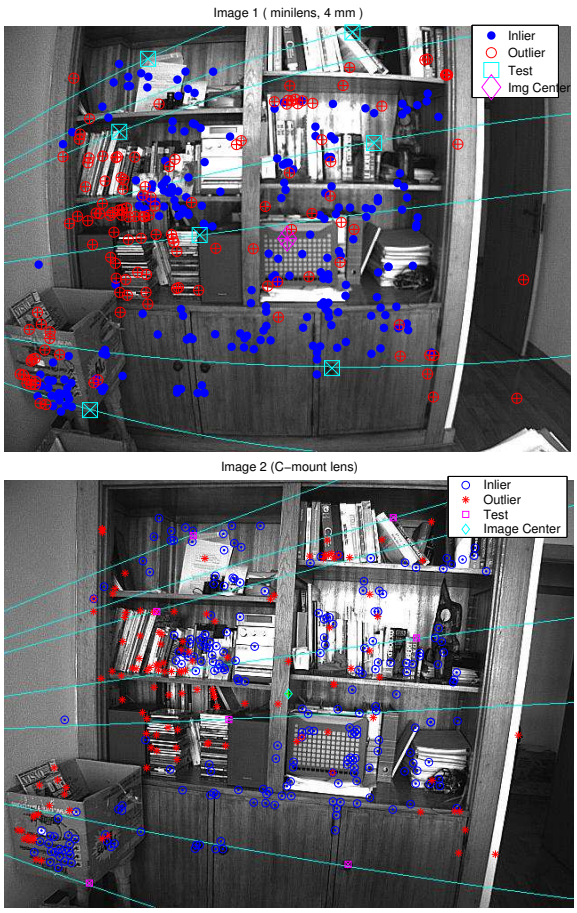

Figure 4. Estimation of the Radial Fundamental Matrix between two 480×640 images acquired by different cameras.

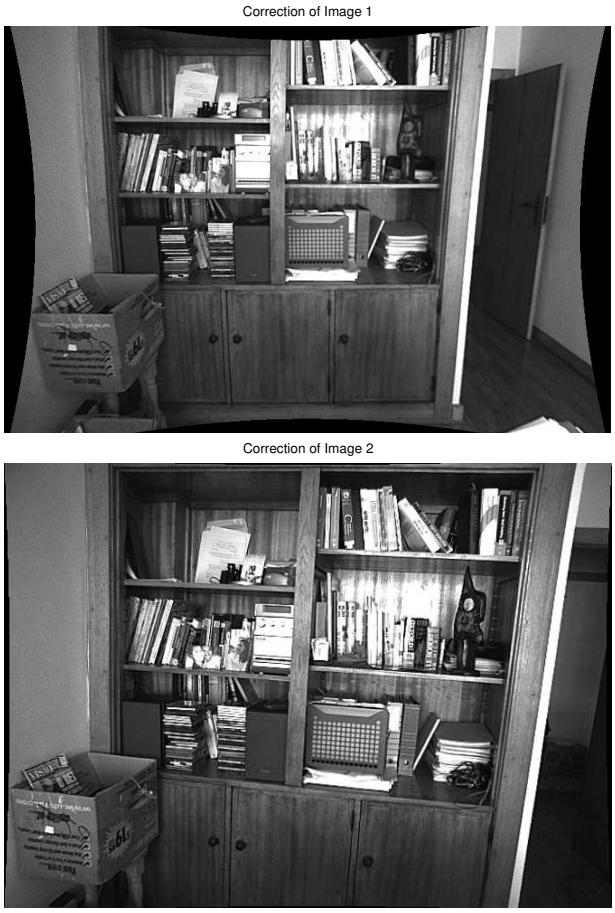

Figure 5. Correction of Radial Distortion. The distortion was estimated in 125.4 pixels for camera 1 and 22.9 pixels for camera 2

stead of being only in the image center [6]. The deviation at the image corner due to the radial distortion was estimated in 125.4 pixels for camera 1 and 22.9 pixels for camera 2 (the corrected images are presented in Fig. 5). In order to test the correctness of the radial fundamental matrix we manually selected 7 point pairs and drew the corresponding epipolar circles (light blue curves). The experiment was repeated for 5 other image pairs acquired with the same cameras. The standard deviation in the distortion estimation was 13.3 pixels for camera 1 and 7.1 pixels for camera 2.

The 288×352 images shown in Fig. 6 were acquired using a cell-phone with a camera. In order to cope with the low resolution and bad quality of the images, we used a modified version of the 15 point algorithm. By assuming equal radial distortion in both views, the lines $\boldsymbol{\mu_x}$ and $\boldsymbol{\mu_y}$, corresponding to the left and right null spaces of $\mathbf{F}$, must intersect OX_0 in the same point (see Fig. 2). This is an additional constraint which helps to improve the estimation. We also increased the threshold for the detection of inliers. The image in the top of Fig. 6 is a linear combination of the two views used in the computation of the radial fundamental matrix. The Lowe detector automatically found 68 correspondences from which 9 were considered outliers (red arrows). According to the estimation results the distortion causes a maximum shift of 19 pixels along the radial direction. This value was used to correct the image at the left-bottom corner of Fig. 6

Remark that in both cases the radial fundamental matrix (RFM) was estimated by only applying the 15-point algorithm. Like the 8-point algorithm used for the standard fundamental matrix, the 15-point algorithm is a linear approach that provides a good initial estimate. Nevertheless the estimation is sub-optimal in terms of error minimization, and the accuracy can always be improved by a subsequent step of non-linear optimization.

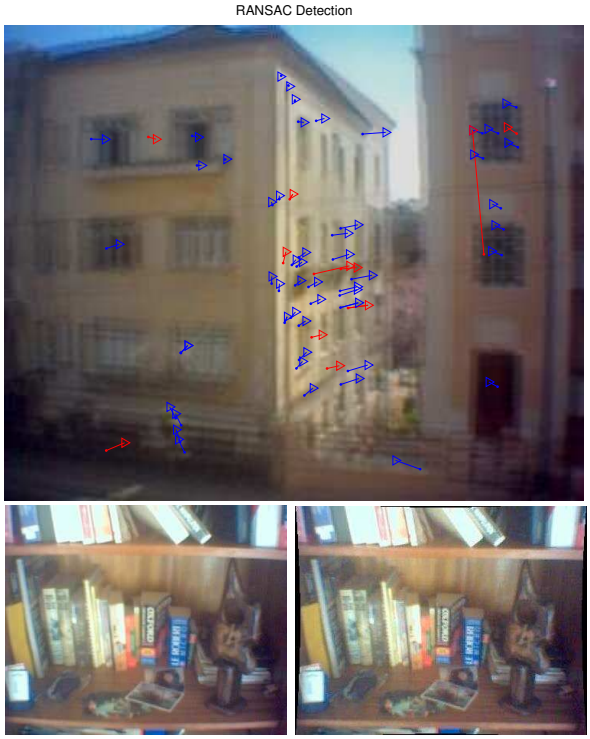

Figure 6. Estimation of the Radial Fundamental Matrix between two views acquired with a Sony-Ericsson cell phone. The distortion estimation is used to correct the image at the left-bottom corner.

6 Conclusions

The fundamental idea of this paper is that by lifting the projective plane to a quadric in the 3D projective space, the radial distortion is expressed as a matrix transformation in the lifted space. An epipolar constraint can be written as a line equation which corresponds to a circle equation in the distorted image plane. Therefore the epipolar constraint in the lifted space can be represented by a new 4x4 matrix that we called radial fundamental matrix. This is very good news since any attempt till now had yielded in the best case a quadratic eigenvalue problem. But even more important is the capability to accommodate different radial distortion coefficients in the left and right image. Because of its simplicity, it is very convenient for use in a RANSAC algorithm.

We would like to point out to the reader, who might think "yet another radial distortion or calibration paper...", the significance of the problem. Camera networks available at minimal cost nowadays are resulting in configurations with only two-view overlap and recording scenes which might not have a single reference landmark or a linear structure. It is also necessary that such networks are deployed rapidly with minimal preparation. It is for such cases that our algorithm shows its potential because it relies only on correspondences without any further prior knowledge on the camera or the scene.

Acknowledgments

The authors acknowledge the generous funding supplied by the Luso-American Foundation for Development. They are also grateful for support through the following grants: NSF-IIS-0083209, NSF-IIS-0121293, NSF-EIA-0324977 and ARO/MURI DAAD19-02-1-0383.

References

[1] J. Barreto and K. Daniilidis. Unifying image plane liftings for central catadioptric and dioptric cameras. *IEEE Int. Workshop on Omnidirectional Vision and Camera Networks*, May 2004.

[2] C. Brauer-Burchardt and K. Voss. A new algorithm to correct fish-eye- and strong wide-angle-lens-distortion from single images. *Int. Conf. on Image Processing*, pages 225–228, October 2001.

[3] D. C. Brown. Decentering distortion of lens. *Photogrametric Engineering*, 1966.

[4] D. Eberly. *3D Game Engine Design (Distance Methods)*. Morgan Kaufman, 2000.

[5] O. Faugeras and Q. Luong. *The Geometry of Multiple Images*. The MIT Press, 2001.

[6] A. Fitzgibbon. Simultaneous linear estimation of multiple-view geometry and lens distortion. *Int. Conf. on Computer Vision and Pattern Recognition*, December 2001.

[7] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. *Int. Conf. on Computer Vision and Pattern Recognition*, December 2001.

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

[9] D. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Journal on Computer Vision*, 2004.

[10] B. Micusik and T. Padjla. Estimation of omnidirectional camera model from epipolar geometry. *Int. Conf. on Computer Vision and Pattern Recognition*, 2003.

[11] G. P. Stein. Lens distortion calibration using point correspondences. *Int. Conf. on Computer Vision and Pattern Recognition*, 1997.

[12] P. Sturm. Mixing catadioptric and perspective cameras. *IEEE Int. Workshop on Omnidirectional Vision and Camera Networks*, May 2002.

[13] R. Y. Tsai. A versatile camera calibration technique for high accuracy 3d machine vision metrology using off the shelf tv cameras and lenses. *Robotics and Automation*, 1987.

[14] R. Willson and S. Shaffer. What is the center of the image. *Int. Conf. on Computer Vision and Pattern Recognition*, 1993.

[15] Z. Zhang. On the epipolar geometry between two images with lens distortion. *Int. Conf. on Pattern Recognition*, 1996.