Parameter-free Radial Distortion Correction with Centre of Distortion Estimation

Richard I. Hartley*
Australian National University
and National ICT Australia†
Richard.Hartley@anu.edu.au

Sing Bing Kang
Interactive Visual Media Group
Microsoft Research
sbkang@microsoft.com

Abstract

We propose a method of simultaneously calibrating the radial distortion function of a camera along with the other internal calibration parameters. The method relies on the use of a planar (or alternatively non-planar) calibration grid, which is captured in several images. In this way, the determination of the radial distortion is an easy add-on to the popular calibration method proposed by Zhang [17]. The method is entirely non-iterative, and hence is extremely rapid and immune from the problem of local minima.

Our method determines the radial distortion in a parameter-free way, not relying on any particular radial distortion model. This makes it applicable to a large range of cameras from narrow-angle to fish-eye lenses. The method also computes the centre of radial distortion, which we argue is important in obtaining optimal results. Experiments show that this point may be significantly displaced from the centre of the image, or the principal point of the camera.

1 Introduction

Radial distortion is a significant problem in the analysis of digital images. Although this problem was widely studied by photogrammetrists, striving for extreme accuracy, it has been largely ignored in the extensive literature of Structure and Motion of the past decade or so. (Less than 5 pages are devoted to this topic in [10].) Almost exclusively, methods have involved iterative techniques such as bundle-adjustment.

At the same time, several different camera models have been proposed for different types of cameras. The most popular radial distortion model is the even-order polynomial model that models radial distortion as scaling by a factor $1 + \kappa_1 r^2 + \kappa_2 r^4 + \ldots$.

In this paper we prefer to ignore the issue of choosing any particular radial distortion model, by adopting a model-free approach. The only assumption we make on the radial distortion function is that it is monotonic. Despite this, it is possible to determine the radial distortion curve. This way, the problem of fitting a parametrized function model (if required) is separated out from the estimation of the distortion curve, and makes our method applicable to all (or at least most) lenses.

Iterative optimization methods can be troublesome, due to lack of convergence, choosing an initial estimate, and determining a stopping criterion. The advantage of our algorithm is that it is entirely non-iterative. This makes it fast and immune from these problems of iterative techniques. Sometimes, the cost of using a non-iterative technique is that it can minimize some arbitrary cost function unrelated to the noise model. Thus, they can be very sensitive to noise. On the other hand, the linear models that we use are closely associated with the optimal model. Although iterative refinement gives some improvement, it is minimal.

Another feature of our method that distinguishes it from most other computer-vision approaches to radial distortion is that we also compute the centre of distortion. The importance of determining this point has long been recognized in the photogrammetry community; to quote Clarke *et al.* [5]: "attention must be paid to even small details, such as the location of the principal point[1], if residuals are to be optimally minimized." By experimentation, we show that the usual assumption that the centre of distortion is at the centre of the image is not safe.

The disadvantage of our method is that it requires a simple calibration grid. We briefly explore the extension of our method to autocalibration methods that do not use a calibration grid. This is indeed possible, and works with exact measurements, but our observation is that it is extremely sensitive to noise.

*This work was done while R.I. Hartley visited Microsoft Research.

†National ICT Australia is funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

[1] Clarke *et al.* use the term principal point to mean the centre of distortion.

2 Prior Work

The work on radial distortion removal is extensive. Photogrammetry methods usually rely on known calibration points or structures (e.g., [3, 4, 6, 15]). Tsai [15] uses corners of regularly spaced boxes of known dimensions for full camera calibration, including radial distortion. Faig [6] only requires that the points used be coplanar. However, Wei and Ma [16] use projective invariants to recover radial distortion coefficients.

Becker and Bove [2] map straight lines onto a unit sphere and find both radial and decentring coefficients that minimize the vanishing point dispersion. (The user has to manually group parallel lines together–each group should have a unique vanishing point.) Swaminathan and Nayar [13] proposed a user-guided self-calibration approach. The distortion parameters are computed from user-selected points along projections of straight lines in the image. Stein [12] describes a more flexible approach, requiring only point correspondences between multiple views. He uses epipolar and trilinear constraints, and searches for radial distortion parameters that minimizes the errors in these constraints.

By incorporating *one* lens distortion parameter into the epipolar constraint involving the fundamental matrix, Fitzgibbon [7] simultaneously solves for both. He then casts the problem as a quadratic-eigenvalue problem, which can then be easily solved using a numerical library. This technique was later generalized to omnidirectional cameras [11]. Independently of our work, Thirthala and Pollefeys [14] propose a linear technique to recover radial distortion. They assume the center of distortion is at the image center. They use the observation that the vectors between the distorted and undistorted points pass through a common point to reduce the problem to estimating a 2D trifocal tensor.

3 Radial distortion model

In our model for radial distortion correction, we ignore decentring distortion, which is commonly due to lack of alignment of different lens elements. This can lead to non-radial components of lens distortion. We assume that distortion is radial. The imaging process therefore is made up of several steps:

1. **Projection.** Points are projected onto the image plane by the ideal camera, via the mapping

$$\tilde{\mathbf{x}}^u = [\mathtt{I}|\mathbf{0}]\mathbf{X}$$

   where $\mathbf{X}$ is expressed in the camera coordinate frame.

2. **Radial distortion.** The distorted point $\tilde{\mathbf{x}}^d$ is given by

$$\tilde{\mathbf{x}}^d = \tilde{\mathbf{e}} + \lambda(\tilde{\mathbf{x}}^u - \tilde{\mathbf{e}})$$

   where λ represents the amount of distortion, and $\tilde{\mathbf{e}}$ is the *centre of distortion.*

3. **Pixel sampling.** The details of the pixel coordinate system are encapsulated in the calibration matrix $\mathtt{K}$. The point $\tilde{\mathbf{x}}^d$ is mapped to the pixel image point

$$\mathbf{x}^d = \mathtt{K}\tilde{\mathbf{x}}^d \ .$$

In this formulation, the tilde in $\tilde{\mathbf{x}}^u$, $\tilde{\mathbf{x}}^d$ and $\tilde{\mathbf{e}}$ indicates that it is measured in geometric focal length units. We may also define $\mathbf{x}^u = \mathtt{K}\tilde{\mathbf{x}}^u$ and $\mathbf{e} = \mathtt{K}\tilde{\mathbf{e}}^u$, to get the undistorted point and centre of distortion in pixel coordinates. Since $\mathtt{K}$ represents an affine transformation of coordinates, and affine transformation preserve length ratios along straight lines, we see that the pixel coordinates are related by the equation

$$\mathbf{x}^d = \mathbf{e} + \lambda(\mathbf{x}^u - \mathbf{e})$$

with the same λ as before.

Radial symmetry. Generally, without decentering distortion, it may be assumed that the distortion factor λ is a function only of the radius $\tilde{r} = ||\tilde{\mathbf{x}}^u - \tilde{\mathbf{e}}||$. It is often assumed that it may be expressed as a function of the "pixel radius", $r = ||\mathbf{x}^u - \mathbf{e}||$. This is not true unless the image has square pixels, however, so that $\mathtt{K}$ represents a scaled Euclidean coordinate transform.

4 Finding distortion centre

Our simplest method for estimating the centre of radial distortion involves the use of a calibration rig, consisting of a plane with several distinguishable points. The positions of the points are assumed known in a Euclidean coordinate frame on the plane. A suitable such rig would be a checkerboard pattern, the vertices of the squares forming our set of points; it is not necessary to recognize a distinguished vertex, since any vertex can serve as the coordinate origin.

We give an intuitive description of our idea before writing the mathematical formalities. The pattern of known points is projected into the image by an ideal non-distorted camera, and then the points are each moved from their "initial" to their "final" position by expansion away from (or towards) a centre of expansion. We can compare this with the motion of points seen by a camera moving forwards towards a scene. In this case, the points also undergo a radial expansion, and in this case the centre of expansion is known as the epipole. It is well known how to find the epipole – we compute the fundamental matrix. The situation is entirely analogous here, and we can compute the centre of radial distortion, by computing the fundamental matrix relating the known coordinates of points on our calibration rig, and the measured positions of the points in the distorted image.

Let $\mathbf{x}_i^c$ be a set of known coordinates on a planar calibration grid, and $\mathbf{x}_i^d$ be the corresponding points in the distorted image. The calibration rig points $\mathbf{x}_i^c$ and the undistorted image points $\mathbf{x}_i^u$ (in pixel coordinates) are related by a homography $\mathtt{H}$, according to $\mathbf{x}_i^u = \mathtt{H}\mathbf{x}_i^c$. Note that the calibration rig points $\mathbf{x}_i^c$ are common to all images and are arbitrarily set once; as a result, $\mathtt{H}$ is not directly linked to a projection of real 3D points.

The points $\mathbf{x}_i^u$ are next distorted radially away from the centre of distortion $\mathbf{e}$, to give

$$\mathbf{x}_i^d = \mathbf{e} + \lambda_i(\mathbf{x}_i^u - \mathbf{e}) \ .$$

We multiply this expression on the left by $[\mathbf{e}]_\times$ (the 3×3 matrix representing the cross product), resulting in $[\mathbf{e}]_\times \mathbf{x}_i^d = \lambda_i [\mathbf{e}]_\times \mathbf{x}_i^u$, where the terms $\mathbf{e}$ disappear when multiplied by $[\mathbf{e}]_\times$. However, since $\mathbf{x}_i^u = \mathtt{H}\mathbf{x}_i^c$, we have

$$[\mathbf{e}]_\times \mathbf{x}_i^d = \lambda_i [\mathbf{e}]_\times \mathtt{H}\mathbf{x}_i^c \ .$$

Finally, multiplying on the left by $\mathbf{x}_i^{d\top}$, and observing that $\mathbf{x}_i^{d\top}[\mathbf{e}]_\times \mathbf{x}_i^d = 0$, we obtain

$$0 = \lambda_i \mathbf{x}_i^{d\top}([\mathbf{e}]_\times \mathtt{H})\mathbf{x}_i^c \ .$$

Writing $\mathtt{F} = [\mathbf{e}]_\times \mathtt{H}$, we have the usual fundamental matrix relation $\mathbf{x}_i^{d\top}\mathtt{F}\mathbf{x}_i^c = 0$. The matrix $\mathtt{F}$ may be called the *fundamental matrix for radial distortion*.

The matrix $\mathtt{F}$ may be computed in the usual way from several point correspondences, and the centre of radial distortion extracted as the left epipole.

Using multiple images. Instead of using a single image for computation of the focus of expansion, we may consider how we may take advantage of several views of a calibration grid. We consider all the points in all images indexed by a single index i, and denote by $k(i)$ the image that the i-th point belongs to. The fundamental equation then becomes

$$\mathbf{x}_i^{d\top}([\mathbf{e}]_\times \mathtt{H}_{k(i)})\mathbf{x}_i^c = \mathbf{x}_i^{d\top}\mathtt{F}_{k(i)}\mathbf{x}_i^c = 0 \ .$$

One may then find the vector $\mathbf{e}$ as the simultaneous left nullspace generator of all the $\mathtt{F}_k$.

5 Determining Homographies

Next, we wish to find the homographies $\mathtt{H}_k$ mapping the calibration grid to the image plane. We can solve for the homography matrix $\mathtt{H}_k$ by factoring the fundamental matrix as $\mathtt{F}_k = [\mathbf{e}]_\times \mathtt{H}_k$. However, a preferable method is to repeat the computation of each $\mathtt{H}_k$ with known epipole. In doing this, it is convenient to change the image coordinates first so that $\mathbf{e}$ is the coordinated origin $(0, 0, 1)^\top$ (homogeneous coordinates). We carry out all subsequent computations in this coordinate system and make an appropriate correction at the end.

Thus, we compute $\mathtt{H}_k$ by solving the equations

$$\mathbf{x}_i^{d\top}\mathtt{F}_k\mathbf{x}_i^c = \mathbf{x}_i^{d\top}([\mathbf{e}]_\times \mathtt{H}_k)\mathbf{x}_i^c = 0$$

individually for each image k using points belonging to that image. Note that if $\mathbf{e} = (0, 0, 1)^\top$, then the final row of $\mathtt{F}_k$ is zero, so we only solve for the 6 other entries of $\mathtt{F}_k$. Then, the homography $\mathtt{H}_k$ can be written down directly from $\mathtt{F}_k$. Namely, $\mathtt{H}_k = [\mathbf{f}_2{}^\top; -\mathbf{f}_1{}^\top; \mathbf{0}]$, where the semicolon means that we stack the rows $\mathbf{f}_i{}^\top$ of $\mathtt{F}$ on top of each other.

Note here that each of $\mathtt{H}_k$ is not uniquely defined by this method, since in fact the last row of $\mathtt{H}_k$ may be arbitrary. Generally, in factoring $\mathtt{F} = [\mathbf{e}]_\times \mathtt{H}$, we may replace $\mathtt{H}$ by $\mathtt{H}_k + \mathbf{e}\mathbf{v}^\top$ for any arbitrary $\mathbf{v}$ without changing the form of the equation. This is because $\mathbf{e}\mathbf{v}^\top$ cancels with $[\mathbf{e}]_\times$. Note that only the third row of $\mathbf{e}\mathbf{v}^\top$ is non-zero.

We now turn to finding the final row of each homography $\mathtt{H}_k$. Various methods of determining the unknown vector $\mathbf{v}$ are available. Essentially this problem is solved in [14] by considering a specific polynomial parametrization of the radial-distortion curve. We prefer a parameter-free method of doing this based on two assumptions:

1. The distortion is radially symmetric. Thus, the radial distortion of an image point depends only on its distance from the centre of distortion.

2. An ordering, or monotonicity condition: the radial distance of points from the radial centre after distortion is a monotonic function of their distance before distortion.

The first condition will not hold in general unless the pixels are square. However, it turns out in practice not to be critical to the success of our method, which works well even with non-square pixels, and in fact the aspect ratio of the pixels falls naturally out of the computation. The second condition is an essential property of any camera, and indeed it would be a strange camera that did not satisfy this condition, since it would mean that more than one ray through the camera centre corresponds to a pixel in the image.

We consider a single homography $\mathtt{H}_k$ and temporarily drop the subscript k. Let $\hat{\mathtt{H}}$ consist of the first two rows of $\mathtt{H}$ so that $\mathtt{H} = [\hat{\mathtt{H}}; \mathbf{v}^\top]$, (this notation means that $\mathbf{v}^\top$ is the final row of $\mathtt{H}$). Now, defining $(\hat{x}^u, \hat{y}^u)^\top = \hat{\mathbf{x}}^u = \hat{\mathtt{H}}\mathbf{x}^c$, we see that $\mathbf{x}^u = [\hat{\mathtt{H}}\mathbf{x}^c; \mathbf{v}^\top\mathbf{x}^c]$, and dehomogenizing, we obtain

$$\mathbf{x}^u = (\hat{x}^u, \hat{y}^u)/(\mathbf{v}^\top\mathbf{x}^c) \ .$$

Thus the effect of $\mathbf{v}^\top$ is to stretch the point $\hat{\mathbf{x}}$ by the factor $1/(\mathbf{v}^\top\mathbf{x}^c)$, which depends on the point $\mathbf{x}^c$.

We now compute the radii of the distorted and undistorted points, setting $r^d = |\mathbf{x}^d|$ and $\hat{r}^u = |\hat{\mathbf{x}}^u|$, and finally $r^u = |\hat{\mathbf{x}}^u/(\mathbf{v}^\top\mathbf{x}^c)| = \hat{r}^u/|\mathbf{v}^\top\mathbf{x}^c|$. More precisely, since we desire

$\mathbf{x}^u$ and $\mathbf{x}^d$ both to lie in the same radial direction (and not on opposite sides of the centre of distortion), we define the radius $\hat{r}^u$ to be positive or negative depending on whether $\mathbf{x}^{d\top}\hat{\mathbf{x}}^u$ is positive or negative. Then we may remove the absolute value on $\mathbf{v}^\top\mathbf{x}^c$, resulting in

$$r^u = \hat{r}^u/(\mathbf{v}^\top\mathbf{x}^c) \ .$$

If we were able to select the correct value of the vector $\mathbf{v}$, then the points (r^d, r^u) would lie along (or with noise, close to) a monotonic curve, as illustrated in the scatter plot of Fig 2. For any other (incorrect) value of $\mathbf{v}$, the scatter plot would be irregularly scattered, because of the different scaling of each value r^u according to $\mathbf{v}^\top\mathbf{x}^c$. It is our goal to find the value of $\mathbf{v}$ that reduces the scatter plot to a monotonic smooth curve. Perhaps surprisingly, this can be accomplished well by simple least-squares techniques.

We begin by ordering the values of r^d, and index them in order by an index i: we write $r_i^d; i = 1, \ldots, N$. The corresponding value of $r_i^u = \hat{r}_i^u/(\mathbf{v}^\top\mathbf{x}^c)$ may be thought of as a discrete function of the values r_i^d. We define the *total squared variation* of this function to be

$$V = \sum_{i=1}^{N-1} (r_{i+1}^u - r_i^u)^2 \ .$$

Note that the ordering of the radii r_i^d is done only *once*.

If the values of r_i^u are a monotonic function of r_i^d, then the total squared variation of this function will be relatively small, compared with that of an irregular function. In fact, it is easily seen that for a monotonic function, $V < (r_{N-1}^u - r_1^u)^2$, and in fact if the values of r_i^u are evenly scattered between r_1^u and r_N^u, then $V \to 0$ as $N \to \infty$. In measurements involving a large number of points, the radii will in general be well scattered and the value of the total squared variation will be small although we do not expect it to vanish completely. Our method therefore is to minimize the total squared variation V as a function of the parameter vector $\mathbf{v}$.

The above discussion extends easily to the case of multiple images. In this case, we consider all the radii of the points in all images together and order them in one single list. The undistorted radius r^u is equal to $\hat{r}^u/(\mathbf{v}_k{}^\top\mathbf{x}^c)$, where k is an index representing the image number. For each image k there will be a different homography $\mathtt{H}_k$, and a different corresponding vector $\mathbf{v}_k$. The total square variation must be minimized over the choice of all $\mathbf{v}_k$.

Minimizing total squared variation. We wish to minimize

$$\sum_{i=1}^{N-1} (r_{i+1}^u - r_i^u)^2$$

$$= \sum_{i=1}^{N-1} \left(\frac{\hat{r}_{i+1}^u}{\mathbf{v}_{k(i+1)}{}^\top\mathbf{x}_{i+1}^c} - \frac{\hat{r}_i^u}{\mathbf{v}_{k(i)}{}^\top\mathbf{x}_i^c} \right)^2 ,$$

where $k(i)$ represents the image number corresponding to the i-th point. As it stands, this is not a linear least-squares problem. However, multiplying each term by $(\mathbf{v}_{k(i+1)}{}^\top\mathbf{x}_{i+1}^c)^2(\mathbf{v}_{k(i)}{}^\top\mathbf{x}_i^c)^2$ leads to

$$\sum_{i=1}^{N-1} \left(\hat{r}_{i+1}^u \mathbf{v}_{k(i)}{}^\top\mathbf{x}_i^c - \hat{r}_i^u \mathbf{v}_{k(i+1)}{}^\top\mathbf{x}_{i+1}^c\right)^2 ,$$

which must be minimized over all $\mathbf{v}_k$. Note that apart from the values of $\mathbf{v}_k$, all other quantities appearing in this expression are known. The minimization problem is one of minimizing the squared norm of the $(N-1)$ dimensional vector with entries

$$\hat{r}_{i+1}^u \mathbf{v}_{k(i)}{}^\top\mathbf{x}_i^c - \hat{r}_i^u \mathbf{v}_{k(i+1)}{}^\top\mathbf{x}_i^c,$$

which are linear in the entries of the vectors $\mathbf{v}_k$. Thus we may write this vector as $\mathtt{A}\mathbf{V}$, where $\mathbf{V}$ is a vector containing all the entries of the $\mathbf{v}_k$, and $\mathtt{A}$ is the matrix of coefficients, which may be constructed easily from the known values of all the $\hat{r}_i^u$ and $\mathbf{x}_i^c$. This problem can be solved by minimizing $\|\mathtt{A}\mathbf{V}\|$ subject to the condition $\|\mathbf{V}\| = 1$. An alternative (which we prefer) is to impose the condition on $\mathbf{v}_{k(N)}$ that ensures that $r_N^d = r_N^u = \hat{r}_N^u/(\mathbf{v}_{k(N)}{}^\top\mathbf{x}_N^c)$, namely that the distorted and undistorted radii are equal for the point most distant from the centre of distortion. This gives a linear equality condition on $\mathbf{v}_{k(N)}$, and leads to a simple linear least-squares estimation problem. (Note that we are free to make this assumption that $r_N^d = r_N^u$ because of the ambiguity between the scale of the distortion and the overall scale represented by the calibration matrix, specifically the focal length.)

Note that we linearized the problem through multiplication by $(\mathbf{v}_{k(i+1)}{}^\top\mathbf{x}_{i+1}^c)^2(\mathbf{v}_{k_i}{}^\top\mathbf{x}_i^c)^2$, which results in an unequal weighting of the individual equations, and means that we do not exactly minimize the total squared variance of the distortion curve. However, this effect is quite benign, since the values of $\mathbf{v}_{k(i)}{}^\top\mathbf{x}_i^c$ represent the depth of the calibration point in the direction of the camera's principal ray. Under the assumption that all these depths are approximately equal, all the equations are weighted approximately equally. The advantage is that it changes a non-linear estimation problem into a linear one. When testing this algorithm, we followed the linear estimation step by non-linear refinement to minimize the true total squared variance, as well as various other geometrically derived conditions designed to lead to a smooth monotonic distortion curve. These will be described in a more detailed version of this paper. Suffice it to say that the improvements achieved were minimal.[2]

[2]To be precise, non-linear refinement caused less than 1.2% decrease in residual in the example of Fig 2, and no improvement in the final result.

5.1 Estimating radial distortion function

In this paper we deliberately do not consider parametrization of the distortion function. This distortion function is often modelled as $r^d = r^u(1 + \kappa_1 r^{u2} + \kappa_2 r^{u4} + ...)$. Such a distortion function does not work well for large distortion such as those given by fish-eye lenses. It is a strength of our method that it does not rely on any particular distortion model, and it has been tested successfully on fish-eye, wide angle, and narrow angle lenses. Of course, ultimately the curve must be approximated by some technique to be useful for image correction.

There is a large literature on the topic of function approximation (see [1] for a good summary). Separating out the problem of determination of a parametrized approximation of the correction curve allows more sophisticated techniques and algorithms to be used. These include approximation of the distortion function by Chebyshev polynomials, or the Remes minimax algorithm. Direct estimation of the polynomial coefficient (κ_i) or use of a Taylor expansion is distinctly ill-conditioned or suboptimal [1]. Further discussion of this topic is deferred to a more detailed version of this paper. The method we preferred for the examples given later involved approximation of the distortion function expressed in terms of angular distortion of points on a spherical rather than planar retina, though that is not a critical choice, except for very wide angle lenses. An alternative is to compute a parameter-free model for the distortion curve by computing the radial displacement at each radial value as a sliding median of local measurements.

5.2 Camera calibration

Estimation of the vectors $\mathbf{v}_k$ allows us to complete the homography matrices $\mathtt{H}_k$. From this, we can compute the true undistorted projection of each of the calibration points $\mathbf{x}^c$ into the image. Given these homographies for at least three images, we are able to compute the calibration matrix of the undistorted cameras, using the algorithm of [17].

5.3 Non-square pixels

If the pixels are not square, then distortion is still radial, but the radial distortion function is not symmetric. Despite this, we found that by simply proceeding with the algorithm as described, estimation of the homographies and distortion curve, followed by internal calibration led to an accurate estimate of the aspect ratio of the pixels. This would allow us to correct for the aspect ratio, and recompute the distortion curve. The explanation for this is that even with errors in the expansion ratio due to non-square pixels, the distortion curve has the correct shape, and led to the correct value of the $\mathbf{v}_k$. The effect of assuming square pixels when they in fact were not leads to a slight broadening of the distortion curve, but not a change to its shape. More investigation of the non-square pixel case is planned. Results were obtained for synthetic data only, and remain anecdotal.

6 Extensions of basic idea

We describe several extensions to this calibration method, based on the general idea of using structure and motion techniques (similar to the use of the fundamental matrix above) to estimate radial distortion. We have implemented all of these methods in some form, and verified their correctness.

6.1 Using non-planar calibration grid

The above method was described in terms of using a known planar calibration grid. However, the method carries over entirely to a non-planar calibration grid. Instead of a fundamental-matrix relation $\mathbf{x}^d\mathtt{F}\mathbf{x}^c = 0$ when the grid is planar, we would obtain an equation $\mathbf{x}^d\mathtt{G}\mathbf{X}^c$ when the calibration points $\mathbf{X}^c$ are known 3D points. One can estimate the matrix $\mathtt{G}$ (which has dimension 3×4) in just the same way one estimates the fundamental matrix. Computation of the final row $\mathbf{v}$ of the projection matrix $\mathtt{G}$ proceeds as before, as does estimation of the distortion curve. The only difference involves the computation of the calibration matrix. Instead of needing three views to compute $\mathtt{K}$, it can be computed from a single view. A single projection matrix $\mathtt{G}$ gives 5 equations in the entries of ω, namely $\mathbf{g}_1{}^\top\omega\mathbf{g}_1 = \mathbf{g}_2{}^\top\omega\mathbf{g}_2 = \mathbf{g}_3{}^\top\omega\mathbf{g}_3$ and $\mathbf{g}_1{}^\top\omega\mathbf{g}_2 = \mathbf{g}_2{}^\top\omega\mathbf{g}_3 = \mathbf{g}_1{}^\top\omega\mathbf{g}_3 = 0$, where $\mathbf{g}_i$ is the i-th column of $\mathtt{G}$. Thus we can solve for ω, and hence $\mathtt{K}$ from a single view, or from many views by linear least-squares techniques.

We tried this method on synthetic data, and verified that it worked with similar accuracy to the planar method. However, we did not try it on real images.

6.2 Doing without calibration grid

It is natural to ask whether we could do without the calibration grid entirely. The method described above used the fundamental matrix as the main algebraic tool. However, the process is analogous to determining the pose of a camera using line correspondences. Indeed, in the case of 3D points, the correspondence equation $\mathbf{x}^{d\top}\mathtt{G}\mathbf{X}^c = 0$ is of precisely the same form as the equation $\mathbf{l}^\top\mathtt{P}\mathbf{X} = 0$ for the image of a point $\mathbf{X}$ to lie on a line $\mathbf{l}$ when projected by camera matrix $\mathtt{P}$. Thus, the determination of the projection matrices and the centre of distortion is similar to that of reconstruction from line correspondences. In fact, the line reprojection problem can be solved using the quadrifocal tensor with 4 views. Corresponding distorted points $\mathbf{x}_j = (x_j^1, x_j^2, x_j^3)^\top$

in image j for $j = 1, \ldots, 4$ will satisfy the relationship

$$\sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{r=1}^{3}\sum_{s=1}^{3} \mathbf{x}_1^p \mathbf{x}_2^q \mathbf{x}_3^r \mathbf{x}_4^s Q_{pqrs} = 0,$$

from which it is possible to compute the quadrifocal tensor Q [9], and hence the projection matrices [8]. There is a slight complication that two solutions exist in both the quadrifocal and trifocal cases (see below), as pointed out in [8].

We verified that this method will work for the present problem also. A total of 80 point correspondences are needed, since each point gives only one equation. The same method works equally for calibration using unknown points on a plane. In this case, we need only 3 views, and a trifocal tensor is used. In this case, 26 points are needed. However, although we verified both practically and theoretically that this method will work, experiments with synthetic data showed the methods to be far too sensitive to noise to be useful, at least with unknown centre of distortion.

It is appropriate to point out the relationship of this work with that of Thirthala and Pollefeys [14]. In our work, we make no assumption on the position of the centre of distortion. Thirthala and Pollefeys, on the other hand, assume that the centre of distortion is known. In this case, assuming it to be at $(0, 0, 1)$, we can immediately set most of the entries of the tensor to zero, leaving only $2^4 = 16$ (in the 3D case) or $2^3 = 8$ (in the planar case) non-zero entries to estimate. This large reduction in the number of unknowns makes it likely that the method will be far more immune to noise. They considered this problem for the case of 2D points, but using 3D points is a natural extension.

7 Results

In this section, we present distortion removal results for a known planar calibration grid. The calibration pattern is a checkerboard with black-and-white squares (left image of Fig 1). The corners of the checkerboard are extracted by finding edges and computing intersection points; sample results are shown on the right of Fig 1. We used 19 images of the checkerboard taken at various poses.

After we applied our technique as described in Section 2 and 4, we obtain the graph of radial shifts as a function of radial position as shown in Fig 2. The best fit 4th degree polynomial (which passes through (0,0)) is superimposed in the graph.

In Fig 3 is shown the distribution of the estimated distortion center in three cases: each image used separately, 5 images randomly chosen from 19, and 10 randomly chosen images. The distortion center estimated using all 19 images is (306.7, 260.5), and the image center is (320, 240). The principal point distributions extracted using the same experiments are shown in Fig 4. The principal point estimated

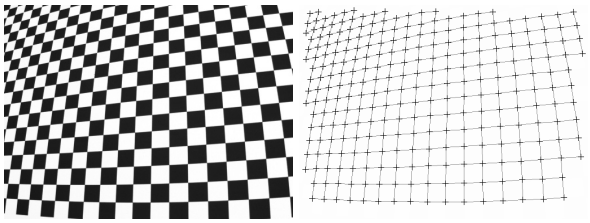

Figure 1. *An image of a checkerboard (left) and the detected grid (right).*

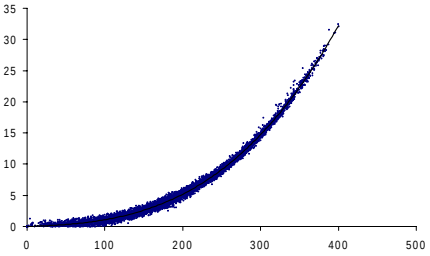

Figure 2. *Graph of radial displacement (distortion) vs. radial position using corner points from 19 checkboard images. The number labels are all in pixels; the image resolution is* 640×480*. In this case, we achieved an RMS error of* 0.4 *pixels in modelling the measured image points.*

using all 19 images is (312.0, 244.8). Note that while the estimated principal point is close to the image center, the estimated distortion center is not.

We used the estimated image noise (based on RMS error fit of $\sigma = 0.4$) and performed Monte Carlo simulation with 1000 random trials. The results are shown in Fig 5. Compare the locations of the estimated distributions for the center of distortion and principal point, and the image center. The mean for the distortion center is (304.3, 259.2) with standard deviation of (0.87, 0.60). The mean for the principal point is (310.6, 244.6) with standard deviation of (0.58, 2.26).

We applied the computed radial distortion mapping to a number of different input images for correction. Two results are shown in Fig 6.

In another set of experiments, we applied our planar calibration technique on three further cameras. These three cameras are the same type (PointGrey Fleas, resolution of 1024×768), use the same type of 4mm lens, and have the same settings. The results based on 10 images of the checkerboard pattern are shown in Table 1. Notice the significant differences in the locations of the estimated distortion center, estimated principal point, and image center. Notice also the significant differences in the estimated distortion center and principal point for the same type of cameras.

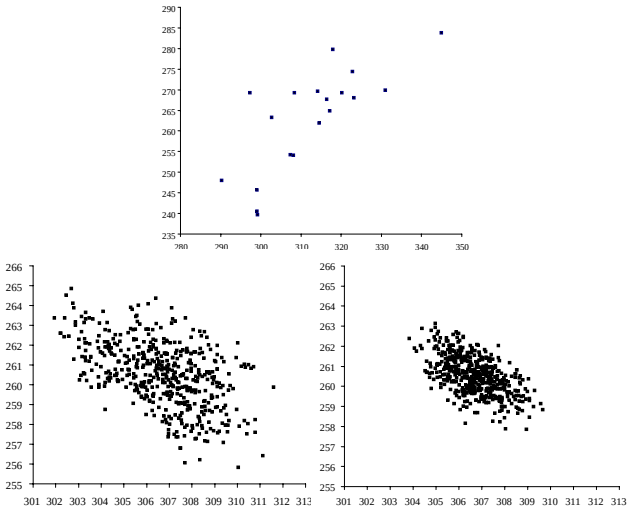

Figure 3. *Distributions of distortion center (from left to right): using single image separately (19 inputs), using 5 random sets (5 of 19), using 10 random sets (10 of 19). Note the different scales in the graphs. The number labels are all in pixels; the image resolution is* 640×480.

#	$\mathbf{c}_d$	$\mathbf{c}_p$	$\mathbf{c}_0$
1	(574.4, 401.5)	(555.6, 377.7)	(512, 384)
2	(545.9, 394.5)	(542.6, 396.6)	(512, 384)
3	(577.8, 392.3)	(574.4, 391.0)	(512, 384)

Table 1. *Comparison of results for three different cameras (same model and same type of 4mm lens).* $\mathbf{c}_d$, $\mathbf{c}_p$, *and* $\mathbf{c}_0$ *are the distortion center, principal point, and image center, respectively.* $\mathbf{c}_d$ *and* $\mathbf{c}_p$ *were computed based on ten images of the same checkerboard pattern at different poses.*

We also applied our technique to a camera with a fish-eye lens with close to 180° field of view. A sample image of the checkerboard pattern (we use 8 calibration images), the extracted grid, and radial distortion results are shown in Fig 7. The image resolution is 1536×1024, and the extracted distortion center is (744.4, 488.2) and principal point is (767.5, 492.4) (compared with the image center (768, 512)). As can be seen, the corrected results look very reasonable.

8 Conclusions

The procedure for radial distortion calibration described in this paper represents a reliable method of determining the centre of distortion and radial distortion function for a wide range of cameras, including fish-eye lenses. At the same time it allows computation of the internal geometric parameters of the camera. As a fast non-iterative procedure, it may be used to initialize a bundle-adjustment algorithm, though for many applications, not requiring extreme precision, it can be used on its own as a stand-alone algorithm for camera calibration.

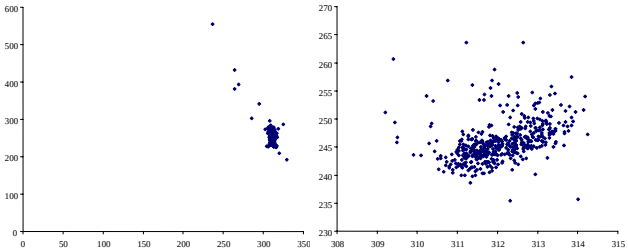

Figure 4. *Distributions of principal point (from left to right): using 5 random sets (5 of 19), using 10 random sets (10 of 19). Note the different scales in the graphs. The number labels are all in pixels; the image resolution is* 640×480.

We argued the necessity of determining the radial distortion centre, and showed that for all cameras that we tried the distortion centre was significantly displaced from the centre of the image, by as much as 30 pixels in a 640×480 image. Our experiments showed that this was enough to cause an extra 0.4 pixels of error for some points in the image (though the RMS error only increased by 0.1 pixel). By contrast, we were able to model the measured image points with RMS error of 0.4 pixels (see the example of Fig 2).

Extensions to auto-calibration techniques involving images or planar or non-planar points are possible, but may be too sensitive to noise to be useful, particularly when the centre of distortion needs to be computed as well.

References

[1] K. E. Atkinson. *An Introduction to Numerical Analysis.* Wiley, 1978.

[2] S. Becker and V. B. Bove. Semiautomatic 3-D model extraction from uncalibrated 2-D camera views. In *SPIE Visual Data Exploration and Analysis II*, vol. 2410, pp. 447–461, 1995.

[3] D. C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3):444–462, 1966.

[4] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.

[5] T.A. Clarke, J.F. Fryer, and X. Wang. The principal point and CCD cameras. *Photogrammetric Record*, 16(92):293–312, 1998.

[6] W. Faig. Calibration of close-range photogrammetry systems: Mathematical formulation. *Photogrammetric Engineering and Remote Sensing*, 41(12):1479–1486, 1975.

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

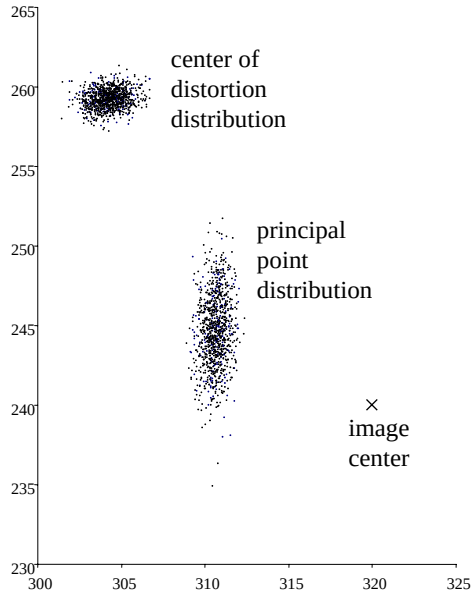

Figure 5. *Monte Carlo simulation (1000 trials with* $\sigma = 0.4$*). The number labels are all in pixels; the image resolution is* 640×480*.*

[8] R. Hartley and F. Schaffalitzky. Reconstruction from projections using Grassman tensors. In *ECCV*, vol. 3, pp. 363–375, 2004.

[9] R. I. Hartley. Computation of the quadrifocal tensor. In *ECCV*, pp. 20–35, 1998.

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

[11] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. *CVPR*, vol. I, pp. 485-490, 2003.

[12] G. Stein, Accurate internal camera calibration using rotation, with analysis of sources of error. In *ICCV*, pp. 230–236, 1995.

[13] R. Swaminathan and S. Nayar. Non-metric calibration of wide-angle lenses and polycameras. In *CVPR*, pp. 413–419, 1999.

[14] S. Thirthala and M. Pollefeys. The radial trifocal tensor: A tool for calibrating the radial distortion of wide-angle cameras. In *CVPR*, vol. 1, pp. 321–328, 2005.

[15] Y. R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, 1987.

[16] G.-Q. Wei and S. D. Ma. Implicit and explicit camera calibration: Theory and experiments. *PAMI*, 16(5):469–480, 1994.

[17] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *ICCV*, pp. 666–673, 1999.

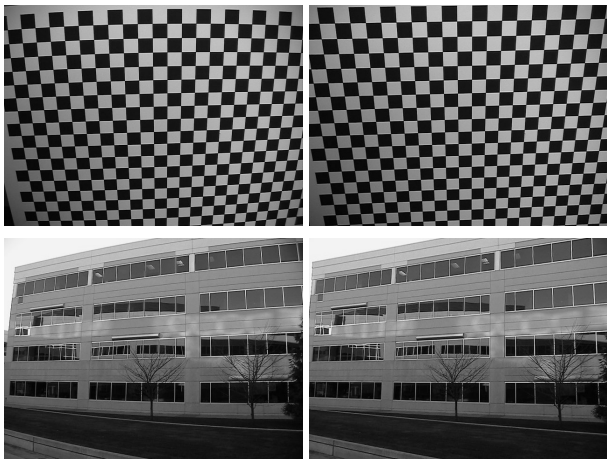

Figure 6. *Distortion removal result: input (left column) and corrected output (right column).*

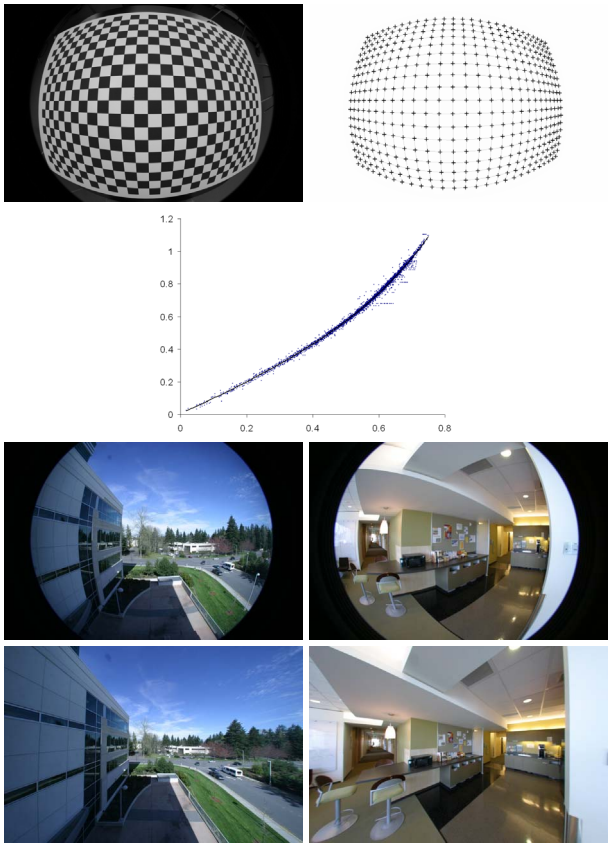

Figure 7. *Fisheye image experiment (from top to bottom): an image of the checkerboard and extracted grid, the radial distortion curve (undistorted versus distorted angle in radians), two inputs, and two radially corrected images.*