Overconstrained Linear Estimation of Radial Distortion and Multi-view Geometry

R. Matt Steele and Christopher Jaynes

Center for Visualization and Virtual Environments,
University of Kentucky, Lexington, Kentucky, USA

Abstract. This paper introduces a new method for simultaneous estimation of lens distortion and multi-view geometry using only point correspondences. The new technique has significant advantages over the current state-of-the art in that it makes more effective use of correspondences arising from any number of views. Multi-view geometry in the presence of lens distortion can be expressed as a set of point correspondence constraints that are quadratic in the unknown distortion parameter. Previous work has demonstrated how the system can be solved efficiently as a quadratic eigenvalue problem by operating on the normal equations of the system. Although this approach is appropriate for situations in which only a minimal set of matchpoints are available, it does not take full advantage of extra correspondences in overconstrained situations, resulting in significant bias and many potential solutions. The new technique directly operates on the initial constraint equations and solves the quadratic eigenvalue problem in the case of rectangular matrices. The method is shown to contain significantly less bias on both controlled and real-world data and, in the case of a moving camera where additional views serve to constrain the number of solutions, an accurate estimate of both geometry and distortion is achieved.

1 Introduction

Radial distortion introduces systematic error into the results of standard linear algorithms (e.g. the Eight Point Algorithm, Direct Linear Transform homography estimation, trifocal tensor estimation) that do not account for it. Many applications require wide-angle lenses, for which distortion can be quite severe. Although a priori modelling of lens distortion [1] can remedy the problem, some computer vision tasks, such as structure and motion recovery from uncalibrated video, preclude offline calibration by definition. Assumptions about the scene structure in order to perform online distortion estimation [2] are often undesirable, and can be error-prone.

Consequently, there has been significant work to obtain distortion estimates based only on image-to-image correspondences and the application of multi-view geometric constraints, that is, exactly the information available in uncalibrated video of an unknown scene. By insightful choice of distortion model, Fitzgibbon [3] is able to express the epipolar constraints for distorted correspondences

A. Leonardis, H. Bischof, and A. Pinz (Eds.): ECCV 2006, Part I, LNCS 3951, pp. 253–264, 2006.

as a quadratic eigenvalue problem (QEP). Solutions may be found via efficient and globally convergent algorithms, yielding an estimate for both the distortion and the epipolar geometry. Some information is lost, however, in multiplication by the transpose matrix in order to make the rectangular eigensystem square. One can see this is the case by noting that no matter how many correspondences are available, the square QEP will still allow multiple solutions. Even the correct choice out of the solutions, that is, the one that best minimizes re-projection error, suffers from a surprising amount of bias due to noise, even when many correspondences are available in a strong geometric configuration.

The standard answer to these deficiencies, refinement via bundle adjustment, can suffer from slow or unreliable convergence when an accurate initialization is unavailable. Consequently, a technique is sought which possesses the highly desirable efficiency and convergence properties of the square QEP while more efficiently exploiting the extra information in all matchpoints to obtain a more accurate estimate of distortion and multi-view geometry.

Section 2 of this paper presents such a technique. Rather than solving a square QEP, a rectangular QEP is constructed with one row for each available correspondence. The rectangular QEP does not have an exact solution, but an optimal approximation may be defined by seeking the closest perturbed eigensystem which does have an exact solution. With the help of results from [4], an efficient algorithm is presented which solves this problem. Building on this contribution, Section 2.3 presents a second, generalized algorithm which supports simultaneous solution of multiple, independent multi-view geometries while enforcing a single, global radial distortion model. Thus, not only are additional correspondences exploited, but also extra view pairs, which need not be interconnected by long-lived feature tracks. Section 3 compares the new algorithms to the previous square QEP method [3] on simulated and real data, revealing striking reductions in estimation variance, and especially bias.

1.1 Related Work

The work builds on a recent tradition of exploring how radial distortion estimation can occur simultaneously with the recovery of multi-view geometry [3, 5, 6]. This tradition is quite different from methods that estimate lens distortion offline [1] or techniques that combine a priori scene knowledge with the results of feature extraction such as plumb-line methods [2]. Instead, more recent efforts don't make assumptions beyond those required for traditional multi-view geometry estimation (e.g. the eight-point algorithm). Much of these efforts have emphasized the importance of simultaneous estimation of both a linear geometric model and the nonlinear distortion parameters. This is an improvement over other methods that have been designed for online radial distortion estimation [7] in that they must deal with each task independently. Independent estimation can lead to bias in the geometric estimate because distorted points are used.

Simultaneous estimation was first explored by Fitzgibbon as a hypothesis generator for RANSAC [3]. The technique was shown to be successful in providing better discrimination between outliers and inliers even in the presence of signif-

icant distortion. Because the radial component is computed simultaneously, the geometric estimate is no longer biased by unmodelled distortion.

In addition to the work of Fitzgibbon, more recent work has presented alternative direct methods of estimating lens distortion parameters without a priori knowledge of scene structure. These methods provide for graceful extension to overdetermined systems, but have other drawbacks. Lifting the correspondences into a higher-dimensional space [8] can allow the distorted epipolar mapping to be represented by linear matrix multiplication in the higher-dimensional space. The lifted matrix, however, has multiple unwanted degrees of freedom, and attempts to appropriately constrain the relationship have not yet been satisfactory. Deriving 1D radial correspondences from the original 2D correspondences allows estimation of a distortion-free multi-view geometry [6] which in turn supports direct estimates of the distortion model. The method requires, however, that all correspondences be constrained to lie on a single plane, or the camera motion be purely rotational.

In addition to work from computer vision, this paper draws on results from the numerical analysis community. Recent results lay the groundwork for an alternative solution based on solving an extended notion of the QEP, one in which the matrices are not square. The concept of pseudospectra, a generalization of eigenvalues for non-square matrices, is discussed and studied in [9]. A non-square analogue of the generalized eigenvalue problem is posed in [4] that builds on these results, and an algorithm is presented for the special case in which only a single, primary pseudoeigenvalue is sought. This line of inquiry informs our approach in solving the rectangular QEP that results from the formulation of the problem studied in this paper.

2 Problem Formulation

Assume that a camera in motion observes an arbitrary scene and that the radial distortion is fixed throughout the image sequence. Under these conditions, the goal is to simultaneously estimate pairwise epipolar relationships as well as the radial distortion coefficient λ. We denote 2D points observed in image i as $\mathbf{x} = (x, y)$ and their correspondences in image j as $\hat{\mathbf{x}}$. Following the notation of [3], each image point, $\mathbf{x}$, is said to arise from a radial distortion model applied to an underlying undistorted point, $\mathbf{p}$.

We shall denote by $\mathtt{F_{ij}}$ the *fundamental matrix* corresponding to the pair of images i and j. The task of this paper, given an image sequence and a set of n image pairings $(i_1 j_1, \ldots, i_n j_n)$ for the sequence, is to derive the fundamental matrices $(\mathtt{F}_{i_1 j_1}, \ldots, \mathtt{F}_{i_n j_n})$ for the view pairs, in addition to the distortion parameter λ governing the radial distortion of points observed in all images.

Traditionally, the eight-point-algorithm can be used to estimate $\mathtt{F}$ for any given pair in the image sequence [10, 11]. By assuming no distortion, a linear system can be derived that utilizes the epipolar constraint described by the fundamental matrix:

$$\hat{\mathbf{p}}^{\mathtt{T}} \mathtt{F} \mathbf{p} = \mathtt{0} \tag{1}$$

The traditional eight-point algorithm requires either offline measurement of λ, an independent estimation of distortion using scene knowledge (e.g. plumb line methods), or restricting matchpoints to a central region of the image where distortion can be neglected. Each of these approaches has significant drawbacks. Offline estimation can be cumbersome and is impossible in the case of archival video. The use of scene knowledge to measure distortion (e.g. [2]) requires that the scene conforms to a priori constraints, and the feature extraction process is typically higher-level and more susceptible to failure than the low-level task at hand. Neglecting potential matchpoints near the periphery of the image is not desirable and can unnecessarily eliminate matchpoints arising from robust features. Many image sequences contain overlap primarily at the periphery of the image, and ignoring matchpoints from these regions leads to unstable estimation of camera geometry.

Given these problems, techniques that support simultaneous estimation of λ are of interest. Recently, Fitzgibbon [3] demonstrated that radial distortion can be cleanly incorporated into Equation 1 by developing a distortion model that only depends on $\mathbf{x}$ and $\hat{\mathbf{x}}$, the measured matchpoints:

$$\mathbf{p} = \frac{1}{1 + \lambda \|\mathbf{x}\|^2} \mathbf{x} \tag{2}$$

Given this division model of distortion, the epipolar constraint is:

$$\begin{aligned} \left(\hat{\mathbf{x}} + \lambda \hat{\mathbf{z}}\right)^{\mathtt{T}} \mathtt{F} \left(\mathbf{x} + \lambda \mathbf{z}\right) &= 0 \\ \hat{\mathbf{x}}^{\mathtt{T}} \mathtt{F} \mathbf{x} + \lambda \left(\hat{\mathbf{z}}^{\mathtt{T}} \mathtt{F} \mathbf{x} + \hat{\mathbf{x}} \mathtt{F} \mathbf{z}\right) + \lambda^2 \hat{\mathbf{z}}^{\mathtt{T}} \mathtt{F} \mathbf{z} &= 0 \end{aligned} \tag{3}$$

where $\mathbf{z}$ is $\begin{bmatrix} 0 & 0 & \|\mathbf{x}\|^2 \end{bmatrix}^{\mathtt{T}}$. Note that Equation 3 is comprised of four terms in $\mathtt{F}$, each possessing the same form as the traditional epipolar constraint.

2.1 The Quadratic Eigenvalue Problem for Estimating λ and $\mathtt{F}$

Simultaneous estimation of λ and $\mathtt{F}$ may be performed by formulating Equation 3 as a quadratic eigenvalue problem (QEP). The QEP is obtained by gathering the vector factors of $\mathtt{F}$, for each term, into a separate design matrix. The elements of $\mathtt{F}$ are extracted into vector $\mathbf{f}$ [3]. This procedure is identical to the method by which the traditional eight-point equations are obtained from Equation 1.

$$\left(\mathtt{D}_1 + \lambda \mathtt{D}_2 + \lambda^2 \mathtt{D}_3\right) \mathbf{f} = 0 \tag{4}$$

Well-known techniques may be employed to solve this QEP for $\mathbf{f}$ and λ when the matrices are square [12]. These techniques cannot be directly applied, however, in the case where there are more than 9 correspondences, and the design matrices are consequently non-square.

In order to solve the QEP for such over-determined problems, Fitzgibbon [3] obtains the normal equations of Equation 4 through left-multiplication by $\mathtt{D}_1^{\mathtt{T}}$. This technique has the virtue of preserving the true solution in the noiseless case. Empirical results have shown, however, that in the presence of noise the solution

to the normal equations suffers from bias and significant variance. Furthermore, the square problem arising from the normal equations admits 10 general solutions, of which 6, in practice, are real regardless of how overdetermined the system becomes. This is counter-intuitive as oftentimes scene geometry and other constraints should support only a single solution.

It may be surprising that the normal equations have proved so problematic, as they usually provide reasonably good results. For example, the standard eight-point algorithm's residual *is* minimized through the normal equations, specifically by computing the eigenvector with smallest-magnitude eigenvalue. This situation appears to parallel that of the QEP, but there is an important difference. In the eight-point algorithm, the normal equations are constructed from the transpose of the entire matrix factor of $\mathbf{f}$. In the QEP, however, λ is not known, and only the component $\mathtt{D}_1$ of the entire matrix factor $\mathtt{D}_p = \mathtt{D}_1 + \lambda\mathtt{D}_2 + \lambda^2\mathtt{D}_3$ is used. If the radial distortion λ were known, then one could solve the normal equations $\mathtt{D}_p^\mathtt{T}\mathtt{D}_p\mathbf{f} = 0$ to obtain an eigenvector that minimizes the residual of Equation 4. The radial distortion λ is not known, however, and approximate normal equations $\mathtt{D}_1^\mathtt{T}\mathtt{D}_p\mathbf{f} = 0$ are solved instead. It is not surprising that this approximation obtains a biased result.

As a driver for RANSAC, one solves minimal problems, in which the matrices are already square, and the techniques of [3] are appropriate. However, in the case where an accurate $\mathtt{F}$ and λ is required directly from a large set of matchpoints, a new approach is desired.

2.2 An Algorithm for Overconstrained Estimation of $\mathtt{F}$ and λ

If we allow $\mathtt{D}_1$, $\mathtt{D}_2$, and $\mathtt{D}_3$ to be rectangular, the problem is overconstrained and typically there will be no solution in the presence of noise. We therefore construct a minimization problem that defines a suitable approximate solution to the QEP of Equation 4. Because noise corrupts the entries of $\mathtt{D}_1$, $\mathtt{D}_2$, and $\mathtt{D}_3$, it is reasonable to seek a solution which involves perturbing those noisy matrices (hopefully removing the noise) in such a way that an exact solution of the perturbed system does exist. This formulation is a constrained optimization problem in which the perturbed system must satisfy Equation 4 exactly (additionally there is the familiar constraint that the eigenvector $\mathbf{f}$ must be nontrivial). The metric to be minimized is the magnitude of the perturbation, given by $\|\tilde{\mathtt{D}}_1 - \mathtt{D}_1\|_F^2 + \|\tilde{\mathtt{D}}_2 - \mathtt{D}_2\|_F^2 + \|\tilde{\mathtt{D}}_3 - \mathtt{D}_3\|_F^2$, where the perturbed matrices are denoted by $\tilde{\mathtt{D}}_1, \tilde{\mathtt{D}}_2$, and $\tilde{\mathtt{D}}_3$, and $\|\cdot\|_F^2$ is the squared Fröbenius norm.

Obvious approaches to this problem, such as general-purpose minimization via e.g. iterative Levenberg-Marquadt or gradient descent, are unlikely to be satisfactory, because they would be equally suited to an error metric which better represents the statistics of observational error (e.g. the Euclidean reprojection error used in standard bundle adjustment). The expectation is that the above optimization problem, while more descriptive than the normal equations, is still simpler than bundle adjustment in a way that will admit a non-iterative algorithm, or a (more) globally convergent one, or one that is faster.

The rectangular QEP may be converted to a linear rectangular generalized eigenvalue problem through a technique similar to the linearization procedure for the square QEP. A new variable $\mathbf{u} = \lambda\mathbf{f}$ is introduced, obtaining the simultaneous linear matrix equations:

$$\begin{aligned} \mathtt{D}_1\mathbf{f} + \lambda\left(\mathtt{D}_2\mathbf{f} + \mathtt{D}_3\mathbf{u}\right) &= 0 \\ \mathbf{u} - \lambda\mathbf{f} &= 0 \end{aligned}$$

This system of equations may be written equivalently as a single matrix equation

$$\left(\begin{bmatrix} \mathtt{D}_1 & \\ & \mathtt{I} \end{bmatrix} - \lambda \begin{bmatrix} -\mathtt{D}_2 & -\mathtt{D}_3 \\ \mathtt{I} & \end{bmatrix}\right) \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \end{bmatrix} = 0 \tag{5}$$

Here $\mathtt{I}$ is the 9×9 identity matrix.

Recent work studying the eigenvalue problem in the case of non-square pencils has shown that this problem can be solved efficiently, and spurious eigenvalues are avoided [4]. We draw on these results to develop an algorithm that simultaneously estimates radial distortion and epipolar geometry while exploiting the additional information that matchpoints afford.

Let

$$\mathtt{A} = \begin{bmatrix} \mathtt{D}_1 & \\ & \mathtt{I} \end{bmatrix}; \quad \mathtt{B} = \begin{bmatrix} -\mathtt{D}_2 & -\mathtt{D}_3 \\ \mathtt{I} & \end{bmatrix}; \quad \mathbf{v} = \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \end{bmatrix} \tag{6}$$

The problem may then be expressed as finding perturbed rectangular matrices $\tilde{\mathtt{A}}$ and $\tilde{\mathtt{B}}$, an eigenvector $\mathbf{v}$ encoding the fundamental matrix, and an eigenvalue λ determining the radial distortion, which minimizes the quantity $\|\tilde{\mathtt{A}} - \mathtt{A}\|_F^2 + \|\tilde{\mathtt{B}} - \mathtt{B}\|_F^2$ subject to the constraint encoded in Equation 6.

The algorithm is initialized with a choice of $\lambda = 0$. Given λ, an updated estimate of the eigenvector $\mathbf{v}$ is obtained by computing the right singular vector corresponding to the smallest singular value of $\mathtt{A} - \lambda\mathtt{B}$. A refinement of λ is then computed; following the result of [4], this refinement is given by the positive root of the scalar quadratic equation

$$\mathbf{v}^{\mathtt{T}}\left(\mathtt{B}^{\mathtt{T}} + \lambda\mathtt{A}^{\mathtt{T}}\right)\left(\mathtt{A} - \lambda\mathtt{B}\right)\mathbf{v} = 0 \tag{7}$$

This procedure is repeated until convergence. See [4] for a proof that the procedure converges to a local minimum. In our experiments the algorithm has converged reliably and swiftly (typically in less than 20 iterations) to the true minimum.

2.3 Simultaneous Solution for Multiple View Pairs

Estimation of λ from a single view pair fails to exploit all of the available information in the common case where many views are available, all at the same fixed (unknown) lens distortion. Given n pairs of views, and their n sets of correspondences $(\{\hat{\mathbf{p}}_1\}, \{\mathbf{p}_1\}), \ldots, (\{\hat{\mathbf{p}}_n\}, \{\mathbf{p}_n\})$, then the n epipolar constraints may be expressed jointly with a single common lens distortion by the matrix equation

$$\left(\begin{bmatrix} \mathtt{A}_1 & & \\ & \ddots & \\ & & \mathtt{A}_n \end{bmatrix} - \lambda \begin{bmatrix} \mathtt{B}_1 & & \\ & \ddots & \\ & & \mathtt{B}_n \end{bmatrix} \right) \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = 0 \tag{8}$$

Each $\mathtt{A}_i$ and $\mathtt{B}_i$ is obtained from Equation 6 as applied to the correspondence set $(\{\hat{\mathbf{p}}_i\}, \{\mathbf{p}_i\})$. Straightforward application of the algorithm of Section 2.2 to Equation 8 leads to a significant problem, however. If $\mathbf{v}_i$ is a non-trivial null-vector of $(\mathtt{A}_i - \lambda \mathtt{B}_i)$, then $[0 \cdots 0\ \mathbf{v}_i^{\mathsf{T}}\ 0 \cdots 0]^{\mathsf{T}}$ will be a nontrivial eigenvector of Equation 8, for all $i \in 1 \ldots n$, as will any linear combination of such eigenvectors. If the algorithm converges to one of these primitive eigenvectors, then the information present in $\mathtt{A}_j$ and $\mathtt{B}_j$ for all $j \neq i$ is ignored and has no impact on the estimation of the radial distortion parameter λ. It is desirable to force each of the $\mathbf{v}_i$ components of the eigenvector of Equation 8 to be individually normalized (and, hence, nonzero) in order to incorporate as much information as possible into the estimate of λ, and to simultaneously obtain a nontrivial estimate for each $\mathtt{F}_i$, the fundamental matrix for each view pair.

In order to accomplish this task, the algorithm discussed in Section 2.2 is modified. Rather than explicitly constructing the large matrices in Equation 8, it suffices to keep track of the individual $\mathtt{A}_i$ and $\mathtt{B}_i$. As before, λ is initialized to 0. The estimate for the eigenvector, however, is not taken from the SVD of the large system. Rather, each component $\mathbf{v}_i$ is estimated individually from $\mathtt{A}_i$ and $\mathtt{B}_i$. Doing this applies the normalization constraint individually to each $\mathbf{v}_i$. The subsequent update of λ is performed as before, in which the equation to be solved is obtained from the combined aggregate matrices and eigenvector.

It is worth noting that, in addition to the crucial property of ensuring that each $\mathtt{F}_i$ is nontrivial, this algorithm also exploits most of the sparse structure of Equation 8. Updating the eigenvector involves only the small, relatively dense matrices $\mathtt{A}_i$ and $\mathtt{B}_i$, and the computational cost is linear in the number of image pairs. The other operation, defined in Equation 7, does formally involve the large sparse matrices, but the matrix-matrix and matrix-vector products may be implemented straightforwardly to take advantage of the block-diagonal structure of $\mathtt{A}$ and $\mathtt{B}$. Again, the cost is linear in the number of image pairs.

3 Experimental Results

We initially study the algorithm using the controlled conditions of a synthetic dataset. In this dataset, feature points were distributed on a regular grid bounded by the unit cube. Two views, each of 640x480 pixels, of this cloud of feature points were synthetically generated and matchpoints between these views are therefore known. Each synthetic camera observed the origin from a distance of 4 units, and the baseline between the views was 30 degrees. The views were synthetically distorted with a known value of λ. The experiment was intended to serve as a baseline that does not involve potentially noisy estimates of feature location that result from feature extraction on real-world data.

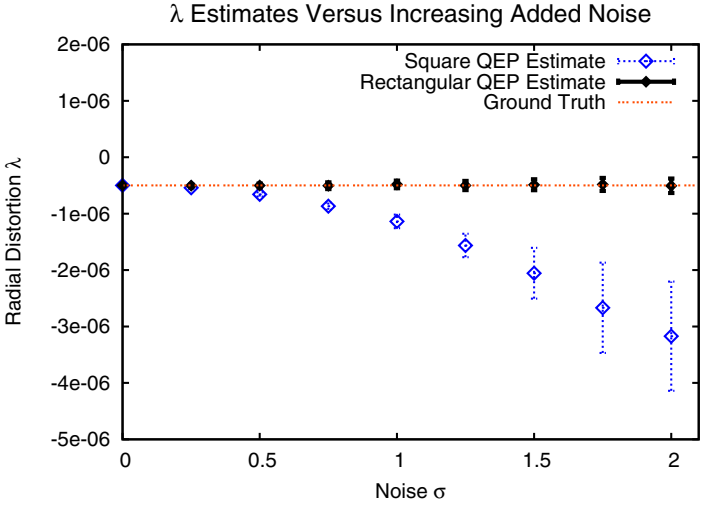


Fig. 1. Accuracy of λ estimation in the presence of increasing Gaussian positional noise. Ground truth is shown as dashed line. Both the rectangular method (solid line) and the square methods (dashed line) are shown for comparison.

The robustness of the estimator with respect to noise was explored by perturbing feature locations with zero mean additive Gaussian noise. Given 75 correspondences under these conditions, $\mathtt{F}$ and λ were estimated. Figure 1 compares the ground truth λ to the estimated λ as noise σ ranged from 0 to 2 pixels. For each noise level, 100 trials were performed and error bars depict one standard deviation. Both the technique described in this work and the method of [3] are shown for comparison.

Notice that the new method exhibits a great reduction in bias at one pixel of error, an amount not uncommon in typical computer vision applications. As error grows as large as two pixels the trend continues.

Given a fixed noise level of 1 pixel reprojection error, it is instructive to study the behavior of the new algorithm as the number of available matchpoints increases. Figure 2 plots λ accuracy as a function of the number of matchpoints used. In this case, random subsets of the available matchpoints were generated over 100 trials for each datapoint. Error bars correspond to one standard deviation.

For purposes of comparison, the new method is also compared to the previously known technique [3]. In order to do so, the earlier approach requires that the rectangular design matrices resulting from the overdetermined set of matchpoints be converted into a square system via the normal equations. Figure 2 shows the behavior of this approach (depicted as a dashed line) as compared directly to the new method (depicted as a solid line).

In the minimal case of 9 correspondences, the two methods produce identical distributions of λ estimates. This is a consequence of two things. First, corresponding trials for the square and rectangular methods received identical input

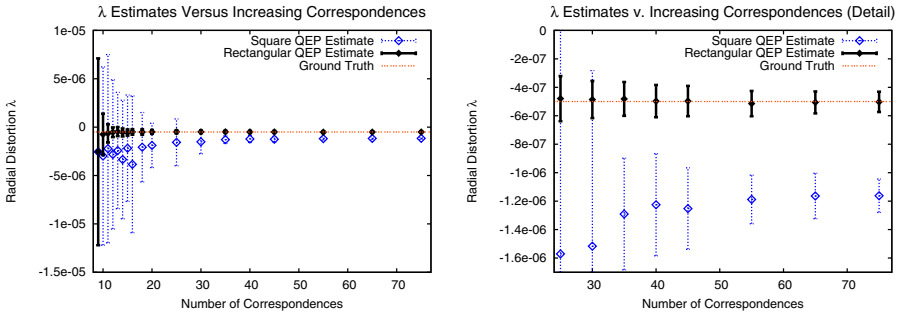


Fig. 2. Accuracy of λ with respect to increasing number of correspondences. (a) A comparison between square (dashed lines) and rectangular QEP (solid lines) from 9 to 75 correspondences. (b) A closeup view of the results. Accuracy from 25 to 75 correspondences.

corrupted by the same noise samples. Second, and more importantly, the rectangular algorithm produces an exact solution which is identical to the result for a square solver for a minimal data set.

Adding just a few additional correspondences dramatically reduces both the bias and variance of the rectangular method, while similar improvements in the results of the square method are not as dramatic. The rectangular method's variance drops significantly below the magnitude of the actual value λ at around 25 correspondences, a point at which the rectangular method's estimate could be said to provide meaningful information. The variance of the rectangular method decreases to about 15% of λ at 75 correspondences, while the square square method's variance is approximately 23% of λ at that point. This difference, while significant, is overshadowed by the dramatic differential in bias at high numbers of correspondences.

3.1 Multiple View-Pair Results

Experiments were performed to provide empirical validation for the case in which a single λ is estimated jointly for multiple view pairs. The setup was similar to the above, except that the baseline for image pairs was reduced to 4 degrees. In all, 8 successive views were generated, and each of the 7 view pairs was obtained via correspondences between adjacent views. Figure 3 shows a plot of the results. The first error bar denotes the mean and standard deviation of λ estimates obtained from 100 trials on the first view pair, each from 75 correspondences corrupted by iid positional Gaussian noise of $\sigma = 1$ pixel. The second error bar represents the results obtained from joint estimation for the first two view pairs; the third bar, from joint estimation for the first three, and so forth.

The variance of the estimates clearly decreases as more pairs are added. The benefits are most dramatic with the addition of the first few view pairs. Although these results may suggest that an online algorithm making use of our new technique could perform well with only a few views, there is no real computational

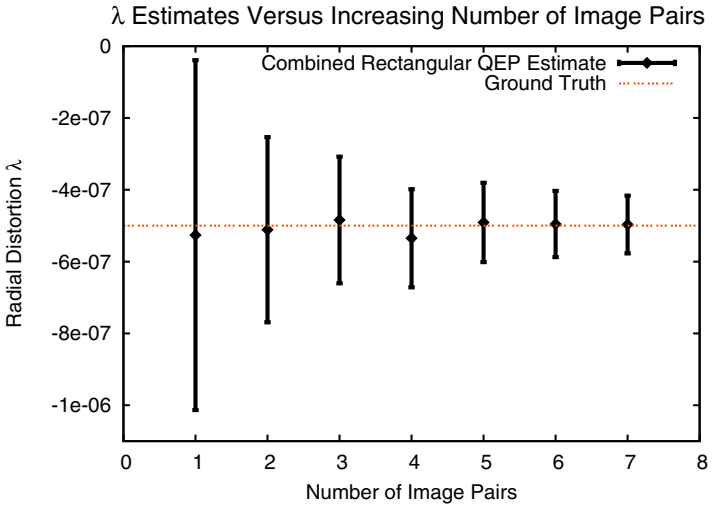


Fig. 3. Accuracy of λ using the rectangular QEP method with respect to increasing number of view pairs. Ground truth is shown as dashed line.

incentive to do so. The computational cost of a single iteration of the multiple view-pair algorithm applied to n pairs is equal to that of the independent algorithm applied separately to the n pairs.

3.2 Real-World Results

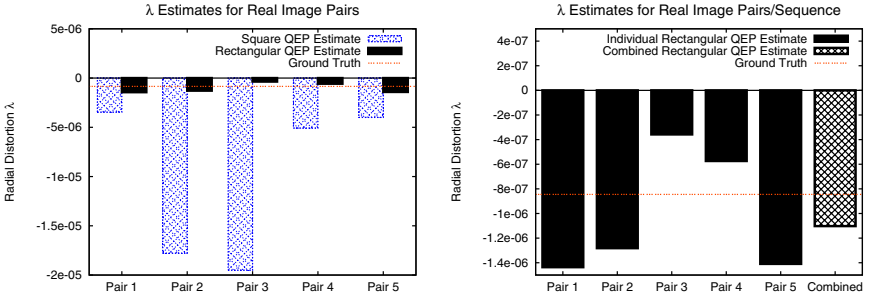
In order to obtain a sense for the algorithm's performance in a practical setting, an experiment was performed on a real image sequence generated from a hand-held camera. Each image was captured at a resolution of 640x480 pixels, and the lens had a nominal focal length of 4 mm. The first and last images in the sequence are shown in Figure 4.

The same camera was also calibrated offline using a well-known method [1]. The iterative technique was constrained to compute the first radial coefficient of the standard multiplicative model. This was then converted to the division model (see Section 2) using standard least squares to obtain a ground-truth estimate of $\lambda = -8.5 \times 10^{-7}$ or a maximum of 54 pixels at the image corner.

The image data was then used to study the behavior of the new algorithm in a real-world context with respect to this ground truth distortion. Proposed matchpoints were generated [13], followed by RANSAC outlier detection based on a square QEP hypothesis generator. An inlier threshold of 1 pixel was employed resulting in an average of approximately 100 inliers from approximately 140 proposals per image pair.

These correspondences were provided to the new method to estimate λ and $\mathtt{F}$ for each image pair independently. In this case, the solution produced by the square method that was known to be closest to the ground-truth estimate was selected as a fair baseline comparison to the new technique. Figure 5 plots the

Fig. 4. First and last images of a real-world sequence used to study the new algorithm. The dataset is composed of six images total captured with a hand-held camera.



(a) (b)

Fig. 5. Accuracy of λ as estimated from each neighboring pair in the real image sequence shown in Figure 4. (a) The square method consistently overestimates distortion, while the new rectangular method obtains dramatically better results. (b) A zoomed-in view of the rectangular estimates in (a), compared with the joint estimate for all 5 pairs, obtained via the multiple view-pairs algorithm of Section 2.3.

distortion estimate (shown as a bar graph for clarity) achieved by both techniques as compared to ground truth (shown as a dashed line).

The results appear to reflect what has already been observed in the simulations. The rectangular method exhibits a large reduction in bias compared to the square method. A close-up view of the results obtained by the new method are shown in Figure 4b. The global estimate of λ derived from all five pairs is also shown. This estimate is more accurate than any of the individual estimates.

4 Conclusion

We have developed a new approach to the simultaneous estimation of radial distortion and multi-view geometry. The method supports any number of correspondences arising from any number of views. In practice, this approach yields

two new algorithms. The first exploits the redundancy of extra correspondences much more effectively than previous methods, while the second introduces an efficient method for estimating a single global λ simultaneously with multiple, independent, multi-view geometries.

These algorithms have been explored in the context of the epipolar geometry, in both simulated and real-world experiments. We find the results demonstrate the striking benefits of the new technique, and lead to more reliable and accurate camera calibration and motion estimation, with a reduced need for a priori knowledge of the scene or camera.

References

1. Tsai, R.: 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 **3** (1987) 323–344
2. Devernay, F., Faugeras, O.: Automatic calibration and removal of distortion from scenes of structured environments. In: SPIE. Volume 2567. (1995) 62–72
3. Fitzgibbon, A.W.: Simultaneous linear estimation of multiple view geometry and lens distortion. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1. (2001) 125–132
4. Boutry, G., Elad, M., Golub, G.H., Milanfar, P.: The generalized eigenvalue problem for non-square pencils using a minimal perturbation approach. (To appear in SIAM Journal on Matrix Analysis and Applications)
5. Micusik, B., Pajdla, T.: Estimation of omnidirectional camera motion from epipolar geometry. In: IEEE Conference on Computer Vision and Pattern Recognition. (2001) 125–132
6. Thirthala, S., Pollefeys, M.: The radial trifocal tensor: A tool for calibrating the radial distortion of wide-angle cameras. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1. (2005) 321–328
7. Stein, G.: Lens distortion calibration using point correspondences. In: IEEE Conference on Computer Vision and Pattern Recognition. (1997) 602–608
8. Claus, D., Fitzgibbon, A.W.: A rational function lens distortion model for general cameras. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1. (2005) 213–219
9. Wright, T.G., Trefethen, L.N.: Eigenvalues and pseudospectra of rectangular matrices. IMA Journal of Numerical Analysis **22** (2002) 501–519
10. Longuet-Higgins, H.C.: A computer algorithm for reconstructing a scene from two projections. Nature **293** (1981) 133–135
11. Hartley, R.: In defense of the eight-point algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 580–593
12. Tisseur, F., Meerbergen, K.: The quadratic eigenvalue problem. SIAM Review **43** (2001) 235–286
13. Nister, D., Naroditsky, O., Bergen, J.: Visual odometry. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1. (2004) 652–659