Practical Improvements to Simultaneous Computation of Multi-View Geometry and Radial Lens Distortion

Ruan Lakemond, Clinton Fookes, Sridha Sridharan
Image and Video Research Laboratory, Queensland University of Technology
GPO Box 2434, 2 George St, Brisbane, Queensland 4001
{*r.lakemond, c.fookes, s.sridharan*}@qut.edu.au

***Abstract*—This paper discusses practical issues related to the use of the division model for lens distortion in multi-view geometry computation. A data normalisation strategy is presented, which has been absent from previous discussions on the topic. The convergence properties of the Rectangular Quadric Eigenvalue Problem solution for computing division model distortion are examined. It is shown that the existing method can require more than 1000 iterations when dealing with severe distortion. A method is presented for accelerating convergence to less than 10 iterations for any amount of distortion. The new method is shown to produce equivalent or better results than the existing method with up to two orders of magnitude reduction in iterations. Through detailed simulation it is found that the number of data points used to compute geometry and lens distortion has a strong influence on convergence speed and solution accuracy. It is recommended that more than the minimal number of data points be used when computing geometry using a robust estimator such as RANSAC. Adding two to four extra samples improves the convergence rate and accuracy sufficiently to compensate for the increased number of samples required by the RANSAC process.**

I. Introduction

A method for simultaneously computing two-view geometry and lens distortion model using point correspondences was introduced in [1]. This method makes use of a single parameter lens distortion model and solves for the geometry and lens in a linear framework. It is useful in solving multi-view geometry problems, such as structure from motion, when it is not possible to properly calibrate the camera and where the images are distorted by significant non-linear lens distortion.

While the solution given in [1] is suitable for minimal solutions (five corresponding points for homographies and nine for epipolar geometry), it does not generalise well to overconstrained problems. An overconstrained solution is presented in [2]. This method involves an iterative process for refining the geometry and distortion coefficient estimates. It is claimed that the solution presented in [2] converges quickly, efficiently and globally and that it is therefore a better solution than alternatives such as Levenberg-Marquardt minimisation on the reprojection error or bundle adjustment. The experimental evidence presented in [2] does not show convincingly that the algorithm indeed possesses these properties over a wide range of conditions.

It is known that the Direct Linear Transform (DLT) approach used to solve two-view geometry problems requires data normalisation [3]. The issue of normalisation in the context of the current problem has not been adequately addressed in the literature. The first contribution of this paper is a suitable normalisation strategy that yields good results and provides a model that is simple to use in practice.

This paper examines the convergence characteristics of the method presented in [2] in Section IV and shows that it can take many iterations to converge for heavily distorted images. In Section V a method is proposed that can reduce the number of iterations to less than 10 in most cases, even in heavily distorted images and with tight convergence constraints. It is shown experimentally in Section VI that this method can be several orders of magnitude faster than that of [2] while yielding superior accuracy. The detailed simulation experiments highlight the sensitivity of these algorithms to noise and to the number of correspondences used in the computation.

The majority of papers on this topic focus on the computation of the fundamental matrix and briefly show how results are applicable to computing a homography. For the sake of variety, this paper will instead focus on the homography.

II. Background

The division model for lens distortion with one parameter is defined as,

$$\mathbf{p} = \mathbf{x} + \lambda r^2(\mathbf{x})\,\mathbf{z},$$

where $\mathbf{x} = \begin{bmatrix} x & y & w \end{bmatrix}^\top$ are the distorted homogenous coordinates with the centre of distortion (COD) at the origin, $\mathbf{p}$ are the undistorted coordinates, $r^2(\mathbf{x}) = w^{-1}\left(x^2 + y^2\right)$ is the distance to the COD squared and $\mathbf{z} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$. In [1] this model is included in the point correspondence relations for a fundamental matrix and for a homography. The homography case is briefly summarised here. Using the discrete linear transform (DLT) method, the homography and lens parameter may be found by solving,

$$[\mathbf{p}_2]_\times \mathbf{H}\mathbf{p}_1 = 0,$$

where $[\mathbf{p}_2]_\times$ is the 3×3 cross product matrix of $\mathbf{p}$. After simplifying so that $w = 1$, expanding the equation and collecting terms, each corresponding point pair yields two equations in the form,

$$\left(\mathbf{D}_1 + \lambda\mathbf{D}_2 + \lambda^2\mathbf{D}_3\right)\mathbf{h} = 0,$$

where the design matrices are of the form,

$$\mathbf{D}_1=\begin{bmatrix} 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1y_2 & y_1y_2 & y_2 \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_2 & -x_2y_1 & -x_2 \end{bmatrix},$$

$$\mathbf{D}_2=\begin{bmatrix} 0 & 0 & 0 & -r_2^2x_1 & -r_2^2y_1 & -r_1^2-r_2^2 & 0 & 0 & r_1^2y_2 \\ r_2^2x_1 & r_2^2y_1 & r_1^2+r_2^2 & 0 & 0 & 0 & 0 & 0 & -r_1^2x_2 \end{bmatrix},$$

$$\mathbf{D}_3=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -r_1^2r_2^2 & 0 & 0 & 0 \\ 0 & 0 & r_1^2r_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and $\mathbf{h}$ is the parameters of the homography, $\mathbf{H}$, in vector form. After combining 9 equations (9 correspondences for the fundamental matrix and 5 for the homography), the design matrices are square and the problem is in the form of a Quadric Eigenvalue Problem (QEP). It is shown that there are at most 6 real solutions to this problem. In the over determined case, the approximate normal equations, $\mathbf{D}_1^\top\left(\mathbf{D}_1+\lambda\mathbf{D}_2+\lambda^2\mathbf{D}_3\right)\mathbf{h}=\mathbf{D}_1^\top\mathbf{D}_\lambda\mathbf{h}=0$ are solved instead. This method is a poor approximation of the proper normal equations, $\mathbf{D}_\lambda^\top\mathbf{D}_\lambda\mathbf{h}=0$ (which cannot be determined without knowing λ) and lead to a biased result.

In [2] the overconstrained case is solved by formulating the problem as a rectangular QEP and linearising the problem through a change of variables. The variable $\mathbf{u}=\lambda\mathbf{h}$ is introduced, yielding the following linear equations:

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

Writing in matrix form gives,

$$\begin{aligned} \left(\begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}-\lambda\begin{bmatrix} -\mathbf{D}_2 & -\mathbf{D}_3 \\ \mathbf{I} & \mathbf{0} \end{bmatrix}\right)\begin{bmatrix} \mathbf{h} \\ \mathbf{u} \end{bmatrix} &= 0, \\ \left(\mathbf{A}-\lambda\mathbf{B}\right)\mathbf{v} &= 0. \end{aligned}$$

The solution is found by iteratively solving for $\mathbf{v}$ in II using singular value decomposition and updating λ by finding the smallest magnitude[1] root of the scalar quadric equation,

$$\mathbf{v}^\top\left(\mathbf{B}^\top+\lambda\mathbf{A}^\top\right)\left(\mathbf{A}-\lambda\mathbf{B}\right)\mathbf{v}=0. \tag{1}$$

It is claimed that this procedure typically converges in less than 20 iterations. This claim is not well supported by the experimental evaluation, since only one synthetic and one real scenario are tested. The convergence rate of this algorithm depends on the amount of distortion and on the data normalisation strategy used, as will be shown in Sections III and IV. One of the contributions of this paper is to explore the convergence properties of this algorithm.

III. Data Normalisation

Data normalisation is a critical step in ensuring good results from the DLT method, however a clear method for normalising data for the simultaneous geometry and lens computation process is not found in the literature. A commonly used approach [3], is to apply a translation,

$$\mathbf{T}\left(t_x,t_y\right)=\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix},$$

and scaling,

$$\mathbf{K}\left(k\right)=\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

to each set of points so that the normalised points ($\mathbf{x}'$) are centred at the origin and have mean distance from the origin of $\sqrt{2}$:

$$\begin{aligned} \mathbf{x}_1' &= \mathbf{K}_1\mathbf{T}_1\mathbf{x}_1, \\ \mathbf{x}_2' &= \mathbf{K}_2\mathbf{T}_2\mathbf{x}_2. \end{aligned}$$

The normalised geometry, $\mathbf{H}'$ or $\mathbf{F}'$, is computed from the normalised correspondences and denormalised according to,

$$\begin{aligned} \mathbf{H} &= \mathbf{T}_2^{-1}\mathbf{K}_2^{-1}\mathbf{H}'\mathbf{K}_1\mathbf{T}_1, \\ \mathbf{F} &= \mathbf{T}_2^\top\mathbf{K}_2^\top\mathbf{F}'\mathbf{K}_1\mathbf{T}_1. \end{aligned}$$

The above method is not well suited to including a division lens parameter in the computation, since it complicates the computation of the radius to the COD and does not account for the scaling of the distortion coefficient. More importantly, the scaling must be the same for both images, since it affects the normalised distortion parameter and both images must have the same distortion. Furthermore, the scale of the normalised distortion has a significant impact on the accuracy of the process. A relatively large distortion value results in slow convergence and reduced stability; a relatively small distortion value results in reduced resolution, depending on machine precision and the thresholds set for determining convergence.

The proposed normalisation consists of a translation to the COD, $\mathbf{T}\left(-x_c,-y_c\right)$, and a scaling $\mathbf{K}\left(k\right)$. This translation keeps the computation of the r^2 term simple, while still nearly centring the data. The scale factor, k, set so that the mean (over both images) distance of the normalised coordinates to the centroid is 1. With this choice of normalisation scale, a severely distorted image will have $\lambda'\in(1.0,0.01)$, i.e. large from a numerical perspective, but still small enough to be correctly determined using equation 1. If it is known that the image is weakly distorted, it is advisable to use a larger value for k in order to increase λ', thereby increasing the precision with which it can be computed. The distorted points are normalisation as,

$$\begin{aligned} \mathbf{x}_1' &= \mathbf{K}\left(k\right)\mathbf{T}\left(-x_c,-y_c\right)\mathbf{x}_1=\mathbf{K}\mathbf{x}_{o1}, \\ \mathbf{x}_2' &= \mathbf{K}\left(k\right)\mathbf{T}\left(-x_c,-y_c\right)\mathbf{x}_2=\mathbf{K}\mathbf{x}_{o2}. \end{aligned}$$

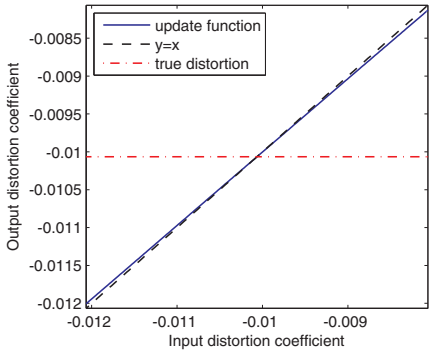
The homography and lens model are computed with the COD at the origin, according to,

$$\begin{aligned} \mathbf{p}_2' &= \mathbf{H}'\mathbf{p}_1', \\ \mathbf{x}_2'-\lambda'r^2\left(\mathbf{x}_2'\right)\mathbf{z} &= \mathbf{H}'\left(\mathbf{x}_1'-\lambda'r^2\left(\mathbf{x}_1'\right)\mathbf{z}\right). \end{aligned}$$

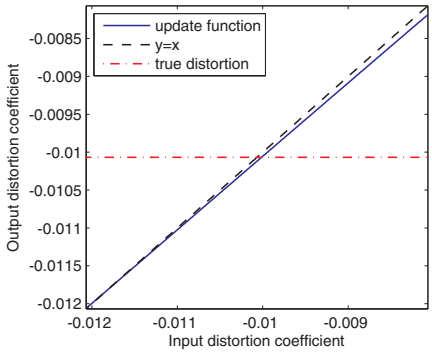
The denormalisation procedure is found by substituting $\mathbf{x}'=\mathbf{K}\mathbf{x}_o$ and collecting terms,

$$\begin{aligned} \mathbf{K}\mathbf{x}_{o2}-\lambda'r^2\left(\mathbf{K}\mathbf{x}_{o2}\right)\mathbf{z} &= \mathbf{H}'\mathbf{K}\mathbf{x}_{o1}-\lambda'r^2\left(\mathbf{K}\mathbf{x}_{o1}\right)\mathbf{H}'\mathbf{z}, \\ \mathbf{x}_{o2}-k^2\lambda'r^2\left(\mathbf{x}_{o2}\right)\mathbf{z} &= \mathbf{K}^{-1}\left(\mathbf{H}'\mathbf{K}\mathbf{x}_{o1}-k^2\lambda'r^2\left(\mathbf{x}_{o1}\right)\mathbf{h}_3'\right), \end{aligned}$$

[1] In [2] it is stated that the positive root of equation 1 is to be used. This is an error, since λ may be negative. Equation 1 typically yields one root with magnitude larger than one and one root with magnitude smaller than one. The magnitude of λ is typically much smaller than 1.



(*a*)



(*b*)

Fig. 1. Response plot or convergence path of the iterative λ and homography update function applied to 1000 normalised correspondences. (*a*) Noiseless case. (*b*) 1% Zero mean Gaussian noise added.

where $\mathbf{h}'_3$ is the third column of $\mathbf{H}'$. This gives the denormalisation by inspection as,

$$\begin{aligned} \mathbf{H} &= \mathbf{K}^{-1}\mathbf{H}'\mathbf{K}, \\ \lambda &= k^2\lambda'. \end{aligned}$$

The counter-translation is not included in the denormalisation since it is redundant for most applications.

Similarly, the denormalisation procedure for the fundamental matrix can be derived as,

$$\begin{aligned} \mathbf{F} &= \mathbf{K}^{\top}\mathbf{F}'\mathbf{K}, \\ \lambda &= k^2\lambda'. \end{aligned}$$

IV. Convergence of the λ Update Process

In [2] it is stated that the iterative procedure for computing the geometry and lens distortion coefficient (see Section II) converges in less than 20 iterations. The convergence properties are not studied in detail and only one geometric scenario is tested. This section presents a detailed study on the characteristics of the method of [2].

Figure 1 plots an example convergence path of the algorithm (output λ plotted against input λ) for a case where 1000 point correspondences are used. Several characteristics of this convergence path are of interest. Firstly, $\lambda_o > \lambda_i$ for $\lambda_i < \lambda_c$ and $\lambda_o < \lambda_i$ for $\lambda_i > \lambda_c$, where λ_i is the input value, λ_o is the output value and λ_c is the convergence point where $\lambda_o = \lambda_i$. This indicates that the process converges monotonically to the point on the graph where the path crosses the $y = x$ line. Secondly, the slope of the line is very close to 1, indicating slow convergence. Ideally the slope would be close to 0, so that any starting value of λ would generate a value close to the final value, but this is unfortunately not the case. Thirdly, where the data is corrupted by noise, the convergence point is shifted away from the true distortion value (the sensitivity to noise depends on the number of correspondences). Lastly, the line is almost exactly straight with slope less than 1. These conditions apply over a large range of input λ values and data conditions and only changes under severe noise or extreme λ values.

The number of iterations required by the iterative update process is therefore approximately linearly dependant on the magnitude of the distortion and may be quite large. Fortunately it is possible to exploit the approximately linear convergence path to predict the convergence point in only a few iterations.

V. An Accelerated λ Update Algorithm

At iteration i the current distortion coefficient estimate, $\lambda_{(i)}$, is used to update the geometry according to equation II and to get an updated distortion estimate, $\lambda_{u(i)}$, according to equation 1. The objective is to estimate $\lambda_{(i+1)}$ that yields $s_{(i+1)} = \lambda_{u(i+1)} - \lambda_{(i+1)} = 0$. This can be done by approximating the update as a linear function,

$$\lambda_u - \lambda = a\lambda - b.$$

The parameters of this linear fit can be extracted from the distortion estimates from the last two iterations,

$$\begin{aligned} s_{(i-1)} &= \lambda_{u(i-1)} - \lambda_{(i-1)} &&= a\lambda_{(i-1)} - b, \\ s_{(i)} &= \lambda_{u(i)} - \lambda_{(i)} &&= a\lambda_{(i)} - b, \\ a &= \tfrac{s_{(i-1)} - s_{(i)}}{\lambda_{(i-1)} - \lambda_{(i)}}, \\ b &= s_i - a\lambda_{(i)}. \end{aligned}$$

The next distortion value is estimated as the point where the update function equals zero,

$$\begin{aligned} a\lambda_{(i+1)} + b &= 0, \\ \lambda_{(i+1)} &= \tfrac{-b}{a}. \end{aligned}$$

The procedure is initialised with $\lambda_{(0)} = 0$ (or any available initial estimate) and $\lambda_{(1)} = \lambda_{u(0)}$. The procedure concludes when $s_{(i)} < \epsilon$. Implementations used for this paper used $\epsilon = 10^{-8}$, which was found to give good precision in combination with the normalisation method defined in Section III, while converging in under 10 iterations in most cases (See Section VI for a detailed evaluation).

The above method will diverge and fail where the slope of the update path is greater than one. To prevent failure in this case, the method reverts to using $\lambda_{(i+1)} = \lambda_{u(i)}$ whenever $\frac{s_{(i)}}{s_{(i-1)}} > 0.999$.

VI. Experimental Evaluation

A simulation was implemented to examine the performance of the algorithm in [2] and the proposed modification presented in Section V. The objective is to compare the two algorithms and to make recommendations on the use of these types of

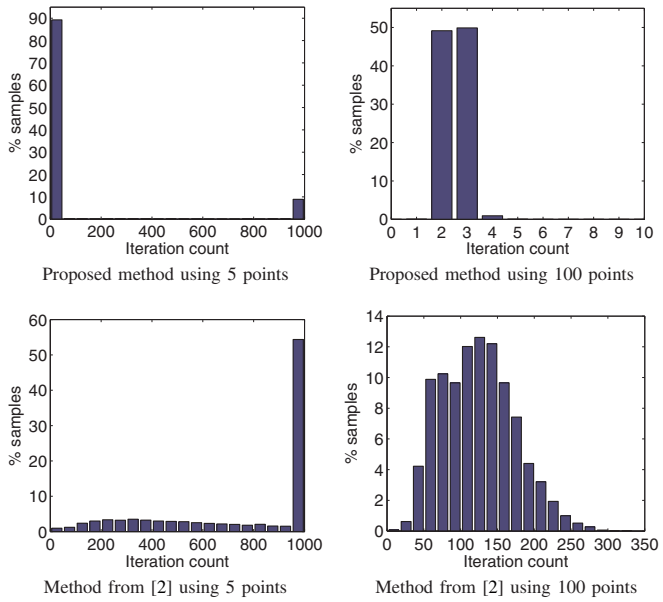


Fig. 2. Histograms of the iteration counts required to complete the Homography and distortion computation process. The lens distortion parameter was set to -0.01 and random noise with a standard deviation of 0.1% of the image width was added.

algorithms in general. The simulation used 100 sets of 1000 randomly generated points with coordinates in the interval $(-1, 1)$. Each point set was mapped to a second image using 100 random homographies and the points were distorted using 7 different distortion coefficients. Zero mean Gaussian noise was added to the distorted points with 10 different variance levels. The geometry and distortion is then recovered using a subsets of the point correspondences ranging from 5 correspondences to all 1000 correspondences. Both algorithms used the normalisation strategy described in Section III. The distortion parameter is initialised to zero at the start of all test trials. A maximum of 1000 iterations are allowed for the geometry computation process and a convergence threshold of 10^{-8} was used. Statistics of the resulting geometry and distortion parameter estimates were collected over the 100 homographies and 100 point sets for each noise, distortion and correspondence count setting. Important results are discussed and summarised in the figures presented in this section.

A. Iteration Count

Figure 2 shows histograms of the number of iterations executed by the geometry computation process in a few example scenarios. The original method of [2] requires a large number of iterations, especially when using small sets of correspondences. This is a result of the relatively small update at each iteration (see Section IV). In contrast, the method proposed in Section V requires several orders of magnitude fewer iterations in most cases. The proposed method typically only requires more than 10 iterations when the update process is not sufficiently stable to predict the convergence point. In these cases, the method of [2] is used until the process begins to converge. Where enough data is available, the proposed method usually completes in fewer than five iteration.

The effect of the number of data points is very clearly shown. Using the minimum of five points does not give a good solution in more than half of the trials for the method of [2]. The proposed method makes a significant improvement, but still fails in 10% of trials. When large numbers of correspondences are available the solution process is much more stable.

Increased noise in the correspondence data typically leads to more iterations. The proposed method completes in less than 10 iterations in most cases, even with severe noise. A distortion value with larger magnitude leads to a linear increase in the number of iterations required by [2]. The severety of the distortion has little effect on the number of iterations required by the proposed method.

B. Noise Sensitivity

Figure 3 plots the 10%, 50% and 90% quantiles of the distortion estimates over a range of noise levels. These quantile plots are used rather than mean and variance plots because the process is prone to outlier results. This result in large variance, which does not give useful information. Geometry estimation algorithms are typically employed in robust estimation algorithms designed to reject outlier data and solutions. The quantile plots give a better indication as to where the majority of solutions lie and are more useful in practice.

The lens distortion estimate is seen to be very sensitive to noise, as previously shown in [1] and [2]. Here it can be seen that the method of [2] yields a slightly biased median result from noisy data. The proposed method shows a similar bias for small distortion values, but yields a much less biased result for large distortion values. Overall, while the distribution of results quickly becomes very large as noise increases, the median result remains reasonably accurate.

C. Number of Corresponding Points

Figure 4 plots the 10%, median and 90% quantiles of the distortion estimates over the number of points used to compute the geometry. Figure 2 shows the iteration counts for two different numbers of correspondences. The number of correspondences clearly has a significant effect on the stability of the geometry computation process. When 100 points or more are used, the convergence is always stable and the proposed method can estimate the convergence point accurately and quickly. With fewer correspondences, the update process is less stable and the incremental update method must sometimes be used. Additionally, the incremental method takes longer to reach convergence or a stable part of the path. Both [2] and the proposed method are affected similarly.

Robust solvers, such as the commonly used Random Sample Consensus (RANSAC) [4], are typically implemented with a kernel that uses the minimal amount of data to estimate trial solutions in order to reduce the number of trials needed to find the best solution. Based on the observations highlighted in Figures 4 and 2, it is recommended that the kernel sample size be increased. Even though this will result in the need for

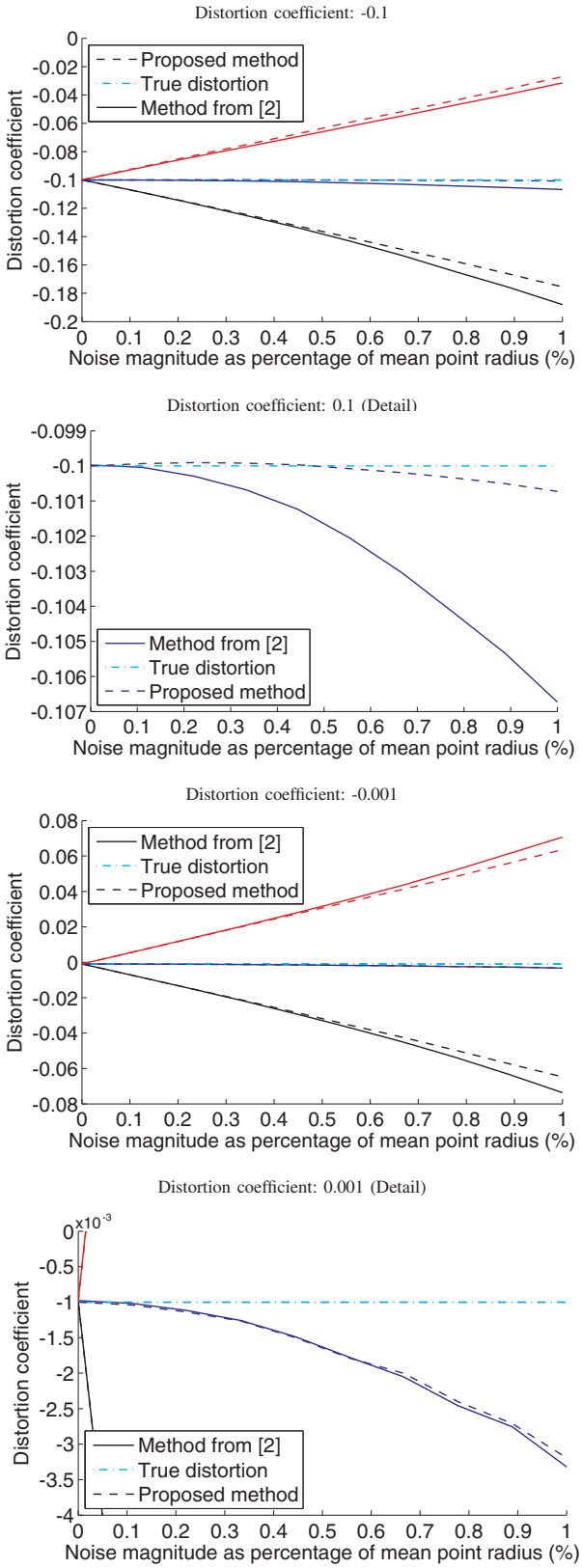


Fig. 3. Plot of the 90%, median and 10% quantiles of distortion coefficients computed using 100 correspondences, plotted over increasing noise.

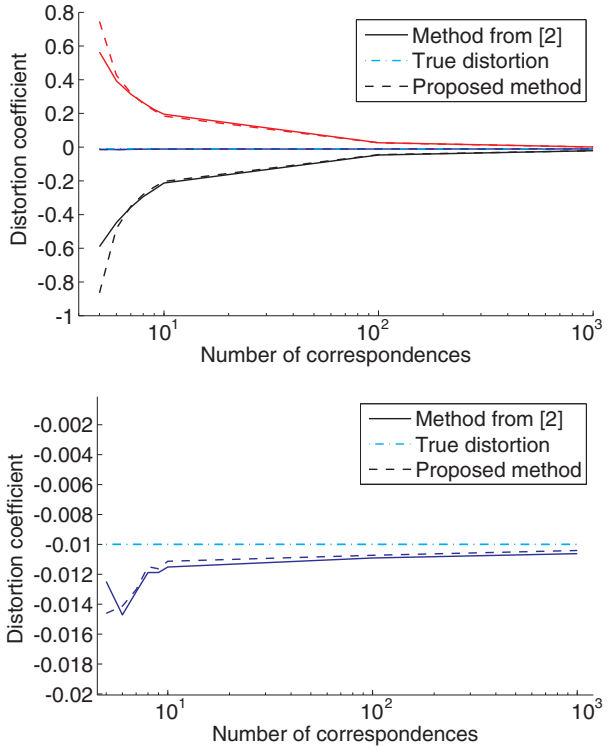


Fig. 4. Plot of the 90%, median and 10% quantiles of distortion coefficients computed using different numbers of correspondences. Both graphs show the same data, each at a different scale.

more trials to be run, the increase in stability and decrease in iteration count will increase the likelihood of finding a good solution in reasonable time.

VII. Real Data

One application of simultaneously computing homographies and lens distortion is in automatically undistorting images taken by a rotating camera. Figure 5 shows some examples. The undistorted images were computed using the following process. Local features were extracted from a sequence of images taken by a rotating camera. The Maximally Stable Extremal Regions extractor [5] and the Scale Invariant Feature Transform descriptor [6] were used to extract features. Features were matched between consecutive frames using nearest neighbour ratio matching [7]. The feature alignment and dense feature extraction method of [8] is then used to get a large set of highly accurate correspondences. A homography and lens model is then computed using RANSAC for each pair of consecutive images using the proposed method. The RANSAC process uses 8 data points to generate solutions instead of the minimum of 5. Though this increases the number of trials required, it greatly increases the stability of solutions. The median lens distortion value from all the estimates is used to undistort the images.

It can be seen from Figure 5 that the distortion can be recovered accurately without any prior knowledge of the lens or camera. The single parameter model yields visually satisfactory results, even for severely distorted images.

Fig. 5. Images distorted by real wide angle camera lenses (left) and automatically undistorted images (right).

VIII. Conclusion

This paper presents a detailed analysis of the performance of algorithms that simultaneously compute multi-view geometry and radial lens distortion using the division model for distortion. Novel additions are made to the existing algorithms to increase computational efficiency and accuracy. A new data normalisation method is presented that is suitable for the problem at hand and allows simple implementation.

It was found that the existing method for overconstrained problems presented in [2] can be slow to converge and can take an impractical number of iterations to reach a good solution if the images are severely distorted. A method is proposed for accelerating the convergence of the algorithm by predicting the convergence point. Due to the linear nature of the convergence path, this method is highly successful and reduces the required number of iterations by as much as two orders of magnitude, depending on the difficulty of the problem.

Detailed simulations show that the slow convergence of the method of [2] leads to poor estimates of the distortion parameter. The method for accelerating convergence proposed in this paper not only reduces computation time, but also improves the estimation accuracy and reduces sensitivity to noise.

Simulation results demonstrate that the accuracy and convergence speed improves dramatically as more correspondences are included in the computation. It is recommended that more than the minimal number of samples be used when implementing this type of algorithm inside a robust estimation system such as RANSAC.

An example application demonstrates how images taken using a rotating camera can automatically be undistorted. The results are visually pleasing and very accurate. These results reaffirm that the single parameter division lens model can be a sufficiently accurate model even when the distortion is severe.

Acknowledgment

This project was supported by Australian Research Council grant number LP0990135. The authors would like to thank Dr Tristan Kleinschmidt for assistance in collecting data.

References

[1] A.W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–125 – I–132, 2001. 1, 4

[2] R. Steele and C. Jaynes. Overconstrained linear estimation of radial distortion and multi-view geometry. In *European Conference on Computer Vision*, volume 3951/2006 of *Lecture Notes in Computer Science*, pages 253–264. Springer, Berlin / Heidelberg, 2006. 1, 2, 3, 4, 6

[3] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, 2 edition, 2003. 1, 2

[4] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 4

[5] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In P.L. Rosin and D.Marshall, editor, *Proc. British Machine Vision Conference*, volume 1, pages 384–393, Cardiff, UK, 2002. 5

[6] D.G. Lowe. Object recognition from local scale-invariant features. In *Proc. International Conference on Computer Vision*, pages 1150–1157, Corfu, Greece, 1999. 5

[7] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 5

[8] R. Lakemond, C. Fookes, and S. Sridharan. Dense correspondence extraction in difficult uncalibrated scenarios. In *Proc. Digital Image Computing: Techniques and Applications*, pages 53–60, Melbourne, Australia, 2009. IEEE Computer Society Conference Publishing Services. 5