Thesis Submitted for the Degree of
Doctor of Philosophy

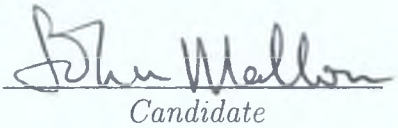
Modelling and Removal of Distortions in Images

Author: John Mallon

Supervisor: Professor Paul F. Whelan

Dublin City University
School of Electronic Engineering
September 2005

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:  ID No.: 97313734

Candidate

Date: 12^{th} September 2005

Acknowledgement

I wish to thank my Ph.D. supervisor Prof. Paul F. Whelan, who afforded me every opportunity, resource and freedom to pursue this work. Also Dr. Ovidiu Ghita for his invaluable help and guidance throughout, all the members of the Vision Systems Group at Dublin City University and Prof. John Barron from the University of Western Ontario for stimulating scientific and non-scientific discussions. I thank my family and Sandra for all their support and healthy distractions.

Contents

Modelling and Removal of Distortions in Images

John Mallon

Abstract

This thesis investigates the compensation and minimisation of distortions in images. Various forms of non-linear lens distortions are modelled and removed. Projective linear distortions are further minimised to give the closest ideal projection from erroneous cameras. Traditional camera calibration treats lens distortion simultaneously with camera pose and lens scaling factors. This often leads to complex algorithms with multi-image requirements, while alternative so called non-metric methods such as straight line techniques, lack ready usability and insufficient precision. Considering the calibration and compensation of non-idealitys separately, allows greater access to error free projections while considerably simplifying subsequent calibrations. The major contributions of this thesis are the precise calibration and removal of lens distortions and the minimisation of perspective distortions. A simple to use technique is proposed for the closed-form calibration of lens distortion based on a single view of a planar calibration chart. Detailed examinations show its accuracy and suitability for all levels of lens distortion. A related method is proposed for the removal of lateral chromatic aberrations in images. Distortion models and approximate inverses are derived to give precise accuracy over all distortion levels including fish-eye lenses. An analysis of calibration patterns is conducted to determine if the choice of pattern can influence the accuracy of the calibration. It is revealed that only specific patterns offer truly bias free control points. Distortion free images are optimally regenerated to minimise pixel scale distortions. This technique is further developed to uniquely minimise perspective distortions, with application to stereo rectification.

Glossary of Acronyms

Acronym	–	Explanation
2D	–	Two Dimensional
3D	–	Three Dimensional
ACA	–	Axial Chromatic Aberration
CA	–	Chromatic Aberration
CCD	–	Charged Coupled Device
DOF	–	Degree Of Freedom
EXIF	–	Exchangeable Image File Format
L-M	–	Levenberg-Marquardt
LCA	–	Lateral Chromatic Aberration
LS	–	Least Square
nc	–	no convergence
ROI	–	Region of Interest
SD	–	Standard Deviation

Chapter 1

Introduction

An optical instrument is required to faithfully produce a geometrically consistent image of a given object, where each point of the latter is imaged as a point in the image. The image is generally formed in accordance with some predefined imaging model, which in this case is assumed to be a projective camera. In this thesis, two related factors that serve to degrade the geometric integrity and quality of an image are considered.

Firstly, the departure of practical optical systems from ideal behaviour, leads to the introduction of aberrations in the resulting images. Two categories of such optical errors are addressed, lens distortion and lateral chromatic aberration. Lens distortion is a well known monochromatic aberration, and is thus present in both colour and greyscale imaging devices. Its nature is predominantly a radial geometric displacement of pixels giving a barrel or a pincushion effect but without loss of image quality. Its chromatic relation, lateral chromatic aberration, arises from the polychromatic nature of light, as it is split into a set of rays or wavelengths upon entering a colour cameras lens. Whilst traversing the optical system light of different wavelengths will follow slightly different paths. Upon reaching the image plane their misaligned recombination introduces chromatic aberration. The focus point of the wavelengths varies both laterally and axially, prompting the distinctions of lateral and axial chromatic aberrations. Lateral chromatic aberration is considered in this work and is characterised by colour dependent shifts in the image plane.

The second source of error arises from the local pixel distortions introduced following aberration removal, and indeed in many image warping applications

such as planar stereo rectification. These distortions are akin to local stretching or warping of pixels, and culminate in both the loss of original image pixels and the degradation of existing ones, through the enlargement or compaction of pixels in the re-sampled image. In addition to poor image quality, geometric distortions may also be introduced. Unlike lens aberrations this form of distortion can, in general, only be minimised.

Optical aberrations form the foundations of this work, from which a selection of related problems are addressed. Through the use of mathematical models, the aberrations, measured at relatively few locations, may be extended to all pixels in the image. To this end, models of these nonlinear optical aberrations are derived from first principles, while their relationship to alternative approximations are shown. A comparison between the derived and existing models is made under the criteria of accuracy and stability. Model inverses are also derived for the specific intent of generating simulation data of high integrity. The aberrations in an image are measured with the aid of planar patterns. Alternative pattern types are analysed to determine which type gives the best quality measurement data. It is found that certain types of patterns and detection methods introduce biasing errors in the recovered image locations. With an appropriate pattern, the measured coordinates are then used to calibrate the aberration models within a least square framework. This calibration is specifically focused on being easy to use and implement, exploiting two geometric priors on the calibration pattern. Using only a single view, a closed-form estimation problem is derived, including full partial derivatives. For lens distortion a comprehensive analysis and comparison of the proposed method is carried out. As the aberration models give no consideration to the optimal formation of new images, a unique means of minimising local pixel distortion is developed. This idea is subsequently expanded to address the unavoidable distortions in planar rectification algorithms.

1.1 Background and Motivation

This section aims to give a brief background to the origins of this work. It also summaries some earlier work that was carried out during this project, but which is not the main focus of this thesis.

1.1.1 Mobile robots

This work originates from work carried out in the area of mobile robotics. A mobile vehicle or platform is a fundamental tool for research and applications in this area. Therefore, a general purpose indoor mobile platform named Mobius (Mobile Vision Autonomous System), shown in Fig. 1.1, was designed and built. Initial design details may be found in Mallon (2001) and subsequently in Mallon et al. (2002*a*). It was designed to sustain agile movement around an indoor laboratory environment using a combination of sensors including multi-camera systems. The design strives to capture the essence of autonomy by ensuring all necessary resources for high level operations are contained onboard the rig. The two most demanding resources, power and computational sources, are included onboard with an additional payload overhead. The drive system comprises of a skid steer system as shown in Fig. 1.2 with encoders for motion control and tracking. Local motion control is effected by two independent micro-controllers with a programmable pole-zero compensator, whose digital outputs are directly interfaced to the stepper motors via digital frequency converters.

Similarly, with all other mobile robot systems the integration based odometric tracking system could not be relied upon for an extended period of time. In an attempt to address this problem, additional 3D sensors, including a stereo vision system, were employed to both help in the fundamental navigation and generate more precise localisation. In a primary study on binocular stereo systems in Ghita et al. (2001), it was found that for many real situations a unique matching solution could not be applied. This forced an increasing dependency on several heuristic constraints to discover plausible matching. Following the addition of a third camera an extra geometric constraint was introduced to reduce the influence of heuristics. The system, reported in Mallon et al. (2002*b*) uses a linear configuration of three equally spaced cameras, mechanically arranged in an effort to align the respective epipolar lines. A feature based matching technique is then applied to calculate a limited set of 3D measures.

It became abundantly clear that there were two large problems with this approach. The first was the virtual impossibility of exactly mechanically positioning two (or one for that matter) cameras relative to the third in the trinocular system. This was further compounded by the slight variations between the three lenses. Secondly, as a broad field of view is required for potential tasks

Fig. 1.1: Mobius, showing its trinocular camera system and payload, (Monitor and optical table)

such as navigation, the use of low focal length lenses with large and variable lens distortions were effectively unavoidable. This distortion effectively removed the possibility of getting any matching points away from the centre of the images. Hence, a journey in camera calibration began.

1.1.2 Calibration: Rectification and Distortion

Without wanting to expend time and resources on the manufacture of precise calibration objects, planar calibration routines based on the work in Zhang (1998), using multiple shots of planar patterns offered an attractive alternative to traditional methods. This offered the correction of lens distortion and an estimate for the camera projection matrices. By decomposing these pro-

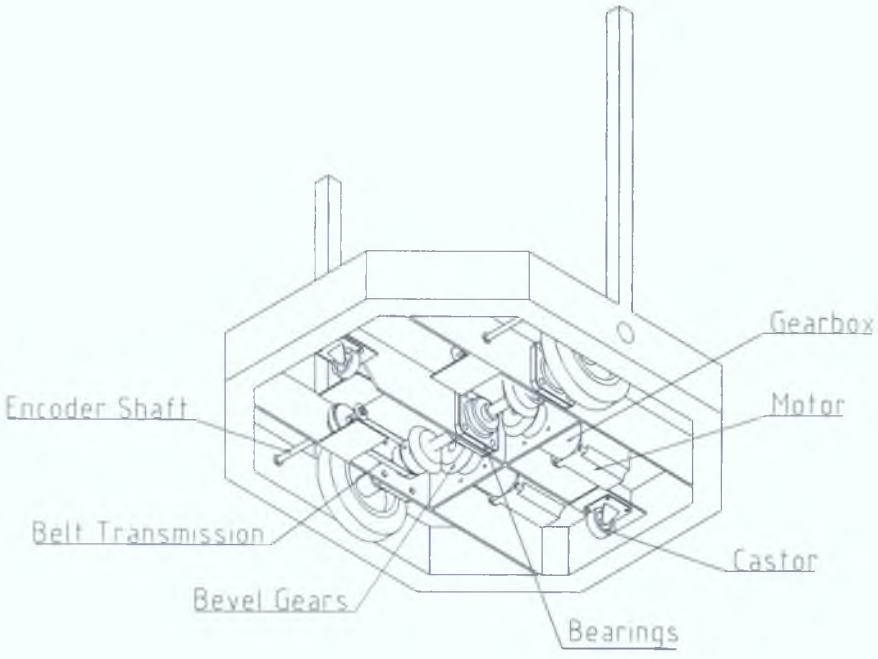


Fig. 1.2: The drive system design of Mobius, viewed from underneath. Stepper motors are coupled to the drive shafts through bevel gears. A toothed belt rotates auxiliary shafts with incremental encoders attached.

jection matrices, rectifying transformations could be obtained, for example using the method proposed in Fusiello et al. (2000). However, this rectification performed poorly, often giving worse alignments than the mechanical setup. As metric rectification was not demanded by the intended application, uncalibrated rectification based on decompositions of the Fundamental Matrix were investigated. Again, using the method proposed in Hartley (1999), the rectification still performed poorly. Additionally, both the calibrated rectification and its uncalibrated relation had a tendency to introduce severe image warping, such as shearing into the new images. Consequently, during the correlation based feature matching or stereo matching phase, the local areas around a point of interest were quite dissimilar, resulting in very few detected correspondences. Regarding the lens distortion removal, it was noticed that the new re-sampled images actually contained a distortion residual of a few pixels.

This thesis details a thorough investigation of these, and related problems. Briefly, these include the uncovering of some useful refinements and simplifications regarding the calibration of distortion. The uncalibrated rectification performance is improved, ultimately matching the level of the noise in

the control points. The unavoidable projective distortions were also uniquely minimised to enhance the multi-image correlation algorithms. From this point some further important issues are addressed and related extensions were developed and tested. These broadly include aberration modelling, the specification of control points for calibration, and the removal of chromatic aberrations in colour images.

1.2 Literature Review

Understandably, a large amount of effort has been directed at these problems. Some of the main publications relevant to this thesis are highlighted, listed under subheadings corresponding to the main chapters. These cover the modelling of optical aberrations in images, calibration methods for these aberrations, the specification of planar control points, chromatic aberration compensations and finally rectification methods and distortions.

1.2.1 Aberration Models

Initially, the photogrammetric community developed methods for modelling and removing lens distortion. Slama (1980) describes the work of Duane C. Brown in modelling lens distortion in arial mapping cameras as the combination of two distinct components, still used and referred to today as radial and decentering lens distortion. These models are functions of the 'plate coordinates' or observed pixel locations in digital terms. The decentering component of this model was further justified by Brown (1966) wherein he advocated the replacement of the thin prism model by a (until then) lesser known model of Conrady (1919) derived by exact ray tracing means. In Brown (1971), formulae for the variation of distortion with focusing distance are revealed. Fryer and Brown (1986) describes some slight modifications to the decentering lens distortion variation with focus distance. These focus variation models are however not used in computer vision as they require the focusing distance to be known in order to correctly select the distortion profile.

The lens distortion model presented in Slama (1980) has become the accepted model in many cases, especially if only low distortion is present. Camera

models for computer vision began including lens distortion factors in an effort to improve accuracy. Tsai (1987) proposed a popular implementation that used only the radial selection of the traditional model as presented in Slama (1980). He commented that a more elaborate model than a plain radial one would not only fail to improve accuracy, but would lead to numerical instability. This was further verified in experiments by Wei and Ma (1994). Truism based methods, generally using straight lines, for example: Prescott and McLean (1997), Swaminathan and Nayar (2000) and Devernay and Faugeras (2001), are forced to use this model as the only data available are the distorted projections.

Despite the apparent widespread adoption of the traditional model, its lack of an analytical inverse makes its use in some calibration techniques awkward. As a consequence its precise form varies from being a function of distorted coordinates (as was originally mooted) to a function of undistorted coordinates. In works such as Weng et al. (1992), lens distortion is presented as a function of undistorted coordinates. Because the undistorted coordinates are unknown, an approximation is made by replacing the undistorted coordinates with distorted ones. He justifies this replacement by reasoning that the re-estimation, carried out with distorted data, will fit equally well. Wei and Ma (1994) also present the traditional model as an approximation of the true one. However, for simplicity, a general third order rational polynomial model is adopted instead. Heikkila and Silven (1997) likewise present the model of lens distortion as a function of undistorted, or in optical terms, Gaussian projections. In Heikkila (2000) the reverse situation is presented, and without any justification, where the same distortion function is now presented as a function of distorted coordinates. Considering the many other conflicting interpretations such as Zhang (1998, 2000), where lens distortion is presented as a function of undistorted data, and Lucchese and Mitra (2003) where it is presented as a function of distorted data, and it is little wonder that in some quarters there is a certain degree of confusion as to the correct interpretation of the lens distortion model and the benefits thereof. Tamaki et al. (2002) identifies this confusion, labelling the models: Distorted-to-Undistorted and Undistorted-to-Distorted.

Alternative functions have been proposed to model lens distortion. In the polynomial approximation vein, Asari et al. (1999) use a general fourth order polynomial to model distortion in endoscopic images, but no analysis of the efficiency of such a model is presented. Shah and Aggarwal (1996) also presents a similar polynomial model including both radial and decentering elements,

while a more general model has been suggested in Kannala and Brandt (2004). Ma et al. (2003) propose a similar model to the traditional one, but to a lower radial order. The advantage gleaned is that a set of inversion solutions become available. A rational polynomial has also been proposed in Heikkila (2000) based on the assumption that distortion follows the traditional form.

Rational models with analytical inverses have also been proposed. Fitzgibbon (2001) presents a single parameter model he calls the divisional model. It offers a good approximation to most distortion profiles, and has been used for high distortion applications in Barreto and Daniilidis (2004), and with slight modifications for fish-eye lenses in Ying and Hu (2004), Brauer-Burchardt and Voss (2001) and Micusik and Pajdla (2003) all reporting adequate performance. The disadvantage of this model is that it cannot model nonlinearities within the distortion profile. Specific fish-eye models have also been proposed. Basu and Licardie (1995) describes a log based model called the Fish-eye transform, which is compared with a polynomial approximation. A hybrid stereographic projection and equisolid angle model has also been proposed for general fish-eye modelling, achieving a sub-pixel fit. However, these models are unsuitable for use with normal perspective camera lens distortion.

Model Inverses

As already described, the lack of an analytical inverse for traditional type distortion models, (both as a function of distorted and undistorted points) is a drawback in many calibration methods. These inverses are required for image or data correction depending on the model assumptions adopted, while also having a important role in the simulation of distortion.

Wei and Ma (1994) propose to use an implicit third order rational polynomial. The resort to such a general model is a direct consequence of the unknown form that a possible inverse might take. Heikkila and Silven (1997) partially address this problem using a fifth order version of this rational polynomial, which is subsequently trimmed of redundant parameters to give a more likely solution form. Good accuracy is presented, but only for very low distortion levels. Its ability to accurately cope outside this range is unknown. A similar form of a suitable inverse approximation has been proposed in Heikkila (2000) based on the inclusion of some terms of a first order Taylor expansion with

the assumed forward distortion model. This model is again tested only for low distortion levels.

Decentering distortion

The practical value of including decentering distortion in the modelling of lens distortion is questionable. Many implementations neglect to model it, seemingly without any adverse consequences. Those that do, find that the associated parameters are very small. Historically, camera calibration authorities in the 1950's had refined their techniques producing accurate estimates for the principal point, but to their dismay discovered that some lenses exhibited an asymmetrical distortion. This was due to slight misalignments of the lens elements and became known as decentering distortion. Initially, a thin prism model was used to model this distortion. With the increase of film resolution and measurement accuracy, Brown (1966) was able to show that the prism model was in exact agreement with the tangential component of decentering distortion but at variance by a factor of three with regard to the radial component. He proposed an alternative model based on previous work of Conrady (1919). Considering that the aerial lenses that were being calibrated could occupy several hundreds of cubic centimeters, were meticulously assembled and extremely expensive, it is unlikely that current low cost, mass produced and small format (e.g C-Mount) are manufactured and assembled to a degree where decentering distortion is negligible.

The ambiguity in the use of decentering distortion originates from the necessity to simultaneously estimate the location of the distortion centre and/or the principal point, with distortion. In the plumb line method of Fryer and Brown (1986) adjustable parameters are carried for the centre point. They found that these parameters were inherently indeterminate if decentering distortion was included in the distortion model. Slama (1980) also offered the opinion that "decentering coefficients also interact to a moderate degree" with the principal point estimation. However, very precise estimates for the principal point were available from the fiducial marks or through laser collimation, allowing informed comparisons to be carried out, e.g. (Brown, 1966). For modern cameras there is no requirement on manufactures to align the lens with the sensor array. Thus without resort to laser collimators the principal point is entirely unknown.

By simultaneously carrying a variable distortion centre point, it was noted by Stein (1993) that a shift in the distortion centre induces decentering like terms in the radial distortion model. This principal is used by many to exclude the explicit modelling of decentering distortion. However, it is unclear how valid this approximation is with increasing decentering distortion. Assuming the existence of decentering distortion, the recovered centre point will thus not match the ideal principal point. The opposite argument is made in Ahmed and Farag (2001) were the centre point is assumed fixed, while decentering elements are included to account for its inevitable misplacement. However, Clarke et al. (1998) showed by experiment that this idea is flawed. He states that the inclusion of decentering can only compensate to a "surprisingly small extent" for shifts in the principal point. Finally, an alternative selection of methods to calculate the principal point are described in Willson and Shafer (1994), while the importance of its estimate in camera calibration is investigated in Hartley and Kaucic (2002) with respect to the estimated focal length.

1.2.2 Calibration Methods

In addition to distortion modelling work, Duane C. Brown also proposed an important method for determining lens distortion based on the truism that straight lines must be imaged as straight lines. This technique, published in (Brown, 1971), and with extensions in (Fryer and Brown, 1986), became known as the 'plumb line' method, where initially fine white thread was stretched by plumb bobs which were stabilised in an oil bath. A comprehensive historical review is given in Clarke and Fryer (1998). This technique was adopted by the machine vision community where simplified versions of the plumb line method are presented, e.g. Prescott and McLean (1997). Haneishi et al. (1995) and Asari et al. (1999) describe a similar truism based correction for the correction of endoscope distortion, using images of co-linear points. Similar, high level distortion is considered in Brauer-Burchardt and Voss (2001), however the real example shown, with a quoted residual distortion of ± 0.7 pixels, resembles a quite benign distortion level. Since these methods only estimate distortion, there are sometimes loosely referred to as non-metric calibration.

An intrinsic problem for these multiple line based methods is that it becomes intractable to form geometric relationships between a distorted line segment and its true projection. An alternating approach is thus employed, as in Dev-

ernay and Faugeras (2001), which iteratively adjusts the distortion parameters in order to minimise the line fitting to the distorted line segments. No meaningful geometric relationship exists between this objective error and the distortion parameters, hence no analytical derivatives are available. This results in slow convergence and can become unstable for elevated distortion levels, unless special steps are taken, as in Swaminathan and Nayar (2000). In this non-metric approach Swaminathan and Nayar (2000) reformulate the objective function in distorted space instead of the usual undistorted space. This is done by performing a further search at each alternation to find the location of a point closest to the considered distorted point, but that lies exactly on the line fitted to the current undistorted point estimates. The reported results show improved robustness to noise for simulated data but no meaningful performance is reported in the real case. A semi-related method has been suggested in Ahmed and Farag (2001) where the curvature of detected lines are used to estimate the parameters of the derivative distortion equation. However, as may be expected, the simulation results show abysmal performance in the presence of noise, while the real results lack a qualitative evaluation.

A more standard means of calibrating distortion is with the simultaneous estimation of a cameras extrinsic and intrinsic parameters. Tsai's method (Tsai, 1987) involves simultaneously estimating, via an iterative numerical optimisation scheme, the single distortion parameter and some internal parameters such as focal length, given the 3D position of a set of control points. The external parameters or position of the camera is already computed in a previous step. The disadvantage of this approach is that it requires known 3D control points and in return offers relatively low accuracy for all but simple distortion profiles. Algorithmic variations on this principal have been proposed by many, including Weng et al. (1992) and Wei and Ma (1994) using more appropriate models for lens distortion. These methods also require known 3D control points. The generation of distortion corrected images is investigated in Heikkila and Silven (1997), while Heikkila (2000) describes a similar technique that requires 3D control points or multiple image sets of 2D control points. An alternative method also based on multiple sets of 2D control points has been advanced in Zhang (1998, 2000) and Sturm and Maybank (1999). This technique addresses distortion through an alternating linear least-squares solution which is then iteratively adjusted in a numerical minimisation including all estimation parameters. Of course the relative complexity of these techniques

is significantly increased by the inclusion of lens distortion.

On the other hand there are many situations where only distortion removal is required, not the full complement of intrinsic and extrinsic parameters. A good example is in the estimation of multiple view geometry in real images, where techniques have been specifically developed to accommodate lens distortion. Zhang (1996) investigates the possibility of simultaneously estimating distortion parameters and the Fundamental Matrix. The results conclude that this is possible if noise is low and distortion is high. Fitzgibbon (2001), Micusik and Pajdla (2003) and Barreto and Daniilidis (2004) use an alternative models for distortion, leading to a polynomial eigenvalue problem and a more reliable estimation of distortion and geometry. Stein (1997) takes the reverse approach and uses the error in Fundamental Matrix estimation as an objective error to estimate distortion parameters.

Alternative methods of distortion calibration exist, where control points correspondences are abandoned in favour of distortion free scenes. These scenes are then imaged by the camera system, whereupon an image alignment process is conducted to correct for distortion. Lucchese and Mitra (2003) describes such a technique, where the distorted image is warped until it registers (in intensity terms) with the reference image. A similar technique using a coarse to fine registration is described in Tamaki (2002) while Sawhney and Kumar (1999) describes a registration technique that does not require an undistorted reference image. Instead, multiple images are registered for the generation of a mosaic, and distortion is simultaneously estimated. These techniques have a very high computational overhead, with twenty minutes quoted in Tamaki (2002).

A final class of non-metric calibration methods are based on distortion induced high-order correlations in the frequency domain. Farid and Popescu (2001) describes such a technique, however its performance is poor in comparison with regular camera calibration techniques and it also appears to be slightly dependent on the image content. Yu (2004) further develops this approach with alternative distortion models and reports accuracy approaching that achieved with regular camera calibration if the source image is of a regular calibration target.

1.2.3 Planar Calibration Targets

There is an abundance of planar charts used within the realms of camera calibration as sources of both 2D and 3D control points. These points are generally constructed on a planar surface by means of some high contrast pattern. In turn, the pattern also facilitates the recovery of the control point projections on the image plane. For example, patterns such as squares in Zhang (1998), Weng et al. (1992), chessboards in Lucchese and Mitra (2002) and circles in Heikkila (2000), Asari et al. (1999) have become popular as they can be readily manufactured to a sufficient precision, and their data points are recoverable through the use of standard image processing techniques.

Naturally, many studies in camera calibration have focused specifically on achieving high calibration accuracy and stability. These works are primarily founded on high precision control points of either 2D or 3D variety, and the accurate detection of their projections. Linear least-square techniques for calibration are improved upon by Tsai (1987) and Weng et al. (1992), who concentrate on improving the calibration accuracy by comprehensively modelling lens distortion and further iteratively optimising the parameters. A comparative study is presented in Salvi et al. (2002). Planar calibration techniques have been proposed by Sturm and Maybank (1999) and Zhang (1998) that place the world coordinate system on the calibration object and thus require only arbitrarily scaled 2D coordinates. These methods, requiring less arduous control point specifications, have contributed largely to the common adoption of planar calibration targets.

All these works assume that the detected image points have zero-mean gaussian distributions in order to correctly converge to the optimal solution through bundle adjustment. Indeed sub-pixel detection methods have been designed for use with specific calibration patterns to give improved accuracy. Peuchot (1992) outlines a method for determining line intersections, while Lucchese and Mitra (2002) describes a local sub-pixel refinement based on surface fitting. The latter is experimentally shown to have zero-mean gaussian errors.

The effects of errors in control points have been investigated in Kopparapu and Corke (1999) where the dependence of camera parameters to inaccurately detected control points are examined. Lavest et al. (1998) advances this problem by considering the error in measured control points. This Error-in-Variables or

Total-Least-Squares approach requires that the errors are random. This may not always be the case. Heikkila (2000) describes a calibration technique using circular control points that are corrected for perspective bias to improve the calibration accuracy. Excluding this one limited case, the biasing influence of the actual calibration patterns (squares, circles, etc.) and associated detection methods have not been addressed so far.

1.2.4 Chromatic Aberration

Chromatic Aberration (CA) can be broadly classified as Axial Chromatic Aberration (ACA) (also known as Longitudinal CA) and Lateral Chromatic Aberration (LCA) (also known as Transverse CA) (Kingslake, 1978). ACA arises from the longitudinal variation of focal position with wavelength along the optical axis. LCA is the variation of image size with wavelength or the vertical off-axis distance of a point from its prescribed point.

Chromatic aberrations have been predominately studied with respect to image formation in the areas of microscopy, photogrammetry and computer vision. Willson (1994) and Willson and Shafer (1991) considers an active lens control system to compensate for chromatic aberration, by separately adjusting three RGB filter lenses to match the colour planes. Their work shows that chromatic aberrations can be compensated in an image by re-alignments of the colour channels. Boult (1992) formulates the compensation of LCA as an image warping problem. No aberration models are employed, focusing solely on the warping problem, and correcting based only on interpolation between control points. Jackowski et al. (1997) presents a similar study on geometric and colour correction in images based on a comparison with a well defined colour calibration chart. The models used are again surface approximations, which are far from optimal solutions, especially since only a limited number of control points are available to estimate the surface parameters. Chromatic aberrations have been addressed by Kuzubek and Matula (2000) where an algorithm for the compensation of both LCA and ACA in fluorescence microscopy is presented, however this technique is not transferrable to images acquired with regular imaging systems. General usage methods, similar to those currently existing for lens distortion, are not available for the calibration of chromatic aberrations.

1.2.5 Rectification distortions

Rectification is known to be a necessary step in stereoscopic analysis. The aligning of epipolar lines allows subsequent algorithms to take advantage of the epipolar constraint, reducing the search space to one dimension. However, the rectilinear mechanical alignment of two (or more) cameras is prohibitively difficult, leading to the development of image warping algorithms to simulate rectilinear images from those of arbitrarily placed cameras. In the uncalibrated case each image can be subjected to a two dimensional projective transformation or planar homography. The homographies can be calculated solely from an analysis of the Fundamental Matrix, to re-orientate the epipolar projections parallel to the horizontal image axis.

Projective rectification has many degrees of freedom. Among these is the problem of finding a rectification that minimises the introduction of distortion in the rectified images. Hartley (1999), Hartley and Zisserman (2003) describes a technique where a rigid rectifying transformation is derived from the Fundamental Matrix. This means that to first order, a points neighborhood undergos rotation and translation only, hence the original and re-sampled images look similar. This criteria is only applied to one of the rectifying homographies, with the result that the second rectified image often contains severe distortions. A related technique has been proposed by Al-Shalfan et al. (2000). Loop and Zhang (1999) consider a stratified decomposition of both rectification homographies in order to minimise projective distortions. This is done by attempting to force affine qualities on the homographies. As image skew and aspect ratio are invariant to affine transforms, they make extra constraints upon the homographies to reduce these distortions. Their approach is not optimal considering only one local region of the image. It is also prone to instabilities when working with real images. Other distortion interpretations have included orthogonality of image corners and maximising image content over the view window (Faugeras and Luong, 2001).

Pollefeys et al. (1999) describe an alternative approach where rectification is considered as a reprojection onto a cylindrical surface instead of a plane, suitable for configurations when the epipole is within or close to an image. Papadimitriou and Dennis (1996) present an approach for convergent stereo geometry, while Isgrò and Trucco (1999) consider rectification directly from point correspondences without explicitly determining the Fundamental Matrix.

These methods however are not focused on the reduction of distortions.

Additionally, the actual rectification performance of many existing planar rectification methods that consider the introduction of distortion, such as (Hartley, 1999, Loop and Zhang, 1999) is often very insufficient when dealing with real images with noisy point correspondences.

1.3 Mathematical notation

Points and vectors are represented by lower case bold symbols, $\mathbf{k} = (k_1, ..., k_n)^T$, with entries k_1, k_2, etc. Point coordinates are predominantly represented in homogeneous form by 3 dimensional vectors, e.g $\mathbf{c} = (u, v, w)^T$. If $w \neq 0$ then this represents the points in $\mathbb{R}^2$ expressed in Euclidean coordinates as $(u/w, v/w)^T$. When the scale has been fixed, i.e. $\mathbf{c} = (u/w, v/w, 1)^T$, these are known as affine points. If $w = 0$, the points are knows as points at ∞ or directions. Points are scale invariant in that $\mathbf{c} = \alpha\mathbf{c}$ ($\alpha \neq 0$). Lines are similarly represented by 3 dimensional column vectors, e.g. $\mathbf{l} = (l_a, l_b, l_c)^T$. Transforms are 3 x 3 matrices of bold uppercase, e.g $\mathbf{T}$, formed of columns $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3]$ with entries $t_{11}, t_{12}, ..., t_{33}$.

Control points detected in a distortion free image are denoted by $\bar{\mathbf{c}} = (\bar{u}, \bar{v}, 1)^T$, where the origin is located in the top left corner of the image. Following a normalisation, these points are referred to as $\mathbf{c} = (u, v, 1)^T$ to reflect the normalisation. The lens centric representation of these points are referred to as $\mathbf{p} = (x, y, 1)^T$, where the origin is located around the intersection of the optical axis and the image array.

The distortion affected counterparts of these points are denoted using a breve, e.g $\breve{\mathbf{p}} = (\breve{x}, \breve{y}, 1)^T$. 2D canonical coordinates of the calibration model are referred to as $\bar{\mathbf{w}}$. The results of fitting the distorted points, $\breve{\mathbf{c}}$ to these model points are referred to as $\bar{\mathbf{c}}$, to distinguish that they do not equal the undistorted points $\bar{\mathbf{c}}$. Lastly the units of detected points are in pixels, referred to as (pix). The units of the normalised control points are referred to as (pix).

1.4 Contributions

In assessing the research described in this thesis, the most important aspects have been identified. The body of work which represents the core of the research effort in this thesis is highlighted. Related work, of lesser impact, but still representing advances in the field are also outlined.

Each of the following topics are addressed in the following chapters and form the backbone contributions of this thesis.

- The models of radial and decentering distortion in an image are derived from fundamental optic equations. In doing so, the apparent conflicting usage of alternative distortion models is resolved. The benefits accruing from its appropriate usage are identified and demonstrated, in comparison with alternative interpretations.
- Regarding the generation of control points for calibration, it is shown theoretically and experimentally that the popular circular type pattern generates an unrecoverable distortion induced bias in the detected control points. This problem has not been previously identified in any of the many calibration articles.
- A highly accurate, non-metric and closed-form calibration method for the calibration of lens distortion is proposed. In contrast with existing methods, it is suitable for use with all levels of lens distortion, is easy to use and implement and requires only a single view of a planar calibration pattern.
- A model based method for the calibration of lateral chromatic aberration is proposed for its compensation. It represents a considerably more accessible method than the few existing approaches.
- A technique is proposed to minimise the projective distortions introduced in planar rectification. It uniquely optimises each transform in order that the rectified images resemble the original images as closely as possible.

The auxiliary contributions are now outlined, which are interspersed throughout the thesis.

- An inverse lens distortion model is derived, displaying much improved accuracy over existing models.
- A model of lateral chromatic aberration is proposed.
- A linear method for computing an affine transform to optimally form a new undistorted image is described.
- An improved method is described for the robust decomposition of the Fundamental Matrix to generate two rectifying projective transforms.

1.5 Thesis Outline

The chapters which follow this introduction are arranged as follows. Chapter 2 describes the origins and the forms of all the aberration models considered. Additionally, an inverse for these functions is proposed. The various models are theoretically and experimentally analysed on real data, highlighting subtleties in their usage.

Chapter 3 aims to investigate if the choice of calibration pattern, and in turn the detection method employed, has any effect on the overall accuracy within calibration. It is found that circular patterns, and those of a similar type, induce a distortion based bias in the detected control points. This comes in addition to a perspective bias. Detailed simulated results confirm the relationship between lens distortion and this bias source, while its magnitude is compared with that of the expected noise and blurring within an image. This bias is finally shown in some real examples, through the use of a hybrid pattern.

Chapter 4, utilising bias free control points, proceeds with the calibration of the lens distortion models presented in chapter 2. A non-metric type solution is proposed to solve the problem, uniquely expressed in a closed-form system. An approach for the generation of distortion free images, using this calibrated lens distortion model, is advocated whereby local pixel distortions are minimised. Comprehensive comparisons between the proposed method and other methods are described, both on extensive simulated data and with real images. The dependence of the method on the assumed geometric constraints is then investigated, considering random and systematic errors in the manufacture of the calibration pattern. An analysis is included regarding the number

of control points required for successful calibration, in addition to the miscalibration resulting from the usage of a bias effected circular type pattern. The accommodation of decentering distortion within the calibration process is highlighted, without recourse to its specific inclusion within the distortion model. Finally, this chapter highlights the violations of a calibrated distortion profile, resulting from the basic lens variables of focusing and aperture settings.

Chapter 5 extends the calibration of lens distortion to the calibration of lateral chromatic aberrations. An efficient means of estimating its presence is described. Following calibration and image correction an evaluation with real images is conducted, clearly identifying the improvement in image quality.

Chapter 6 extends the minimisation of distortions during re-sampling idea, first encountered in chapter 4, to the problem of planar rectification of stereo pairs. The approach uniquely optimises each transform to ensure the rectified images resemble the originals as closely as possible. Additionally, an improved method for the decomposition of the Fundamental Matrix into two rectifying transforms is described. The main body of the thesis is closed in Chapter 7, where a summary of the research work conducted and a review of the results achieved are presented. A list of the publications stemming from this work is also provided.

Chapter 2

Geometric Aberration Modelling

All lens systems introduce a degree of optical error in an image. The departure of practical optical systems from ideal behaviour is known as aberrations. The aberrations affecting images are broadly segregated by the nature of light the lenses are designed to capture. Monochromatic aberrations both deteriorate the image quality (spherical aberration, coma, astigmatism) and deform the image, for example distortion. Colour systems are additionally affected by chromatic aberration, arising from the fact that refracting is a function of frequency or colour. In general is is impossible to design a system which is free from all aberrations. This leads lens manufacturers to consider aberration compensation as an optimisation between different types.

Fig. 2.1: Crop from an image affected with barrel distortion. White dot shows the centre of the original image, about which distortion is radially distributed.

Distortion is a well known monochromatic aberration that affects both colour and greyscale imaging devices. Its nature is a radial geometric displacement of the light ray intersections with the image plane, but without loss of image quality. Physically, it arises from the fact that different portions of the lens have different focal lengths and magnifications. Radial distortion is perceptually categorised as barrel, as demonstrated in Fig. 2.1, or pincushion. Decentering distortion is historically related to the misalignments of individual lens elements and generates both radial and tangential components. Naturally, conducting accurate measurements over such distorted images is impossible without knowing distortion compensation factors.

In a colour camera's lens, polychromatic light is split into a set of rays or wavelengths. Whilst traversing the optical system light of different wavelengths will follow slightly different paths. Upon reaching the image plane their misaligned recombination introduces chromatic aberration. Chromatic aberrations are moving out of the sub-pixel range with the advent of high resolution arrays, giving rise to noticeable colour fringes at edges and high contrast areas. This gives the overall impression of poor quality or definition. Many consumer cameras display this aberration. For scientific applications, it is akin to the effects of colour shifts and blurring, that contravene the imaging models.

This chapter is concerned with the mathematical modelling of distortion and chromatic aberrations in images. A distortion model is derived from first principles for both radial and decentering distortion with the aim of resolving the confusion over the model usage. The theoretical relationships between this model and alternative interpretations are formally defined, while their modelling capabilities are experimentally investigated over a broad range of real lens distortion levels. The derived model is shown to be a good general model, outperforming alternative functions in displaying both stable and accurate performance over the entire range of possible or practical distortion levels.

The absence of a suitable means to approximate lateral chromatic aberration in images is also addressed, in the derivation of an appropriate parametric model. Finally, the problem of determining the possible form of an inverse to the distortion model is addressed. An approximate inverse function is subsequently formulated, giving superior accuracy over existing approaches.

2.1 Lens Distortion

On the image plane, ideal image points are denoted in Euclidian space as $\mathbf{p} = (x, y)^T$ while actual observed points are $\breve{\mathbf{p}} = (\breve{x}, \breve{y})^T$. Within the accuracy of Gaussian optics or perfect projection $\mathbf{p} = \breve{\mathbf{p}}$. Referring to Fig. 2.2, on the plane of the exit pupil ξ and η are x, y measurements related through a constant of lateral magnification to the coordinates of $\mathbf{P}'$. The aberration of the wave elements as a consequence of the preceding optics, causes an optical ray other than the prescribed gaussian one. The wave aberration can be expressed in polynomial form, derived from Seidel perturbation eikonals as (Born and Wolf, 1980):

$$\phi^{(4)} = -B\rho^4 - C\kappa^4 - Dr^2\rho^2 + Er^2\kappa^2 + F\rho^2\kappa^2, \tag{2.1}$$

where $r^2 = x^2 + y^2$, $\rho^2 = \xi^2 + \eta^2$ and $\kappa^2 = x\xi + y\eta$. Each coefficient represents a primary Seidel aberrations: spherical aberration (B), astigmatism (C), field curvature (D), distortion (E) and coma (F). The aberration function is a series approximation of the actual wavefront surface, and further terms can be added to closer approximate the aberrations. Equation 2.1 shows a fourth order approximation.

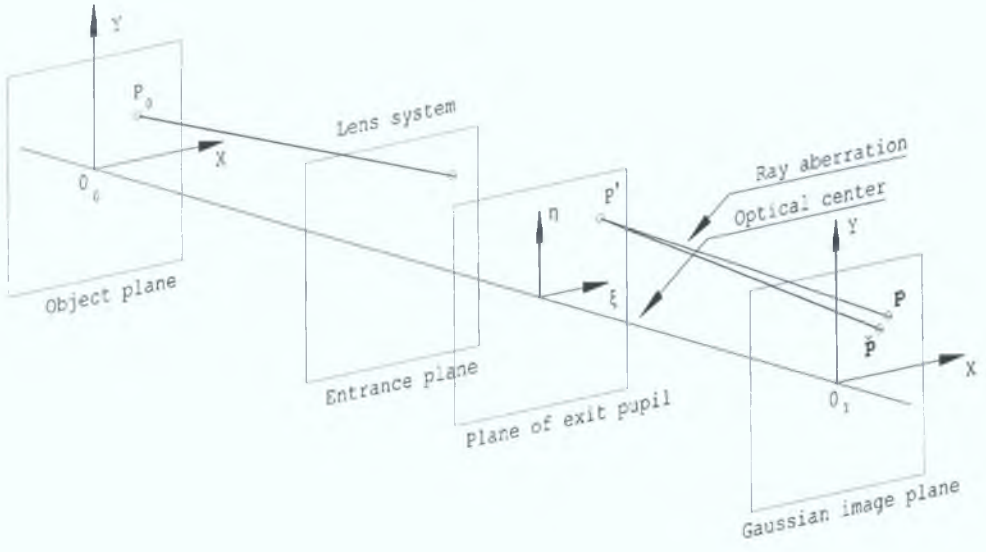


Fig. 2.2: Formation of an image in a general lens system.

In a general system containing a number of surfaces, the primary aberrations equal the sum of the corresponding aberration coefficients associated with the

individual surfaces of the system. By this reasoning aberrations are compensated for optically, by the addition of appropriate extra lens elements. In a two surface system, where (x_1, y_1) represent the space of the intermediately formed image, the combined aberration function for the system is (Born and Wolf, 1980):

$$\phi = \phi_1 + \phi_2 + (x_0 - x_1)(\xi_2 - \xi_1) + (y_0 - y_1)(\eta_2 - \eta_1).$$

If decentering or misalignments of the surfaces is considered, subsequent image deformation may be approximated by perturbing the intermediately formed image by $x_1 \rightarrow x_1 + \lambda$ and $y_1 \rightarrow y_1 + \mu$ as demonstrated in Fig. 2.3. This leads to:

$$\phi = \phi_1 + \phi_2 + \lambda\frac{\partial\phi_2}{\partial x_1} + \mu\frac{\partial\phi_2}{\partial y_1}, \text{ where } \xi_2 - \xi_1 = \frac{\partial\phi_2}{\partial x_1} \text{ and } \eta_2 - \eta_1 = \frac{\partial\phi_2}{\partial y_1}.$$

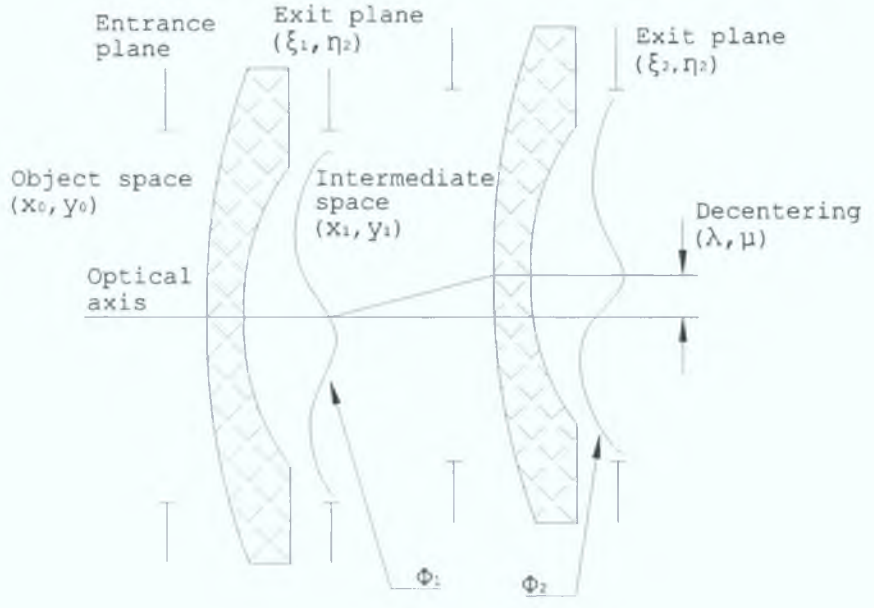


Fig. 2.3: Introduction of decentering lens distortion through misalignments of the optical surfaces.

As the primary interested in distortion aberration the additional aberrations will be ignored. Considering only the distortion component of the wave aberration equation, and replacing the arguments by their gaussian values, the corresponding wave aberration for the combined surfaces to a fourth order approximation is obtained:

$$\phi = k_1 r^2 \kappa^2 + \lambda k_1(\xi(3x^2 + y^2) + 2\eta xy) + \mu k_1(\eta(3y^2 + x^2) + 2\xi xy).$$

The constant $k_1 = E_1 + E_2$ is the sum of the individual lens contributions. The combined decentering effects of multiple lens elements also sums in such a linear fashion, as all arguments are evaluated using their gaussian values. The altered wavefront is the root of all aberrations formed on the image by distorting the ray projections. These ray aberrations are evaluated as the shift from the predicted gaussian coordinates as Born and Wolf (1980):

$$\Delta x = x - \breve{x} = \frac{\partial \phi}{\partial \xi} \quad \text{and} \quad \Delta y = y - \breve{y} = \frac{\partial \phi}{\partial \eta}.$$

Evaluating this using the sixth order approximation of ϕ, results in the combined model for distortion in terms of Cartesian coordinates is:

$$\begin{aligned} \mathcal{D}(\mathbf{p}, \mathbf{k}) &= \begin{pmatrix} \mathcal{D}_x(\mathbf{p}, \mathbf{k}) \\ \mathcal{D}_y(\mathbf{p}, \mathbf{k}) \end{pmatrix} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}, \\ &\approx \begin{pmatrix} k_1 x r^2 + k_2 x r^4 + \lambda k_1 (3x^2 + y^2) + 2\mu k_1 xy + \cdots \\ \quad \lambda k_2 (5x^4 + 6x^2 y^2 + y^4) + \mu k_2 (4x^3 y + 4xy^3) \\ k_1 y r^2 + k_2 y r^4 + 2\lambda k_1 xy + \mu k_1 (3y^2 + x^2) + \cdots \\ \quad \lambda k_2 (4x^3 y + 4xy^3) + \mu k_2 (5y^4 + 6x^2 y^2 + x^4) \end{pmatrix}, \end{aligned} \tag{2.2}$$

where higher orders of λ and μ are ignored. In general for most wide angle lens imaging systems, radial distortion is the predominant observable distortion. For this reason its profile can be more closely approximated by a higher order wave aberration function. Taking a general high order wave aberration approximation $\phi^{(8+)} = k_1 r^2 \kappa^2 + k_2 r^4 \kappa^2 + k_3 r^6 \kappa^2 + \ldots$, results in the general lens distortion approximation:

$$\mathcal{D}(\mathbf{p}, \mathbf{k}) \approx \begin{pmatrix} k_1 x r^2 + k_2 x r^4 + k_3 x r^6 + \ldots \\ k_1 y r^2 + k_2 y r^4 + k_3 y r^6 + \ldots \\ +(p_1(3x^2 + y^2) + 2p_2 xy)(1 + p_3 r^2 + \ldots) \\ +(2p_1 xy + p_2(3y^2 + x^2))(1 + p_3 r^2 + \ldots) \end{pmatrix}, \tag{2.3}$$

In this function the radial component is represented by k_1, k_2 and k_3 while the distortions introduced by decentering correspond to p_1, p_2 and p_3. These are combined into the parameter vector $\mathbf{k} = (k_1, k_2, k_3, \ldots p_1, p_2, \ldots)^T$.

Result 1. *The **Forward** model of distortion, derived from the wave aberration equation is defined as a function of gaussian or undistorted coordinates:*

$$\breve{\mathbf{p}} = \mathbf{p} + \mathcal{D}(\mathbf{p}, \mathbf{k}).$$

Proof. See equations 2.1 through 2.3. □

This model of lens distortion has the same form as the traditional one used by Brown (1971) and many others. The important difference is that this model is a function of gaussian points or ideal distortion free coordinates. It can be expected that since each radial coefficient, $k_1, k_2...$ corresponds to a higher order wave approximation, the dominating terms in the distortion model are the low order coefficients.

The derived decentering model has the same form as that of Conrady (1919) as promoted by Brown (1966). The actual level of decentering distortion in currently used camera systems is questionable, with most works concluding that it is of little significance. This is understandable as it is shown above to be related to the actual misalignments and the level of radial distortion: $p_1 = \lambda k_1$ and $p_2 = \mu k_1$, where λ and μ are the approximations of the resulting x and y shifts. Thus if the radial distortion is small, then the expected decentering distortion should also be small. Higher order approximations of decentering distortion, using p_3 are unlikely to be required, and were not used in Brown (1966). Note that the misalignments of lens elements also introduces elements of other aberrations such as coma, also noted in Conrady (1919), but these are not tracked through in the equations here. Decentering distortion is investigated further in chapter 4.

2.2 Taylor Expansion

Consider the formation of a distorted image coordinates using the forward model in Result 1. This relation is a function of unobservable data $\mathbf{p}$, and for some calibration methods such as those based on the straightness of lines, where the location of the true line is absolutely unknown, it offers no direct route to calibrate distortion [1], as an analytical inverse is not available.

However, taking the Taylor expansion of Result 1, about the known distorted

[1] One indirect means of using this forward model in conjunction with a straight line methodology, is to reformulate the problem in distorted space. This is possible (theoretically at least) if for each distorted line its undistorted counterpart is parameterised, giving two DOF. With one more DOF the exact location of the undistorted point may be solved in order to correctly match the point in distorted space.

locations, $\breve{\mathbf{p}}$ gives:

$$\mathbf{p} = \breve{\mathbf{p}} - \mathcal{D}(\breve{\mathbf{p}}, \mathbf{k}) + \frac{\partial \mathcal{D}(\breve{\mathbf{p}}, \mathbf{k})}{\partial \breve{x}}(x - \breve{x}) + \frac{\partial \mathcal{D}(\breve{\mathbf{p}}, \mathbf{k})}{\partial \breve{y}}(y - \breve{y}) + H.O.T. \quad (2.4)$$

Taking the first term in this Taylor expansion gives:

Result 2. *The **Reverse** model of distortion is defined as a function of distorted coordinates as follows:*

$$\mathbf{p} = \breve{\mathbf{p}} - \mathcal{D}(\breve{\mathbf{p}}, \mathbf{k}).$$

This is the commonly used approximation for distortion in terms of known distorted coordinates and from now on is referred to as the reverse model. It is theoretically possible that this polynomial will approximate the distortion profile equally as well as its forward model counterpart. This is known from the fundamental theorem of approximation theory, due to Weierstrass, which states that on a finite interval, and given an error $\varepsilon > 0$, there exists an algebraic polynomial p for which

$$|f(x) - p(x)| < \varepsilon.$$

However, since the first and higher order terms of the Taylor expansion are neglected, their contribution must be absorbed by the parameters in $\mathcal{D}(\breve{\mathbf{p}}, \mathbf{k})$. Thus, it can be expected that the higher order coefficients, i.e. $k_2, k_3, \ldots$ will assume ever increasing values when fit to a general distortion profile. These large parameter values give rise to poor conditioning in the estimation equations, which in turn lead to sensitive estimates for the distortion coefficients $\mathbf{k}$. The estimated parameters then become increasingly sensitive to perturbations in the raw calibration data, with different levels of either noise or location leading to different estimates for distortion. This situation is clearly undesirable. This explains to some degree why calibration techniques such as Tsai (1987), that use this reverse model of distortion, state that a high order model leads to numerical instability. Finally, a special case is encountered with low levels of distortion. In this scenario high order coefficients in the forward distortion model will have little contribution. Therefore, the higher order terms in the reverse model will also have little influence. The reverse model in this case can be considered equally suitable as the forward model, and gives stable results.

In summary, regarding the polynomial approximation of lens distortion in images, the forward model derived to Result 1 as a function of undistorted

coordinates is theoretically capable of approximating all possible distortion profiles. Its parameters are bounded making it conducive for use in estimation schemes, giving robust parameter estimations. The alterative interpretation of this model as a function of distortion coordinates (reverse model described in Result 2) has been shown to be the first term in a Taylor expansion of the forward model. Theoretically, it has equivalent fitting power, but the solution becomes increasingly unstable with increasing distortion level. It is specific for use only with low distortion levels. These results aim to dispel the confusion over the two different interpretations.

2.2.1 Inverse approximation

Considering the forward model as an appropriate general distortion approximation, the alternative reverse model can be viewed as an inverse approximation to this model. If more terms of the Taylor expansion (equation 2.4) are included a better inverse approximation can be formulated. In this sense, the inverse is formulated by an analysis of the forward model (which is itself only an approximation), and not on the actual lens distortion profile. Such an inverse is useful for reconstructing undistorted points given the forward model as in Heikkila and Silven (1997), and proves very useful for generating synthetic data.

Taking the first two terms of the Taylor expansion (equation 2.4) and rearranging:

$$x = \breve{x} + \frac{-\mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k}) - \frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}(y-\breve{y})}{1+\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}}$$
$$y = \breve{y} + \frac{-\mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k}) - \frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}(x-\breve{x})}{1+\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}}.$$

Eliminating the unknown coordinates by mutual substitution gives:

$$\begin{aligned} x &= \breve{x} + \frac{-\mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k}) - \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}} + \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}}{1+\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}+\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}+\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}-\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}} \\ y &= \breve{y} + \frac{-\mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k}) - \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}} + \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}}{1+\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}+\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}+\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}-\frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}\frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}}}. \end{aligned} \tag{2.5}$$

Ignoring all products in equation 2.5, results in the model proposed in Heikkila

(2000) as:

$$\mathbf{p} = \breve{\mathbf{p}} - \frac{\mathcal{D}(\breve{\mathbf{p}}, \mathbf{k})}{1 + \frac{\partial \mathcal{D}_x(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{x}} + \frac{\partial \mathcal{D}_y(\breve{\mathbf{p}},\mathbf{k})}{\partial \breve{y}}} \tag{2.6}$$

Specifically evaluating equation 2.5 with the distortion, $\mathcal{D}(\mathbf{p}, \mathbf{k})$, approximated radially to fifth order (i.e. $\mathbf{k} = (k_1, k_2)^T$) results in the following [2]:

$$\begin{aligned} x &= \breve{x} - \frac{k_1\breve{x}\breve{r}^2 + k_2\breve{x}\breve{r}^4 + k_1^2\breve{x}\breve{r}^4 + 2k_1k_2\breve{x}\breve{r}^6 + k_2^2\breve{x}\breve{r}^8}{1 + 4k_1\breve{r}^2 + 6k_2\breve{r}^4 + 3k_1^2\breve{r}^4 + 8k_1k_2\breve{r}^6 + 5k_2^2\breve{r}^8} \\ y &= \breve{y} - \frac{k_1\breve{y}\breve{r}^2 + k_2\breve{y}\breve{r}^4 + k_1^2\breve{y}\breve{r}^4 + 2k_1k_2\breve{y}\breve{r}^6 + k_2^2\breve{y}\breve{r}^8}{1 + 4k_1\breve{r}^2 + 6k_2\breve{r}^4 + 3k_1^2\breve{r}^4 + 8k_1k_2\breve{r}^6 + 5k_2^2\breve{r}^8} \end{aligned} \tag{2.7}$$

Equation 2.7 describes the possible form of an inverse approximation to the forward model. However, in this state it is effectively a first order Taylor approximation. In an effort to afford the function some freedom, in the hope of including the influences of the previously ignored higher order terms, the function is adjusted to be linear in parameters:

$$\begin{aligned} \mathbf{p} &= \breve{\mathbf{p}} - \mathcal{D}^*(\breve{\mathbf{p}}, \mathbf{a}), \\ &= \breve{\mathbf{p}} - \frac{\breve{\mathbf{p}}(a_1\breve{r}^2 + a_2\breve{r}^4 + a_3\breve{r}^6 + a_4\breve{r}^8)}{1 + 4a_5\breve{r}^2 + 9a_6\breve{r}^4 + 8a_7\breve{r}^6 + 5a_8\breve{r}^8}. \end{aligned} \tag{2.8}$$

This significantly improves the inverse accuracy, while also allowing the parameters to be linearly estimated. A slightly simplified version of this model, where a_7 and a_8 were not included in the denominator, was used directly on distorted coordinates in Mallon and Whelan (2004) to calibrate distortion.

Given knowledge of the forward model parameters, the unknown parameters are solved using a set of N undistorted tie-points $\mathbf{p}_N$, spread equally over the

[2] Decentering distortion is not considered further in the inverse to simplify presentation complexity. For completeness the evaluation of equation 2.5 including decentering distortion with $\mathbf{k} = (k_1, k_2, p_1, p_2)^T$ results in the addition of the following to the denominator of equation 2.7:

$$\psi = k_3(8\breve{x}+12k_1\breve{x}\breve{r}^2+16k_2\breve{x}\breve{r}^4+4k_3(3\breve{x}^2-\breve{y}^2))+k_4(+8\breve{y}+12k_1\breve{y}\breve{r}^2+16k_2\breve{y}\breve{r}^4+4k_4(3\breve{y}^2-\breve{x}^2)),$$

with the following added to the numerators:

$$\begin{aligned} \varphi_x = (k_3 + 2k_3^2\breve{x})(3\breve{x}^2 + \breve{y}^2) + 2k_4\breve{x}\breve{y}(1 + 2k_1\breve{r}^2 + k_2\breve{r}^4) + (k_1k_3 + k_2k_3\breve{r}^2)(5\breve{x}^2 + \breve{y}^2)\breve{r}^2 \\ +16k_3k_4\breve{x}^4\breve{y} + 2k_4^2\breve{x}(\breve{x}^2 - 3\breve{y}^2), \end{aligned}$$

$$\begin{aligned} \varphi_y = (k_4 + 2k_4^2\breve{x})(3\breve{y}^2 + \breve{x}^2) + 2k_3\breve{x}\breve{y}(1 + 2k_1\breve{r}^2 + k_2\breve{r}^4) + (k_1k_4 + k_2k_4\breve{r}^2)(5\breve{y}^2 + \breve{x}^2)\breve{r}^2 \\ +16k_3k_4\breve{x}\breve{y}^4 + 2k_3^2\breve{x}(\breve{y}^2 - 3\breve{x}^2). \end{aligned}$$

entire image surface, and a corresponding set of distorted points $\breve{\mathbf{p}}_N$. The system of equations are formed as $\mathbf{e} = \mathbf{Ta}$, where $\mathbf{e} = (x_1 - \breve{x}_1, y_1 - \breve{y}_1, \ldots, x_i - \breve{x}_i, y_i - \breve{y}_i, \ldots, x_N - \breve{x}_N, y_N - \breve{y}_N)^T$ and

$$\mathbf{t}_{xi} = (\breve{x}_i\breve{r}_i^2, \breve{x}_i\breve{r}_i^4, \breve{x}_i\breve{r}_i^6, \breve{x}_i\breve{r}_i^8, 4e_{xi}\breve{r}_i^2, 9e_{xi}\breve{r}_i^4, 8e_{xi}\breve{r}_i^6, 5e_{xi}\breve{r}_i^8)^T$$
$$\mathbf{t}_{yi} = (\breve{y}_i\breve{r}_i^2, \breve{y}_i\breve{r}_i^4, \breve{y}_i\breve{r}_i^6, \breve{y}_i\breve{r}_i^8, 4e_{yi}\breve{r}_i^2, 9e_{yi}\breve{r}_i^4, 8e_{yi}\breve{r}_i^6, 5e_{yi}\breve{r}_i^8)^T$$
$$\mathbf{T} = (\mathbf{t}_{x1}, \mathbf{t}_{y1}, \ldots, \mathbf{t}_{xi}, \mathbf{t}_{yi}, \ldots, \mathbf{t}_{xN}, \mathbf{t}_{yN})^T$$

The parameter vector is now estimated in a least squares sense as:

$$\hat{\mathbf{a}} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{e}. \tag{2.9}$$

Finally, the further simplification of equation 2.7 by taking only the k_1 terms results in:

$$\mathbf{p} = \breve{\mathbf{p}} - \frac{\breve{\mathbf{p}}k_1\breve{r}^2}{1 + 4k_1\breve{r}^2},$$

and replacing the four in the denominator by one, the divisional model as described in Fitzgibbon (2001) and Micusik and Pajdla (2003) is arrived at:

$$\mathbf{p} = \frac{\breve{\mathbf{p}}}{1 + k_1\breve{r}^2}, \text{with associated inverse } \breve{\mathbf{p}} = \frac{1 - \sqrt{1 - 4k_1r^2}}{2k_1\mathbf{p}}. \tag{2.10}$$

In summary, an inverse approximation to the forward model has been proposed, based on the form of its Taylor expansion. This may be directly used on distorted coordinates, but offers more potential in the generation of synthetic images, as chapter 3 demonstrates. The relationship of an alternative model proposed in Heikkila (2000) to the Taylor expansion of the forward model is also presented. As a byproduct, the origins of the divisional distortion model were also shown to be closely linked to this Taylor expansion.

2.3 Lateral Chromatic Aberration

Chromatic Aberration (CA) can be broadly classified as Axial Chromatic Aberration (ACA) (also known as Longitudinal CA) and Lateral Chromatic Aberration (LCA) (also known as Transverse CA). ACA arises from the longitudinal variation of focal position with wavelength along the optical axis. LCA is the variation of image size with wavelength or the vertical off-axis distance of a

point from its prescribed point. In an image it is identified by a radially dependent misalignment of the colour planes. Considering the modelling of LCA, it can be specified from the contributions of two separate factors: the chromatic variation of distortion and lateral colour distortion as outlined in Kingslake (1978)

Monochromatic aberrations such as distortion, are in general not largely affected by polychromatic light. The chromatic variation of distortion is however detectable, especially on large pixel arrays. This distortion is lateral in nature and can be modelled comprehensively by a simplified forward distortion model. In many cases a significant decentering distortion is apparent.

In addition to the chromatic variation of distortion there is an additional lateral colour distortion that is due to the refraction index variation of the lens elements. The refraction index is quite linear within the visible spectrum (Kingslake, 1978), resulting in the addition of an extra first order term that does not appear in the chromatic distortion equation. Deviations from linear behaviour are naturally accounted for in the chromatic distortion equation. Thus, the combined LCA for a specific frequency or colour plane (g), can thus be modelled as a function of another frequency (f) by the addition of the chromatic variation of distortion and the lateral colour distortion as:

$$\begin{aligned} \mathcal{C}^g(\mathbf{p}_f, \mathbf{c}_g)_x &= c_1 x_f + c_2 x_f r_f^2 + c_3(3x_f^2 + y_f^2) + 2c_4 x_f y_f \\ \mathcal{C}^g(\mathbf{p}_f, \mathbf{c}_g)_y &= c_1 y_f + c_2 y_f r_f^2 + 2c_3 x_f y_f + c_4(3y_f^2 + x_f^2), \end{aligned} \tag{2.11}$$

where $\mathbf{c}_f = (c_1, c_2, c_3, c_4)^T$ is the parameter vector.

2.4 Experiments: Distortion Modelling Performances

This section aims to evaluate the performance of various distortion modelling functions. The performance and suitability of a particular model is characterised by measuring its residual fitting error and examining its estimation integrity by looking at the scale and uncertainty of the parameters. The models are fit to a sample set of increasingly distorted images, using the iterated least square method of Levenberg-Marquardt (L-M) (Walter and Pronzato, 1997), with Jacobi estimated from finite differences. The models are fitted, as

in chapter 4, using an image of a chessboard target, who's ideal projection is simultaneously estimated. The residuals are then evaluated on separate images, with the same camera settings. A selection of the calibration and evaluation samples are illustrated in Fig. 2.4.

The chessboard intersections are calculated using edge intersections within a small local neighborhood around the control point, (refer to chapter 3). Sample numbers 1 and 2 were taken with a Fuji FinePix 6900, with 6MegaPixel resolution. Sample number 2 used an additional wide angle lens adaptor (Fuji WL-FX9). Sample numbers 3 through 12 were taken with Nikon fish-eye lens (FC-E8) attached to a Nikon Coolpix E4500 with 4MegaPixel resolution. Distortion was varied by incremental zooming.

Only radial components are considered for fitting, while decentering modelling is examined in conjunction with the centre of distortion in Chapter 4. The radial distortion models considered in detail are the:

- Derived forward model of Result 1, for both four and five radial parameters.
- The reverse model of Result 2, estimated for four and five radial parameters
- A general fourth order polynomial function similar to Asari et al. (1999) with the form: $r = \breve{r}+k_1\breve{r}+k_2\breve{r}^2+k_3\breve{r}^3+k_4\breve{r}^4$, $\theta = \breve{\theta}$, where $\breve{r} = (\breve{x}^2+\breve{y}^2)^{1/2}$ and $\breve{\theta} = \text{atan2}(\breve{y}, \breve{x})$.
- The divisional model of equation 2.10.

The estimation procedures proceed by defining a minimisation criterion for each control point i as e_i. The control points are firstly pre-scaled by the average of the image width and height. Parameter uncertainties are estimated by calculating the Fisher information matrix as:

$$\mathbf{F}(\mathbf{k}) = \sum_{i=1}^{n_t} \frac{1}{\hat{\sigma}^2}\mathbf{H}(e_i, \mathbf{k}), \text{ where } \hat{\sigma}^2 = \frac{1}{n_t - n_p}\sum_{i=1}^{n_t} e_i^2$$

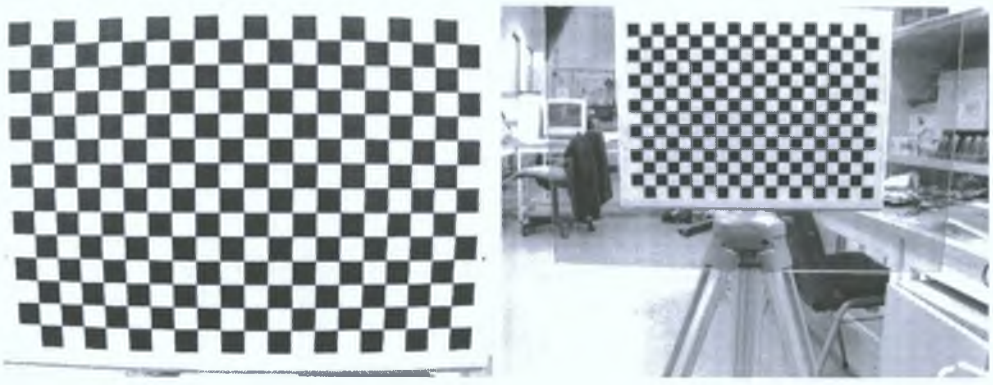
and $\mathbf{H}(e_i, \mathbf{k})$ is the Hessian matrix of the system upon convergence, with n_t the number of control points and n_p the total number of estimated parameters. The standard deviation of each parameter can be computed as:

$$\hat{\sigma}_{\mathbf{k}}^2 = \text{diag}(\sqrt{\mathbf{F}^{-1}(\mathbf{k})})$$

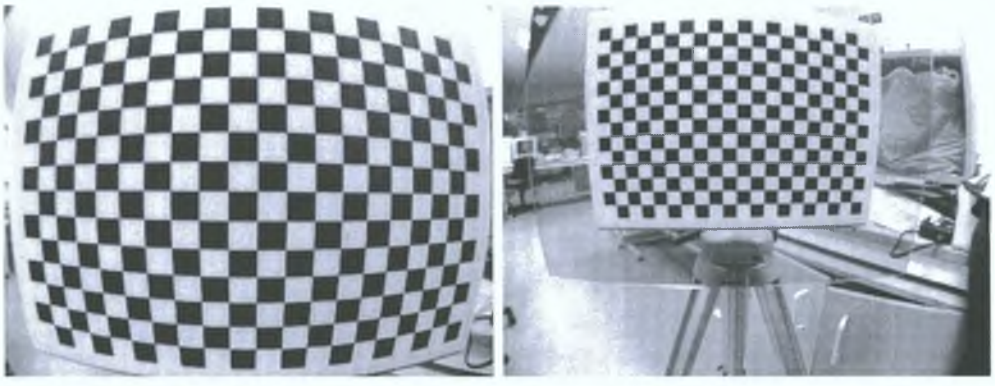
Full details of the estimation procedure for the forward model are described in Chapter 4

Fig. 2.5 shows the parameter estimates and uncertainties for the forward model with four parameters, when applied on the sample set of real distortion profiles. This shows the higher order coefficients of the forward model increasing with distortion. The associated parameter uncertainties however remain relatively constant for all distortion values. This indicates that the system of equations are non-singular, leading to stable estimation of the parameters. The magnitude of the parameters can therefore be considered bounded, depending only on the pre-scaling applied on the control points.

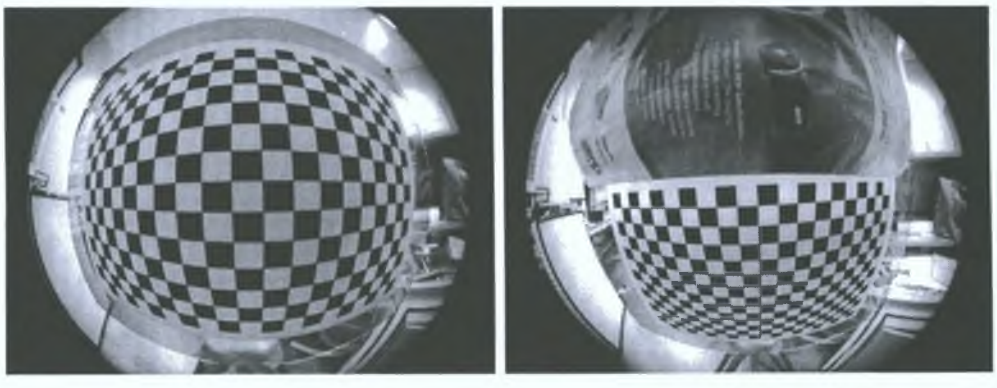
The alternative reverse model parameter estimates and uncertainties are illustrated in Fig. 2.6, also for a four parameter model. In contrast with the forward model the magnitudes of the estimated parameters rapidly increase with increasing distortion. The parameter uncertainties also steadily increase to much greater magnitudes, indicating that the system of equations in this case are likely to become unstable or singular. Indeed, this is the case for sample numbers ten through twelve where the estimation routine failed to properly converge.



(a) Sample number 1



(b) Sample number 6



(c) sample number 12

Fig. 2.4: The first, middle and last calibration and evaluation images used to evaluate the various distortion models.

Table 2.1: Estimated parameters for the forward and reverse models on two sets of calibration data.

Forward Model					
	k_1	k_2	k_3	k_4	k_5
Image 1	-0.4207	0.2446	-0.1095	0.0290	-0.0033
Image 2	-0.4177	0.2365	-0.1014	0.0255	-0.0027
Difference	-0.0030	0.0081	-0.0081	0.0035	-0.0006
Reverse Model					
Image 1	0.4537	0.0922	0.7604	-0.8773	0.6308
Image 2	0.4737	0.0037	0.8985	-0.9075	0.5823
Difference	-0.0200	0.0885	-0.1381	0.0303	0.0485

These practical differences are further illustrated in figures 2.7 and 2.8, where the parameter vector is increased to five coefficients. The results are compiled only for samples where the reverse model converged, i.e. samples one through nine. The addition of the extra parameter causes the higher order reverse model parameters and uncertainties to rapidly increase in value, as Fig. 2.8 shows. However, the addition of the extra parameter to the forward model does not induce this instability, as demonstrated in Fig. 2.7.

The invariance of the forward and reverse model to changes in the location of the control points is now investigated. Two images are taken from slightly differing viewpoints as Fig. 2.9(a) shows. The models are estimated with both sets of data and the respective variation in parameters are compared. Table 2.1 shows the parameter estimates and differences, with the reverse model parameters varying considerably in comparison with the forward model parameters. Fig. 2.9(b) illustrates the effect of these parameter variations, by comparing the respective distortion profiles. This shows that the reverse is significantly more affected by slight changes in control points than the forward model, making it less attractive for practical usage.

A general fourth order polynomial, similar to those proposed in Asari et al. (1999) and Shah and Aggarwal (1996) is also evaluated for comparison purposes in Fig. 2.10. This model is a function of distorted coordinates, with both the parameter values and uncertainties steadily increasing with distortion. It can be noticed that certain parameters contribute little to the modelling, e.g. k_1. The divisional model, having only one parameter exhibits good stability, making it suitable for use with all distortion levels.

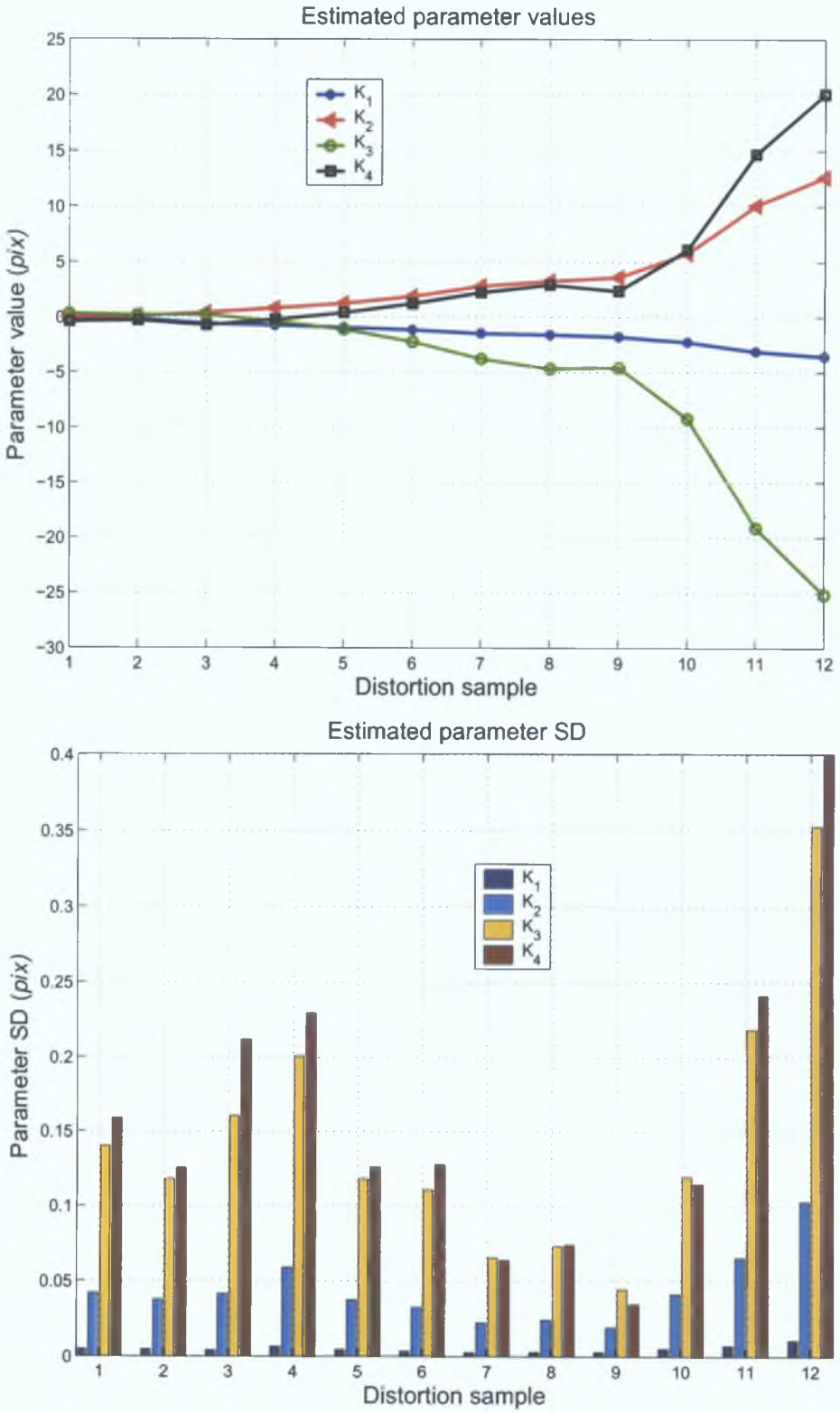


Fig. 2.5: Forward model parameters, $\mathbf{k} = (k_1, k_2, k_3, k_4)^T$, and uncertainties when applied to sample set of increasingly distorted real images.

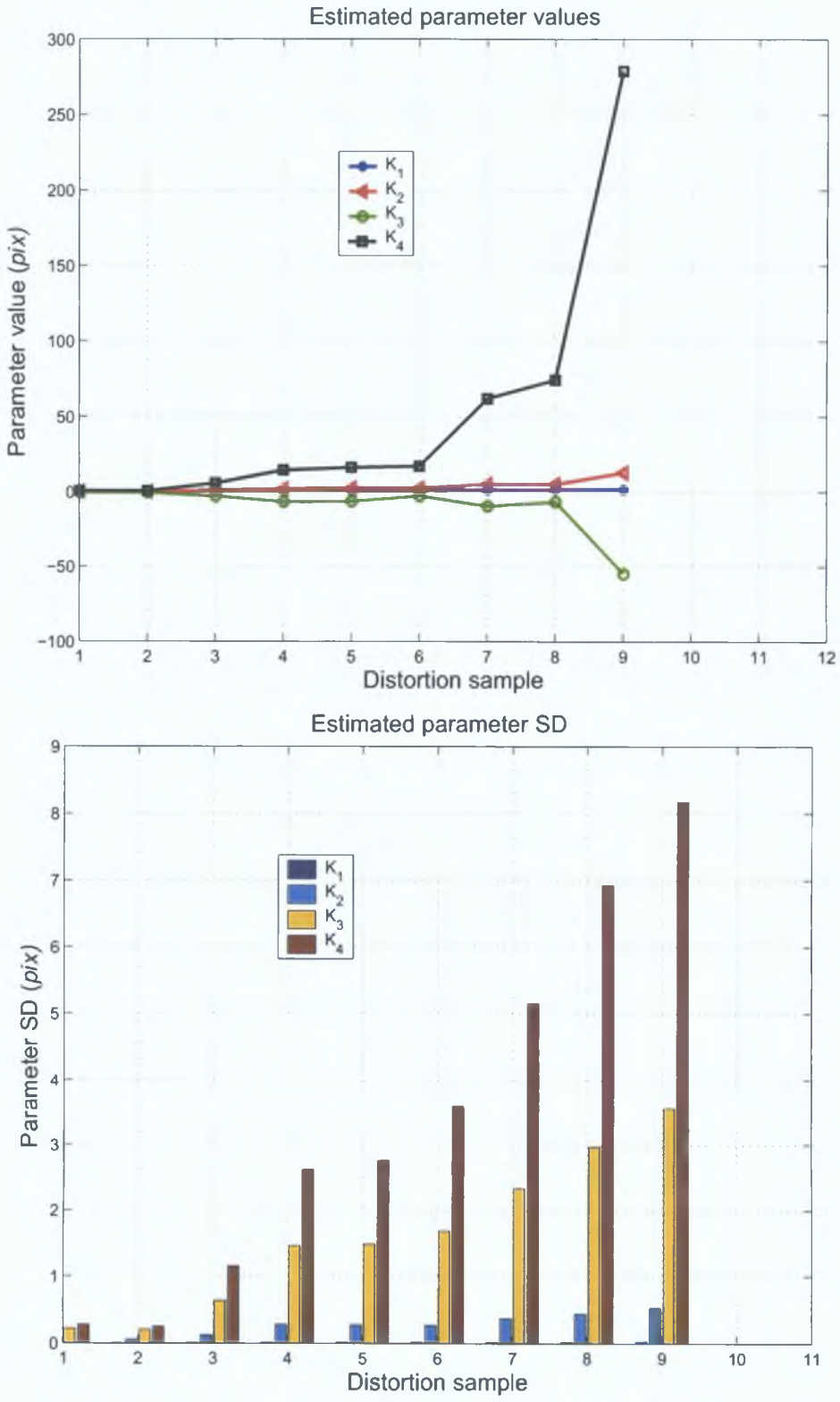


Fig. 2.6: Reverse model parameters, $\mathbf{k} = (k_1, k_2, k_3, k_4)^T$, and uncertainties when applied to sample set of increasingly distorted real images.

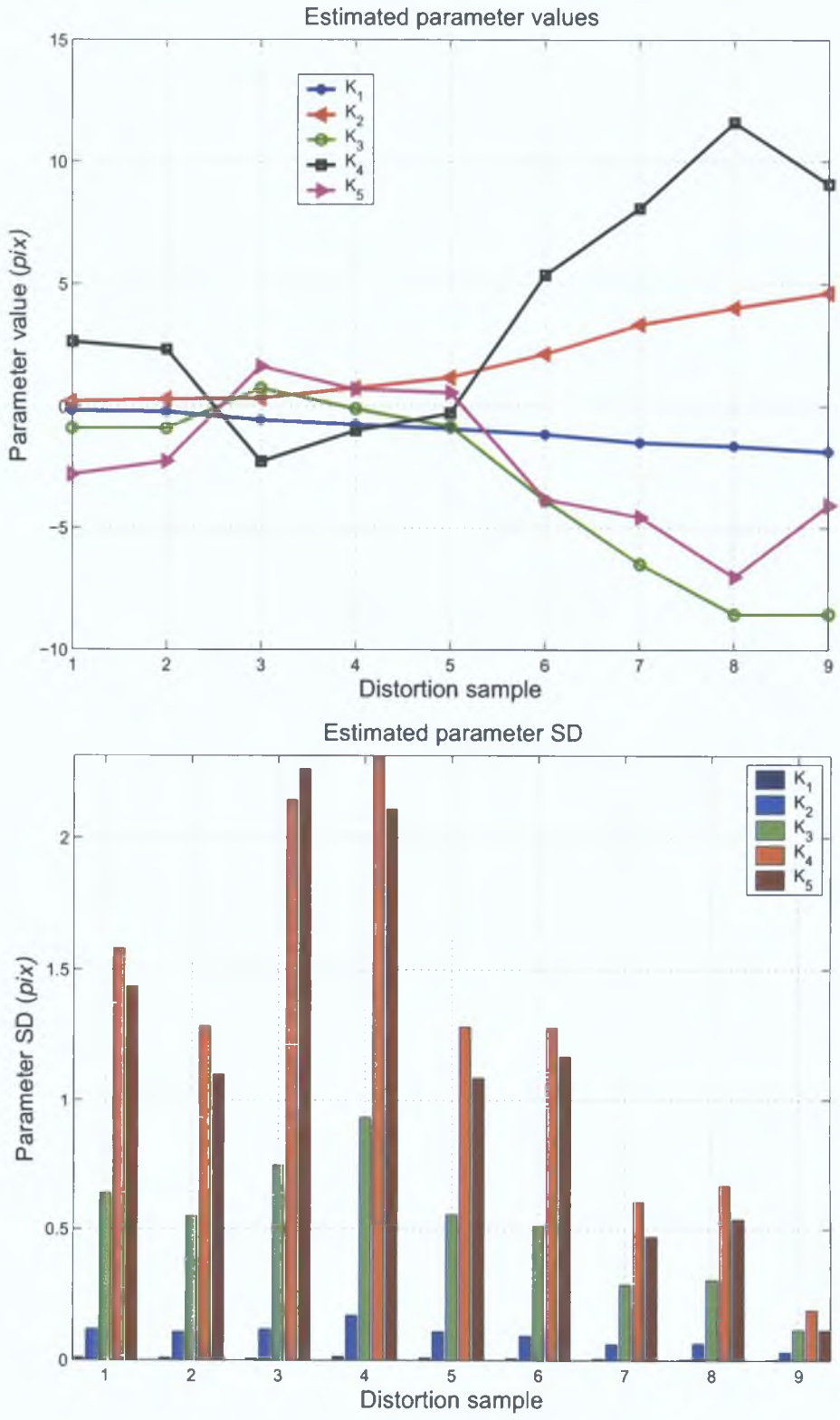


Fig. 2.7: Forward model parameters, $\mathbf{k} = (k_1, k_2, k_3, k_4, k_5)^T$, and uncertainties when applied to sample set of increasingly distorted real images.

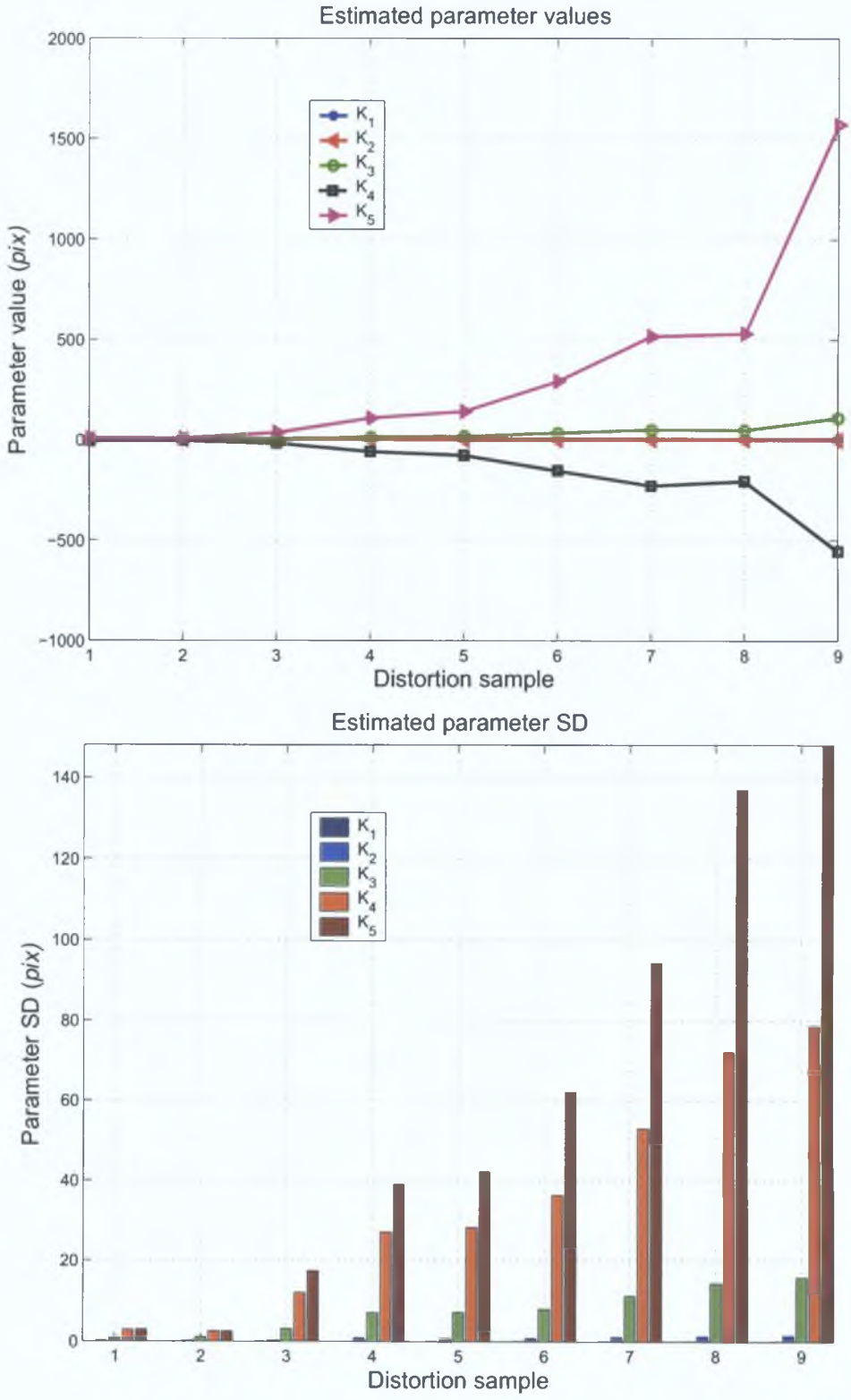
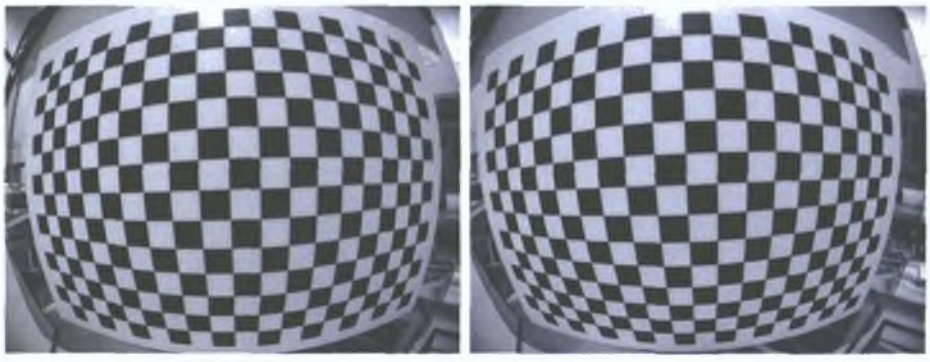
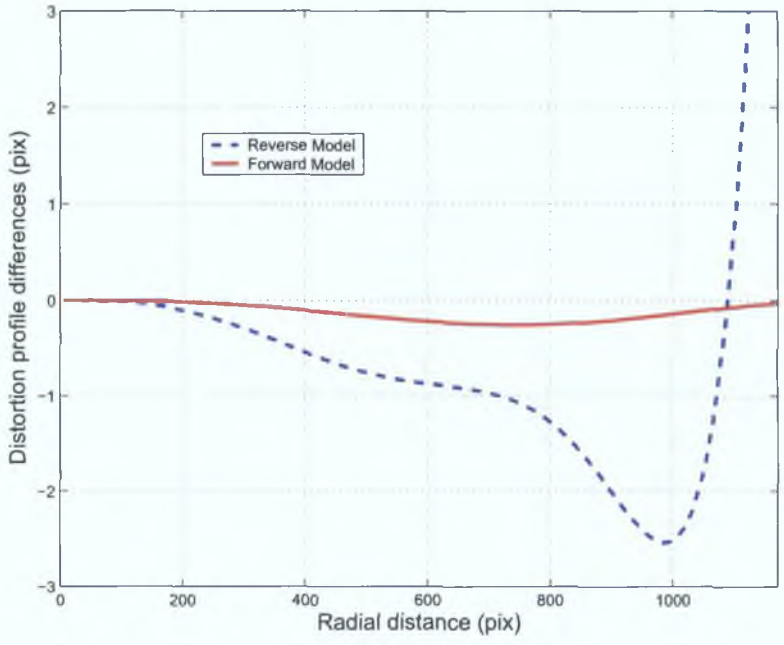


Fig. 2.8: Reverse model parameters, $\mathbf{k} = (k_1, k_2, k_3, k_4, k_5)^T$, and uncertainties when applied to sample set of increasingly distorted real images.



(a)



(b)

Fig. 2.9: Test for parameter invariance to calibration data locations. Two images taken with slightly differing viewpoints are shown in 2.9(a). The resulting profile variations for the forward and reverse models are shown graphically in 2.9(b).

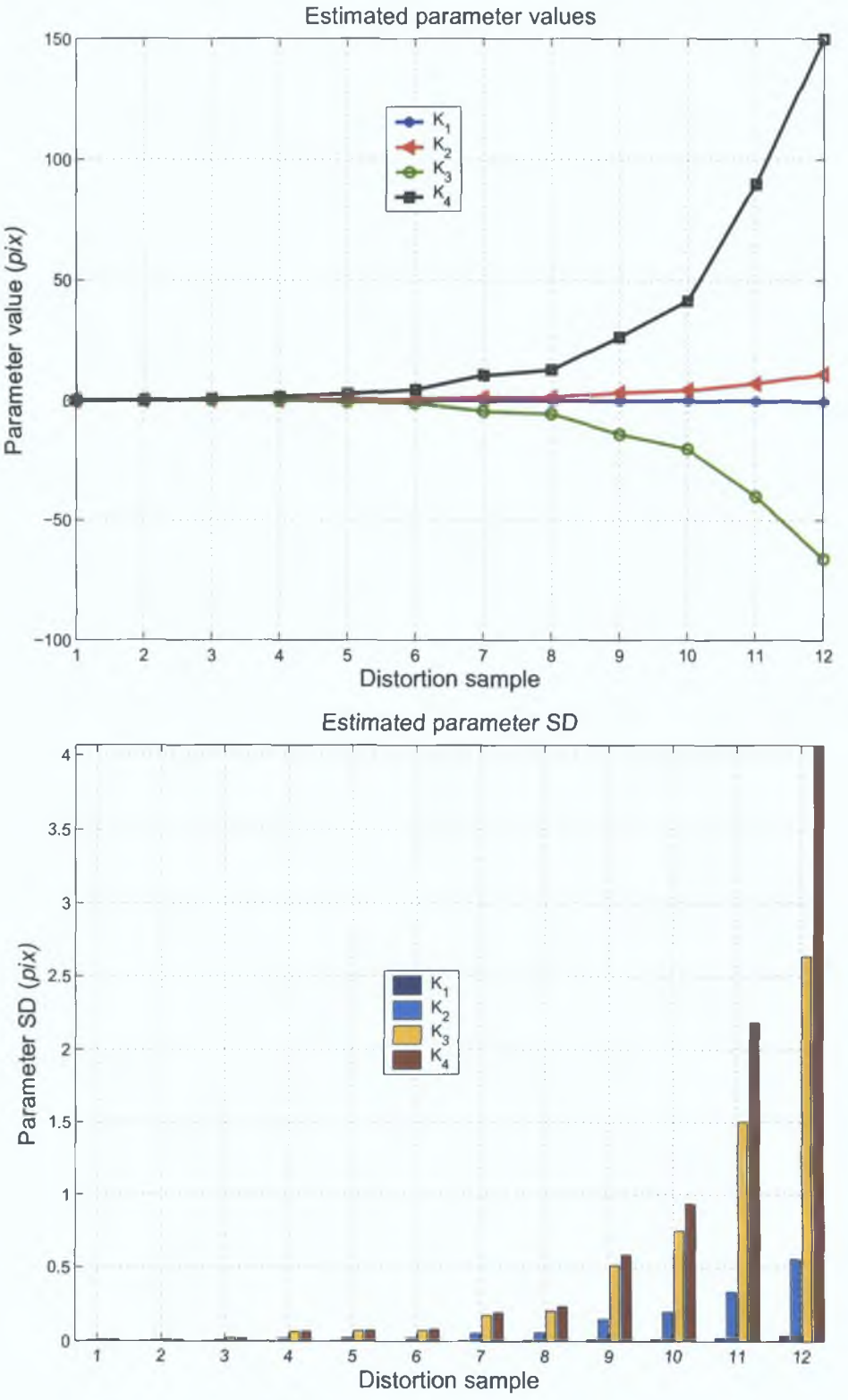


Fig. 2.10: General polynomial model parameters, $\mathbf{k} = (k_1, k_2, k_3, k_4)^T$, and uncertainties when applied to sample set of increasingly distorted real images.

The alternative models are now compared over the sample range by measuring the residual distortion in the set of evaluation images. The parameter space for all polynomial models is set to four. Fig. 2.11 shows the mean and standard deviation (SD) of the residual distortion remaining in the evaluation images, for the four models. The residuals are measured by taking a radially weighted homography (as described in Appendix A) to the corrected evaluation image [3]. The resulting Euclidean means and SD are then calculated.

The forward distortion model exhibits low mean and SD in residual distortion over the sample range. The alternative reverse model also exhibits a similar characteristic, however this model failed to converge for samples ten through twelve. The general polynomial model shows erratic behaviour for the upper range of distortion. This is due to the combined effects of the need for a higher order polynomial, unnecessary parameters and parameter or system instability. The divisional model shows good accuracy up to relatively high distortion levels.

The divisional model offers a very good and robust approximation to a constantly increasing distortion profile such as that observed in fish eye lenses. Hence, the popularity of this model for use with such lenses, e.g. Brauer-Burchardt and Voss (2001). However, many lenses that are optically compensated for distortion display a complex distortion profile. For such cases, the divisional model is unable to accurately model the lens nonlinearities. An example of one such case is shown in Fig. 2.12, taken with a low focal length lens. The divisional model residuals are compared to those of the forward model, clearly highlighting the shortcomings of the divisional model in this case. For this reason the divisional model cannot be considered as a general model if high accuracy is required.

[3]This radially weighted homography is only suitable for application to very mild distortion, such as residual distortion, and in general may not be a very stable criterion for measuring normal distortion levels or for its calibration

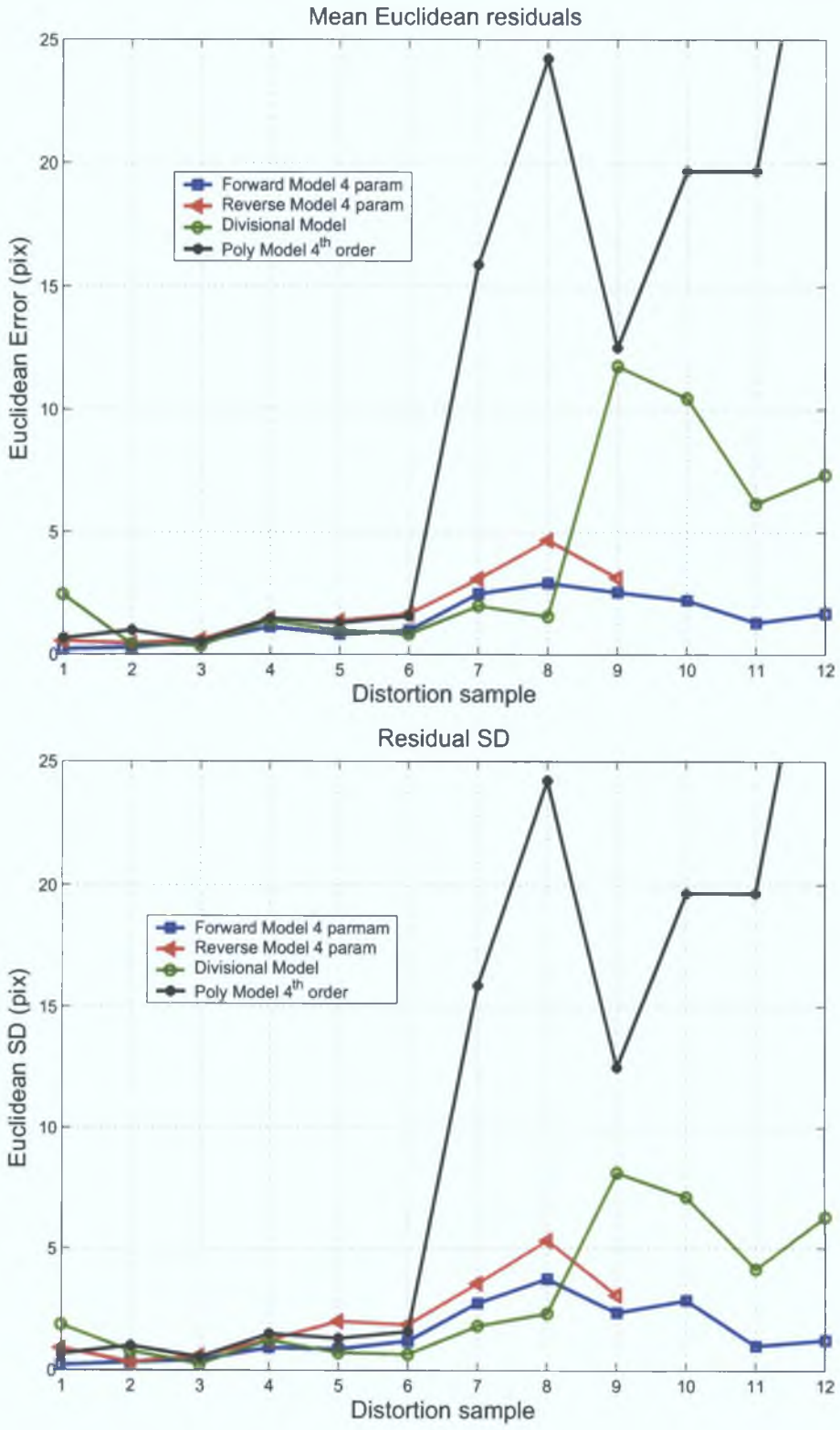


Fig. 2.11: Comparison of the residual fitting error of four possible distortion models.

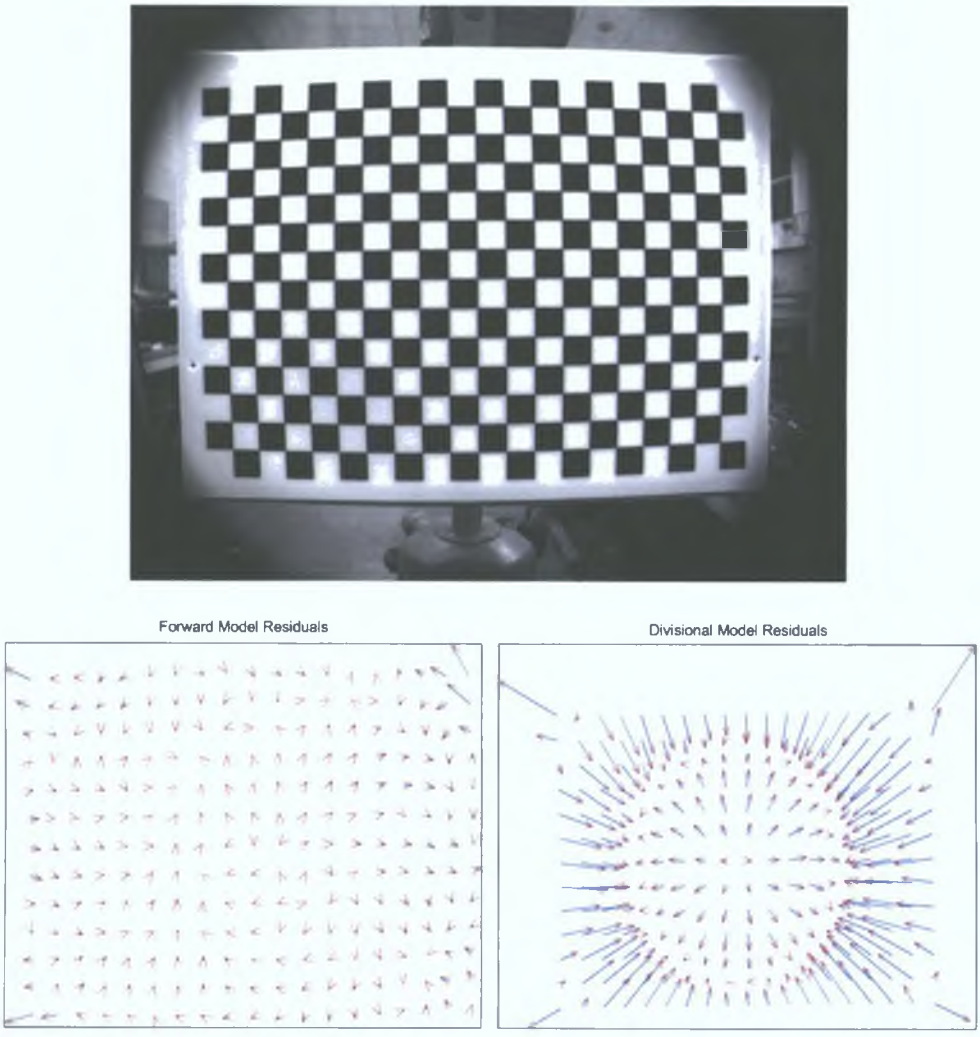


Fig. 2.12: Shows an image of a calibration pattern, with forward model and divisional model residuals. The image is taken with a Kodak Megaplus digital camera with a 3/4" CCD and fitted with a low cost Computar 6mm 1/2" lens (Format size mismatch results in the dark rimming). Divisional model parameter $-0.284(pix)^2$ and the first parameter of the forward model is $-0.3741(pix)^3$. Residual vector plot scale is $\times 100$.

In summary, these results, demonstrate that the forward model is a more appropriate approximation to radial lens distortion than the commonly used reverse interpretation. It is a better general model, giving stable results up to a very high level of distortion, see for example Fig. 2.4. The experiments show that an increased quantity of parameters does not lead to instability. In contrast, the reverse model is suited only for low levels of distortion, with high distortion and/or number of parameters increasingly leading to instabilities. These results agree with the theoretical investigation of these models.

The residual comparison between the parametric models show that the forward and reverse models achieve roughly equal accuracy, as expected. The general polynomial model fails to achieve this level of accuracy for larger distortion values, due to the need for a higher order function. Finally, the divisional model is shown to be a good model for constantly increasing distortion such as fish-eye lenses, but fails to model more complex distortion nonlinearities as present in many perspective lenses.

2.4.1 Inverse Performance

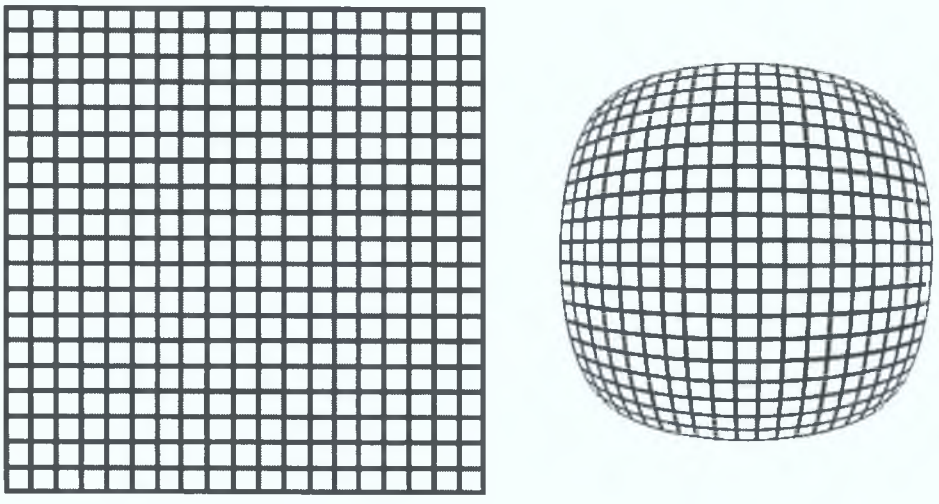
The accuracy of the proposed inverse (equation 2.8) is compared with two alternative formulations. The technique referred to as Method 1, is taken from Heikkila (2000)[4] and is also described in equation 2.6. Method 2 is taken from Heikkila and Silven (1997). [5] Method 1 is estimated using the iterative L-M technique with finite difference Jacobian, while Method 2 is estimated linearly as described in Heikkila and Silven (1997).

Separate sets of data are used for the parameter estimation and residual measurements. Fig. 2.13(a) shows the 20×20 grid used for residual measurements, completely covering the image window of 475×475 pixels. The estimation data consists of a 30×30 grid with the same coverage. Distortion is simulated using two radial parameters in the forward model with the ranges $k_1 = -2 \rightarrow 2$ and $k_2 = -0.8k_1$. The parameters, $\mathbf{a}$, of the proposed inverse model are estimated using the LS method of equation 2.9. Pixel coordinates are pre-scaled by 670

[4]This formulation originates from the inclusion of specific terms from the Taylor expansion of the assumed forward model.

[5]This formulation has a similar form to that of Method 1 but originates from the step-by-step refinement of a general firth order rational polynomial.

with the distortion centre at the image centre. The extreme barrel distortion of $k_1 = -2$, $k_2 = 1.6$ is simulated in Fig. 2.13(b), giving a visual impression of the distortion levels considered.



(a) (b)

Fig. 2.13: Evaluation data, represented by the line intersections in undistorted state 2.13(a), and in distorted state in 2.13(b) according with the forward model with parameters $k_1 = -2$, $k_2 = 1.6$.

The results are shown in Fig. 2.14. including the mean and SD over the considered range of distortion. These show the proposed inverse improves upon the inverse accuracy by an average of roughly 5000 times smaller than Method 1 and 100 times smaller than Method 2. This is a significant improvement, achieving a sub-pixel accuracy better than the lower bound of sub-pixel feature detectors over the entire range of distortion levels, (see chapter 3).

A second experiment is conducted to clarify the smooth interpolation of the inverse solution between the tie points. A uniformly distributed random set of 3000 points are generated to cover the image window of 475×475. These points are then distorted by applying the forward model with parameters $k_1 = -2$ and $k_2 = 1.6$, resulting in severe barrel distortion, as illustrated in Fig. 2.13(b). Following application of the estimated inverse solution, the error magnitudes are calculated and represented in the histograms of Fig. 2.15. These show that the error for over 2500 of the 3000 points lie within the histogram bin of $\pm 1 \times 10^{-3}$ for both the x and y directions. A small quantity, < 20 samples lie in the bins stretching from $\pm(0.01 \rightarrow 0.05)$. This inversion error naturally reduces as the distortion becomes less severe.

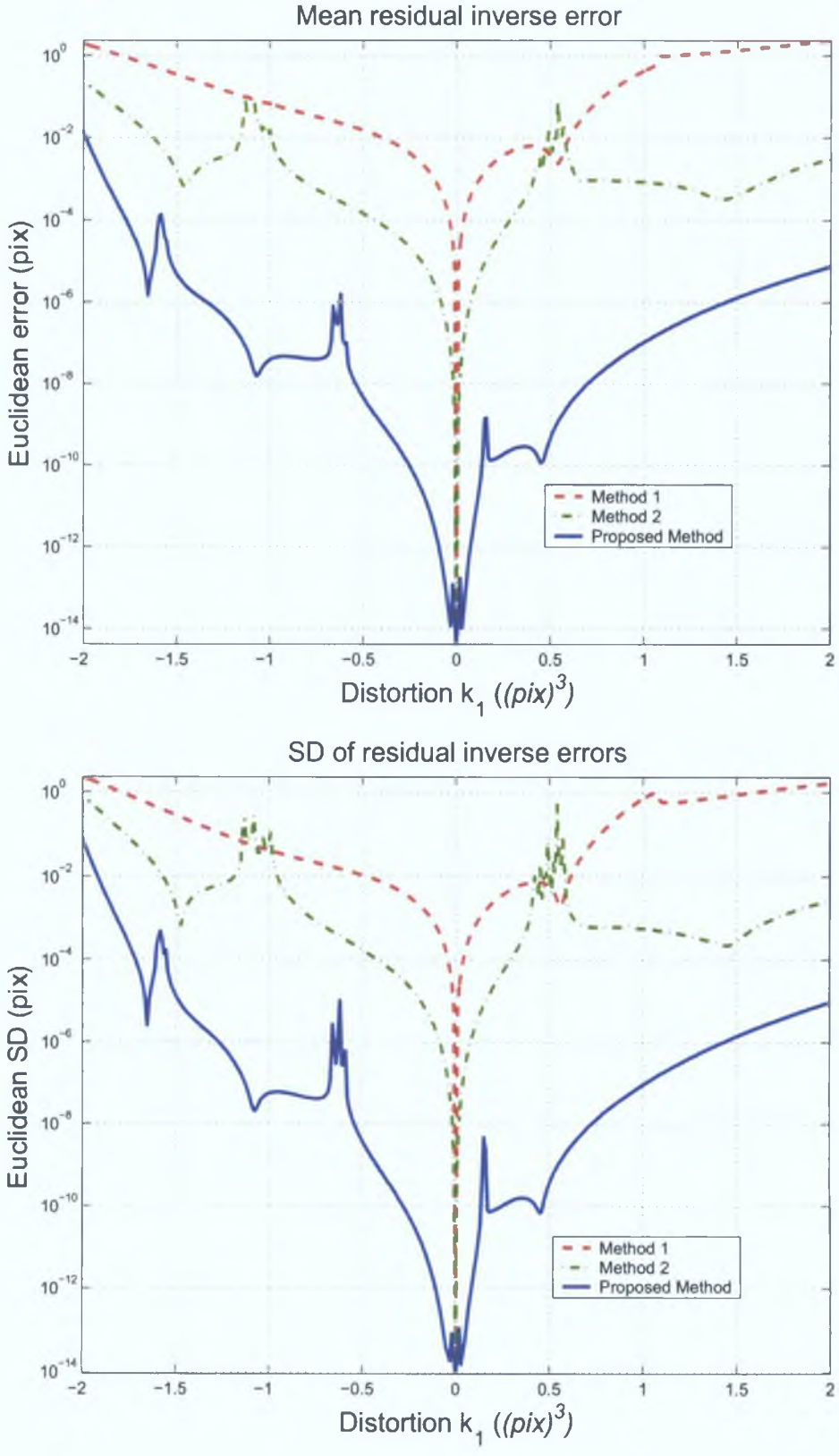


Fig. 2.14: Inverse Residuals

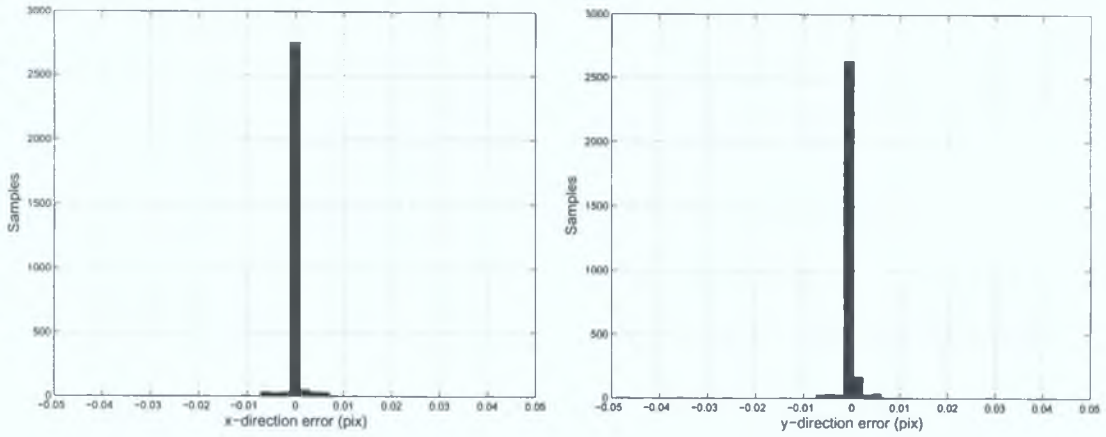


Fig. 2.15: Inverse Residuals

2.5 Discussion

This chapter is primarily concerned with the modelling of lens distortion in images. No previous work was found in the literature that considers the origins of the assumed parametric models, while alternative interpretations of the traditionally used model have led to a degree of confusion in some cases. To resolve this issue, a lens distortion model is derived from the optic wave equation, resulting in a model with the same form as the traditionally assumed one. The derived model includes radial and decentering components, where the decentering distortion is approximated by a simple displacement of the intermediately formed image. This model is a function of gaussian points, or in image terms, undistorted coordinates, and is referred to as the forward model.

An analysis of this forward model for the dominant radial type distortion shows that its Taylor expansion can be used to relate the alternative interpretation of this model (as a function of distorted coordinates). A theoretical argument is made that this model, referred to as the reverse model, will possibly match the forward model in terms of residual fitting errors, but will suffer in terms of stability as higher order coefficients will rapidly increase as distortion levels rise (in contrast with the forward model where higher order coefficients are likely to remain bounded). This increase in parameter magnitude subsequently led to poor conditioning in the set of estimation equations. As a result the set of equations become more sensitive to perturbations in location or noise content of the control points used in the calibration process.

This result is borne out in the experiments, where the reverse model parameters significantly increase as the distortion level rises. The associated confidence intervals in these parameters also increases indicating that the system of equation is less stable. Two manifestations of this instability are demonstrated, one in the failure to reach convergence due to singularities in the co-factor matrix for high distortion levels. The second example demonstrates the increased sensitivity of the reverse model to a slight change in control point location, as compared with the forward model relative invariance. The conclusion of this investigation is that the derived forward model is experimentally validated.

Alternative distortion models are also considered. The divisional model is shown to be related to the forward model, and shows good distortion modelling up to relatively high levels. However, it is unable to precisely model certain complex distortion profiles, commonly encountered in perspective camera lenses. A general polynomial model is also considered, but it is shown to be suitable only for low levels of distortion.

A disadvantage of the forward model is its lack of an analytical inverse. The form of such an inverse is also unknown, making the polynomial approximation of such an inverse effectively a trial and error exercise. This issue is also addressed in the formulation of an appropriate inverse approximation form. This is based on the inclusion of extra terms from the Taylor expansion of the forward model. A linear inverse function using this form is then proposed and is compared with existing alternatives. These comparisons indicate that the proposed solution improves the inversion accuracy by orders of magnitude over the range of likely encounter-able distortion levels.

Finally, the lack of a parametric model for lateral chromatic aberration is addressed with the derivation of an appropriate function. This model is composed of two elements, the chromatic variation of distortion and the lateral colour distortion. This is further investigated in Chapter 5.

Chapter 3

Choosing a Calibration Pattern

There is an abundance of planar charts used within the realms of camera calibration as sources of both 2D and 3D control points. These points are generally constructed on a planar surface by means of some high contrast pattern. In turn, the pattern also facilitates the recovery of the control point projections on the image plane. Patterns such as squares, chessboards and circles have become popular as they can be readily manufactured to a sufficient precision, and their data points are recoverable through the use of standard image processing techniques.

In real cameras, an image of the calibration pattern is likely to undergo two types of transformation: a projective transformation as a consequence of relative 3D position, and a nonlinear transformation due to various lens distortions. The control point invariance to errors resulting from these two transformations is based on a combination of the pattern employed, and the detection method used. As a consequence, for any theoretical combination of calibration pattern and detection method, two possible sources of bias in control point recovery have been identified, which are simply termed: Perspective bias and Distortion bias. In practice, the presence of these bias sources is primarily governed by the type of pattern used, which in turn dictates the appropriate detection methods. This chapter considers two calibration patterns from the camera calibration literature, each with two associated control point detection methods. These are the popular chessboard and circular patterns with respective detection methods of edge intersections, corners, centroids and conic fitting. The underlying biasing principles naturally extend to other similar patterns

and detection methods. The importance of acquiring bias free data has often been ignored in calibration articles, the result being that camera models and accuracy are not reliably estimated.

The main aim is to establish which pattern offers the best precision in control point recovery. The primary concern in this regard is to obtain bias free data, as this is clearly essential for obtaining uncorrupted estimates from calibration algorithms. It is shown theoretically and experimentally, with both real and simulated data, that circle centroid detected points are corrupted by both perspective bias and distortion bias, with greater distortion bias magnitude in a typical camera. However, only perspective bias compensation has been considered in the literature, most prominently by Heikkila (2000), using adjusted conic centroids. It is shown that the compensation of distortion bias from such circular pattern points is difficult, without knowledge of the true distortion free image. Real cameras systems offer no information on this front. As a consequence, the many calibration articles using such patterns, have their claims of high accuracy significantly compromised by their choice of calibration pattern.

The analysis is primarily conducted on simulated images with known ideal control points as shown in Fig. 3.1. Images and control points are synthesised with both nonlinear and projective transformations. Details of the image synthesis precision is presented ensuring no additional errors are introduced from this stage. The emphasis is on automatic point recovery where all points are observable in the image. Two sub-pixel detection methods for each pattern are described, each having unique bias invariant properties. The sources of biases are theoretically identified for each pattern and are subsequently verified on the simulated images. A comparative study of each method with respect to blurring and noise serves to show that the biasing magnitudes are significantly greater than the expected detection accuracy or noise floor. Finally, examples of perspective and distortion bias in a real images are shown. Overall, this chapter emphatically shows that the choice of pattern and detection technique is much more important than previously realised to achieve bias free control points for real cameras affected by lens distortion.

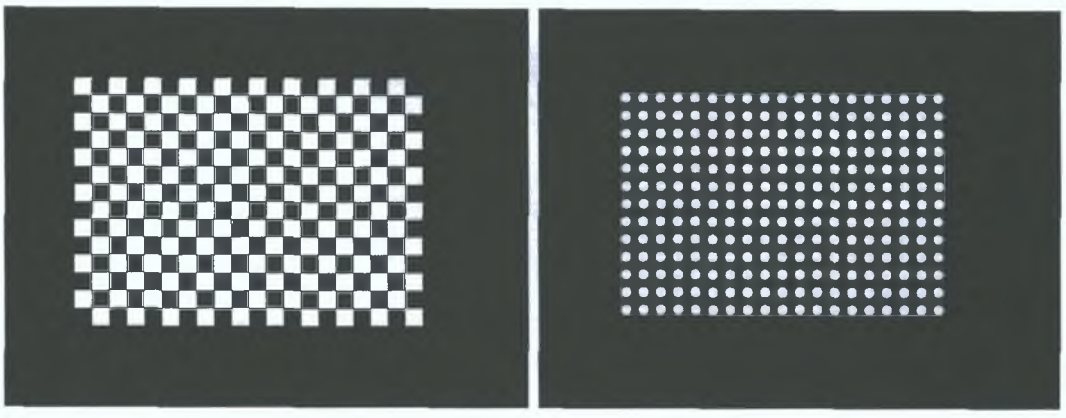


Fig. 3.1: The two classes of patterns used for this study, chessboard and circles, shown in their canonical form. Image dimensions are 2560 x 1920 pixels. 247 control points are synthesised. Chessboard squares have dimensions 85 x 85 pixels, while circle diameters are 51 pixels. The centre or principal point is at (1280,960).

3.1 Pattern and control point synthesis

Two factors influence the recovery of control points: the camera lens effects and the relative positioning of the calibration object. The error invariance to these transformations is based on the type of pattern employed and the detection methods used. Two popular patterns are chosen for this study, each with sufficiently different characteristics to illustrate all the possible sources of control point bias. Biasing aspects of other pattern types can be understood by comparison with the principles introduced here. The chosen chessboard and circular patterns are illustrated in canonical form in Fig. 3.1. The sizes of the circles and chessboard squares, and the actual number of control points are chosen as typical practical values. [1]

In analysis each pattern is subjected to Gaussian blurring, additive gaussian noise, pincushion and barrel distortion and random placements. The recovered points are then be compared with their true locations. For the distortion and positioning effects, both control points and the corresponding image must be transformed. Geometric image re-sampling is carried out by mapping from

[1]The size of the patterns are chosen to nominally occupy a large portion of the central image region. The circle diameters are chosen so as to give a reasonable spacing between circles. The number of control points are chosen as over 200 which is recommended in Section 4.3.1

the transformed image to the original. This involves calculating for every pixel in the transformed image, the corresponding pixel coordinate in the original image, effectively requiring an inverse mapping. The transformed image intensity is then calculated based on the standard bilinear interpolation around this coordinate.

3.1.1 Pattern positioning

The image perception of various 3D positions of the control points, $\mathbf{p} = (x, y, 1)^T$, (in homogeneous form) are simulated using a pseudo randomly generated homography $\mathbf{H}$, giving $\mathbf{p} = \mathbf{H}\bar{\mathbf{p}}$, where $\bar{\mathbf{p}}$ are the canonical representation of the control points. This homography is generated by a combination of a 3D rotation and translation, whose values are drawn randomly from a specific range. This range limit ensures that the transformed image lies roughly within the image window and that its apparent 3D position simulates a likely view of the calibration object. The corresponding image re-sampling is calculated using $\mathbf{H}^{-1}$. To quantify the effect of the homography the conditioning of its Jacobian is taken at each control point:

$$\text{cond}(\mathbf{J}(\mathbf{H}, \mathbf{p})), \quad \text{where} \quad \mathbf{J}(\mathbf{H}, \mathbf{p}) = \begin{pmatrix} \frac{\partial x}{\partial \bar{x}} & \frac{\partial x}{\partial \bar{y}} \\ \frac{\partial y}{\partial \bar{x}} & \frac{\partial y}{\partial \bar{y}} \end{pmatrix}.$$

The condition number in this case measures the distance from an orthogonal transform, thus effectively quantifying the perspective element of the homography. The mean of these values is taken as a measure of the net effect of the perspective transform.

3.1.2 Simulating lens distortion

General radial lens distortion is approximated in an image according to the forward model described in chapter 2 as:

$$\breve{\mathbf{p}} = \mathbf{p} + \mathcal{D}(\mathbf{p}, \mathbf{k}), \text{ where } \mathcal{D}(\mathbf{p}, \mathbf{k}) = \begin{pmatrix} x(k_1 r^2 + k_2 r^4 + 1) \\ y(k_1 r^2 + k_2 r^4 + 1) \end{pmatrix}, \tag{3.1}$$

where $\mathbf{p} = (x, y, 1)^T$ are the undistorted image coordinates with $r^2 = x^2 + y^2$ and $\breve{\mathbf{p}} = (\breve{x}, \breve{y}, 1)^T$ are the corresponding distorted coordinates. The distortion parameter, (k_1), is used as an index, with values varying through $\pm 2(pix^3)$,

with $k_2 = -k_1$. These values are applied to pixel coordinates normalised by the average of the image width and height, denoted by (pix). In order to re-sample a distorted image an inverse of (3.1) is required. The approximate linear model is used as proposed in chapter 2 equation 2.8. The parameters of this model are linearly estimated using a dense collection of points covering the image window. The inversion accuracy for the distortion range under consideration is shown in Fig. 3.2. These residuals are orders of magnitude lower than the precision of the sub-pixel point detection algorithms, (see Section 3.3), ensuring that no additional source of error is introduced from the simulated images.

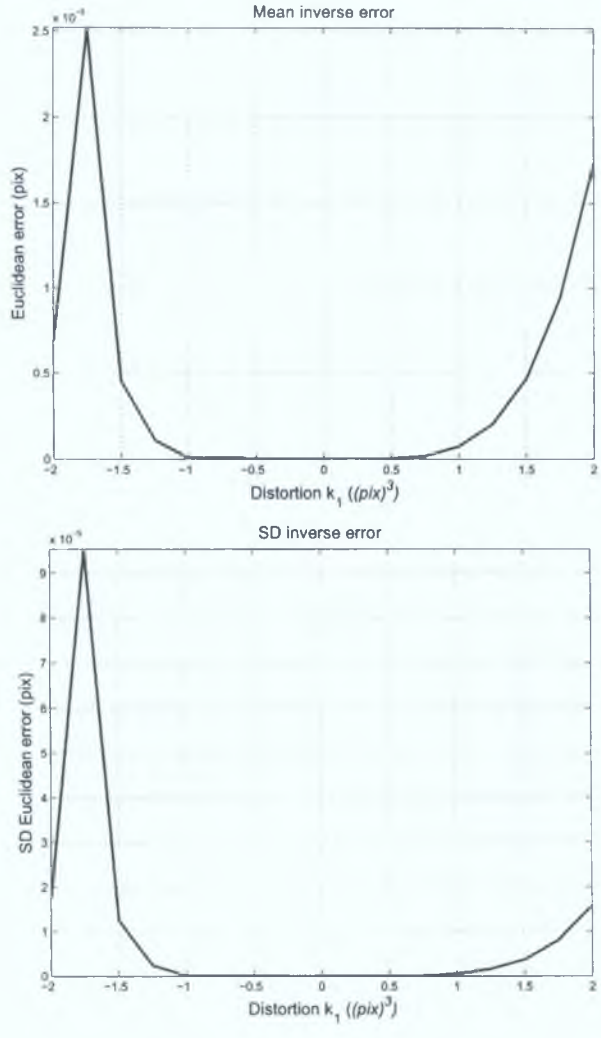


Fig. 3.2: Mean and SD inverse distortion residuals after the fitting of inverse distortion approximation. Levels are orders of magnitude below that of the control point detection accuracy, see Section 3.3

3.2 Control point recovery

For each pattern, two sub-pixel detection methods are described. The circle detection methods are based on the centroid extraction and ellipse fitting. Square detection methods are based on edge intersections and refining an initial corner solution with local surface fitting. It will be shown that the detection methods are subject to two sources of biasing. It is theoretically shown how they arise as a result of perspective viewing and lens distortion.

Centroid extraction methods are shown to be compromised by both perspective bias and distortion bias. As is currently known, conic fitting techniques can be adjusted to alleviate perspective bias only. Edge fitting techniques also suffer from distortion bias, while only local surface fitting offers both perspective and distortion bias free recovery of control points.

3.2.1 Circle pattern detection

Given an image of a circular pattern and following some basic image processing, the centroids of the circles are simply calculated as:

$$\mathbf{c_{cen}} = (x_{cen}, y_{cen}, 1)^T = \left(\frac{\sum_{x_f \in F} x_f \cdot I(\mathbf{p_f})}{\sum_{x_f \in F} I(\mathbf{p_f})}, \frac{\sum_{y_f \in F} y_f \cdot I(\mathbf{p_f})}{\sum_{y_f \in F} I(\mathbf{p_f})}, 1 \right)^T, \quad (3.2)$$

where $I(\mathbf{p_f})$ is the intensity at point $\mathbf{p_f}$ and F is the set of pixels deemed to belong to the circle. It is known that if the calibration plane is not parallel with the image plane, a bias is introduced into $\mathbf{c_{cen}}$.

Perspective bias

Considering a calibration plane in a general 3D position, a homography $\mathbf{H}$ can be computed between the pixel coordinates of the control points and the ideal canonical position as: $\mathbf{p} = \mathbf{H}\bar{\mathbf{p}}$, where $\mathbf{p}$ are the locations of the control points in the image, and $\bar{\mathbf{p}}$ are the ideal canonical position of the control points. The conic approximation to the edge points of the pattern, $\bar{\mathbf{p}}_{edge}$, can be estimated linearly as: $\bar{\mathbf{p}}_{edge}^T \bar{\mathbf{Q}} \bar{\mathbf{p}}_{edge}$ (Hartley and Zisserman, 2003). The centre of the conic is then calculated from the conic $\bar{\mathbf{Q}}$ as: $\bar{\mathbf{c}}_{conic} = \bar{\mathbf{Q}}^{-1}[0, 0, 1]^T$. For a general 3D position these centers transform to $\mathbf{c}_{conic} = \mathbf{H}\bar{\mathbf{Q}}^{-1}[0, 0, 1]^T$. However, in

an image only the conic $\mathbf{Q}$ is available, but it may be related to $\bar{\mathbf{Q}}$ through $\mathbf{Q} = \mathbf{H}^{-T}\bar{\mathbf{Q}}\mathbf{H}^{-1}$. Thus the unbiased estimates for the centers of the conics undergoing a general perspective transform $\mathbf{H}$ is given by:

$$\mathbf{c}_{conic} = \mathbf{Q}^{-1}\mathbf{H}^{-T}[0, 0, 1]^T. \tag{3.3}$$

The extent of this biasing influence is simulated in Fig. 3.3 for random

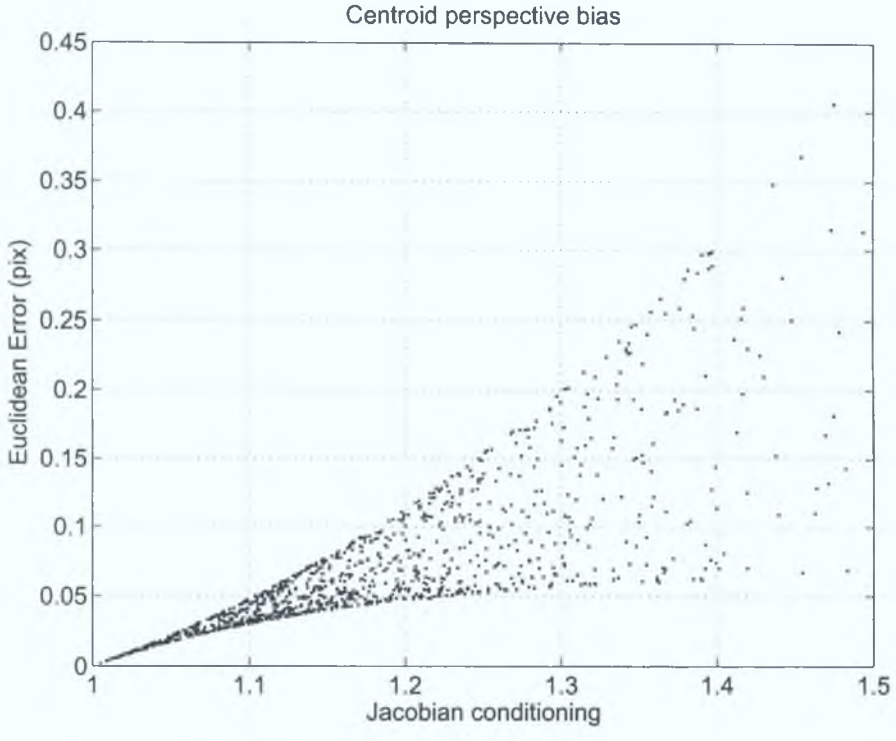


Fig. 3.3: Mean Euclidean error as a consequence of 1000 random positions of the circular pattern. Orthogonal like homographies induce a low conditioning (close to one) while the perspective bias increases with increasing conditioning. The average diameter of the circles is around 50 pixels.

perspective views described in Section 3.1.1. In many algorithms, especially when lens distortion is a factor, the value of $\mathbf{H}$ or equivalently, the elements of the camera projection matrix, are not known exactly beforehand. This forces the algorithm to iteratively update the estimates of the control points. This re-estimation of the control points is an added complication, increasing the number of iterations and degrading derivative information.

Distortion bias

The second major drawback of circle patterns and their detection methods, is that they are also subject to bias from lens distortion. Lens distortion

introduces a nonlinear shape warping to the area of the conic. This warping subsequently biases the centre point of the conic. The extent of the bias is dependent on the amount of lens distortion and the area or radius of the conic, as illustrated in Fig. 3.4 and 3.5.

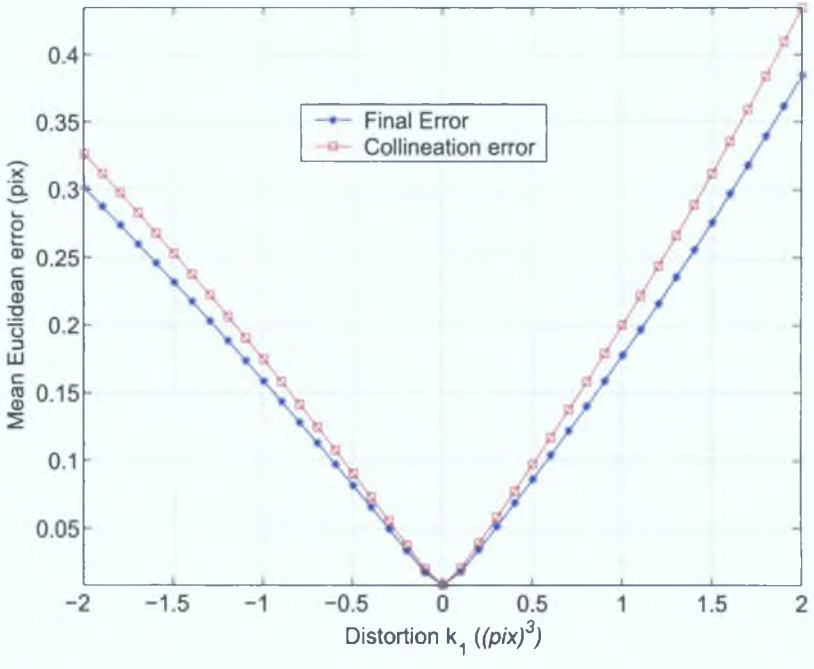


Fig. 3.4: Simulated mean Euclidean distortion bias for circle pattern (circle dia 50 pix) over the considered range of distortion. The collineation error resulting from $\mathbf{H}_{\mathcal{D}}$ is actually reduced by the conic ill-fitting error.

This bias results from the combined effects of two error sources. Firstly, the nonlinear nature of distortion warps the conic so that it is no longer a true conic. Certain sections of the conic become elongated or compressed, all culminating in the introduction of a bias from the eventual conic fitting. Tracking the equations for only one term of distortion (k_1) the second order least squares conic fitting: $\breve{\mathrm{p}}^T \breve{\mathrm{C}} \mathrm{p}$, is performed on a sixth order section. This leads to an ill-fitting bias, the extent of which is illustrated in Fig. 3.4. The analytical compensation for such bias is not possible without exact knowledge as to the true undistorted state of the control points.

The second error source is from the distortion induced local perspective transform, resulting from the conic fitting. Considering a general distortion free conic $\mathbf{Q}$, its least square counterpart in distorted space is calculated by minimising the algebraic expression: $\sum_i (\breve{\mathrm{p}}_i, \breve{\mathbf{Q}} \mathrm{p}_i)^2$, where $\breve{\mathbf{Q}}$ must be a real proper conic. These two conics are related through a set of collineations or homo-

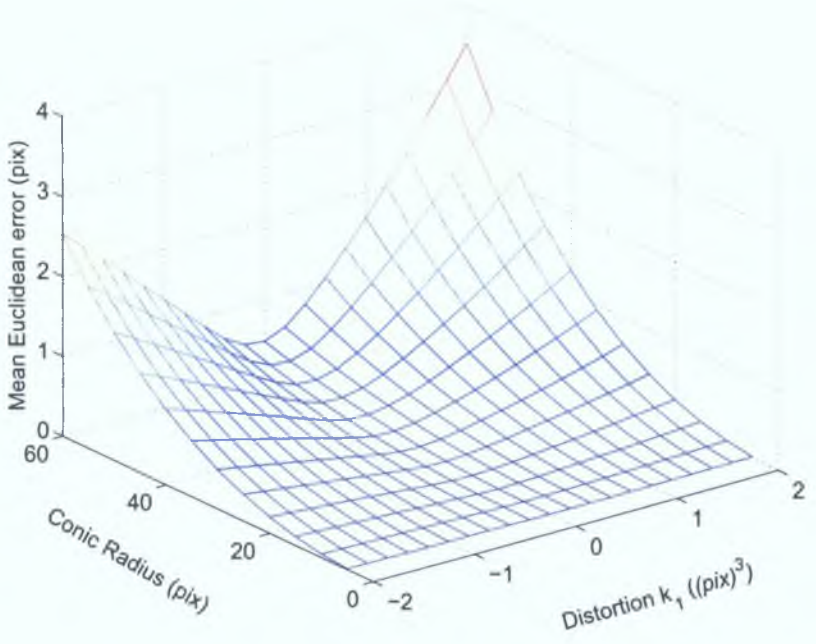


Fig. 3.5: Simulated mean Euclidean distortion bias in patterns of various circle radii, considered over a range of distortion levels.

graphies $\mathbf{H}_{\mathcal{D}}$ that map $\mathbf{Q}$ to $\breve{}$ through $\mathbf{H}_{\mathcal{D}}^{T}\mathbf{Q}\mathbf{H}_{\mathcal{D}} = \breve{}$. This collineation may be found by taking the orthogonal matrices $\mathbf{U}$ and $\breve{}$ that diagonalise $\mathbf{Q}$ and $\breve{}$: $\mathbf{U}^{T}\mathbf{Q}\mathbf{U} = \mathbf{\Lambda}$ and $\breve{}^{T}\breve{} = \breve{}$, where $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, \lambda_3)$ and $\breve{} = diag(\breve{\lambda}_1, \breve{\lambda}_2, \breve{\lambda}_3)$, and by choosing $\mathbf{\Lambda_0} = diag(\sqrt{\breve{\lambda}_1/\lambda_1}, \sqrt{\breve{\lambda}_2/\lambda_2}, \sqrt{\breve{\lambda}_3/\lambda_3})$. The collineation $\mathbf{H}_{\mathcal{D}}$ is then formed as $\mathbf{H}_{\mathcal{D}} = \breve{}\ {}_0\mathbf{U}$. However, the equation $\mathbf{H}_{\mathcal{D}}^{T}\mathbf{Q}\mathbf{H}_{\mathcal{D}} = \breve{}$ is not unique as it provides only five of the necessary eight independent constraints. For simulation purposes a unique solution is obtained in least square sense by minimising $\sum_i (\breve{\mathrm{p}}_i, \mathbf{H}_{\mathcal{D}}\mathbf{p}_i)^2$.

The homography $\mathbf{H}_{\mathcal{D}}$ introduces a local perspective bias that we call the collineation error. Compensation for this bias cannot be applied in real cameras as the undistorted points are always unobservable. The contribution of the collineation error source is simulated in Fig. 3.4, revealing that it is dominant, and is actually reduced by the ill-fitting error. Naturally, distortion bias is heavily dependent on the size of the feature. This is examined in Fig. 3.5 for a range of distortion levels and circle diameters. Compared with the perspective bias simulated in Fig. 3.3, a typical low focal length lens, where k_1 is roughly in the region of $-0.3 \rightarrow -0.7 pix^3$, distortion bias is likely to be greater in magnitude than perspective bias. This is verified in Section 3.3. Note also that distortion bias is not limited to conic fitting, and is present to the same extent regardless of detection mode, centroids or conic fitting.

3.2.2 Chessboard pattern detection

Given an image of a chessboard pattern, initial estimates of the location of the intersections can be gathered using standard corner detection methods. These estimates are generally within a few pixels of the true locations. We describe two existing means of refining these initial solutions using edge information (Willson, 1994), (Li and Lavest, 1996) and surface fitting (Lucchese and Mitra, 2002). Additionally, we address the filtering of initial corner estimates to ensure they lie upon a square intersection, and the automatic ordering of these coordinates to correspond with the canonical point representation.

Primary detection

Initial estimates for the location of chessboard type intersections are obtained using standard corner detectors such as those described in Lucchese and Mitra (2002), Whelan and Molloy (2000), Jain et al. (1995). For real situations where the background scene registers candidate corners a further refinement step is necessary to remove false hits. A small $N \times N$ region of interest, Ψ, centered on the candidate corner is first thresholded using the mean gray level of Ψ. A symmetry measure t_Ψ can then be calculated as:

$$t_\Psi = \sum_{y=0}^{N}\sum_{x=0}^{N} \mathcal{O}(x,y), \text{where } \mathcal{O}(x,y) = \begin{cases} a & \text{if } \begin{array}{c}\Psi(x,y) = \Psi(N-x,N-y) \\ \Psi(x,y) \neq \Psi(x,N-y)\end{array} \\ b & \text{otherwise.} \end{cases},$$

where a is positive and b is negative. We obtain good performance using $a = 6$ and $b = -1$ with $N = 9$. High values of the symmetry measure t_Ψ indicate the corner is situated on a chessboard intersection.

Edge intersections

A fitting function that models line intersections is formulated. In order to use it with a chessboard pattern the edges or intensity derivatives in a medium sized local region Ψ, centered on the initial estimate are first calculated. The function is then fit using the L-M non-linear iterative technique:

$$\min_{\mathbf{h}} \| h_1 e^{-h_2^2((x-h_5)\cos h_3 + (y-h_6)\sin h_3)^2} + h_1 e^{-h_2^2((x-h_5)\sin h_4 + (y-h_6)\cos h_4)^2} - 2h_1 e^{-h_2^2((x-h_5)^2+(y-h_6)^2)} - \Psi(x,y) \|^2 ,$$

where the intersection point is (h_5, h_6), h_1 is the height of the derivative profile, h_2 is the width of the profile and h_3 and h_4 are the edge directions. The process is illustrated in Fig. 3.6. As lines project to lines under perspective transformations, this detection method is invariant to perspective bias. However, under lens distortion, it is clear that lines project to curves, with the result that this method is affected by distortion bias. Consequently, an analytical proof is not perused.

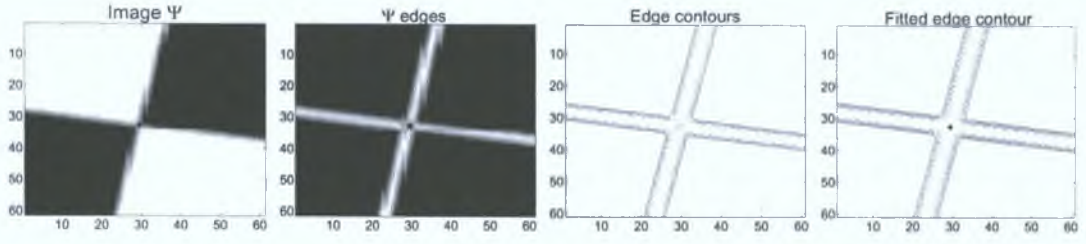


Fig. 3.6: Control point refinement based on edge fitting. The first image shows the selected ROI. The second image shows the detected edges. The third image shows the edge image contours. The final image shows the function fit contours from which the control point is calculated.

Corners

A category of sub-pixel refinement is based on surface fitting of intensity around a corner point (Lucchese and Mitra, 2002). For each initial location estimate, a small region of interest Ψ is considered for fitting. Following blurring, a quadratic function can be linearly fit to the resulting intensity profile, as demonstrated in Fig. 3.6 by minimising:

$$\min_{\mathbf{s}} \|s_1x^2 + s_2xy + s_3y^2 + s_4x + s_5y + s_6 - \Psi(x, y)\|^2.$$

The intersection point or saddle point is derived from this surface as the intersection of the two lines $2s_1x + s_2y + s_4 = 0$ and $s_2x + 2s_3y + s_5 = 0$. The process is illustrated in Fig. 3.7. In practice, the small patch Ψ can effectively be considered a single point, especially in light of the detection accuracy and noise floor. As points project to points under both projective and lens distortion transformations, this method has the desirable properties of being invariant to both perspective and distortion bias.

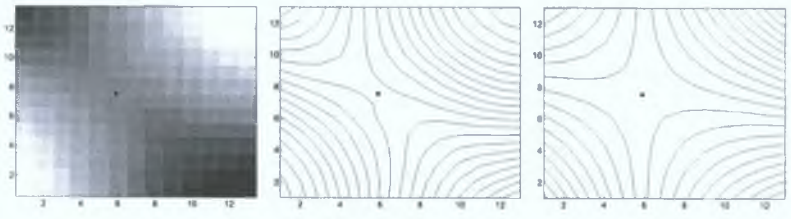


Fig. 3.7: The saddle refinement process. First the blurred ROI with marked saddle point is shown. Middle image is a contoured image of the intensity profile. Last image shows the intensity profile of the surface fit from which the saddle point is calculated.

3.2.3 Ordering

After extraction, the coordinates of the control points need to be appropriately ordered to ensure one-to-one correspondence with the arbitrary scaled canonical positions of the control points. For a pattern containing a matrix of control points of $N \times M$, the entire indexed set of points are collected into a three vector $\mathbf{d}_{(0\ldots N\times M)} = [x_1 \ldots x_{N\times M}, y_1 \ldots y_{N\times M}, 1]^T$. For most of the analysis in this thesis $N = 13$ and $M = 19$ giving a total of 247 points. Existing ordering techniques in Lucchese (2005) require prior user initialisation. The fully automatic algorithm described here has been extensively tested and performs successfully for different distortion levels and resolutions. For severely distorted images, such as fish-eye lenses, a prior rough un-distortion of $\mathbf{d}_{(0\ldots N\times M)}$ is required.

A brief description of the algorithm is as follows. For each row N, the topmost left and right coordinates in $\mathbf{d}_{(0\ldots N\times M)}$ are found. For all but the most acute angles of view, these points correspond to the top left and right points of the calibration pattern. Their indexes i and j are found by: $i = \min(\mathbf{x} \times \mathbf{y})$ and $j = \min(\mathbf{y}/\mathbf{x})$. A line though these points is then computed as $L = \mathbf{d}_i^T \bigwedge \mathbf{d}_j^T$, where $\bigwedge$ is the cross product. The $M - 2$ most closest points to this line are then found. Finally, these M points are ordered with increasing x coordinate, and are not considered in further iterations. This is performed N times until all the data has been ordered.

3.3 Experiments

Three sets of experiments are conducted on the synthesised test images described in Section 3.1. Two of the experiments verify and quantify the perspective and distortion bias for each detection method and pattern. A side by side accuracy evaluation for noise and blurring is presented, to give a benchmark from which to access the magnitude of both bias sources. Finally, real examples of distortion and projective bias are presented, based on a combination pattern of circles and squares. For simplicity detection method labels are shortened, circle centroids are referred to as centroids, conic centroids as conics, chessboard edge intersections as edges and chessboard corner saddle refinements are referred to as corners.

3.3.1 Noise and Blurring

The performance of each method is examined for a range of Gaussian blurring. Fig. 3.8 shows the mean and standard deviation of the Euclidean errors computed using the true locations. These test image patterns are also projectively transformed so as conic based compensation can be accessed. This shows that the expected detection errors remain relatively constant with respect to blurring. Excluding bias corrupted centroids, these errors are roughly in the pixel range of $0.02 \rightarrow 0.04$, and lower for the edge based method.

The robustness of the detection methods to noise is presented in table 3.1, for additive normally distributed noise. The upper level, $\sigma = 20$ (pix), represents severe noise unlikely to be encountered in typical calibration shots. Typical values for noise in images are in the range $\sigma = 5$ to $\sigma = 10$ pix.

3.3.2 Positioning Bias

The detection patterns and methods are examined for a range of nine different projective transformations as shown in Fig. 3.9. The perspective bias of centroid detection on circle patterns can be observed, and increases with perspective severity. These values correspond with their simulated counterparts in Section 3.2.1. These basic statistics do not convey that these errors are not randomly distributed, and are in fact biased. Fig. 3.10 shows one sample of

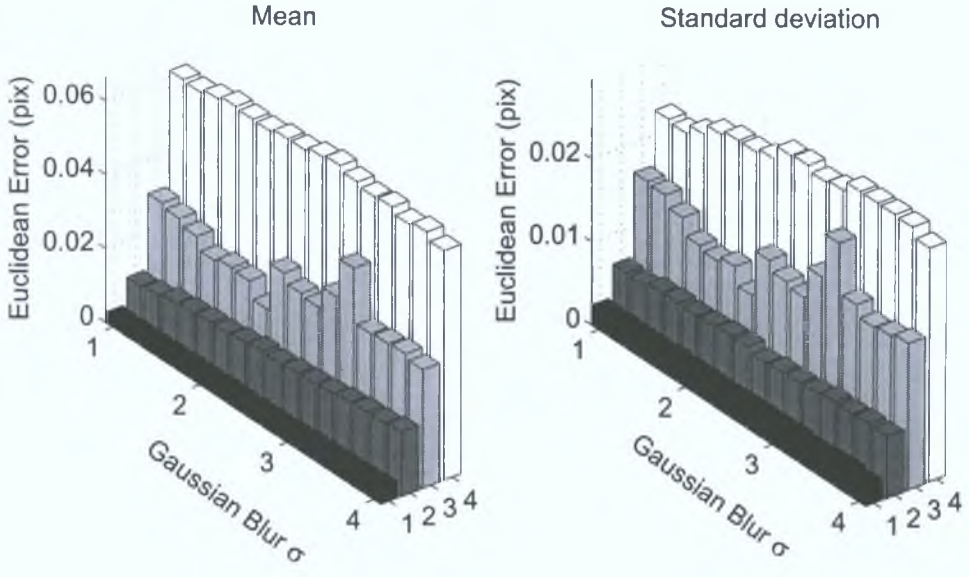


Fig. 3.8: The mean and standard deviation for the four methods over a range of gaussian blur levels. 1 ⇒ edges, 2 ⇒ corners, 3 ⇒ conics and 4 ⇒ centroids. Images are subject to a perspective transform.

this circle centroid bias compared with bias free adjusted conic fitting.

3.3.3 Distortion Bias

The patterns and detection methods are evaluated without perspective warping over the range of distortion levels. External sources of simulation error have been shown to have insignificant levels in Section 3.1.2. Fig. 3.11 shows the mean Euclidean error of circle pattern detection methods steadily increases

Table 3.1: Euclidean errors (Mean & SD) with respect to additive gaussian noises. Errors are compiled over one hundred independent trials and are conducted on perspective and distortion free images.

Noise σ (pix) =	Centroids		Conics		Corners		Edges	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	0	0	0.0281	0.0141	0.0051	0.0026	0.0014	0.0007
5	0.0012	0.0039	0.0369	0.0198	0.0144	0.0076	0.0067	0.0035
10	0.0220	0.0122	0.0541	0.0287	0.0279	0.0149	0.0134	0.0070
15	0.0355	0.0188	0.0657	0.0350	0.0420	0.0221	0.0200	0.0104
20	0.0447	0.0233	0.0841	0.0805	0.0568	0.0298	0.0265	0.0139

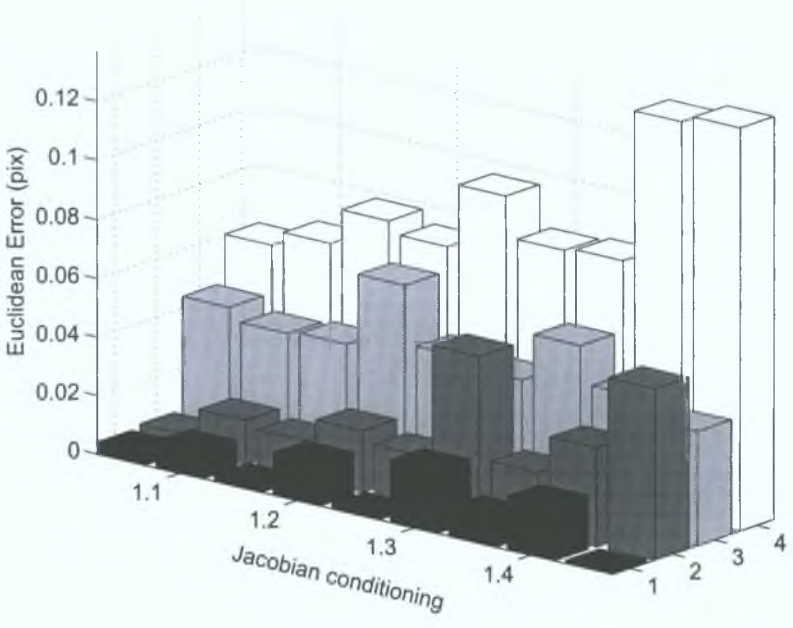


Fig. 3.9: The mean Euclidean error for the four detection methods (1 ⇒ edges, 2 ⇒ corners, 3 ⇒ conics, 4 ⇒ centroids) simulated over various degrees of perspective transforms. Centroid bias is clearly shown by the large mean Euclidean error.

with distortion level. This is in excellent agreement with the simulated distortion bias of Section 3.2.1. Edge based detection, because of its line fitting, reduces distortion bias somewhat, more so for the pincushion variety due to the fixed windowing size and distortion induced image expansion. Fig. 3.12 shows the distortion simulated images, each with associated detection method, for one sample of distortion.

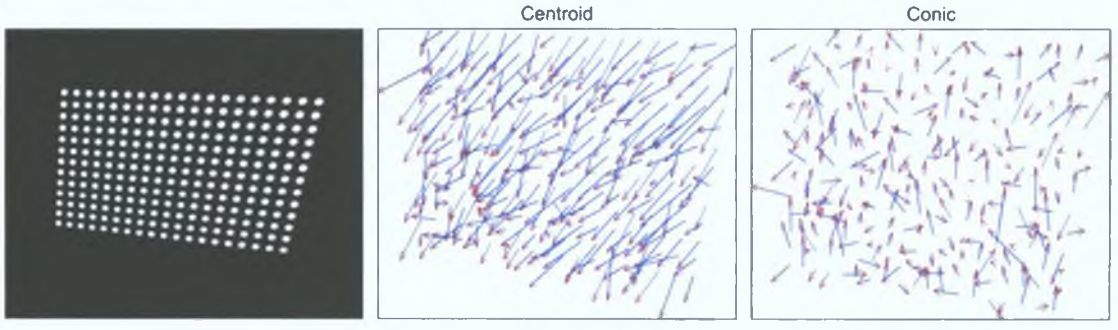


Fig. 3.10: Sample of one perspective transformation (cond. = 1.25). Vector plots reveal the centroid bias. Residual scale = ×2000.

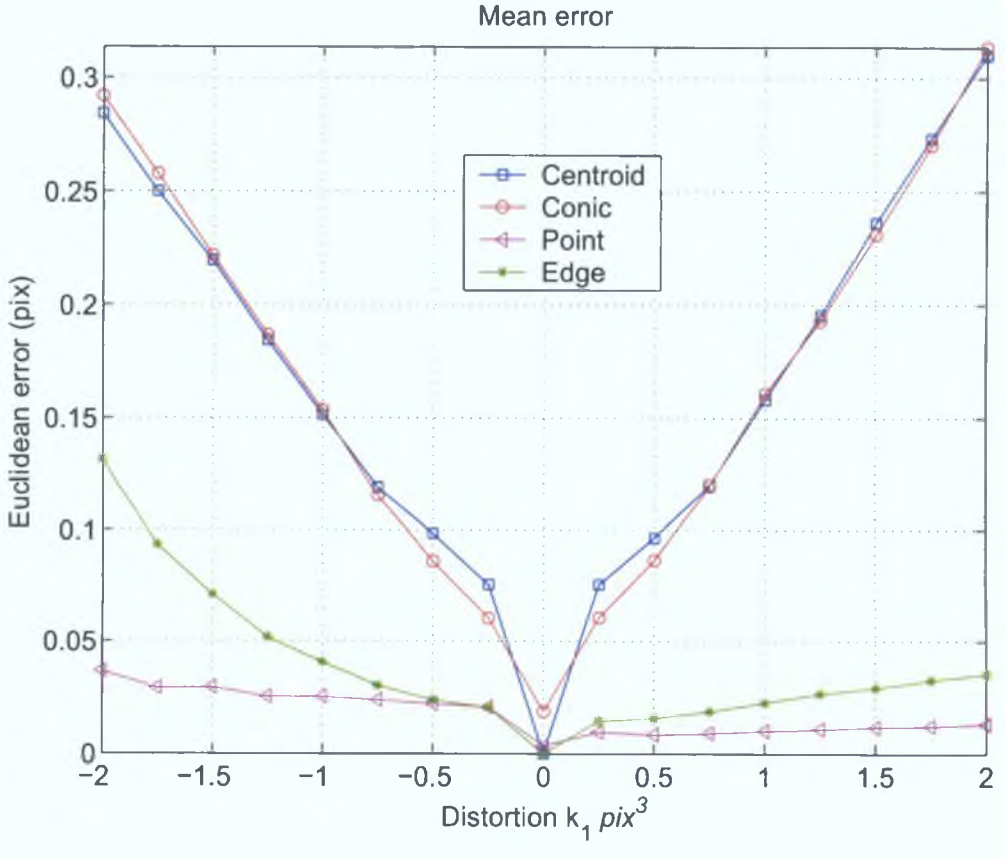


Fig. 3.11: Illustrates the distortion induced bias in control points for four detection methods.

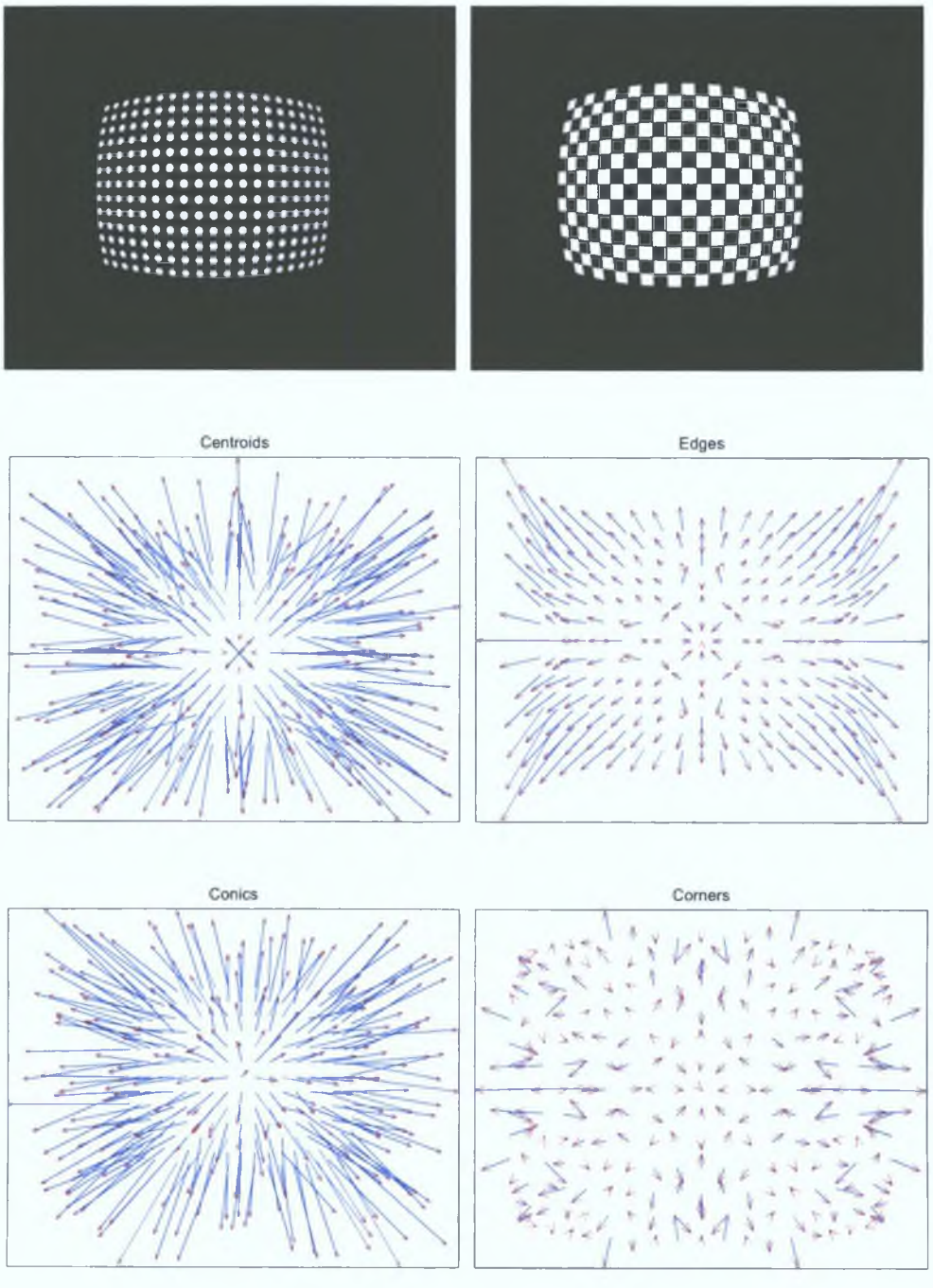


Fig. 3.12: Left column images show the circle pattern and associated detection method errors. Right column images show the chessboard and associated methods. Distortion level is $k_1 = -1.5pix^3$. Residual scale = $\times 2000$.

3.3.4 Bias in real images

Real examples of distortion and perspective bias are shown, measured on a special pattern that combines both circles and chessboards. This pattern, shown in Fig. 3.13, consists of three greyscale levels from which the two types of control points are extracted. Both the circle centroids and the square intersections are located at exactly the same locations. The circle control points are formed by thresholding above the mid grey level, where the circle integrity is ensured by the outer white rim. The square intersections are evaluated with a small ROI using the original intensity profile.

Two different degrees of lens distortion are examined, both with and without perspective bias. [2] Fig. 3.13 shows low level distortion bias, which is swamped by perspective bias as shown in Fig. 3.14. Fig. 3.15 shows an increased distortion bias that is dominant over perspective bias (Fig. 3.16) for an image affected by a greater level of distortion. This comprehensively demonstrates that such biases are not limited to the simulated case, and are equally prevalent in real images.

[2] Camera placement is done manually, and undoubtedly some degree of perspective bias is included.

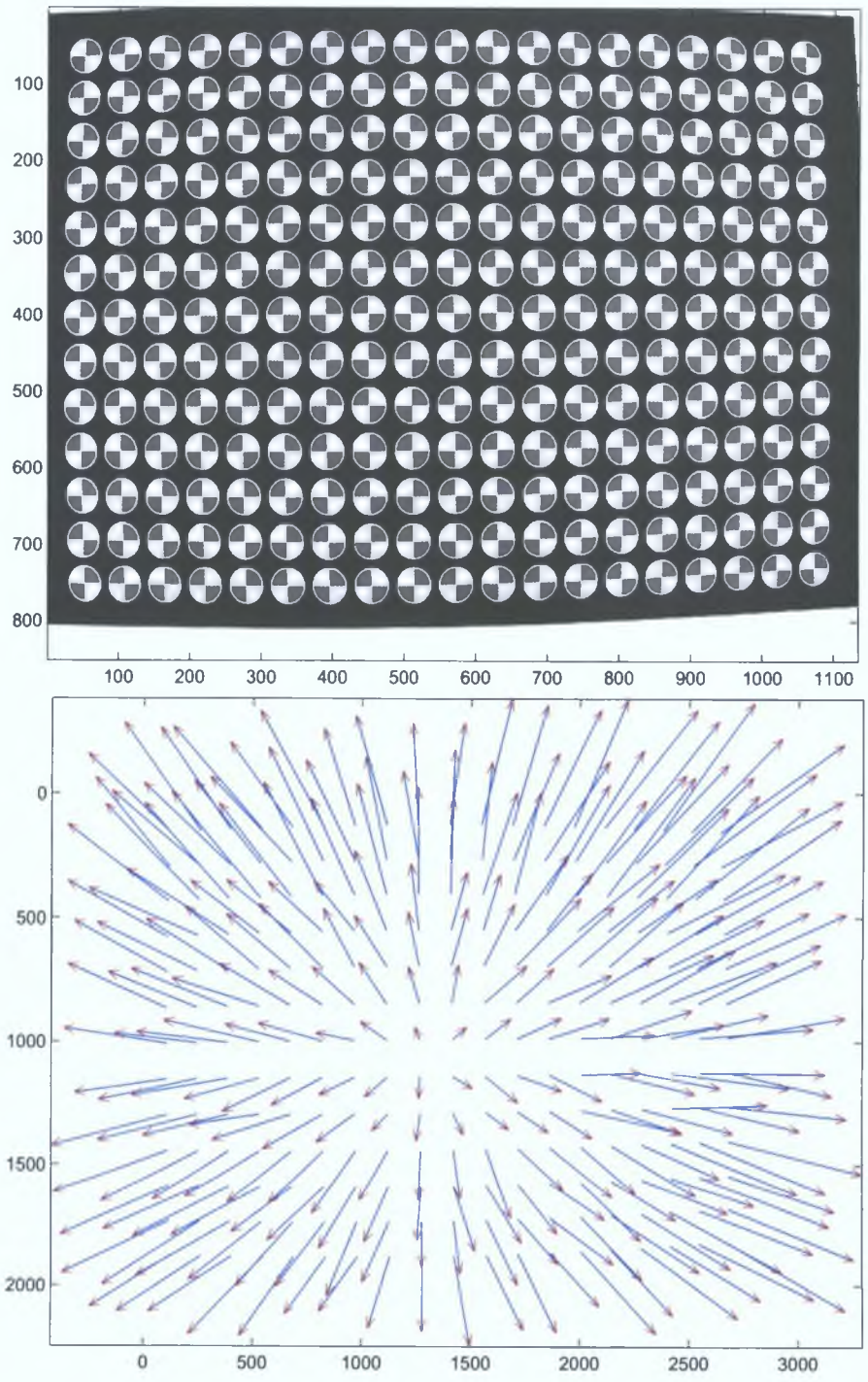


Fig. 3.13: Example of bias in real image (Fuji FinePix size: 2832×2128), evaluated using a specially designed pattern. Circle dia $\simeq 120$ pix, with estimated primary distortion term: $k_1 = -0.2$. Vector plot shows Centroid and Corner differences, revealing the distortion bias (scale is $\times 500$). Residual mean and SD are 0.817(0.269).

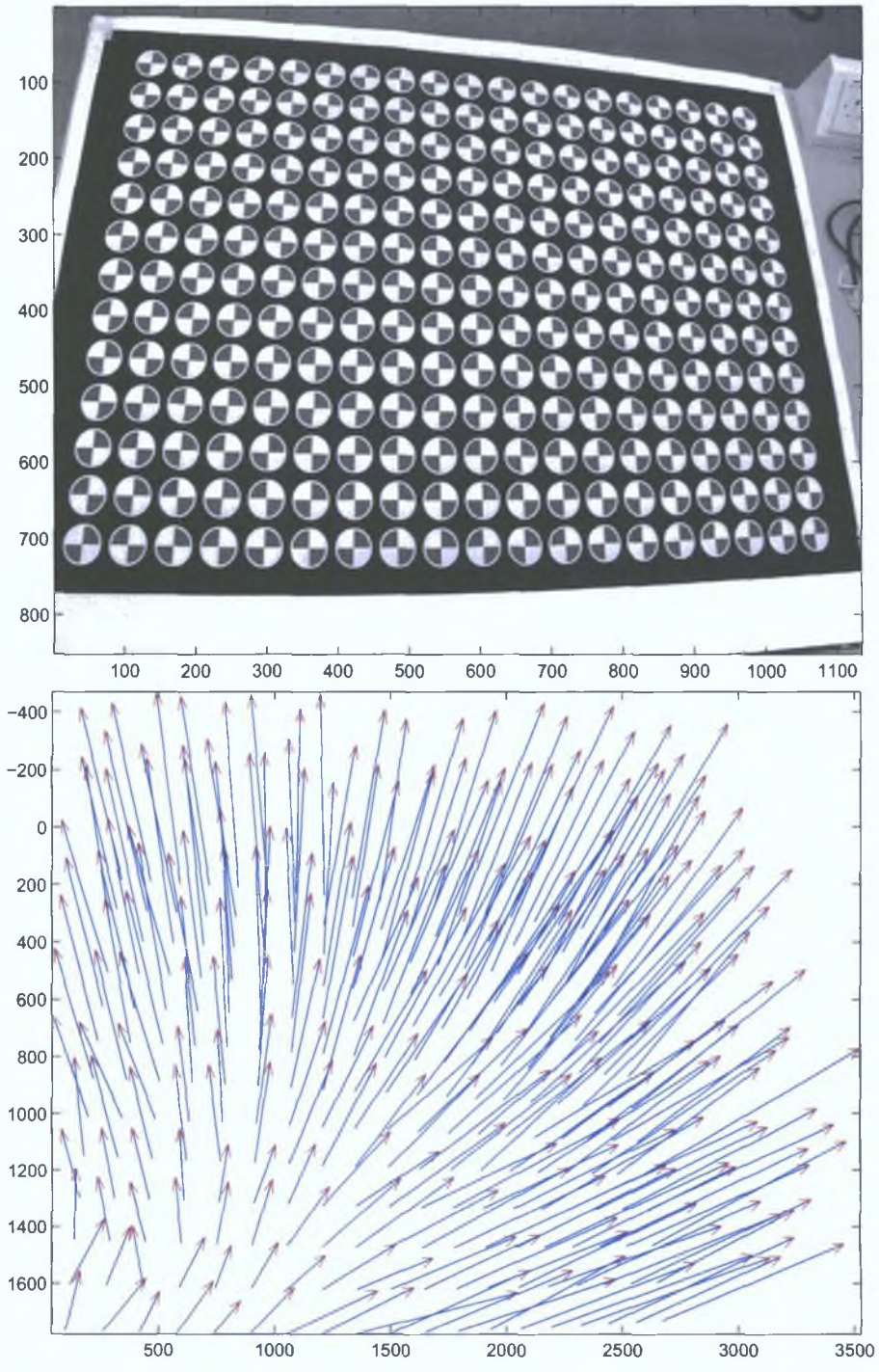


Fig. 3.14: Second example of bias in real image with the same camera and settings as Fig. 3.13 (Fuji FinePix size: 2832 × 2128). Vector plot shows Centroid and Corner differences, revealing the dominance of projective bias over distortion bias (scale is ×500). Residual mean and SD are 1.227(0.391).

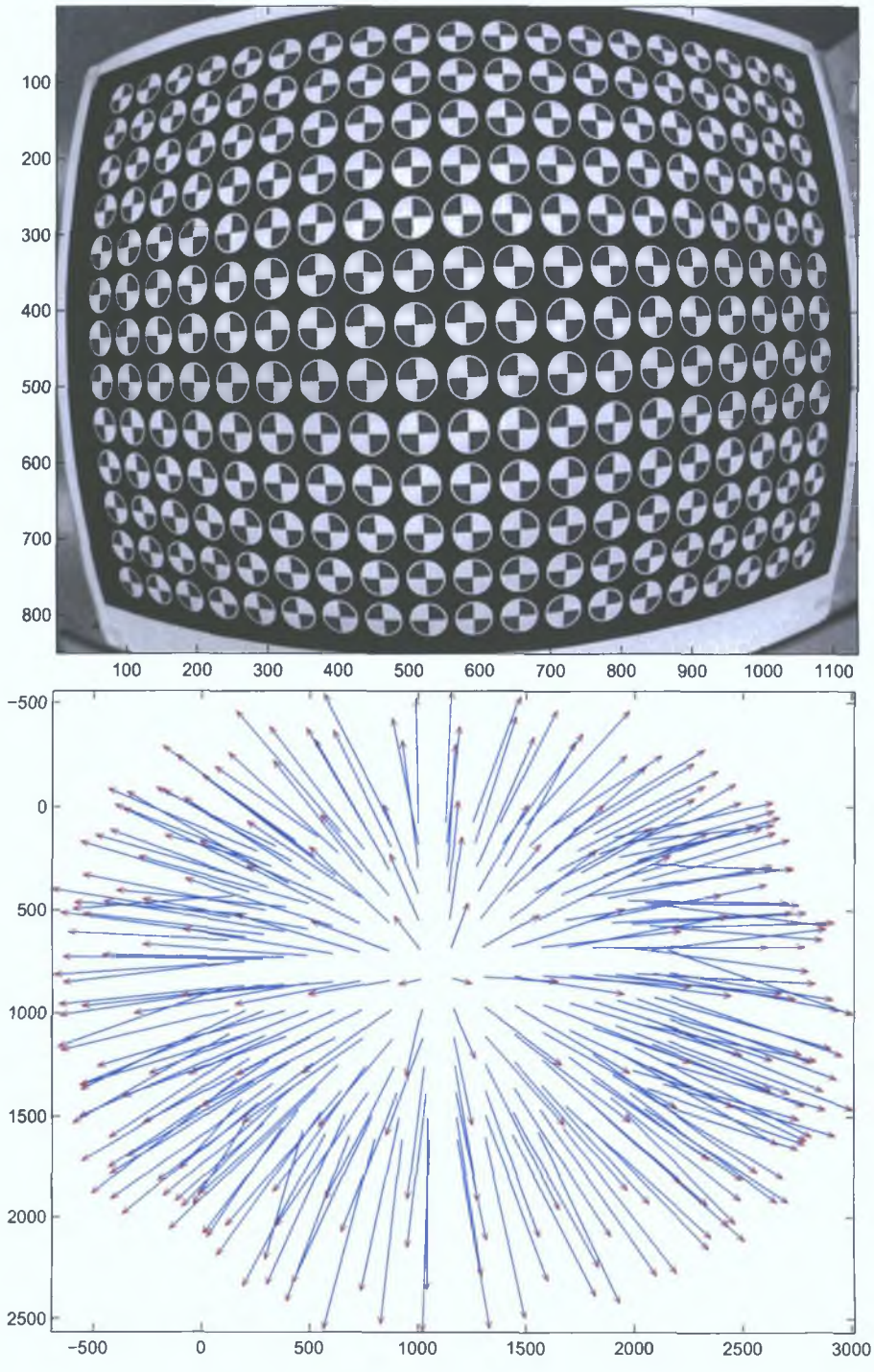


Fig. 3.15: Example of bias in real image (Nikon coolpix with FC-E8 lens size: 2272×1704). Circle dia $\simeq 110$ pix, with estimated primary distortion term: $k_1 = -1.2$. Vector plot shows Centroid and Corner differences, revealing the distortion bias (scale is $\times 500$). Residual mean and SD are 1.475(0.322).

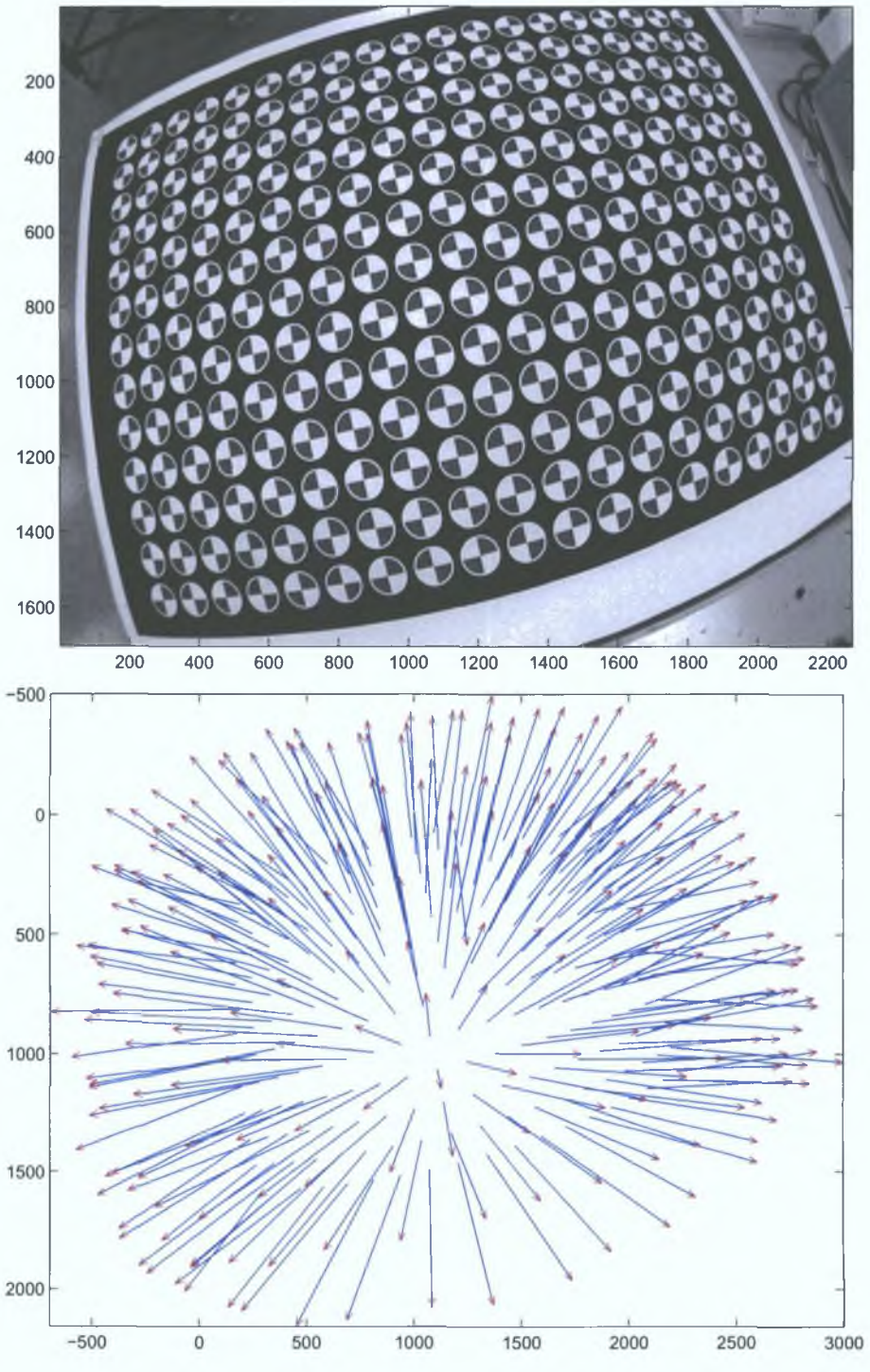


Fig. 3.16: Example of bias in real image with the same camera and settings as Fig. 3.15 (Nikon coolpix with FC-E8 lens size: 2272×1704). Vector plot shows Centroid and Corner differences, revealing the almost complete dominance of distortion bias (scale is $\times 500$). Residual mean and SD are 1.272(0.2711).

3.3.5 Chessboard detection noise

It is clear that a chessboard pattern with corner detection offers far superior data than circular patterns and methods. Naturally, throughout this thesis, control points are specified using a chessboard pattern. Of the two types of detection methods considered for use with this pattern, only corner detection offers bias free coordinates. However, as described in Section 3.2.2, the edge based method described is based on a large region of interest around the rough initial location. In practice, this region can be reduced to the same size as the surface fitting corner method, thus reducing distortion bias to negligible levels. This has the adverse consequences of making this line intersection method more susceptible to noise. In experiments, it was found that this line intersection method was more robust to poor initial corner estimates and various lens blurring. On the downside, the line based method in its current implementation is very much slower than the surface corner fitting. In experiments, each method was utilised.

It is therefore useful to assess the levels of noise in control points recovered with these two methods. The precision of the control points is also highly dependent on the camera and lens system employed. Predominantly, three digital cameras are used in this work, and are briefly described in table 3.2[3]. The recovered control point noise is assessed by taking multiple images of the calibration pattern with varying lighting conditions. The results are presented as a global noise measure for both small ROI line and corner detection methods in Figs. 3.17, 3.18 and 3.19 for the Kodak, Nikon and Fuji cameras respectively.

The Kodak camera shows high point accuracy, though this is influenced somewhat by the lower camera resolution. The line based method shows slightly better point localisation. Similarly for the Nikon camera, the edge based method slightly outperforms the corner method. For the high resolution Fuji camera, the line based method shows considerably lower noise. All errors display a zero mean normal distribution.

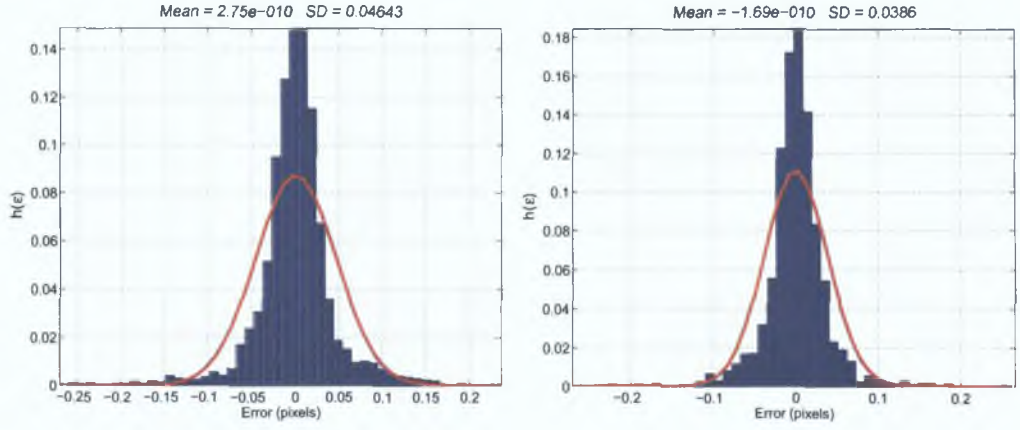
The noise dependence on the control point position within the image is finally examined, to access if the control point noise is random with respect to image location. This is conducted by taking the mean x and y errors for each control point over multiple images. The results are presented in vector format in Fig.

[3]These are the three cameras and various lens attachments that are available in the lab

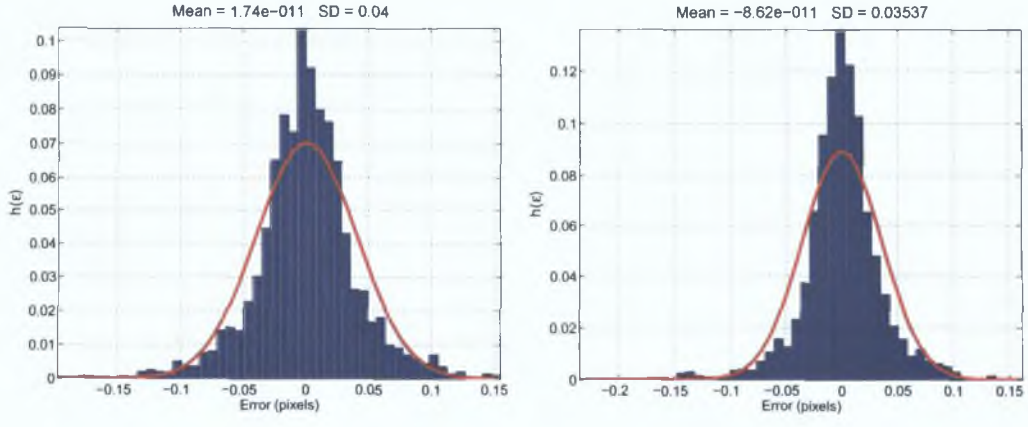
Table 3.2: Description of the three main cameras used for experiments in this work.

Camera Type	Resolution (pix)	Lens system
Kodak MegaPlus 4.1i	1312 × 1032	Computar 6mm 1/2"
Nikon CoolPix E4500	2272 × 1704	Nikon FC-E8 fish eye
Fuji Finepix 6900	2832 × 2128	Fuji WL-FX9 wide converter

3.20, for all three cameras. These plots show that there is no location dependent pattern observable for any of the cameras, and are henceforth considered random.

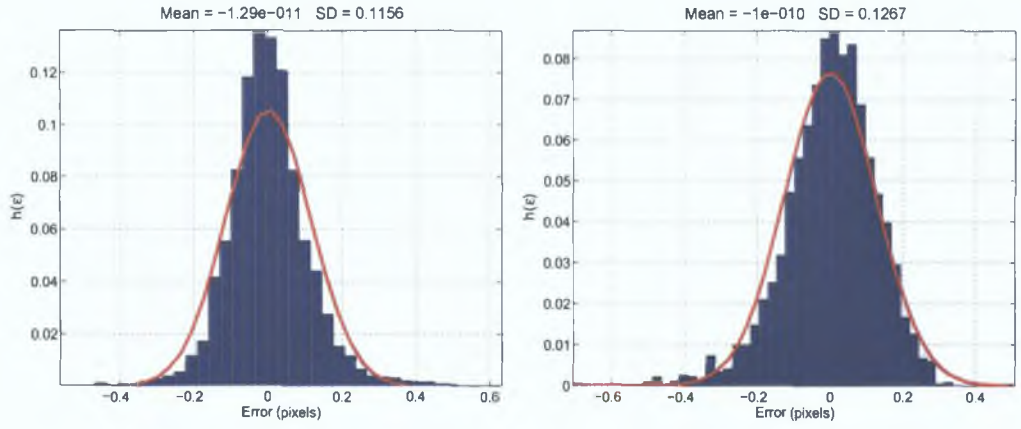


(a) Line Based Detection: Left x, Right y

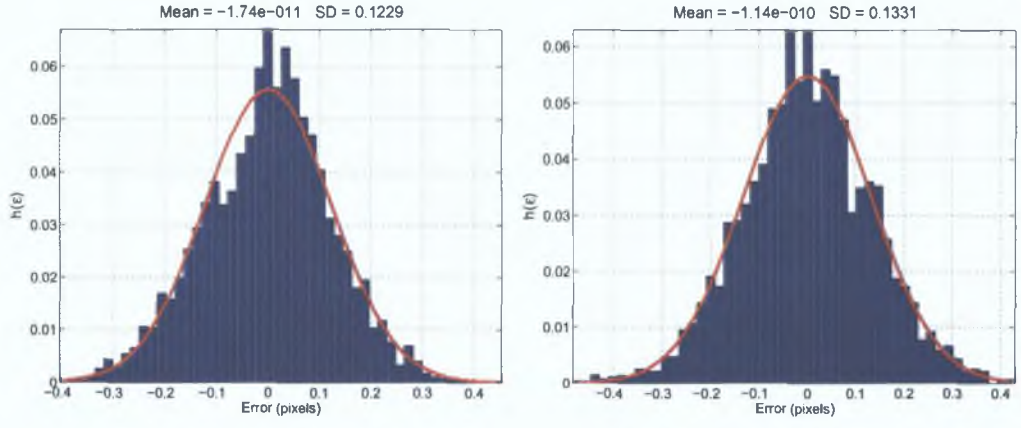


(b) Corner Based Detection: Left x, Right y

Fig. 3.17: Histogram of noise in control points using Kodak camera as described in table 3.2 for x and y directions including fitted normal distributions.

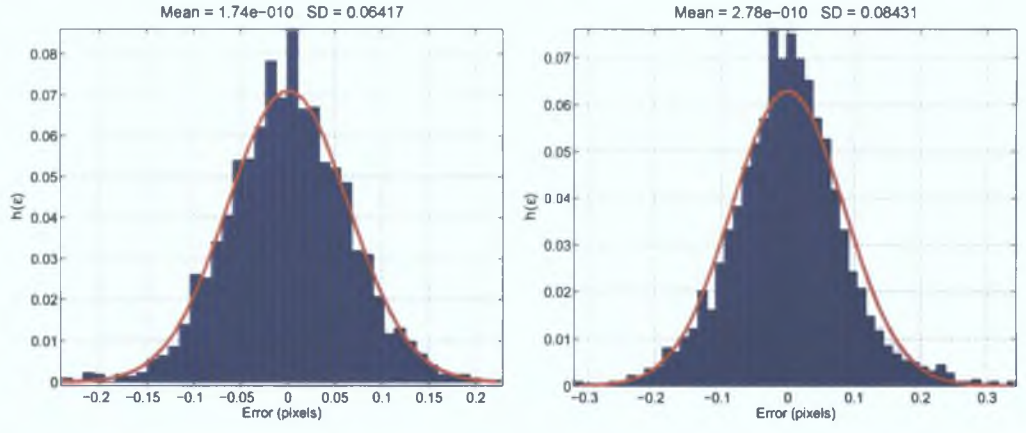


(a) Line Based Detection: Left x, Right y

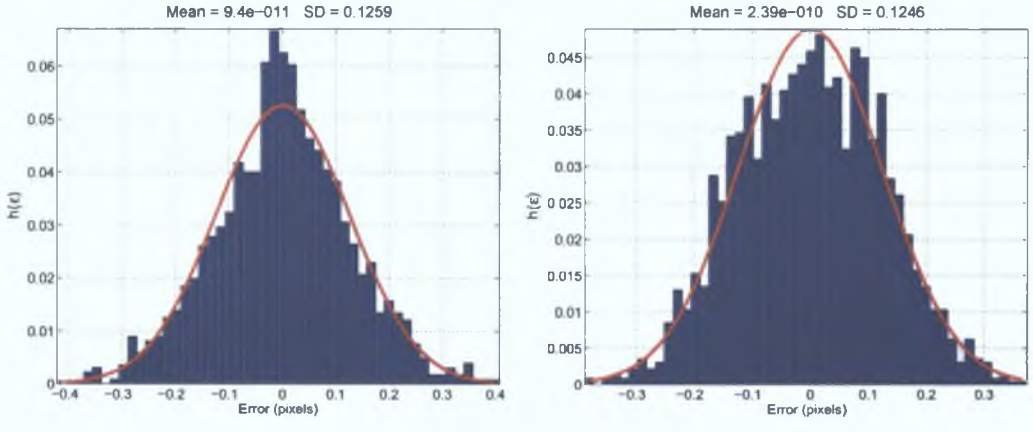


(b) Corner Based Detection: Left x, Right y

Fig. 3.18: Histogram of noise in control points using Nikon camera as described in table 3.2 for x and y directions including fitted normal distributions.

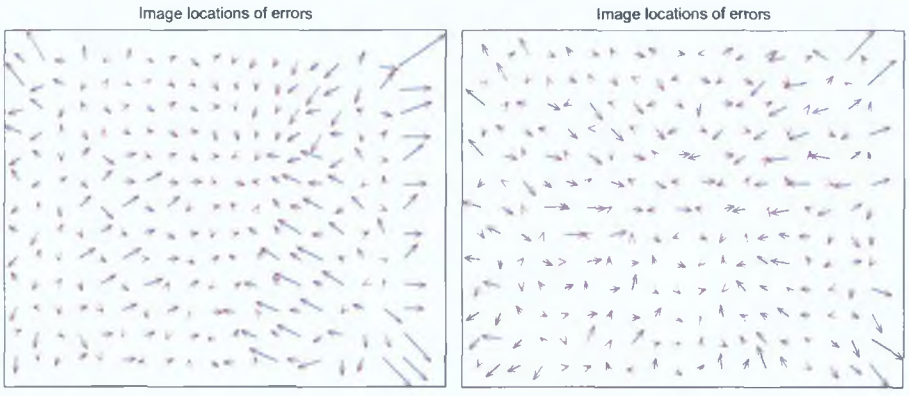


(a) Line Based Detection: Left x, Right y

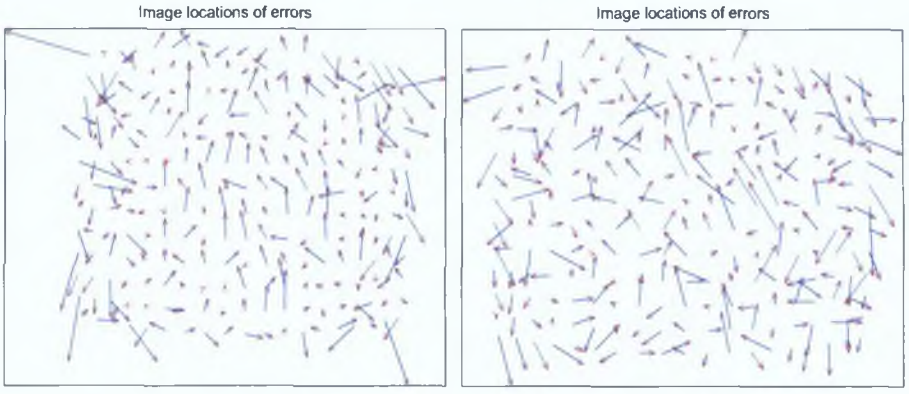


(b) Corner Based Detection: Left x, Right y

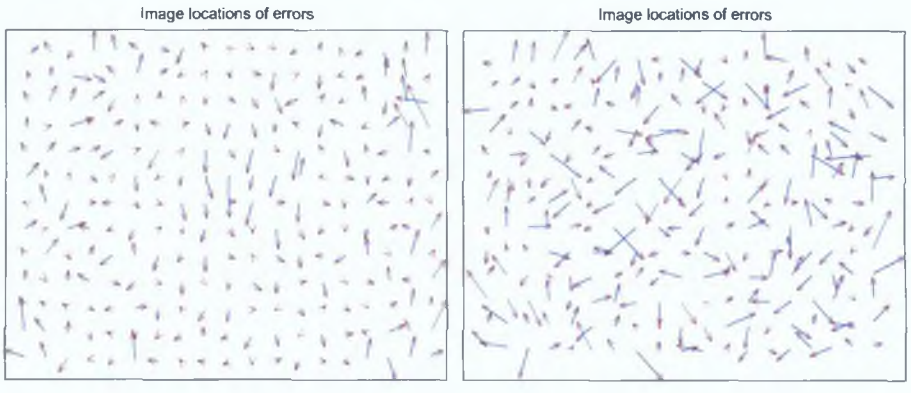
Fig. 3.19: Histogram of noise in control points using Fuji camera as described in table 3.2 for x and y directions including fitted normal distributions.



(a) Kodak camera. Left: Line errors, Right: Corner errors



(b) Nikon camera. Left: Line errors, Right: Corner errors



(c) Fuji camera. Left: Line errors, Right: Corner errors

Fig. 3.20: Location dependence of control point errors for three cameras. Vectors indicate the mean x and y noise magnitudes for multiple images. Vector scale = $\times 1000$.

3.4 Discussion

This chapter deals with control point recovery from planar calibration charts, by investigating if the choice of pattern can improve the overall detection precision. This accuracy is examined with respect to perspective transformations and lens distortion. Initially, pattern synthesis issues are detailed, in particular the generation of accurate lens distortion in images. Two representative types of patterns are considered: circles and chessboards, each with two common methods of control point recovery: centroids, conic fitting, edge approximation and corner points. We show theoretically and experimentally that compensated conic fitting, edge approximation and corner points are invariant to perspective bias, while only corner points are invariant to distortion bias. Simulated and real results indicate that distortion induced bias has a significant magnitude. Even for low distortion levels, roughly $\pm 0.3(pix^3)$, the biasing influence of distortion is greater than the noise/blur floor, and is more significant than the likely perspective bias encountered with normal calibration views. Thus, the compensation for perspective bias only with large conic area, is clearly not sufficient to acquire bias free control points. This has important implications for the field of high accuracy camera calibration.

It is clearly demonstrated that chessboard patterns are superior to circular type patterns in the generation of bias free control points. Therefore, throughout this thesis chessboard patterns are used this purpose. For the selection of cameras used, the control point noise from the chessboard detection methods is presented. It is very important to note that this study primarily considers a nominal pixel conic diameter of 50 pixels. It can be noticed throughout that for lower conic diameters the influence of biasing is significantly reduced. For example, a conic diameter of 10 pixels will induce negligible bias in comparison with the expected detection accuracy. Indeed, this diameter is the recommended rule of thumb for use of use of circular patterns in order to avoid incurring their biasing potential. The conclusions for this chapter must be regarded with this in mind.

Finally, some external factors that influence the precision of control points are not considered in this chapter. These factors relate to the manufacture of the calibration target, in terms of its planarity and the spatial precision of the particular pattern. The chessboard calibration target used in this thesis is fixed

to a glass substrate, while the black and white pattern was precisely manufactured using a high resolution xy table. Consequently, the errors introduced from the pattern imprecision are considered negligible. Chapter 4 further examines the influence of errors in the pattern precision, on the calibration of distortion.

Chapter 4

Distortion Calibration

Lens distortion is a thorn in the side of many relatively straightforward image analysis tasks. It comprehensively degrades the accuracy of measurements made in real images, where pixels in a normal perspective camera with a low focal length may move up to 30 pixels. In terms of the removal of such distortions, it is clear that, currently, the only precise means of calibrating and removing distortion is in conjunction with a full calibration of the internal and external parameters. The importance of distortion in this regard has seen the utter dominance of these more complicated algorithms at the expense of those that do not consider distortion. Despite this, many applications do not require the full complement of internal camera parameters and the relative orientation in relation to some calibration target. Selections of these, such as multiple view geometry estimation, in cases prefer to explicitly include lens distortion factors at the expense of extra complexity.

Algorithms for internal and external camera calibration, multiple view geometry estimation, etc. are rendered considerably simpler by the lack of lens distortion. However, despite the long term existence of non-metric calibration techniques to enact this removal of distortion, they have not been adopted. This can only be due to the poor performance of such methods, their limited circumstances of usage, and their own relatively complex practical implementation. This chapter comprehensively addresses the issue, in the description of an easy to use and highly precise method for calibrating lens distortion of all levels in perspective cameras.

The algorithm advanced in this chapter can be considered non-metric[1] as no internal or external camera parameters are explicitly available. The method exploits two geometric priors on the planarity of the control points and their known structure up to an arbitrary scale. Using a single view of this calibration pattern, an error function is formed in distorted space using the general distortion model. The necessary undistorted coordinates are simultaneously estimated through a general projective transform. The partial derivatives of the quadratic cost function are computed for all estimated parameters, allowing the closed-form computation of the cost gradients for minimisation. The re-sampling of an aberration free image is interpreted as a function in distortion free coordinate space, hence the calibrated forward distortion model may be applied directly. Lastly, automatic re-scaling is applied to balance the creation and loss of pixels in this re-sampling.

In comparison with the current benchmark for calibrating distortion through full camera calibration, the proposed approach has numerable advantages. Only one input image is required for the calibration, ideally taken in a roughly fronto parallel position. The distortion and associated parameters are estimated in a closed-form solution[2] with full partial derivatives, giving a computational advantage over current numerical techniques. The accuracy of the proposed method matches and surpasses that of complete calibration methods in many cases, while this accuracy is offered over a general coverage of all possible distortion levels in perspective cameras. It also offers an alternative means of dealing with distortion for the many tasks that do not require the full complement of camera parameters. Additionally, the subsequent calibration of internal and/or external camera parameters becomes much simpler in the absence of lens distortion.

A least square solution to the calibration problem is described in Section 4.1 through the description of the undistorted coordinates by a unique error homography. Full partial derivatives are given for use in a Gauss-Newton iterative solution described in Section 4.1.1. Section 4.1.2 describes a novel means of minimising the introduction of extra local pixel distortions in the generation of

[1]Non-metric in this context referrers to the unavailability of any internal or external camera parameters or orientations.

[2]Closed-form in this sense referrers to the direct geometric relationship between the criterion and the underlying model. This means that it is absolutely correct at all times not just upon convergence conditions.

an undistorted image. A comprehensive examination of the proposed method is given in Section 4.2 including comparisons with two popular full calibration techniques and one non-metric method. These comparisons are conducted on simulated and actual data over a extensive range of distortion. The calibration requires the input of at least one view of a calibration pattern. The dependence of the proposed method on this control data is examined in Section 4.3. Included is an description of the required quantity of control points for successful calibration in Section 4.3.1 which is extended to a multiple view variant of the algorithm. Two geometric priors are assumed on this calibration target, that of planarity and of known (up to an arbitrary scale) canonical coordinates for the pattern defined control points. The impact of violations of these constraints are investigated in Section 4.3.2. Additionally, it is assumed that the image projection of this calibration pattern has been detected free of distortion and projective bias (refer to chapter 3). The influence of bias is further examined for distortion calibration in Section 4.3.3. A comprehensive analysis of decentering distortion with the proposed technique is presented in Section 4.4. Lastly, variable lens parameters that influence the observed distortion are examined in Section 4.5 for the fundamental lens functions of focusing and aperture settings.

4.1 Least squares simultaneous calibration

The calibration technique pursued here requires one view of a calibration pattern. Two geometric priors are assumed on this calibration target, that of planarity and of known (up to an arbitrary scale) canonical coordinates for the pattern defined control points. Additionally, it is assumed that all control points are observable within the image window and are appropriately ordered as described in chapter 3.

Considering the bias free detection of chessboard intersections with image coordinates $\bar{\mathbf{c}} = (\bar{u}, \bar{v}, 1)^T$ and units of pixels (pix), a transform is required to convert these measurements into the lens centered coordinate space. The lens centric coordinate system can be thought of as the actual metric measurements of the ray intersections with the surface of the sensor pickup (e.g. CCD array), centered on the optical axis of the lens. As most sensors are not aligned accurately with respect to the lens, the centre of the sensor array cannot be

assumed to contain the optical axis. Indeed it is unclear if lens distortion is actually centered on the optical axis or not. Additionally, the actual physical size and resolution of the array is variable from camera to camera. Fortunately, the precise metric values of the ray intersections are not required since they only influence the scale of the distortion parameters as follows:

Result 3. *Arbitrary scaling of lens centric coordinates, or equivalently varying sensor resolution, causes the parameter values of distortion to change, not the number of parameters:*

$$\lambda(\breve{\mathbf{p}} - \mathbf{p}) = \lambda\mathcal{D}(\mathbf{p}, \mathbf{k}) = \mathcal{D}(\lambda\mathbf{p}, \mathbf{k}_\lambda), \tag{4.1}$$

where λ is the scaling factor and $\mathbf{k}_\lambda$ is the scaled parameter vector.

Proof.

$$\begin{aligned}\lambda\mathcal{D}(\mathbf{p}, \mathbf{k}) &= \mathcal{D}(\lambda\mathbf{p}, \mathbf{k}_\lambda), \\ &= \lambda \begin{pmatrix} k_1\lambda^2 x r^2 + k_2\lambda^4 x r^2 + \dots \\ k_1\lambda^2 y r^2 + k_2\lambda^4 y r^2 + \dots \end{pmatrix}, \\ &= \lambda\mathcal{D}(\mathbf{p}, \mathbf{k}),\end{aligned}$$

as $\mathbf{k}_\lambda$ absorbs the scaling factors as: $\mathbf{k}_\lambda = (\lambda^2 k_1, \lambda^4 k_2, \dots)^T$ □

As a consequence, the recovered control points $\bar{\mathbf{c}}$ are normalised roughly to unit length by the average of the image width and height, $\breve{\mathbf{c}} = (\bar{u}, \bar{v}, w)^T$, to standardise the scaling of the distortion parameters and improve the conditioning of the estimation equations. The required lens centric transform is then defined as:

$$\breve{\mathbf{p}} = \begin{pmatrix} s & 0 & -u_o \\ 0 & 1 & -v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{u} \\ \bar{v} \\ w \end{pmatrix} = \mathbf{A}\breve{\mathbf{c}} \tag{4.2}$$

where s is the applied compensation for non-square pixels and the symmetrical centre of distortion is $\mathbf{t} = (u_o, v_o, 1)^T$. These parameters need to be estimated in the calibration. The units of $\breve{\mathbf{p}}$ are normalised pixels, (pix).

Given the normalised coordinates, $\breve{\mathbf{c}}$, of an arbitrarily orientated planar target, a planar transform or homography (Hartley and Zisserman, 2003) can be computed as shown in Fig. 4.1 by $\breve{\mathbf{c}} = \mathbf{H_d}\bar{\mathbf{w}}$, where $\bar{\mathbf{w}} = (\bar{x}, \bar{y}, 1)^T$ are the arbitrarily scaled 2D coordinates of the planar data points. It follows that:

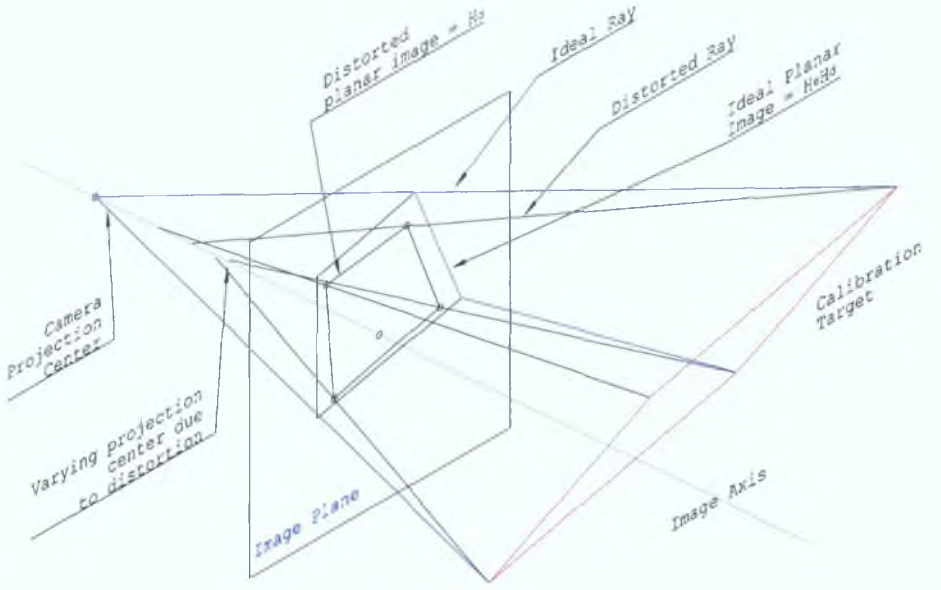


Fig. 4.1: A graphical interpretation of the formation of a distorted image and its undistorted counterpart.

Result 4. *The true (distortion free) projections of planar points, viewed from an arbitrary position, can be related to their distorted counterparts through the application of a general projective transform* $\mathbf{H_e}$*:*

$$\min \sum (\mathbf{c}, \mathbf{H_e}\breve{\mathbf{c}})^2 \quad \textit{giving} \quad \mathbf{c} = \mathbf{H_e}\tilde{\mathbf{c}} \tag{4.3}$$

Proof. $\tilde{\mathbf{c}}$ and $\breve{\mathbf{c}}$ are equivalent in a Least Squares sense, giving rise to equivalent $\mathbf{H_e}$ with the alternate objectives:

$$\min \sum (\mathbf{c}, \mathbf{H_e}\breve{\mathbf{c}})^2 = \min \sum (\mathbf{c}, \mathbf{H_e}\tilde{\mathbf{c}})^2,$$

since $\tilde{\mathbf{c}} = \mathbf{H_d}\bar{\mathbf{w}}$, where $\mathbf{H_d}$ originates from $\min \sum (\breve{\mathbf{c}}, \mathbf{H_d}\bar{\mathbf{w}})^2$, and $\bar{\mathbf{w}}$ are the canonical coordinates of the control points with unknown scale. □

This represents the key element in expressing distortion in a closed-form solution, enabling the recovery of the previously unavailable data $\mathbf{c}$, through the simultaneous estimation of $\mathbf{H_e}$. Without loss of generality, $\mathbf{H_e}$ is scaled so that h_{33} is one. In practice, no further constraints are available on $\mathbf{H_e}$ as the estimate of $\mathbf{H_d}$ is corrupted due to noise. A general form for $\mathbf{H_e}$ is then required to absorb these inaccuracies.

For each observed control point $\breve{\mathbf{c}}_i$ the following error function can be formed using the general forward distortion model (Result 1, Chapter 2):

$$\mathbf{e}_i(\breve{\mathbf{c}}_i, \mathbf{\Phi}) = \mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}_i + \mathcal{D}(\mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}_i, \mathbf{k}) - \mathbf{A}\breve{\mathbf{c}}_i, \tag{4.4}$$

where the full parameter vector is $\mathbf{\Phi} = (h_{11}, h_{12}, \ldots, h_{32}, s, u_o, v_o, k_1, k_2, \ldots)^T$ with $\mathbf{\Phi} \in \mathbb{R}^{n_\Phi}$.

4.1.1 Solving the problem

A parameter counting exercise reveals that for a parameter vector of length n_Φ a minimum of $m = \text{ceil}(n_\Phi/2)$ control point observations are required, where ceil() is a rounding towards $+\infty$. Given at least $n \geq m$ observations a Least square solution is obtained by:

$$\min \sum_{i=1}^{n} \mathbf{e}_i(\breve{\mathbf{c}}_i, \hat{\mathbf{\Phi}})^2 = \begin{pmatrix} \mathbf{e}_i(\breve{x}_i, \hat{\mathbf{\Phi}})^2 \\ \mathbf{e}_i(\breve{y}_i, \hat{\mathbf{\Phi}})^2 \end{pmatrix}. \tag{4.5}$$

This problem is nonlinear in parameters particularly due to $\mathcal{D}(\mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}_i, \mathbf{k})$ requiring a nonlinear optimisation solution. One way of solving this is to linearise equation 4.5 with some initial parameter value of $\mathbf{\Phi}_0$, resulting in an iterative Gauss-Newton scheme (Walter and Pronzato, 1997) which can be solved using many robust least square techniques (Golub and Loan, 1996):

$$\hat{\mathbf{\Phi}}_{\mathbf{k}+1} = \hat{\mathbf{\Phi}}_\mathbf{k} - \lambda \left(\frac{\partial \mathbf{e}^T(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}} \frac{\partial \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \right)^{-1} \frac{\partial \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k}), \tag{4.6}$$

where $\lambda \leq 1$ ensures a decrease in cost at each step. The full partial derivatives are given in Appendix B. It has been shown in Section 3.3.5 that the data covariances are equal, so no covariance matrix is required in this solution (covariance is effectively equal to an identity matrix).

An initial estimate for the parameters of $\mathbf{A}$ are $s = 1$, $u_o = m_x$ and $v_o = m_y$ where m_x and m_y are the normalised coordinates of the image centre. An initial estimate for the parameters h_{11}, h_{22} and k_1 may be obtained directly from the linear solution of:

$$(h_{diag}, k_i)^T = \mathbf{T}^+\mathbf{b}, \quad \text{where} \quad \mathbf{T} = \begin{pmatrix} \tilde{x}_1 & \tilde{x}_1\tilde{r}_1^2 \\ \tilde{y}_1 & \tilde{y}_1\tilde{r}_1^2 \\ & \vdots \\ \tilde{y}_n & \tilde{y}_n\tilde{r}_n^2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \breve{u}_1 - m_x \\ \breve{v}_1 - m_y \\ \vdots \\ \breve{v}_n - m_y \end{pmatrix}, \tag{4.7}$$

where $\tilde{x} = \tilde{u} - m_x$, $\tilde{y} = \tilde{v} - m_y$, $\tilde{r}^2 = \tilde{x}^2 + \tilde{y}^2$ and $(.)^+$ denotes the pseudo inverse. The parameter vector including three terms for radial distortion is then initialised at:

$$\mathbf{\Phi} = (h_{diag}, 0, -m_x, 0, h_{diag}, -m_y, 0, 0, 1, -m_x, -m_y, k_i, 0, 0)^T. \tag{4.8}$$

An overview of the algorithm is now presented.

Algorithm 1 The algorithm for estimating general lens distortion based on one view of a calibration pattern

---

Objective

Given $n \geq m$ control points with image coordinates $\bar{\mathbf{c}}$, from a single view of a planar calibration target with known structure, $\bar{\mathbf{w}}$, determine the parameters of the forward lens distortion model.

Outline

1. **Normalisation** Observed image coordinates $\bar{\mathbf{c}}$ are normalised by the average of the image width and height giving $\breve{\mathbf{c}}$. A least square homography $\mathbf{H_d}$ is formed giving: $\breve{\mathbf{c}} = \mathbf{H_d}\bar{\mathbf{w}}$

2. **Initialisation** Parameter vector is initialised (equation 4.8) by directly solving equation 4.7.

3. **Minimisation** Distortion parameters are estimated by minimising equation 4.5 through the iteration of equation 4.6 until convergence. Additionally the undistorted coordinates are recovered as: $\mathbf{p} = \mathbf{H_e}\mathbf{H_d}\bar{\mathbf{w}}$.

4.1.2 Balancing pixel warping

Resampling or image warping computes new samples on a target image from original samples on the source image. Two frequently used and well known interpolation filters are nearest neighbor and bilinear interpolation (Keys, 1981, Heckbert, 1989). The formation of a new undistorted image $\mathcal{I}(\mathbf{x})$ can be formed knowing the forward distortion model parameters from the original image $\breve{\mathcal{I}}(\mathbf{x})$ as:

$$\mathcal{I}(\mathbf{x}) = \mathcal{R}\left(\breve{\mathcal{I}}(\mathbf{x} + \mathcal{D}(\mathbf{x}, \hat{\mathbf{k}}))\right), \tag{4.9}$$

where $\mathcal{R}(.)$ is the interpolation method. Throughout this thesis standard bilinear interpolation is used, which is available in many image processing such

as Intel Image Processing Library (IPL, 2000). This formulation has the important advantage that the calibrated forward distortion model can be used directly, without need for an inverse function as is required in Heikkila and Silven (1997) and Heikkila (2000).

One problem with the modelling of distortion according to equation 4.4 is that it does not consider the optimal formation of a new distortion free image. Thus, for barrel distortion, image warping according to equation 4.9 results in the stretching of central pixels to occupy the viewing window, and hence the loss of outer perimeter pixels. The reverse case is noticed in the literature, such as Micusik and Pajdla (2003) and Fitzgibbon (2001), where the new image is compressed into the viewing window, due to the use of an inverse distortion model in the re-sampling. This image stretching effectively results in a reduction of window size, as illustrated in Fig. 4.2, which rises two issues. Firstly, there is a reduction in the effective field of view or angle of view. This is counter productive as wide angle lenses are chosen for their wide angle of view. Secondly, there is an increase in noise and a degradation of image quality in comparison to the original image, due to the creation of new pixels by zooming-like action of the warping.

To address this issue, the image warping of equation 4.2 is modified, by taking into account local pixel distortions. These pixel distortions are the result of the increasing pixel area warping, radially from the centre of the lens distortion. By introducing the ability to compress pixels in the resampling process, the overall pixel distortions can be balanced or minimised. This is accommodated by introducing an appropriate scaling matrix $\mathbf{S}$ into the resampling:

$$\mathcal{I}(\mathbf{x}) = \mathcal{R}\left(\breve{\mathcal{I}}(\mathbf{S}\mathbf{x} + \mathcal{D}(\mathbf{S}\mathbf{x}, \hat{\mathbf{k}}))\right), \tag{4.10}$$

where the scaling matrix contains one variable s_1 as:

$$\mathbf{S} = \begin{pmatrix} s_1 & 0 & s_1\hat{u}_o - \hat{u}_o \\ 0 & s_1 & s_1\hat{v}_o - \hat{v}_o \\ 0 & 0 & 1 \end{pmatrix},$$

and $(\hat{u}_o, \hat{v}_o)$ is the estimated distortion centre.

A measure of the local pixel distortion can be compiled by taking the Jacobian of the local area around point $\mathbf{p}$ as:

$$\mathbf{J}(\mathbf{p}) = \begin{pmatrix} 1 + \frac{\partial \mathcal{D}_x(\mathbf{p},\mathbf{k})}{\partial x} & \frac{\partial \mathcal{D}_x(\mathbf{p},\mathbf{k})}{\partial y} \\ \frac{\partial \mathcal{D}_y(\mathbf{p},\mathbf{k})}{\partial x} & 1 + \frac{\partial \mathcal{D}_y(\mathbf{p},\mathbf{k})}{\partial y} \end{pmatrix}. \tag{4.11}$$

This gives the the size ratio of the newly created pixel at the location $\mathbf{p}$ to its original location $\breve{\mathbf{p}}$. Letting λ_1 and λ_2 be the eigenvalues of $\mathbf{J}(\mathbf{p})$ then the net measure of the local pixel distortion may be obtained by taking its determinant as $\det(\mathbf{J}(\mathbf{p})) = \lambda_1\lambda_2$. Since the eigenvalues of an orthogonal transform are equal to one, the ideal value of the determinant causing no net size change is also one. A compression of pixel size results in a determinant of less than one, while the enlarging of pixel size results in a determinant greater than one.

Pixel distortion can therefore be minimised by choosing a scaling parameter s_1 that minimises:

$$\min \sum_{i=1}^{n} \left(\det(s_1 \mathbf{I}_{2\times 2}\mathbf{J}(\mathbf{p}_i)) - 1\right)^2. \tag{4.12}$$

As $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) \quad \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times m}$ and $\det(s_1\mathbf{I}_{2\times 2}) = s_1^2$ this minimisation can be solved for linearly as follows:

$$s_1 = \sqrt{\frac{1}{\frac{1}{n}\sum_{i=1}^{n} \det(\mathbf{J}(\mathbf{p}_i))}}. \tag{4.13}$$

The point set $\mathbf{p}_i$ may be chosen as a grid covering the entire image area. The density of this grid has a very minor influence on the overall scaling parameter s_1.

This improved resampling, by considering the local pixel distortions, is illustrated in Fig. 4.2. It can be observed that the scaling solution avoids including peripheral areas of the original image as these areas would induce large pixel distortion after image warping. The balance that is reached is a clear improvement in terms of the resulting field of view over the original unscaled method. Note, that this formulation assumes that the new image dimensions are the same as the original. An alternative means of implementing the same effective field of view is to enlarge the destination image. The optimal enlargement in this case is again described by the scaling matrix $\mathbf{S}$ solved with equation 4.13.

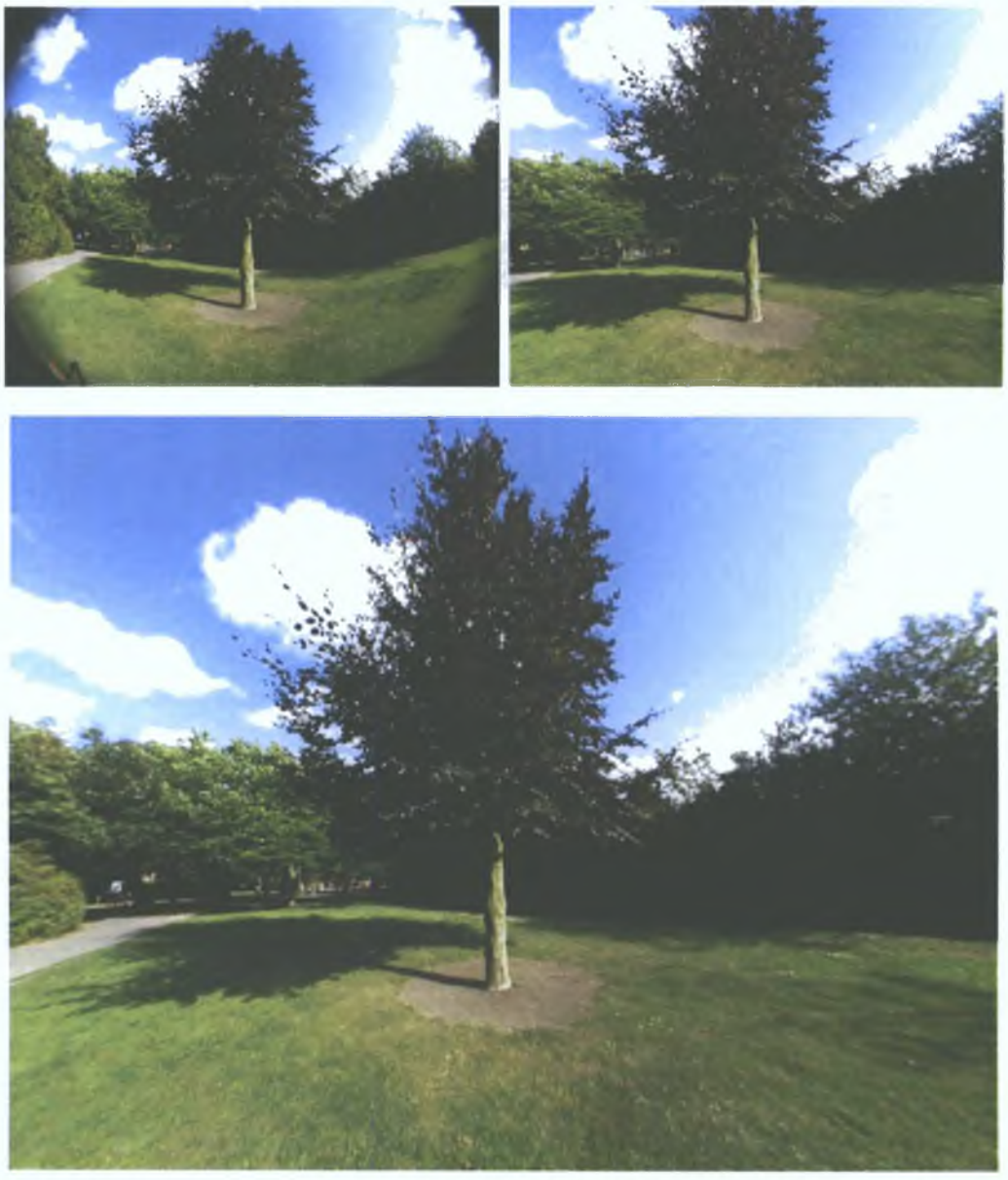


Fig. 4.2: Top: Original distorted image, and resampled image according to equation 4.9. Bottom image, adjusted resampling according to equation 4.10, showing the larger field of view.

4.2 Accuracy assessment

The proposed algorithm is compared with similar alternatives from the literature that take similar input data. The assessment is conducted in terms of evaluating how accurate the estimated parameters are in describing lens distortion throughout the image. In general it is not enough to merely find the best value for the parameters with respect to the objective criteria. It is also important to evaluate the uncertainty attached to this result, taking into account the uncertainty in the data. Three methods are taken from the literature for comparison purposes. These include two full calibration methods and one truism based method. The proposed algorithm is referred to as Method 1 throughout.

From a least squares or nonlinear least squares estimation problem the residual error vector may be calculated. However, instead of calculating this residual an alternative one is compiled directly using new data. Under the assumption that data error has a zero mean normal random distribution $\mathcal{N}(0, \sigma_n^2)$ (see chapter 3 for detection patterns and methods that violate this assumption), the estimation algorithm is asymptotically efficient, implying that with enough data points the exact solution may be recovered. The parameter uncertainty can be estimated from the Fisher information matrix:

$$\mathbf{F}(\hat{\mathbf{\Phi}}) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial \mathbf{e}^T(\check{\mathbf{c}}_i, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \hat{\mathbf{\Phi}}} \frac{\partial \mathbf{e}(\check{\mathbf{c}}_i, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \hat{\mathbf{\Phi}}^T}. \tag{4.14}$$

$\mathbf{F}^{-1}(\hat{\mathbf{\Phi}})$ is then used to characterise the uncertainty in the parameters, by forming an estimate of the associated Standard Deviation (SD) as the square root of the i^{th} diagonal element as:

$$\mathrm{SD}(\hat{\Phi}_i) = \sqrt{\mathrm{diag}_i(\mathbf{F}^{-1}(\hat{\mathbf{\Phi}})} \tag{4.15}$$

In the special case where all the σ_n^2's are equal and possibly unknown:

$$\hat{\sigma}^2 = \frac{1}{n-m} \sum_{i=1}^{n} (\check{\mathbf{p}} - \mathbf{p} - \mathcal{D}(\mathbf{p}, \hat{\mathbf{k}}))^2 \tag{4.16}$$

$\mathbf{F}(\hat{\mathbf{\Phi}})$ is then approximated by:

$$\mathbf{F}(\hat{\mathbf{\Phi}}) = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^{n} \frac{\partial \mathbf{e}^T(\check{\mathbf{c}}_i, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \hat{\mathbf{\Phi}}} \frac{\partial \mathbf{e}(\check{\mathbf{c}}_i, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \hat{\mathbf{\Phi}}^T}. \tag{4.17}$$

4.2.1 Comparison Methods

With the proposed algorithm referred to as Method 1, the initial comparison technique is referred to as Method 2, which is taken from the popular technique of Zhang (1998, 2000), and available on the web at (Zhang, 2005). This full calibration technique takes multiple views of a planar pattern as input. Although distortion calibration is not the main focus of this technique (no decentering and a low order radial (2 parameters) model is used), it provides a useful comparison for lower distortion levels. It is not expected to be highly accurate for larger distortion levels. The method estimates distortion by the numerical solution of the back projection problem in n views with m control points in each view. The available implementation does not return the parameter standard deviations.

Method 3 is also a full calibration method, taken from Heikkila (2000) and available in a Matlab implementation on the web at (Heikkila, 2005). This method also requires multiple views of a planar calibration object, though 3D coordinates with one image can also be used. Distortion is given prominent focus in this implementation, with the reverse model assumed including decentering distortion. A simplified approximation of its Taylor expansion is used as an inverse (see chapter 2 for a full description of this model). The implementation returns the parameter SD's.

Method 4 is a truism based method, similar to that described in Haneishi et al. (1995), and others such as Asari et al. (1999) and Devernay and Faugeras (2001), where the estimation is based on the criterion that straight lines in object space must be imaged as straight lines. The input data is formed from the collinear coordinates of the chessboard pattern similarly with Haneishi et al. (1995) and Asari et al. (1999). This method is implemented by first calculating the lines joining the horizontal chessboard intersections as follows:

- Initialise the principal point, and 4 parameters of the reverse distortion model.
- A moment matrix is formed representing the algebraic line fit to the

input data as:

$$\mathbf{M} = \begin{pmatrix} \sum_{i=1}^{n} \breve{x}_i^2 & \sum_{i=1}^{n} \breve{x}_i\breve{y}_i & \sum_{i=1}^{n} \breve{x}_i \\ \sum_{i=1}^{n} \breve{x}_i\breve{y}_i & \sum_{i=1}^{n} \breve{y}_i^2 & \sum_{i=1}^{n} \breve{y}_i \\ \sum_{i=1}^{n} \breve{x}_i & \sum_{i=1}^{n} \breve{y}_i & n \end{pmatrix}.$$

- Calculate the LS line fit from the eigenvector associated with the smallest eigenvalue.
- Using this line and the input data form the error criterion.
- Minimise this using the L-M method for the sum of all lines.

Parameter SD's are available upon convergence. A summary of the different properties of each method is given in table 4.2.1. The majority of these methods do not include decentering distortion. In order to compare like-with-like decentering distortion is not included in the distortion model of the proposed algorithm, Method 1. An detailed analysis of decentering distortion is made in Section 4.4.

Table 4.1: Properties of the proposed algorithm (Method 1) in comparison to three alternative techniques.

Criteria	Method 1	Method 2	Method 3	Method 4
# of Views	1	Multi	Multi	1
View Data	Planar	Planar	Planar	Line
Solution	Closed	Numerical	Numerical	Numerical
Dist. Model	Forward	Forward	Reverse	Reverse
Cam. Cal.	×	✓	✓	×
Param. SD	✓	×	✓	✓
Online	–	✓	✓	×

4.2.2 Comparisons on Simulated data

Comparisons on simulated data are made with respect to varying control point locations and increasing levels of control point noise. Distortion is simulated using a non-standard formulation in order that the resulting profile does not exactly match with any of the calibration models. Considering the divisional model as presented in chapter 2, extra nonlinear terms can be added as follows:

$$\breve{\mathbf{c}} = \frac{\mathbf{c}}{1 - k_1 r^2 - k_2 r^4}, \tag{4.18}$$

where $\mathbf{c} = (u, v, 1)^T$ are the undistorted points, $\breve{\mathbf{c}}$ are the distorted counterparts and $r^2 = (u - u_o)^2 + (v - v_o)^2$. The centre of distortion $\mathbf{t} = (u_o, v_o, 1)^T$ is fixed at $u_o = 732.33$ and $v_o = 812.21$. Distortion parameter k_1 is varied in the range of $-2 \rightarrow 1.4$ with k_2 specified as $k_2 = -k_1/2$, representing a wide range of barrel and pincushion distortions. The lowest level of distortion is $k_1 = 0.01$, effectively showing on the graphs as zero.

The calibration data is comprised of a 10×10 planar grid of equally spaced points (spacing = 100 (pix)). Random positions of these points are simulated as in Chapter 3 using a 3D rotation randomly drawn from a specific range of rotation angles, with limited translation. Noise is added to these control points with a normal distribution $\mathcal{N}(0, \sigma^2)$, where σ^2 is in the range $0 \rightarrow 1$ (pix), well in excess of the expected noise range (see Section 3.3). The evaluation data consists of a 20×20 grid (grid spacing = 75 (pix)) covering the entire image window.

Random control point positions

The proposed algorithm is compared with the three alternative methods for robustness to control point positions. One hundred random control point positions are generated for each distortion level and the accuracy and stability of each method is evaluated. Accuracy is measured using the estimated parameters on the evaluation data set. The stability of each estimated parameter is quantified by it variance throughout the one hundred samples.

The accuracy of the four methods are compared in Fig. 4.3. The ability to correctly model distortion is primarily linked with the underlying distortion model. As Fig. 4.3 shows the reverse model in Method 3 and 4 is more unstable than the forward model used in Method 1 and 2 . This follows from the analysis in chapter 2 of the properties of these models. Method 4 is shown to be highly dependent on the location of the control points. Method 1 outperforms all other methods with much lower distortion residuals, indicating that the proposed algorithm is not dependent on the the location of control points and its use of a higher order model than Method 2 does not lead to instabilities.

Each method involves estimating the distortion centre. Fig. 4.4 shows the x and y errors for the estimated centre point for the four methods over 100

random placements of the calibration data and subject to varying distortion. This shows that the centre point in the line based method is quite sensitive to the location of the calibration data. The performance of Method 2 and 3 are roughly similar (excluding the peak errors at $k_1 = 0.01$ for Method 3). The proposed, method shows a very low centre point error in comparison. This demonstrates the high stability of the algorithm and the high accuracy afforded by the use of an appropriate distortion model.

The parameter values and standard deviations of each method are shown in Figs. 4.5 - 4.8. These show the variation of the parameters with distortion and their integrity with respect to varying control point positions. As expected the methods using the forward distortion model (Methods 1 and 2) show bounded parameter values. In contrast the parameter values for methods using the reverse model are much larger. The line based method shows large parameter and SD values, indicating in conjunction with its poor residual accuracy that this algorithm fails to converge for many of the position samples.

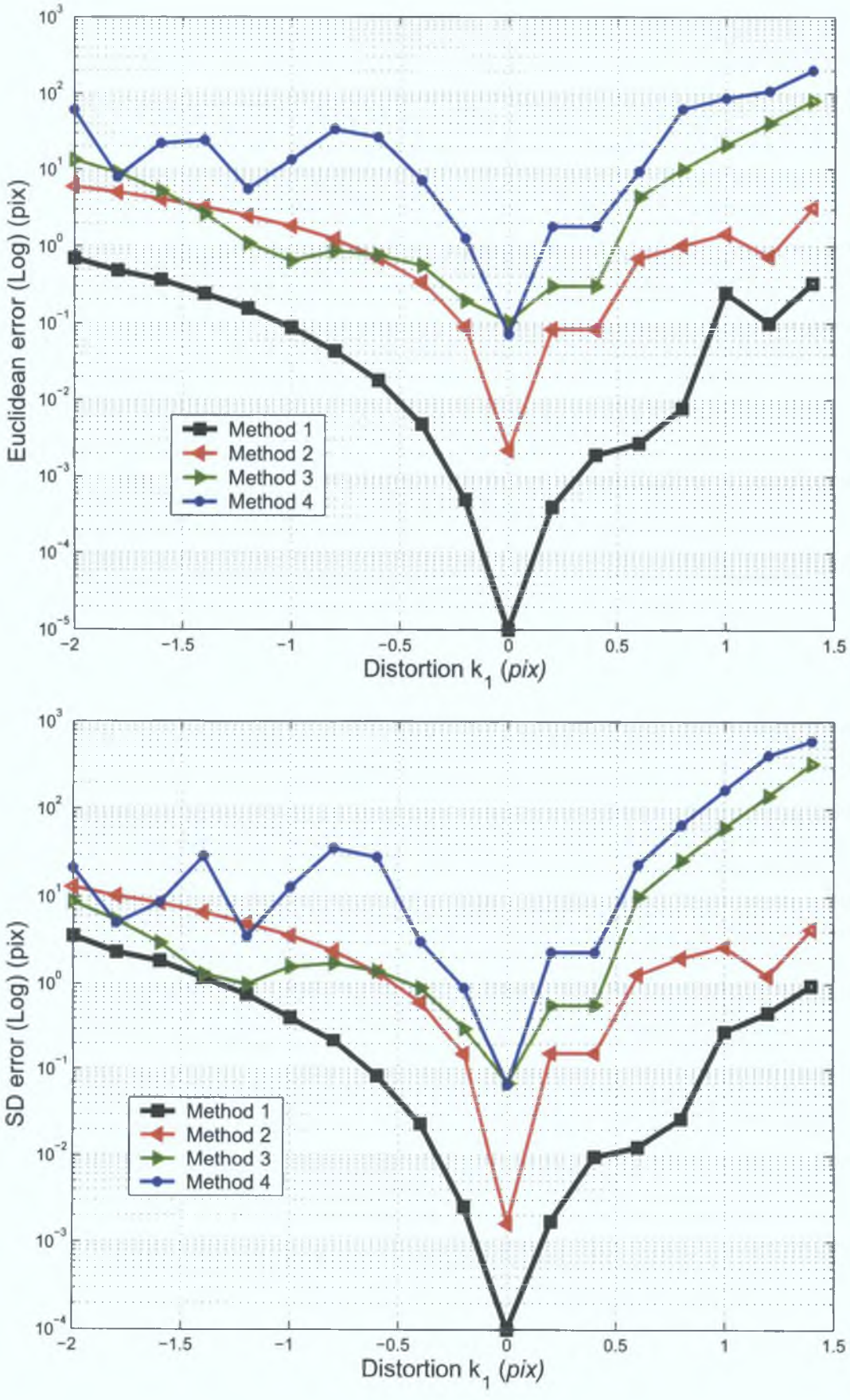


Fig. 4.3: Mean and SD errors for 4 methods computed over 100 random positions of control points.

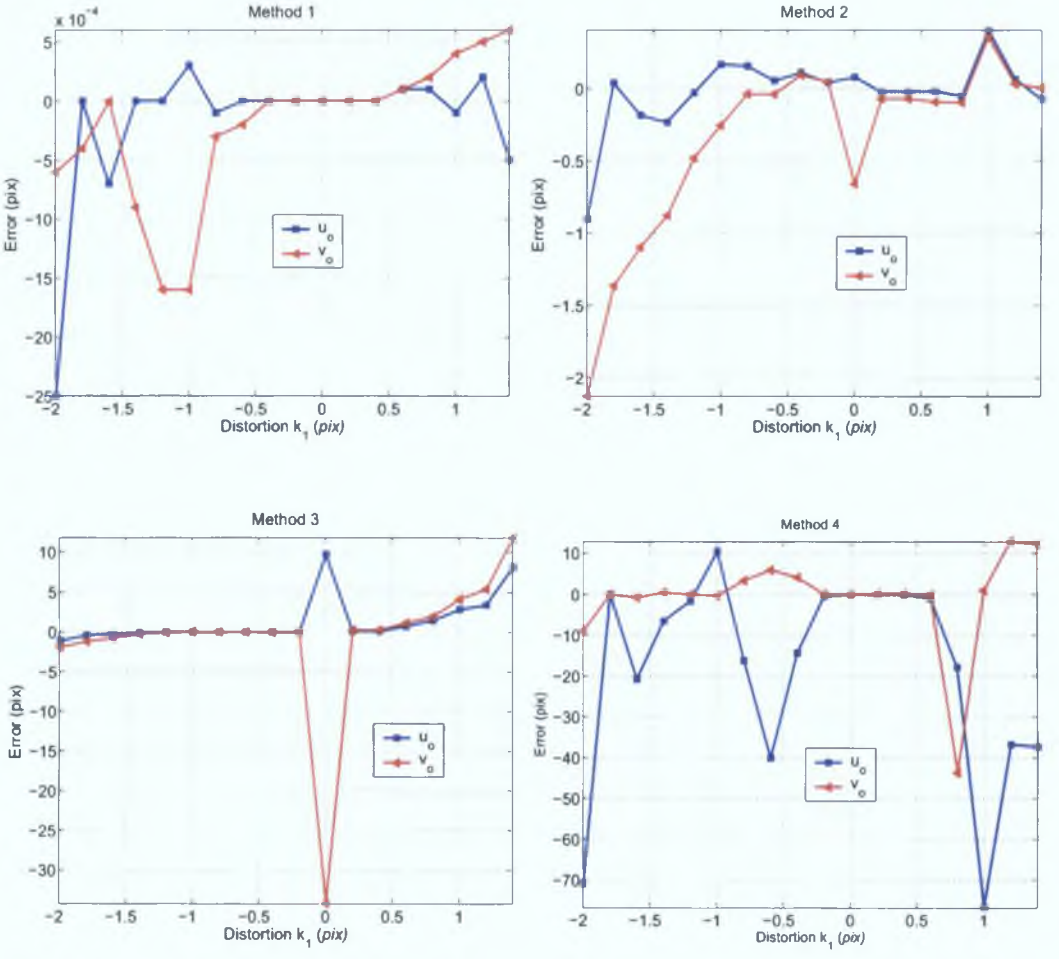


Fig. 4.4: Illustrates the mean estimated distortion centre less the ideal value, for 4 methods over the range of distortion.

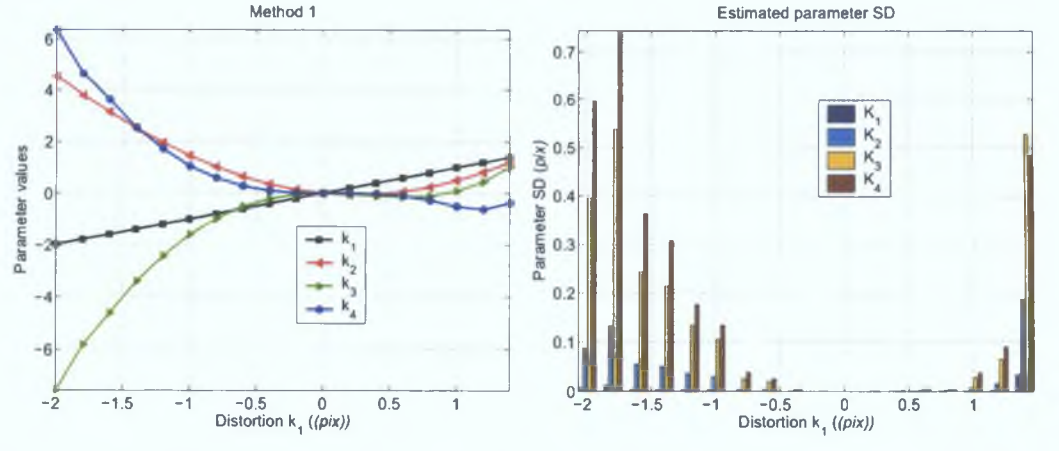


Fig. 4.5: Mean and SD of parameter values for Method 1 over the 100 randomly chosen control point locations.

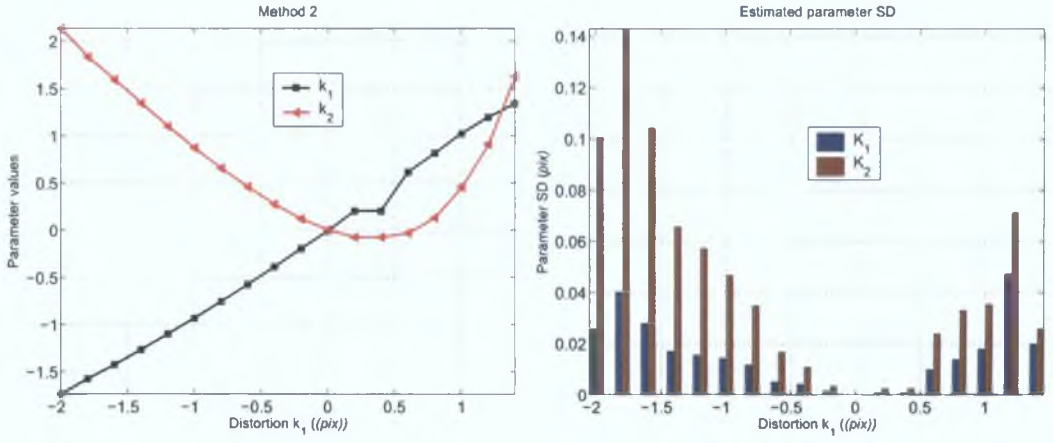


Fig. 4.6: Mean and SD of parameter values for Method 2 over the 100 randomly chosen control point locations.

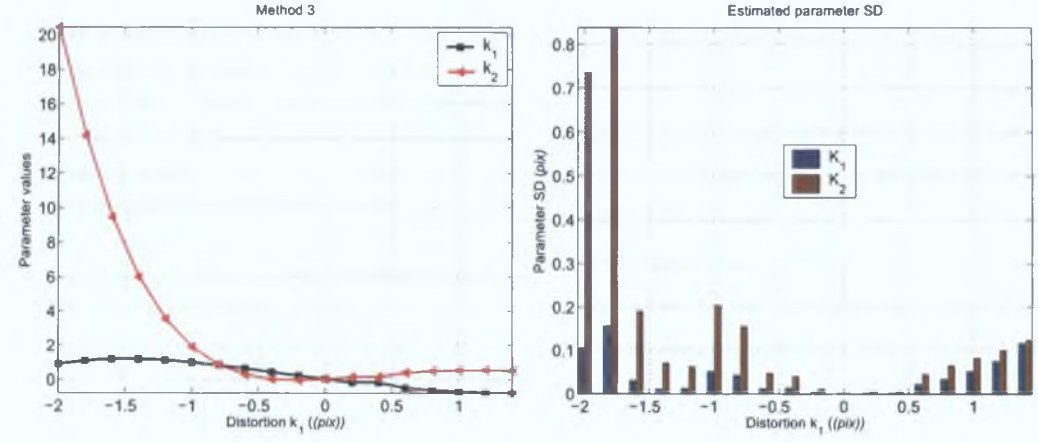


Fig. 4.7: Mean and SD of parameter values for Method 3 over the 100 randomly chosen control point locations.

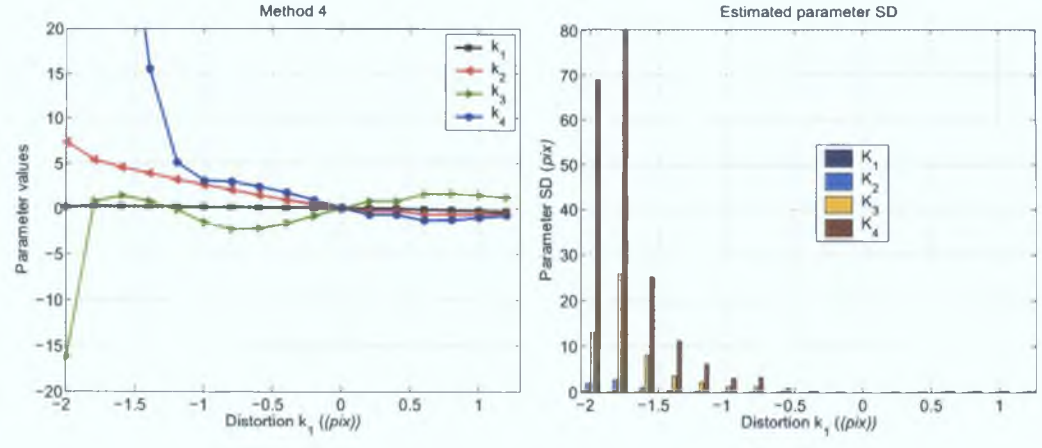


Fig. 4.8: Mean and SD of parameter values for Method 4 over the 100 randomly chosen control point locations.

Random noise in control points

The proposed algorithm is compared with the three alternative methods for robustness to noise in control point positions. The 3D position of the control points are fixed throughout. For each simulated distortion level, noise with a distribution of $\mathcal{N}(0, \sigma^2)$ is added for 100 samples. The noise variance σ^2 is then varied through $0 \rightarrow 1$ (pix). As this generates a significant quantity of data, only the residual distortion is presented, for the range of barrel distortion only.

The residual errors are presented in tables 4.2 - 4.5. Table 4.2 present the residual results for the proposed method, (Method 1). It can be noted from these results that for zero distortion levels, a larger residual error is induced than for the other distortion levels. This is due to the incorrect classification of noise in the data as actual distortion. This issue of low distortion is investigated further in Section 4.2.4. Excluding this zero level distortion the residual SD increases in direct proportion with the induced noise. All residuals remain in the subpixel range. It can be confidently concluded that the proposed algorithm copes successfully with noise.

In comparison, Method 2 shows a larger residual error. As this method also uses the forward model of distortion, its increase in error is proportional with the increase in noise variance. Method 3 uses the reverse model of distortion, and the results in table 4.4 show that the residual error is larger than that of Method 2. However, in terms of robustness to noise, the algorithm appears to cope successfully. The final line based method, Method 4, also uses the reverse model, but to a higher order. As the high order reverse model has been previously shown to have instability tendencies in Chapter 2, it is no surprise that the increase in noise levels induces a large increase in residual error. Indeed for some samples the algorithm failed to converge (nc), while convergence is questionable in some other cases.

Table 4.2: Residual errors for Method 1 in mean(SD) format.

k_1 / σ^2	0	0.2	0.4	0.6	0.8	1
0	0(0)	0.2358 (0.2202)	0.5061 (0.4798)	0.8739 (0.8051)	1.5543 (1.0397)	2.3643 (1.4354)
-0.2	0.0004 (0.0007)	0.2207 (0.1993)	0.4007 (0.3500)	0.6144 (0.5485)	0.9469 (0.8381)	1.0401 (0.9469)
-0.4	0.0033 (0.0057)	0.2039 (0.1715)	0.4246 (0.3671)	0.6096 (0.5305)	0.7837 (0.7404)	1.0038 (0.9216)
-0.6	0.0113 (0.0196)	0.1773 (0.1591)	0.3546 (0.3709)	0.5482 (0.5126)	0.7843 (0.7370)	0.9144 (0.8469)
-0.8	0.0263 (0.0466)	0.1848 (0.1855)	0.3558 (0.3197)	0.5494 (0.5390)	0.7712 (0.6792)	0.8102 (0.8446)
-1	0.0499 (0.0901)	0.1833 (0.1913)	0.3257 (0.3662)	0.5350 (0.5250)	0.7027 (0.6526)	0.8904 (0.8558)
-1.2	0.0826 (0.1530)	0.1803 (0.2138)	0.3262 (0.3651)	0.4948 (0.5114)	0.6774 (0.6311)	0.7748 (0.8862)
-1.4	0.1248 (0.2372)	0.1821 (0.2901)	0.3137 (0.3756)	0.483 (0.4655)	0.6680 (0.6826)	0.7415 (0.8500)
-1.6	0.1762 (0.3445)	0.2234 (0.3605)	0.3437 (0.4360)	0.4798 (0.5407)	0.6200 (0.7346)	0.7337 (0.8322)
-1.8	0.2361 (0.4759)	0.2687 (0.5049)	0.3326 (0.4851)	0.4739 (0.6025)	0.5936 (0.7165)	0.7481 (0.8524)

Table 4.3: Residual errors for Method 2 in mean(SD) format.

k_1 / σ^2	0	0.2	0.4	0.6	0.8	1
0	0.0025 (0.0016)	0.1526 (0.1198)	0.3297 (0.2589)	0.5066 (0.3979)	0.7312 (0.5524)	0.8521 (0.6454)
-0.2	0.1121 (0.0785)	0.1815 (0.147)	0.2912 (0.234)	0.4473 (0.3451)	0.575 (0.4306)	0.7237 (0.5438)
-0.4	0.4181 (0.2926)	0.4264 (0.3242)	0.4594 (0.3658)	0.5329 (0.4435)	0.6743 (0.5251)	0.7927 (0.6187)
-0.6	0.8721 (0.6176)	0.8904 (0.6345)	0.8861 (0.6883)	0.925 (0.724)	0.9884 (0.7998)	0.9729 (0.7942)
-0.8	1.4332 (1.0378)	1.4383 (1.0536)	1.4341 (1.0654)	1.4801 (1.096)	1.4469 (1.1413)	1.4627 (1.1953)
-1	2.0662 (1.5455)	2.0657 (1.5471)	2.0639 (1.5758)	2.0765 (1.614)	2.0903 (1.6185)	2.0906 (1.6943)
-1.2	2.7427 (2.1389)	2.7477 (2.1405)	2.7282 (2.1498)	2.7617 (2.2101)	2.7334 (2.1747)	2.7651 (2.2068)
-1.4	3.4406 (2.8204)	3.4373 (2.8374)	3.4266 (2.8337)	3.4349 (2.84)	3.4473 (2.8369)	3.4424 (2.8875)
-1.6	4.143 (3.5953)	4.1414 (3.5914)	4.1349 (3.5985)	4.1594 (3.6123)	4.1526 (3.629)	4.1353 (3.6114)
-1.8	4.8384 (4.47)	4.8398 (4.4729)	4.8404 (4.4859)	4.8372 (4.4832)	4.8528 (4.4594)	4.8579 (4.5038)
-2	5.336 (6.7399)	5.3365 (6.7389)	5.3557 (6.7436)	5.3478 (6.7374)	5.3728 (6.7714)	5.3571 (6.7408)

Table 4.4: Residual errors for Method 3 in mean(SD) format.

k_1 / σ^2	0	0.2	0.4	0.6	0.8	1
0	0.1908 (0.1231)	0.1919 (0.1468)	0.2644 (0.2165)	0.3791 (0.3083)	0.4921 (0.3912)	0.7031 (0.501)
-0.2	0.2701 (0.1806)	0.2847 (0.229)	0.3948 (0.3118)	0.5356 (0.4345)	0.6672 (0.5251)	0.8367 (0.6663)
-0.4	0.8474 (0.5649)	0.8419 (0.5923)	0.8771 (0.6841)	0.9321 (0.754)	1.0319 (0.8708)	1.1468 (0.9374)
-0.6	1.4005 (0.9605)	1.3625 (0.969)	1.3911 (1.0212)	1.3634 (1.0931)	1.38 (1.1461)	1.5862 (1.3061)
-0.8	1.5987 (1.2166)	1.5795 (1.2214)	1.5728 (1.2878)	1.4801 (1.3207)	1.6442 (1.4604)	1.6977 (1.4215)
-1	1.127 (1.2317)	1.1225 (1.2612)	1.1614 (1.281)	1.2035 (1.3069)	1.3363 (1.4139)	1.5164 (1.4935)
-1.2	0.583 (0.948)	0.6356 (0.9737)	0.7598 (1.0423)	0.9992 (1.0914)	1.2612 (1.2694)	1.4331 (1.327)
-1.4	3.0758 (1.2845)	3.0391 (1.3077)	3.0691 (1.3832)	3.2173 (1.5767)	3.3565 (1.7272)	3.4087 (1.8816)
-1.6	7.2372 (2.9217)	7.2639 (2.9396)	7.3176 (3.0208)	7.5021 (3.2005)	7.4749 (3.3129)	7.8262 (3.5051)
-1.8	13.2071 (5.4237)	13.2457 (5.4568)	13.2545 (5.4956)	13.3984 (5.599)	13.7115 (5.8024)	13.8121 (6.0262)
-2	19.2323 (8.4785)	19.2652 (8.5073)	19.4367 (8.6819)	19.522 (8.7784)	19.6985 (9.048)	19.9261 (9.2825)

Table 4.5: Residual errors for Method 4 in mean(SD) format. nc = no convergence.

k_1 / σ^2	0	0.2	0.4	0.6	0.8	1
0	0.0841 (0.0343)	55.4104 (9.1837)	nc	26.7411 (6.019)	128.5432 (15.6178)	nc
-0.2	1.4546 (0.5785)	1.7232 (1.0687)	2.5774 (1.8105)	3.4364 (2.4439)	5.3523 (3.7368)	6.3437 (4.433)
-0.4	2.6048 (1.0238)	2.478 (1.4911)	3.5198 (2.2741)	4.6648 (3.2089)	6.1555 (4.4794)	8.7572 (6.2686)
-0.6	3.7093 (1.4552)	3.6468 (2.0818)	4.0407 (2.7672)	5.8312 (4.0605)	nc	nc
-0.8	4.9811 (1.9633)	4.8992 (2.7498)	12.1224 (11.0923)	14.9092 (12.64)	12.6745 (9.6749)	38.743 (15.412)
-1	6.5494 (2.6018)	55.4478 (52.3379)	38.7723 (35.1774)	29.9265 (26.8411)	nc	nc
-1.2	8.422 (3.3727)	8.2952 (4.7627)	9.4199 (6.9322)	13.0092 (9.9833)	17.7393 (13.4921)	19.7825 (15.228)
-1.4	10.4553 (4.2195)	10.0689 (5.2757)	10.4045 (7.0633)	11.1341 (8.1748)	nc	nc
-1.6	12.3429 (5.0315)	11.7013 (5.5733)	10.587 (6.9299)	9.719 (6.9631)	14.3564 (10.7096)	15.7847 (12.1106)
-1.8	13.6184 (5.6291)	12.8806 (5.8626)	9.4004 (5.4976)	8.837 (6.6854)	10.8022 (8.1776)	10.7765 (8.2311)
-2	12.4561 (5.4377)	12.6685 (6.3563)	13.6445 (7.9549)	16.1739 (9.8614)	20.3563 (12.8011)	28.1754 (15.7858)

Discussion

The comparison of the proposed technique with the line based method, shows a very considerable improvement. The line based method fails to reach this performance despite using a high order model. This is mainly due to the fact the that no geometric relationship exists between the objective error and the distortion. The ill conditioning induced by the reverse model at high distortion levels hampers its performance further, all culminating in the tendency to lodge in local minima and in some cases fail to converge due to singularities in the estimation moment or co-factor matrix. The poor estimation performance can be additionally observed in the widely varying estimates for the centre of distortion. It is not suitable for larger distortion levels.

In comparison with the full calibration technique of Method 3, the proposed technique shows a significant performance increase. This method used the reverse model and hence fails to achieve comparable residual results. This algorithm is restricted to low levels of distortion as an increase in distortion model order would lead directly to increased sensitivity and ultimately poorer performance. The side-by-side analysis with the full calibration technique of Method 2, primarily indicates the proposed method achieves smaller residual error. This may be expected as both techniques use the forward distortion model, only to a higher order in Method 1. Despite the use of this higher order model, no adverse consequences arise from its use such as instabilities in the estimation problem. In fact, the comparisons on the recovered centers of distortion show that the proposed method achieves far superior stability in its estimation than Method 2.

The experiments conducted on simulated data clearly show that the proposed algorithm outperforms the comparison techniques on an accuracy front. Though this improvement in accuracy may be attributed to the use of an appropriate distortion model, the estimation algorithm still performs equally as well as the methods that require multiple input of data points. The simulations also show that the proposed algorithm at least matches the stability of full calibration methods, and surpasses them in many cases such as in the estimation of the distortion centre. In summary it is shown to be suitable for distortion levels of all kinds, with good invariance to control point location and noise.

Table 4.6: Real distortion samples, including the number of iterations for each method where available. (nc) = no convergence

Sample	Camera	Resolution	M. 1	M. 3	M. 4
1	Fuji + WL-FX9	2832 × 2128	11	11	18
2	Nikon E4500 + FC-E8 (FL 32mm)	2272 × 1704	8	15	35
3	Kodak Megaplus + 6mm Computar	1312 × 1032	16	11	24
4	Nikon E4500 + FC-E8 (FL 24mm)	2272 × 1704	9	20	27
5	Nikon E4500 + FC-E8 (FL 21.5mm)	2272 × 1704	9	71	25
6	Nikon E4500 + FC-E8 (FL 17.8mm)	2272 × 1704	13	nc	24
7	Nikon E4500 + FC-E8 (FL 14.6mm)	2272 × 1704	15	nc	140

4.2.3 Comparisons on Real images

Comparisons on real data are compiled with seven different samples of distortion, arranged in increasing order. Table 4.6 gives a brief description of the camera type and image resolution of each sample. Three images were input into the full calibration techniques of Methods 2 and 3, while the first of these was used in the single image methods. As an example, sample number 5 from this data set is shown in Fig. 4.9. One additional image is used for the residual distortion evaluation of all four methods. These residuals are compiled, with one exception, by undistorting the evaluation image and estimating a radially weighted homography (described in appendix A) on this data. The residuals for the line based Method 4 are compiled directly by undistorting the data. A radially weighted homography is then computed on these coordinates. This avoids the addition of extra inaccuracies through the use of an inverse approximation to undistort the image. All chessboard intersection points are estimated from an initial guess using the nonlinear line intersection method described in chapter 3 with a small local support in order to avoid the introduction of distortion bias.

The number of iterations required for each method to converge are also presented in table 4.6. This shows that there is a less computational overhead with the proposed method. It should also be noted that no optimisation is conducted within the Gauss-Newton method (used by Method 1), unlike the comparison methods which use a combination of a Gradient Decent and Gauss-Newton (LM) to speed up convergence.

Firstly, each method is examined for the residual distortion remaining after

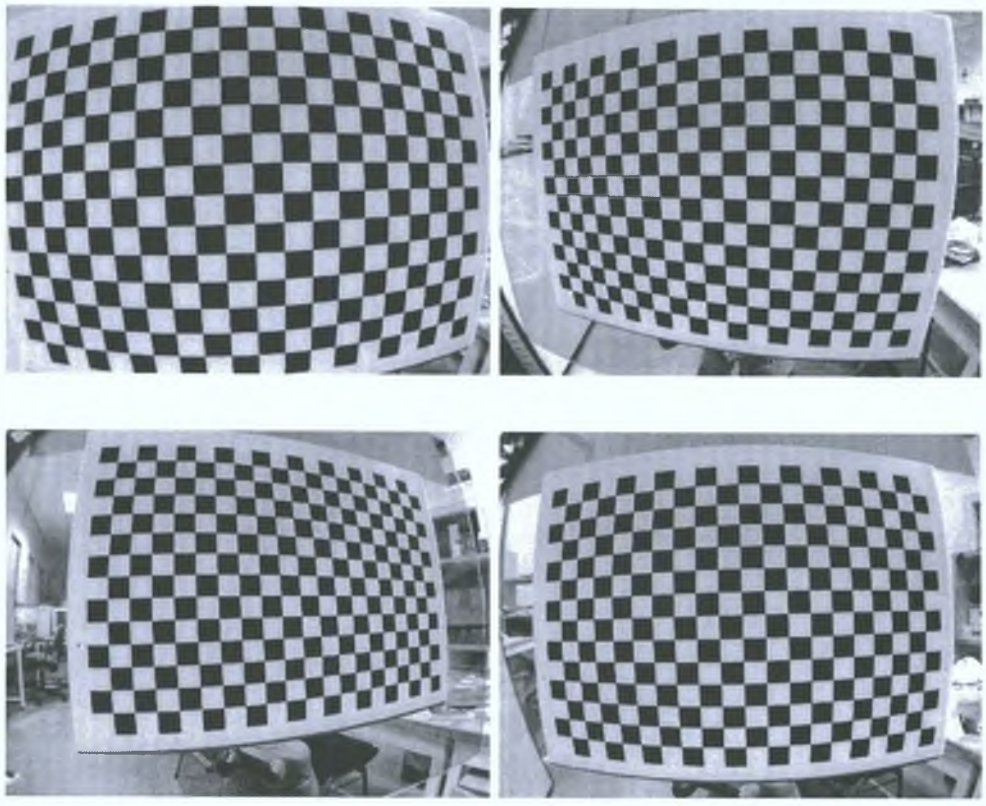


Fig. 4.9: Sample No. 5. Three calibration images and one evaluation image (bottom right).

Table 4.7: Distortion residuals on real examples (pix).

Sample	Method 1	Method 2	Method 3	Method 4
1	0.3164 (0.3362)	0.4088 (0.3402)	10.356 (8.6270)	0.3667 (0.3111)
2	0.4569 (0.3323)	0.5868 (0.9809)	0.9490 (0.7739)	11.8077 (8.0165)
3	0.2646 (0.2566)	0.2673 (0.2650)	0.4456 (0.4318)	0.3121 (0.3217)
4	0.5556 (0.7865)	1.4319 (0.8993)	16.4172 (8.7168)	3.9692 (2.7197)
5	0.8690 (0.6615)	2.4142 (1.5129)	21.724 (11.311)	0.8294 (2.0553)
6	1.8964 (1.6195)	9.3508 (7.7634)	nc	2.0041 (1.6612)
7	4.6403 (3.5282)	32.625 (26.973)	nc	5.9567 (7.3236)

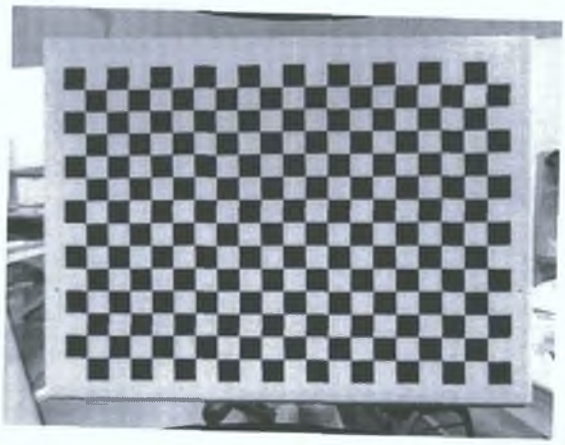
Table 4.8: Comparison of Method 1 and 2 both using the same distortion model (i.e two parameter radial model) (pix).

Sample	Method 1	Method 2
1	0.3734 (0.3200)	0.4088 (0.3402)
2	0.3578 (0.2839)	0.5868 (0.9809)
3	0.2629 (0.2836)	0.2673 (0.2650)
4	1.5406 (1.0859)	1.4319 (0.8993)
5	2.6919 (1.9254)	2.4142 (1.5129)
6	7.6891 (6.8891)	9.3508 (7.7634)
7	16.535 (13.799)	32.625 (26.973)

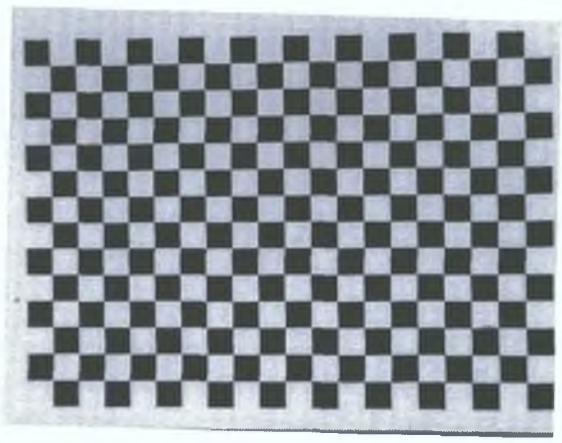
correction. These residuals are presented in Table 4.7 in mean (SD) format, with pixel units. This shows that the proposed method achieves a lower residual error than any of the comparison methods. Method 3 and 4 again show slightly erratic performances. As the improvement in performance of the proposed method may be attributable to the higher order model used, an extra comparison was conducted using a two parameter distortion model in Method 1. This is the same model as used in Method 2. The results are presented in table 4.8 for comparison with those of Method 2. These results show that the proposed method still achieves an overall lower residual error. This comes despite the fact that Method 2 uses three times the amount of input data than Method 1 in these experiments. The undistorted images for Sample 5 are shown in Fig. 4.10. A residual distortion is observable in the image undistorted with Method 3.

The values and uncertainties of each estimated parameter is presented for Methods 1-4 in Figs. 4.11 - 4.14 respectively. Method 4 shows as expected large parameter values and uncertainty for larger distortion levels. Method 3, using a two parameter version of the reverse model exhibits a similar trend, only with a lower magnitude. Method 2, using a two parameter version of the forward model, shows bounded values. Method 1 also shows bounded parameter magnitudes with small error bands. This confirms the results obtained on simulated data, and indicates the suitability of a high order forward model for distortion calibration.

The estimates of the distortion centre for each method are presented in table 4.9. These show that Method 1 recovers the centre point very close to that



(a) Method 1



(b) Method 2

(c) Method 3

Fig. 4.10: Corrected evaluation image of Sample No. 5.

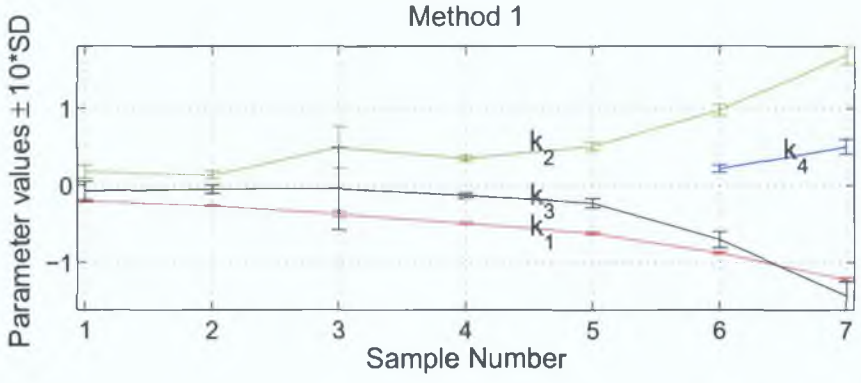


Fig. 4.11: Method 1 parameters with scaled 95 % uncertainty bound. k_4 was included to illustrate its bounded magnitude.

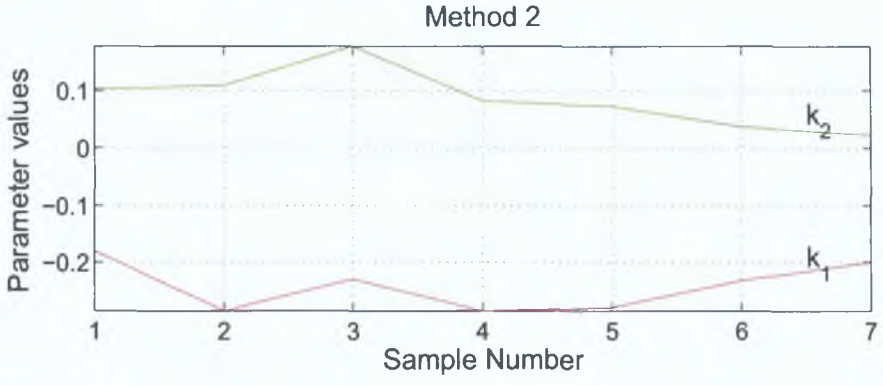


Fig. 4.12: Method 2 parameters.

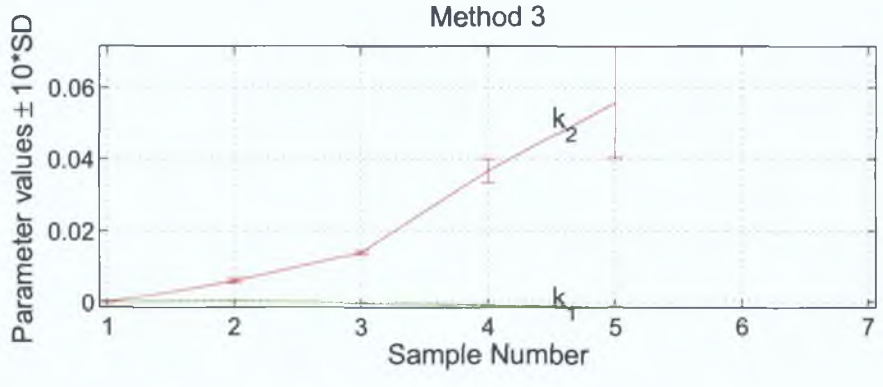


Fig. 4.13: Method 3 parameters with scaled 95 % uncertainty bound.

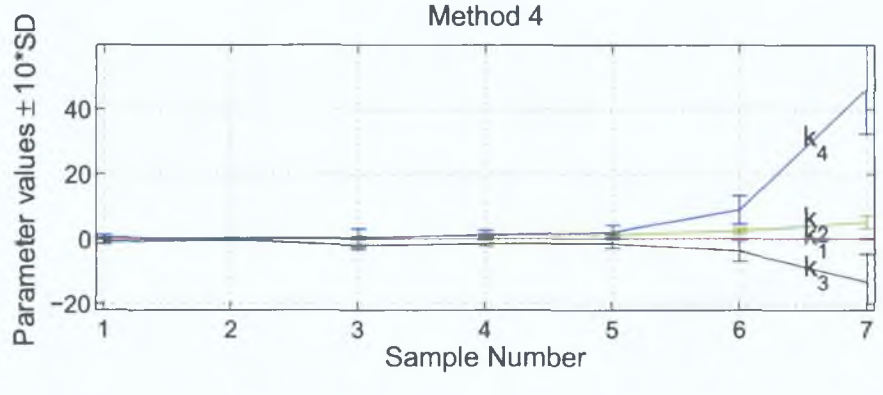


Fig. 4.14: Method 4 parameters with scaled 95 % uncertainty bound.

Table 4.9: Centre point estimates form each method (pix)

Sample		Method 1	Method 2	Method 3	Method 4
1	x	1374.7 (1.41)	1382.0	1378.2 (0.72)	1360.8 (3.21)
	y	1063.2 (1.01)	1062.5	1052.3 (1.18)	1065.3 (0.62)
2	x	1100.2 (0.70)	1107.6	1117.5 (0.91)	693.1 (25.1)
	y	887.8 (0.50)	890.3	883.0 (0.89)	910.8 (1.39)
3	x	618.3 (1.52)	618.2	615.6 (0.23)	617.3 (0.493)
	y	501.0 (0.58)	501.8	502.6 (0.22)	503.1 (0.23)
4	x	1113.3 (0.43)	1115.7	1148.4 (2.75)	1193.3 (9.06)
	y	879.7 (0.31)	881.8	864.8 (3.29)	883.4 (0.48)
5	x	1114.3 (0.41)	1115.4	1101.6 (3.45)	1117.4 (5.39)
	y	878.1 (0.29)	878.2	852.8 (4.27)	878.5 (0.28)
6	x	1117.6 (0.37)	1134.3	nc	1129.7 (3.75)
	y	873.2 (0.26)	870.1	nc	871.4 (0.28)
7	x	1120.3 (0.34)	1093.4	nc	1123.5 (1.81)
	y	873.1 (0.26)	861.4	nc	873.2 (0.21)

of the full calibration Method 2. This indicates a close relationship between the centre point of distortion and the principal point (as recovered by Method 2). This relationship is examined further in Section 4.4.1. Method 3 recovers a similar principal point except in the samples where the distortion is poorly modelled, e.g. sample no. 5. Method 4 shows much more volatile centre point estimation, caused by the instability that afflicts this method. On the contrary, Method 1, also employing only one input image, consistently recovers a robust distortion centre, as demonstrated by the small uncertainties associated with these estimates.

Discussion

Following the results obtained with simulated data, the experiments with real images aim to confirm these findings. In this sense Method 4 is again shown to have erratic accuracy and a tendency for high uncertainty in its parameters. In comparison, Method 1 significantly improves upon the performance of Method 3, from a distortion removal point of view. As in the simulated case there is also a clear improvement in the accuracy of Method 1 in comparison with Method 2.

It was slightly unclear if the proposed algorithm would match the performance of Method 2 using a lower order distortion model. This was investigated with the comparison of both methods using the same distortion model. It revealed that on average Method 1 achieved a slightly lower distortion residual. This indicates the robustness of the proposed approach, considering that Method 2 uses, in this case, three times the amount of input data. The centre point is reliably estimated, with very similar locations to that of the full calibration techniques that require more than one image for its recovery. However, this is not the case for the line based Method 4, which shows unreliable centre point estimates.

In conclusion, the experiments with real and simulated data clearly demonstrate that the proposed method outperforms all comparison methods in terms of accuracy in calibrating and removing distortion. This level of accuracy is achieved while using less input data, requiring only one view. The parameter estimates are shown to be reliable, and lend themselves to a well conditioned problem. Additionally, the real experiments show that there is less computational overhead than the comparison methods. These factors in combination with the more accessible closed form solution, appropriate distortion modelling and unique minimisation of pixel distortions in resampling, make this method a highly suitable non-metric method for removing lens distortion of all levels in perspective cameras.

4.2.4 Low distortion lenses

As seen in the experiments with simulated data, the accuracy of the proposed algorithm is weakest for very low distortion levels with noisy control points. The performance of the algorithm is now evaluated with a selection of large focal length lenses to assess both the levels of distortion in these lenses and the behaviour of the proposed technique in such circumstances.

The unknown affine transform $\mathbf{A}$ may be equated with the scaled internal camera parameter matrix (assuming zero skew), where a minimum of two views are required to solve for three unknowns (Zhang, 1998). It is important therefore to highlight that these parameters are solely dependent on the distortion present in the image, where $\mathbf{t}$ is the apparent centre of distortion and bears no direct relationship with the principal point. Any frame grabber stretching

Table 4.10: Low distortion lens details. All are C-Mount lenses and are attached to a Kodak MegaPlus digital camera.

Sample ID	Make	Focal Length	Distortion
a	Cosmicar	25mm	Pincushion
b	Computar	25mm	Barrel
c	Computar	50mm	Pincushion
d	Computar	50mm	Barrel
e	Computar	55mm Tele	Pincushion

of the x or y coordinate pixels, or non-square pixels are accounted for by the variable s.

Five lenses are tested for, with a typical distortion of less than one pixel. These lenses are described in table 4.10 with the accompanying type of distortion. They are all C-Mount lenses and are attached in turn to a Kodak MegaPlus digital camera. Due to the mounting, it is unlikely that the centre of distortion resides near the centre of the image array. Each image is calibrated using the proposed method, with a three parameter model for distortion. The distortion residuals before and after calibration are presented in table 4.11. This shows that there is an appreciable reduction in lens nonlinearities following calibration. A selection of the distortion residuals before and after calibration are shown in Figs. 4.15 and 4.16. These results show that the proposed technique does indeed correctly converge to model distortion, even to very small levels.

The estimated variance in the recovered distortion centers are presented in table 4.12. As expected due to the very low distortion levels, its location has an associated increase in uncertainty. However, overall these levels are low in relation to the image size (1312×1032).

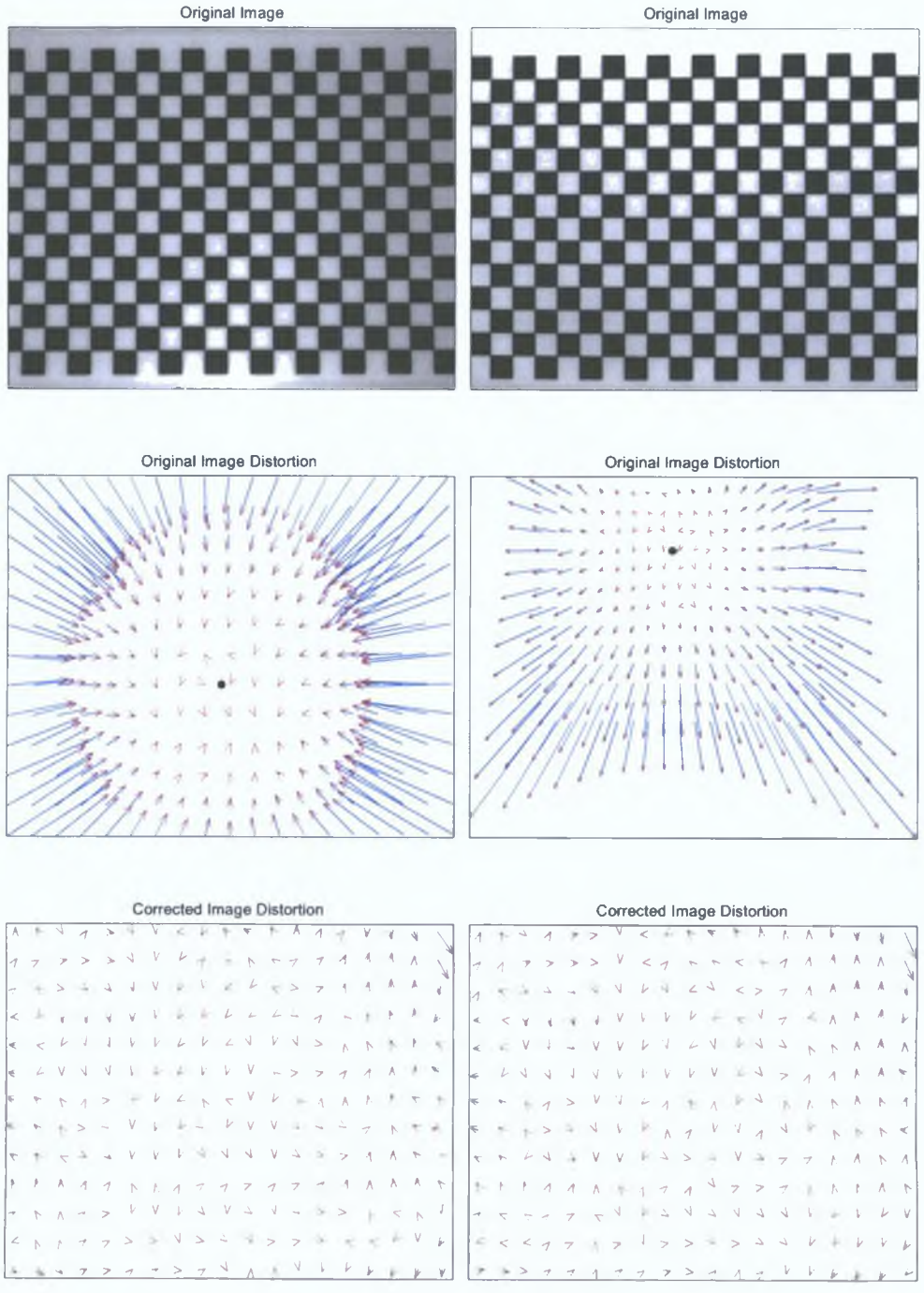


Fig. 4.15: Left column: Sample a. Right column Sample b. Both are 25 mm lenses. Sample a contains pincushion distortion, while sample b contains barrel distortion. Black dot represents the estimated distortion centre. The vector fields of distortion are scaled $\times 50$.

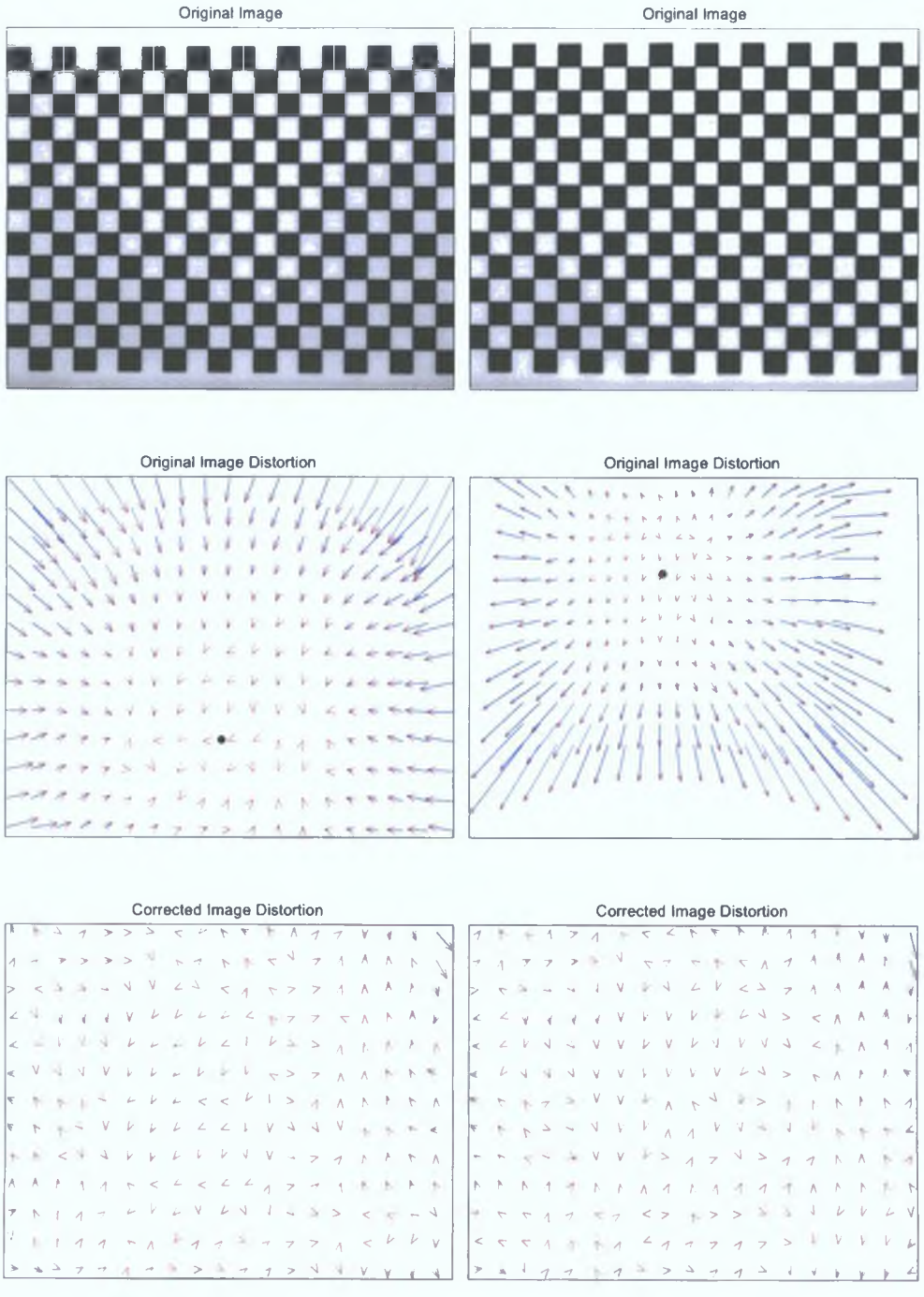


Fig. 4.16: Left column: Sample c. Right column Sample d. Both are 50 mm lenses. Sample c contains pincushion distortion, while sample d contains barrel distortion. Black dot represents the estimated distortion centre. The vector fields of distortion are scaled $\times 50$.

Table 4.11: Residuals before and after calibration for low distortion samples (pix)

Sample	Before Calibration		After Calibration	
	Mean	SD	Mean	SD
a	1.0577	0.9685	0.0823	0.0712
b	1.1208	1.0975	0.0837	0.0742
c	0.4358	0.3898	0.0819	0.0711
d	0.8715	0.8313	0.0814	0.0707
e	0.6099	0.6252	0.0786	0.0698

Table 4.12: Estimated centre point uncertainties (SD) (pix)

Coordinate	a	b	c	d	e
x	4.7062	6.2169	11.1108	7.7288	6.6478
y	3.6325	5.1892	15.0059	6.0099	5.9465

The proposed algorithm is based on the premise that at least some level of distortion is present in an image. This is generally easily satisfied using normal general purpose lenses. The behaviour of the technique with very low distortion lenses is thus investigated to determine its stability under such conditions. It is shown for a selection of low distortion lenses, roughly in the sub-pixel range, the algorithm successfully models the distortion which in turn improves the linearity of the image following distortion compensation.

4.3 Dependence on control points

The proposed algorithm is based on some assumptions on the geometry of the calibration target and the reliable detection of this target in an image. This section investigates these constraints, to assess the impact of their non compliance on the calibration and removal of distortion in images.

The issue of bias free control point detection in an image has largely been addressed in Chapter 3, while the robustness of the proposed distortion calibration to random error in pixel coordinates has been investigated in Section 4.2.2. This section now examines the necessary quantity of control points required for successful calibration. It is subsequently shown that this quantity is directly related to a multi-image input variation of the proposed algorithm.

The constraints on the calibration target are that of planarity and precision in the chessboard pattern. For general usage, the calibration pattern might only be manufacturable to a specific precision, using (for example) a laser printer, which can result in the introduction of errors. The dependence of the calibration technique on these errors is investigated. It is also possible that in general usage, the pattern may not be exactly planar. This results in a systematic type error, for which the errors resulting from the proposed technique are quantified. Finally, the mis-calibration resulting from bias inducing patterns such as circular type features is highlighted.

4.3.1 Number of control points required

In Section 4.1.1 it was shown that a minimum of m control points are required to solve the system of equations, where $m = \text{ceil}(n_\Phi/2)$ and n_Φ is the length of the parameter vector. For the basic radial distortion model with two parameters at total of eight control points are required, where each control point yields two constraints. According to the principal of Maximum-Likelihood there is an exponential relationship between the convergence to the true solution and the quantity of calibration data, assuming normally distributed data errors. The aim now is to outline the number of control points required to converge sufficiently close to the optimum solution, under varying noise conditions. It has previously been shown in Section 4.2.2 that the algorithm is sufficiently invariant to the location of these control points.

This investigation is primarily conducted with simulated data, where lens distortion is simulated as in Section 4.2.2. The number of control points n are varied from 8 up to 500. Their locations are chosen from a uniform random distribution covering the entire image window, while the residuals are compiled over 100 independent trials of these locations. The residuals are computed with the usual regular grid covering the image window. Only samples where the algorithm has converged are included, the majority of these non-convergence situations naturally occurring with low numbers of control points. As shown in Section 3.3.5 the expected control point noise is less than $\sigma = 0.15$ (pix) with a normal distribution. The control points are thus corrupted with normal noise with excessive standard deviations of $\sigma = 0.15$, 0.3 and 0.5 pixels to estimate a worst case senecio.

Fig. 4.17 shows the error convergence with increasing number of control points for three levels of distortion. In all these examples, there is a rapid increase in performance up to 100 control points, at which point the error begins to level off. From 200 onwards the improvement in error is negligible. The level of distortion primarily effects the final error level, but does not drastically alter the convergence shape. The increase in noise levels naturally slows the rate of convergence, but again becomes insignificant after 200 control points. In this light, the calibration pattern used in this work yields 247 control points which in conjunction with noise levels lower than $\sigma = 0.15$, provides ample data for correct convergence.

Multiple Input Images

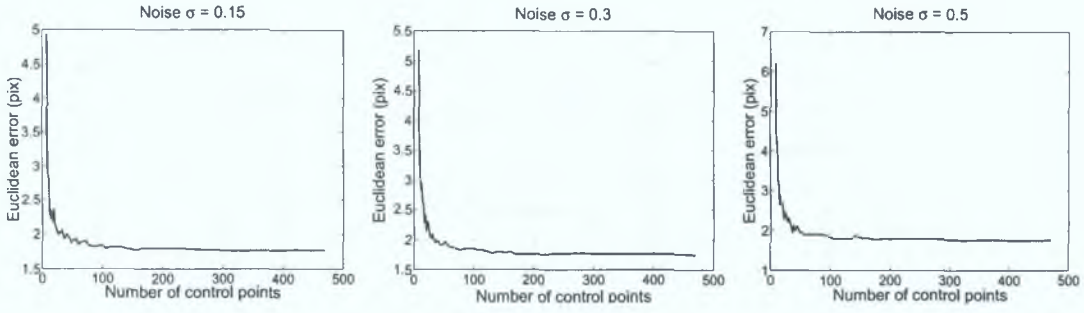
An alternate means of increasing the quantity of data, without increasing the density of control points on the calibration pattern, is to use multiple input images. Considering q images with n control points each, the objective error function of equation 4.4 can be modified as:

$$\begin{aligned} &\min \sum_{j=1}^{q} \sum_{i=1}^{n} \mathbf{e}_{ij}(\check{\mathbf{c}}_{ij}, \hat{\mathbf{\Phi}}_j)^2 \quad \text{with} \\ &\mathbf{e}_{ij}(\check{\mathbf{c}}_{ij}, \hat{\mathbf{\Phi}}_j) = \mathbf{H}_{\mathbf{e}j}\bar{\mathbf{c}}_{ij} + \mathcal{D}(\mathbf{H}_{\mathbf{e}j}\bar{\mathbf{c}}_{ij}, \mathbf{k}) - \mathbf{A}\check{\mathbf{c}}_{ij}, \end{aligned} \tag{4.19}$$

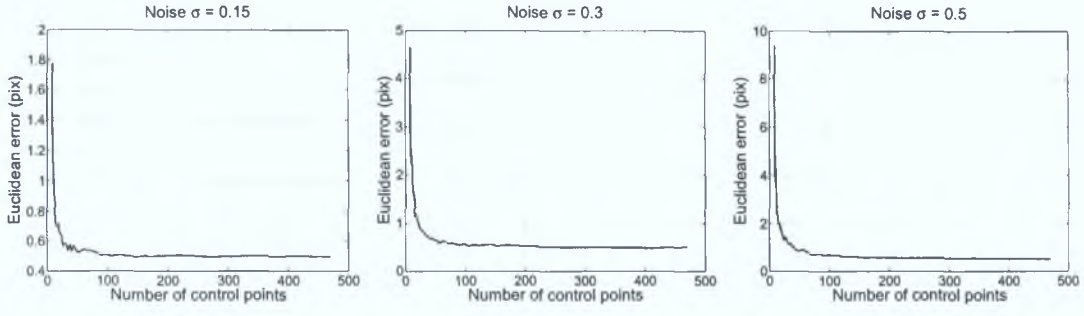
where for every additional image, the dimensions of the parameter space increases by eight. A solution to avoid this increase was sought by finding the relationship between the entries in $\mathbf{H_e}$ with $\mathbf{H_d}$ and the distortion function. However, an analytical solution was not found, nor practical, due to the poor estimation of $\mathbf{H_d}$.

The effective equivalence of using fewer control points on multiple images, and the use of single images of higher point density, is demonstrated by experiments on real images. Seven input images, each with 35 control points [3] are compared to the performance of one planar view (image number 1 of 7) with the full complement of 247 control points. For evaluation purposes the noise in the detected control points (maximum expected SD= 0.15) is amplified by adding Gaussian noise of SD= 0.55 pixels. The residual errors are measured with a separate image using the full complement of 247 control points and the usual

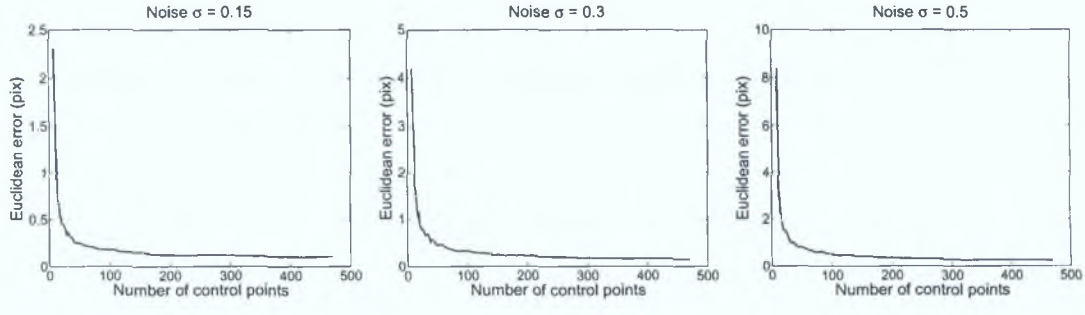
[3] The subset of 35 control points are taken from the 19×13 chessboard pattern by taking every third point.



(a) Distortion=-0.5



(b) Distortion=-0.3



(c) Distortion=-0.1

Fig. 4.17: Convergence properties of proposed algorithm with respect to the number of control points.

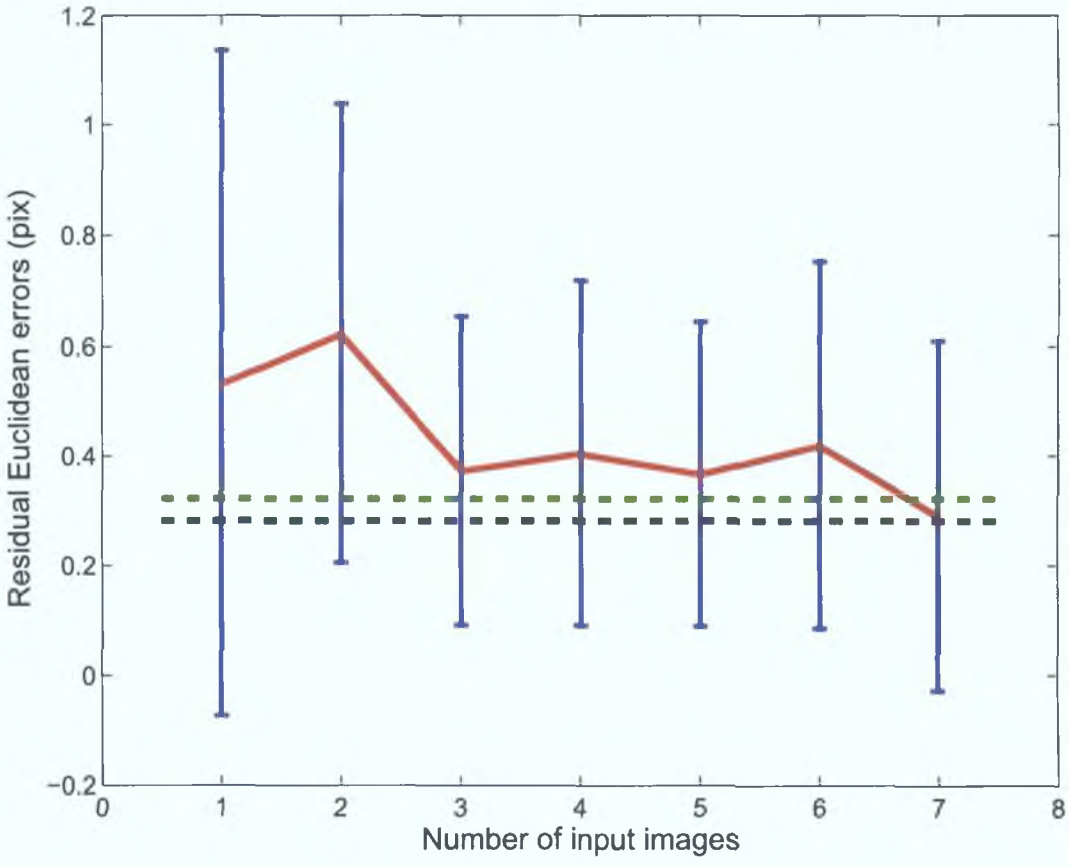


Fig. 4.18: Errors for multiple input images with 35 control points per image, including standard deviation error bands (red trace). Noise was synthetically added to real point coordinates to amplify the errors for comparison. The green reference line shows the mean error from a single input image with 247 control points with noise added. The black reference line represents the noise free solution error.

weighted homography approximation (Appendix A). Fig. 4.18 shows the drop in error (Red trace) and SD as the point count and image number increases. Upon reaching 245 control points (from 7 images) the error converges to the single image level (Green line), evaluated with 247 points. Due to the amplified noise content, this level is slightly larger than the solution obtained without the additional noise (Black line). This examination demonstrates that the addition of extra input images gives an equivalent performance to a similar increase in the quantity of control points in a single view, save for the enlarged parameter space dimensions.

4.3.2 Errors in calibration target

In the above sections, and in chapter 3, it is assumed that the calibration pattern is both perfectly planar, with a very high spatial accuracy for the chessboard intersections. The calibration model is then simply specified up to scale as an equally spaced matrix of points. To best satisfy these requirements a chessboard pattern is manufactured using an high precision xy table originally designed for PCB board manufacture. Black and white layers are formed with a very thin black vinyl sheet attached to a sheet (1.5mm) of white PVC. Using a knife attachment with the xy table, a series of horizontal and vertical slits are made on the black vinyl. The contrast pattern is formed by removing alternative squares of the black vinyl to reveal the white PVC backing. Finally, the PVC sheet is adhered to a glass backing plate to ensure planarity. For evaluation purposes, such high precision is necessary,

For general usage however, it is significantly easier to use a conventional printer to print the chessboard pattern, and fix it to a wall or table. Most printers have a habit of not sucking the paper in perfectly straight or at exactly the same velocity (worn roller slippage, friction resistance), resulting in non co-linearity of the chessboard rows and columns. Considering these errors as random deviations from the assumed model, knowledge of the expected performance of the algorithm in such circumstances is useful. Although sheet glass is cheap and readily available, it is useful to assess the performance of the algorithm where the planarity constraint is not fully met. This is examined for a slightly cylindrical calibration target.

Random noise in calibration pattern

This experiment is carried out using real data. Model imprecision is simulated by adding zero mean Gaussian noise with standard deviation ranging from 0.01 to 1mm. The size of the each calibration square is 14×14 mm so this signifies a considerable error. For each noise level 100 trials were conducted, and the average and SD values for the radial distortion parameters are computed. These variations are shown in Fig. 4.19. Naturally, the variation of the distortion parameters increases with increasing noise. The corresponding mean Euclidean and SD errors are shown in Fig. 4.20. These results show that there is roughly a one-to-one transfer in error. Note that it is possible

to improve the performance if considerable random errors are expected in the calibration target, using a Total Least Square (TLS), or Error in Variables (EIV) technique such as Lavest et al. (1998).

Systematic errors in calibration pattern

Non-planarity of the calibration target is now considered. Since the calibration target is a sheet, e.g. printed on a sheet of hard paper, it has a natural tendency to bend along either its horizontal or vertical axis, giving rise to a cylindrical type shape. It is unusual to encounter significant bending in both directions as this causes the material to kink. A cylindrical shape can be simulated by displacing the z coordinate of the calibration model as $\bar{\mathbf{w}} = (\bar{x}, \bar{y}, 1 + p\bar{x}^2)^T$, where p dictates the degree of bending. The extent to which this displacement is picked up by the camera depends on its distance from the target and the focal length, e.g. a camera very far away with a large focal length may be considered as having parallel projection where the z displacements are invisible, while the opposite situation is encountered with a fish-eye lens. This experiment is conducted as before with a wide angle lens at approximately 200mm from the 250 × 170 mm calibration chart and a pixel resolution of 2830 × 2128. The cylindrical distortion observed for varying levels of p is then used in the distortion calibration. Fig. 4.21 shows the induced error from this bending, measured as the maximum deviation from the planar position. Roughly, a one-to-one proportional increase in error is again observed. Bending in excess of 1mm is generally visible to the eye, and may be alleviated by the user. The planarity induced errors in lower focal length lenses will be lower than this level.

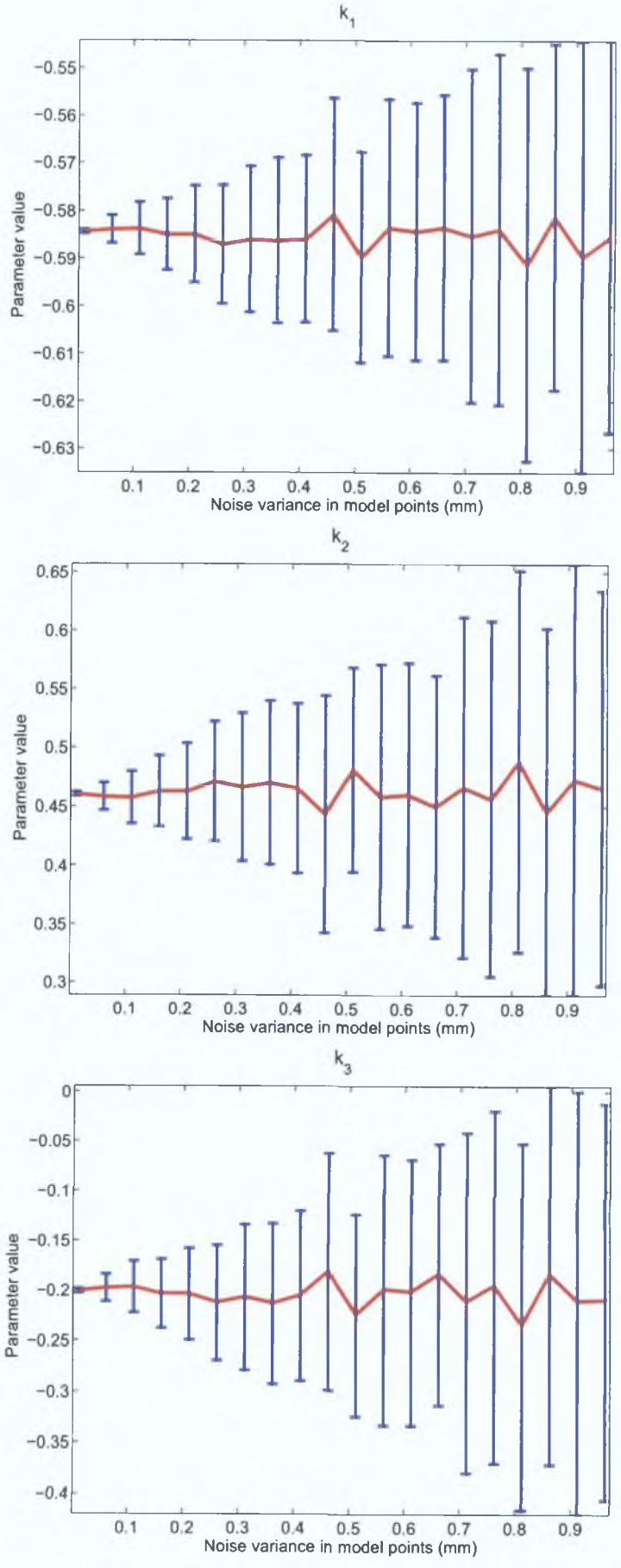


Fig. 4.19: The estimated parameter mean (Red trace) and SD bounds for random noise in the calibration target.

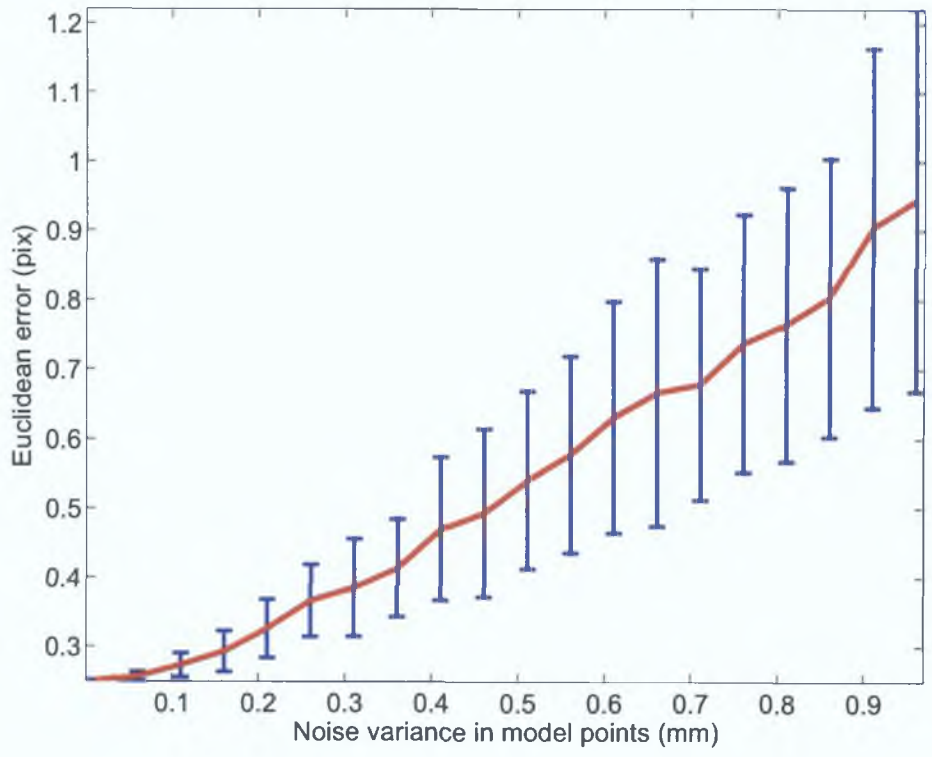


Fig. 4.20: The mean Euclidean and SD errors for increasing noise in the calibration pattern points.

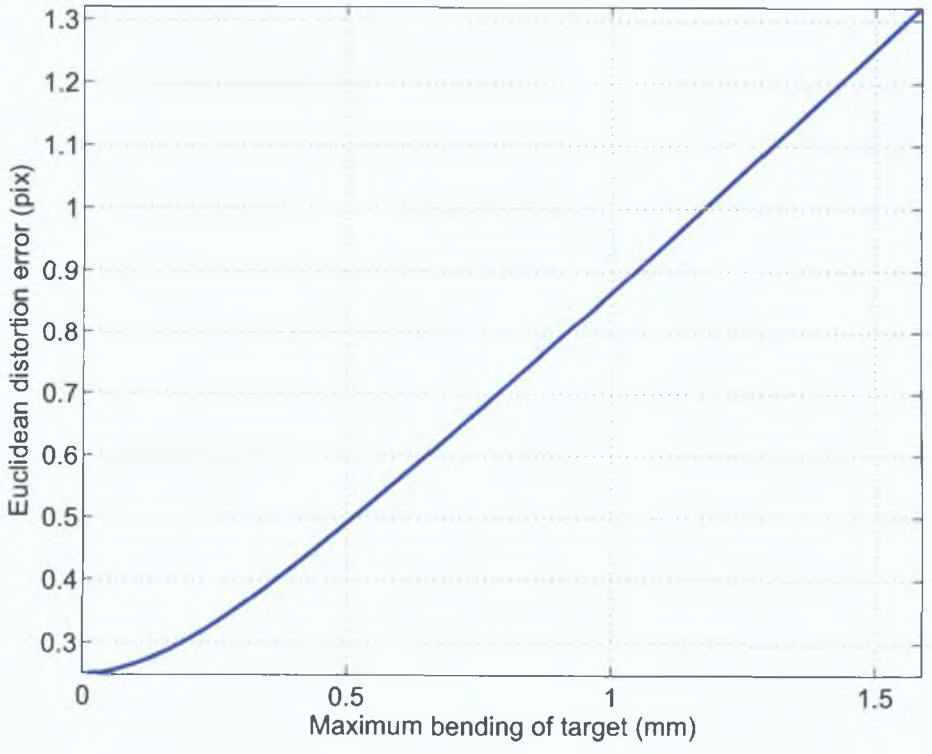


Fig. 4.21: The mean Euclidean errors for increasing cylindrical bending of the calibration chart.

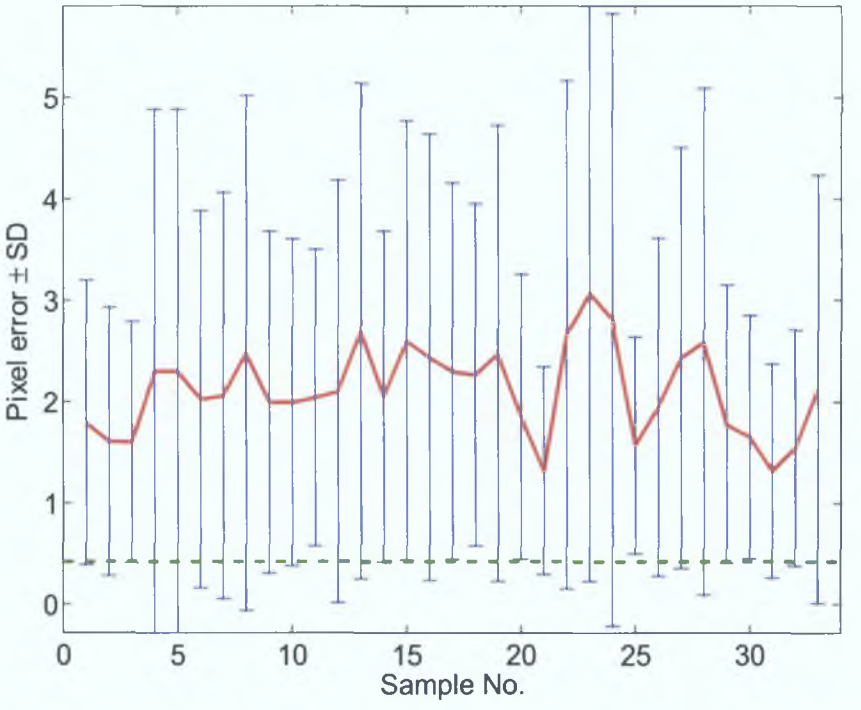


Fig. 4.22: The mean and SD errors resulting from the calibration of distortion with multiple samples of a laser printed pattern. This is compared to the mean level(0.42 pix) achieved using the precisely manufactured pattern (green line).

Low accuracy printed patterns

As a high quality calibration pattern is not always available, it is useful to investigate how well the calibration can be achieved using a pattern printed on an A4 sheet using a standard office laser printer[4]. To this end, the distortion calibration for 33 separate printed patterns is compared to that obtained with the high accuracy pattern. The relative positions of the camera is fixed throughout. A Fiji FinePix 6900 is used with a wide angle lens (displaying mild distortion) and a pixel resolution of 2832×2128. The intersections of the chessboards are extracted using the saddle point refinement method. This data is then used in the calibration routine, from which the residuals are compiled using the radially weighted homography. The results are shown in Fig. 4.22 in comparison with the mean error obtained using the high accuracy pattern. It can be seen that an average error of over 2.1 pixels is incurred if low quality printed patterns are used. Although these results are linked to the particular camera and indeed laser printer used, it gives a strong indication of the expected calibration accuracy for printed patterns.

[4]Laser printer in this experiment is HP LaserJet 6mp, feeding standard 80 g/m^2 paper.

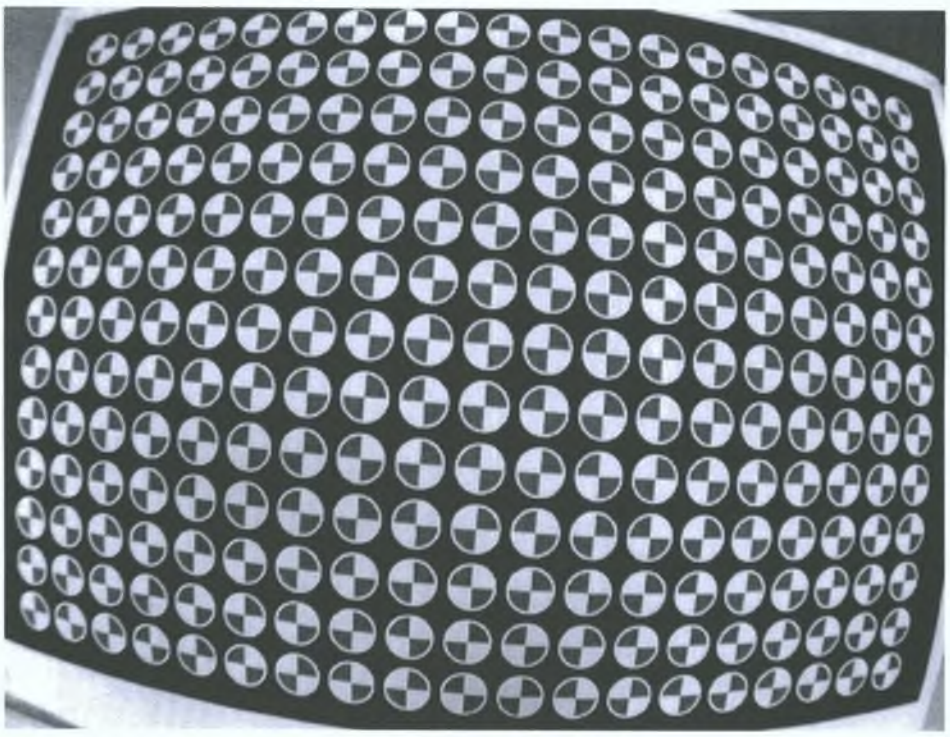


Fig. 4.23: Input image used to compare distortion calibration with circular control points and square intersection control points.

4.3.3 Biasing influence of circular calibration patterns

The calibration of distortion with circular control points is compared with that of square chessboard type control points to quantify the influence distortion bias has on the removal of lens distortion. The hybrid pattern is again used as shown in Fig. 4.23, from which the centroids and chessboard intersections are extracted. Following calibration, the estimated parameters are used to correct an image of the standard chessboard pattern. These points are then approximated using a radially weighted homography, from which the residual distortion is estimated. These residuals are shown in Fig. 4.24 for circular correction, and in Fig. 4.25 for the square correction. These vector plots clearly show the underestimation of radial distortion using circular control points. The mean and standard deviation of the circular and square residuals are 0.8939(0.6993) and 0.3996(0.3101) respectively. The estimated parameters for radial distortion are $\mathbf{k}_c = (-0.5773, 0.4463, -0.1941)^T$ and $\mathbf{k}_s = (-0.5808, 0.4491 - 0.1937)^T$ for circular and square type control points respectively.

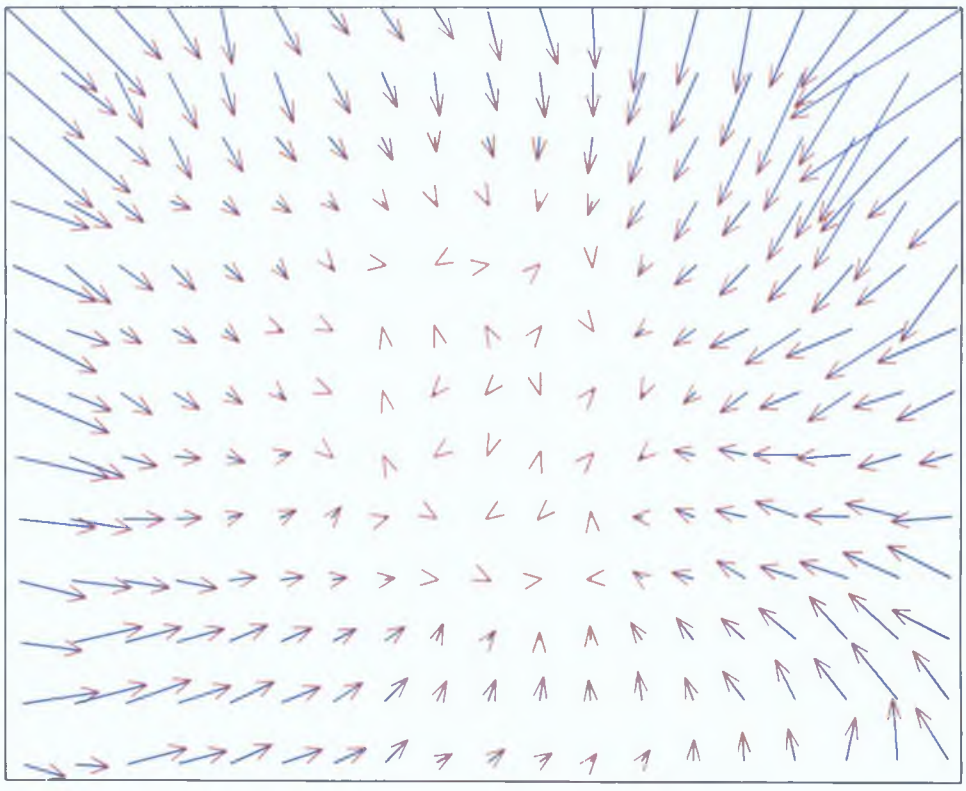


Fig. 4.24: Residual distortion following correction with parameters estimated using circular control points of Fig. 4.23

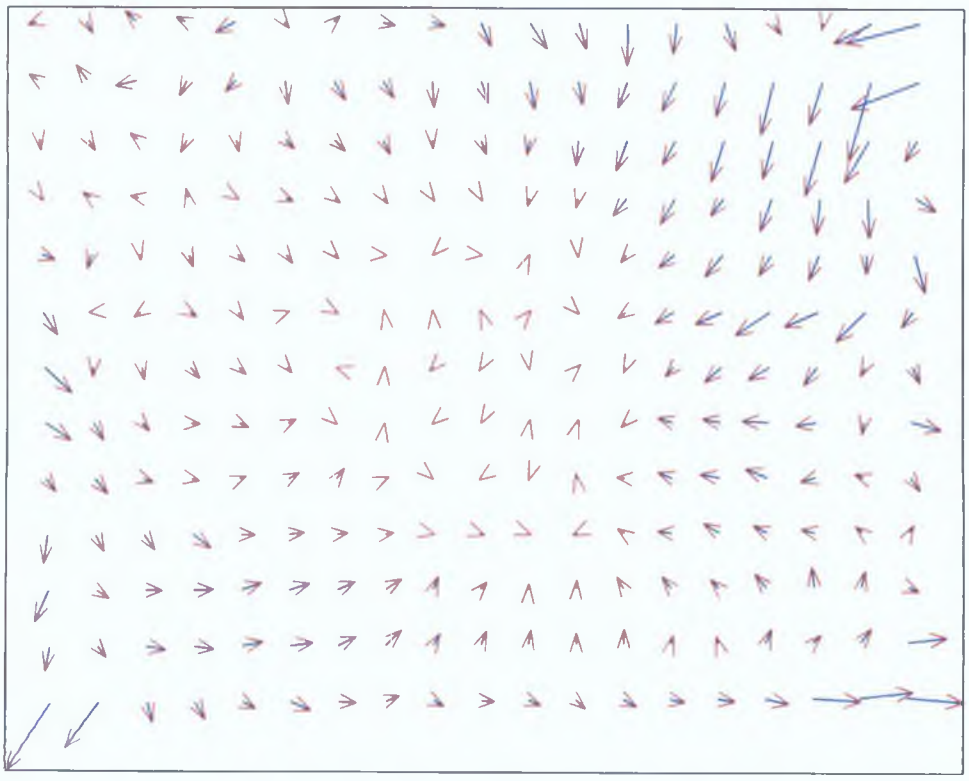


Fig. 4.25: Residual distortion following correction with parameters estimated using square intersection control points of Fig. 4.23.

4.3.4 Discussion

This section investigates, among other things, the degree to which the assumptions on the geometry of the calibration chart must be satisfied. These assumptions relate to the planarity of the calibration target and known relative 2D coordinates of the pattern up to an arbitrary scale. Errors in these 2D locations are simulated by adding random noise of varying degrees. The results, evaluated on a high resolution sensor, indicate that there is roughly a direct transfer from these errors, in millimeters, to distortion errors in pixels. Systematic errors such as non-planarity are investigated, with the results indicating that the algorithm performs similarly to the pattern imprecisions, with 1mm bending inducing a little less than a mean 1 pixel error. Lastly, the miscalibration of distortion from circular control points corrupted with distortion bias is investigated. This demonstrates for the example considered, that there is a considerable induced error compared with the square based control point calibration. In the light of other error sources, it equates to roughly 1mm bending or 1mm of imprecision in the control pattern. Such a pattern would represent a fairly poorly manufactured calibration target.

Additionally, this section also demonstrates that there is no advantage gained by using multiple images for distortion calibration over single image calibration, assuming roughly equal numbers of control points in each data set. A sufficient data set size has been identified as containing roughly 200 or more control points, considering normal levels of noise. Thus, single view calibration with a sufficiently sized data set, and accurate bias free detection, will lead to optimal distortion calibration.

In summary, the following observations can be made regarding the practical implementation of this algorithm:

- To reach an optimal estimation for distortion, the minimum number of control points required is roughly in the region of 200. This is largely independent of the level of distortion, and does not assume very precisely detected control points, (up to the region of SD=0.5 pixels).
- Multiple images may be used to generate extra control points but they offer no further advantages.
- The requirements for precision in the 2D pattern and its planarity are

roughly equal, with an approximate one-to-one transfer between millimeter inaccuracies and pixel errors.

- Circular control points, susceptible to distortion bias, lead to a significant underestimation for distortion, equivalent to a very poorly manufactured chart.

4.4 Decentering Distortion

Thus far all experimentation was carried out using a model that excludes decentering distortion. This facilitated the side by side comparison with alternative methods that in the majority do not consider its inclusion. Also, since radial distortion is clearly the dominant distortion type, the performance of the algorithm was investigated with respect to it only.

The remaining residuals following the inclusion of decentering distortion model, as presented in Section 2.1, into the distortion calibration equations are now investigated. The comparisons are made on real lenses (a selection of those in Section 4.2.3, table 4.6) and compared with the performance of the decentering free models in Section 4.2.3. These results are presented in table 4.13, showing that, surprisingly, there is no overall improvement in the residuals. Investigating this further for a low cost lens that is known to contain misaligned elements, i.e. sample 3, the residual vector fields after calibration with and without decentering are shown in Fig. 4.26. This shows that vector field including decentering distortion displays a larger decentering type residual field than the vector field that does not consider its modelling. These results at first glance are at odds with the expected behaviour.

An investigation of the decentering parameter values p_1 and p_2, reveals that they have large associated uncertainties, in some cases the uncertainty is greater than the actual value. The values and uncertainties are presented in table 4.14 for the samples considered. The main assumption for calibrating distortion is that the functional model correctly models the real distortion effects. Since the addition of decentering distortion does not serve to reduce the residuals, it may be assumed that the decentering distortion is either not there or has been accounted for by a combination of other parameters. Considering that modern lenses are mass produced, cheap and of small format, it is unlikely

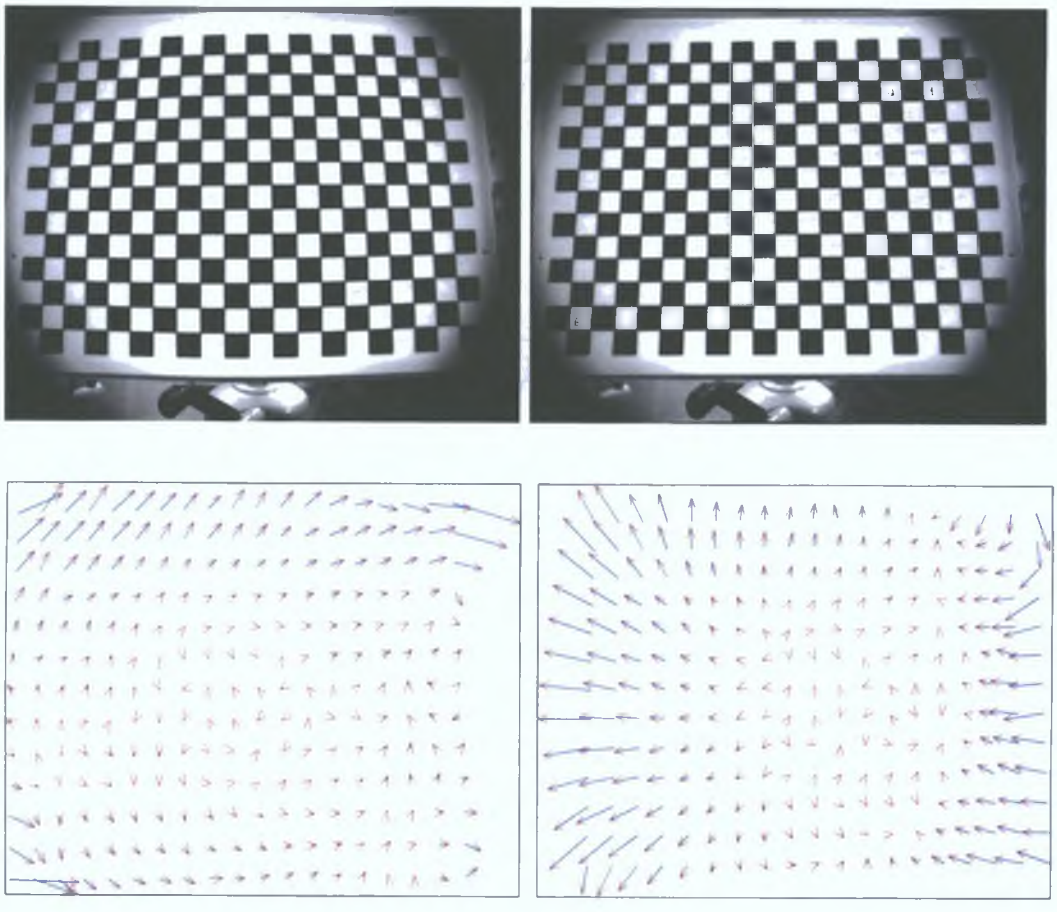


Fig. 4.26: Residuals for sample number 3. Images show uncorrected and corrected images. Left vector plot depicts the residuals following distortion removal without the explicit consideration of decentering distortion. Right vector plot show the residuals remaining following calibration with the inclusion of decentering distortion.

Table 4.13: Comparison of remaining distortion residuals following calibration with and without the explicit inclusion of decentering distortion models. Residuals without decentering are taken directly from table 4.7.

Sample no.	Without decentering	With decentering
1	0.3164 (0.3362)	0.3264 (0.3225)
2	0.4569 (0.3323)	0.4158 (0.3636)
3	0.2646 (0.2566)	0.3194 (0.2921)
4	0.5556 (0.7865)	0.5029 (0.4886)
5	0.8690 (0.6615)	0.8795 (0.6602)

Table 4.14: Decentering parameter values and associated uncertainties for results on real images.

Sample	$p_1 \times 10^{-4}$	SD $p_1 \times 10^{-4}$	$p_2 \times 10^{-4}$	SD $p_2 \times 10^{-4}$
1	-4.1567	1.4226	7.3960	1.8937
2	1.0257	2.1559	8.8403	3.7863
3	-31.1631	1.2181	3.6948	1.2933
4	11.6121	1.1331	-10.124	2.1927
5	0.3762	1.2660	5.4786	1.7113

that decentering distortion is not a factor. Therefore, this distortion must be accounted for by some combination of the other parameters.

It is very useful to look at the correlations between the estimated parameters. These correlations may be obtained directly from the covariance matrix resulting from the iterated estimation procedure upon convergence. Again considering the Fisher information matrix in equation 4.17, an approximate correlation coefficient between the ith and kth parameters is given by:

$$c_{ik} = \frac{(\mathbf{F}^{-1})_{ik}}{(\mathbf{F}^{-1})_{ii}^{1/2}(\mathbf{F}^{-1})_{kk}^{1/2}}, \quad \text{where} \quad -1 \leq c_{ik} \leq 1. \tag{4.20}$$

Decentering distortion is historically linked with the estimation of the principal point or the intersection of the optical axis with the image surface. The correlation coefficients also indicate a strong link between these parameters. The correlation between the decentering parameters and the centre point for a typical example, i.e. sample 3, are $p_1 \sim 0.8379 \sim c_x$ and $p_2 \sim 0.9536 \sim c_y$. Indeed it has been shown by Stein (1993) that a variable principal point induces decentering like elements through the standard radial distortion model.

Considering that the error homography $\mathbf{H_e}$ may, at least, introduce a variable

centre point by the simple adjustment of $h_3 \rightarrow h_3 + \Delta_x$ and $h_6 \rightarrow h_6 + \Delta_y$, resulting in the following:

Result 5. *A translation of the undistorted points induces decentering distortion and a small residual through the radial distortion functional:*

$$\mathcal{D}(\mathbf{p}+\boldsymbol{\Delta},\mathbf{k}) = \mathcal{D}(\mathbf{p},\mathbf{k}) + \begin{pmatrix} k_1\Delta_x(3x^2+y^2) + 2k_2\Delta_y xy + \cdots \\ 2k_1\Delta_x xy + k_2\Delta_y(3y^2+x^2) + \cdots \end{pmatrix} + E$$

Proof. Considering the formulation of distortion as $\mathcal{D}(\mathbf{H_e}\bar{\mathbf{c}},\mathbf{k}) = \mathcal{D}(\mathbf{p},\mathbf{k})$, and introducing a pure translation into $\mathbf{H_e}$ results in the shifting: $x \rightarrow x+\Delta_x$ and $y \rightarrow y+\Delta_y$. The two parameter radial distortion model then becomes:

$$\begin{aligned}\mathcal{D}_x(\mathbf{p}+\boldsymbol{\Delta},\mathbf{k}) =& k_1 x r^2 + k_2 x r^4 + k_1\Delta_x(3x^2+y^2) + 2k_1\Delta_y xy + \\ & k_2\Delta_x(5x^4+6x^2y^2+y^4) + k_2\Delta_y(4x^3y+4xy^3) + \\ & k_1\left((3\Delta_x^2+\Delta_y^2)x + 2\Delta_x\Delta_y y + \Delta_x\Delta_y^2 + \Delta_x^3\right) + \\ & k_2\left((5\Delta_x^4+6\Delta_x^2\Delta_y^2+\Delta_y^4)x + (4\Delta_x^3\Delta_y + 4\Delta_x\Delta_y^3)y + \right. \\ & \left. \Delta_x^5 + 2\Delta_x^3\Delta_y^2 + \Delta_x\Delta_y^4 + \ldots\right)\end{aligned}$$

and similarly for the y component of the distortion equation. Radial distortion is modelled as before, while $k_1\Delta_x$, $k_1\Delta_y$, $k_2\Delta_x$ and $k_2\Delta_y$ exactly match the decentering terms derived from the wave aberration equation 2.2 in Section 2.1. The basic introduction of a translation element into $\mathbf{H_e}$ thus models decentering distortion with the additional introduction of a small residual E. This residual may additionally be approximated by a small affine transformation $E = \boldsymbol{\Delta}_\mathbf{H}\mathbf{p}$. □

Therefore the objective error $\mathbf{e}$ may be re-written, implicitly modelling decentering distortion as:

$$\mathbf{e}(\check{\mathbf{c}},\boldsymbol{\Phi}) = \mathbf{H_e}\bar{\mathbf{c}} + \mathcal{D}(\mathbf{H_e}\bar{\mathbf{c}},\mathbf{k}) + \boldsymbol{\Delta}_\mathbf{H}\mathbf{H_e}\bar{\mathbf{c}} - \mathbf{A}\check{\mathbf{c}} \tag{4.21}$$

In the least square estimation, the small decentering residual $\boldsymbol{\Delta}_\mathbf{H}\mathbf{H_e}\bar{\mathbf{c}}$ is absorbed by a combination of $\mathbf{H_e}$ and $\mathbf{A}$ in the search for a global minimum. It should also be remembered that these equations are only approximations to the actual lens distortion, so such small adjustments are likely to have a negligible impact.

The effectiveness of this formulation in modelling decentering distortion is demonstrated by simulating radial and decentering distortion, and using this

data for calibration with the proposed method. Recalling from Section 2.1, that decentering parameters must be considered in conjunction with the radial distortion parameters, ie. $p_1 = k_1\Delta_x$ and $p_2 = k_1\Delta_y$. Thus to increase the effects of decentering a relatively low radial distortion level is chosen at $k_1 = -0.1$. The radial distortion level is set relatively low in order to induce the effect of larger decentering distortion. The decentering parameters are varied within the range $p_1 = p_2 = -0.01 \rightarrow 0.01$. Considering a normal 10mm square CCD element, this equates to an maximum misalignment of an individual lens element by 1mm which is fairly significant, for example, a small C-Mount lens may have an entire lens diameter of 20mm. Following calibration, the residuals are measured as before with a radially weighted homography and are shown in Fig. 4.27. This shows that there is a small pixel residual over the entire range of simulated decentering levels. When considered with respect to the the typical noise in the detected pixels if $\sigma = 0.15$ (pix) this residual becomes insignificant. Fig. 4.28 shows the decentering residuals when control point detection noise is included.

The advantage of using a perspective transform to model decentering distortion through the normal radial distortion equations is now investigated. A comparison is made with the formulation given by Stein (1993), which uses only a variable centre point for modelling decentering distortion. In this solution the reverse distortion model is used, which may be used in the following objective error to calibrate distortion:

$$\mathbf{e}(\breve{\mathbf{c}}, \mathbf{\Phi}) = \breve{\mathbf{c}} + \mathcal{D}(\breve{\mathbf{p}} + \mathbf{\Delta}, \mathbf{k}) - \mathbf{H}_e\bar{\mathbf{c}} \tag{4.22}$$

Using the same simulation data as in Fig. 4.27, the residuals resulting from this formulation are measured. These are shown in Fig. 4.29, demonstrating that this method is much less effective at modelling decentering distortion compared to the proposed one, which utilises a perspective transform. Significantly, the errors are much greater than the typical control point noise.

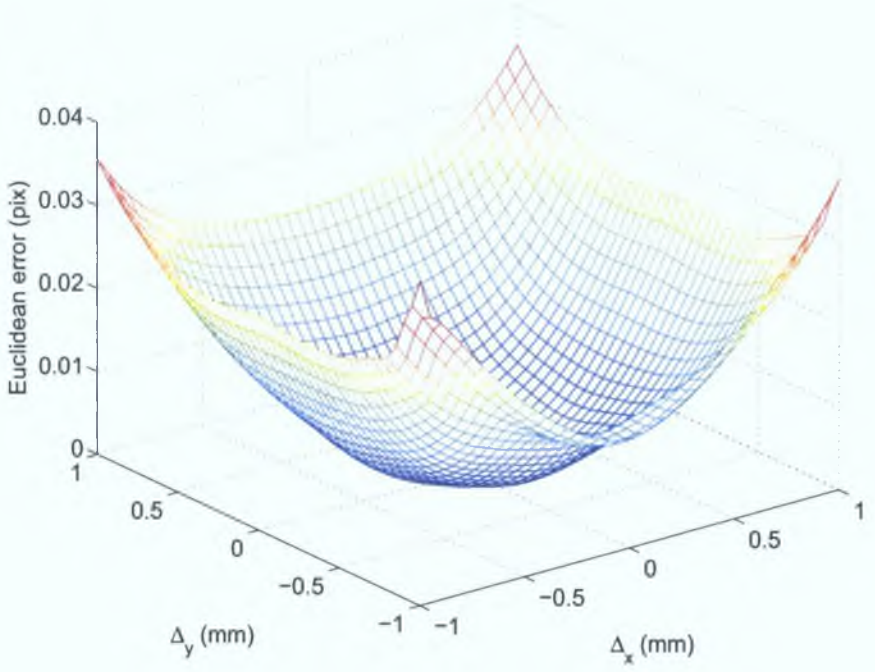


Fig. 4.27: The decentering residuals following calibration with the proposed method. Original data is simulated with radial and decentering distortion. The calibration model is radial only. The residual distortion (pix) is shown for the estimate.

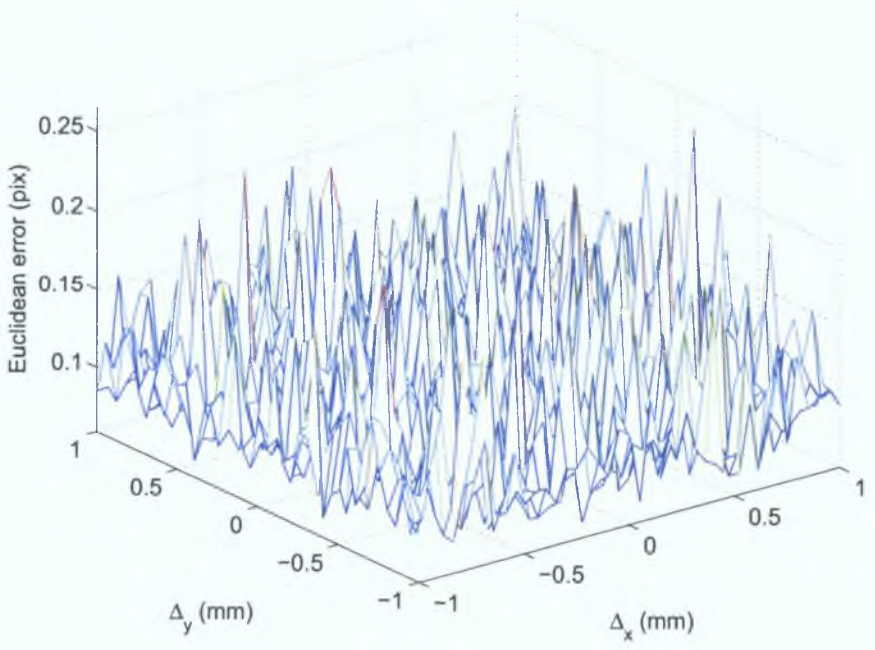


Fig. 4.28: The decentering residuals following calibration with the proposed method. Typical control point detection noise is added to the simulated radial and decentering data. Following calibration the small decentering distortion residual in Fig. 4.27 is negligible in comparison.

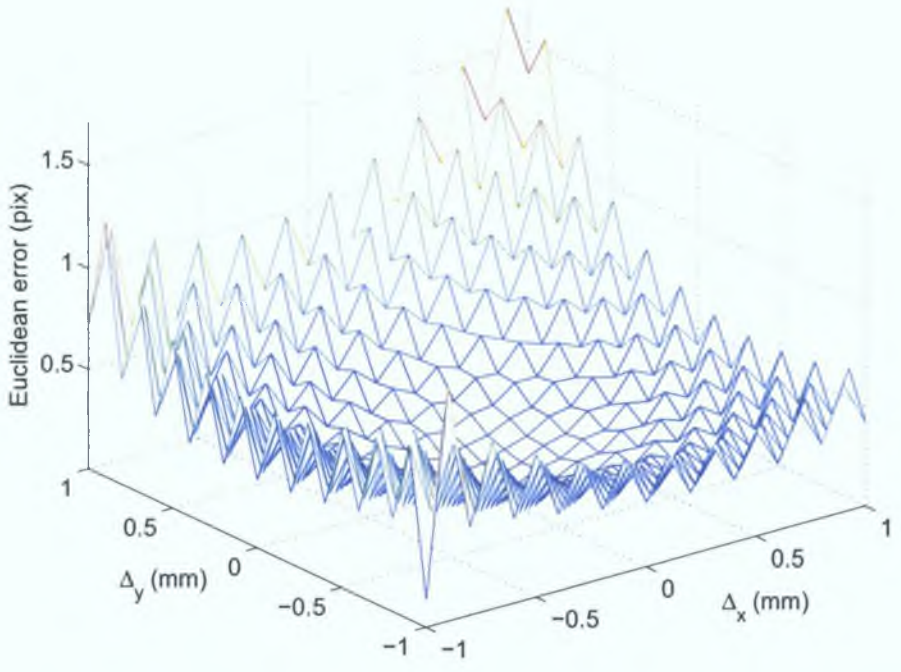


Fig. 4.29: The decentering residuals following calibration with the method proposed by Stein (1993), where decentering is approximated solely by a variable centre point.

4.4.1 Distortion centre and the principal point

It has been proposed, for example in Ahmed and Farag (2001), that the inclusion of decentering distortion parameters will compensate for an error in the centre point. This idea may be used to fix the centre point, say at the image centre, thus removing the necessity to estimate it. This idea has been shown to be flawed by Clarke et al. (1998) through simple experiments. It is slightly curious why, if a variable centre point accounts for decentering distortion to a high degree, does decentering distortion not properly compensate for a fixed centre point?

Considering the objective error $\mathbf{e}$ with a fixed centre point that is displaced from the true centre point by $\boldsymbol{\Delta} = (\Delta_x, \Delta_y)^T$ the objective function can be written as:

$$\mathbf{e}(\breve{\mathbf{c}} + \boldsymbol{\Delta}, \boldsymbol{\Phi}) = \mathbf{H}_\mathbf{e}\tilde{\mathbf{c}} + \boldsymbol{\Delta} + \mathcal{D}_R(\mathbf{H}_\mathbf{e}\tilde{\mathbf{c}} + \boldsymbol{\Delta}, \mathbf{k}) - \mathbf{A}\breve{\mathbf{c}} - \boldsymbol{\Delta},$$

where $\mathcal{D}_R(.)$ refers to the radial distortion component. The decentering component is likewise concisely referred to as $\mathcal{D}_D(.)$. Decentering distortion is explicitly added to compensate for the the miscalculation of the radial distor-

tion at the incorrect centre point as:

$$\mathbf{e}(\breve{\mathbf{c}} + \boldsymbol{\Delta}, \boldsymbol{\Phi}) = \mathbf{H}_\mathbf{e}\bar{\mathbf{c}} + \mathcal{D}_R(\mathbf{H}_\mathbf{e}\bar{\mathbf{c}} + \boldsymbol{\Delta}, \mathbf{k}) - \mathcal{D}_D(\mathbf{H}_\mathbf{e}\bar{\mathbf{c}} + \boldsymbol{\Delta}, \mathbf{k}) - \mathbf{A}\breve{\mathbf{c}}.$$

Using Result 5 this gives:

$$\mathbf{e}(\breve{\mathbf{c}} + \boldsymbol{\Delta}, \boldsymbol{\Phi}) \approx \mathbf{H}_\mathbf{e}\bar{\mathbf{c}} + \mathcal{D}_R(\mathbf{H}_\mathbf{e}\bar{\mathbf{c}}, \mathbf{k}) + \mathcal{D}_D(\mathbf{H}_\mathbf{e}\bar{\mathbf{c}}, \mathbf{k}) - \mathcal{D}_D(\mathbf{H}_\mathbf{e}\bar{\mathbf{c}} + \boldsymbol{\Delta}, \mathbf{k}) - \mathbf{A}\breve{\mathbf{c}}.$$

Clearly, the decentering distortion terms do not properly cancel each other, with one being evaluated at a different location. Thus with an increase in the centre point error $\boldsymbol{\Delta}$, there is a corresponding rise in the modelling error. This trend is shown in experiments with real images. The centre point is fixed at various intervals with a ± 40 pixel distance from the true location on an image size of 1312×1032 pixels. The calibration is then carried out with this fixed centre point with and without the inclusion of decentering terms. Fig. 4.30 shows the distortion residual resulting from the calibration without decentering elements. A larger error is accumulated due to the mis-calibration of radial distortion about an incorrect centre. The residuals resulting from the calibration with decentering terms are shown in Fig. 4.31 showing a decrease in error. However, the remaining residual is unacceptable, except for very small displacements of the centre point. In reality, such accurate placement is not possible.

To assess the likely magnitude of decentering distortion in some of the lenses used, the distortion centre point is compared with the estimated principal point, computed using the method proposed in Sturm and Maybank (1999). A maximum of nine images of the calibration pattern are taken from different locations, from which the principal point is estimated incrementally, for example, for image number five, five homographies are input in the internal calibration algorithm and so on. The centre point is estimated independently for each sample, with the proposed method. The results are shown in Fig. 4.32. For the integral Fuji lens, the x coordinate overall assumes slightly lower values, indicating that there is a slight negative x directional decentering distortion or a negative p_1 parameter. This corresponds with the approximations in table 4.14. For the integral Nikon lens, there is also a negative x directional distortion and a very slight negative y element. Finally, for the 6mm lens that was known to have decentering elements, a large negative x or p_1 direction value is observed, with relatively constant distortion in the y or p_2 direction. This is again in good agreement with the estimates in table 4.14.

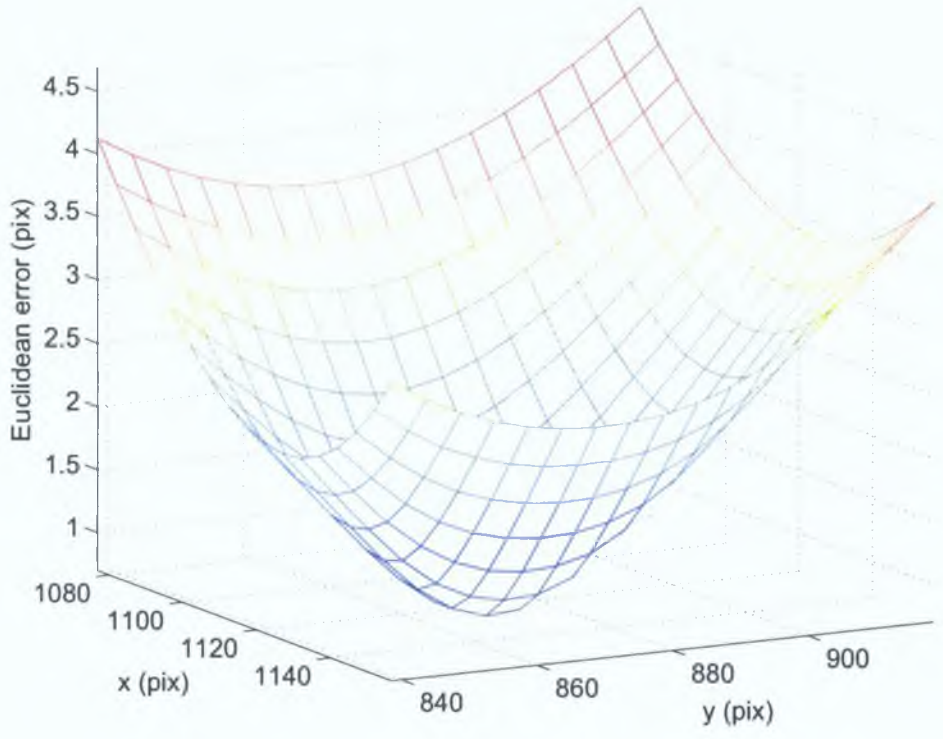


Fig. 4.30: Residual distortion errors following calibration without the explicit inclusion of decentering distortion. The ideal centre point is located at (1116,878) with ±40 pixel displacement of the centre point.

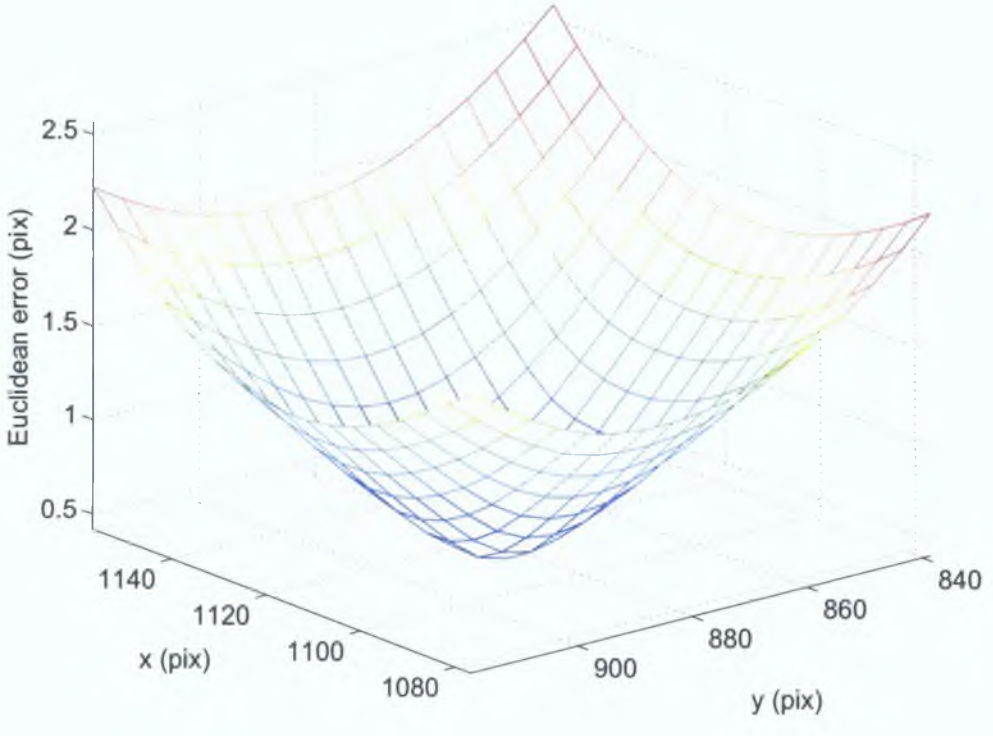
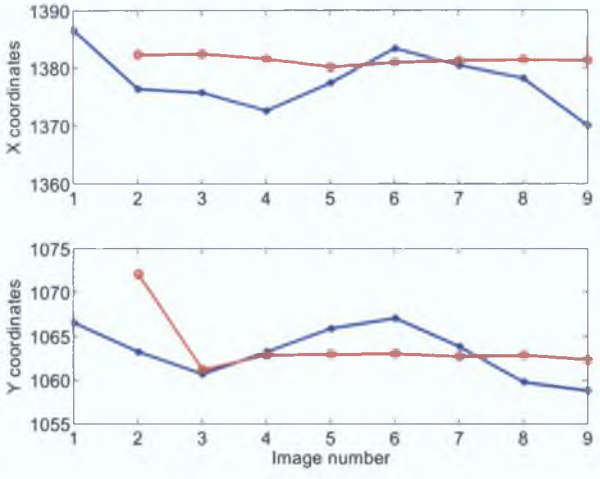
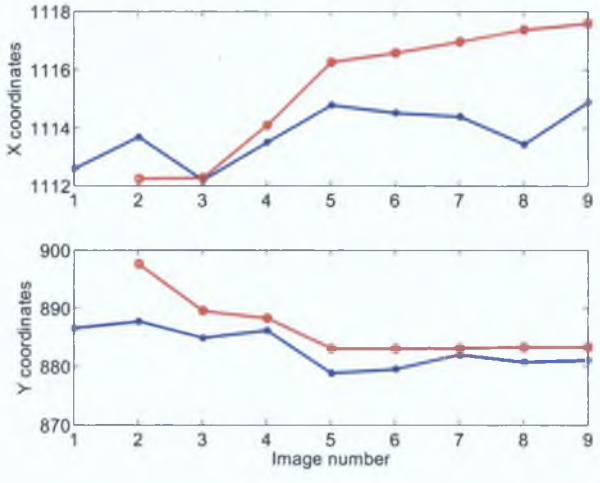


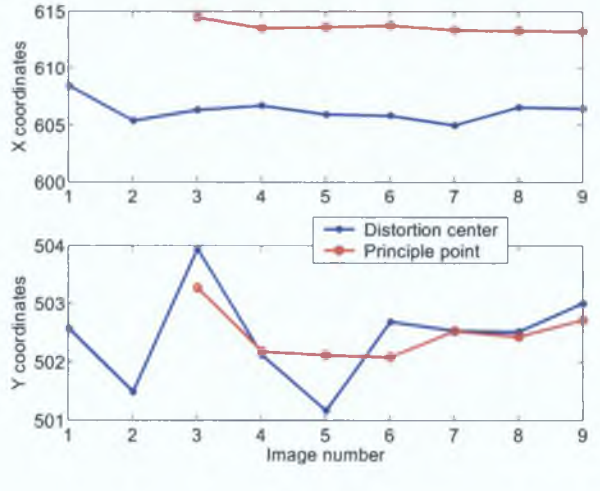
Fig. 4.31: Residual distortion error following calibration with a fixed centre point and the inclusion of decentering terms to help compensate for the centre point errors.



(a) Fuji (Sample no. 1)



(b) Nikon (Sample no. 2)



(c) Kodak (Sample no. 3)

Fig. 4.32: Comparison between principal point and distortion centers.

In summary, the proposed method of calibrating lens distortion implicitly includes decentering distortion. It is effectively calibrated by the LS process, through the use of projectively adjusted points in the standard radial distortion equations. This explains why the addition of decentering distortion did not improve the distortion residuals, with the redundant parameters reflected in the high correlations. Also, it clarifies the reason why the centre point is likely to be indeterminable if decentering distortion elements are explicitly included as reported in (Brown, 1971). From the simulated results it has been shown that the proposed method models decentering to a higher accuracy than the previous approximation in Stein (1993). Crucially, this precision is better than the typical noise in the detected control points.

The relationship between the distortion centre, and decentering distortion is examined further. It is shown theoretically and experimentally that the inclusion of decentering distortion parameters does not compensate for a fixed and erroneous centre point. An extra illustration of the modelling of decentering distortion by a variable centre point is given by comparing it with the estimated principal point using an alternative method. This shows that the integral lenses on the digital cameras used show relatively little decentering distortion. The magnitude of decentering distortion in a low cost 6mm lens is demonstrated by the large shift in the distortion centre in comparison to the principal point.

4.5 Focusing variation of distortion

So far this study has been dealing with entirely fixed lenses. Most basic lenses allow the facility to alter the focusing distance and the aperture opening. The focusing is implemented by a relative movement between the lens elements and/or the sensor. It is known that changing the focusing distance will impact on the distortion content. This section aims to identify the nature and magnitude of the distortion variation resulting from the variable focusing distance. The influence of the aperture variation of the lens is also examined.

It is known from Brown (1971) that there is a variation of distortion with object focusing distance. Considering s as the distance of a focusing plane,

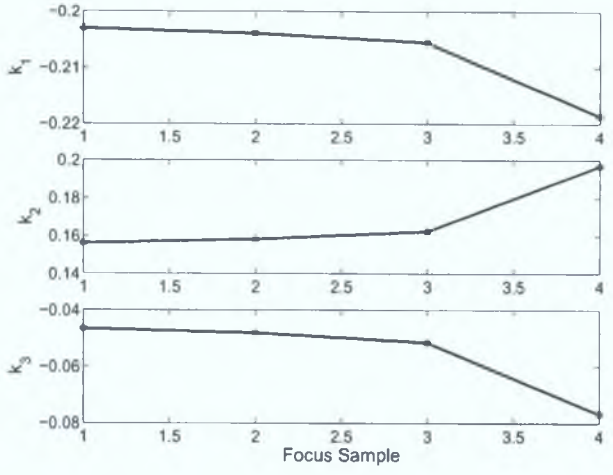
the distortion variation at another distance s' may be interpolated with:

$$\mathcal{D}(\mathbf{p},\mathbf{k})_{s,s'} = \gamma_{s,s'}^2 k_1 r^3 + \gamma_{s,s'}^4 k_2 r^5 + \gamma_{s,s'}^6 k_3 r^7 + \cdots , \tag{4.23}$$

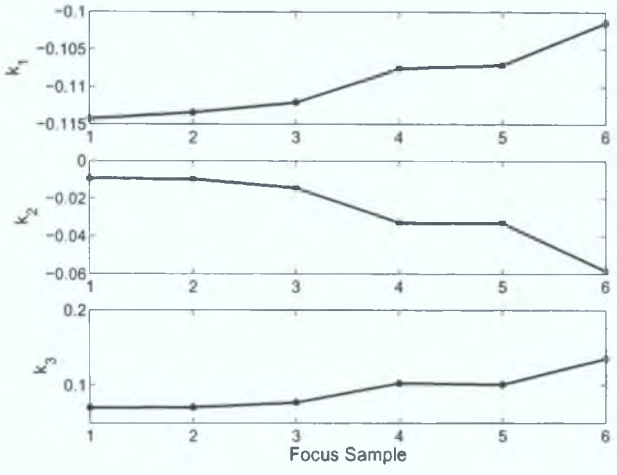
where $\gamma_{s,s'} = \frac{s-f}{s'-f}\frac{s'}{s}$ with f as the focal length of the lens. By calibrating distortion for at least two different focus settings, the distortion at other locations may be approximated using equation 4.23

The variation of distortion is investigated with three different lenses, a 6mm Computar and a Fuji 6.8mm lens that were used before and described in table 4.6, and a Nikon 7.85mm lens. From Chapter 3 it has been shown that the edge based intersection method is quite invariant to blurring. This allows the extraction of control points at a number of different focusing distances, while the calibration pattern distance remains fixed. These control points are then used in the calibration routine. Unfortunately, it is not possible to get a meaningful measure of the focusing distance. The Fuji camera allows the manual manipulation of the focusing distance, which was stepped through with equal spacing from far to near. The Nikon camera uses a motorized focusing adjustment, though the individual step increments are not available. It is stepped from near to far. Finally, the 6mm lens uses a screw adjustment, which was again varied linearly from near to far. The distortion parameter variations over the focusing samples are presented in Fig. 4.33. This shows that distortion parameters for the far to near focus variations follow a form similar to equation 4.23, where the negative k_1 and k_3 parameters becomes more negative while the positive k_2 parameter increases, and the opposite for the near to far examples. The Fuji example increases most for the final sample, indicating that the dial and the focusing distance are not that linearly connected. Excluding the fourth sample, the Nikon camera distortion variation shows a smooth variation in parameters. The 6mm lens displays the greatest magnitude in parameter variation while also showing a steady parameter variation. These variations in parameters are now examined in pixel terms throughout the image.

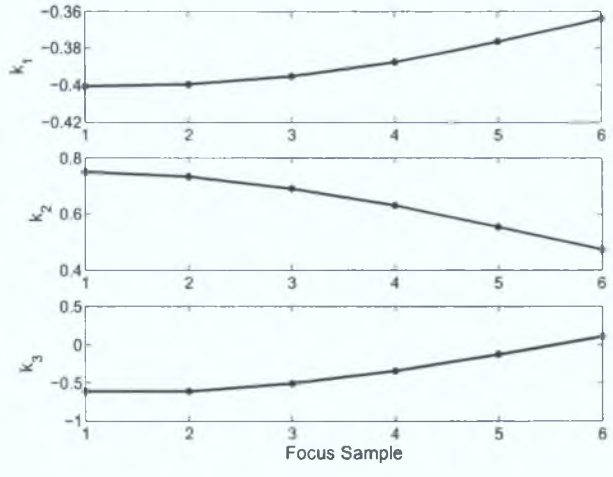
To quantify the actual pixel manifestation of these variations, the first and last focusing positions for each camera are compared. The vector fields in Fig. 4.34 show these variations throughout the image field. The Fuji example shows a radial field, with a slight decentering influence. The Nikon field is predominantly radial, while the 6mm lens exhibits a considerable decentering element. This x direction decentering content has previously been identified as shown in Fig. 4.32.



(a) Fuji 6.2mm

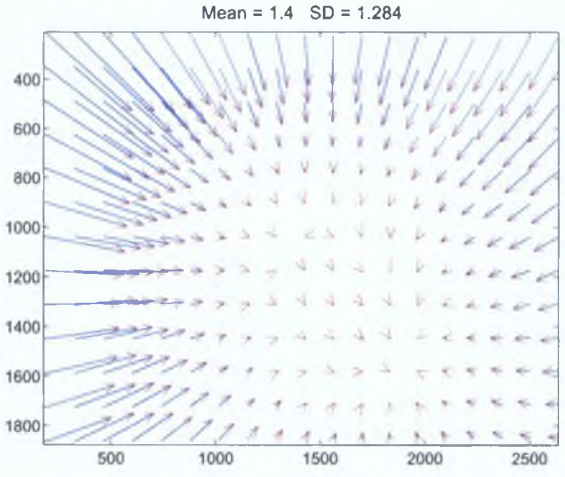


(b) Nikon 7.8mm

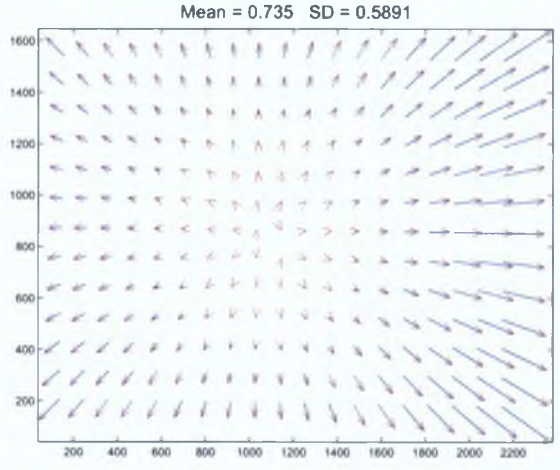


(c) Computar 6mm)

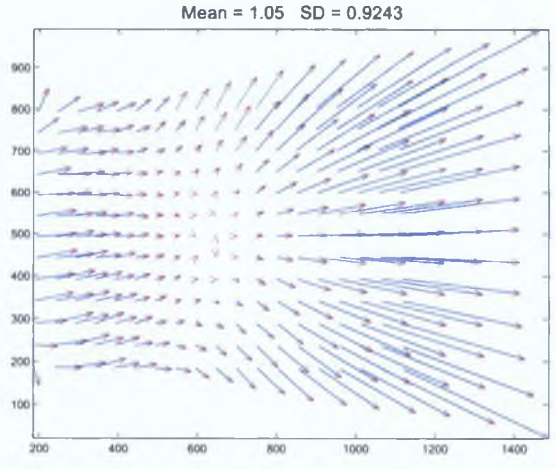
Fig. 4.33: Variation of distortion parameters with focus position.



(a) Fuji 6.2mm



(b) Nikon 7.8mm



(c) Computar 6mm)

Fig. 4.34: Vector plot showing the variation of distortion between the first position and last focus positions. Residual scale is $\times 100$ throughout.

4.5.1 Aperture variation of distortion

In Section 2.1, it was assumed that the aperture plane was fixed. Most lenses allow the aperture to be varied, controlling the amount of light entering the lens. With an increasingly open aperture, there is direct increase in the light rays intersecting the CCD array. Altering this quantity will thus theoretically affect the distortion content as these rays traverse different portions of the lens. To what extent this happens is unclear however. Thus, this subsection experimentally investigates the actual variation of lens distortion with aperture setting.

The same cameras and lenses that were utilised in Section 4.5 are again employed. On this occasion the lens focal lengths and the camera positions were fixed, while the aperture was varied manually. The digital cameras allow the F-stops to be changed, but no such quantities are generally marked for low cost C-Mount lenses. Thus, the exposure time was used as a base for comparison, taken from the EXIF tags for the digital camera images, and from the acquisition software for the C-Mount lens images. The distortion is again calibrated using the proposed method, including two distortion parameters for simplicity in the presentation of results.

Fig. 4.35 traces the variation of the distortion parameters for the three different lenses as the aperture is altered. Clearly, there is an aperture related variation in the image distortion. The Fuji and Nikon lens distortions increase steadily and slightly trail off toward the closed aperture position. The C-Mount lens distortion rapidly increased up to 50 ms exposure time, and then settles off. The variations are lower than that encountered for the variable focus experiments. In terms of the pixel manifestation of these differences, the maximum distortion variations are chosen for comparison in Fig. 4.36. This worst case scenario shows that the aperture variation in pixel terms is less than the focus variation differences. Indeed for the Nikon lens, the difference in the distortion profiles is close to the normal observed control point noise, effectively rendering it insignificant. Again the larger distortion lens displays a greater tendency for aperture influenced distortion variance.

In summary, the focusing distance alters the lens distortion profile. As an example, a high resolution image showed a mean pixel difference of 1.4 over the considered focusing range. However, since the focusing distance or indeed

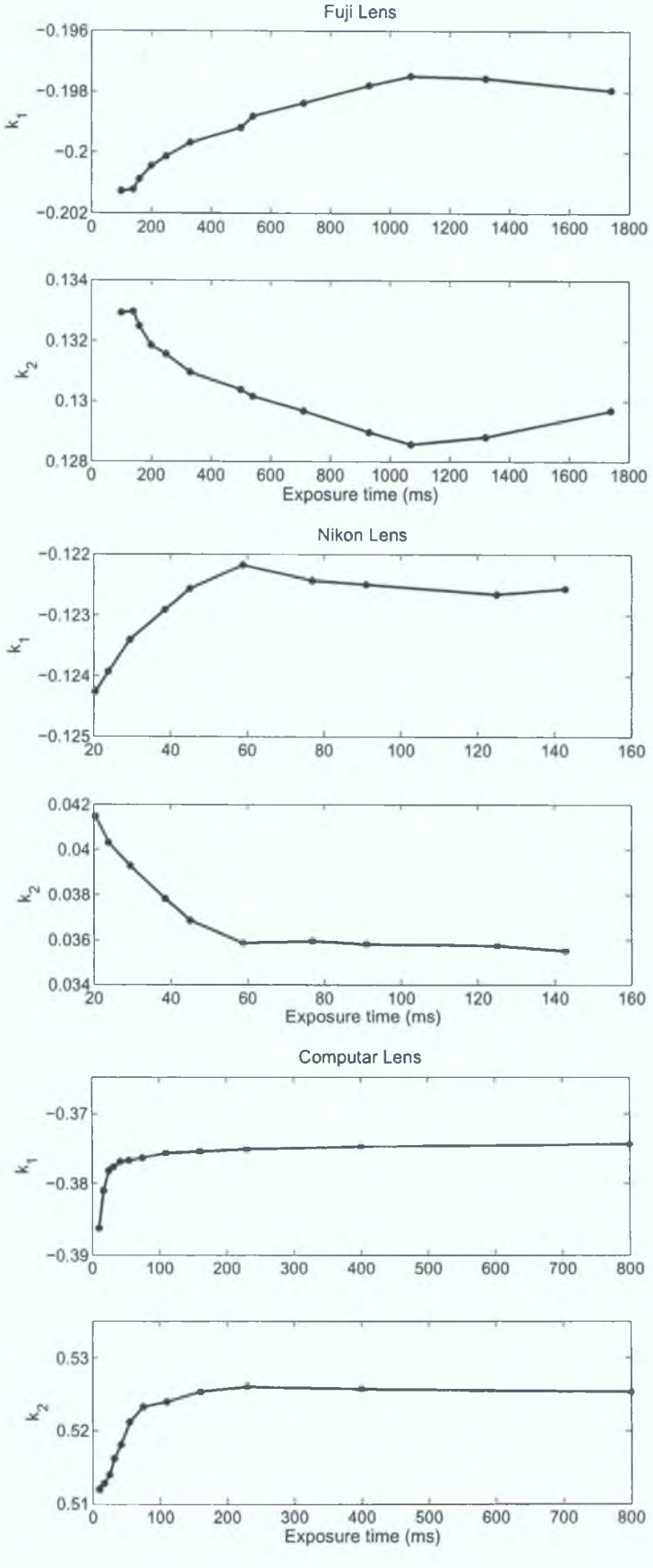
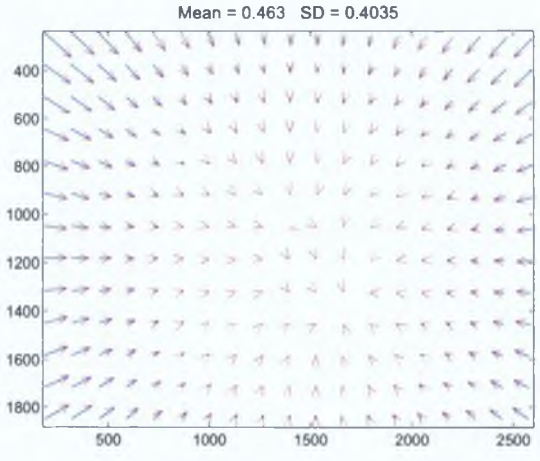
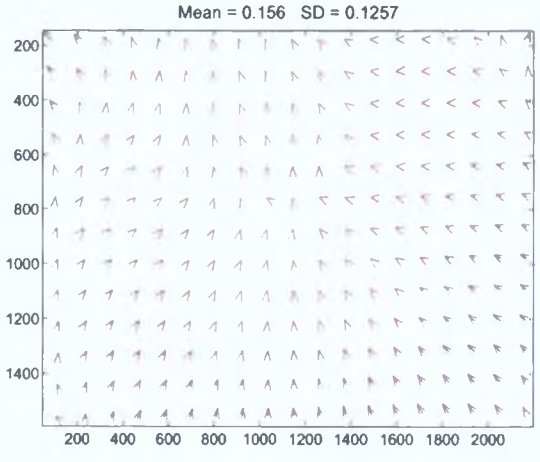


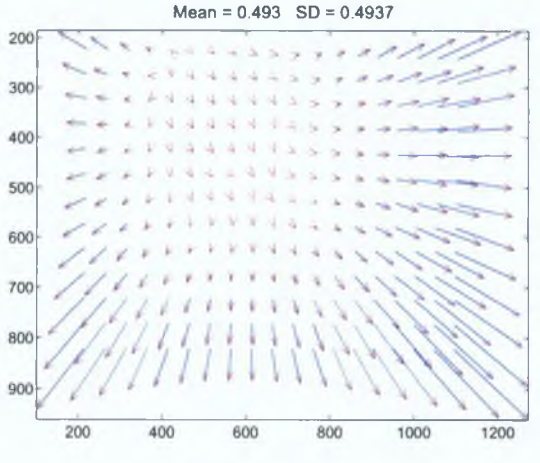
Fig. 4.35: Variation of distortion parameters with aperture position, described in terms of exposure time.



(a) Fuji 6.2mm



(b) Nikon 7.8mm



(c) Computar 6mm)

Fig. 4.36: Vector plot showing the variation of distortion between the maximum parameter variations over the range of aperture positions. Residual scale is $\times 100$ throughout.

any other meaningful focus measure is unavailable in most basic lenses, the variation equation 4.23 can not be used. Higher distortion lenses naturally show a greater variation than low distortion lenses. High distortion is generally encountered in wide angle lenses which tend to be used as fixed focus lenses in many applications. The extreme fish-eye lenses are usually fixed at infinity focus. The aperture setting also influences the image lens distortion profile, but to a lesser degree than the change in focusing. In the same high resolution example, the distortion varied by a mean of 0.45 pixels over the range of aperture settings. Of course, in many applications the aperture setting remains constant.

4.6 Discussion

This chapter has detailed a non-metric technique to compensate for geometric lens distortion in images. It is based on a single view of a calibration target, from which an appropriate distortion model is calibrated by exploiting two geometric constraints on the targets planarity and known pattern distribution. The calibration procedure advances the following contributions.

- The use of the forward distortion model in calibration, in comparison to existing non-metric approaches which cannot avail of its advantages (see Chapter 2).
- The parameter estimation problem presented as a closed-form system of equations, with full partial derivatives included.
- A novel means of linearly identifying a new image scaling to minimise the local distortion of pixels in the compensation of distortion in the image array.

An extensive set of experiments are conducted on the proposed method, while its performance is compared with respect to three alternative techniques for calibrating distortion. The comparisons are conducted on both simulated and real data and demonstrate that the proposed method is highly accurate, and stable. It is shown to at least match, and often surpass the performance of the full camera calibration techniques, and consistently outperforms other non-metric methods. Its application to potentially problematic sub-pixel distortion

lenses revels that the algorithm retains its stability and accuracy, allowing such images to be further improved. The dependence of the proposed method on the planar calibration target is comprehensively addressed. For reliable calibration at least 200 control points are recommended. It is shown that multiple input views offer no advantage save in the supply of additional control points. The degree to which the two geometric constraints on the calibration target must be satisfied is also examined. Regarding the required precision of the pattern, it is shown that for the 250×170 mm sized pattern used, there is roughly a direct transfer of error from millimeters to pixels with a 6MegaPixel camera resolution. Concerning the required planarity of the pattern, a similar relationship is found, where roughly 1mm bending of the chart will induce a mean of 1 pixel error, again for a 6MegaPixel resolution. Out of interest, the use of circular type patterns is shown to induce a similar 1 pixel error. Finally, the influence of lens focusing and aperture settings on the distortion profile is examined. This shows that the maximum mean distortion variations due to focusing, for example in a 6MegaPixel camera, is 1.4 pixels and roughly 0.5 pixels for the maximum aperture variation. Without readily available measures of the focusing distance or the aperture, compensations for these variations cannot be applied.

In conclusion, the proposed approach has the inherent advantage of being generally applicability to all levels of lens distortion, from very mild up to fish-eye. It implicitly incorporates quite an effective means of modelling decentering lens distortion, which adjusts as required. Lastly, it is straightforward to implement and use, overall offering an attractive and viable alternative to current distortion calibration and removal solutions.

Chapter 5

Lateral Chromatic Aberration Removal

An optical instrument is required to faithfully produce a geometrically consistent image of a given object. The departure of practical optical systems from this ideal (Gaussian or first order) behaviour is due to aberrations. In general it is impossible to design a system which is free from all aberrations. This leads lens manufacturers to consider aberration compensation as an optimisation between different types. This chapter is concerned with chromatic aberrations that have recently become more amplified due to the higher resolution sensors currently employed in many consumer and scientific cameras. By compensating for these aberrations as a post process in the image array, higher quality images can be produced without recourse to expensive optics.

In a colour camera's lens, polychromatic light is split into a set of rays or wavelengths. Whilst traversing the optical system light of different wavelengths will follow slightly different paths. Upon reaching the image plane their misaligned recombination introduces chromatic aberration. Chromatic Aberration (CA) can be broadly classified as Axial Chromatic Aberration (ACA) (also known as Longitudinal CA) and Lateral Chromatic Aberration (LCA) (also known as Transverse CA). ACA arises from the longitudinal variation of focal position with wavelength along the optical axis. LCA is the variation of image size with wavelength or the vertical off-axis distance of a point from its prescribed point. In an image it is identified by a radially dependent misalignment of the colour planes. Chromatic aberrations are moving out of the sub-pixel range with

the advent of high resolution arrays, giving rise to noticeable colour fringes around edges and high contrast areas. This gives the overall impression of poor quality or definition. Many consumer cameras display this aberration. For scientific applications, it is akin to the effects of colour shifts and blurring, that contravene the imaging models. The digital compensation of LCA through image warping is considered here. There are two main aspects of digital compensation in images: determining what quantity of warp to apply, and the actual implementation of the warp. The main contribution deals with the former problem, which has currently not been addressed, by considering the modelling and model calibration of LCA in images.

The proposed compensation is achieved by realigning the colour planes through image warping. Using the LCA model derived in Section 2.3, a more precise and concise means of extending the aberration, measured over a limited set of control points, to every pixel in the colour plane is facilitated. LCA is initially measured by extracting the intersections of a chessboard pattern on each colour plane. No special planarity constraints or canonical representation of the pattern is required and it can be imaged without knowing its 3D position. Measurement errors are filtered by non-linear least square fitting of the proposed LCA model. The partial derivatives of the quadratic cost functions are given allowing the closed-form computation of the gradients and Hessian matrices used by the optimisation algorithms. This gives a computational advantage over numerical estimation techniques. Detailed results clearly demonstrate the successful compensation of LCA for test images and for real scenes.

5.1 Geometrical Theory of Aberrations

Optically, aberrations are compensated for by adding lens elements with appropriate properties. Chromatic aberration is typically eliminated for two selected wavelengths, but only at the centre and some zonal region. These lenses are known as achromatised. Lenses corrected for three different wavelengths are known as apochromatic while superachromatic lenses are corrected for four wavelengths. Of interest are the the remaining chromatic aberrations, known as the secondary spectrum. No distinctions are made between types of corrected lenses, as the derived model is generally applicable. Willson (1994) and

Willson and Shafer (1991) show that chromatic aberrations can be compensated in an image by re-alignments of the colour channels.

As described in Section 2.3 Lateral Chromatic Aberration can be considered as the sum of two aberrations: lateral colour distortion due to the refraction index of the lens elements and the chromatic variation of distortion (Kingslake, 1978). An appropriate model for the chromatic variation of distortion in one colour plane g may then be described relative to another colour plane f as:

$$\begin{aligned} \mathcal{C}^g(\mathbf{p}_f, \mathbf{c}_g)_x &= c_1 x_f + c_2 x_f r_f^2 + c_3(3x_f^2 + y_f^2) + 2c_4 x_f y_f \\ \mathcal{C}^g(\mathbf{p}_f, \mathbf{c}_g)_y &= c_1 y_f + c_2 y_f r_f^2 + 2c_3 x_f y_f + c_4(3y_f^2 + x_f^2), \end{aligned} \tag{5.1}$$

where $\mathbf{c}_f = (c_1, c_2, c_3, c_4)^T \in \mathbb{R}^4$ is the parameter vector and $\mathbf{p}_f = (x_f, y_f, 1)^T$ are the homogeneous lens centric coordinates in the f colour plane.

5.2 Model Calibration

Lateral chromatic aberration is modelled for a specific frequency according to equation 5.1. The actual secondary spectrum is difficult to exactly quantify, but manifests itself by misalignments in the colour planes as demonstrated by Willson (1994). These planes typically match the RGB filters of a typical colour sensor, though other colour representations can be used, as the methods are general. If one colour plane is taken as a reference, chromatic aberration can be compensated for by realigning the other planes with this reference. This reference colour is chosen as the Green (G) channel, as it is midway within the visible spectrum and is dominant in the standard Bayer array used in digital cameras[1].

5.2.1 Measuring lateral chromatic aberrations

Chromatic aberration has been previously measured by Kuzubek and Matula (2000) using florescent dyed beads. These are then imaged in 3D, when their centroids are estimated. From these centroids the LCA and ACA are measured. This approach is only suited to fluorescent microcopy, but the measured LCA

[1]The impact of correlations between the colour channels due to the interpolation of the raw sensor data is not addressed in this work.

exhibits a similar profile to the results obtained using the proposed approach. Willson (1994), measures chromatic aberration by comparing the location of edges detected on three colour planes. In this chapter lateral chromatic aberration is again measured by detecting the intersections of a chessboard pattern for each of the colour planes. These are automatically extracted by the two stage process of initial detection and sub-pixel refinement as outlined in Section 3.2.2.

5.2.2 Chromatic parameter estimation

The pattern intersection points are represented in pixel coordinates as $\bar{\mathbf{c}}_f = (\bar{u}_f, \bar{v}_f, 1)^T$ for a certain colour plane f. Given the average of the image width and height as w, the intersection coordinates are normalised by scaling $\mathbf{c}_f = (\bar{u}_f, \bar{v}_f, w)^T = (u_f, v_f, 1)^T$. This does not affect the chromatic distortion calibration, following from Result 3. The required transform taking the points $\mathbf{c}_f$ to the lens centric coordinates $\mathbf{p}_f$ is then defined as:

$$\mathbf{p}_f = \begin{pmatrix} s & 0 & -u_o \\ 0 & 1 & -v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_f \\ v_f \\ 1 \end{pmatrix} = \mathbf{A}\mathbf{c}_f \tag{5.2}$$

where s is the applied compensation for non-square pixels and the unknown symmetrical centre of the aberration is $\mathbf{t}_f = (u_o, v_o, 1)^T$. These parameters need to be estimated in the calibration. The units of $\mathbf{p}_f$ are normalised pixels, (pix). It should be noted that Result 5, regarding the modelling of the decentering distortion with a variable centre point, may not be used in this case. This is due to the extra radial term c_1, which under an incorrect centre point gives rise to a directional bias. Thus the centre point and the decentering elements must be explicitly estimated.

The lateral misalignments between the red and green planes are modelled as a function of the green plane, following equation 5.1 as:

$$\begin{pmatrix} C^r(\mathbf{p}_g, \mathbf{c}_r)_x \\ C^r(\mathbf{p}_g, \mathbf{c}_r)_y \end{pmatrix} = \begin{pmatrix} c_1 x_g + c_2 x_g r_g^2 + c_3(3x_g^2 + y_g^2) + 2c_4 x_g y_g) \\ c_1 y_g + c_2 y_g r_g^2 + 2c_3 x_g y_g + c_4(3y_g^2 + x_g^2) \end{pmatrix}, \tag{5.3}$$

and similarly for the difference between the blue and green planes. For each detected intersection point, two equations are formed. It is sufficient to follow

these equations with respect to the red/green planes only:

$$e(\mathbf{p}_r|\mathbf{p}_g, \hat{\mathbf{\Phi}}_r) = \begin{pmatrix} e_x(\mathbf{p}_r|\mathbf{p}_g, \hat{\mathbf{\Phi}}_r) \\ e_y(\mathbf{p}_r|\mathbf{p}_g, \hat{\mathbf{\Phi}}_r) \end{pmatrix} = \begin{pmatrix} u_g + \mathcal{C}^r(\mathbf{p}_g, \mathbf{c}_r)_x - u_r \\ v_g + \mathcal{C}^r(\mathbf{p}_g, \mathbf{c}_r)_y - u_r \end{pmatrix}, \quad (5.4)$$

where the parameter vector to be estimated is $\hat{\mathbf{\Phi}}_r = (u_o, v_o, s, c_1, c_2, c_3, c_4)^T$, where $\hat{\mathbf{\Phi}}_r \in \mathbb{R}^7$.

A parameter counting exercise reveals that a minimum of 4 control points are required to solve this system of equations. Given at least $n \geq 4$ observations a least square solution is obtained by:

$$\min \sum_{i=1}^{n} \mathbf{e}_i(\mathbf{p}_r|\mathbf{p}_g, \hat{\mathbf{\Phi}}_r)^2. \quad (5.5)$$

This quadratic cost function may be linearised by performing a first order expansion of the error around the last iterative estimate $\hat{\mathbf{\Phi}}^k$, resulting in a Gauss-Newton scheme that can be iterated utilising many robust least square techniques (Golub and Loan, 1996):

$$\hat{\mathbf{\Phi}}^{k+1} = \hat{\mathbf{\Phi}}^k - \lambda \left(\frac{\partial e^T(\hat{\mathbf{\Phi}}^k)}{\partial \hat{\mathbf{\Phi}}} \frac{\partial e(\hat{\mathbf{\Phi}}^k)}{\partial \mathbf{\Phi}^T} \right)^{-1} \frac{\partial e^T(\hat{\mathbf{\Phi}}^k)}{\partial \mathbf{\Phi}} e(\hat{\mathbf{\Phi}}^k), \quad (5.6)$$

where $\lambda \leq 1$ ensures a decrease in cost at each step. The partial derivatives used in the closed-form calculation are given as:

$$\frac{\partial e(\hat{\mathbf{\Phi}}^k)}{\partial \mathbf{\Phi}^T} = \begin{pmatrix} \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial \mathbf{\Phi}^T} \\ \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial \mathbf{\Phi}^T} \end{pmatrix} = \begin{pmatrix} \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial u_o}, \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial v_o}, \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial s}, x_g, x_g r_g^2, 3x_g^2 + y_g^2, 2x_g y_g \\ \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial u_o}, \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial v_o}, \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial s}, y_g, y_g r_g^2, 2x_g y_g, 3y_g^2 + x_g^2 \end{pmatrix},$$

with

$$\begin{pmatrix} \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial u_o} \\ \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial u_o} \end{pmatrix} = \begin{pmatrix} c_1 + c_2(3x_g^2 + y_g^2) + 6c_3 x_g + 2c_4 y_g \\ 2c_2 x_g y_g + 2c_3 y_g + 2c_4 x_g \end{pmatrix},$$

$$\begin{pmatrix} \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial v_o} \\ \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial v_o} \end{pmatrix} = \begin{pmatrix} 2c_2 x_g y_g + 2c_3 y_g + 2c_4 x_g \\ c_1 + c_2(x_g^2 + 3y_g^2) + 2c_3 x_g + 6c_4 y_g \end{pmatrix},$$

$$\begin{pmatrix} \frac{\partial e_x(\hat{\mathbf{\Phi}}^k)}{\partial s} \\ \frac{\partial e_y(\hat{\mathbf{\Phi}}^k)}{\partial s} \end{pmatrix} = \begin{pmatrix} c_1 u_g + 3c_2 x_g^2 u_g + 6c_3 x_g u_g + 2c_4 u_g y_g \\ 2c_3 y_g u_g + 2c_4 x_g u_g \end{pmatrix}.$$

Table 5.1: Description of the cameras used for the experiments

	Make and Model	Focal	Resolution
Cam 1	Nikon E4500	7.8mm	2272 × 1704
Cam 2	Fuji FinePix 6900	7.8mm	2832 × 2128
Cam 3	Fuji FinePix 6900 + WL-FX9	6.2mm	2832 × 2128

Equation 5.6 is iterated until $\hat{\mathbf{\Phi}}^{k+1} - \hat{\mathbf{\Phi}}^{k}$ falls below a preset threshold. The parameter vector can be simply initialised as $\mathbf{\Phi}_r^0 = (-.5, -.5, 1, 0, 0, 0, 0)^T$. Following calibration the colour planes are realigned using bilinear interpolation as described in equation 4.9.

5.3 Experiments

Chessboard patterns and real images are used to measure the effects of LCA compensation. Three different commercial digital cameras are used to capture the test images, briefly described in table 5.1. The pattern used for calibration is shown in Fig. 5.1. No canonical coordinates are required for calibration, hence no precise constraints are needed on the planarity or precision of the pattern. A second lower density chessboard pattern (test image), shown in Fig. 5.1, is used for independent validation of the proposed LCA model and the resulting realignments. Parameter uncertainties are also investigated, while shots of an outdoor scene are used to demonstrate the typical improvement in image quality following LCA compensation.

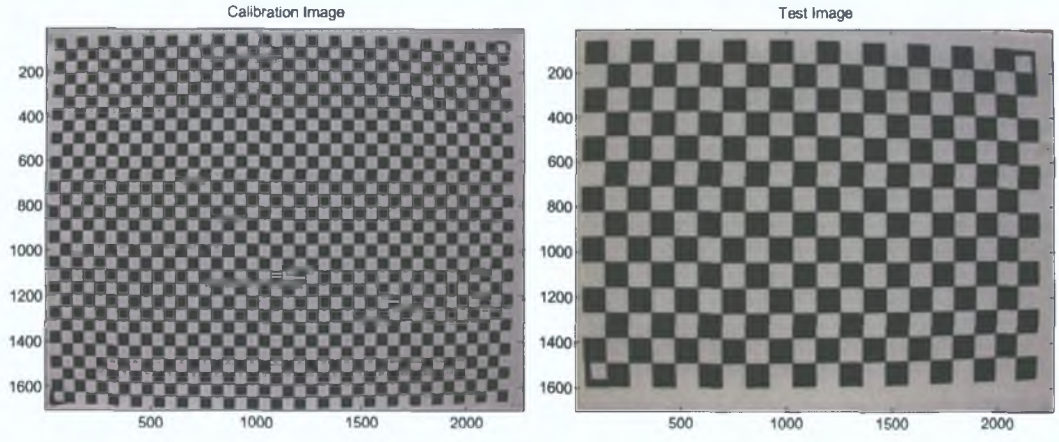


Fig. 5.1: Chessboard patterns used for calibration (calib image) and testing (test image) taken with cam 1, see tables 5.2 and 5.3.

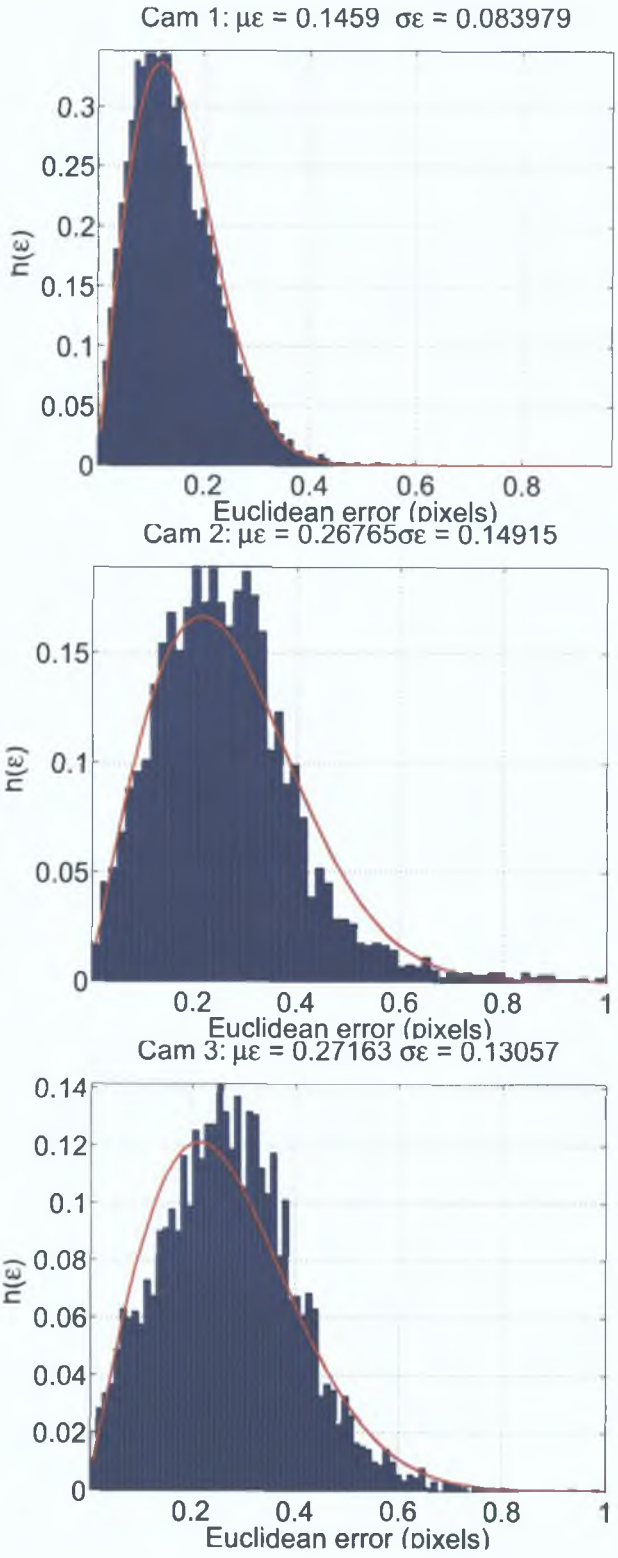


Fig. 5.2: Histogram of sub-pixel detection errors for three different cameras with their fitting with Rayleigh PDF. Errors are estimated using multiple shots of the calibration pattern.

Table 5.2: Colour plane misalignments (in pixels) before calibration in mean (SD) format for three different cameras. R/G and B/G are the red and blue misalignments with reference to the green channel.

		Cam 1	Cam 2	Cam 3
Calib Image	R/G	0.5707 (0.2113)	0.5496 (0.2308)	1.1834 (0.4125)
	B/G	0.4110 (0.2635)	0.7374 (0.6361)	0.5665 (0.3848)
Test Image	R/G	0.5355 (0.2225)	0.5413 (0.2035)	0.9729 (0.2866)
	B/G	0.4877 (0.2925)	1.1630 (0.8971)	0.8378 (0.7956)

The intersections of the chessboard patterns are firstly determined for each colour plane. The typical sub-pixel detection accuracy of the techniques outlined in Section 3.2.2 are shown in Fig. 5.2 for the three cameras used in the experiments.

5.3.1 Evaluation with real images

The colour plane misalignments before calibration for the two chessboard patterns are presented in table 5.2. Following calibration, the known LCA models are used to warp the colour planes so as to register the red and blue colour planes with the green channel. The Euclidean registration residuals remaining following this re-registration are presented in table 5.3, showing a significant decrease in misalignments. These residuals are of a similar magnitude to the sub-pixel detection accuracy, thus validating both the proposed LCA model and the effectiveness of the proposed calibration algorithm.

The contribution of the decentering LCA component is now evaluated. The results presented in table 5.4 show the Euclidean registration residuals following compensation based on a model without decentering elements. The increase in these residuals compared with those of the full calibration model indicated that although radial chromatic aberration is predominant, there is a varying element of decentering aberration depending on the lens employed. The inclusion of decentering elements in the LCA description gives a more general and accurate model of lateral chromatic aberration in an image.

More details of the colour plane misalignments before and after calibration are presented for one example (Cam 1) from table 5.2 and 5.3. Fig. 5.3 shows

Table 5.3: Colour plane misalignments (in pixels) following calibration and colour plane warping in mean (standard deviation) format for three different cameras.

		Cam 1	Cam 2	Cam 3
Calib Image	R/G	0.1202 (0.0636)	0.1401 (0.0733)	0.1846 (0.0722)
	B/G	0.1376 (0.0734)	0.1658 (0.0947)	0.1543 (0.0925)
Test Image	R/G	0.1788 (0.1062)	0.1625 (0.0784)	0.2044 (0.1149)
	B/G	0.1879 (0.1110)	0.3092 (0.2146)	0.3202 (0.2419)

Table 5.4: Colour plane misalignments (in pixels) following calibration and warping using a model without tangential elements in mean (standard deviation) format for three different cameras.

		Cam 1	Cam 2	Cam 3
Calib Image	R/G	0.1828 (0.0904)	0.1615 (0.0858)	0.2196 (0.1208)
	B/G	0.2131 (0.1117)	0.1805 (0.1087)	0.1507 (0.0754)
Test Image	R/G	0.1864 (0.1110)	0.2022 (0.1019)	0.1886 (0.1449)
	B/G	0.2071 (0.1334)	0.3670 (0.3070)	0.3761 (0.3029)

the distribution of colour plane misalignments before and after compensation for LCA for the calibration pattern in Fig. 5.1. The corresponding Euclidean vector representation of these misalignments for the test image, before and after compensation, are illustrated in Fig. 5.4. These show that the remaining misalignments are random in nature (with magnitude similar to the detection noise), indicating the successful modelling and compensation of LCA.

Model parameter analysis

To determine the suitability of the model, in terms of redundant parameters, and its stability, it is useful to look at the parameter uncertainties and the parameter correlations. These measures may be computed directly from the iterative estimation scheme as described in equations 4.17 and 4.20. In order to concisely describe these results two extreme examples are taken from the calibration data in Section 5.3.1. These are the red/green alignments in Cam 1 and Cam 3. The parameter values and uncertainties are presented in table 5.5. The parameter c_1, related to the lateral colour distortion, takes a large role in the calibration in Cam 1. However, in all estimations its estimated

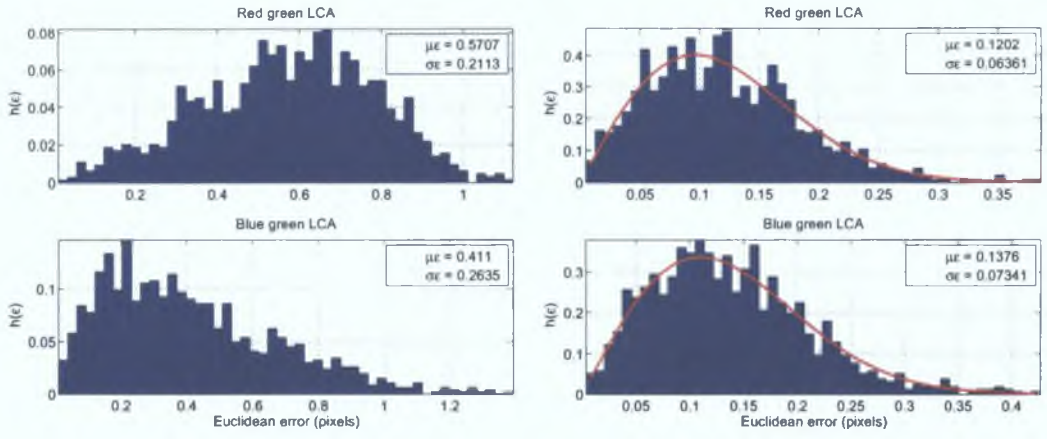


Fig. 5.3: Histograms of Euclidean misalignments computed for chessboard intersections on the calibration image with Cam 1. Left column shows the R/G and B/G differences before compensation, while the right column shows those detected following calibration with fitted Rayleigh PDF's.

uncertainty value remains low indicating a stable system of equations. Parameter c_2, related to the chromatic variation of distortion proves useful for all calibrations and also exhibits low uncertainties. The decentering parameters c_3 and c_4 play a variable role, clearly related to the camera or lens employed. The centre point estimates and uncertainties are presented for completeness.

In lens distortion calibration, it was found that there was a high correlation between the centre point and decentering parameters. It was argued that this relationship does not exist to the same extent when using the chromatic aberration model due to its additional lower order term. This is easily investigated experimentally by looking at the correlations between these parameters. Table 5.6 shows the extreme values of these correlations, where Cam 1 shows a reasonable to strong link between the parameters. From table 5.5 it is noted that there was little use made of the decentering terms so this is not unusual. Cam 3 on the other hand shows complete independence between the decentering and centre point parameters. The correlations are thus more camera related, and it can be assumed that these parameters are independent, unlike the analogous lens distortion case.

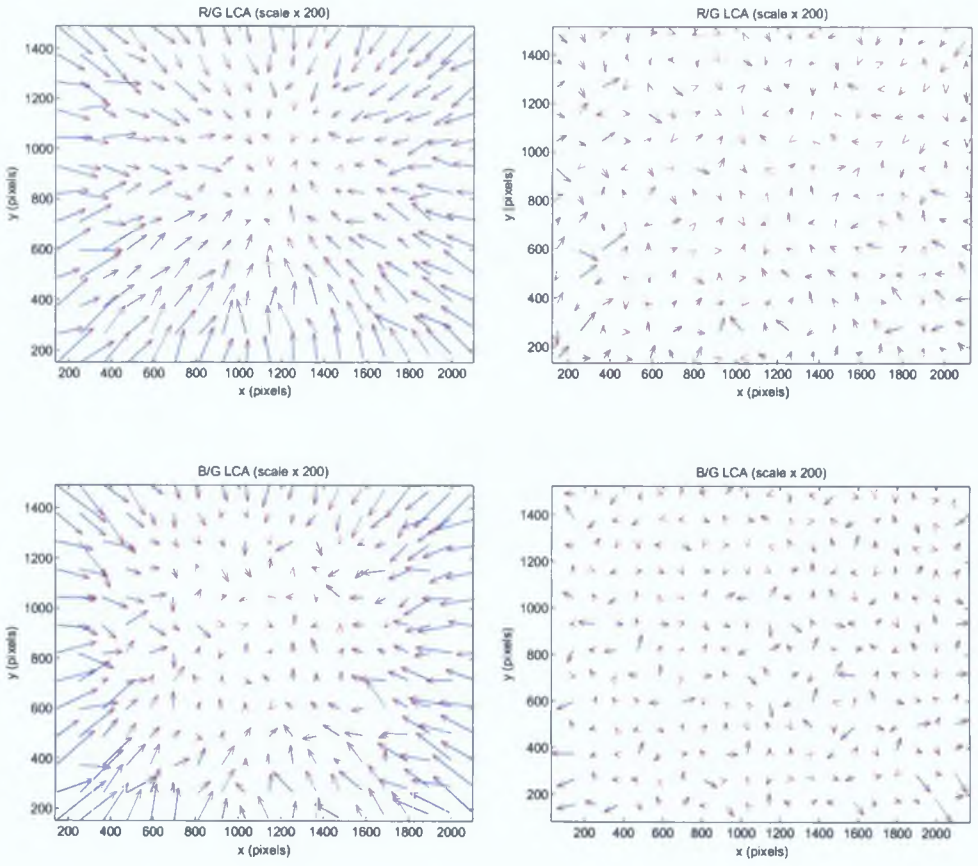


Fig. 5.4: Euclidean vector plots of colour plane misalignments before (left column) and after (right column) LCA compensation, evaluated on the test chessboard pattern with Cam 1.

Table 5.5: Parameter values and uncertainties for the selection of reg/green calibration with Cam 1 and Cam 3

Parameter	Cam 1	Cam 3
$c_1 \times 10^{-3}$	82.730 (0.0094)	2.021 (0.0152)
$c_2 \times 10^{-3}$	-11.164 (0.0142)	-4.970 (0.1023)
$c_3 \times 10^{-4}$	-0.0693 (0.0252)	2.188 (0.0633)
$c_4 \times 10^{-4}$	-0.0716 (0.0349)	1.111 (0.0891)
u_o	1163 (5.33)	1335 (4.13)
v_o	903 (5.38)	1074 (4.18)

Table 5.6: Parameter correlations for decentering and centre point parameters.

	u_o	v_o
c_3	-0.666	-0.012
c_4	-0.009	-0.702
u_o	1	0.004
v_o	0.004	1

(a) Cam 1

	u_o	v_o
c_3	-0.0717	0.012
c_4	-0.008	-0.340
u_o	1	0.0001
v_o	0.0001	1

(b) Cam 3

5.3.2 Examples

To access the improvement in image quality a selection of examples are shown for each camera referred to in tables 5.2 and 5.3. For each camera a region of interest (ROI) is selected in the test image and an outdoor scene image. Images taken with Cam 1 are presented in Figs. 5.5 and 5.6, Cam 2 in Figs. 5.7 and 5.8, while Cam 3 examples are presented in Figs. 5.9 and 5.10. The associated colour histograms for the test ROI's of Figs. 5.5, 5.7 and 5.9 show that for the uncorrected image, two colour paths exist between the black and white squares of the test pattern. This is due to the additional colour fringing introduced by the LCA around regions of high contrast. The colour histograms for the corrected images show that following compensation there is only one colour path between dark and bright squares, indicating the successful removal of LCA from these images. The real examples of Figs. 5.6, 5.8 and 5.10 show a similar behaviour. Additional colour paths can be seen in the original images, while following calibration and compensation these extra colors are removed. These results clearly indicate that the proposed method of automatically calibrating and removing LCA in images leads to a significant increase in image quality.

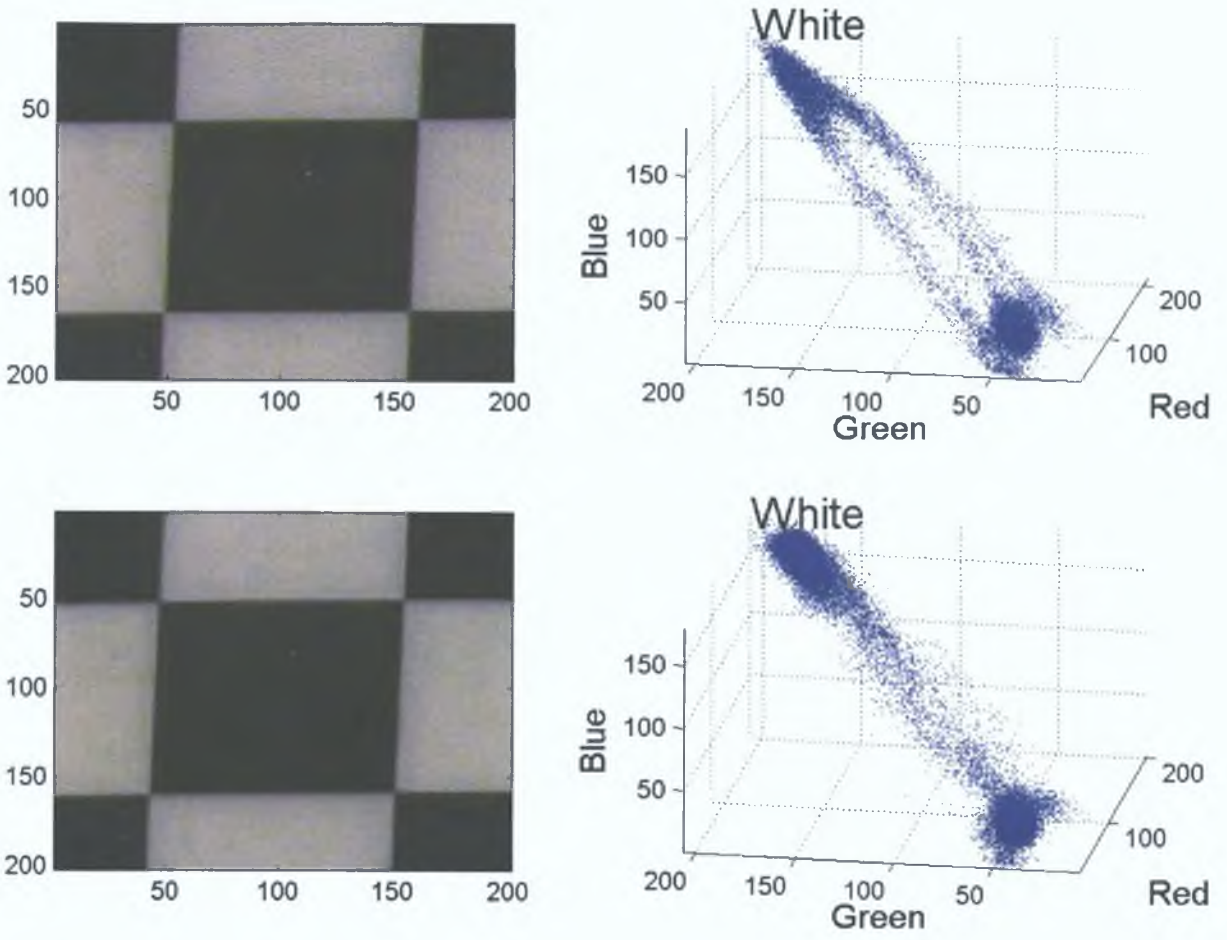


Fig. 5.5: Test image for Cam 1. Top row shows crop with associated colour histogram before compensation. Two additional colour paths are formed by the colour fringing between dark and bright regions. Second row shows the corresponding LCA compensated crop, where the colour fringing has been cancelled.

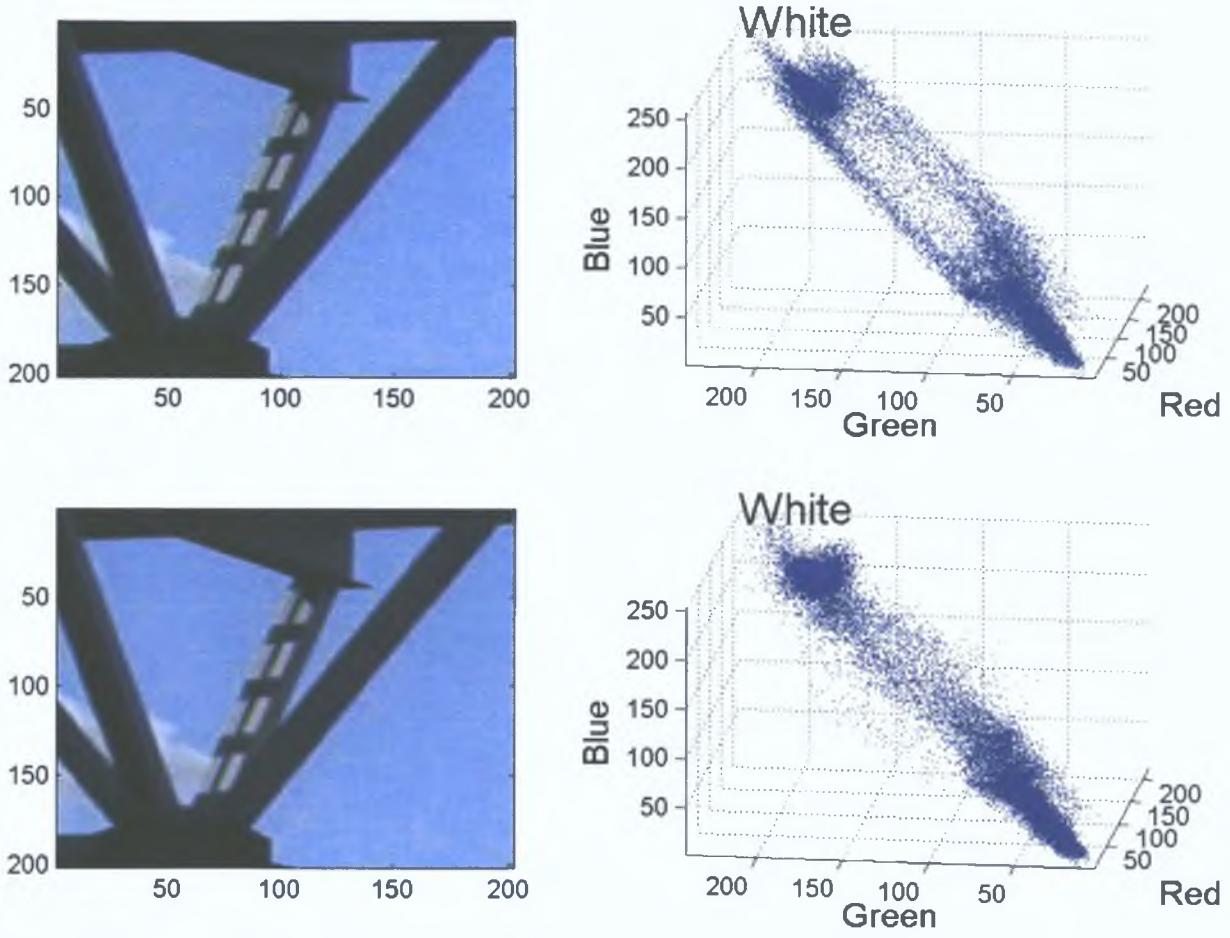


Fig. 5.6: Outdoor image for Cam 1. Top row shows crop with associated colour histogram before compensation. Additional colour paths are formed by the colour fringing. Second row shows the corresponding LCA compensated crop, where the colour fringing has been removed.

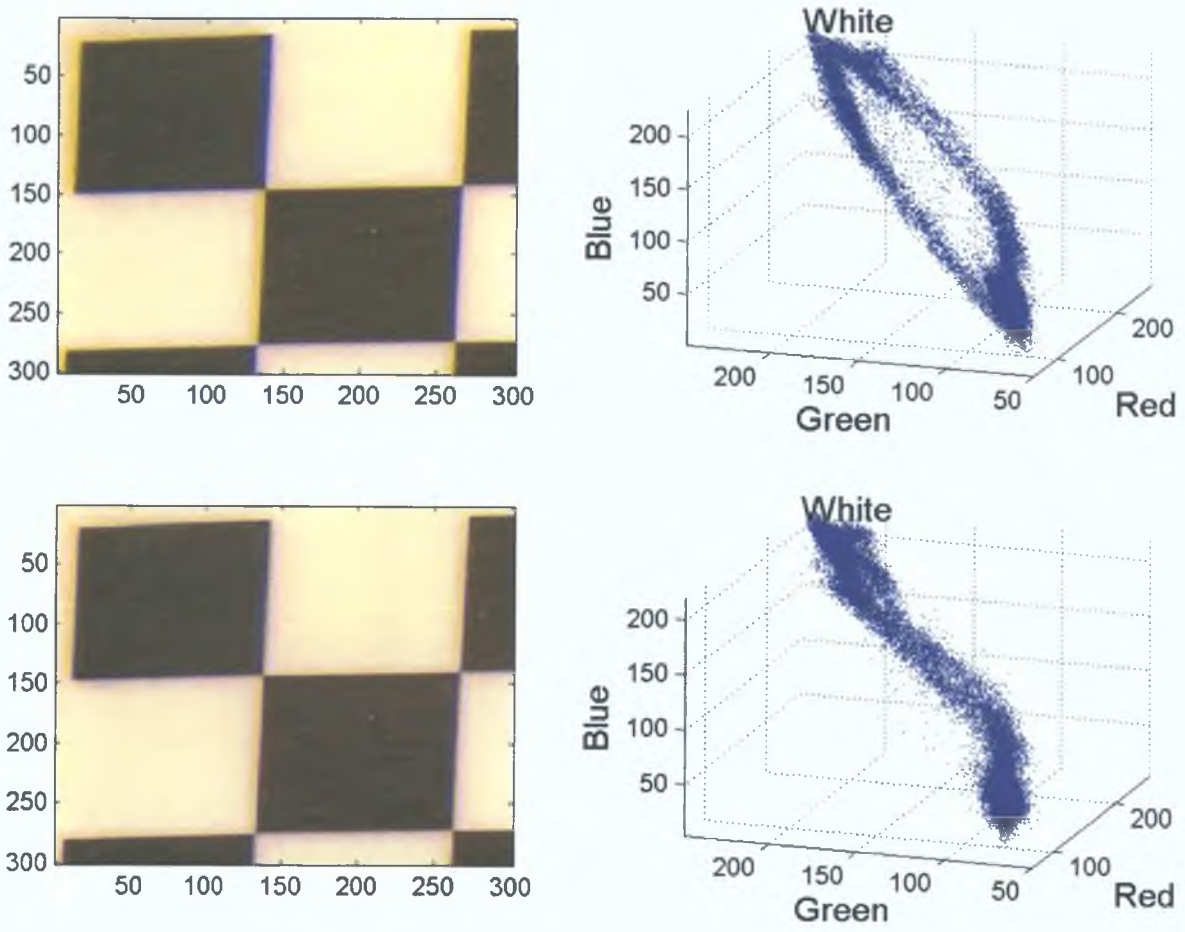


Fig. 5.7: Test image for Cam 2. Top row shows crop with associated colour histogram before compensation. Two additional colour paths are formed by the colour fringing between dark and bright regions. Second row shows the corresponding LCA compensated crop, where the colour fringing has been cancelled.

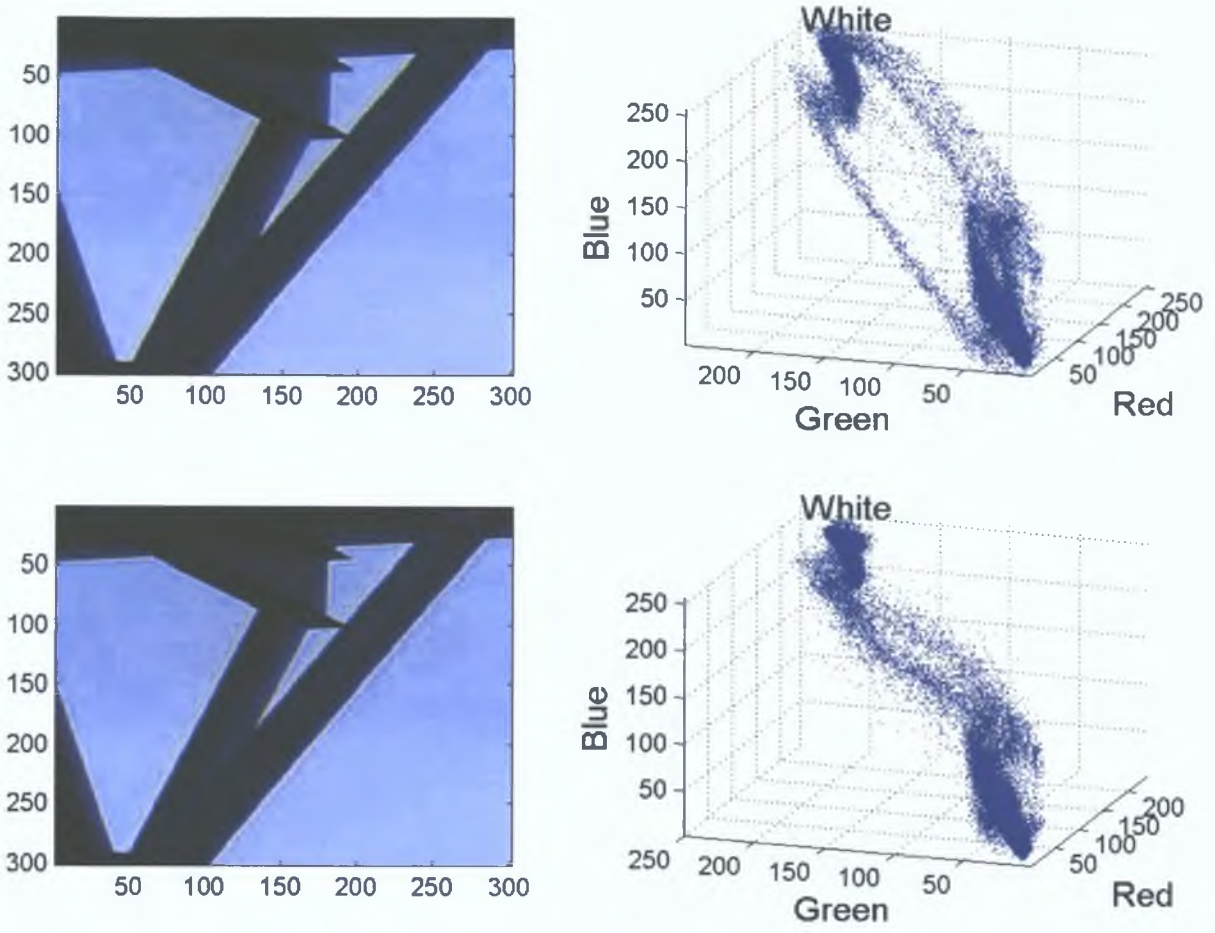


Fig. 5.8: Outdoor image for Cam 2. Top row shows crop with associated colour histogram before compensation. Additional colour paths are formed by the colour fringing. Second row shows the corresponding LCA compensated crop, where the colour fringing has been removed.

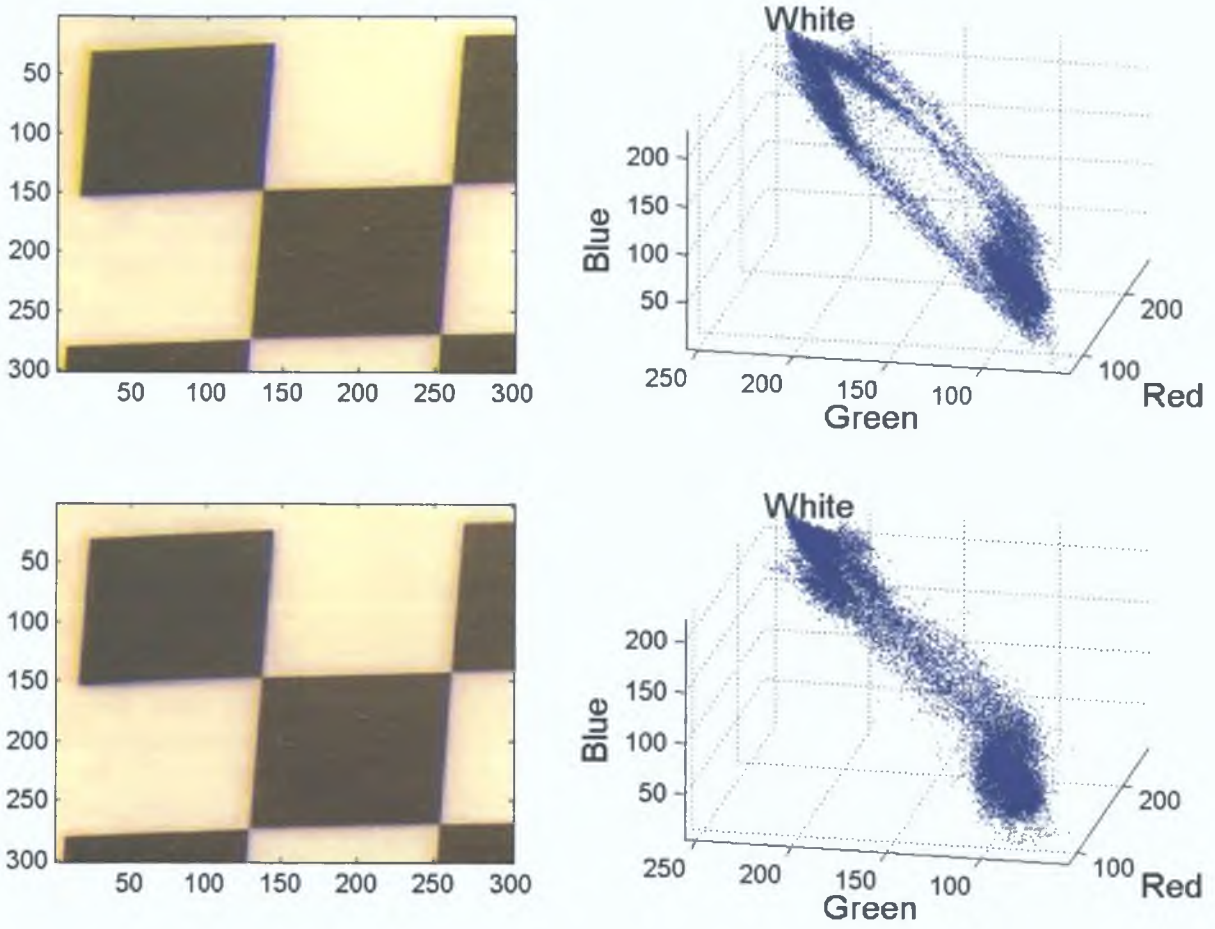


Fig. 5.9: Test image for Cam 3. Top row shows crop with associated colour histogram before compensation. Two additional colour paths are formed by the colour fringing between dark and bright regions. Second row shows the corresponding LCA compensated crop, where the colour fringing has been cancelled.

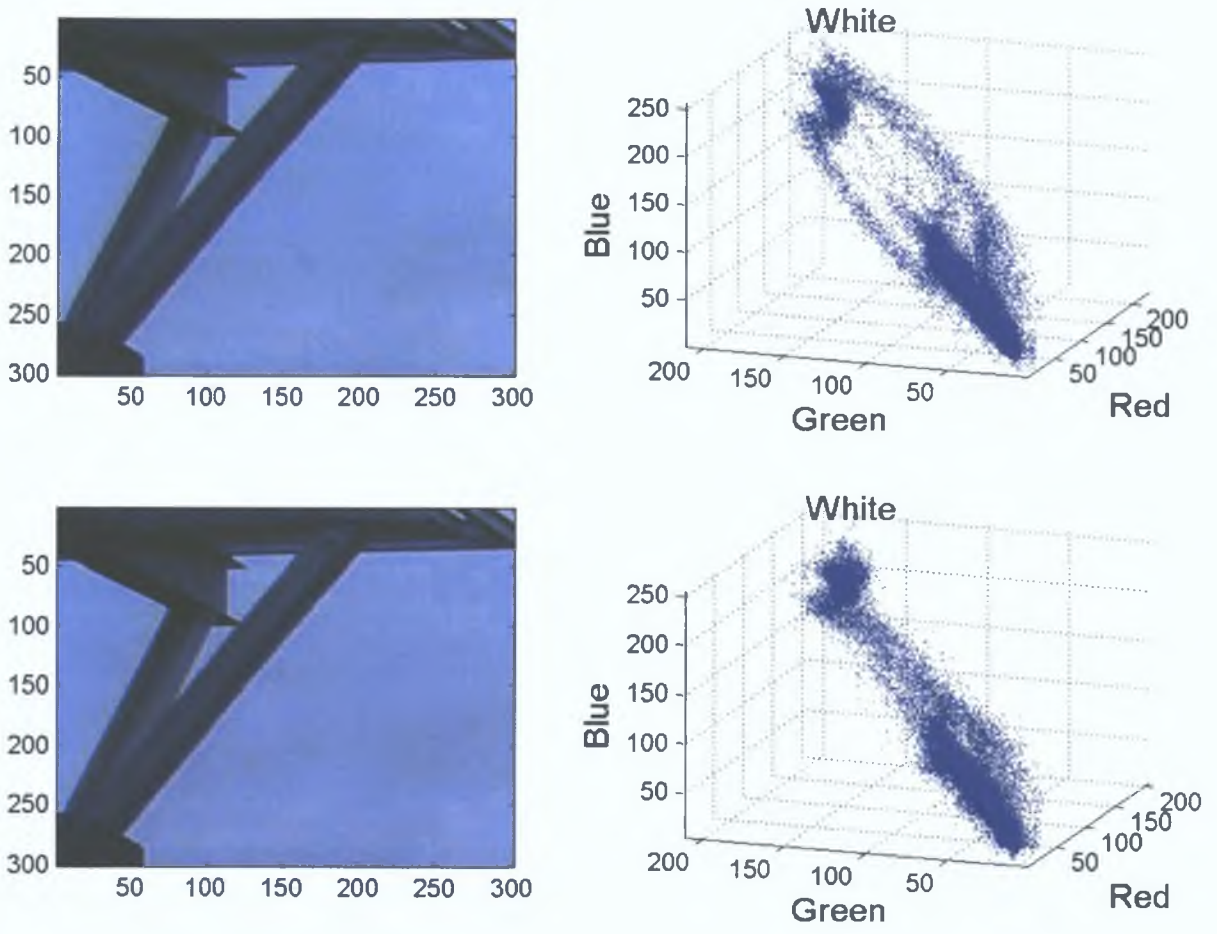


Fig. 5.10: Outdoor image for Cam 3. Top row shows crop with associated colour histogram before compensation. Additional colour paths are formed by the colour fringing. Second row shows the corresponding LCA compensated crop, where the colour fringing has been removed.

5.4 Discussion

This chapter proposes a new model based method of compensating for lateral chromatic aberration in images, offering a usable alternative to active lens control techniques and data interpolation methods. The main contributions are in the derivation of lateral chromatic aberration models and their subsequent parameter estimation techniques. The chromatic calibration technique is easy to use, based on a single view of a chessboard pattern without any strict geometric constraints. This fully automated method is presented in a closed-form allowing faster and simpler estimation. Compensated images are formed by re-sampling the originals based on these calibrated models generating higher quality aberration free images. Model validation is carried out indicating strong global agreement with detected LCA. Examples on selections of real images demonstrate the higher quality achievable with such aberration removal. These show that the additional colours that LCA introduces are removed following compensation with the proposed methods, ultimately giving superior quality colour images.

Chapter 6

Distortion Minimisation in Planar Rectification

Rectification is known to be a useful step in stereoscopic analysis. The aligning of epipolar lines allows subsequent algorithms to take advantage of the epipolar constraint, reducing the search space to one dimension. It is known and easily demonstrated that the rectilinear mechanical alignment of two cameras is prohibitively difficult. The method developed in this chapter aims to simulate rectilinear images from those of arbitrarily placed cameras. This involves subjecting the images to a two dimensional projective transformation or planar homography, while simultaneously minimising projective distortions. This work has been published in Mallon and Whelan (2005).

There are obvious advantages to specifying transformations that minimise the introduction of distortions, or equivalently maximise the similarities between the original image and the transformed one. Recall from Section 4.1.2, that a unique scaling was applied to the new undistorted image in order to globally minimise the local pixel distortions. This idea is now expanded to 2D projective rectifying transformations who's solutions are not fully constrained.

This work follows on from Hartley (1999), where a rigid transformation is derived from the Fundamental Matrix. In this uncalibrated case the resulting depth reconstruction is determined up to a projective transformation (Hartley and Zisserman, 2003). Many applications requiring such relative depth measures exist, including view synthesis (Ng et al., 2002) and robotic navigation (Faugeras and Luong, 2001). The homographies are calculated solely from

an analysis of the Fundamental Matrix, to re-orientate the epipolar projections parallel to the horizontal image axis. Undetermined parameters of the homographies operating on the x coordinate are then specified to maximise viewpoint similarities between the original and rectified images, thus reducing distortional effects of the homographies and improving stereo matching. The rectification is therefore described by a reprojection onto two planes with relative degrees of freedom about the vertical axis only.

The main contribution of this chapter is the proposal of a novel technique to reduce rectification distortions for the maximisation of viewpoint similarities between the original and rectified images. Previous distortion interpretations have included orthogonality of image corners and maximising image content over the view window (Faugeras and Luong, 2001). Loop and Zhang (1999) consider distortion by attempting to force affine qualities on the homographies. As skew and aspect ratio are invariant to affine transforms, they make extra constraints upon the homographies to reduce these types of distortion. Their approach is not optimal as only one local region of the image is considered. The proposed approach in contrast considers all regions of the image, enforcing first order orthogonal qualities in a natural way through Singular Value Decomposition.

This chapter also presents an improved method for the computation of robust matching homographies, from a real Fundamental Matrix estimated from noise affected points. This results in a rectification error equal to that of the Fundamental Matrix error, significantly improving upon the alignment of epipolar lines compared to similar methods such as Hartley (1999), Hartley and Zisserman (2003), Al-Shalfan et al. (2000) and Loop and Zhang (1999).

6.1 Epipolar geometry

Given two images of a scene, let $\mathbf{m}$ and $\mathbf{m}'$ be the projections of some 3D point $\mathbf{M}$ in images $\mathcal{I}$ and $\mathcal{I}'$ respectively. The intrinsic projective geometry between the two views is defined as:

$$\mathbf{m}'^{\mathrm{T}}\mathbf{F}\mathbf{m} = 0, \tag{6.1}$$

where the Fundamental Matrix $\mathbf{F}$ (Faugeras and Luong, 2001, Hartley and

Zisserman, 2003, Armangué and Salvi, 2003) is a 3 x 3 matrix of rank 2. Given at least 8 point matches it is possible to determine the matrix (Hartley, 1997). The Fundamental Matrix maps points in $\mathcal{I}$ to lines in $\mathcal{I}'$, $\mathbf{Fm} = \mathbf{l}'$ upon which corresponding points lie. The image in $\mathcal{I}$ of the camera centre, $\mathbf{c}'$, is termed epipole $\mathbf{e} = (e_u, e_v, 1)^T$ and similarly for $\mathcal{I}$ the image of $\mathbf{c}$ is $\mathbf{e}' = (e'_u, e'_v, 1)^T$:

$$\mathbf{Fe} = 0 = \mathbf{F}^T\mathbf{e}'.$$

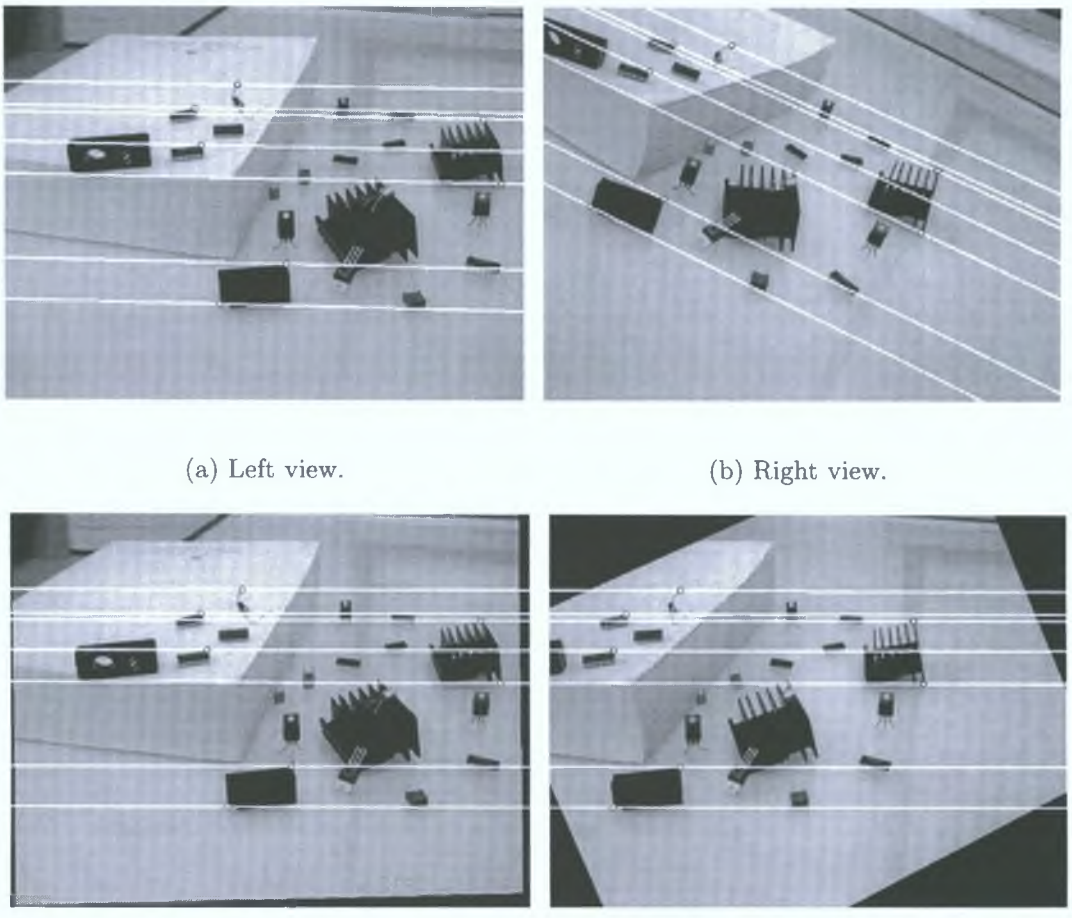
The epipoles $\mathbf{e}$ and $\mathbf{e}'$ can be simply computed from the Singular Value Decomposition of $\mathbf{F} = \mathbf{U}diag(0, \sigma_1, \sigma_2)\mathbf{V}^T$ where $\mathbf{U} = (\mathbf{e}', \mathbf{u_1}, \mathbf{u_2})$, $\mathbf{V} = (\mathbf{e}, \mathbf{v_1}, \mathbf{v_2})$, σ_1 and σ_2 are the typically non-zero singular values. All the epipolar lines in the respective images pass through the epipoles. In this chapter it is assumed that the Fundamental Matrix has been found, which requires at least 8 point matches for linear estimation. In addition the origin of the images is considered to be $(0, 0)$, generally the top left corner.

6.2 Rectification

Image rectification is the process of re-aligning corresponding epipolar lines to become collinear and parallel with the x axis as illustrated in Fig. 6.1. For a stereo sensor, mechanical adjustments of this calibre are difficult to achieve. However, given a description of the projective geometry between the two views, projective transformations can be applied resulting in rectified images. The projective transformations are uniquely chosen to minimise distortions and maintain as accurately as possible the structure of the original images. This helps during subsequent stages, such as matching, ensuring local areas are not unnecessarily warped.

Rectification can be described by a transformation that sends the epipoles to infinity, hence the epipolar lines become parallel with each other. Additionally, it is ensured that corresponding points have the same y coordinate by mapping the epipoles in the direction $\mathbf{e} = (1, 0, 0)^T$ or equivalently $\mathbf{e} = (e_u, 0, 0)^T$. The Fundamental Matrix for such a rectified pair of images is:

$$\bar{\mathbf{F}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$



(a) Left view. (b) Right view.

(c) Rectified left. (d) Rectified right.

Fig. 6.1: Example of the rectification procedure. The original images are shown in 6.1(a) and 6.1(b) overlaid with their respective epipolar lines. After rectification these lines become collinear and parallel with the image x axis, as shown 6.1(c) and 6.1(d).

The desired homographies give new image coordinates as $\bar{\mathbf{m}} = \mathbf{H}\mathbf{m}$ and $\bar{\mathbf{m}}' = \mathbf{H}'\mathbf{m}'$. It follows from equation (6.1) that $\bar{\mathbf{m}}'^T\bar{\mathbf{F}}\bar{\mathbf{m}} = 0$ and $\mathbf{m}'^{\mathrm{T}}\mathbf{H}'^{\mathrm{T}}\bar{\mathbf{F}}\mathbf{H}\mathbf{m} = 0$ resulting in a set of constraints relating $\mathbf{H}$ to $\mathbf{H}'$:

$$\mathbf{H}'^{\mathrm{T}}\bar{\mathbf{F}}\mathbf{H} = \mathbf{F}. \tag{6.2}$$

The homographies satisfying equation (6.2) are not unique, but similarly to Hartley Hartley (1999) $\mathbf{H}$ is chosen to transform the epipole $\mathbf{e}$ to infinity:

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 \\ -e_v/e_u & 1 & 0 \\ -1/e_u & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ h_{21} & 1 & 0 \\ h_{31} & 0 & 1 \end{pmatrix}. \tag{6.3}$$

The determinant of the Jacobian, $\det(\mathbf{H}) = 1/(1 - x/e_u)$, gives an indication

of the changes or warping of local areas. At the origin the transformation appears orthogonal ($\det(\mathbf{H}) = 1$), while in general e_u is large in comparison to the image size. This ensures $\mathbf{H}$ does not cause severe perspective distortion.

6.2.1 Matching Homography

Considering equation (6.2), it is clear that for an ideal Fundamental Matrix there are no applicable constraints on the first row of $\mathbf{H}'$. Thus the matching transformation $\mathbf{H}'$ is specified with the form:

$$\mathbf{H}' = \begin{pmatrix} 1 & 0 & 0 \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & h'_{33} \end{pmatrix}.$$

Evaluating equation (6.2), it is now proposed to estimate the entities of $\mathbf{H}'$ by the elementary comparison of entries in equation (6.4), where α represents the (optional) arbitrary scale difference. The constraints on $\mathbf{H}'$ are:

$$\begin{pmatrix} (h_{21}h'_{31} - h_{31}h'_{21}) & h'_{31} & -h'_{21} \\ (h_{21}h'_{32} - h_{31}h'_{22}) & h'_{32} & -h'_{22} \\ (h_{21}h'_{33} - h_{31}h'_{23}) & h'_{33} & -h'_{23} \end{pmatrix} = \alpha \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}. \tag{6.4}$$

Assuming an imperfect $\mathbf{F}$ matrix[1], the solution for $\mathbf{H}'$ can be robustly found in a least squares sense from equation (6.4) by the SVD of $\mathbf{B}\widehat{\mathbf{p}} = \mathbf{0}$, where $\widehat{\mathbf{p}} = (h'_{21}, h'_{22}, h'_{23}, h'_{31}, h'_{32}, h'_{33}, \alpha)^T$. Computing $\mathbf{H}'$ in a least square sense using all the entries of $\mathbf{F}$ significantly improves the rectification accuracy in contrast with Hartley (1999) solution $\mathbf{H}' = \mathbf{H}([\mathbf{e}]_\times \mathbf{F} + \mathbf{e}\mathbf{e}'^T)$, and others (see Section 6.4).

6.3 Reducing Rectification Distortions

The application of $\mathbf{H}$ and $\mathbf{H}'$ does indeed rectify the images as required. However, as can be noted above, the first rows of the homographies are undeter-

[1] F matrix computed with a limited set of noisy points.

mined. This results naturally from the Fundamental Matrix, which does not encapsulate any information about the position of the x coordinate. Weng et al. (1993) shows that only one component of the image position of a point is used by the epipolar constraint. The projective transformations in general introduces distortions in the rectified images, specifically skewness and aspect/scale distortions. However, it is possible to specify the first rows of both homographies without invalidating the constraints used to compute them, $\mathbf{AHe} = \mathbf{He}$, and similarly for the primed counterparts giving:

$$\mathbf{H'^T A'^T \bar{F} A H} = \mathbf{K'^T \bar{F} K} = \mathbf{F}, \tag{6.5}$$

where $\mathbf{K} = \mathbf{AH}$. $\mathbf{A}$ and $\mathbf{A}'$ are transformations of affine form:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The creation and loss of pixels as a result of the application of transformation $\mathbf{K}$, can be quantified in the local area of point $\mathbf{p}$ by any norm of the Jacobian:

$$\mathbf{J}(\mathbf{K}, \mathbf{p}) = \begin{pmatrix} \frac{\partial \bar{x}}{\partial x} & \frac{\partial \bar{x}}{\partial y} \\ \frac{\partial \bar{y}}{\partial x} & \frac{\partial \bar{y}}{\partial y} \end{pmatrix}.$$

Let $\sigma_1(\mathbf{J})$ and $\sigma_2(\mathbf{J})$ be the non zero singular values of $\mathbf{J}$ in descending order. Ideally, an orthogonal transform that neither creates or destroys pixels will have singular values equal to one. In general $\sigma_1(\mathbf{J}) > 1$ for a transformation that overall creates extra pixels, and $\sigma_1(\mathbf{J}) < 1$ for an overall compression of pixels within a local region.

The search for the best compromise of the affine pair a_{11} and a_{12} to maintain orthogonality and perspective of the original image can thus be expressed by searching for the singular values that are closest to one. The Wielandt-Hoffman theorem (Golub and Loan, 1996) for singular values states that if $\mathbf{A}$ and $\mathbf{E}$ are matrices in $\mathbb{R}^{m \times n}$ with $m \geq n$, then:

$$\sum_{k=1}^{n} (\sigma_k(\mathbf{A} + \mathbf{E}) - \sigma_k(\mathbf{A}))^2 \leq \|\mathbf{E}\|_F^2,$$

where $\|\mathbf{E}\|_F$ is the Frobenius norm of $\mathbf{E}$. This indicates that if $\mathbf{A}$ is perturbed by $\mathbf{E}$, the corresponding perturbation in any singular value of $\mathbf{A}$ will be less than that of the Frobenius magnitude of $\mathbf{E}$. This means that the relationship between entries in a matrix and its singular values is a smooth function, making them very suitable for iterative search techniques.

6.3.1 Minimisation

The search is conducted by evaluating the singular values of the Jacobian at various points over the image. These points, $\mathbf{p}_i$ can be simply specified as a grid covering the image area or as the corners of the image. The function to be minimised is then expressed as:

$$f(\hat{a}_{11}, \hat{a}_{12}) = \sum_{i=1}^{n} [(\sigma_1(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2 + (\sigma_2(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2]. \tag{6.6}$$

This functional is minimised using the Nelder and Mead simplex search method which converges, on average, after 50 iterations. Finite derivative methods can also be applied as the function inherently has smooth derivatives. Since a_{13} is an x direction shift it does not introduce any distortion. It can be chosen automatically to centre the rectified image in the old image window if desired. The same procedures equally apply to the primed counterpart image. The rectification is determined solely on the estimate of the $\mathbf{F}$ matrix. This has the advantage that no point correspondences are explicitly needed. Thus the rectification is invariant to the location or quantity of the point set, which overall tends towards a more consistent result.

6.4 Experiments

A selection of nine real examples are presented [2]. The performance of the proposed rectification is quantified using various metrics, and compared side-by-side with two popular methods from the literature, Hartley (1999), Hartley and Zisserman (2003) and Loop and Zhang (1999). The examples feature a range of Fundamental Matrix accuracy levels, and requiring various transformation complexity. The images were taken with a digital camera with 640×480 pixel resolution, over random unknown baselines. The lens parameters remained unchanged throughout and lens distortion has been removed using the method outlined in chapter 4. The images are of indoor and outdoor scenes with relatively low and high depth of scene respectively. The Fundamental Matrix was calculated using the linear normalised eight point method (Hartley, 1997)

[2] These real images are used as opposed to those in the literature for two reasons. Firstly, the availability of suitable data sets is limited, and secondly the relative simple geometry of some of these examples.

using manually matched points. The examples are available at the VSG code archive web page, `http://www.eeng.dcu.ie/~vsl/vsgcode.html` including data.

6.4.1 Error Metrics

The rectification technique is based solely on the estimation of the Fundamental Matrix. Therefore, and according to equation (6.5), the rectification performance is directly related with the integrity of the Fundamental Matrix. A direct evaluation of the accuracy of the Fundamental Matrix is given by the perpendicular distance from a point to its epipolar line. Considering the corresponding points $\mathbf{m} = (u, v, 1)^T$ and $\mathbf{m}' = (u', v', 1)^T$, the epipolar line in $\mathcal{I}$ is given by $\mathbf{l} = \mathbf{F}^\mathbf{T}\mathbf{m}' = (l_a, l_b, l_c)^T$. The perpendicular line through $\mathbf{m}$ is: $\mathbf{l}^\perp = (l_b, -l_a, (l_a v - l_b u))^T$ and the intersection point: $\mathbf{p}^\perp = \mathbf{l} \wedge \mathbf{l}^\perp = (u^\perp, v^\perp, 1)^T$, where $\wedge$ is the cross product. The Fundamental Matrix error is the distance $E_f = ((u^\perp - u)^2 + (v^\perp - v)^2)^{\frac{1}{2}}$. The rectification precision is then evaluated as: $E_r = \|(\mathbf{Km})_2^T - (\mathbf{K}'\mathbf{m}')_2^T\|$.

In general, it is not possible to avoid all distortions in a perspective transformation. Distortions in this case are defined as departures from the original image structure, such as skewness and relative scale changes. These factors can be quantified by measuring the proportional sizes and orthogonality of the transformed images. Thus $a = (w/2, 0, 1)^T$, $b = (w, h/2, 1)^T$, $c = (w/2, h, 1)^T$ and $d = (0, h/2, 1)^T$ are defined as four cross points of the image where w and h are the image width and height. These points are then transformed, whereupon the vectors $\bar{x} = \bar{b} - \bar{d}$ and $\bar{y} = \bar{c} - \bar{a}$ are formed. The orthogonality is then given as the angle of the upper left corner between the $\bar{x}$ and $\bar{y}$ vectors (ideally 90^o). This angle is: $E_o = \cos^{-1}\left(\frac{\bar{x}\cdot\bar{y}}{|x||y|}\right)$. The aspect ratio can be used to measure the relative size of the image. Taking the corner to corner length ratio by redefining the points $a = (0, 0, 1)^T$, $b = (w, 0, 1)^T$, $c = (w, h, 1)^T$ and $d = (0, h, 1)^T$, the vectors $\bar{x}$ and $\bar{y}$ are formed as before. The aspect ratio (ideally unity) is then defined as: $E_a = \left(\frac{\bar{x}^T\bar{x}}{\bar{y}^T\bar{y}}\right)^{\frac{1}{2}}$.

Table 6.1: Presents the error metrics described in Section 6.4.1 for nine different real world samples. The errors are presented in mean (standard deviation) format where applicable. Orthogonality is ideally 90^o while the ideal aspect ratio is 1. The results are compared with two alternative techniques of Hartley (1999), Hartley and Zisserman (2003) andLoop and Zhang (1999), (nc) = no convergence. A selection of examples (Boxes, Roof, Yard and Drive) are shown in figures 6.2 6.3 6.4 and 6.5.

Sample	F Mat. E_f Mean (SD)	Method	Orthogonality E_o		Aspect Ratio E_a		Rectification E_r	
			H′	H	H′	H	Mean	SD
Lab	4.861 (2.732)	Proposed	89.87	89.01	0.9960	0.9835	4.45	2.41
		Loop (nc)	-	-	-	-	-	-
		Hartley	99.07	96.56	1.1635	1.1111	23.02	4.53
Boxes*	0.5068 (0.3630)	Proposed	88.78	89.33	0.9785	0.9889	0.44	0.33
		Loop	97.77	95.69	1.1279	1.0900	4.35	9.20
		Hartley	86.56	94.99	0.9412	1.0846	33.36	8.65
Slates	0.5987 (0.4593)	Proposed	89.12	89.13	0.9852	0.9855	0.59	0.56
		Loop	37.29	37.15	0.2698	0.2805	1.14	3.84
		Hartley	89.96	88.54	1.0000	0.9769	2.27	5.18
Junk	0.3437 (0.2832)	Proposed	90.78	91.62	1.0233	1.0274	0.11	0.32
		Loop (nc)	-	-	-	-	-	-
		Hartley	102.67	99.60	1.3074	1.2466	14.38	8.51
Hall	1.9829 (1.2124)	Proposed	90.00	90.03	1.0003	1.0006	1.81	1.39
		Loop	91.14	91.58	1.0194	1.0271	4.92	2.40
		Hartley	102.56	90.48	1.2353	1.0081	2.59	2.75
Roof*	1.6422 (1.7085)	Proposed	88.35	88.23	1.1077	0.9700	1.96	2.95
		Loop	69.28	87.70	0.6665	1.0497	0.84	11.01
		Hartley	122.77	80.89	1.5256	0.8552	11.89	18.15
Arch	0.3244 (0.3123)	Proposed	91.22	90.26	1.0175	1.0045	0.22	0.33
		Loop	95.40	98.94	1.0991	1.1662	131.3	20.63
		Hartley	100.74	93.05	1.2077	1.0546	39.21	13.85
Yard*	0.6365 (0.4776)	Proposed	89.91	90.26	0.9987	1.0045	0.53	0.54
		Loop	133.62	134.27	2.1477	2.4045	8.91	13.19
		Hartley	101.95	91.91	1.2303	1.0335	48.19	11.49
Drive*	0.5684 (0.7568)	Proposed	90.44	90.12	1.0060	1.0021	0.18	0.91
		Loop	98.73	101.42	1.1541	1.2052	10.41	3.24
		Hartley	107.66	90.87	1.3491	1.015	3.57	3.43

6.4.2 Rectification Performance

The rectification performance is concerned with quantifying only the y or row alignment of corresponding points over the images. Referring to table 6.1, of interest is the Fundamental Matrix error E_f, which is sufficiently characterised by its mean and standard deviation. As the rectification is only based on the Fundamental Matrix, its error E_f represents the minimum expected rectification error E_r. The Fundamental error E_f ranges from 4.8(2.7) to 0.32(0.31) pixels in the examples. This is matched in every instance by the method proposed here, where E_r ranges from 4.5(2.4) to 0.11(0.32).

In comparison, the Hartley (1999) technique fails to match any of these results[3], with E_r ranging from 48.2(11.4) to 2.6(2.7) (see table 6.1). It stems from a lack of robustness in the method used to match the homographies. In comparison to Loop and Zhang (1999) methods, for some cases this rectification did not converge (nc), while convergence is questionable for the *Arch* example. As this technique is scale invariant, suitable scaling for each example was chosen manually. Disregarding the non-convergence cases the rectification error E_r ranges from 10.4(3.2) to 1.4(3.8). The poor alignment for these alternative techniques can additionally be seen in the figures 6.2, 6.3, 6.4, and 6.5. It can be seen that the rectification process described here significantly outperforms the comparison techniques. Its convergence to the minimum error E_f in every case demonstrates good robustness.

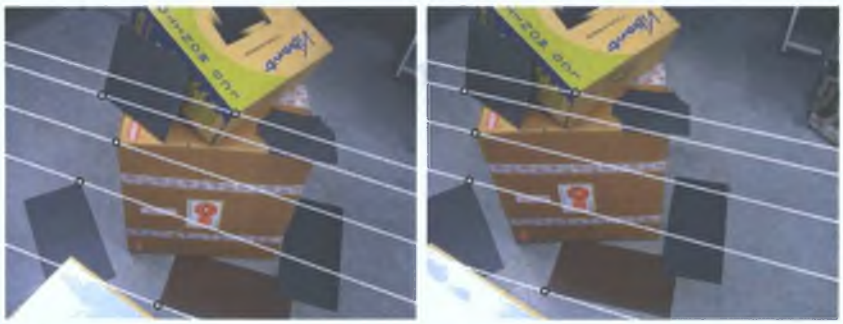
[3]Theoretically Hartleys method guarantees horizontal epipolar lines. From an examination of the experiments in Figs. 6.2 to 6.5 it can be concluded that the right image satisfies this constraint, but not the left image. The left image is rectified by the matching transform $\mathbf{H}'$, which is computed in Hartley (1999) according to $\mathbf{H}' = (\mathbf{I} + \mathbf{H}\mathbf{e}\mathbf{a}^T)\mathbf{H}\mathbf{M}$. This simplifies to $\mathbf{H}' = \mathbf{H}\mathbf{M}$ if $\mathbf{a}^T = (1, 0, 0)$. From Result A4.4 and 9.14 in Hartley and Zisserman (2003), $\mathbf{M}$ is defined as $\mathbf{M} = [\mathbf{e}]_\times \mathbf{F}$ or equivalently $\mathbf{M} = [\mathbf{e}]_\times \mathbf{F} + \mathbf{e}\mathbf{v}^T$ for any vector $\mathbf{v}$. However, the constraint on $\mathbf{M}$ that $\mathbf{M}\mathbf{e}' = \mathbf{e}$ is not well satisfied in practice. This may be due to the normal epipole inaccuracies, i.e. that $\mathbf{F}\mathbf{e} = 0$ is generally not exactly 0. The specific reason as to why the matching homography $\mathbf{H}$ does not entirely match (thus yielding horizontal epipolar lines) is unclear, though it must be as a result of the effects of these aforementioned inaccuracies.

6.4.3 Distortion Reduction

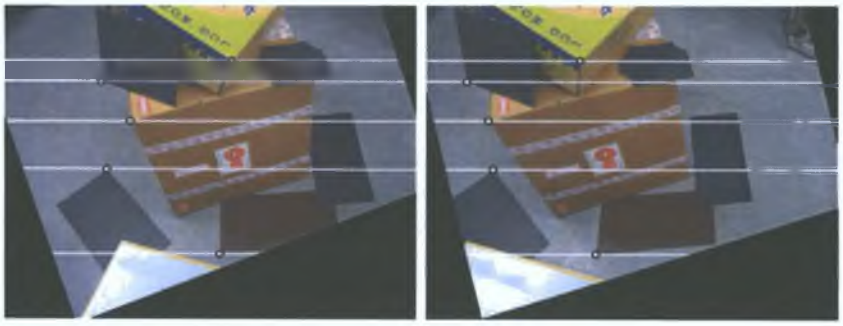
The rectification process introduces a necessary distortion to the images that realigns them horizontally relative to each other. It is possible to specify the x position of pixels with an affine shearing transform that leaves the rectification unaffected. To this end the aim is to preserve, as much as possible, the original viewpoint of each camera in order to avoid introducing unnecessary distortions. The distortion reduction criteria of equation (6.6) therefore strives to minimise local pixel warping throughout the image. Table 6.1 and figures 6.2, 6.3, 6.4, and 6.5 show the performance on the set of test images. The orthogonality E_o and aspect ratio E_a are of interest and are calculated for both homographies. Orthogonality gives a intuitive measure of the distortion level. Taking the examples in table 6.1, the average absolute orthogonal angle error for the proposed method is 0.8^o for both $\mathbf{H}$ and $\mathbf{H}'$.

In comparison Hartley (1999), Hartley and Zisserman (2003) method introduces a significant quantity of distortion. This is expected as distortion is not considered, except to minimise disparity. As the results show, this can introduce serious warping. The average absolute orthogonal angle error for $\mathbf{H}$ and $\mathbf{H}'$ from table 6.1 is 4.2^o and 12^o. Loop and Zhang's methods Loop and Zhang (1999) explicitly consider distortion, defined in a similar way to the proposed method, by preserving perpendicularity and aspect ratio. The results in table 6.1 show that their method at no point matches our results. The average absolute orthogonal angle error for this method is 18^o and 20^o for $\mathbf{H}$ and $\mathbf{H}'$. This is because their criteria is defined only for the midpoint of the image. An optimal estimate for one point does not mean it will be optimal for all image points, and indeed this is the case. In contrast, our method considers the local areas over the entire image to preserve orthogonality giving superior results.

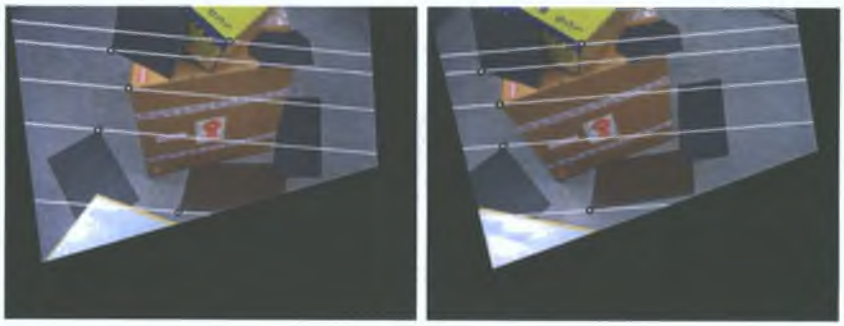
The rectification detailed above is based on the application of planar 2D projective transformations. As evident from Section 6.2, these techniques are not applicable for configurations where the epipole is within an image. In this case it is possible to cause severe image distortions, even splitting connected regions by their application. In general, the geometry that results in such undesirable situations, such as forward translation, is not generally encountered in a stereo-like setup.



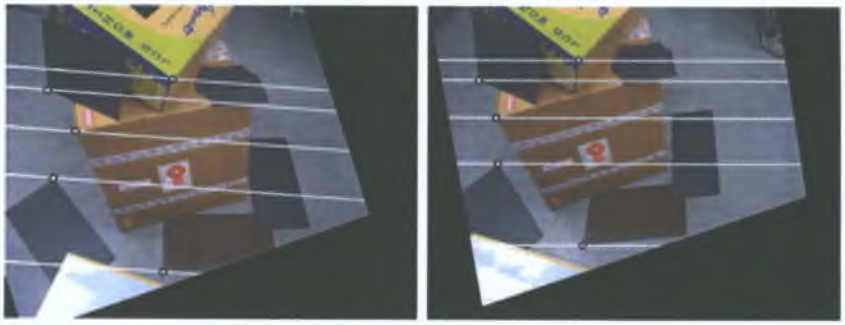
(a) Originals for Boxes example, Left and Right



(b) Proposed Rectification Method



(c) Loop and Zhang (1999) Method



(d) Hartley (1999) Method

Fig. 6.2: Boxes example including epipolar lines, see table 6.1 for more details.

(a) Originals for Roof example, Left and Right

(b) Proposed Rectification Method

(c) Loop and Zhang's Method

(d) Hartley's Method

Fig. 6.3: Roof example including epipolar lines, see table 6.1 for more details.

(a) Originals for Yard example, Left and Right

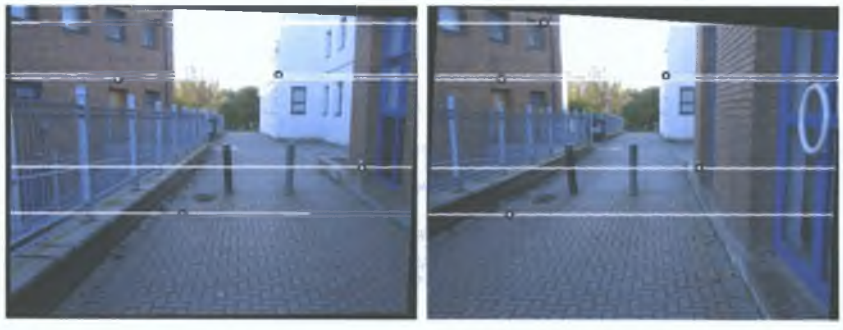
(b) Proposed Rectification Method

(c) Loop and Zhang's Method

(d) Hartley's Method

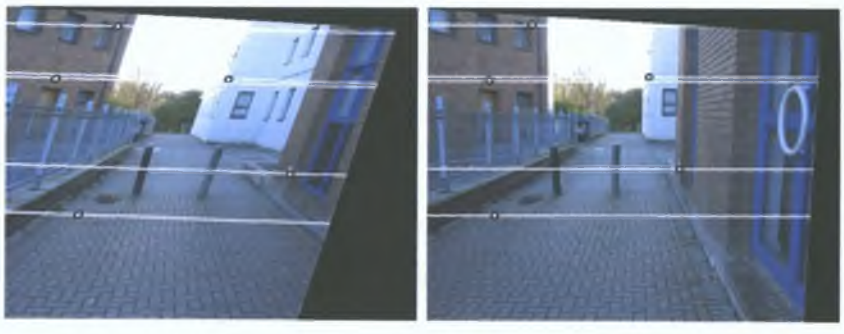
Fig. 6.4: Yard example including epipolar lines, see table 6.1 for more details.

(a) Originals for Drive example, Left and Right



(b) Proposed Rectification Method

(c) Loop and Zhang's Method



(d) Hartley's Method

Fig. 6.5: Drive example including epipolar lines, see table 6.1 for more details.

6.5 Discussion

This chapter describes a robust method for uncalibrated planar rectification for a pair of stereo images taken from distinct viewpoints. It is simple to implement and based solely on the estimated Fundamental Matrix. A much improved method is given for the computation of matching perspective transformations, with experimental results showing that the rectification accuracy, or epipolar alignment, is equal to the error in the Fundamental Matrix estimation. A novel technique is described to reduce the inevitable perspective distortions. This ensures that the rectified images resemble the originals as closely as possible, virtually eliminating unnatural skews and scaling. This has clear advantages for subsequent processing steps. The distortion minimisation is carried out by searching through a SVD for the best first order approximation of an orthogonal-like transform throughout the image window. Detailed comparison results clearly indicate much improved performance for both the rectification process and the distortion reduction techniques.

Chapter 7

Conclusion and Future Work

A typical CCD sensor easily produces well in excess of over one million 8-bit measurements every time an image is acquired, while in comparison, the required quantity of information is typically much less. However, depending on the application, this information is generally encoded throughout the image array and often over multiple images. On top of this decoding problem, extra complexity is added by the fact that most optical systems naturally distort the imaged geometry of the scene. This means that an objects geometry is warped depending on the viewing location or position within an image.

Two types of such errors are considered in this thesis, lens distortion and chromatic aberration. The compensation of lateral chromatic aberration, through the realignment of the image colour planes, may influence some specific high precision colour applications, but currently its main practical usage pertains to image enhancement in digital photography. Lens distortion removal or compensation is currently used extensively in the many applications requiring the measurement of pixel or object locations. The introduction of a method for the sole calibration of distortion, as outlined in this thesis, potentially allows much greater access to distortion calibration, without encountering the overheads associated with more traditional calibration methods. The method also incorporates a unique means of balancing the unavoidable compression and expansion effects following the pixel relocations. This is mindful of the overall end applications for which the images may be used. Naturally, good image quality is universally beneficial. This idea is not limited to one specific case, and is further extended to minimising projective distortions in a planar rec-

tification application. Again this considers the bigger picture including the further applications, in this case typically stereo which depends on the similarities between local areas over multiple images.

It can be seen from the literature that accurate, but easy to use, techniques prove more popular than equally accurate but complex methods, and indeed the lack of such methods was part of the original motivation behind this thesis. The calibration of optical and further projective distortions is thus carried out within a framework of low complexity, general applicability and high accuracy. This chapter presents an outline of the findings and contributions of this thesis. A list of the publications which have resulted from this work is also included, as well as some auxiliary work and presentations. Finally, a brief discussion is included, highlighting some specific extensions and directions for future research in this area.

7.1 Summary of contributions

The main findings and contributions resulting from this thesis are identified and summarised below.

7.1.1 Aberration modelling

Previous works in lens distortion calibration give little reason for the choice of distortion models. Consequently, a degree of confusion has arisen in some cases, as to the appropriate model usage and the benefits thereof. Chapter 2 aims to redress this situation, wherein the origins and form of the distortion models are derived, and the relationships between different alternative interpretations are highlighted. The benefits of each model are compared on real data in terms of accuracy and stability. The result is that the forward model offers the most general applicability and suitability for distortion estimation schemes. The alternative reverse model shows greater instability, ultimately manifesting itself with increased sensitivity in calibration, such as sensitivity to a change or error in the input data. This shows that there are significant advantages to be obtained with the use of the appropriate model.

The origins and performance of alternative models are also considered, includ-

ing the popular divisional model. This model is shown to perform well for fish-eye type distortion, but cannot capture the nonlinearities often encountered in perspective camera lenses. A high accuracy linear inverse approximation to the forward model is also described, as outlined in Mallon and Whelan (2004). Lastly, this chapter proposes a parametric model for modelling lateral chromatic aberration in images.

7.1.2 Calibration patterns

Chapter 3 also addresses another neglected, but important aspect in camera calibration, that of generating control points. Two popular types of planar patterns, circles and chessboards, are considered to determine if the choice of pattern and in turn the detection method, influences the precision of the detected control points. It is clearly demonstrated that chessboard patterns are superior to circular type patterns in the generation of bias free control points.

The respective accuracies are examined with regard to perspective transformations and lens distortion. For each pattern two types of detection methods are compared: centroids, conic fitting, edge approximation and corner points. It is shown theoretically and experimentally that compensated conic fitting, edge approximation and corner points are invariant to perspective bias, while only corner points are truly invariant to distortion bias. Simulated and real results indicate that distortion induced bias has a significant magnitude. Even for low distortion levels the biasing influence of distortion is likely to be greater than the noise/blur floor, and is more significant than the likely perspective bias encountered with normal calibration views. Thus, the current compensations for perspective bias only, are clearly not sufficient to acquire bias free control points. Ultimately, it is demonstrated that the choice of pattern and detection method will significantly impact on the overall accuracy of any control point based calibration procedure.

7.1.3 Calibration of lens distortion

Using this high accuracy data, high accuracy distortion calibration is addressed in Chapter 4. A non-metric technique (no other camera parameters are directly

available) using a single view input is proposed. It uses the forward distortion model in a closed-form least squares solution to accurately calibrate lens distortion of all levels. Additionally, a method to linearly estimate a transform to minimise the introduction of distortions in the re-sampling of distortion free images is also introduced. The method is straightforward to implement and use and offers a viable alternative to current distortion calibration solutions.

An exhaustive set of experiments are conducted on the proposed method with real and simulated data. These are conducted to fully characterise the method and position itself with regard to potential application areas. In comparisons with existing non-metric and full calibration methods, the proposed method at least matches, and often surpasses, the best of these methods. The method is shown to be stable for very severe distortion, right down to sub-pixel levels. The dependence of the proposed method on the planarity and accuracy of the calibration chart is investigated. This shows that the relationship between the calibration accuracy and errors in the calibration chart are roughly proportional. Naturally, the rate depends on the pattern size and the camera resolution. As an example, for a 6MegaPixel resolution camera and a 250×170 mm pattern, the induced error from 1 mm imprecisions in planarity or pattern precision is roughly 1 pixel. It is shown that the use of bias corrupted control points from a circular pattern induces a similar level of error. For successful calibration, it is recommended that at least 200 control points are present. Multiple input views may be used to attain this quantity, but they offer no further advantage. The variable lens parameters which violate the distortion calibration are also investigated. The extent of the distortion variations with the basic lens variables of focusing and aperture setting are shown. Additionally, the modelling of decentering distortion by a variable projective transform, through the radial distortion equation, is shown to surpass previous approximations and naturally accommodates this type of distortion as required.

7.1.4 Compensation for lateral chromatic aberrations

Chapter 5 proposes a model based method for compensating lateral chromatic aberration in colour images. This compares with active lens control and rough image interpolation methods, offering a much more usable alternative. It is based on a single view of a chessboard pattern, but without any geometric or model constraints. The chromatic model is calibrated in a closed-form least

squares fashion, from which the colour planes are re-aligned to negate the effects of the aberration. A selection of examples before and after compensation show the superior quality images achievable with such aberration removal. It is not limited to image quality concerns, and also offers a contribution in the powerful but possibly slightly under-used area of colour vision systems.

7.1.5 Distortion minimisation in rectification

Finally, Chapter 6 describes a direct, self-contained methods for planar rectification of stereo pairs with particular focus on minimising projective distortions. The method, presented in Mallon and Whelan (2005), is based solely on an examination of the Fundamental Matrix, where an improved method is given for the derivation of two rectifying transforms. A approach to uniquely optimised each transform in order to minimise perspective distortions is proposed. This ensures that the rectified images resemble the originals as closely as possible. Detailed results show that the rectification precision exactly matches the estimation error in the Fundamental Matrix calculation. In tests, the remaining perspective distortions are shown to be much less than alternative methods that also consider their minimisation. This means that for follow on applications such as stereo matching, the rectified images maintain, as much as possible, the structure of the original images to leave the local matching regions virtually invariant.

7.2 Publications arising

The following list of publications stem directly from the work conducted during this project, including background work on robotics, external presentations and submitted work. All papers are full length and peer reviewed.

Calibration and Removal of Lateral Chromatic Aberration in Images J. Mallon and P.F. Whelan *Pattern Recognition Letters*, (Accepted)

Projective Rectification from the Fundamental Matrix J. Mallon and P.F. Whelan *Image and Vision Computing*, Volume 23, Issue 7, pp 643-650, 2005.

Precise Radial Un-distortion of Images J. Mallon and P.F. Whelan *In Proc. ICPR2004 - 17th International Conference on Pattern Recognition*, Cambridge, UK, pp 18-21, August 2004.

Robust 3-D Landmark Tracking using Trinocular Vision J. Mallon, O. Ghita and P.F. Whelan *In Proc. SPIE OPTO-Ireland*, Galway, Ireland, 2002.

An Integrated Design Towards the Implementation of an Autonomous Mobile Robot J. Mallon, O. Ghita, P. F. Whelan *In Proc. OPTIM 2002 - 8th International Conference On Optimization of Electrical and Electronic Equipment*, Brasov, Romania, 2002.

Epipolar line extraction using feature matching O. Ghita, J. Mallon and P.F. Whelan *In Proc. Irish Machine Vision and Image Processing Conference*, Maynooth, Ireland, pp 87-95, 2001.

Auxiliary publications and presentations

Computational approach for depth from focus O. Ghita, P. F. Whelan and J. Mallon, *Journal of Electronic Imaging* Volume 14 issue 2, pp 1-8, 2005

Mobile robotics and trinocular vision J Mallon, *Presentation to Australian Centre for Field Robotics* University of Sydney, Australia, January 2003.

Submitted publications [1]

Which Pattern? Biasing Aspects of Planar Calibration Patterns and Detection Methods J. Mallon and P.F. Whelan *Pattern Recognition*, (Under review: Submitted July 12, 2005)

7.3 Direction for future research

Several extensions and related topics to this work are worthy of further investigation. This section proposes a few directions for future research.

[1] At the time of writing these are under review

7.3.1 Chessboard detection

Considering the popularity of planar calibration targets in many applications, including the calibration of distortion as presented in this thesis, there is a need for dedicated methods for automatically extracting accurate estimates, and ordering, of control points from these targets. It has been proven in this thesis that chessboard targets give much higher quality data than alternative patterns. Thus, the detection methods should be focused on these types of patterns. The detection consists of three subproblems, initial estimation, accurate refinement, and data ordering.

It was found during the experimentation, that in the presence of image blur, large lens distortion or acute pattern positions, that the regular corner detectors failed to register all the intersections of the chessboard squares. This problem can be overcome by taking into account the known intensity structure around a intersection point. False hits can be filtered as described in this thesis, Section 3.2.2.

It was also found that the saddle point refinement process is fast but very dependent on the level of blurring, distortion and the initial location guess. Conversely, the edge based approach is quite computationally expensive, but accurate when applied with a relatively small ROI, regardless of blurring, distortion or initial location [2]. A new technique is required that combines the benefits of each of these methods. It should be fast, invariant to image blurring and distortion and able to cope with poor initialisation. It should also be possible to return a confidence measure regarding the accuracy to which the intersection point is detected.

Lastly, for many calibration methods, a comparison with a calibration model is required, invariably requiring the ordering of the detected data to form the correct correspondences. A basic method is outlined in Section 3.2.3, but only works unaided for relatively low distortion. For higher levels an initial estimate of distortion is required to undistort the points. It may be possible to automate this process regardless of distortion level for example using a Delaunay triangulation.

[2] Within reason, at maximum of a few pixels in difference is typical

7.3.2 Axial chromatic aberration

Chapter 5 described the compensation of lateral chromatic aberration only. As mentioned there is also a axial component to this aberration that introduces a similar radially and colour dependent blurring. Although the correction for lateral chromatic aberration gives a significant improvement, the complete chromatic aberration would represent another step forward. There is no reason that it cannot be modelled using the same model as its lateral cousin, and may also be measured using the chessboard pattern. Considering that the location of the intersection points of the chessboard are already detectable, the axial chromatic aberration introduces a sort of radial smearing of the intensity with the result that the local intensity area is not balanced. This imbalance may be measured, directly corresponding with the axial chromatic aberration. Its image compensation may not be so straightforward, requiring a programmable de-blurring filter with a variable origin. The compensation for all chromatic aberrations may facilitate more colour applications, for example in underwater imaging, where additional chromatic aberrations are introduced due to the extra medium (water-glass-air). For general image enhancement, the successful compensation of chromatic aberrations offers significant improvements in the perceived image quality.

7.3.3 Model based image warping

Currently the image re-sampling process is carried out based on the known distortion function using bilinear interpolation to calculate the new pixel intensity. This method is point-wise correct, but because of the use of supporting pixel intensities is not area-wise correct. The interpolation area should not be based on a regular patch, instead it should be based on a distortion warped patch that transforms into a regular patch in the corrected image. This may be addressed by either using the distortion function in order to manipulate the interpolation patch or through a distortion based biasing of the regular shaped patch intensities.

7.3.4 Optimal image rectification

The method outlined in chapter 6 optimally estimates the undetermined parameters of a planar rectifying transform. However, the actual rectifying component, based on Hartley (1999), only considers the optimal formation of one of these rectifying pairs. This indeed may not even be the most optimal formation. Alternative methods such as Loop and Zhang (1999) are unstable, while their motives of imposing affine qualities are not optimal either. There is a need for a formulation that will also lead to the formation of rectifying transforms that specifically minimise perspective distortions. It may also be beneficial to bypass the explicit computation of relative geometry, i.e the Fundamental Matrix or trifocal tensor.

7.4 Concluding remarks

The techniques presented in this thesis have a broad scope and influence for many tasks that at least require some consideration of distortions. However, the evaluations are carried out independently of such applications as it is considered that a stand alone investigation effectively allows the extrapolation of performance for any potential application. As an example, the rectification procedure is evaluated using specific metrics that demonstrate satisfactory rectification, without recourse to stereo algorithms and their inherent drawbacks. Similarly, the aberration modelling and calibration pattern contributions have a broader scope than what they were used for in this thesis. These were not explicitly investigated, but their benefits may be appreciated in light of the detailed investigations presented.

Appendices

Appendix A

Radially Weighted Homography

Following the calibration of geometric distortions in an image, a meaningful measure of the residual aberration is required to assess the success of the procedure. This primarily refers to the aberration models' fitting ability. A useful means of generating data to carry out this evaluation, is to use corrected images of planar calibration patterns. The points recovered from these corrected images are referred to as $\bar{\mathbf{c}} = (\bar{u}, \bar{v}, 1)^T$.

Ideally, the exact residuals would be formed by the difference between these points and the perfect projection points. In the clear absence of such ideal projections, an approximation is required. One such approximation can be formed by taking a homography between the canonical model of the calibration pattern, $\bar{\mathbf{w}} = (x_w, y_w, 1)^T$ [1], and the recovered points $\bar{\mathbf{c}}$, by minimising the algebraic distance, $\sum(\bar{\mathbf{c}}, \mathbf{H}\bar{\mathbf{w}})$. This results in a least square fit to the points $\bar{\mathbf{c}}$, with the effect of minimising the resulting residuals. Thus, the true residuals are guaranteed to be greater or equal to these.

A better approximation of the true projected points can be obtained by noting that geometric aberrations are radially dependent about the optical axis. Thus, it can be safely assumed that points close to the optical axis will exhibit less residual magnitude than outer points. By appropriately weighting the data in favour of central points, giving less influence to outer points, a better approximation of true residuals can be gleaned. A Gaussian type weighting is applied over the image window as shown in Fig. A.1 by minimising the

[1] The calibration pattern is assumed to have a high degree of planarity with high spacial measurement accuracy.

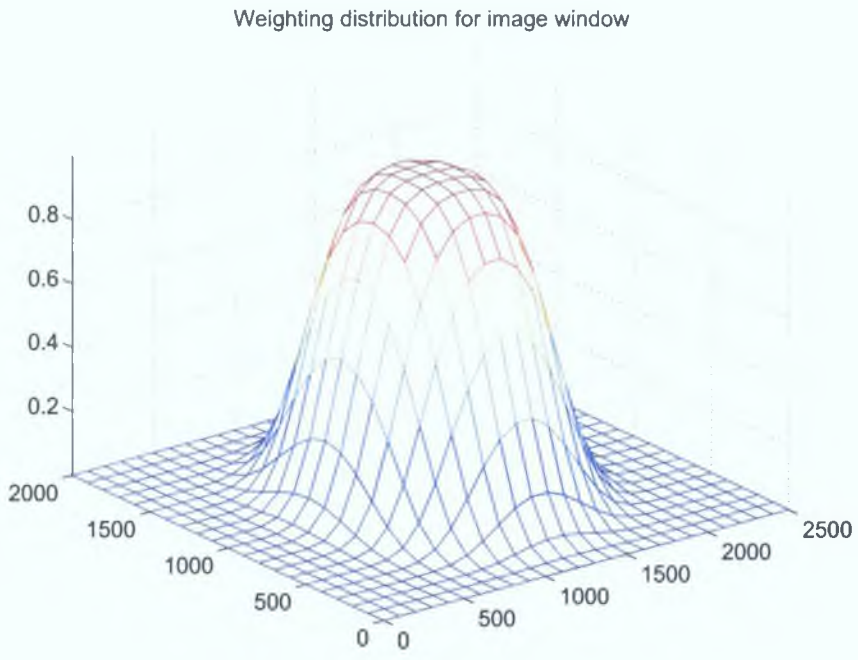


Fig. A.1: Central weighting of an image window of dimensions 2500 × 2000

algebraic quantity[2]:

$$\sum_{i}^{n} \mathcal{C}_i(\bar{\mathbf{c}}_i, \mathbf{H}\bar{\mathbf{w}}_i)^2, \text{ where } \mathcal{C}_i = \exp\left(-\left(\frac{(\bar{u}_i - c_x)^2 + (\bar{v}_i - c_y)^2}{s^2}\right)^2\right),$$

(c_x, c_y) is the optical axis or image centre, s is the spread and n is the number of points. Good values for s are around 0.15 times the image width.

Following a data normalisation procedure that scales both sets of data within a unit circle centered at the origin, the solution is obtained by stacking the equations in the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x} = (h_1, h_2, ...h_8)^T$ with $h_9 = 1$. Rearranging, including the weighting, the solution is obtained as:

$$\mathbf{x} = (\mathbf{A}^T\mathcal{C}\mathbf{A})^{-1}\mathbf{A}^T\mathcal{C}\mathbf{b}.$$

The appropriateness of this measure is demonstrated by evaluating on a range of synthetic residuals, generated by simulating with very small distortion values as shown in Fig. A.2. The dominant distortion is high order i.e. (k_2) with $k_1 = .5k_2$. As can be seen, the weighted Homography serves to approximate the actual residuals much better than the regular un-weighted version.

[2]A tradeoff between robustness and accuracy is made here. The data is first pre-scaled to lie within a unit circle centered at the origin to improve conditioning. As is known this improves the precision of the estimate, while maintaining the robustness and speed of a closed form estimation.

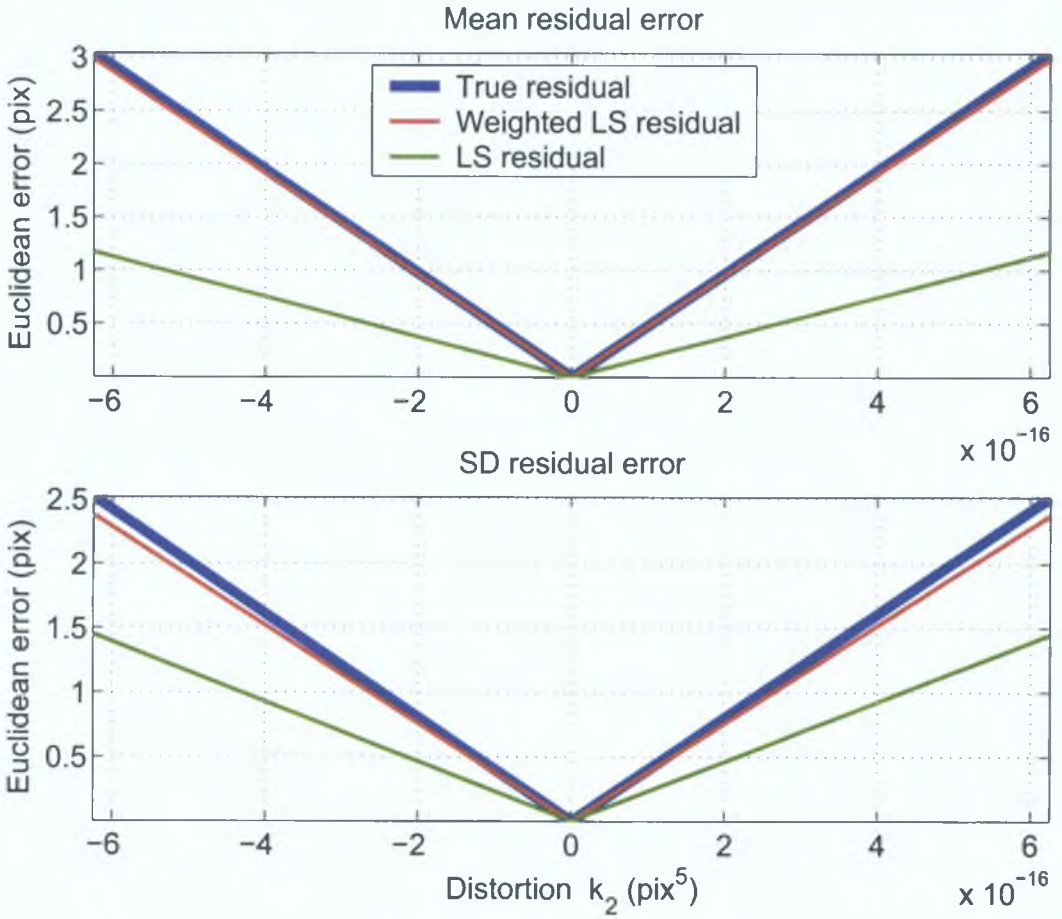


Fig. A.2: Euclidean means and SD of residuals computed with regular Homography estimation (Blue) and Weighted Homography estimation (Red). Simulations are carried out with low levels of distortion. Results show that Weighted homography estimation significantly improves the approximation of true residuals in comparison with the regular Least Square (LS) method (Green).

Appendix B

Closed-form estimation of distortion

This appendix computes the partial derivatives for use in the iterative estimation of distortion in chapter 4. Each control point $\breve{\mathbf{c}}_i = (\breve{u}_i, \breve{v}_i, 1)^T$ contributes to the formation of the following objective function which must be minimised:

$$\mathbf{e}_i(\breve{\mathbf{c}}_i, \mathbf{\Phi}) = \mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}_i + \mathcal{D}(\mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}_i, \mathbf{k}) - \mathbf{A}\breve{\mathbf{c}}_i, \tag{B.1}$$

where the parameter vector is $\mathbf{\Phi} = (h_{11}, h_{12}, \ldots, h_{32}, s, u_o, v_o, k_1, k_2, \ldots, p_1, p_2)^T$. Performing a first order expansion of the error $\mathbf{e}(\breve{\mathbf{c}}, \mathbf{\Phi})$ around the last estimate $\hat{\mathbf{\Phi}}_k$ results in a Gauss-Newton scheme that can be iterated utilising many robust least square techniques (Golub and Loan, 1996):

$$\hat{\mathbf{\Phi}}_{\mathbf{k}+1} = \hat{\mathbf{\Phi}}_\mathbf{k} - \lambda \left(\frac{\partial \mathbf{e}^T(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}} \frac{\partial \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \right)^{-1} \frac{\partial \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k}), \tag{B.2}$$

If the forward distortion model $\mathcal{D}(\hat{\mathbf{p}}, \hat{\mathbf{k}})$ is computed with three radial distortion parameters [1] where $\hat{\mathbf{p}} = \mathbf{H}_\mathbf{e}\tilde{\mathbf{c}}$ are the estimated lens centric coordinates of the ideal projections. The derivatives are given exactly as:

$$\frac{\partial \mathbf{e}(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} = \begin{pmatrix} \dfrac{e_x(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \\ \dfrac{e_y(\breve{\mathbf{c}}, \hat{\mathbf{\Phi}}_\mathbf{k})}{\partial \mathbf{\Phi}^T} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial e_x}{\partial \mathbf{H}_\mathbf{e}}, & -\breve{u}, 1, 0, \hat{x}\hat{r}^2, \hat{x}\hat{r}^4, \hat{x}\hat{r}^6 \\ \dfrac{\partial e_y}{\partial \mathbf{H}_\mathbf{e}}, & 0, 0, 1, \hat{y}\hat{r}^2, \hat{y}\hat{r}^4, \hat{y}\hat{r}^6 \end{pmatrix}, \tag{B.3}$$

[1] Decentering distortion is automatically included in this formulation due to the variable projective transformation $\mathbf{H}_\mathbf{e}$, as described in Section 4.4.1.

where

$$\begin{pmatrix} \frac{\partial e_x}{\partial \mathbf{H_e}} \\ \frac{\partial e_y}{\partial \mathbf{H_e}} \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + k_1 \left((3\hat{x}^2 + \hat{y}^2) \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + 2\hat{x}\hat{y} \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) + \dots \\ \frac{\partial \hat{y}}{\partial \mathbf{H_e}} + k_1 \left(2\hat{x}\hat{y} \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + (\hat{x}^2 + 3\hat{y}^2) \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) + \dots \\ k_2 \left((5\hat{x}^4 + 6\hat{x}^2\hat{y}^2 + \hat{y}^4) \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + (4\hat{x}^3\hat{y} + 4\hat{x}\hat{y}^3) \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) + \dots \\ k_2 \left((4\hat{x}^3\hat{y} + 4\hat{x}\hat{y}^3) \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + (\hat{x}^4 + 6\hat{x}^2\hat{y}^2 + 5\hat{y}^4) \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) + \dots \\ k_3 \left((7\hat{x}^6 + 15\hat{x}^4\hat{y}^2 + 9\hat{x}^2\hat{y}^4 + \hat{y}^6) \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + (6\hat{x}^5\hat{y} + 12\hat{x}^3\hat{y}^3 + 6\hat{x}\hat{y}^5) \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) \\ k_3 \left((6\hat{x}^5\hat{y} + 12\hat{x}^3\hat{y}^3 + 6\hat{x}\hat{y}^5) \frac{\partial \hat{x}}{\partial \mathbf{H_e}} + (\hat{x}^6 + 9\hat{x}^4\hat{y}^2 + 15\hat{x}^2\hat{y}^4 + 7\hat{y}^6) \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \right) \end{pmatrix},$$

where

$$\begin{pmatrix} \frac{\partial \hat{x}}{\partial \mathbf{H_e}} \\ \frac{\partial \hat{y}}{\partial \mathbf{H_e}} \end{pmatrix} = \begin{pmatrix} \tilde{x}, \tilde{y}, 1, 0, 0, 0, -\hat{x}\tilde{x}, -\hat{x}\tilde{y} \\ 0, 0, 0, \tilde{x}, \tilde{y}, 1, -\hat{y}\tilde{y}, -\hat{x}\tilde{y} \end{pmatrix} \Big/ \; h_{31}\tilde{x} + h_{32}\tilde{y} + 1$$

Bibliography

Ahmed, M. T. and Farag, A. A. (2001), Differential methods for nonmetric calibration of camera lens distortion, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', Vol. 2, pp. 477–482.

Al-Shalfan, K. A., Haigh, J. G. B. and Ipson, S. S. (2000), 'Direct algorithm for rectifying pairs of uncalibrated images', *Electronics Letters* **36**(5), 419–420.

Armangué, X. and Salvi, J. (2003), 'Overall view regarding fundamental matrix estimation', *Image and vision computing* **21**, 205–220.

Asari, K. V., Kumar, S. and Radhakrishnan, D. (1999), 'New approach for nonlinear distortion correction in endoscopic images based on least squares estimation', *IEEE Transactions on Medical Imaging* **18**(4), 345–354.

Barreto, J. P. and Daniilidis, K. (2004), Wide area multiple camera calibration and estimation of radial distortion, *in* 'Omnivis2004', pp. 5–16.

Basu, A. and Licardie, S. (1995), 'Alternative models for fish-eye lenses', *Pattern Recognition Letters* **16**(4), 433–441.

Born, M. and Wolf, E. (1980), *Principles of Optics*, sixth edn, Permagon.

Boult, T.E.; Wolberg, G. (1992), Correcting chromatic aberrations using image warping, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', pp. 684 – 687.

Brauer-Burchardt, C. and Voss, K. (2001), A new algorithm to correct fish-eye- and strong wide-angle-lens-distortion from single images, *in* 'IEEE International Conference on Image Processing', Vol. 1, pp. 225–228.

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

Brown, D. C. (1971), 'Close-range camera calibration', *Photogrammetric engineering* **37**, 855–866.

Clarke, T. A. and Fryer, J. G. (1998), 'The development of camera calibration methods and models', *Photogrammetric Record* **16**(91), 51–66.

Clarke, T. A., Fryer, J. G. and Wang, X. (1998), 'The principle point and ccd cameras', *Photogrammetric Record* **16**(92), 293–312.

Conrady, A. (1919), 'Decentering lens systems', *Monthly notices of the royal astronomical society of america* **79**, 951–954.

Devernay, F. and Faugeras, O. (2001), 'Straight lines have to be straight', *Machine Vision and Applications* **13**(1), 14–24.

Farid, H. and Popescu, A. C. (2001), 'Blind removal of lens distortion', *Journal of the optical society of america* **18**(9), 2072–2078.

Faugeras, O. and Luong, Q.-T. (2001), *The Geometry of Multiple Images*, The MIT Press.

Fitzgibbon, A. W. (2001), Simultaneous linear estimation of multiple view geometry and lens distortion, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', Vol. 1, pp. 125–132.

Fryer, J. G. and Brown, D. C. (1986), 'Lens distortion for close-range photogrametry', *Photogrammetric engineering and remote sensing* **52**(1), 51–58.

Fusiello, A., Trucco, E. and Verri, A. (2000), 'A compact algorithm for rectification of stereo pairs', *Machine Vision and Applications* **12**(1), 16–22.

Ghita, O., Mallon, J. and Whelan, P. F. (2001), Epipolar line extraction using feature matching, *in* 'Proceedings of the irish machine vision conference', pp. 87–97.

Golub, G. H. and Loan, C. F. V. (1996), *Matrix computations*, 3 edn, John Hopkins University press.

Haneishi, H., Yagihashi, Y. and Miyake, Y. (1995), 'A new method for distortion correction of electronic endoscope images', *IEEE Transactions on Medical Imaging* **14**(3), 548–555.

Hartley, R. I. (1997), 'In defense of the 8-point algorithm', *IEEE Trans. Pattern Analysis and Machine Intelligence* **19**(6), 580–593.

Hartley, R. I. (1999), 'Theory and practice of projective rectification', *International Journal of Computer Vision* **35**(2), 115–127.

Hartley, R. I. and Zisserman, A. (2003), *Multiple View Geometry in Computer Vision*, second edn, Cambridge University Press.

Hartley, R. and Kaucic, R. (2002), Sensitivity of calibration to principal point position, *in* 'Proceedings of 7th European Conference on Computer Vision', Vol. 2, pp. 433 – 446.

Heckbert, P. S. (1989), Fundamentals of texture mapping and image warping, Master's thesis, University of California, Berkeley.

Heikkila, J. (2000), 'Geometric camera calibration using circular control points', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(10), 1066–1077.

Heikkila, J. (2005), 'Camera calibration toolbox for matlab', http://www.ee.oulu.fi/ jth/calibr/.

Heikkila, J. and Silven, O. (1997), A four-step camera calibration procedure with implicit image correction, *in* 'Computer Vision and Pattern Recognition', pp. 1106 – 1112.

IPL (2000), 'Intel image processing library', http://developer.intel.com.

Isgrò, F. and Trucco, E. (1999), On projective rectification, *in* 'Proceedings IEE Conference on Image Processing and Analysis', pp. 42–46.

Jackowski, M., Goshtasby, A., Bines, S., Roseman, D. and Yu, C. (1997), 'Correcting the geometry and color of digital images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(10), 1152–1158.

Jain, R., Kasturi, R. and Schunck, B. G. (1995), *Machine vision*, McGraw-Hill.

Kannala, J. and Brandt, S. (2004), A generic camera calibration method for fish-eye lenses, *in* 'International Conference on Pattern Recognition', Vol. 1, pp. 10–13.

Keys, E. G. (1981), 'Cubic convolution interpolation for digital image processing', *IEEE Transactions on Acoustics, Speech, and Signal Processing* **29**(6), 11531160.

Kingslake, R. (1978), *Lens Design fundamentals*, academic press.

Kopparapu, S. K. and Corke, P. (1999), The effect of measurement noise on intrinsic camera calibration parameters, *in* 'IEEE International conference on Robotic and Automation', pp. 1281–1286.

Kuzubek, M. and Matula, P. (2000), 'An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy', *Journal of Microscopy* **200**(3), 206–217.

Lavest, J. M., Viala, M. and Dhome, M. (1998), Do we really need an acccurate calibration pattern to achieve a reliable camera calibration, *in* 'European conference on computer vision', pp. 158–174.

Li, M. and Lavest, J. M. (1996), 'Some aspects of zoom lens camera calibration', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(11), 1105 – 1110.

Loop, C. and Zhang, Z. (1999), Computing rectifying homographies for stereo vision, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition', Vol. 1, pp. 125–131.

Lucchese, L. (2005), 'Geometric calibration of digital cameras through multi-view rectification', *Image and Vision computing* **23**, 517–539.

Lucchese, L. and Mitra, S. K. (2002), Using saddle points for subpixel feature detection in camera calibration targets, *in* 'Asia-Pacific Conference on Circuits and Systems', Vol. 2, pp. 191 – 195.

Lucchese, L. and Mitra, S. K. (2003), Correction of geometric lens distortion through image warping, *in* 'International Symposium on Image and Signal Processing and Analysis', Vol. 1, pp. 516 – 521.

Ma, L., YangQuan, C. and Moore, K. L. (2003), Flexible camera calibration using a new analytical radial undistortion formula with application to mobile robot localization, *in* 'IEEE International Symposium on Intelligent Control', pp. 799 – 804.

Mallon, J. (2001), Mobius: An integrated design towards the implementation of an autonomous mobile robot, Technical report, Vision Systems Lab, DCU, Electronic engineering, Dublin City University.

Mallon, J., Ghita, O. and Whelan, P. F. (2002*a*), An integrated design towards the implementation of an autonomous mobile robot, *in* 'Proceedings of the 8th International Conference On Optimization of Electrical and Electronic Equipment', Vol. 2, pp. 577–582.

Mallon, J., Ghita, O. and Whelan, P. F. (2002*b*), Robust 3-d landmark tracking using trinocular vision, *in* 'OPTO-Ireland: SPIE's Regional Meeting on Optoelectronics, Photonics and Imaging, Galway'.

Mallon, J. and Whelan, P. F. (2004), Precise radial un-distortion of images, *in* '17th International Conference on Pattern Recognition', Vol. 1, pp. 18–21.

Mallon, J. and Whelan, P. F. (2005), 'Projective rectification from the fundamental matrix', *Image and Vision Computing* **23**(7), 643–650.

Micusik, B. and Pajdla, T. (2003), Estimation of omnidirectional camera model from epipolar geometry, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', Vol. 1, pp. 485–490.

Ng, K. C., Trivedi, M. and Ishiguro, H. (2002), 'Generalized multiple baseline stereo and direct virtual view synthesis using range-space search, match, and render', *International Journal of Computer Vision* **47**(1-3), 131–147.

Papadimitriou, D. V. and Dennis, T. J. (1996), 'Epipolar line estimation and rectification for stereo image pairs', *IEEE transactions on image processing* **5**(4), 672–676.

Peuchot, B. (1992), Accurate subpixel detectors, *in* 'IEEE conference on Engineering in Medicine and Biology Society', pp. 1958 – 1959.

Pollefeys, M., Koch, R. and Gool, L. V. (1999), A simple and efficient rectification method for general motion, *in* 'Proceedings of the International Conference on Computer Vision', Vol. 1, pp. 496–501.

Prescott, B. and McLean, G. F. (1997), 'Line-based correction of radial lens distortion', *Graphical Models and Image Processing* **59**(1), 39–47.

Salvi, J., Armangue, X. and Batlle, J. (2002), 'A comparative review of camera calibrating methods with accuracy evaluation', *Pattern Recognition* **35**, 1617–1635.

Sawhney, H. S. and Kumar, R. (1999), 'True multi-image alignment and its application to mosaicing and lens distortion correction', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(3), 235–243.

Shah, S. and Aggarwal, J. K. (1996), 'Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation', *Pattern Recognition* **29**(11), 1775–1788.

Slama, C. C. (1980), *Manual of photogrammetry*, 4 edn, American society of photogrammetry.

Stein, G. P. (1993), Internal camera calibration using rotation and geometric shapes, Master's thesis, MIT.

Stein, G. P. (1997), Lens distortion calibration using point correspondences, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', pp. 17–19.

Sturm, P. F. and Maybank, S. J. (1999), On plane-based camera calibration: A general algorithm, singularities, applications, *in* 'IEEE conference on computer vision and pattern recognition', Vol. 1, pp. 432–437.

Swaminathan, R. and Nayar, S. K. (2000), 'Nonmetric calibration of wide-angle lenses and polycameras', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(10), 1172–1178.

Tamaki, T. (2002), 'Correcting distortion of image by image registration with implicit function theorem', *International journal of Image and Graphics* **2**(2), 309–329.

Tamaki, T., Yamamura, T. and Ohnishi, N. (2002), Unified approach to image distortion, *in* 'International conference on pattern recognition', Vol. 2, pp. 584–587.

Tsai, R. Y. (1987), '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 Automatomation* **3**(4), 323–344.

Walter, E. and Pronzato, L. (1997), *Identification of parametric models*, Springer.

Wei, G. Q. and Ma, S. D. (1994), 'Implicit and explicit camera calibration: theory and experiments', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(5), 469–480.

Weng, J., Ahuja, N. and Huang, T. S. (1993), 'Optimal motion and structure estimation', *IEEE Transactions on pattern analysis and machine intelligence* **15**(9), 864–884.

Weng, J., Cohen, P. and Herniou, M. (1992), 'Camera calibration with distortion models and accuracy evaluation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(10), 965–980.

Whelan, P. F. and Molloy, D. (2000), *Machine Vision Algorithms in Java: Techniques and Implementation*, Springer (London).

Willson, R. (1994), Modeling and Calibration of Automated Zoom Lenses, PhD thesis, Carnegie Mellon University.

Willson, R. G. and Shafer, S. A. (1991), Active lens control for high precision computer imaging., *in* 'IEEE International conference on robotics and automation.', pp. 2063–2070.

Willson, R. G. and Shafer, S. A. (1994), 'What is the center of the image', *Journal of the optical society of america* **11**(11), 2946–2955.

Ying, X. and Hu, Z. (2004), Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model?, *in* 'European Conference on Computer Vision (ECCV 2004)'.

Yu, W. (2004), 'Image-based lens geometric distortion correction using minimization of average bicoherence index', *Pattern Recognition* **37**(6), 1175–1187.

Zhang, Z. (1996), On the epipolar geometry between two images with lens distortion, *in* 'International Conference on Pattern Recognition', Vol. 1, pp. 407 – 411.

Zhang, Z. (1998), A flexible new technique for camera calibration, Technical report, Microsoft research.

Zhang, Z. (2000), 'A flexible new technique for camera calibration', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), 1330 – 1334.

Zhang, Z. (2005), 'A flexible new technique for camera calibration', http://research.microsoft.com/ zhang/calib/.