The Normalised Image of the Absolute Conic and its application for zooming camera calibration

Jean-Yves Guillemaut *, John Illingworth

*School of Electronics and Physical Sciences, University of Surrey,
Guildford, Surrey, GU2 7XH, United Kingdom*

Abstract

We present a novel technique for calibrating a zooming camera based on the invariance properties of the *Normalised Image of the Absolute Conic (NIAC)*. We show that the camera parameters independent of position, orientation and zooming are determined uniquely by the NIAC, and we exploit these invariance properties to develop a stratified calibration method that decouples the calibration parameters. The method is organised in three steps: i) computation of the NIAC, ii) computation of the focal length for each image, iii) computation of the orientation and the position of the camera. The method requires a minimum of three views of a single planar grid. Experiments with synthetic and real data suggest that the method is competitive with other state-of-the-art plane-based zooming calibration methods in the scenarios considered.

Key words: Camera Calibration, Zooming Camera, 2D Calibration Object, Image of the Absolute Conic, Invariants.

1 Introduction

The ability to zoom is of considerable interest in computer vision as it enables to focus on selected parts of a scene. However this also requires more complex calibration techniques. In the case of motorised zoom lenses, the relationship between lens control parameters and camera parameters can be determined from calibration at a series of sampled lens settings. For example, in [1], the parameter values estimated at the sampled positions are stored in a look-up table, from which parameters corresponding to new settings are derived by interpolation. Chen et al.

---

* Corresponding author. Tel.: +44 1483 683958; fax: +44 1483 686031.
  Email addresses: `jy.guillemaut@ieee.org` (Jean-Yves Guillemaut),
`j.illingworth@surrey.ac.uk` (John Illingworth).

[2] accelerated this process by using instead an adaptive algorithm, where samples are selected automatically based on required accuracy. In [3], Willson and Shafer model parameter variations by fitting a polynomial at the sampled values, thus producing a more compact representation but at the cost imposing smooth parameter variations. A more general algorithm based on neural networks was proposed in [4].

For accurate results, the previous methods may require a dense sampling over the full range of all lens settings, which can be a demanding task. In addition, they cannot handle discontinuities in parameter variations, and require the use of motorised lenses with indexed position settings, which is not the case for all camera technologies. Self-calibration methods relax all these assumptions by calibrating the camera directly from the same images used for the vision task. The concept was introduced in [5] by Faugeras et al. in the case of cameras with fixed lens settings, and then generalised to zooming cameras by Pollefeys et al. in [6]. The approach is very attractive, however there exists a number of critical motion sequences for which the solution is ambiguous [7,8]. An example which occurs frequently in practice is a rotating and zooming camera for which specific algorithms have been developed [9,10]. To simplify the self-calibration process, some methods have tried reducing the number of parameters to estimate. For example, in [11] Sturm has shown that pre-calibration can be used to model the interdependency between the zooming camera parameters, and thereby reduce self-calibration to the estimation of a single parameter. Generally, self-calibration techniques rely on sufficient and accurate point correspondences, and require good initial values. Convergence problems and noise often limit the accuracy of such techniques [12].

One of the reasons why calibration of a zooming camera is difficult is that it increases significantly the number of parameters to estimate, often resulting in a larger scale non-linear optimisation problem which can be ill-conditioned. Previous works, in the case of non-zooming cameras, have taken advantage of invariants to decouple the camera parameters into simpler sub-problems and thus guarantee that the number of unknowns of each sub-problem is constant. Examples include Vanishing Points (VPs) [13–21], which are invariant to translation, and the Image of the Absolute Conic (IAC) [22,20,23–28], which is invariant to translation and rotation. In this paper, the invariance properties of the IAC are extended to zooming, by defining the *Normalised Image of the Absolute Conic (NIAC)*, which characterises uniquely the camera parameters independent of position, orientation and zooming. The invariance properties are used to define a stratified calibration which sequentially estimates: i) the invariant intrinsic parameters (computed through the NIAC), ii) the focal length and iii) the extrinsic parameters for each image.

The calibration pattern required by this method consists of a simple 2D grid which can be printed off a standard printer. Only one view of the pattern per zoom setting is required; this means that the calibration pattern could for example be placed beside the object of interest and the camera allowed to move and zoom freely around

2

the object. In total, a minimum of three views is required. Plane-based calibration methods such as in [22,20,23–28] are frequently used because they offer a good compromise in terms of flexibility and accuracy. In comparison, a full 3D calibration pattern is more difficult to produce and also practically difficult to position in a scene if required to be simultaneously visible in all views, while simpler linear patterns such as in [29–33] are difficult to apply to zooming cameras as they require a large number of frames for each zoom setting.

Our main contributions are the following. Firstly we propose a theoretical and experimental analysis of the effects of assuming a fixed principal point (PP) when calibrating a zooming camera. Such an assumption has often been made in the literature with usually no or very little justification. We demonstrate that the image plane error induced by making this assumption depends not only on the amplitude of the PP variation but also on the scene depth relief relative to its distance from the camera and we derive a maximum bound for this error. Our second contribution is the introduction of the concept of the NIAC, which complements the existing hierarchy of invariants, and from which we derive a novel calibration algorithm for zooming cameras. Several implementations are discussed. It will be noted that under the NIAC framework, the method presented in [27,28] is actually a special case of our algorithm although coming from totally different perspectives.

2 Zooming camera model

We adopt a general pinhole camera model where 3D scene points $\boldsymbol{P}_i = [X, Y, Z, 1]^\top$ are mapped into 2D image points $\boldsymbol{p}_i = [u, v, 1]^\top$ by a perspective projection

$$
\boldsymbol{p}_i \sim K_1 F [R \quad \boldsymbol{t}] \boldsymbol{P}_i \quad \text{with} \quad K_1 = \begin{bmatrix} 1 & -\cot\theta & u_0 \\ & \frac{r}{\sin\theta} & v_0 \\ & & 1 \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix}. \quad (1)
$$

The symbol $\sim$ denotes equality up to a non-zero scale factor. $K_f = K_1 F$ represents the intrinsic parameters, which include the coordinates of the PP (u_0, v_0), the focal length f in pixel units, the aspect ratio r (usually close to 1), and the angle θ between the axes of the image reference frame ($s = -f\cot\theta$ is called the skew parameter). The latter parameter is introduced here for generality only, standard cameras having zero skew, i.e. $\theta = \pi/2$ rad. The extrinsic parameters represent the camera orientation (rotation matrix R) and position (translation vector $\boldsymbol{t}$) with respect to the world reference frame. The camera centre C is defined mathematically as the homogeneous vector generating the right null-space of the projection matrix $M = K_1 F [R \quad \boldsymbol{t}]$, i.e. $\boldsymbol{C} = [(-R^\top \boldsymbol{t})^\top, 1]^\top$.

It is well known that the PP as well as the camera centre can move while zooming,

3

therefore the considered model is clearly only an approximation of the rather complicated physical effects of zooming. Such motions are due mainly to changes in the alignments of the lens elements induced by zooming, and have been previously studied in [34,35,3,36]. While fully aware of these motions, we are interested in showing that a model with a fixed PP can yield acceptable results and we would like to establish under what circumstances this applies. There are two main motivations for considering a fixed PP. Firstly it allows calibration in scenarios where calibration with a varying PP would otherwise not be possible, such as calibration from multiple views of a single plane each captured at a different position, orientation and zoom factor [1] (this will be referred to as scenario 1 in the results section). Secondly, even in situations where a variable PP can be computed (e.g. in scenario 2 in the results section), considering a simpler model can still be accurate. In fact it can even be more advantageous as the resulting algorithm can be less prone to convergence or numerical errors than more sophisticated algorithms which involve a larger number of parameters. The main result from this section is to show that the image plane error induced by assuming a fixed PP can be expressed as a function of the amplitude of the PP variation and the relative depth relief, and to give a maximum bound on this error. Experimental results are given to support the claim that a fixed PP assumption is acceptable under these circumstances.

2.1 Theoretical analysis of fixed principal point assumption

Let us consider a camera centred at the origin of the world reference frame and oriented such that the vertical image axis, the horizontal image axis, and the optical axis are aligned respectively with the X, Y and Z world axis. Under this assumption, the camera projection matrix is given by:

$$
M_0 = \begin{bmatrix} f & -f\cot\theta & u_0 \\ & f\frac{r}{\sin\theta} & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & 0 \\ & 1 & 0 \\ & & 1 & 0 \end{bmatrix}.
\tag{2}
$$

Changing the camera zoom setting results primarily in a change of focal length Δf, as well as in a change of camera position $(\Delta t_X, \Delta t_Y, \Delta t_Z)$ and PP coordinates $(\Delta u_0, \Delta v_0)$. Note that the orientation of the camera, which is related to the orientation of the image plane in space, remains unchanged as the CCD chip re-

---

[1] Generally, a model with varying PP requires a minimum of two planar grids captured with constant zoom setting in order to define a sufficient number of constraints on the camera parameters.

mains fixed while zooming. The projection matrix becomes

$$
M = \begin{bmatrix} f + \Delta f & -(f + \Delta f)\cot\theta & u_0 + \Delta u_0 \\ & (f + \Delta f)\frac{r}{\sin\theta} & v_0 + \Delta v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \Delta t_X \\ & 1 & \Delta t_Y \\ & & 1 \ \Delta t_Z \end{bmatrix} ,
\tag{3}
$$

and it can be shown that the projection of a point $\boldsymbol{P} = (X, Y, Z, 1)^\top$ is

$$
\boldsymbol{p} = M\boldsymbol{P} = \begin{bmatrix} \frac{f+\Delta f}{Z+\Delta t_Z}(X + \Delta t_X - Y\cot\theta - \Delta t_Y\cot\theta + \Delta u_0\frac{Z+\Delta t_Z}{f+\Delta f}) + u_0 \\ \frac{f+\Delta f}{Z+\Delta t_Z}((Y + \Delta t_Y)\frac{r}{\sin\theta} + \Delta v_0\frac{Z+\Delta t_Z}{f+\Delta f}) + v_0 \\ 1 \end{bmatrix} .
\tag{4}
$$

If we now consider the simplified model, where zooming is represented by a change of focal length Δf, as well as a change of camera position $(\Delta t'_X, \Delta t'_Y, \Delta t_Z)$, but no change in the PP coordinates, the projection of the same point $\boldsymbol{P}$ is

$$
\boldsymbol{p}' = \begin{bmatrix} \frac{f+\Delta f}{Z+\Delta t_Z}(X + \Delta t'_X - Y\cot\theta - \Delta t'_Y\cot\theta) + u_0 \\ \frac{f+\Delta f}{Z+\Delta t_Z}(Y + \Delta t'_Y)\frac{r}{\sin\theta} + v_0 \\ 1 \end{bmatrix} .
\tag{5}
$$

Imposing $\boldsymbol{p} = \boldsymbol{p}'$, we obtain, after simplification, that the two models are equivalent if and only if the condition

$$
\begin{cases} \Delta t'_X = \Delta t_X + (\Delta u_0 + \Delta v_0\frac{\cos\theta}{r})\frac{Z+\Delta t_Z}{f+\Delta f} , \\ \Delta t'_Y = \qquad \Delta t_Y + \Delta v_0\frac{\sin\theta}{r}\frac{Z+\Delta t_Z}{f+\Delta f} , \end{cases}
\tag{6}
$$

is satisfied simultaneously for all scene points. This condition can be satisfied exactly if all scene points are located in a plane parallel to the image plane (because Z is the same for all scene points), or only approximately otherwise by replacing Z by, for example, an average depth value $\bar{Z}$ computed over all scene points. In this case, the validity of the fixed PP assumption will depend on how closely the conditions defined in Eq. (6) are satisfied. Below, we compute explicitly the effect of such an approximation on the image point error (for simplicity the rest of the derivation is done in the case of a zero-skew camera). We have

$$
\boldsymbol{p}' - \boldsymbol{p} = \frac{f + \Delta f}{Z + \Delta t_Z} \begin{bmatrix} \Delta t'_X - \Delta t_X - \Delta u_0\frac{Z+\Delta t_Z}{f+\Delta f} \\ (\Delta t'_Y - \Delta t_Y)r - \Delta v_0\frac{Z+\Delta t_Z}{f+\Delta f} \\ 0 \end{bmatrix} .
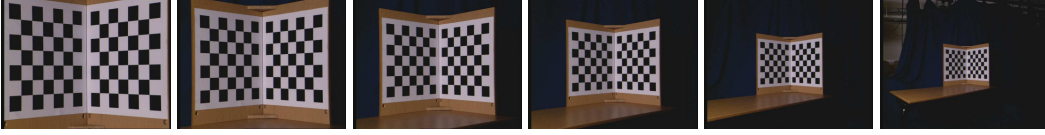\tag{7}
$$

5

Fig. 1. Some of the images used for experimental validation of the PP assumption.

After substitution of Eq. (6) in the previous equations and simplification, we obtain

$$\boldsymbol{p}' - \boldsymbol{p} = \frac{f + \Delta f}{Z + \Delta t_Z} \begin{bmatrix} \Delta u_0 \frac{\bar{Z} + \Delta t_Z}{f + \Delta f} - \Delta u_0 \frac{Z + \Delta t_Z}{f + \Delta f} \\ \Delta v_0 \frac{\bar{Z} + \Delta t_Z}{f + \Delta f} - \Delta v_0 \frac{Z + \Delta t_Z}{f + \Delta f} \\ 0 \end{bmatrix} = \frac{\bar{Z} - Z}{Z + \Delta t_Z} \begin{bmatrix} \Delta u_0 \\ \Delta v_0 \\ 0 \end{bmatrix}. \quad (8)$$

from which we conclude that the squared image plane error for this point is

$$\|\boldsymbol{p} - \boldsymbol{p}'\|^2 = (\Delta u_0^2 + \Delta v_0^2)\frac{(\bar{Z} - Z)^2}{(Z + \Delta t_Z)^2}. \quad (9)$$

This expression provides a maximum bound on the error as the product of the amplitude of the PP motion and the depth relief relatively to the scene distance from the camera. This suggests that assuming a fixed PP is a valid assumption if this product is small. For example, for a camera which exhibits PP variations of 5 pixels along each axis for the zooming range considered and which sees a scene of 1 m depth located 10 m away from the image plane, the maximum image plane error will be $(5^2 + 5^2) \times 0.5^2/10^2 = 0.125$, while the average or root mean squared (RMS) image plane error computed over all points will be far smaller than this.

2.2 *Experimental validation*

Two experiments were carried out, the first to show the motion of the PP, the second to verify that a fixed PP model, although less accurate than a variable PP model, can yield acceptable results. The camera used is a Sony DXC-9100P equipped with a Fujinon S12×5BRM-38 zooming lens which has a 5–60 mm focal length range. The resolution of the images is 720×576 pixels. The camera is mounted on a tripod, and its pose and orientation are kept constant during the experiment, so that variations in the parameters are due only to zooming. The camera is pointing at a calibration grid made of two orthogonal square grids of size 420 mm and located approximately 2500 mm away from the camera. Images are acquired for 36 different zoom settings (see Fig. 1). We deliberately chose a large range of zoom factors.

Firstly, the position of the PP is determined separately for each zoom setting by applying the Gold Standard camera calibration algorithm described in [37]. It can be observed in Fig. 2 that the PP describes an approximately linear motion in the
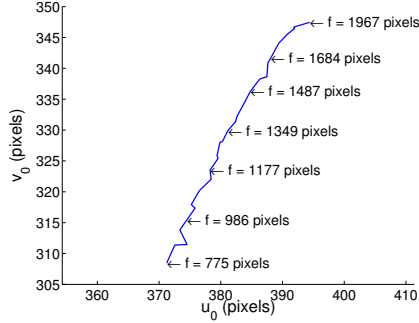
Fig. 2. Displacement of the PP while the camera is zooming.

image whose amplitude is approximately 25 pixels along the horizontal axis and 40 pixels along the vertical axis. This is a rather large PP motion.

Having established that the PP moves during zooming it remains to quantify the impact when this is not explicitly accounted for in the camera model. To do this, we did a similar experiment to the one proposed earlier by Willson and Shafer in [34], in which they looked at calibrating camera data with models of increasing complexity. We consider the following models where zooming is accounted for by allowing variation in: 1) only the focal length f, 2) the focal length f and the PP coordinates (u_0, v_0), 3) the focal length f, the PP coordinates (u_0, v_0), and the t_Z camera position parameter (these three models were considered in [34]), 4) the focal length f and the (t_X, t_Y) camera position parameters, and 5) the focal length f and the (t_X, t_Y, t_Z) camera position parameters (the last two models have not been considered by Willson), where $\boldsymbol{t} = (t_X, t_Y, t_Z)^\top$. It should be noted that all these models are approximations of the physical process of zooming. Here we are interested in analysing the performance of each model, and more particularly to show that the last model which assumes a fixed PP is acceptable.

The experiment starts by calibrating all the camera parameters for the largest zoom setting, then the zoom factor is reduced and the camera re-calibrated for each new setting, allowing only the variable camera parameters of the model chosen to vary. For each model, the Root Mean Squared (RMS) reprojection error is computed and measures the accuracy of the model. The minimum obtainable error is given by performing a full calibration, where all parameters are allowed to vary independently for each setting. In all cases, the calibration is done using the Gold Standard algorithm described in [37]. The linear method is used to initialise the varying zoom parameters, which are then refined by non-linear optimisation. Note that because the scale of the grid varies due to zooming, the minimum obtainable error is expected to decrease slightly over the range of zoom settings.

The results are shown in Fig. 3. The simplest model that allows only the focal length to vary is, not surprisingly, the least accurate and results in an increase in the reprojection error that is unacceptably high. The remaining models which model second order zoom variations such as PP motion or camera centre motion produce
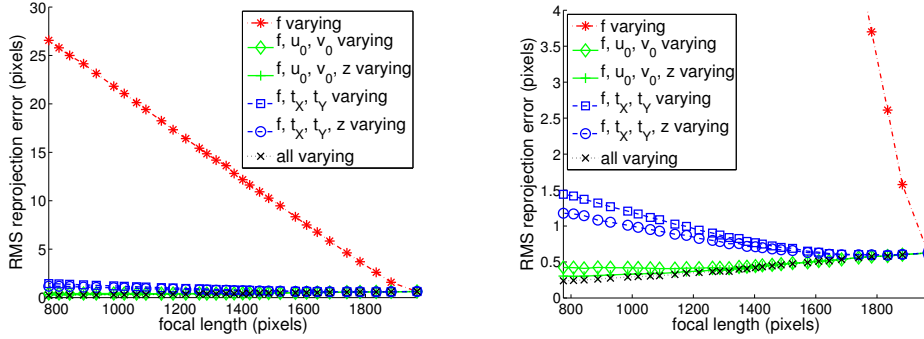
7

Fig. 3. Comparison of the accuracy of the different zooming camera models. The graph on the right hand-side is a magnification of the lower part of the graph on the left hand side.

results that are very dramatically better than the simplest model. The variable PP models yield errors which are closest to the minimum obtainable error (this was observed earlier in [34]), while the variable camera centre models yield slightly larger errors which however remain acceptable. With the best fixed PP model, the error increases by less than a factor of two and does not exceed the order of 1 pixel over the range considered. This can still be accurate enough for many applications.

It is important to stress that the loss in accuracy due to having a simpler model is balanced by a gain in flexibility of the calibration method (possibility to use simpler calibration patterns because fewer parameters must be estimated) and also the possibility to define simpler invariants. The ability to exploit invariants may ultimately translate into an improvement in accuracy when multiple images are used, which may not be observed with a more complex camera model. If the use of a fixed PP model proves too inaccurate in a specific application it can still be useful as an initial estimate that can be refined and extended to a more general camera model including varying PP in a bundle-adjustment.

3   A novel invariant: the NIAC

This section defines a novel geometric entity called the Normalised Image of the Absolute Conic (NIAC) which encapsulates the camera parameters invariant to translation, rotation and focal length variation. The concept of NIAC was first introduced in our preliminary work in [38], where it was inferred from a geometric construction requiring several observations of a pair of calibrated planes. In this paper, this assumption is relaxed and a more elegant formulation allowing calibration from observations of a single plane is proposed.
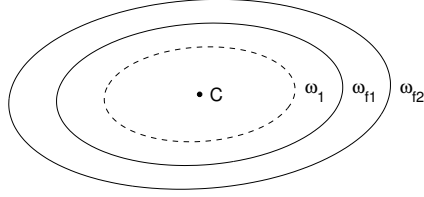
Fig. 4. Illustration of the transformation of the IAC while the camera is zooming. The different IACs ω_{f_i} are all centred in the PP C and homothetic. The one with focal length 1 is chosen as a reference and called the NIAC ω_1.

3.1 Invariance properties of the IAC

For a given focal length f, the IAC is defined by the conic coefficient matrix $\omega_f = (K_f K_f^\top)^{-1}$ or equivalently by the following equation (see [39] pp. 211–212):

$$(u - u_0)^2 + \frac{1}{r^2}(v - v_0)^2 + 2\frac{cos\theta}{r}(u - u_0)(v - v_0) = -f^2 . \tag{10}$$

It appears immediately that any given IAC is centred at the PP and that f is related to only the scale of the IAC. Under the fixed PP model defined previously, zooming produces a one-parameter family of IACs which can be parametrised by the focal length f. The effect of varying f is illustrated in Fig. 4. The set of IACs obtained is *homothetic* (curves are related by an expansion or geometric contraction) and concentric, with the centre at the PP (u_0, v_0) of the camera.

3.2 The NIAC

We define the NIAC as *the IAC corresponding to a focal length of* 1. The NIAC is an imaginary conic represented by the symmetric matrix $\omega_1 = (K_1 K_1^\top)^{-1}$. By construction ω_1 is invariant to the position, orientation and change in the focal length of the camera. It has four degrees of freedom, corresponding to the camera parameters u_0, v_0, r and θ. In terms of invariant, the NIAC can be considered as the natural extension of the IAC to zooming cameras. Because the NIAC encapsulates all the intrinsic parameters invariant to zooming, calibrating these parameters is equivalent to estimating ω_1. Once ω_1 is known, K_1 and therefore the intrinsic parameters invariant to zooming can be recovered from the Cholesky factorisation [40].

4 Application to camera calibration

Before describing the novel camera calibration method, a brief reminder of the principle of plane-based camera calibration using the IAC is given. The main idea is to replace the computation of the calibration matrix K representing the intrinsic

parameters, by the estimation of the IAC. The absolute conic being an imaginary object, it is a priori not directly observable. However, it has been shown in [23,24] that it is possible to compute the image of two remarkable points belonging to it from the observation of a planar calibration pattern. These two points are called *circular points*, and we give a summary of their computation below.

4.1   Computation of the circular points

Let us suppose that the camera is pointing at a planar calibration pattern. By definition, the circular points of this plane are the two points of intersection with the absolute conic. For simplicity and without loss of generality, it is assumed that the calibration plane is located in the plane $Z = 0$, in which case the points of intersection with the absolute conic are the two points $\boldsymbol{I} = [1, i, 0]^{\top}$ and $\boldsymbol{J} = [1, -i, 0]^{\top}$. Because the plane is marked with known control points, it is also possible to compute the homography H between the calibration plane and its image, from which we can derive that the images of the two circular points are: $\boldsymbol{P} = H\boldsymbol{I} = \boldsymbol{h}_1 + i\boldsymbol{h}_2$ and $\boldsymbol{Q} = H\boldsymbol{J} = \boldsymbol{h}_1 - i\boldsymbol{h}_2$, where $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are the first and second column vectors of H. Both points lie on the IAC. In the case of a camera with constant intrinsic parameters, each image of a calibration plane provides two such points on the IAC. A general conic being uniquely defined by five points, three plane observations are sufficient (in practice two are sufficient because of the additional zero-skew constraint) in order to determine uniquely the IAC, and K.

In the case of a zooming camera, it is necessary to consider a more general invariant such as the NIAC. The calibration algorithm proposed can be broken into three stages. In the first stage, the invariant intrinsic parameters encapsulated in the NIAC are computed; such parameters are the coordinates of the PP, the aspect ratio and the skew parameter. This is the most complicated stage of the method. The next stages concentrate, separately for each image, first on the computation of the focal length, then on the computation of the extrinsic parameters, i.e. position and orientation.

4.2   Computation of K_1

The matrix K_1 represents the intrinsic parameters of the camera which are invariant to a change in position, orientation and zooming. These parameters are characterised uniquely by the NIAC $\omega_1 = (K_1 K_1^{\top})^{-1}$. Like the IAC, the NIAC is an imaginary conic, it is therefore not directly observable, and a special construction is needed. As for the IAC, the information is provided by the observation of a sufficient number of calibration planes, which provide a set of pairs of images of the circular points. However, this time there exists as many different IACs as there are pairs of images of circular points, therefore a more elaborate strategy is needed.

We start by observing that, if the parameters from K_1 are known, it is possible to define a normalised image reference frame in which the NIAC is a unit circle centred at the origin. In this normalised image reference frame, the camera has effectively a unit aspect ratio, zero skew, and its PP is at the origin. Such a reference frame is obtained by applying an image transformation T which is composed of a shear transformation along the X axis (to eliminate the skew), a scaling along the Y axis (to correct the aspect ratio), and a translation (to map the PP to the origin). The transformation obtained is parametrised by four parameters t_1, t_2, t_3 and t_4 ($t_3 \neq 0$):

$$
T = \begin{bmatrix} 1 & 0 & t_2 \\ & 1 & t_4 \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & t_3 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & t_1 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_2 \\ & t_3 & t_4 \\ & & 1 \end{bmatrix}. \tag{11}
$$

The main idea of the method is that calibration can be reformulated in terms of identifying the unique transformation T which maps the NIAC to a unit circle centred at the origin. Because all IACs are homothetic and concentric, T maps the set of IACs into a set of concentric circles centred at the origin. We show that such a configuration can be characterised uniquely by the perpendicular bisectors to the chords defined by the pairs of images of circular points on the IACs. The result is stated below and the concept is illustrated in Fig. 5.

Result 1 *Consider n ($n \geq 3$) concentric homothetic conics centred at C. Take one chord on each conic such that no chord passes through C, and no two chords are parallel. The perpendicular bisectors to the chords are concurrent [2] in C if and only if the conics are circular.*

Proof This is a direct result from geometry resulting from the observation that in a non-degenerate ellipse (the homothetic and concentric conics being all of the form $(KK^\top)^{-1}$ they are clearly ellipses) the only centre of symmetry is the ellipse centre and the only axes of symmetry are its axes. $\square$

Given a pair of images of circular points $P = HI = h_1 + ih_2$ and $Q = HJ = h_1 - ih_2$, with $h_1 = [h_{11}, h_{21}, h_{31}]^\top$ and $h_2 = [h_{12}, h_{22}, h_{32}]^\top$, it is shown in Appendix A that after mapping by T, the equation of the perpendicular bisector is:

$$
l = [-(d_1 + t_1 d_2), -t_3 d_2, (m_1 + t_1 m_2 + t_2)(d_1 + t_1 d_2) + (t_3 m_2 + t_4) t_3 d_2]^\top, \tag{12}
$$

with $m_1 = \frac{1}{h_{31}^2 + h_{32}^2}(h_{31}h_{11} + h_{32}h_{12})$, $d_1 = h_{32}h_{11} - h_{31}h_{12}$, $m_2 = \frac{1}{h_{31}^2 + h_{32}^2}(h_{31}h_{21} + h_{32}h_{22})$ and $d_2 = h_{32}h_{21} - h_{31}h_{22}$. Note that although the circular points P and Q are imaginary points, their perpendicular bisector l defines a real line in the image plane. It follows that calibrating K_1 is equivalent to finding the unique values of the parameters t_1, t_2, t_3 and t_4 for which the perpendicular bisectors l are concurrent

---

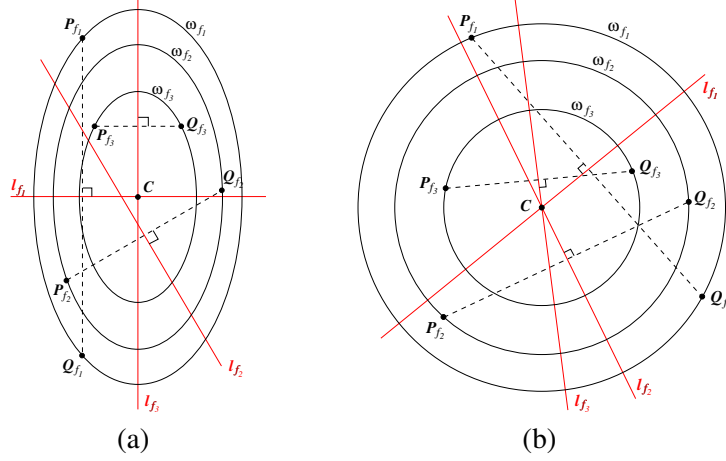[2] Three of more lines are concurrent if they meet at one point.

Fig. 5. ω_{f_1}, ω_{f_2} and ω_{f_3} are three concentric homothetic conics centred at C. On each conic ω_{f_i}, the points P_{f_i} and Q_{f_i} represent the images of the circular points. They define a chord on each conic. We assume that none of the chords passes through C and that no two chords are parallel. The perpendicular bisectors to the chords are represented by the lines l_{f_1}, l_{f_2} and l_{f_3}. In the general case where the conics are non-circular (a), the perpendicular bisectors pass through the centre C if and only if the chord is parallel to an axis of the conic (l_{f_1} and l_{f_3}). Because there are only two axis, l_{f_1}, l_{f_2} and l_{f_3} cannot be concurrent at C. The only case where l_{f_1}, l_{f_2} and l_{f_3} are concurrent at C is when the conics are circular (b).

at the origin. Once this transformation has been estimated, the NIAC is given by $\omega_1 = T^\top T$ and the intrinsic parameters by $K_1 = T^{-1}$.

In practice, due to image noise, the perpendicular bisectors l will not be exactly concurrent at the origin, but we can seek an optimum solution which minimises the distance between these lines and the origin. Two algorithms for estimating these parameters are presented below; the first one minimises the exact geometric distance between these lines and the origin, while the second one minimises a simpler distance, which we call algebraic, and which turns out to be a close approximation of the former geometric distance.

4.2.1 Non-linear solution minimising a geometric distance

The first method proposed finds the solution which minimises the sum of squared distances $\sum d_{\text{geom}}^2$ computed over all views and where d_{geom} represents the distance between the line l defined in Eq. (12) and the image origin. d_{geom} is such that

$$d_{\text{geom}}^2 = \frac{[(m_1 + t_1 m_2 + t_2)(d_1 + t_1 d_2) + (t_3 m_2 + t_4)t_3 d_2]^2}{(d_1 + t_1 d_2)^2 + (t_3 d_2)^2}. \tag{13}$$

In the case of a zero-skew camera, we have $t_1 = 0$, and the previous expression simplifies to

$$d_{\text{geom}}^2 = \frac{[(m_1 + t_2)d_1 + (t_3 m_2 + t_4)t_3 d_2]^2}{d_1^2 + (t_3 d_2)^2}. \tag{14}$$

12

A minimum of four images is required to determine uniquely the fixed intrinsic parameters in the case of a general camera. With a zero-skew camera, three images are sufficient, because t_1 is known to be zero. Minimising such cost functions requires non-linear techniques such as the Levenberg-Marquardt algorithm [40]. Experiments showed good convergence properties, however it is a priori not guaranteed that the convergence of the algorithm to the correct solution may not be affected by a local minimum in the vicinity of the solution. In practice, a reasonable initialisation which gives good results is to choose the PP at the image centre and an aspect ratio of one. Alternatively, the method defined in the next paragraph can be used for initialisation.

4.2.2 *Linear solution minimising an algebraic distance*

Contrary to non-linear methods, linear methods are usually simpler to implement, because they do not need any initialisation and do not suffer from convergence problems. However, because the distance minimised may lack physical meaning, they are usually not so accurate. In this case, the following algebraic constraint is defined by requiring the origin to lie on the line l:

$$[0, 0, 1]\boldsymbol{l} = (m_1 + t_1 m_2 + t_2)(d_1 + t_1 d_2) + (t_3 m_2 + t_4)t_3 d_2 = 0\,. \quad (15)$$

In practice, considering this equation does not present any advantage over the previous method because the equation remains non-linear in the case of a general camera. However, in the case of a zero-skew camera, the unknown values u_0, v_0 and r are related to the entries of T by

$$T = K_1^{-1} = \begin{bmatrix} 1 & 0 & -u_0 \\ & \frac{1}{r} & -\frac{v_0}{r} \\ & & 1 \end{bmatrix}\,, \quad (16)$$

and the following substitution can be carried out: $t_1 = 0$, $t_2 = -u_0$, $t_3 = \frac{1}{r}$ and $t_4 = -\frac{v_0}{r}$. This leads to the equation

$$d_1 r^2 u_0 + d_2 v_0 - m_1 d_1 r^2 = m_2 d_2\,, \quad (17)$$

which is linear in the unknowns $r^2 u_0$, v_0 and r^2. This equation is similar to the one obtained by Gurdjos et al. in [27,28] using the centre-line constraint. The different equations (one per view) can be combined and a least-square solution can be computed by using for example the pseudo-inverse. Although algebraic distances are often less accurate than geometric distance, the distance considered here has the interesting property of being an extremely good approximation of the geometric distance defined in Eq. (13) after suitable normalisation has been applied. This will be further discussed in Section 4.5.1.

13

4.3 Computation of F_1

Computing F_1 is a simple matter of finding the isotropic scaling factor f which maps the NIAC into a conic passing through the images of the two circular points for each image. This is a straightforward least-squares problem (see for example [27] for a solution). It should be noted that the determination of f is not possible if the optical axis of the camera is perpendicular to the calibration plane.

4.4 Computation of R and t

Once all the intrinsic parameters are known, the computation of the extrinsic parameters is done in a standard way (see for example [23]).

4.5 Practical considerations

4.5.1 Normalisation

Normalisation is carried out before computing the homographies between the calibration plane and the image plane. The technique employed is described in [37] and consists in normalising world points and images points such that their centroids coincide with the origin and the average distance from the origin is $\sqrt{2}$. In the case of the non-linear method, no extra minimisation is required as the distance minimised in Eq. (13) defines a geometric distance in the image space. In the case of the linear method, we normalise each term in [27,28] by the inverse of $\sqrt{d_1^2 + d_2^2}$. This choice has been mathematically justified in [27], and we can observe that in the case of an aspect ratio close to one the normalisation introduced is very close to the denominator in Eq. (14). This means that the system is well conditioned and approximates very closely the more complicated geometric distance with the advantage of not requiring non-linear minimisation. It will be shown in the result section that the two methods produce very similar results (although not identical). In comparison, the method described in [24] relies heavily on normalisation and is more prone to stability problems.

4.5.2 Degenerate Configurations

It has been seen earlier that a minimum of three views of the calibration plane is necessary in the case of a standard zero-skew camera, while four views are required in the case of a non standard skewed camera. In addition, the three following assumptions have been made during the discussion: i) the optical axis of the camera is not perpendicular to the calibration plane (Section 4.3 and Appendix A), ii) the

chords defined by the pairs of images of the circular points do not pass through the PP of the camera (Result 1), iii) and no such chords are parallel (Result 1). After observing that the chords are the vanishing lines of the calibration planes observed, it is straightforward to show that ii) corresponds to the case where the optical axis of the camera is parallel to the calibration plane, while iii) corresponds to the case where two cameras are related by a translation and/or a rotation along an axis parallel to the calibration plane. This characterises all the degenerate configurations.

5 Results

In this section, the methods presented earlier are tested and evaluated. A comparison with two of the methods (fixed or variable PP) presented in [24] is also given. When referring to these methods, the following terminology is adopted: *Sturm & Maybank (fixed PP)* and *Sturm & Maybank (variable PP)* denote the Sturm and Maybank method with fixed and variable PP respectively, *linear NIAC* and *nonlinear NIAC* denote the methods based on the NIAC which minimise, respectively, an algebraic distance and a geometric distance. All methods assume zero-skew, which is normally the case of most cameras. It should be noted that the *linear NIAC* method is identical to the method derived by Gurdjos et al. using the centre line constraint in [27,28]. Gurdjos et al. refer to a theorem of projective geometry (Poncelet's theorem) to characterise the locus of the PP when a zero-skew pinhole camera is zooming. Although this is different from the NIAC concept, both methods result in the same linear system of equations in the case of a zero-skew camera.

In all experiments, the camera is pointing at a planar calibration grid (see Fig. 6 in the case of real data). The position, orientation and zoom vary for each frame. For each image, the homography is computed using the Direct Linear Transform (DLT) method as described in [37], with the appropriate normalisation, and then the different calibration algorithms are applied. Two experimental scenarios are considered. In the first case, hereafter referred to as scenario 1, it is assumed that all images of the calibration grid are acquired with independent zoom settings. This is the most general scenario; note that in this case it is not possible to compute a variable PP, therefore the *Sturm & Maybank (variable PP)* method cannot be applied (it requires several observations of at least two planar patterns). The second case, which we will call scenario 2, assumes that pairs of images are acquired with the same zoom factor; all the methods previously mentioned can be compared in this case.

Two different evaluation criteria are considered to measure the accuracy of the calibration methods. The first evaluation criterion is the RMS estimation error defined by: $\epsilon_{\text{est}} = \sqrt{\frac{1}{N} \sum_i \left(\frac{\overline{x}_i - x_i}{\overline{x}_i} \right)^2}$, where $\overline{x}_i$ are the ground truth parameters and x_i are the estimated parameters. The RMS value is computed for each intrinsic parameter.
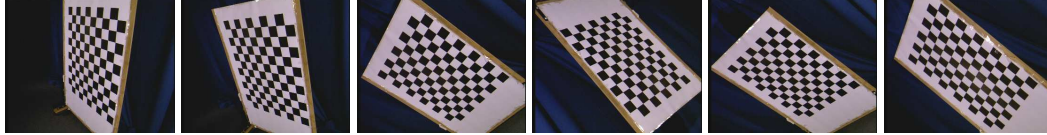
Fig. 6. Real images used for calibration. Each image illustrates a different zoom setting varying from approximately 560 pixels to approximately 2210 pixels. While zooming, the operator also adjusted the camera pose so as to maximise the size of the grid in the image.

Note that because some of the intrinsic parameters are varying while zooming, a relative error is computed. This is a good measure of how closely the estimated camera parameters match the ground truth camera parameters. The second evaluation criterion is the RMS reprojection error $\epsilon_{\text{rep}} = \sqrt{\frac{1}{N'} \sum_i \|\boldsymbol{p}_i - K_1 F [R \quad \boldsymbol{t}] \boldsymbol{P}_i \|^2}$. Both ϵ_{est} and ϵ_{rep} are computed over a large number of experiments (1000) in order to obtain statistically meaningful results.

In order to provide an unbiased evaluation, the RMS reprojection error must be computed on points different from the ones used for calibration, otherwise the RMS reprojection error would correspond to a residual error, which is well known to be a poor measure of the quality of the solution obtained (see [37], Chapter 4). For example, the *Sturm & Maybank (variable PP)* method would be expected to always lead to lower residuals because it has two extra degrees of freedom compared to the other methods and can therefore fit the data better, which does not mean it computes more accurately the camera parameters. In our evaluation methodology, we generate extra images which are used for evaluation only; one such image is generated for each zoom setting by varying the extrinsic parameters while maintaining the intrinsic parameters constant. In the case of synthetic data, the computation of the reprojection error on the evaluation images is straightforward because the exact values of the extrinsic parameters used to generate these images are known. In the case of real data, the computation of the reprojection error is not immediately obvious for the evaluation images because the values of the extrinsic parameters are not readily available (we do not know the pose of the camera for such images). However, the extrinsic parameters can be computed iteratively by finding the values which minimise the reprojection error (i.e. we find the optimum pose); this provides a good estimate of the reprojection error. The optimisation is carried out with the Levenberg-Marquardt algorithm [40].
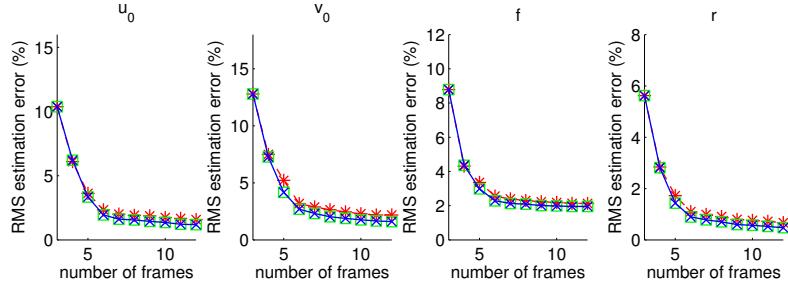
5.1 Synthetic data

The calibration pattern consists of a square grid of size $20\,\text{cm} \times 20\,\text{cm}$ which contains 10×10 control points. The grid coincides with the plane $Z = 0$ of the world reference frame. We simulate a camera with variable PP, where the different intrinsic parameters are defined by the following equations parametrised by the focal length f: $u_0 = 384 + \lambda(\frac{f - f_{\min}}{f_{\max} - f_{\min}} - \frac{1}{2})$ pixels, $v_0 = 247 + \lambda(\frac{f - f_{\min}}{f_{\max} - f_{\min}} - \frac{1}{2})$ pixels,

$r = 1.167$ and $\theta = 90°$ (zero-skew). This simulates a camera where the PP follows a linear motion of amplitude λ along both camera axes while the camera is zooming. This is a plausible model for a zooming camera (see observations made in Section 2.2). For each frame, the focal length is assigned a random value between $f_{\min} = 476$ pixels and $f_{\max} = 1428$ pixels, following a uniform distribution on this interval. The camera centre is located on a 0.5 m radius sphere centred at the middle of the calibration grid. The position and orientation of the camera is generated by applying the following Euler transformation. A random rotation is applied successively around the Z axis (rotation), the X axis (precession) and finally the Z axis (nutation). The rotation around the X axis is constrained between $30°$ and $70°$, so as to be in the optimum condition required by the *Sturm & Maybank* methods. Under such conditions, the grid occupies the whole image at the maximum zoom factor. Some Gaussian noise with standard deviation of 1 pixel is added in the coordinates of each imaged point so as to simulate image noise. For both scenarios, two different amplitudes for the PP motion are considered: i) $\lambda = 5$ pixels (lens with small PP motion) and ii) $\lambda = 50$ pixels (lens with larger PP motion). In each case, calibration results are shown with respect to the number of image frames considered. The results obtained from 1000 experiments are shown in Fig. 7 (independent images) and Fig. 8 (pairs of images).

For all methods the error decreases when the number of images used increases. The *linear NIAC* and *non-linear NIAC* methods are indistinguishable without magnifying the graphs. Theoretically, it has already been observed in [27,28] by Gurdjos et al., that the distance defined by the centre-line constraint, which is equivalent to the one minimised by the *linear NIAC*, is a very good approximation of the optimum geometric distance (here represented by the *non-linear NIAC* method) if the aspect ratio is close to one. Our results confirm this experimentally. If the aspect ratio differed more significantly from one, however, it is expected that the non-linear algorithm would become more accurate than the linear one. In general, it can be observed that the NIAC methods and the *Sturm & Maybank (fixed PP)* method lead to very close results; in fact the differences are marginal, which suggests that in this case the NIAC methods are competitive with a state-of-the-art method assuming a fixed PP.

In the case of the second scenario where pairs of images are acquired with the same zoom factor, the *Sturm & Maybank (variable PP)* method seems to perform slightly worse than the other methods. This may first appear surprising, as we may have expected that including PP variations would allow more accurate camera calibration. This can be explained by the significant increase in the number of unknowns, caused by the introduction of the variable PP, which affects the conditioning of the system of equations. This was also observed previously in [27,28]. One advantage of the *Sturm & Maybank (variable PP)* over the other methods, however, is that its performance remains unaffected by the range of PP variation. If we were to increase further the range of PP variation, it is likely that this method would lead to more accurate results than the other methods for which the error would become
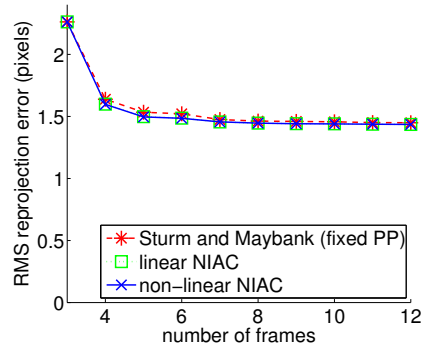
(a) RMS estimation error for each parameter ($\lambda = 5$ pixels).
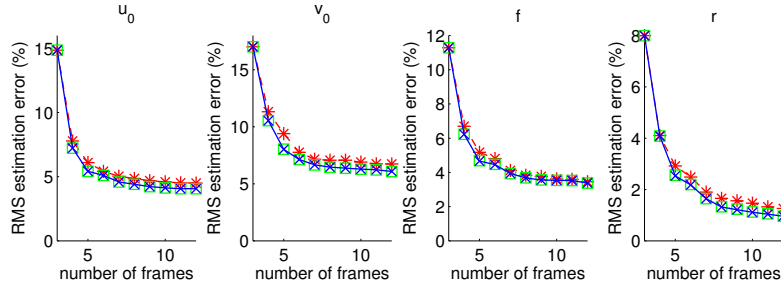
(b) RMS estimation error for each parameter ($\lambda = 50$ pixels).
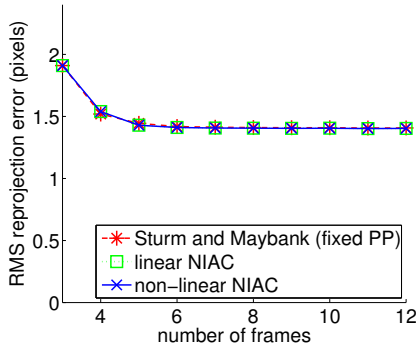
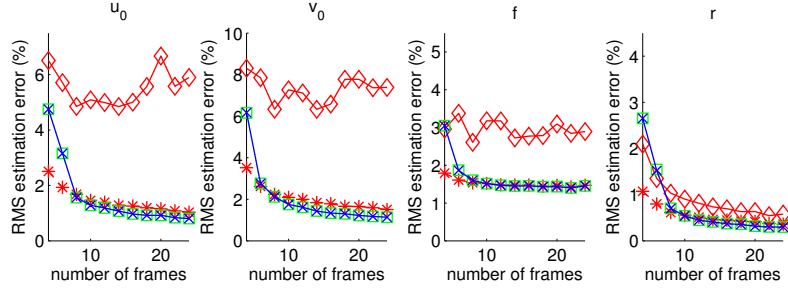(c) RMS reprojection error ($\lambda = 5$ pixels).

(d) RMS reprojection error ($\lambda = 50$ pixels).

Fig. 7. Results with synthetic data in the case of independent images (scenario 1). Two different amplitudes are considered for the PP motion: $\lambda = 5$ pixels and $\lambda = 50$ pixels.

unacceptably large. However, in the case of a modest range of PP motion, the benefit of introducing a variable PP seems to be outbalanced by the deterioration of the conditioning of the system of equations to solve, which makes it acceptable to consider a fixed PP instead.
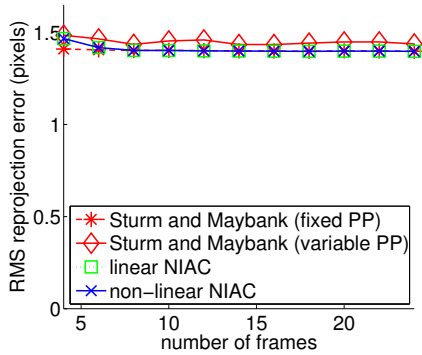
5.2 Real data

The camera used is a Sony DXC-9100P equipped with a Fujinon S12×5BRM-38 zooming lens which has a 5–60 mm focal length range. The lens exhibits very low lens distortion which can be ignored during calibration. We captured a sequence
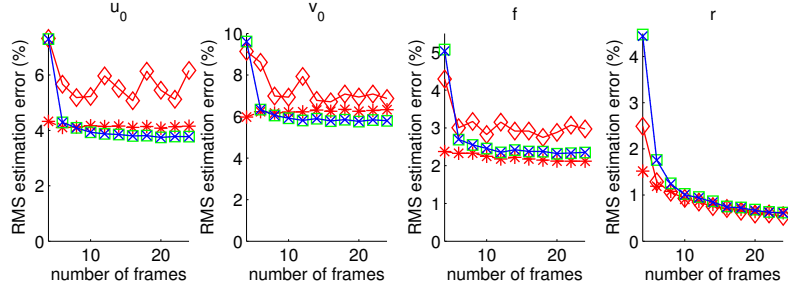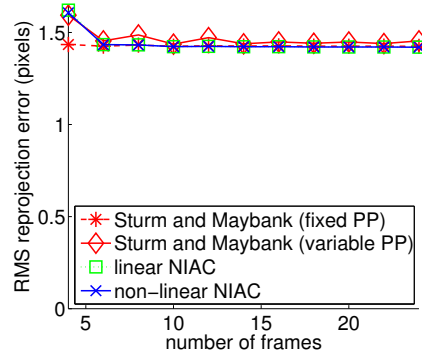
18

(a) RMS estimation error for each parameter ($\lambda = 5$ pixels).

(b) RMS estimation error for each parameter ($\lambda = 50$ pixels).

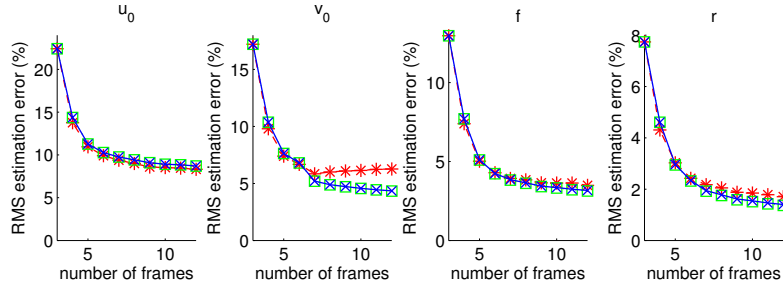

(c) RMS reprojection error ($\lambda = 5$ pixels).

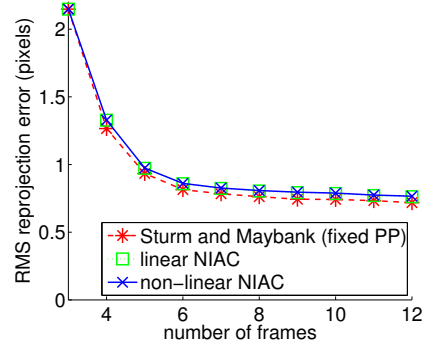(d) RMS reprojection error ($\lambda = 50$ pixels).

Fig. 8. Results with synthetic data in the case of pairs of images (scenario 2). Two different amplitudes are considered for the PP motion: $\lambda = 5$ pixels and $\lambda = 50$ pixels.

of 30 images of the grid shown in Fig. 6. The camera is hand-held, and the zoom settings are changed manually by the operator holding the camera. Each group of five successive images were acquired with a constant zoom setting, varying only the position and orientation of the camera. We show in Fig. 6 one image for each of the six zoom settings. The method described in [23] is applied to groups of five images acquired with the same zoom factor in order to compute the ground truth camera intrinsic parameter values. Then the calibration experiments are repeated 1000 times, selecting a random subset of images each time. The results obtained are shown in Fig. 9 (independent images) and Fig. 10 (pairs of images).

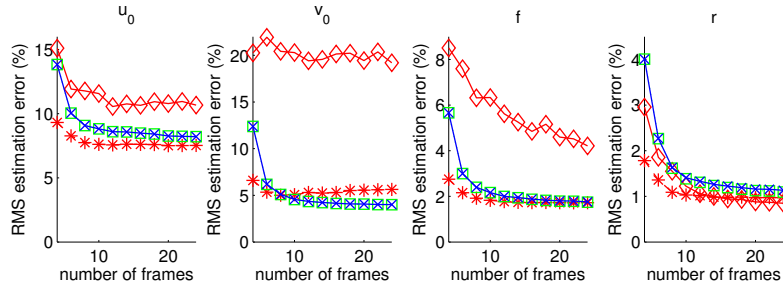The conclusions that can be drawn from these experiments are similar to the ones

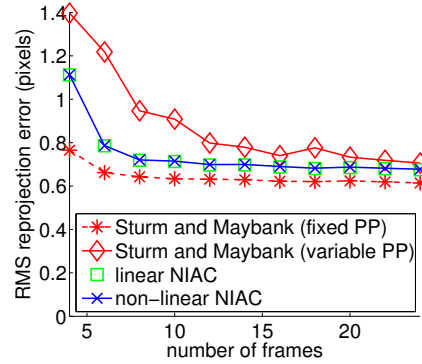(a) RMS estimation error for each parameter.

(b) RMS reprojection error.

Fig. 9. Results with real data in the case of independent images (scenario 1).

(a) RMS estimation error for each parameter.

(b) RMS reprojection error.

Fig. 10. Results with real data in the case of pairs of images (scenario 2).

made in the case of synthetic data. In the case of independent images, all methods appear to perform closely. In the case of pairs of images, the *Sturm & Maybank (fixed PP)* method is the best performing method, followed closely by the *linear NIAC* and *non-linear NIAC* which seem to perform equally well, and then by the *Sturm & Maybank (variable PP)* which is the least accurate. For this particular camera, the introduction of the variable PP does not seem beneficial. This suggests that the range of PP motion is not excessively large, in which case the errors due to the increase in complexity of the system of equations and the deterioration of the conditioning of the system are larger than the gain in accuracy due to modelling the PP motion in the *Sturm & Maybank (variable PP)* method. It is not clear why the *Sturm & Maybank (fixed PP)* method performs slightly better than the NIAC method in this particular experiment. In any case, it should be pointed out that the differences of performance between the different methods are relatively small.

It has been chosen here to restrict the experiments to cameras presenting zero-skew, because this is the case of most cameras encountered in real life. It is worth noting however that if this is not the case, the *non-linear NIAC* method usually becomes the most accurate (see experiments reported in [41]) because it is the only one able to handle camera skew. Although this may appear to be of limited interest given the current camera technologies, we believe this is a nice feature, which could become useful in situations where non-conventional sensors are considered.

6   Conclusions

Firstly, we demonstrated that the image plane error induced by assuming a fixed principal point (PP) model remains small as long as the product of the amplitude of the PP variation and the relative depth relief is small. This is supported by an experimental verification. The implication for camera calibration is that assuming a fixed PP can be acceptable as long as the previous condition is satisfied.

Secondly, we introduced a new mathematical object called the Normalised Image of the Absolute Conic (NIAC), and proposed a novel plane based calibration technique for a zooming camera based on this concept. The NIAC is a mathematical representation of the intrinsic camera parameters invariant to zooming, translation and rotation of the camera. In practice, it can be estimated from a minimum of three images of a planar calibration grid taken from arbitrary positions, orientations and zoom settings. A simplified linear algorithm and a non-linear algorithm have been proposed.

Results with synthetic and real images showed that the two algorithms are competitive with existing state-of-the-art plane-based camera calibration methods. The performance reported are similar to that of other methods assuming a fixed PP, and, as long as the range of PP variation is modest, better than that of the Sturm and

Maybank method assuming a variable PP.

The NIAC method presents the following advantages compared to other calibration methods: i) fixed computational cost due to the use of invariants to decouple the camera parameters, ii) minimisation of an exact geometric distance in the case of the non-linear algorithm, iii) a simpler linear method minimising a very good approximation of the geometric distance, iv) increased generality with the possibility to include the camera skew (in practice, circumstances requiring camera skew computation are rare).

An apparent limitation of the method is the assumption of a fixed PP. We would like to point out that there are common circumstances (calibration of a zooming camera from a single moving calibration plane) where this is the only possible model. Otherwise, these parameters could be included in a final bundle-adjustment where PP motion, as well as lens distortion if required, are computed. In this case, the NIAC method can be used to provide a good initialisation to the non-linear optimisation algorithm required for the bundle-adjustment.

A   Equation of the perpendicular bissector to a chord on the IAC

We consider a pair of images of circular points $\boldsymbol{P} = \boldsymbol{H}\boldsymbol{I} = \boldsymbol{h}_1+\mathrm{i}\boldsymbol{h}_2$ and $\boldsymbol{Q} = \boldsymbol{H}\boldsymbol{J} = \boldsymbol{h}_1 - \mathrm{i}\boldsymbol{h}_2$ on the the IAC, with $\boldsymbol{h}_1 = [h_{11}, h_{21}, h_{31}]^\top$ and $\boldsymbol{h}_2 = [h_{12}, h_{22}, h_{32}]^\top$. We want to compute the equation of the perpendicular bissector to this chord after the image transformation T has been applied.

We first observe that the mid-point of $[\boldsymbol{P}\boldsymbol{Q}]$ is the point:

$$\boldsymbol{M} \sim \frac{1}{h_{31}^2 + h_{32}^2}(h_{31}\boldsymbol{h}_1 + h_{32}\boldsymbol{h}_2) = [m_1 \quad m_2 \quad 1]^\top, \qquad (\text{A.1})$$

and the direction of the line $(\boldsymbol{P}\boldsymbol{Q})$ is represented by the point at infinity:

$$\boldsymbol{D} \sim h_{32}\boldsymbol{h}_1 - h_{31}\boldsymbol{h}_2 = [d_1 \quad d_2 \quad 0]^\top. \qquad (\text{A.2})$$

This can be verified by noting that the points $\boldsymbol{M}$, $\boldsymbol{D}$, $\boldsymbol{P}$ and $\boldsymbol{Q}$ are aligned (they are linear combinations of the base vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$) and that they are harmonic (their cross ratio is -1). It should be noted that the denominator in the expression of M is non-zero if and only if the optical axis of the camera is not orthogonal to the image plane. Four new parameters m_1, m_2, d_1 and d_2 have been introduced in the two previous equations.

After transformation by T, $\boldsymbol{M}$ and $\boldsymbol{D}$ are mapped respectively into

$$\boldsymbol{M'} = T[m_1 \quad m_2 \quad 1]^\top = [m_1 + t_1 m_2 + t_2 \quad t_3 m_2 + t_4 \quad 1]^\top, \qquad \text{(A.3)}$$

$$\boldsymbol{D'} = T[d_1 \quad d_2 \quad 0]^\top = [d_1 + t_1 d_2 \quad t_3 d_2 \quad 0]^\top. \qquad \text{(A.4)}$$

Noting that a normal vector to $(\boldsymbol{P'Q'})$ is $\boldsymbol{N} = [-t_3 d_2, d_1 + t_1 d_2, 0]^\top$, we conclude that the perpendicular bisector to the chord is represented by the equation

$$\boldsymbol{l} = [-(d_1 + t_1 d_2) \quad -t_3 d_2 \quad (m_1 + t_1 m_2 + t_2)(d_1 + t_1 d_2) + (t_3 m_2 + t_4)t_3 d_2]^\top. \quad \text{(A.5)}$$

References

[1] K. Tarabanis, R. Y. Tsai, D. S. Goodman, Calibration of a computer-controlled robotic vision sensor with a zoom lens, Computer Vision, Graphics, and Image Processing 59 (2) (1994) 226–241.

[2] Y.-S. Chen, S.-W. Shih, Y.-P. Hung, C.-S. Fuh, Simple and efficient method of calibrating a motorized zoom lens, Image and Vision Computing 19 (14) (2001) 1099–1110.

[3] R. G. Willson, S. Shafer, What is the center of the image?, Journal of Optical Society of America A 11 (11) (1994) 2946–2955.

[4] M. T. Ahmed, A. A. Farag, A neural approach to zoom-lens camera calibration from data with outliers, Image and Vision Computing 20 (9-10) (2002) 619–630.

[5] O. D. Faugeras, Q.-T. Luong, S. J. Maybank, Camera self-calibration: Theory and experiments, in: Proc. European Conference on Computer Vision, 1992, pp. 321–334.

[6] M. Pollefeys, R. Koch, L. J. Van Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, in: Proc. IEEE International Conference on Computer Vision, 1998, pp. 90–95.

[7] P. F. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 1100–1105.

[8] P. F. Sturm, Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length, Image and Vision Computing 20 (5-6) (2002) 415–426.

[9] L. de Agapito, E. Hayman, I. D. Reid, Self-calibration of rotating and zooming cameras, International Journal of Computer Vision 45 (2) (2001) 107–127.

[10] Y. Seo, K. S. Hong, About the self-calibration of a rotating and zooming camera: Theory and practice, in: Proc. IEEE International Conference on Computer Vision, Vol. 1, 1999, pp. 183–189.

[11] P. F. Sturm, Self-calibration of a moving zoom-lens camera by pre-calibration, Image and Vision Computing 15 (8) (1997) 583–589.

[12] S. Bougnoux, From projective to Euclidean space under any practical situation, a criticism of self-calibration, in: Proc. IEEE International Conference on Computer Vision, 1998, pp. 790–796.

[13] B. Caprile, V. Torre, Using vanishing points for camera calibration, International Journal of Computer Vision 4 (2) (1990) 127–140.

[14] L. L. Wang, W. H. Tsai, Computing camera parameters using vanishing-line information from a rectangular parallelepiped, Machine Vision and Applications 3 (3) (1990) 129–141.

[15] T. Echigo, A camera calibration technique using three sets of parallel lines, Machine Vision and Applications 3 (3) (1990) 159–167.

[16] W. Chen, B. C. Jiang, 3-D camera calibration using vanishing point concept, Pattern Recognition 24 (1) (1991) 57–67.

[17] L. L. Wang, W. H. Tsai, Camera calibration by vanishing lines for 3-d computer vision, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (4) (1991) 370–376.

[18] P. Beardsley, D. Murray, Camera calibration using vanishing points, in: Proc. British Machine Vision Conference, 1992, pp. 416–425.

[19] R. Cipolla, D. P. Robertson, E. G. Boyer, Photobuilder - 3D models of architectural scenes from uncalibrated images, in: Proc. IEEE International Conference on Multimedia Computing and Systems, Vol. 1, 1999, pp. 25–31.

[20] D. Liebowitz, A. Criminisi, A. Zisserman, Creating architectural models from images, in: Proc. Eurographics, Vol. 18, 1999, pp. 39–50.

[21] J.-Y. Guillemaut, A. S. Aguado, J. Illingworth, Using points at infinity for parameter decoupling in camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2) (2005) 265–270.

[22] D. Liebowitz, A. Zisserman, Metric rectification for perspective images of planes, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1998, pp. 482–488.

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

[24] P. F. Sturm, S. J. Maybank, On plane-based camera calibration: A general algorithm, singularities, applications, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1999, pp. 432–437.

[25] X. Meng, Z. Hu, A new easy camera calibration technique based on circular points, Pattern Recognition 36 (5) (2003) 1155–1164.

[26] C. Matsunaga, K. Kanatani, Calibration of a moving camera using a planer pattern: Optimal computation, reliability evaluation and stabilization by model selection, in: Proc. European Conference on Computer Vision, Vol. II, 2000, pp. 595–609.

[27] P. Gurdjos, R. Payrissat, Plane-based calibration of a camera with varying focal length: the centre line constraint, in: Proc. British Machine Vision Conference, Vol. 2, 2001, pp. 623–632.

[28] P. Gurdjos, A. Crouzil, R. Payrissat, Another way of looking at plane-based calibration: the centre circle constraint, in: Proc. European Conference on Computer Vision, Vol. IV, 2002, pp. 252–266.

[29] Z. Zhang, Camera calibration with one-dimensional objects, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (7) (2004) 892–899.

[30] P. Hammarstedt, P. Sturm, A. Heyden, Degenerate cases and closed-form solutions for camera calibration with one-dimensional objects, in: Proc. IEEE International Conference on Computer Vision, Vol. 1, 2005, pp. 317–324.

[31] F. C. Wu, Z. Y. Hu, H. J. Zhu, Camera calibration with moving one-dimensional objects, Pattern Recognition 38 (5) (2005) 755–765.

[32] F. Qi, Q. Li, Y. Luo, D. Hu, Camera calibration with one-dimensional objects moving under gravity, Pattern Recognition 40 (1) (2007) 343–345.

[33] F. Qi, Q. Li, Y. Luo, D. Hu, Constraints on general motions for camera calibration with one-dimensional objects, Pattern Recognition 40 (6) (2007) 1785–1792.

[34] R. G. Willson, S. Shafer, A perspective projection camera model for zoom lenses, in: Proc. Conference on Optical 3-D Measurement Techniques, 1993.

[35] R. G. Willson, Modeling and calibration of automated zoom lenses, in: Proc. SPIE #2350: Videometrics III, 1994, pp. 170–186.

[36] R. G. Willson, Modeling and calibration of automated zoom lenses, Ph.D. thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (January 1994).

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

[38] J.-Y. Guillemaut, A. S. Aguado, J. Illingworth, Calibration of a zooming camera using the Normalized Image of the Absolute Conic, in: Proc. International Conference on 3-D Digital Imaging and Modeling, 2003, pp. 225–232.

[39] O. D. Faugeras, Q.-T. Luong, The geometry of multiple images, The MIT Press, 2001.

[40] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes in C, 2nd Edition, Cambridge University Press, 1992.

[41] J.-Y. Guillemaut, Contributions to image-based object reconstruction: geometric and photometric aspects, Ph.D. thesis, School of Electronics and Physical Sciences, University of Surrey, Guildford, Surrey, UK (September 2005).