Machine Vision and Applications (2000) 12: 84–97

Machine Vision and Applications

Algorithms for coplanar camera calibration

Chanchal Chatterjee[1], **Vwani P. Roychowdhury**[2]

[1] Tiernan Communications, 5751 Copley Drive, San Diego, CA 92111, USA; Fax: (858) 587-0257, e-mail: Cchatterj@hotmail.com
[2] Electrical Engineering Department, 7335C Boelter Hall, UCLA, Los Angeles, CA 90095-1594, USA; Tel: (310) 206-4975, Fax: (310) 794-1592, e-mail: vwani@ee.ucla.edu

Received: 30 September 1998 / Accepted: 12 January 2000

Abstract. Coplanar camera calibration is the process of determining the extrinsic and intrinsic camera parameters from a given set of image and world points, when the world points lie on a two-dimensional plane. Noncoplanar calibration, on the other hand, involves world points that do not lie on a plane. While optimal solutions for both the camera-calibration procedures can be obtained by solving a set of constrained nonlinear optimization problems, there are significant structural differences between the two formulations. We investigate the computational and algorithmic implications of such underlying differences, and provide a set of efficient algorithms that are specifically tailored for the coplanar case. More specifically, we offer the following: (1) four algorithms for coplanar calibration that use linear or iterative linear methods to solve the underlying nonlinear optimization problem, and produce sub-optimal solutions. These algorithms are motivated by their computational efficiency and are useful for real-time low-cost systems. (2) Two optimal solutions for coplanar calibration, including one novel nonlinear algorithm. A constraint for the optimal estimation of extrinsic parameters is also given. (3) A Lyapunov type convergence analysis for the new nonlinear algorithm. We test the validity and performance of the calibration procedures with both synthetic and real images. The results consistently show significant improvements over less complete camera models.

Key words: Camera calibration – Coplanar calibration – Lens distortion – Nonlinear optimization

1 Introduction

There is a vast literature on noncoplanar camera calibration where the world points lie on a three-dimensional (3D) surface. However, not much importance is given to coplanar camera calibration, where the world points lie on a two-dimensional (2D) plane. This is in spite of the fact that there are several industrial and military applications that require coplanar calibration [16,26,27]. In this study, we survey the coplanar camera calibration problem, and in the process present new and provably convergent algorithms for the solution of extrinsic (external camera geometric) and intrinsic (internal camera geometric and camera optics) parameters.

Coplanar camera calibration is used frequently in commercial and military applications. In metrology problems [16,27], we commonly measure the dimensions of small (semiconductor) objects under high magnification, where the part dimensions are much greater than their surface elevations. In remote sensing of the earth's surface [26], the image data is acquired from either an aircraft or a spacecraft platform, whose distance from the earth's surface may be much larger than the surface topography. In both these situations, the calibration problems are essentially coplanar.

An optimal solution of all camera calibration parameters requires nonlinear optimization. In our discussions, "optimality" refers to the orthonormality constraints that the extrinsic (rotation) parameters must satisfy (see Sect. 2). In some respects, the coplanar calibration problem is more complex than the noncoplanar case. While the number of parameters to be estimated is the same in both cases, the number of usable orthonormality constraints is fewer for the coplanar case. We have available only three orthonormality constraints (see Sect. 3) in comparison to six constraints (see Sect. 2) for the noncoplanar case. Besides, some standard techniques in photogrammetry such as the absolute and exterior orientation methods [16,23] can be used to solve the noncoplanar extrinsic parameters (see [8]), while such methods are inapplicable in the coplanar case.

In this study, we extend a number of linear and nonlinear methods of noncoplanar calibration to the coplanar case. The extensions involve novel steps that are special to the coplanar problem. We also present new linear and nonlinear methods of coplanar calibration. We offer a rigorous proof of convergence for the new nonlinear method. In summary, we describe two types of algorithms for coplanar camera calibration: (1) four linear and iterative linear algorithms which produce efficient but sub-optimal solutions, and (2) two nonlinear algorithms which produce optimal solutions, and including a novel nonlinear algorithm with an analytical proof of convergence.

Correspondence to: C. Chatterjee

In the next four sections, we review the following aspects of coplanar calibration: (1) image center parameters, (2) scale factor parameter, (3) lens distortion parameters, and (4) methods to solve the coplanar camera calibration problem.

1.1 Image center parameter

Ideally the image center is the intersection of the optical axis of the camera-lens system with the camera's sensing plane. For real lenses, optical axis is not so easily defined, and different definitions of image centers [30] depend on whether the lens has fixed or variable parameters, and on how the variable parameters are mechanically implemented. Examples are (1) for a simple lens, there can be two axes of symmetry – optical and mechanical. The optical axis is the straight line joining the centers of curvature of the two surfaces of the lens, whereas the mechanical axis is determined by the centerline of the machine used to grind the lens' edge. The angle between these axes is called *decentration* [30]. (2) In a compound lens, the optical axes of multiple lens elements may not be accurately aligned due to decentration of each lens element, resulting in multiple possibilities for the optical axis. (3) In adjustable and variable focal length lenses, the misalignment between the optical and mechanical axes change as the spacing between the lens elements are changed.

There are both offline and algorithmic methods of computing the image center parameters. Some offline methods are: (1) measuring the center of the radial lens distortion [20,30], (2) determining the normal projection of a viewing point onto the imaging plane [28,30], (3) measuring the center of the camera's field of view [30], (4) passing a laser beam through the lens assembly and matching the reflection of the beam from the lens with the center of the light spot in the image [20,30], (5) measuring the center of $\cos^{4th}$ radiometric falloff or the center of vignetting/image spot [30], (6) changing the focal length of a camera-lens system to determine image center from a point invariant in the image [20,30]. Most algorithmic methods [7,8,10,11,15] usually use the orthonormality condition of the extrinsic parameters or use nonlinear minimization [20,29,30] to compute the image center. Since a smaller number of constraints are available in the coplanar case, we cannot compute the image center and scale factor parameters from the orthonormality conditions alone.

1.2 Scale factor parameters

Another intrinsic parameter commonly considered is the scale factor s. Array sensors such as CCD/CID sensors acquire the video information line by line, where each line of video signal is well separated by the horizontal sync of the composite video. Usually, the vertical spacing between lines perfectly matches that on the sensor array, giving us no scale factor in the vertical direction. The pixels in each line of video signal are re-sampled by the ADC, which in reality, samples the video lines with a rate different than the camera and causes the image to be scaled along the horizontal direction, i.e., $s \neq 1$. Hence, the problem of determining scale factor s.

Some researchers [20,27,30] suggest that s can be approximately determined from the ratio of the number of sensor elements in the image column direction to the number of pixels in a line as sampled by the processor. However, due to timing errors, inconsistency of the ADC, and possible tilt of the sensor array, this is not so accurate. A more accurate estimate of s is the ratio of the frequency that sensor elements are clocked off of the CCD to the frequency at which the ADC samples.

We found a number of methods to compute the scale factor parameter by offline techniques. These methods are (1) measuring the frequency of the stripes generated by the interface of ADC-clock and camera-clock that create the scale factor problem [20], (2) measuring the distortion in an image of a perfect circle into an ellipse [25], (3) computing the power spectrum of the image of two sets of parallel lines [1], (4) counting the grid points in an image of a grid pattern [6].

All of the above-mentioned methods can be used to obtain initial estimates of the image center and scale factor parameters for the algorithms discussed here. In addition, we have given algorithmic methods to obtain the image center and scale factor parameters as a part of the parameter estimation procedure.

1.3 Lens distortion parameters

Another intrinsic parameter commonly studied is the effect of image distortion due to a nonlinear lens system. Many researchers [17,20,22,29] have observed that ignoring lens distortion is unacceptable in doing measurements. However, most of these studies are for the noncoplanar case only. For example, Beyer [3] demonstrated the effects of higher order radial and tangential distortion models in a noncoplanar setup. By using a first-order radial model, accuracy in image space of $1/7^{th}$ of the pixel spacing is obtained, and by using a third-order radial and first-order decentering distortion model, this accuracy is enhanced to $1/46^{th}$ of the pixel spacing. Faugeras and Toscani [10] and Weng et al. [29] used wide-angle lenses and also found that adding nonradial distortion components improved accuracy for noncoplanar calibration. In this study, we include radial, tangential and thin-prism lens distortions in the coplanar calibration model.

1.4 Methods for coplanar camera calibration

In the literature of camera calibration, there are plenty of methods of noncoplanar camera calibration that can be extended to the coplanar case. However, the extensions involve novel steps pertaining to the coplanar problem. Most of these methods transform the collinearity condition equations such that the parameters are estimated by solving linear equations. Many earlier studies using this approach primarily considered the extrinsic parameters, although some intrinsic parameters, such as image center, are considered only for the noncoplanar case. These methods are usually efficient. However, the solutions are sub-optimal and they generally

ignore nonlinear lens distortion. Yakimovsky and Cunningham [32] and Ganapathy [11] used the pin-hole model and treated some combinations of parameters as single variables in order to formulate the problem as a linear system. However, in this formulation, the variables are not completely linearly independent, yet are treated as such.

Grosky and Tamburino [15] used a linear method after correcting lens distortion with a bivariate polynomial model. The coplanar case is considered, although it is necessary for the user to supply up to three additional constraints. Tsai [27] and Lenz and Tsai [20] considered a two-stage algorithm using a radial alignment constraint (RAC) in which most parameters are computed in closed form. A small number of parameters such as the focal length, depth component of the translation vector, and radial lens distortion parameters are computed by an iterative scheme. If the image center is unknown, it is determined by a nonlinear approach [20] based on minimizing the RAC residual error. Although the solution is efficient, and radial lens distortion parameters are computed, the method does not estimate tangential lens distortion parameters. Furthermore, by taking the ratio of the collinearity condition equations, the method discards radial *information* in these equations. For example, Weng et al. [29] suggest that ignoring radial information can result in a less reliable estimator. Moreover, the solution is sub-optimal. However, Lenz and Tsai [20] have offered one of the few algorithms for the coplanar case.

Analytical photogrammetry offers various methods [4,5, 9,13,16,18,23] to solve the noncoplanar calibration problem, most of which are based on a parametric recursive procedure known as the nonlinear least squares method [16,23]. Using the Euler angles for the rotation matrix (see Eq. 38), two collinearity condition equations are obtained for each observation. The nonlinear equations are linearized using Newton's first-order approximation. The linear equations are solved and updated by an iterative procedure. Several iterations of the solution must be made to eliminate errors due to the linearization procedure. Initial estimates are a prerequisite for this recursive procedure to succeed. Faig [9], Wong [31] and Malhotra and Karara [18] used the above method for a general solution of all noncoplanar calibration parameters. The direct linear transformation method of Abdel-Aziz and Karara [18] also uses this technique. Similar formulations are used by Beyer [3] and Weng et al. [29]. These methods yield optimal solutions and very accurate parameter estimates. However, many researchers [15,17,27] observe that these methods require a considerable computational effort. We extend the photogrammetric formulation to the coplanar calibration problem.

In Sect. 2, we describe the camera calibration model. Section 3 describes the linear and iterative linear methods of coplanar camera calibration. Section 4 discusses two nonlinear methods of coplanar calibration including a new method with its analytical proof of convergence. Section 5 has the experimental results. Section 6 has the concluding remarks.

2 Camera calibration model

We discuss the camera calibration model consisting of extrinsic and intrinsic parameters. The extrinsic parameters consist of a 3×3 rotation matrix R which defines the camera orientation, and the 3×1 translation vector $\mathbf{t}$ which defines the position of the camera center. The intrinsic parameters consist of the effective focal length f of the camera, center of the image array (i_0, j_0), horizontal scale factor s of the image array, radial lens distortion parameters $\{k_1, k_2, \ldots, k_{r_0}\}$, tangential lens distortion parameters $\{p_1, p_2\}$, and thin-prism lens distortion parameters $\{s_1, s_2\}$.

Table 1. Camera calibration parameters discussed in this study

Parameters	Type	Description
R	Extrinsic	Rotation matrix for camera orientation
$\mathbf{t}$		Translation vector for position of camera center
f	Intrinsic	Focal length
i_0, j_0	Intrinsic	Image center displacement
s		Scale factor
$k_1, \ldots, k_{r_0}$	Intrinsic	Radial lens distortion
p_1, p_2		Decentering lens distortion
s_1, s_2		Thin-prism lens distortion

2.1 Extrinsic parameters and focal length

Here, we describe the geometry of the calibration system. Since the world points lie on a 2D plane, we assume, without loss of generality, that it is the z-coordinate that is unimportant. The geometry involves three coordinate systems (see Fig. 1): (1) a world coordinate system (X_w, Y_w, Z_w) centered around a point O_w and including a point (x, y) lying on the (X_w, Y_w) plane, (2) a camera coordinate system (X_c, Y_c, Z_c) with origin at optical center O_c with Z_c-axis the same as the optical axis, and (3) a 2D image array system (I, J) centered at a point O_s in the image plane, with (I, J) axes aligned with (X_c, Y_c), respectively, and including a point (i, j). Let f be the effective focal length of the camera. The collinearity condition equations are [13,18,23]

$$f\left(\frac{r_{11}x + r_{12}y + t_1}{r_{31}x + r_{32}y + t_3}\right) = i, \quad \text{and} \tag{1}$$

$$f\left(\frac{r_{21}x + r_{22}y + t_2}{r_{31}x + r_{32}y + t_3}\right) = j, \tag{2}$$

where

$$R = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{bmatrix} = \begin{bmatrix} r_{11}\ r_{12}\ r_{13} \\ r_{21}\ r_{22}\ r_{23} \\ r_{31}\ r_{32}\ r_{33} \end{bmatrix} \tag{3}$$

is a 3×3 rotation matrix which defines the camera orientation, and $\mathbf{t}^T = [t_1 t_2 t_3]$ is a translation vector which defines the position of the camera center. An important constraint in calibration algorithms is the *orthonormality constraint* of the rotation matrix R given by

$$R^T R = R R^T = I. \tag{4}$$

The extrinsic and focal length parameters to be calibrated are $\mathbf{b} = \{R, \mathbf{t}, f\}$.

2.2 Lens distortion parameters

As a result of imperfections in the design and assembly of lenses, the image of a plane object lies, in general, on a

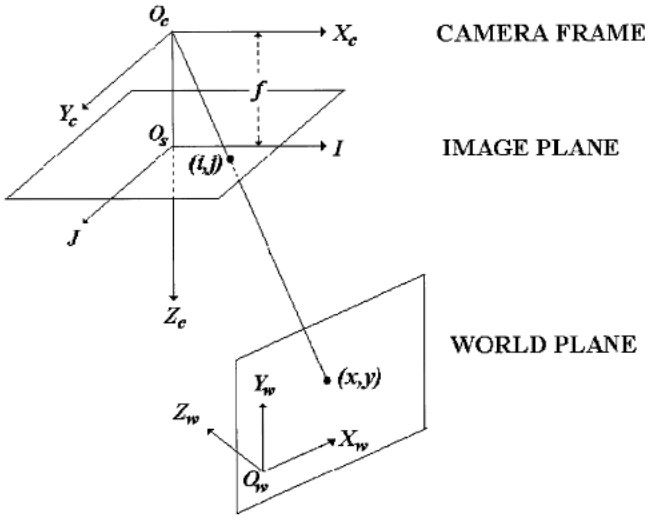

Fig. 1. Mapping of world point (x, y) to image point (i, j)

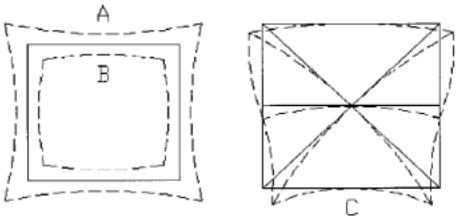

Fig. 2. Lens distortions: (**A**) radial pincushion, (**B**) radial barrel, and (**C**) tangential

slightly curved field [23] (see Fig. 2), wherein objects at the edge of the field of view appear somewhat smaller or larger than they should. Types of lens distortions commonly seen are radial [4,23] and tangential [5,23]. Two common radial distortions are pincushion and barrel distortions. Pincushion distortion results, for example, when a lens is used as a magnifying glass, whereas barrel distortion results when the object is viewed through a lens at some distance from the eye. Tangential distortions are usually caused by (a) decentering of the lens (decentering distortion) [4,5,9,24,29], and (**b**) imperfections in lens manufacturing or tilt in camera sensor or lens (thin-prism distortion) [9,29]. One of the effects of tangential distortion is that a straight line passing through the center of the field of view may appear in the image as weakly curved line (see Fig. 2). Clearly, these distortions are disturbing in applications where the ultimate task is to map a 2D object in uniform scale from its acquired image.

One commonly used model for correcting lens distortion is that developed by Brown [4,5]. Let (δ_i, δ_j) be the corrections for geometric lens distortions present in distorted image coordinates (i_d, j_d). Let (i, j) be the ideal undistorted image coordinates of a 2D point (x, y). With $r_{\mathbf{d}}^2 = i_{\mathbf{d}}^2 + j_{\mathbf{d}}^2$, δ_i and δ_j are expressed by the following series [4,5,23,29]:

$$\begin{aligned}\delta_i = i_d &\left(k_1 r_d^2 + k_2 r_d^4 + \ldots + k_{r_0} r_d^{2r_0}\right) \\ &+ \left(p_1(r_d^2 + 2i_d^2) + 2p_2 i_d j_d\right)\left(1 + p_3 r_d^2 + \ldots\right) \\ &+ \left(s_1 r_d^2 + \ldots\right), \end{aligned} \tag{5}$$

$$\begin{aligned}\delta_j = j_d &\left(k_1 r_d^2 + k_2 r_d^4 + \ldots + k_{r_0} r_d^{2r_0}\right) \\ &+ \left(2p_1 i_d j_d + p_2(r_d^2 + 2j_d^2)\right)\left(1 + p_3 r_d^2 + \ldots\right) \\ &+ \left(s_2 r_d^2 + \ldots\right). \end{aligned} \tag{6}$$

Here, r_0 is the order of the radial distortion model. Terms including coefficients $\{k_1, k_2, k_3, \ldots, k_{r_0}\}$ account for radial distortion, $\{p_1, p_2, p_3, \ldots\}$ represent decentering distortions, and $\{s_1, s_2, \ldots\}$ represent thin-prism distortions. We define the distortion vector $\mathbf{d}$ as

$$\mathbf{d}^T = [\, k_1 \; k_2 \; \ldots \; k_{r_0} \; p_1 \; p_2 \; s_1 \; s_2 \,]. \tag{7}$$

Image coordinates are corrected for lens distortion by the expressions below:

$$i = i_d + \mathbf{w}^T \mathbf{d}, \text{ where}$$
$$\mathbf{w} = \left[\, i_d r_d^2 \; i_d r_d^4 \; \ldots \; i_d r_d^{2r_0} \; r_d^2 + 2i_d^2 \; 2i_d j_d \; r_d^2 \; 0 \,\right], \text{ and} \tag{8}$$
$$j = j_d + \tilde{\mathbf{w}}^T \mathbf{d}, \text{ where}$$
$$\tilde{\mathbf{w}} = \left[\, j_d r_d^2 \; j_d r_d^4 \; \ldots \; j_d r_d^{2r_0} \; 2i_d j_d \; r_d^2 + 2j_d^2 \; 0 \; r_d^2 \,\right]. \tag{9}$$

The lens distortion parameter to be calibrated is $\mathbf{d}$.

2.3 Image center and scale factor parameters

Consider a frame buffer image point (i_f, j_f) with respect to the center O_s of the image buffer. Let the actual image center be at (i_0, j_0). Let (i_d, j_d) be the location of the distorted image point with respect to (i_0, j_0). Let s be the horizontal scale factor of the image. Then, we obtain [10,11,20,27,29]

$$i_d = s^{-1}\left(i_f - i_0\right) \text{ and } j_d = \left(j_f - j_0\right). \tag{10}$$

The parameters to be calibrated are $\mathbf{m} = \{i_0, j_0, s\}$.

2.4 Discussion

Figure 3 below provides the steps for transforming the frame buffer image point (i_f, j_f) to the ideal undistorted image point (i, j) and to the 2D world point (x, y).

Given frame buffer image coordinates (i_f, j_f), the distorted image coordinates (i_d, j_d) obtained from Eq. 10 are functions of the image center and scale factor parameters $\mathbf{m}$, i.e.,

$$i_d = i_d(\mathbf{m}) \text{ and} j_d = j_d(\mathbf{m}). \tag{11}$$

Finally, the ideal undistorted image coordinates (i, j) obtained from (i_d, j_d) by applying Eqs. 8 and 9 are functions of the lens distortion parameters $\mathbf{d}$. Thus, the complete transformation from frame buffer coordinates (i_f, j_f) to ideal undistorted coordinates (i, j) is a function of both $\mathbf{m}$ and $\mathbf{d}$:

$$i = i(\mathbf{d}, \mathbf{m}) \text{ and } j = j(\mathbf{d}, \mathbf{m}). \tag{12}$$

3 Linear and iterative linear algorithms for coplanar calibration

Since the world points lie on a 2D plane, we assume, without loss of generality, that it is the z-coordinate of the world points that is unimportant. Due to this assumption, the last column of the rotation matrix R is not available

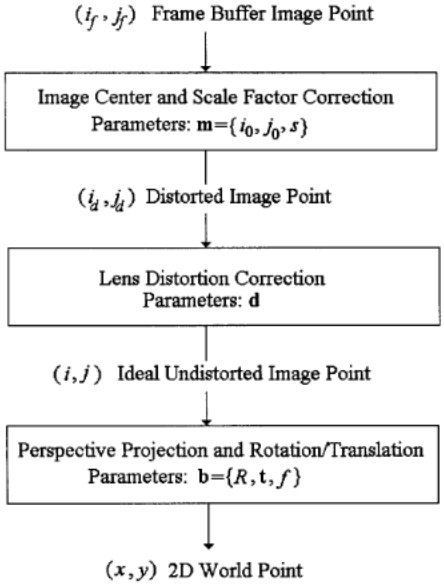

Fig. 3. Steps for transforming frame buffer image point (i_f, j_f) to ideal undistorted image point (i, j) and to 2D world point (x, y)

in the collinearity condition Eqs 1 and 2. Instead of all the orthonormality constraints described in Eq. 4, we have available three constraints:

$$r_{11}^2 + r_{21}^2 + r_{31}^2 = 1, r_{12}^2 + r_{22}^2 + r_{32}^2 = 1,$$
$$r_{11}r_{12} + r_{21}r_{22} + r_{31}r_{32} = 0. \tag{13}$$

We discuss four methods of parameter computation by solving linear equations. The first method is due to Tsai [27], and uses the ratio of the two collinearity condition Eqs. 1 and 2. The remaining three methods utilize both collinearity condition equations. These are coplanar extensions of Ganapathy's [11], Grosky and Tamburino's [15], and Chatterjee *et al.* [8] noncoplanar solutions. Table 2 shows the parameters computed by these methods.

In our discussion of the algorithms, some of the steps may fail for the degenerate case $r_{31} = r_{32} = 0$. This case will happen if the camera axis is perpendicular to the world $X_w - Y_w$ plane. A slight tilt in the camera axis with respect to the $X_w - Y_w$ plane will avoid this condition. Thus, we assume $r_{31} \neq 0$ and $r_{32} \neq 0$ in all our algorithms. If these conditions do not hold, then a unique solution of f and t_3 may not exist in steps 5 and 6 of Tsai's algorithm, and step 2 of Ganapathy's and Grosky-Tamburino's algorithms. Instead, a solution for f/t_3 can be obtained.

3.1 Tsai's method

Tsai [27] designed a method that includes the radial lens distortion parameters in his camera model and provided a two-stage algorithm for estimating the camera parameters. In his method, most parameters are computed in a closed form, and a small number of parameters such as focal length, the depth component of the translation vector, and the radial lens distortion parameters are computed by an iterative scheme.

Tsai introduced an RAC, which results from the observation that the vectors (i_d, j_d), (i, j), and $(r_{11}x+r_{12}y+t_1, r_{21}x+r_{22}y+t_2)$ in Fig. 1 are radially aligned when the image center is chosen correctly. Algebraically, the RAC states that

$$(i_d, j_d)//(i, j)//(r_{11}x + r_{12}y + t_1, r_{21}x + r_{22}y + t_2), \tag{14}$$

where // denotes parallel vectors. Here, we assume that the image center (i_0, j_0) and scale factor s are either ignored or known in advance by any one of the offline methods described in Sect. 1. The frame buffer image coordinates (i_f, j_f) can be corrected for these parameters by Eq. 10 to obtain the distorted image coordinates $i_d = s^{-1}(i_f - i_0)$ and $j_d = j_f - j_0$. The method considers radial lens distortion only. By applying Eqs. 8 and 9 to the collinearity condition equations, we obtain

$$f\left(\frac{r_{11}x + r_{12}y + t_1}{r_{31}x + r_{32}y + t_3}\right) = i = i_d\left(1 + k_1 r_d^2 + \ldots + k_{r_0} r_d^{2r_0}\right)$$
$$= i_d\left(1 + \mathbf{v}_d^T \mathbf{d}\right), \tag{15}$$

and

$$f\left(\frac{r_{21}x + r_{22}y + t_2}{r_{31}x + r_{32}y + t_3}\right) = j = j_d\left(1 + k_1 r_d^2 + \ldots + k_{r_0} r_d^{2r_0}\right)$$
$$= j_d\left(1 + \mathbf{v}_d^T \mathbf{d}\right), \tag{16}$$

where $r_d^2 = i_d^2 + j_d^2$, $\mathbf{v}_d^T = [r_d^2 \; r_d^4 \; \ldots \; r_d^{2r_0}]$, and $\mathbf{d}^T = [k_1 \; k_2 \; \ldots \; k_{r_0}]$ are the radial lens distortion parameters. Then, by dividing the first equation by the second, we obtain

$$\frac{r_{11}x + r_{12}y + t_1}{r_{21}x + r_{22}y + t_2} = \frac{i_d}{j_d}. \tag{17}$$

From Eq. 15, we obtain the following objective function:

$$J(\mathbf{b}) = \sum_{all \; calib.points} (r_{11}j_d x + r_{12}j_d y + j_d t_1 - r_{21}i_d x$$
$$- r_{22}i_d y - i_d t_2)^2, \tag{18}$$

where $\mathbf{b} = \{r_{11}, r_{12}, r_{21}, r_{22}, t_1, t_2\}$ are the extrinsic parameters. Tsai's algorithm consists of the following steps.

1. By minimizing J in Eq. 18, we compute the extrinsic parameters from the following linear equation:
$$\left[j_d x \; j_d y \; j_d \; -i_d x \; -i_d y\right] \mathbf{b} = [i_d], \tag{19}$$
where
$$\mathbf{b}^T = \left[\frac{r_{11}}{t_2} \; \frac{r_{12}}{t_2} \; \frac{t_1}{t_2} \; \frac{r_{21}}{t_2} \; \frac{r_{22}}{t_2}\right].$$
We use the linear least squares method to solve for $\mathbf{b}$.
2. Compute t_2 from the following equations derived from the constraints in Eq. 13:
$$|t_2| = \sqrt{\frac{S - \sqrt{S^2 - 4\left(b_1 b_5 - b_4 b_2\right)^2}}{2\left(b_1 b_5 - b_4 b_2\right)^2}}. \tag{20}$$
or
$$|t_2| = \frac{2}{\sqrt{(b_1 + b_5)^2 + (b_2 - b_4)^2} + \sqrt{(b_1 - b_5)^2 + (b_2 + b_4)^2}},$$
where $b_1, \ldots, b_5$ are obtained in Step 1 as
$$b_1 = r_{11}/t_2, \; b_2 = r_{12}/t_2, \; b_3 = t_1/t_2, \; b_4 = r_{21}/t_2,$$
$$b_5 = r_{22}/t_2 \text{ and } S = b_1^2 + b_2^2 + b_4^2 + b_5^2.$$

Table 2. Coplanar calibration parameters computed by the linear methods

Method	Parameters Computed	Parameters Not Computed
Tsai	$R, \mathbf{t}, f, k_1, \ldots, k_{r0}$	$i_0, j_0, s, p_1, p_2, s_1, s_2$
Ganapathy	$R, \mathbf{t}, f$	$i_0, j_0, s, k_1, \ldots, k_{r0}, p_1, p_2, s_1, s_2$
Grosky & Tamburino	$R, \mathbf{t}, f, s$	$i_0, j_0, k_1, \ldots, k_{r0}, p_1, p_2, s_1, s_2$
Chatterjee et al.	$R, \mathbf{t}, f, s, k_1, \ldots, k_{r0}, p_1, p_2, s_1, s_2$	i_0, j_0

3. Determine the sign of t_2. For any (x, y) and corresponding (i_d, j_d), if $Sign(b_1x+b_2y+b_3)$ and $Sign(i_d)$ are same, then $Sign(t_2) = +1$, otherwise $Sign(t_2) = -1$.
4. Determine the extrinsic parameters:

$$r_{11} = t_2 b_1, \;\; r_{12} = t_2 b_2, \;\; r_{13} = -\sqrt{1 - r_{11}^2 - r_{12}^2}, \tag{21}$$

$$r_{21} = t_2 b_4, \;\; r_{22} = t_2 b_5,$$
$$r_{23} = Sign(r_{11}r_{21} + r_{12}r_{22})\sqrt{1 - r_{21}^2 - r_{22}^2},$$

$$r_{31} = r_{12}r_{23} - r_{22}r_{13}, \;\; r_{32} = r_{21}r_{13} - r_{11}r_{23},$$
$$r_{33} = r_{11}r_{22} - r_{21}r_{12}, \;\; t_1 = t_2 b_3.$$

Ignoring lens distortion, i.e., $\mathbf{d} = \mathbf{0}$, compute approximate values of focal length f and depth component t_3 of translation vector from the following equation (derived from Eqs. 15 and 16):

$$\begin{bmatrix} (r_{11}x + r_{12}y + t_1) & -i_d \\ (r_{21}x + r_{22}y + t_2) & -j_d \end{bmatrix} \begin{bmatrix} f \\ t_3 \end{bmatrix} = \begin{bmatrix} i_d(r_{31}x + r_{32}y) \\ j_d(r_{31}x + r_{32}y) \end{bmatrix}. \tag{22}$$

Here, $r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2$ are obtained from Step 4.
5. Compute accurate estimates of focal length f, depth component t_3 of translation vector, and radial lens distortion parameters $\mathbf{d}$ from the following nonlinear equations by using a standard nonlinear minimization scheme such as the steepest descent [21]:

$$f(r_{11}x + r_{12}y + t_1) - t_3 i_d - \mathbf{d}^T \mathbf{v}_d i_d (r_{31}x + r_{32}y) - t_3 \mathbf{d}^T \mathbf{v}_d i_d - i_d(r_{31}x + r_{32}y) = 0, \tag{23}$$

$$f(r_{21}x + r_{22}y + t_2) - t_3 j_d - \mathbf{d}^T \mathbf{v}_d j_d (r_{31}x + r_{32}y) - t_3 \mathbf{d}^T \mathbf{v}_d j_d - j_d(r_{31}x + r_{32}y) = 0,$$

where $\mathbf{v}_d$ is defined in Eqs. 15 and 16, and r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2 are obtained from Step 4.

The method is computationally efficient and also includes the radial lens distortion parameters. This makes the method widely applicable to a variety of applications. However, the solution is sub-optimal and does not impose the orthonormality constraints Eq. 13. Furthermore, the image center and scale factor parameters are not included in the solution. Lenz and Tsai [20] proposes a nonlinear method for solving the image center parameters by minimizing the RAC residual error. Moreover, by taking the ratio of the collinearity condition equations (see Eq. 17), the method only considers the tangential component of the collinearity equations and ignores the radial component. Furthermore, the tangential lens distortion parameters are ignored. Besides, our experiments reveal poorer estimates of parameters with this method.

3.2 Ganapathy's method

We extend the noncoplanar solution of Ganapathy [11] to the coplanar case. Here, we assume that the image center (i_0, j_0) and scale factor s are either ignored or known in advance by any one of the offline methods described in Sect. 1. We further assume that lens distortion is ignored or corrected in advance. We use the following unconstrained objective function derived from Eqs. 1 and 2:

$$J(\mathbf{b}) = \sum_{all\ calib.points} (fr_{11}x + fr_{12}y + ft_1 - r_{31}ix - r_{32}iy - it_3)^2 + \sum_{all\ calib.points} (fr_{21}x + fr_{22}y + ft_2 - r_{31}jx - r_{32}jy - jt_3)^2, \tag{24}$$

where $\mathbf{b} = \{r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2, t_3, f\}$ are the extrinsic parameters and focal length. Here (i, j) are ideal undistorted image coordinates corrected for lens distortion, image center, and scale factor by using Eqs. 8,9 and 10. The method consists of the following steps.

1. Minimizing J in Eq. 24, we obtain the following linear equation in unknown parameter vector $\mathbf{b}$:

$$\begin{bmatrix} x & y & 0 & 0 & -ix & -iy & 1 & 0 \\ 0 & 0 & x & y & -jx & -jy & 0 & 1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} i \\ j \end{bmatrix}, \tag{25}$$

where

$$\mathbf{b}^T = \left[\frac{fr_{11}}{t_3} \;\; \frac{fr_{12}}{t_3} \;\; \frac{fr_{21}}{t_3} \;\; \frac{fr_{22}}{t_3} \;\; \frac{r_{31}}{t_3} \;\; \frac{r_{32}}{t_3} \;\; \frac{ft_1}{t_3} \;\; \frac{ft_2}{t_3} \right].$$

We use the linear least squares method to solve for $\mathbf{b}$.
2. Impose the orthonormality constraints Eq. 13, and obtain the parameters below:

$$f^2 = -\frac{b_1 b_2 + b_3 b_4}{b_5 b_6}, \tag{26}$$

$$|t_3| = \sqrt{\frac{f^2}{b_1^2 + b_3^2 + f^2 b_5^2}} \text{ or } \sqrt{\frac{f^2}{b_2^2 + b_4^2 + f^2 b_6^2}},$$

where $\mathbf{b}^T = \left[b_1 \; b_2 \; b_3 \; b_4 \; b_5 \; b_6 \; b_7 \; b_8 \right]$ is obtained in Step 1.
3. The sign of t_3 can be determined from the camera position with respect to the world coordinate system. As drawn in Fig. 1, the sign of t_3 is negative (positive) if the origin of the world coordinate system O_w is in front (behind) the camera. An algorithmic method to determine the sign of t_3 is given in [27].
4. Obtain the remaining parameters from the following equations:

$$r_{11} = \frac{t_3 b_1}{f},\ r_{12} = \frac{t_3 b_2}{f},\ r_{21} = \frac{t_3 b_3}{f},\ r_{22} = \frac{t_3 b_4}{f},$$
$$t_1 = \frac{t_3 b_7}{f},\ t_2 = \frac{t_3 b_8}{f} \tag{27}$$

$$r_{31} = t_3 b_5,\ r_{32} = t_3 b_6,\ r_{13} = r_{21} r_{32} - r_{31} r_{22},$$
$$r_{23} = r_{31} r_{12} - r_{11} r_{32},\ r_{33} = r_{11} r_{22} - r_{21} r_{12}.$$

The method is very efficient, but it produces a suboptimal solution including an ambiguous solution for t_3. Furthermore, the image center, scale factor, and lens distortion parameters are not considered.

3.3 Grosky and Tamburino's method

We extend the noncoplanar solution of Grosky and Tamburino [15] to the coplanar case. In this algorithm, we assume that lens distortion can be ignored or corrected in advance, and the image center parameters (i_0, j_0) are ignored or computed in advance by applying an offline method in Sect. 1.1. Ignoring image center and lens distortion, from Eq. 10, we get $i = s^{-1} i_f$ and $j = j_f$. From the collinearity conditions (Eqs. 1 and 2), we obtain

$$f\left(\frac{r_{11}x + r_{12}y + t_1}{r_{31}x + r_{32}y + t_3}\right) = s^{-1} i_f,$$

and

$$f\left(\frac{r_{21}x + r_{22}y + t_2}{r_{31}x + r_{32}y + t_3}\right) = j_f. \tag{28}$$

From Eq. 28, we obtain the following objective function:

$$J(\mathbf{b}, s) = \sum_{all\ calib.points} \left(sfr_{11}x + sfr_{12}y + sft_1 - r_{31} i_f x - r_{32} i_f y - i_f t_3\right)^2 + \sum_{all\ calib.points} \left(fr_{21}x + fr_{22}y + ft_2 - r_{31} j_f x - r_{32} j_f y - j_f t_3\right)^2, \tag{29}$$

where $\mathbf{b} = \{r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2, t_3, f\}$ are the extrinsic parameters, focal length, and scale factor. The method consists of the following steps.

1. Minimizing J in Eq. 29, we obtain the following linear equation in unknown parameter vector $\mathbf{b}$:

$$\begin{bmatrix} x & y & 0 & 0 & -i_f x & -i_f y & 1 & 0 \\ 0 & 0 & x & y & -j_f x & -j_f y & 0 & 1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} i_f \\ j_f \end{bmatrix}, \tag{30}$$

where

$$\mathbf{b}^T = \left[\frac{sfr_{11}}{t_3}\ \frac{sfr_{12}}{t_3}\ \frac{fr_{21}}{t_3}\ \frac{fr_{22}}{t_3}\ \frac{r_{31}}{t_3}\ \frac{r_{32}}{t_3}\ \frac{sft_1}{t_3}\ \frac{ft_2}{t_3}\right].$$

We use the linear least squares method to solve for $\mathbf{b}$.

2. Impose the orthonormality constraints Eq. 13, and obtain the parameters below:

$$s = \sqrt{\frac{b_5 b_6 \left(b_2^2 - b_1^2\right) + b_1 b_2 \left(b_5^2 - b_6^2\right)}{b_5 b_6 \left(b_3^2 - b_4^2\right) - b_3 b_4 \left(b_5^2 - b_6^2\right)}}, \tag{31}$$

$$f = \sqrt{\frac{b_3 b_4 \left(b_1^2 - b_2^2\right) - b_1 b_2 \left(b_3^2 - b_4^2\right)}{b_5 b_6 \left(b_2^2 - b_1^2\right) + b_1 b_2 \left(b_5^2 - b_6^2\right)}}, \tag{32}$$

$$|t_3| = \frac{sf}{\sqrt{b_1^2 + b_3^2 s^2 + b_7^2 s^2 f^2}}, \tag{33}$$

where $\mathbf{b}^T = \left[b_1\ b_2\ b_3\ b_4\ b_5\ b_6\ b_7\ b_8\right]$ is obtained from Step 1.

3. The sign of t_3 can be determined from the camera position with respect to the world coordinate system. As drawn in Fig. 1, the sign of t_3 is negative (positive) if the origin of the world coordinate system O_w is in front (behind) the camera. An algorithmic method to determine the sign of t_3 is given in [27].
4. Obtain the remaining parameters from the following equations:

$$r_{11} = \frac{t_3 b_1}{sf},\ r_{12} = \frac{t_3 b_2}{sf},\ r_{21} = \frac{t_3 b_3}{f},\ r_{22} = \frac{t_3 b_4}{f},$$
$$t_1 = \frac{t_3 b_7}{sf},\ t_2 = \frac{t_3 b_8}{f}, \tag{34}$$

$$r_{31} = t_3 b_5,\ r_{32} = t_3 b_6,\ r_{13} = r_{21} r_{32} - r_{31} r_{22},$$
$$r_{23} = r_{31} r_{12} - r_{11} r_{32},\ r_{33} = r_{11} r_{22} - r_{21} r_{12}.$$

The method is very efficient, but it produces a suboptimal solution. It also computes the scale factor parameter s. However, the image center, and lens distortion parameters are not computed.

3.4 Chatterjee et al. method

We extend Chatterjee et al.'s noncoplanar algorithm [8] to the coplanar case. The method also includes the lens distortion parameters in the estimation algorithm. However, it assumes that the image center parameters (i_0, j_0) are ignored or computed in advance by an offline method in Sect. 1.1. A starting estimate of scale factor parameter s is required. After ignoring image center and scale factor, we correct for lens distortion by using Eqs. 8 and 9 to obtain ideal undistorted image coordinates which are functions of $\mathbf{d}$ (see Eq. 12) as $i = i(\mathbf{d})$ and $j = j(\mathbf{d})$, where $\mathbf{d}^T = [k_1 k_2 \ldots k_{r_0} p_1 p_2 s_1 s_2]$. Both radial and tangential lens distortions are considered. The algorithm is based on the following unconstrained objective function:

$$J(\mathbf{b}, d) = \sum_{all\ calib.points} \left(fr_{11}x + fr_{12}y + ft_1 - r_{31} ix - r_{32} iy - it_3\right)^2 + \sum_{all\ calib.points} \left(fr_{21}x + fr_{22}y + ft_2 - r_{31} jx - r_{32} jy - jt_3\right)^2, \tag{35}$$

where $\mathbf{b} = \{r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2, t_3, f\}$. The method consists of the following iterative algorithm consisting of two linear least squares steps. We start the algorithm with $\mathbf{d} = 0$. We iterate between Steps 1 and 2.

1. Compute $\mathbf{b}$ with $\mathbf{d}$ held constant. Correct the image coordinates for lens distortion $\mathbf{d}$ by using Eqs. 8 and 9.

From Eq. 35, we obtain the following linear equation in unknown parameter vector **b**:

$$\begin{bmatrix} x\ y\ 0\ 0\ -ix\ -iy\ 1\ 0 \\ 0\ 0\ x\ y\ -jx\ -jy\ 0\ 1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} i \\ j \end{bmatrix}, \tag{36}$$

where

$$\mathbf{b}^T = \left[\frac{fr_{11}}{t_3}\ \frac{fr_{12}}{t_3}\ \frac{fr_{21}}{t_3}\ \frac{fr_{22}}{t_3}\ \frac{r_{31}}{t_3}\ \frac{r_{32}}{t_3}\ \frac{ft_1}{t_3}\ \frac{ft_2}{t_3} \right].$$

We use the linear least squares method to solve for **b**.

2. Compute **d** with **b** held constant. From Eqs. 8 and 9 and Eq. 35, we obtain the following linear equation in unknown parameter vector **d**:

$$(b_5x + b_6y + 1) \begin{bmatrix} \mathbf{w}^T \\ \tilde{\mathbf{w}}^T \end{bmatrix} \mathbf{d} = \begin{bmatrix} (b_1x + b_2y + b_7) - i_d(b_5x + b_6y + 1) \\ (b_3x + b_4y + b_8) - j_d(b_5x + b_6y + 1) \end{bmatrix}, \tag{37}$$

where **w** and $\tilde{\mathbf{w}}$ are defined in Eqs. 8 and 9 respectively, and **b** is obtained from Step 1.

At the end of the iterations, all extrinsic parameters and focal length are obtained from Steps 2–4 in Sect. 3.2. Although the method is iterative, it is efficient because each step involves linear equations only. However, the method gives us a sub-optimal solution, and does not compute the image center parameters. Our experiments suggest, that the method performs quite well when the scale factor is unknown. At the end of iterative Steps 1 and 2, we use Steps 2–4 of Sect. 3.3 to obtain the scale factor as well as all extrinsic parameters and focal length.

4 Nonlinear algorithms for coplanar calibration

In this section, we describe two algorithms to estimate all calibration parameters such that the orthonormality conditions Eq. 13 are fully satisfied. Both methods are based on nonlinear optimization, and yield optimal solutions. The first method is a coplanar extension of the noncoplanar methods in photogrammetry [13,18,23,29] in which an unconstrained objective function, with transcendental terms, is used. The second method is a novel technique that uses an objective function that is constrained by the orthonormality conditions Eq. 13. We refer to this method as the *Constrained Optimization* algorithm. For both methods, we partition the parameter space into blocks of parameters, and use the Gauss-Seidel [2,14] technique of nonlinear minimization. In the Gauss-Seidel method, the objective function is iteratively minimized for each parameter block with the remaining held constant. This procedure is particularly attractive because (1) of its easy implementation, (2) it reduces the nonlinear minimization to a smaller parameter space, and (3) it lends itself to a comprehensive theoretical convergence analysis.

From the perspective of optimization theory, a nonlinear procedure should satisfy the following desirable features: (1) use a good starting estimate obtained by an alternative procedure, (2) an efficient optimization scheme that satisfies all constraints, and (3) a proof of convergence at least to a local minimum solution. Here, we address all these issues in our nonlinear algorithms. Any one of the linear methods discussed in Sect. 3 can be used as an initialization step to obtain starting estimates of parameters. We also provide an analytical proof of convergence for the Constrained Optimization algorithm in Sect. 4.3.

4.1 Photogrammetric method

This method is a coplanar extension of the noncoplanar algorithms in analytical photogrammetry [13,16,18,23,29]. Here, we represent the rotation matrix R by Euler angles ω, ϕ, and κ [16,23], where ω is the clockwise rotation angle (tilt) around X_w-axis, ϕ (pan) around Y_w-axis, and κ (swing) around Z_w-axis (Eq. 38; see top of next page). The collinearity condition equations (1,2) can then be represented in terms of the Euler angles $\{\omega, \phi, \kappa\}$ to obtain an unconstrained objective function (Eq. 39; see top of next page) where $\mathbf{b} = \{\omega, \phi, \kappa, t_1, t_2, t_3, f\}$, $\mathbf{d}^T = [k_1, \ldots, k_{r_0}, p_1, p_2, s_1, s_2]$, and $\mathbf{m} = \{i_0, j_0, s\}$. The frame buffer image coordinates (i_f, j_f) are corrected for image center and scale factor parameters **m** by using Eq. 10, and for lens distortion parameters **d** by using Eqs. 8 and 9 to obtain ideal undistorted image coordinates (i, j). Thus, the ideal undistorted image coordinates (i, j) are functions of both **m** and **d** as shown in Eq. 12. We start the algorithm with an initial estimate of **m**, and with $\mathbf{d} = \mathbf{0}$. The algorithm iterates between the following three steps.

1. Compute **b** with **d** and **m** held constant. Correct the frame buffer image coordinates (i_f, j_f) by image center and scale factor parameters **m** by using Eq. 10, and then by lens distortion parameters **d** by using Eqs. 8 and 9 to obtain ideal undistorted image coordinates (i, j). Then, minimize J in Eq. 39 with respect to **b**. This is an unconstrained nonlinear minimization problem, whereby the objective function is minimized by a nonlinear procedure such as steepest descent or quasi-Newton methods [2,21].
2. Compute **d** with **b** and **m** held constant. Correct the frame buffer image coordinates (i_f, j_f) by image center and scale factor parameters **m** by using Eq. 10 to obtain distorted image coordinates (i_d, j_d). Then, from Eqs. 39 and 8 and 9, we obtain the following unconstrained linear equation in unknown parameter vector **d**:

$$\begin{bmatrix} \mathbf{w}^T \\ \tilde{\mathbf{w}}^T \end{bmatrix} \mathbf{d} = \begin{bmatrix} c_1 - i_d \\ c_2 - j_d \end{bmatrix}, \tag{40}$$

where **w** and $\tilde{\mathbf{w}}$ are defined in Eqs. 8 and 9, respectively, and **b** is obtained from Step 1,

$$c_1 = f\left(\frac{(\cos\phi\cos\kappa)x + (\sin\omega\sin\phi\cos\kappa + \cos\omega\sin\kappa)y + t_1}{(\sin\phi)x + (-\sin\omega\cos\phi)y + t_3} \right),$$

and

$$c_2 = f\left(\frac{(-\cos\phi\sin\kappa)x + (-\sin\omega\sin\phi\sin\kappa + \cos\omega\cos\kappa)y + t_2}{(\sin\phi)x + (-\sin\omega\cos\phi)y + t_3} \right).$$

3. Compute **m** with **b** and **d** held constant. This is an unconstrained nonlinear minimization problem whereby the objective function J in Eq. 39 is minimized with respect to $\mathbf{m} = \{i_0, j_0, s\}$. The frame buffer image coordinates (i_f, j_f) are corrected for lens distortion by Eqs. 8 and 9 as

$$R = \begin{bmatrix} \cos\phi\cos\kappa & \sin\omega\sin\phi\cos\kappa + \cos\omega\sin\kappa & -\cos\omega\sin\phi\cos\kappa + \sin\omega\sin\kappa \\ -\cos\phi\sin\kappa & -\sin\omega\sin\phi\sin\kappa + \cos\omega\cos\kappa & \cos\omega\sin\phi\sin\kappa + \sin\omega\cos\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \tag{38}$$

$$J(\mathbf{b},\mathbf{d},\mathbf{m}) = \sum_{all\ \ calib.points} \left(i(\mathbf{d},\mathbf{m}) - f\left(\frac{(\cos\phi\cos\kappa)x + (\sin\omega\sin\phi\cos\kappa + \cos\omega\sin\kappa)y + t_1}{(\sin\phi)x + (-\sin\omega\cos\phi)y + t_3} \right) \right)^2 + \sum_{all\ \ calib.points} \left(j(\mathbf{d},\mathbf{m}) - f\left(\frac{(-\cos\phi\sin\kappa)x + (-\sin\omega\sin\phi\sin\kappa + \cos\omega\cos\kappa)y + t_2}{(\sin\phi)x + (-\sin\omega\cos\phi)y + t_3} \right) \right)^2, \tag{39}$$

$i(\mathbf{m}) = i_d(\mathbf{m}) + \mathbf{w}(\mathbf{m})^T\mathbf{d}$ and $j(\mathbf{m}) = j_d(\mathbf{m}) + \tilde{\mathbf{w}}(\mathbf{m})^T\mathbf{d}$,

where $\mathbf{w}$ and $\tilde{\mathbf{w}}$ defined in Eqs. 8 and 9 respectively are functions of $\mathbf{m}$. Also, $i_d = s^{-1}(i_f - i_0), j_d = (j_f - j_0)$ and $r_d^2 = s^{-2}(i_f - i_0)^2 + (j_f - j_0)^2$ are all functions of $\mathbf{m} = \{i_0, j_0, s\}$. The steepest descent or quasi-Newton methods [21] may be used to solve for $\mathbf{m}$ in the nonlinear objective function J in Eq. 39.

At the end of the iterations, the extrinsic rotation matrix R is computed from (ω, ϕ, κ) by using Eq. 38. The algorithm yields an optimal solution and solves all calibration parameters. However, it requires precise starting estimates to converge correctly, since the periodic transcendental terms in the objective function may lead to false minima. The method is also computationally complex.

4.2 Constrained-optimization method

This algorithm uses the following constrained objective function obtained from the collinearity condition Eqs. 1 and 2:

$$J(\mathbf{b},\mathbf{d},\mathbf{m}) = \sum_{all\ \ calib.points} (fr_{11}x + fr_{12}y + ft_1 - r_{31}ix - r_{32}iy - it_3)^2 + \sum_{all\ \ calib.points} (fr_{21}x + fr_{22}y + ft_2 - r_{31}jx - r_{32}jy - jt_3)^2.$$

under orthonormality constraints (Eq. 13), (41)

where $\mathbf{b} = \{r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2, t_3, f\}$, $\mathbf{d}^T = [k_1, \ldots, k_{r_0}, p_1, p_2, s_1, s_2]$, $\mathbf{m} = \{i_0, j_0, s\}$. The frame buffer image coordinates (i_f, j_f) are corrected for image center and scale factor parameters $\mathbf{m}$ by using Eq. 10, and for lens distortion parameters $\mathbf{d}$ by using Eqs. 8 and 9 to obtain ideal undistorted image coordinates (i, j). Thus, the ideal undistorted image coordinates (i, j) are functions of both $\mathbf{m}$ and $\mathbf{d}$ as shown in Eq. 12. Furthermore, the orthonormality constraints Eqs. 13 are imposed *during the optimization process*. We define the parameter block $\mathbf{b}$ containing extrinsic parameters and focal length as

$$\mathbf{b}^T = \left[\frac{fr_{11}}{t_3} \ \frac{fr_{12}}{t_3} \ \frac{fr_{21}}{t_3} \ \frac{fr_{22}}{t_3} \ \frac{r_{31}}{t_3} \ \frac{r_{32}}{t_3} \ \frac{ft_1}{t_3} \ \frac{ft_2}{t_3} \right]. \tag{42}$$

We can now write the three constraints in Eq. 13 in terms of the elements of $\mathbf{b}$. Let

$$\mathbf{b}^T = \left[b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6 \ b_7 \ b_8 \right].$$

Then, the constraint for coplanar camera calibration is

$$h(\mathbf{b}) = (b_5b_1 + b_6b_2)(b_6b_1 - b_5b_2) + (b_5b_3 + b_6b_4)(b_6b_3 - b_5b_4) = 0. \tag{43}$$

We start the algorithm with an initial estimate of $\mathbf{m}$ (see Sects. 1.1 and 1.2), and with $\mathbf{d} = 0$. The method consists of iterations between the following three steps.

1. Compute $\mathbf{b}$ with $\mathbf{d}$ and $\mathbf{m}$ held constant. Correct the frame buffer image coordinates (i_f, j_f) by image center and scale factor parameters $\mathbf{m}$ by using Eq. 10, and then by lens distortion parameters $\mathbf{d}$ by using Eqs. 8 and 9 to obtain ideal undistorted image coordinates (i, j). Then, minimize J in Eq. 41 with respect to $\mathbf{b}$ under constraint 43. We solve for $\mathbf{b}$ from the following constrained linear equation:

$$\begin{bmatrix} x & y & 0 & 0 & -ix & -iy & 1 & 0 \\ 0 & 0 & x & y & -jx & -jy & 0 & 1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} i \\ j \end{bmatrix}, \tag{44}$$

under constraint

$$h(\mathbf{b}) = (b_5b_1 + b_6b_2)(b_6b_1 - b_5b_2) + (b_5b_3 + b_6b_4) \times (b_6b_3 - b_5b_4) = 0.$$

Any constrained minimization method such as quasi-Newton with penalty function or augmented Lagrangian [21] can be used.

2. Compute $\mathbf{d}$ with $\mathbf{b}$ and $\mathbf{m}$ held constant. Correct the frame buffer image coordinates (i_f, j_f) by image center and scale factor parameters $\mathbf{m}$ by using Eq. 10 to obtain distorted image coordinates (i_d, j_d). Then from Eqs. 41 and 8 and 9, we obtain the following unconstrained linear equation in unknown parameter vector $\mathbf{d}$:

$$(b_5x + b_6y + 1) \begin{bmatrix} \mathbf{w}^T \\ \tilde{\mathbf{w}}^T \end{bmatrix} \mathbf{d} = \begin{bmatrix} (b_1x + b_2y + b_7) - i_d(b_5x + b_6y + 1) \\ (b_3x + b_4y + b_8) - j_d(b_5x + b_6y + 1) \end{bmatrix}, \tag{45}$$

where $\mathbf{w}$ and $\tilde{\mathbf{w}}$ are defined in Eqs. 8 and 9 respectively, and $\mathbf{b}$ is obtained from Step 1.

3. Compute $\mathbf{m}$ with $\mathbf{b}$ and $\mathbf{d}$ held constant. This is an unconstrained nonlinear minimization problem where the objective function J in Eq. 41 is minimized with respect to $\mathbf{m} = \{i_0, j_0, s\}$. The frame buffer image coordinates

(i_f, j_f) are corrected for lens distortion by Eqs. 8 and 9 as:

$$i(\mathbf{m}) = i_d(\mathbf{m}) + \mathbf{w}(\mathbf{m})^T\mathbf{d} \text{ and } j(\mathbf{m}) = j_d(\mathbf{m}) + \tilde{\mathbf{w}}(\mathbf{m})^T\mathbf{d},$$

where $\mathbf{w}$ and $\tilde{\mathbf{w}}$ defined in Eqs. 8 and 9, respectively are functions of $\mathbf{m}$. Also, $i_d = s^{-1}(i_f - i_0), j_d = (j_f - j_0)$ and $r_d^2 = s^{-2}(i_f - i_0)^2 + (j_f - j_0)^2$ are all functions of $\mathbf{m} = \{i_0, j_0, s\}$. The steepest descent or quasi-Newton methods [21] may be used to solve for $\mathbf{m} = \{i_0, j_0, s\}$ in the following nonlinear equations:

$$\begin{aligned}\left(i_d(\mathbf{m}) + \mathbf{w}(\mathbf{m})^T\mathbf{d}\right)(b_5x + b_6y + 1) + (b_1x + b_2y + b_7) &= 0,\\ \left(j_d(\mathbf{m}) + \tilde{\mathbf{w}}(\mathbf{m})^T\mathbf{d}\right)(b_5x + b_6y + 1) + (b_3x + b_4y + b_8) &= 0,\end{aligned} \tag{46}$$

where $\mathbf{b}$ is obtained from Step 1, and $\mathbf{d}$ from Step 2.

At the end of the iterations, all extrinsic parameters and focal length are obtained from Steps 2–4 in Sect. 3.2. The method is computationally complex, but it produces an optimal solution and computes all calibration parameters.

4.3 Convergence analysis for the constrained optimization algorithm

A convergence analysis of the Constrained Optimization algorithm is necessary because the constraint $h(\mathbf{b}) = 0$ in Eq. 43 is nonconvex. The lemmas and theorem below discuss: (1) conditions for the objective function J in Eq. 41 to (strictly) decrease in successive iterations, (2) a closed-form expression for the amount of decrease in each iteration, and (3) convergence of the nonlinear algorithm to a local minimum solution. Proofs for all lemmas and theorem 1 are given in the Appendix.

Lemma 1. *The Constrained Optimization algorithm reduces the objective function J in Step 1.*

Lemma 2. *The Constrained Optimization method strictly reduces the objective function J in Step2.*

Lemma 3. *The Constrained Optimization algorithm reduces the objective function J in Step 3.*

Theorem 1. *Suppose that J has a strict local minimum point $\mathbf{x}^* = (\mathbf{b}^*, \mathbf{d}^*, \mathbf{m}^*)$. Let $\{\mathbf{x}_k = (\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k)\}$ be the sequence generated by the Constrained Optimization algorithm. If $\mathbf{x}_0$ is sufficiently close to $\mathbf{x}^*$, then $\{\mathbf{x}_k\}$ converges to $\mathbf{x}^*$.*

5 Experimental results

We test the algorithms on two sets of data: (1) synthetic data corrupted with known noise, and (2) real data obtained from a calibration setup. We evaluate the accuracy of each algorithm according to the square root of the mean square error in both image components:

$$\text{image error (IE)} = \sqrt{\frac{1}{N}\left(\sum_{n=1}^{N} s^{-2}\,(i_n - i_n(\mathbf{b},\mathbf{d},\mathbf{m}))^2 + \sum_{n=1}^{N}(j_n - j_n(\mathbf{b},\mathbf{d},\mathbf{m}))^2\right)}, \tag{47}$$

where (i_n, j_n) are points measured from the image and $(i_n(\mathbf{b},\mathbf{d},\mathbf{m}), j_n(\mathbf{b},\mathbf{d},\mathbf{m}))$ are computed from calibration parameters $(\mathbf{b},\mathbf{d},\mathbf{m})$.

5.1 Experiments with synthetic data

We generated synthetic data with a known set of extrinsic and intrinsic camera parameters. We first produced a 10×10 (i.e., $N = 100$) grid of points to simulate the image points. The image is 512 columns and 480 rows. We placed the image points in the 10×10 grid at equal distance from each other and from the edges of the image in both column and row dimensions. We generated the rotation matrix R from Euler angles $\omega = \phi = \kappa = 15^\circ$ according to Eq. 38. We placed the optical center $\mathbf{c} = O_c$ at (5,3,15). The translation vector $\mathbf{t}$ is obtained as $\mathbf{t} = -R\mathbf{c} = (-2.99, -6.22, -14.54)$, and focal length $f = 300$. World points are obtained from the collinearity conditions (Eqs. 1 and 2). Given image points (i, j), the world points (x, y) is obtained from (Eqs. 1 and 2) as

$$x = \frac{(p_8v - p_5)(p_3 - u) - (p_8u - p_2)(p_6 - v)}{(p_8v - p_5)(p_7u - p_1) - (p_8u - p_2)(p_7v - p_4)}, \tag{48}$$

$$y = \frac{(p_7u - p_1)(p_6 - v) - (p_7v - p_4)(p_3 - u)}{(p_8v - p_5)(p_7u - p_1) - (p_8u - p_2)(p_7v - p_4)},$$

where

$$p_1 = \frac{fr_{11}}{t_3}, p_2 = \frac{fr_{12}}{t_3},\ p_3 = \frac{ft_1}{t_3},\ p_4 = \frac{fr_{21}}{t_3} p_5 = \frac{fr_{22}}{t_3},$$
$$p_6 = \frac{ft_2}{t_3},\ p_7 = \frac{r_{31}}{t_3},\ \text{and}\ p_8 = \frac{r_{32}}{t_3}.$$

Lens distortion is then added to this image points according to first- and second-order radial distortion models with coefficients $k_1 = 10^{-7}$ and $k_2 = 10^{-14}$. We next add image center and scale factor parameters to the image points with $i_0 = 5, j_0 = 4$ and $s = 1$. Finally, we add independent quantization noise from a uniform distribution on the interval $(-0.5, 0.5)$ pixel to the image points. With the calibration parameters described before, we generated 100 test data sets, where each data set contains $N = 100$ (10×10 grid) image and world points and different quantization noise.

We use all four linear methods and two nonlinear methods to estimate the calibration parameters and image error (see Eq. 47). For each algorithm, from the 100 data sets, we obtained 100 estimates of each parameter and 100 image errors. We next compute their mean and standard deviation. Finally, we compute the relative error of each parameter as follows:

$$\text{relative error of parameter}\quad \mathbf{a} = \frac{\|\mathbf{a} - \hat{\mathbf{a}}\|}{\|\mathbf{a}\|}, \tag{49}$$

where $\mathbf{a}$ is the true parameter value (described above), and $\hat{\mathbf{a}}$ is the estimated parameter value which is the mean of the 100 parameter estimates. Image error (IE) is also computed as the mean of the image errors from 100 data sets.

Table 3. Relative errors Eq. 49 of parameters and standard deviations of parameter estimates for synthetic data

	Tsai	Ganapathy	Grosky	Chatterjee	Photogrammetry	Constrained
r_{11}	0.0045 (0.0002)	0.0045 (0.0002)	0.0024 (0.0001)	0.0024 (0.0001)	0.0004 (0.0004)	0.0005 (0.0020)
r_{12}	0.0137 (0.0002)	0.0135 (0.0002)	0.0156 (0.0002)	0.0153 (0.0002)	0.0025 (0.0002)	0.0014 (0.0020)
r_{13}	0.0880 (0.0011)	0.0104 (0.0004)	0.0176 (0.0004)	0.0195 (0.0004)	0.0026 (0.0005)	0.0029 (0.0035)
r_{21}	0.0142 (0.0003)	0.0137 (0.0002)	0.0065 (0.0002)	0.0068 (0.0002)	0.0009 (0.0003)	0.0006 (0.0031)
r_{22}	0.0039 (0.0003)	0.0035 (0.0002)	0.0037 (0.0002)	0.0034 (0.0002)	0.0007 (0.0003)	0.0005 (0.0030)
r_{23}	0.0410 (0.0010)	0.0093 (0.0006)	0.0278 (0.0008)	0.0252 (0.0008)	0.0051 (0.0009)	0.0059 (0.0116)
r_{31}	0.0463 (0.0011)	0.0100 (0.0005)	0.0251 (0.0006)	0.0253 (0.0006)	0.0041 (0.0008)	0.0048 (0.0063)
r_{32}	0.0713 (0.0010)	0.0093 (0.0005)	0.0258 (0.0007)	0.0223 (0.0007)	0.0050 (0.0007)	0.0072 (0.0092)
r_{33}	0.0018 (0.0004)	0.0014 (0.0003)	0.0037 (0.0003)	0.0035 (0.0003)	0.0007 (0.0004)	0.0014 (0.0015)
t_1	0.0812 (0.0028)	0.0827 (0.0022)	0.0805 (0.0021)	0.0790 (0.0021)	0.0124 (0.0022)	0.0055 (0.0072)
t_2	0.0313 (0.0026)	0.0307 (0.0022)	0.0382 (0.0025)	0.0386 (0.0025)	0.0065 (0.0021)	0.0092 (0.0686)
t_3	0.0075 (0.0534)	0.0054 (0.0317)	0.0295 (0.0373)	0.0241 (0.0378)	0.0043 (0.0492)	0.0057 (0.4486)
f	0.0054 (1.0867)	0.0011 (0.6747)	0.0427 (0.8259)	0.0311 (0.8579)	0.0065 (0.8863)	0.0066 (0.9297)
s			0.0094 (0.0003)	0.0093 (0.0003)	0.0035 (0.0042)	0.0051 (0.2675)
i_0					0.1630 (0.0023)	0.1994 (0.0048)
j_0					0.0018 (0.0003)	0.0004 (0.0032)
k_1	0.8473 (0.0000)			0.0934 (0.0000)	0.0296 (0.0000)	0.0177 (0.0000)
k_2	280.76 (0.0000)			6.7209 (0.0000)	3.3463 (0.0000)	2.8410 (0.0000)
IE	5.1504 (1.0381)	0.7797 (0.0284)	0.6960 (0.0283)	0.4023 (0.0213)	0.3087 (0.0225)	0.3230 (0.0237)

Table 4. Parameter estimates and their standard deviations for real data

	Tsai	Ganapathy	Grosky	Chatterjee	Photogrammetry	Constrained
r_{11}	−0.1533(0.0003)	−0.1536(0.0002)	−0.1533(0.0003)	−0.1535(0.0003)	−0.1540(0.0001)	−0.1534(0.0002)
r_{12}	0.9748 (0.0006)	0.9758 (0.0006)	0.9740 (0.0003)	0.9754 (0.0004)	0.9756 (0.0001)	0.9746 (0.0006)
r_{13}	0.1618 (0.0037)	0.1637 (0.0014)	0.1668 (0.0019)	0.1581 (0.0024)	0.1564 (0.0005)	0.1600 (0.0014)
r_{21}	−0.9721(0.0005)	−0.9723(0.0002)	−0.9732(0.0003)	−0.9719(0.0003)	−0.9726(0.0001)	−0.9724(0.0002)
r_{22}	−0.1782(0.0006)	−0.1779(0.0003)	−0.1781(0.0002)	−0.1780(0.0003)	−0.1776(0.0001)	−0.1779(0.0001)
r_{23}	0.1525 (0.0031)	0.1432 (0.0026)	0.1456 (0.0017)	0.1541 (0.0019)	0.1502 (0.0006)	0.1512 (0.0010)
r_{31}	0.1775 (0.0028)	0.1684 (0.0027)	0.1715 (0.0016)	0.1784 (0.0017)	0.1743 (0.0006)	0.1760 (0.0010)
r_{32}	−0.1339(0.0039)	−0.1375(0.0013)	−0.1400(0.0020)	−0.1300(0.0026)	−0.1290(0.0005)	−0.1324(0.0014)
r_{33}	0.9750 (0.0006)	0.9761 (0.0005)	0.9752 (0.0003)	0.9753 (0.0003)	0.9762 (0.0001)	0.9749 (0.0006)
t_1	15.025(0.0107)	15.039 (0.0075)	15.012 (0.0165)	15.036 (0.0178)	15.514 (0.0028)	15.096 (0.0532)
t_2	14.795(0.0079)	14.796 (0.0050)	14.809 (0.0028)	14.790 (0.0035)	15.159 (0.0018)	14.861 (0.0023)
t_3	−142.37(2.4527)	−134.26(2.2340)	−136.74(1.5386)	−144.28(1.5070)	−141.38(0.3628)	−143.15(0.2941)
f	1158.71(20.08)	1093.00 (18.06)	1112.13 (12.66)	1177.30 (12.50)	1153.57 (3.358)	1162.60 (4.448)
s			1.0027 (0.0011)	0.9991 (0.0013)	1.0000 (0.0003)	1.0006 (0.0005)
i_0					3.9983 (0.0034)	3.6464 (0.0092)
j_0					3.0018 (0.0030)	2.8708 (0.0047)
k_1	$-2.3252\text{X}10^{-9}$			$2.9753\text{X}10^{-8}$	$1.0465\text{X}10^{-8}$	$0.9753\text{X}10^{-8}$
k_2	$6.5583\text{X}10^{-13}$			$3.4495\text{X}10^{-14}$	$1.9450\text{X}10^{-13}$	$1.7532\text{X}10^{-13}$
IE	2.3148 (1.8947)	0.9164 (0.2113)	0.9146 (0.2101)	0.3690 (0.1147)	0.2034 (0.0093)	0.2355 (0.0084)

The results of this experiment are given in Table 3. For the Tsai, Chatterjee and nonlinear methods, we fit a second-order radial lens distortion model to the calibration data. We start all algorithms with image center $i_0 = j_0 = 0$, scale factor $s = 1$, and lens distortions $k_1 = k_2 = 0$.

We observe the following from Table 3.

1. The Tsai method is sensitive to the scale factor parameter s, and the image center parameters i_0 and j_0. Although the actual image center is $i_0 = 5$, $j_0 = 4$, we start the algorithms with $i_0 = j_0 = 0$. Consequently, the Tsai method produces a larger image error. By taking the ratio of the collinearity condition equations 1 and 2, Tsai has discarded the radial information in these equations. This is also reflected in the higher image error.
2. The Ganapathy and Grosky methods are very similar in parameter accuracy and in image error. The Grosky method produces the additional scale factor parameter. Both these methods ignore lens distortion and image center parameters.
3. The Chatterjee method also computes the lens distortion parameters, and produces better image error among all linear methods. However, the Chatterjee method is iterative linear, whereas the Ganapathy and Grosky methods are single-step linear methods. The Tsai method uses nonlinear minimization for the lens distortion parameters.
4. Accuracy of parameters is comparable for all four linear methods. All linear methods ignore the image center parameters. Furthermore, all four linear methods are sub-optimal, i.e., they do not satisfy the orthonormality conditions (Eq. 13).
5. The best estimates of all parameters (by an order of magnitude) are obtained from the nonlinear (Photogrammetric and Constrained Optimization) methods. Furthermore, the nonlinear methods also satisfy the orthonormality constraints (Eq. 13), and are, therefore, optimal. Moreover, the nonlinear methods compute all parameters

including the image center parameters. The image errors are also better for the nonlinear algorithms.
6. The variations in the parameter estimates are small (as shown by the standard deviations). This demonstrates that the algorithms are consistent and repeatable in estimating parameters.

5.2 Experiments with real data

Real data is generated from test calibration points created by accurately placing a set of 36 dots in a square grid of 6×6 dots on a flat surface. The center to center distance between the dots is 7.875 mm. The diameter of each dot is 3.875 mm. The calibration pattern is mounted on a custom-made calibration stand. The centroid pixel of each dot is obtained by image processing to subpixel accuracy. Although we observed high lens distortion for wide angle lenses, we used a 35–70 mm zoom lens because of its frequent use in applications, and a depth of field that can focus within a range of 0–60 mm. We used an off-the-shelf camera in a monoview setup. The camera resolution is 512×480 pixels and the digitizer gives digital images with 16 bits/pixel. We acquired 10 images of the calibration points, thereby generating 10 data sets with $N = 36$ image and world points in each data set. For each algorithm, we estimated the parameters and computed image errors from 10 data sets. We next computed their mean and standard deviation.

For the Tsai, Chatterjee and nonlinear methods, we fit a second-order radial lens distortion model to the calibration data. We start all algorithms with image center $i_0 = j_0 = 0$, scale factor $s = 1$, and lens distortions $k_1 = k_2 = 0$. The results are shown in Table 4. As seen with the synthetic data, image error improves due to the nonlinear algorithms as compared to the linear methods. Furthermore, the estimates are consistent and repeatable for all algorithms as shown by their standard deviations.

6 Concluding remarks

In this paper, we discussed six methods of coplanar camera calibration. The methods range in computational complexity and accuracy of estimates. The four linear methods discussed in Sect. 3 are computationally efficient, but they lack the accuracy of the parameter estimates and produce sub-optimal solutions. Besides, these methods cannot compute all calibration parameters. The two nonlinear methods discussed in Sect. 4 are computationally complex, but the parameter estimates are very accurate and the solutions are optimal.

Appendix

Proof of Lemma 1

With $\mathbf{d}$ and $\mathbf{m}$ held constant, the objective function J in Eq. 41 has the form

$$\mathbf{b}^T Q\mathbf{b} - 2\mathbf{b}^T\alpha \text{ under constraint } h(\mathbf{b}) = 0, \tag{50}$$

where

$$Q = \sum_{n=1}^{N} A_n^T A_n, \quad \alpha = \sum_{n=1}^{N} A_n^T \mathbf{u}_n,$$

$$A_n = \begin{bmatrix} x_n & y_n & 0 & 0 & -i_n x_n & -i_n y_n & 1 & 0 \\ 0 & 0 & x_n & y_n & -j_n x_n & -j_n y_n & 0 & 1 \end{bmatrix},$$

and

$$\mathbf{u}_n = \begin{bmatrix} i_n \\ j_n \end{bmatrix}.$$

This is a constrained minimization problem that can be described in terms of Lagrangian L and Lagrange multiplier λ as

$$L(\mathbf{b}, \lambda) = \mathbf{b}^T Qb - 2\mathbf{b}^T\alpha + \lambda h(\mathbf{b}). \tag{51}$$

At $(k+1)^{th}$ iteration after Step 1:

$$\begin{aligned} &J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) - J(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k) \\ &\quad = \mathbf{b}_{k+1}^T Q_k \mathbf{b}_{k+1} - 2\mathbf{b}_{k+1}^T \alpha_k - \mathbf{b}_k^T Q_k \mathbf{b}_k + 2\mathbf{b}_k^T \alpha_k. \end{aligned} \tag{52}$$

Here, we assume that the solution $\mathbf{b}_{k+1}$ obtained in Step 1 satisfies the first-order necessary conditions $h(\mathbf{b}_{k+1}) = 0$ and $\nabla_{\mathbf{b}} L(\mathbf{b}_{k+1}, \lambda_{k+1}) = 0$. Further, note that in every iteration, $\mathbf{b}_k$ satisfies $h(\mathbf{b}_k) = 0$. Thus, from Eq. 52 we obtain

$$\begin{aligned} &J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) - J(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k) \\ &\quad = L(\mathbf{b}_{k+1}, \lambda_{k+1}) - L(\mathbf{b}_k, \lambda_{k+1}). \end{aligned} \tag{53}$$

Here, L is the Lagrangian in Eq. 51. Since $\mathbf{b}_{k+1}$ satisfies the first-order necessary conditions, we have

$$\{\mathbf{b}_{k+1}, \lambda_{k+1}\} = \arg\min_{\{\mathbf{b},\lambda\}} L(\mathbf{b}, \lambda). \tag{54}$$

Let us denote $L(\mathbf{b}_{k+1}, \lambda_{k+1})$ as $L_{\min}(k)$. Then, from Eq. 54 we have, $L(\mathbf{b}_{k+1}, \lambda_{k+1}) - L_{\min}(k) \geq 0$ giving us $J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) \leq J(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k)$. In order to obtain an expression for $L_{\min}(k)$, note that from the first order necessary conditions we obtain

$$\nabla_{\mathbf{b}} L(\mathbf{b}_{k+1}, \lambda_{k+1}) = 2(Q_k \mathbf{b}_{k+1} - \alpha_k + \lambda_{k+1} \nabla_{\mathbf{b}} h(\mathbf{b}_{k+1})) = 0. \tag{55}$$

Note that $\mathbf{b}^T \nabla_b h(\mathbf{b}) = 4h(\mathbf{b}) = 0$. Then, from Eq. 55 we have: $\mathbf{b}_{k+1}^T \alpha_k = \mathbf{b}_{k+1}^T Q_k \mathbf{b}_{k+1}$. Combining these results, we obtain: $L(\mathbf{b}_{k+1}, \lambda_{k+1}) = L_{\min}(k) = -\mathbf{b}_{k+1}^T Q_k \mathbf{b}_{k+1}$. □

Proof of Lemma 2

With $\mathbf{b}$ and $\mathbf{m}$ held constant, the objective function J in Eq. 37 has the form

$$\mathbf{d}^T P\mathbf{d} - 2\mathbf{d}^T\beta, \tag{56}$$

where

$$P = \sum_{n=1}^{N} B_n^T B_n, \quad \beta = \sum_{n=1}^{N} B_n^T \mathbf{v}_n,$$

$$B_n = (b_5 x_n + b_6 y_n + 1) \begin{bmatrix} \mathbf{w}_n^T \\ \tilde{\mathbf{w}}_n^T \end{bmatrix},$$

$$\mathbf{v}_n = \begin{bmatrix} (b_1 x_n + b_2 y_n + b_7) - i_n (b_5 x_n + b_6 y_n + 1) \\ (b_3 x_n + b_4 y_n + b_8) - j_n (b_5 x_n + b_6 y_n + 1) \end{bmatrix}$$

(See Eq. 45).

Here P is positive definite for $(r_0 + 4)$ linearly independent calibration points. At $(k+1)^{th}$ iteration after Step 2:

$$J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) - J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) = \mathbf{d}_{k+1}^T P_{k+1} \mathbf{d}_{k+1} - 2\mathbf{d}_{k+1}^T \beta_{k+1} - \mathbf{d}_k^T P_{k+1} \mathbf{d}_k + 2\mathbf{d}_k^T \beta_{k+1}. \quad (57)$$

Let $P_{k+1} = Z_{k+1}^T Z_{k+1}$ be the Cholesky decomposition of P_{k+1}, and let $\gamma_{k+1} = Z_{k+1}^{-T} \alpha_{k+1}$, where Z_{k+1}^{-T} exists since P_{k+1} is positive definite. Then, from Eq. 57, we have

$$J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) - J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) = ||Z_{k+1}\mathbf{d}_{k+1} - \gamma_{k+1}||^2 - ||Z_{k+1}\mathbf{d}_k - \gamma_{k+1}||^2. \quad (58)$$

Note that $Z_{k+1}\mathbf{d}_{k+1} = \gamma_{k+1}$ is the least square solution of $\mathbf{d}_{k+1}$ in Step 2 making $||Z_{k+1}\mathbf{d}_{k+1} - \gamma_{k+1}|| = 0$. Clearly,

$$J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) - J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) = -||Z_{k+1}\mathbf{d}_k - \gamma_{k+1}||^2, \quad (59)$$

and $J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) < J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k)$ for $\mathbf{d}_{k+1} \neq \mathbf{d}_k$. □

Proof of Lemma 3

Computation of $\mathbf{m}$ with $\mathbf{b}$ and $\mathbf{d}$ held constant is an unconstrained minimization problem. Clearly, $J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_{k+1}) \leq J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k)$, where

$$J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) - J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_{k+1}) = g_k(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) \geq 0. \quad \square$$

Proof of Theorem 1

The constrained minimization of the objective function $J(\mathbf{b}, \mathbf{d}, \mathbf{m})$ in Eq. 41 can be described in terms of Lagrangian L and Lagrange multiplier λ as

$$L(\mathbf{b}, \mathbf{d}, \mathbf{m}, \lambda) = J(\mathbf{b}, \mathbf{d}, \mathbf{m}) + \lambda h(\mathbf{b}). \quad (60)$$

From Lemmas 1–3, we have

$$J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_{k+1}) \leq J(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) < J(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k) \leq J(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k)$$

for all k and $\mathbf{d}_{k+1} \neq \mathbf{d}_k$. Furthermore, $J(\mathbf{x}_{k+1}) = J(\mathbf{x}_k) - \mathbf{z}_k$, where

$$\mathbf{z}_k = (L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1}) - L_{\min}) + \|Z_{k+1}\mathbf{d}_k - \gamma_{k+1}\|^2 + g_k(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k) > 0.$$

Here, $J(\mathbf{x}_{k+1}) < J(\mathbf{x}_k)$ and $J(\mathbf{x}_{k+1}) \geq 0$ implies that there exists $J^* < \infty$, such that $J(\mathbf{x}_k)$ converges to J^*. Furthermore, $J(\mathbf{x}_k) = J(\mathbf{x}_0) - \sum_{i=0}^{k-1} \mathbf{z}_i$. Since $J(\mathbf{x}_k) \to J^*$, $\sum_{i=0}^{\infty} \mathbf{z}_i < \infty$, which implies that $\mathbf{z}_k \to 0$ as $k \to \infty$.

Assume that $\mathbf{x}_0$ is sufficiently close to $\mathbf{x}^*$ such that $S = \{\mathbf{x} = (\mathbf{b}, \mathbf{d}, \mathbf{m}) : J(\mathbf{x}) \leq J(\mathbf{x}_0)$ and $h(\mathbf{b}) = 0\}$ is compact, and J is uniquely minimized on S by $\mathbf{x}^*$. By the above, we have that $\{\mathbf{x}_k\} \subset S$.

We claim that there exists a subsequence $\{\mathbf{x}_{k_n}\}$ of $\{\mathbf{x}_k\}$ such that $\mathbf{x}_{k_n} \to \mathbf{x}^*$. To see this, suppose $\{\mathbf{x}_k\}$ is bounded away from $\mathbf{x}^*$. Note that $\mathbf{x}^*$ is the unique minimum point in S satisfying the first-order necessary conditions; that is, there exists λ^* such that

$$\nabla_\lambda L(\mathbf{b}^*, \mathbf{d}^*, \mathbf{m}^*, \lambda^*) = 0,$$
$$\nabla_\mathbf{b} L(\mathbf{b}^*, \mathbf{d}^*, \mathbf{m}^*, \lambda^*) = 0,$$
$$\nabla_\mathbf{d} L(\mathbf{b}^*, \mathbf{d}^*, \mathbf{m}^*, \lambda^*) = 0,$$
$$\nabla_\mathbf{m} L(\mathbf{b}^*, \mathbf{d}^*, \mathbf{m}^*, \lambda^*) = 0.$$

Note that

$$\nabla_\lambda L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1}) = 0,$$
$$\nabla_\mathbf{b} L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1}) = 2(Q_k \mathbf{b}_k - \alpha_k + \lambda_{k+1} \nabla_\mathbf{b} h(\mathbf{b}_k)),$$
$$\nabla_\mathbf{d} L(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1}) = 2Z_{k+1}^T (Z_{k+1}\mathbf{d}_k - \gamma_{k+1}).$$

Since $\nabla_\lambda L$, $\nabla_\mathbf{b} L$ and $\nabla_\mathbf{d} L$ are continuous, and $(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k) \in S$, we have that $\nabla_\mathbf{b} L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1})$ and $\nabla_\mathbf{d} L(\mathbf{b}_{k+1}, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1})$ are bounded away from $\mathbf{0}$. Therefore, it is clear that $||Z_{k+1}\mathbf{d}_k - \gamma_{k+1}||^2$ is bounded away from 0. Similarly, $L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}_k, \lambda_{k+1}) - L_{\min}(k)$ (see Lemma 1) is bounded away from 0. Hence, $\mathbf{z}_k = ||Z_{k+1}\mathbf{d}_k - \gamma_{k+1}||^2 + (L(\mathbf{b}_k, \mathbf{d}_k, \mathbf{m}k, \lambda_{k+1}) - L_{\min}(k)) + g_k(\mathbf{b}_{k+1}, \mathbf{d}_{k+1}, \mathbf{m}_k)$ is bounded away from 0, which contradicts $\mathbf{z}_k \to 0$, proving our claim.

Since J is continuous, $J(\mathbf{x}_{k_n}) \to J(\mathbf{x}^*)$, which implies $J^* = J(\mathbf{x}^*)$. Hence, $J(\mathbf{x}_k) \to J(\mathbf{x}^*)$, which in turn implies that $\mathbf{x}_k \to \mathbf{x}^*$, since $\mathbf{x}^*$ is the unique local minimum of J on S. □

References

1. Bani-Hasemi A (1991) Finding the Aspect-Ratio of an Imaging System. Proceedings Computer Vision and Pattern Recognition Conference, Maui, Hawaii, June 1991, pp. 122–126
2. Bertsekas DP, Tsitsiklis JN (1989) Parallel and Distributed Computation. Prentice Hall Englewood Cliffs, NJ.
3. Beyer HA (1992) Accurate Calibration of CCD-Cameras. Proc. Computer Vision Pattern Recognition Conf., Champaign, Ill., pp. 96–101
4. Brown DC (1966) Decentering Distortion of Lenses. Photogramm Eng 32: 444–462
5. Brown DC (1971) Close-Range Camera Calibration. Photogramm Eng 37: 855–866
6. Caprile B, Torre V (1990) Using Vanishing Points for Camera Calibration. Int J Comput Vision 4: 127–140
7. Cardillo J, Sid-Ahmed MA (1991) 3D Position Sensing Using a Passive Monocular Vision System. IEEE Trans Pattern Anal Mach Intell 13(8): 809–813
8. Chatterjee C, Roychowdhury VP, Chong EKP (1997) A Nonlinear Gauss-Seidel Algorithm For Noncoplanar And Coplanar Camera Calibration With Convergence Analysis. Comput Vision Image Understanding 67(1): 58–80
9. Faig W (1975) Calibration of Close-Range Photogrammetric Systems: Mathematical Formulation. Photogramm Eng Remote Sensing 41(12): 1479–1486
10. Faugeras OD, Toscani G (1986) The Calibration Problem for Stereo. Proceedings Computer Vision and Pattern Recognition Conference, Miami Beach, Fla., pp 15–20
11. Ganapathy S (1984) Decomposition of Transformation Matrices for Robot Vision. In: Proc. International Conf. on Robotics and Automation, pp 130–139
12. Gennery DB (1979) Stereo-Camera Calibration. Proceedings Image Understanding Workshop, pp. 101–108
13. Ghosh SK (1979) Analytical Photogrammetry, Pergamon Press, New York
14. Golub GH, VanLoan CF (1983) Matrix Computations, Johns Hopkins University Press, Baltimore, MD
15. Grosky WI, Tamburino LA (1990) A Unified Approach to the Linear Camera Calibration Problem. IEEE Trans Pattern Anal Mach Intell 12(7): 663–671

16. Haralick RM, Shapiro LG (1992) Computer and Robot Vision, Vol. 2, Addison-Wesley, Reading, Mass.
17. Ito M, Ishii A (1994) A non-iterative procedure for rapid and precise camera calibration. Pattern Recognition 27(2) 301–310
18. Karara HM Ed. (1989) Non-Topographic Photogrammetry, 2nd ed., American Society for Photogrammetry and Remote Sensing
19. Lai JZC (1993) On the Sensitivity of camera calibration. Image and Vision Computing 11(10): 656–664
20. Lenz RK, Tsai RY (1988) Techniques for Calibrating the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology. IEEE Trans Pattern Anal Mach Intell 10(5): 713–720
21. Luenberger D (1984) Linear and Non-Linear Programming, Addison-Wesley, Mass.
22. Luh JYS, Klaasen JA (1985) A Three-Dimensional Vision by Off-Shelf System with Multi-Cameras. IEEE Trans Pattern Anal Mach Intell 7(1): 35–45
23. Manual of Photogrammetry, 4th ed., American Society of Photogrammetry, 1980
24. Nomura Y et al. (1992) Simple Calibration Algorithm for High-Distortion-Lens Camera. IEEE Trans Pattern Anal Mach Intell 14(11): 1095–1099
25. Penna MA (1991) Camera Calibration: A Quick and Easy Way to Determine the Scale Factor. IEEE Trans Pattern Anal Mach Intell 13(12): 1240–1245
26. Richards JA (1993) Remote Sensing Digital Image Analysis - An Introduction, Springer New York
27. Tsai RY (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 Automation RA-3(4): 323–343
28. Wang LL, Tsai WH (1991) Camera Calibration by Vanishing Lines for 3D Computer Vision. IEEE Trans Pattern Anal Mach Intell 13(4): 370–376
29. Weng J, Cohen P, Herniou M (1992) Camera Calibration with Distortion Models and Accuracy Evaluation. IEEE Trans Pattern Anal Mach Intell 14(10)
30. Willson RG, Shafer SA (1993) What is the Center of the Image? Technical Report CMU-CS-93-122, Carnegie Mellon University, Pittsburgh, Pa
31. Wong KW (1975) Mathematical Formulation and Digital Analysis in Close-Range Photogrammetry. Photogramm Eng Remote Sensing 41(11): 1355–1375
32. Yakimovsky Y, Cunningham R (1978) A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras. Comput Graphics Image Process 7: 195–210
33. Zhang Z (1999) A Flexible New Technique for Camera Calibration. ICCV 1999
34. Strum P, Maybank S (1999) On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications. CVPR 1999

Chanchal Chatterjee received the B. Tech degree in Electrical Engineering from the Indian Institute of Technology, Kanpur, India, and the M.S.E.E. degree from Purdue University, West Lafayette, Indiana, in 1983 and 1984, respectively. Between 1985 and 1995, he worked at Machine Vision International and Medar Inc. both at Detroit, Michigan. In 1996, he received the Ph.D. degree in Electrical and Computer Engineering from Purdue University, West Lafayette, Indiana. He is currently with Tiernan Communications at San Diego, California. Dr. Chatterjee is also affiliated with BAE Systems at San Diego, CA. His areas of interest include digital video, image processing, computer vision, neural networks, and adaptive algorithms and systems for pattern recognition and signal processing.

Vwani P. Roychowdhury received the B. Tech degree from the Indian Institute of Technology, Kanpur, India and the Ph.D. degree from Stanford University, Stanford, Calif, in 1982 and 1989, respectively, all in Electrical Engineering. He is currently a professor in the Department of Electrical Engineering at the University of California, Los Angeles. From August 1991 until June 1996, he was a faculty member at the School of Electrical and Computer Engineering at Purdue University. His research interests include parallel algorithms and architectures, design and analysis of neural networks, application of computational principles to nanoelectronics, special-purpose computing arrays, VLSI design and fault-tolerant computation. He has coauthored several books, including Discrete Neural Computation: A Theoretical Foundation (Prentice Hall, 1995) and Theoretical Advances in Neural Computation and Learning (Kluwer Academic Publishers, 1994).