



**Master of Science**
**<<Artificial Intelligence and Visual Computing>>**

# UNIVERSITY OF WEST ATTICA & UNIVERSITY OF LIMOGES

## FACULTY OF ENGINEERING

### DEPARTMENT OF INFORMATICS AND COMPUTER ENGINEERING

**Master Thesis**

# State-of-the-Art Methods for Camera Calibration in Computer Vision and Photogrammetry

**Student: Christos Chiampopoulos**
**(aivc20014)**

**Supervisor: Elli Petsa**

**Athens, October 2023**

## Members of the examining committee

**This Master Thesis has been approved by the following examining committee**

| | Name | Position | Digital Signature |
|---|---|---|---|
| 1 | Lazaros Grammatikopoulos | Associate Professor | Lazaros Grammatikopoulos Digitally signed by Lazaros Grammatikopoulos Date: 2023.10.09 16:39:06 +03'00' |
| 2 | Vassilios Krassanakis | Assistant Professor | Vasileios Krasanakis Digitally signed by Vasileios Krasanakis Date: 2023.10.09 15:40:51 +03'00' |
| 3 | Elli Petsa | Professor | Eleni Petsa Digitally signed by Eleni Petsa Date: 2023.10.09 14:33:46 +03'00' |

# ABSTRACT

The camera calibration process has been a prominent subject of research in photogrammetry and computer vision for decades. It is an essential part of numerous applications in which it directly affects accuracy and geometric quality of results. However, it does not have a single optimal solution; thus, many methods have been proposed to handle the different applications. With the increasing adoption of digital photogrammetric applications, there comes a need for robust and flexible techniques that operate autonomously. That is what camera autocalibration, or alternatively self-calibration, aims to address. Extended research has already been and is still being done in this field in such a short amount of time so that keeping up with current trends can be difficult. That is why reviews and surveys are important to aggregate all the information. This thesis is a review focused on selected multi-image autocalibration methods. In the first part, the fundamentals of the camera calibration process and the traditional calibration methods are analyzed. A selection of state-of-the-art automated traditional methods are analyzed and compared. Then, the fundamentals of camera self-calibration and its conventional methods are reviewed. Finally, a selection of state-of-the-art self-calibration methods are analyzed and compared.

# Table of Contents

# Table of Figures



# Table of Tables

# Section 1: Introduction

The interest in computer vision and photogrammetry continues to grow as the field consistently advances (Forlani et al., 2015) and becomes more accessible to the broader public, as a result of the decreased cost and increased availability of camera equipment and pertinent algorithms. The drawback of the prominence of low-cost digital cameras is the decrease in image quality as well as in geometric stability, as observed by Läbe & Förstner (2004). Subsequently, a subject of interest for researchers has been the tackling of such problems.

Camera calibration is an essential process for a plethora of photogrammetric applications, like 3D reconstruction (Fathi & Brilakis, 2014; Sturm, 2000b; Wang et al., 2005; Wilczkowiak et al., 2001), robot vision (Davis et al., 2017; Ito, 1990; Kroeger et al., 2019), digital elevation models (Garcia & de Oliveira, 2021; Setiawan et al., 2013), augmented reality (Baratoff et al., 2002; Gibson et al., 2002), autonomous driving (Martins et al., 2020; Song et al., 2016; Xu et al., 2019), and sports (Choi & Seo, 2011; Wu et al., 2020). This process expresses the computation of the camera intrinsic (focal length, principal point, pixel size, lens distortions) and extrinsic (translation, rotation) parameters by observing a single or multiple images. These parameters are then used for estimating the spatial coordinates of the objects projected onto the camera. As a result, the accuracy of the computed camera parameters directly affects the accuracy of the estimated world geometry.

When analyzing the camera calibration process there are multiple different factors that have to be considered depending on the application being examined. There is no single camera calibration method to cover all possible photogrammetric applications, therefore research on the topic has been broad and diverse in terms of approaches. The factor of system cost is important for the calibration method since researchers and commercial users usually have to deal with a limited budget. It is for this reason that the different proposed methods have to be reviewed and critically evaluated, in order to refine knowledge, update on recent developments, and identify shortcomings. Such a review may too be beneficial for new researchers regarding the current state of the subject.

In the literature, camera calibration techniques are typically divided into three main categories: traditional calibration methods, self-calibration methods, and active-vision methods (Long & Dongri, 2019; Song et al., 2013; Wang et al., 2010). The main distinctive

characteristic of traditional calibration methods, also referred to as photogrammetric calibration methods, is the requirement of some type of calibration target whose geometry is known, in contrast to self-calibration methods that do not require calibration targets. Active-vision methods are an intermediate between these two categories, as they do not require a known object but a known camera motion. Recently, a large variety of approaches based on deep learning technology have also been suggested (Liao et al., 2023). Since these approaches do not accurately fit within the aforementioned categories, it is preferred to classify them in their own category which will not be dealt with in the present thesis.

While the traditional target-based methods tend to perform with great precision, they are often not possible or difficult to reproduce in real-world scenarios. On that account, self-calibration methods, also referred to as autocalibration, present a valuable tool for practical applications. Additionally, the ability to automate these methods further expedites their process. A major benefit of automating the calibration process is that photogrammetry becomes more accessible to everyone (Romero & Quintana, 2023). Accordingly, camera autocalibration becomes more desirable than other methods that require proper user input.

The principal proposition of autocalibration is that the camera parameters can be extracted from a set of images that have no predetermined reference points. The advantage of this method is that users do not have to construct and include specific objects while capturing the photographs. In addition, the ability to use photographs that were not initially captured for the purpose of photogrammetric applications opens up new possibilities. However, when considering the automation of the calibration process it is notable that the traditional calibration methods have also been partially automated as far as the calibration object detection goes (Fraser, 2013). A basic overview of these approaches is included as part of this review, with the goal to contrast and compare the advantages and disadvantages of using them over self-calibration methods.

The thesis is structured in the following way. In Section 2, the fundamentals of camera calibration are explained, specifically the camera model. In Section 3, the traditional calibration methods are analyzed and compared. Section 4 contains a review of various state-of-the-art automated approaches based on traditional calibration methods. Each

method is given a brief overview and at the end of the section, they are summarized and compared. Section 5 is the survey of self-calibration methods, where the fundamentals of self-calibration are explained and the conventional methods are analyzed and compared. In Section 6, various state-of-the-art self-calibration methods are reviewed in the same way as in Section 4. Section 7 contains the conclusions, where a brief summary of the observations of this review is given, and future considerations are made.

## Section 2: Camera Calibration

The objective of the camera calibration process is to calculate the internal and external parameters of a camera by examining an associated image. These parameters define the projection of the 3D world geometry to the 2D image plane. The transformation of 3D world coordinates to 2D image coordinates has been described by Tsai (1987) as a 4-step transformation, as illustrated in Figure 1. The first step is a rigid body transformation from real-world coordinates (X,Y,Z) to camera coordinates ($X_c$,$Y_c$,$Z_c$). The second step is perspective projection of the camera coordinates ($X_c$,$Y_c$,$Z_c$) to image coordinates (x,y), assuming an ideal (no distortions) pinhole camera model (Figure 2). The transformation in this step is not invertible because the depth information is lost. Step three is the non-linear transformation caused by lens distortion and step four is the image coordinate to pixel coordinate transformation.

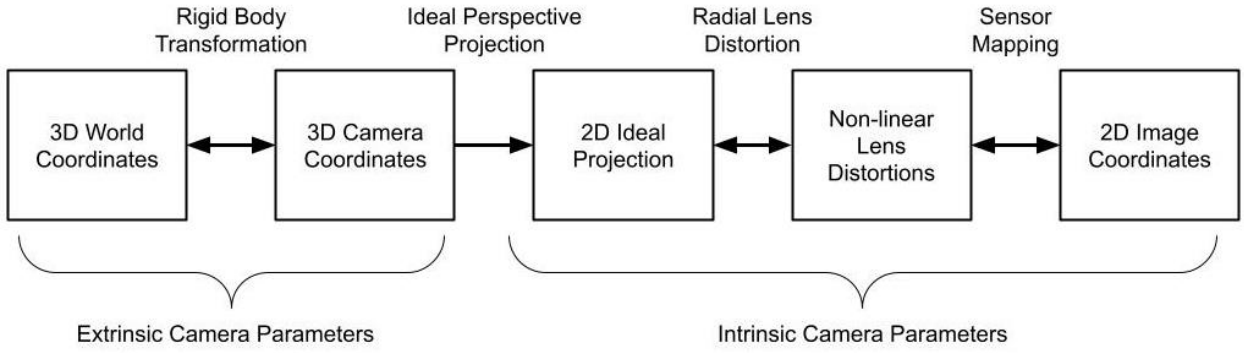


*Figure 1: Transformation from 3D world coordinates to 2D image coordinates (Tsai, 1987)*

The first step of the transformation is dependent on the extrinsic parameters of the camera, while the rest of the steps are dependent on its intrinsic parameters. Extrinsic parameters refer to the camera orientation in relation to the world and consist of a rotation matrix **R** and a translation vector **t**. Intrinsic parameters refer to the internal geometry of the camera system and consist of the focal length f, the principal point offset O'($x_o$,$y_o$), the pixel scale factors ($m_x$,$m_y$), the skew coefficient γ, and the lens distortion coefficients.

Equation (2.1) represents the perspective transformation for a *pinhole* camera model (Figure 2), where **P** is the projection matrix (or camera matrix) and **K** is the intrinsic matrix (camera matrix):

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (2.1)$$

$$K = \begin{bmatrix} f \cdot m_x & \gamma & x_o & 0 \\ 0 & f \cdot m_y & y_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad (2.2)$$

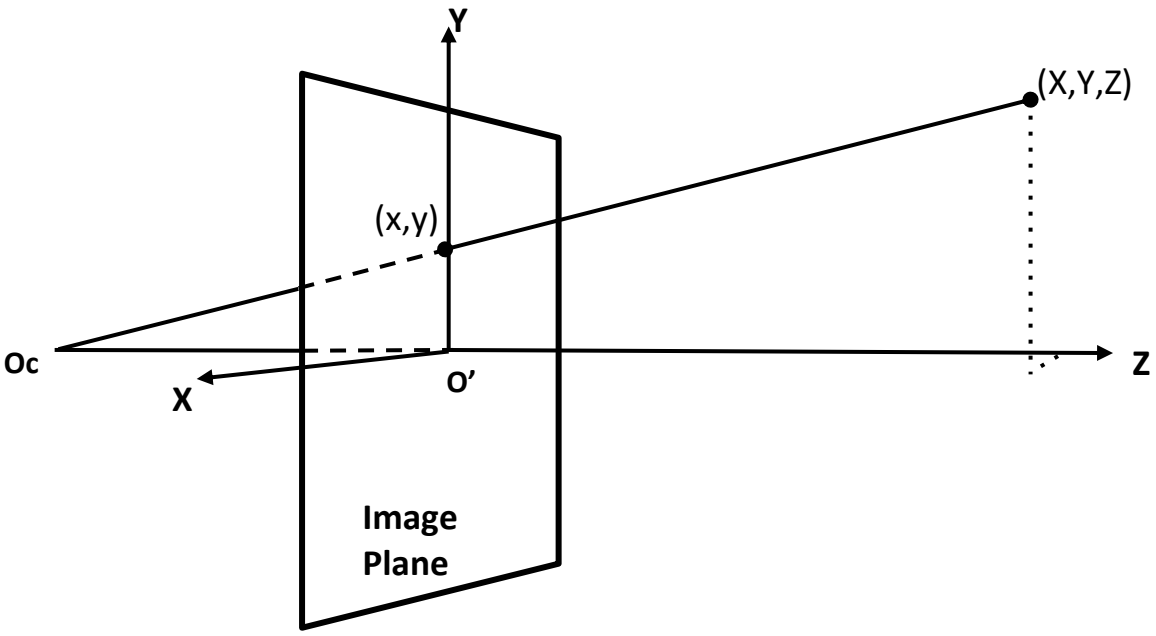


*Figure 2: Ideal pinhole camera model*

Out of all of these variables, the lens distortion coefficients are the only *non-linear* parameters, and for this reason some methods omit them during the initial calibration process and attempt to correct the errors afterwards by adopting the bundle adjustment technique (Triggs et al., 2000). The following subsections define the camera parameters in a more detailed way in order to better define the problem of camera calibration.

## 2.1 Extrinsic Parameters

Extrinsic camera parameters are also referred to as exterior orientation and represent the camera pose as a set of six parameters, three for the translation and three for the rotation around the three camera axes. The camera translation is defined by the ground coordinates of the center of projection $O_c(X_c,Y_c,Z_c)$, the rotation by the 3x3 rotation matrix $\mathbf{R}(r_{ij})$. The camera extrinsic parameters determine the camera pose, which is indispensable in order to understand the motion of the camera between multiple images. The 4x4 extrinsic matrix is deconstructed into its elements as:

$$[R|T] = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (2.3)$$

## 2.2 Intrinsic Parameters

Intrinsic camera parameters represent the geometrical and physical aspects of the camera. Because these parameters depend on the camera model, there are several features to be considered for the calibration. In general, one may say the more parameters considered for the camera calibration the better the accuracy of the reprojection will be; however, “over-parameterization” should always be avoided.

The most basic feature of an ideal pinhole camera model is the focal length f (also found in literature as camera constant c), which is basically the physical distance of the image plane from the pinhole. For the case of camera with thin lens, the distance between the lens and the image plane is the principal distance c, while the focal length represents the distance when the lens is focused at infinity. For this reason, the parameter considered for the camera calibration is the principal distance c (also found in literature as camera constant c). The relationship between focal length f and principal distance c of a thin lens is given by Equation (2.4), where S is the distance of the lens from the projected object (focusing distance):

$$\frac{1}{f} = \frac{1}{S} + \frac{1}{c} \qquad (2.4)$$

Another important phenomenon to be considered is that, as a rule, the trace O'($x_o$,$y_o$) of the center of projection on the image plane (principal point) does not coincide with the center of the image coordinates (x = y = 0). This can be corrected with a simple displacement of the image coordinates origin to the principal point O'($x_o$,$y_o$). Abiding by these characteristics, the projection matrix K' of an ideal lens camera with principal point offset is given by Equation (2.5):

$$\boldsymbol{K'} = \begin{bmatrix} c & 0 & x_o \\ 0 & c & y_o \\ 0 & 0 & 1 \end{bmatrix} \qquad (2.5)$$

Because the main focus of modern photogrammetry is, of course, on digital cameras, the structure of the camera sensor has also to be accounted for. Particularly, the shape of the pixels of a real sensor is not perfectly square but rectangular (difference in scale between the two axes). This scale difference can be calculated either by denoting a scale factor for each direction, $m_x$ and $m_y$, or by normalizing the scale for one axis and defining the differential scale factor m (aspect ratio) for the other axis. Additionally, the camera sensor may demonstrate a skewness instead of being perfectly orthogonal. Considering the angle θ as the angle formed between the two axis of the image plane, the skewed coordinates $x_s$,$y_s$ can be determined by Equation (2.6). Figure 3 illustrates the skewness of the image plane. For simplification of the calibration matrix, the skewness coefficient γ is used for the transformations. These parameters represent the affine distortions of the camera. The camera matrix of an affine sensor **K**'' is presented in Equation (2.7).

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} 1 & -\cot(\theta) \\ 0 & \frac{1}{\sin(\theta)} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2.6)$$

$$\boldsymbol{K''} = \begin{bmatrix} c & \gamma & x_o \\ 0 & c(1+m) & y_o \\ 0 & 0 & 1 \end{bmatrix} \qquad (2.7)$$

where $\gamma = -c \cdot tan(\theta)$

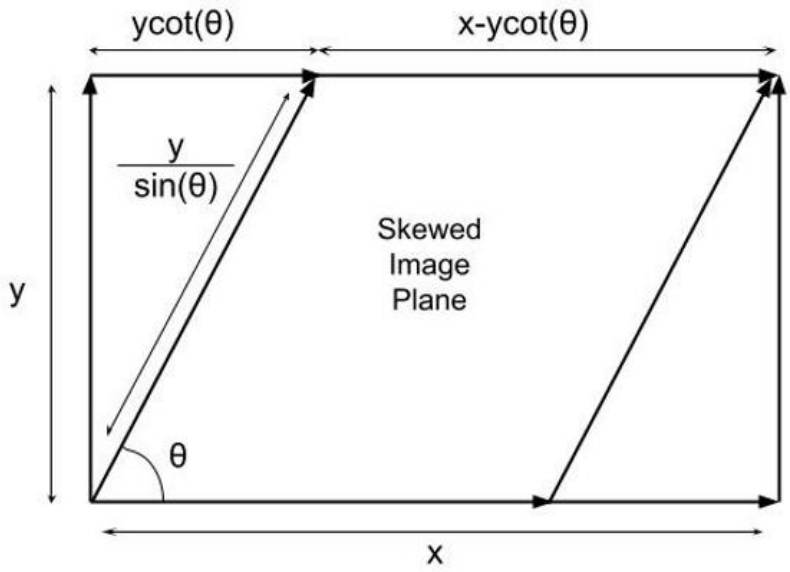


*Figure 3: Image plane with skewness*

## 2.3 Lens Distortions

Up to this point, only the linear transformations have been taken into account. By assuming a distortion free camera model, a straight world line is expected to be a straight line in the image plane. In reality, the camera lens distorts the image because of imperfections of the lens shape and the lens to sensor composition. According to Weng et al. (1992), there are three types of lens distortion. The first is the radial distortion caused by the imperfect curvature of the lens and is indicated by the light rays bending more when near the edges of the image than they do at its center. This distortion is expressed as a radial displacement of the image points. When the displacement is directed inwards, it is referred to as pincushion distortion, while when the displacement is directed outwards it is referred to as barrel distortion. A geometric representation of the radial distortions is illustrated in Figure 4. The polynomial expression in Equation (2.8) expresses the symmetric radial distortion for a centered lens, where ρ is the radial distance from the center of the image (in fact from the principal point) and $k_i$ for i=1,2,3… are the radial distortion coefficients.

$$\Delta r = k_1\rho^3 + k_2\rho^5 + k_3\rho^7 + \cdots \qquad (2.8)$$

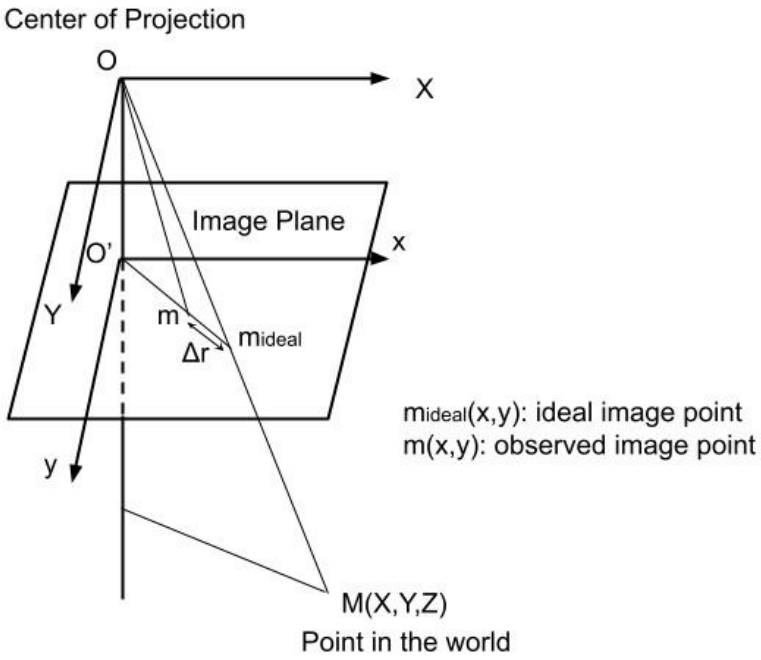


*Figure 4: Pinhole camera projection model with radial distortions*

The second type of lens distortion is the decentering distortion. This is a non-symmetric distortion that can be caused by the misalignment of the lens elements in a camera. This type of distortion is a combination of *non-symmetric* radial distortion and tangential distortion. The effects of radial and tangential distortion are illustrated by Figure 5. The displacement caused by decentering distortion can by calculated by Equation (2.9) for the x axis and (2.10) for the y axis.

$$\Delta t_x = p_1(r^2 + 2(x - x_0)^2) + 2p_2(x - x_0)(y - y_0) \qquad (2.9)$$

$$\Delta t_y = p_2(r^2 + 2(y - y_0)^2) + 2p_1(x - x_0)(y - y_0) \qquad (2.10)$$

where $r^2 = (x - x_0)^2 + (y - y_0)^2$ and $p_1$, $p_2$ are the decentering distortion coefficients.

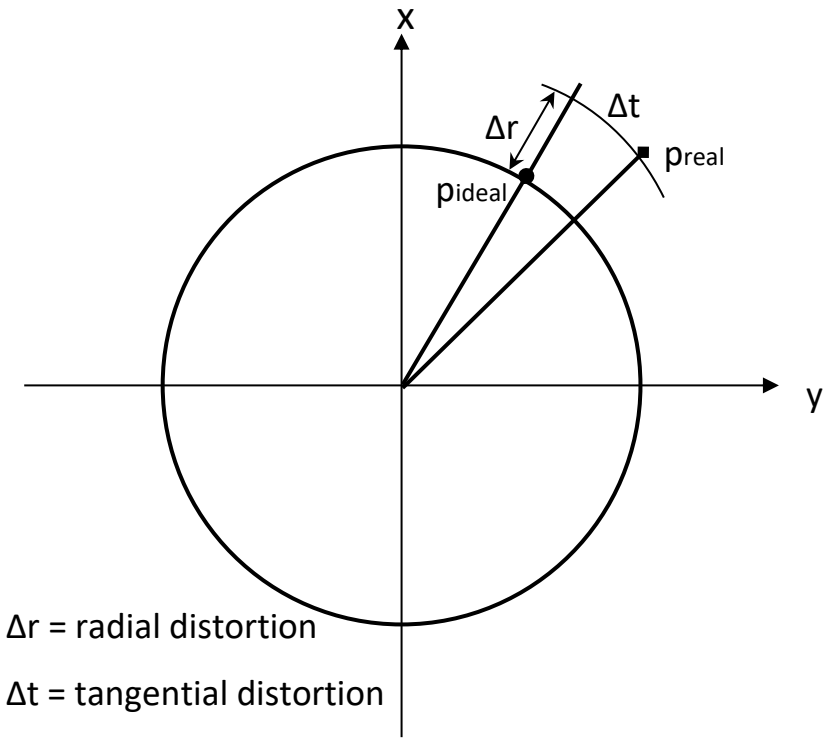


*Figure 5: Representation of radial and tangential distortion*

The third type of distortion is the thin prism distortion caused by a slight tilt between the lens and the camera, as illustrated by Figure 6. This type of distortion also causes both radial distortion and tangential distortion. For simplification purposes, the aggregate of these non-linear distortions can be combined in the lens distortion correction terms Δx(x,y) for the x axis and Δy(x,y) for the y axis. The complete camera matrix is thus expressed:

$$\boldsymbol{K} = \begin{bmatrix} c & \gamma & x_o + \Delta x \\ 0 & c(1+m) & y_o + \Delta y \\ 0 & 0 & 1 \end{bmatrix} \qquad (2.11)$$

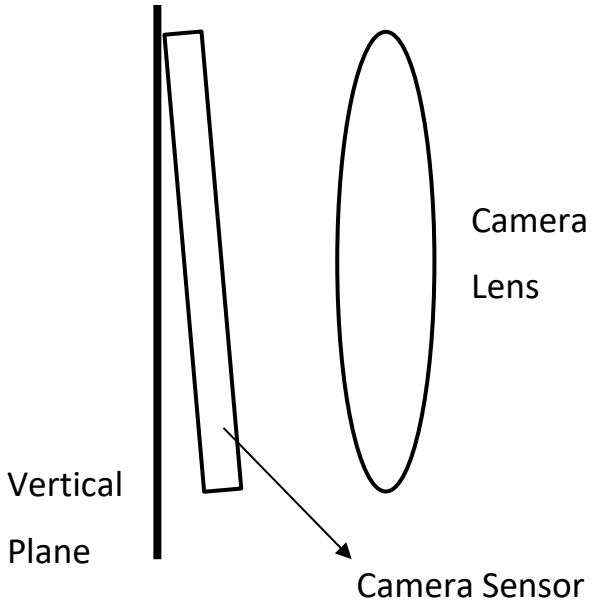


*Figure 6: Camera assembly where sensor and lens are not parallel*

# Section 3: Traditional Calibration Methods

The proposition of traditional camera calibration methods is that if an object with known geometry is depicted in images, then the camera parameters can be calculated by simple mathematical transformations. The real-world coordinates of the known object are associated with the image coordinates so that the unknown projection matrix can be solved for. The most widely adopted algorithmic approaches representing these methods are the direct linear transform (DLT) (Abdel-Aziz & Karara, 2015), Tsai's two-step method (Tsai, 1987), dual-plane method (Martins et al., 1981; Wei & De Ma, 1991), and Zhang's calibration method (Zhang, 2000).

## 3.1 Single Image Approach

Single-image calibration presupposes the existence of an array of well-distributed points of known world coordinates (3D test-fields). Such test-fields may be appropriate for either terrestrial or even for aerial cameras. Clearly, such arrays are difficult to create and, even more so, to maintain. Hence, they are now rarely used. Of interest here, however, is mainly the algorithmic treatment of single images for the purposes of calibration.

The DLT algorithm (initially presented by Abdel-Aziz & Karara as early as 1971) is a convenient simplified approach to the problem of single-image camera calibration. The 11 independent elements of the projection matrix are calculated using linear equations for a set of 6 or more reference world points. The process is not iterative (although it can be subsequently refined iteratively); therefore, calculations are fast and simple. This algorithm can output accurate results but only after the lens distortions have been corrected (Heikkila & Silven, 1997).

The basis of the DLT method is the well-known collinearity condition. Equation (3.1) is a rendition of the transformation of the world coordinates to image coordinates, and Equations (3.2) and (3.3) express the transformations for each coordinate:

$$\begin{bmatrix} x \\ y \\ -c \end{bmatrix} = \lambda \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X - X_c \\ Y - Y_c \\ Z - Z_c \end{bmatrix} \qquad (3.1)$$

where λ is a scale factor

$$x + c\frac{a_{11}(X-X_c)+a_{12}(Y-Y_c)+a_{13}(Z-Z_c)}{a_{31}(X-X_c)+a_{32}(Y-Y_c)+a_{33}(Z-Z_c)} = 0 \qquad (3.2)$$

$$y + c\frac{a_{21}(X-X_c)+a_{22}(Y-Y_c)+a_{23}(Z-Z_c)}{a_{31}(X-X_c)+a_{32}(Y-Y_c)+a_{33}(Z-Z_c)} = 0 \qquad (3.3)$$

The DLT method permutates the Equations (2.2) and (2.3) to (2.4) and (2.5), where $L_i$ for i=1…11 are the DLT parameters.

$$x + \frac{L_1X+L_2Y+L_3Z+L_4}{L_9X+L_{10}Y+L_{11}Z+1} = 0 \qquad (3.4)$$

$$y + \frac{L_5X+L_6Y+L_7Z+L_8}{L_9X+L_{10}Y+L_{11}Z+1} = 0 \qquad (3.5)$$

For n number of control points in the 3D space with known coordinates ($x_n$,$y_n$), the Equation (3.6) associates the resulting system of equations. The camera calibration is then resolved by least-square estimation, as denoted by Equation (3.7).

$$\begin{bmatrix} x_1 \\ y_1 \\ \dots \\ x_n \\ y_n \end{bmatrix} = \lambda \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1X_1 & -x_1Y_1 & -x_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1X_1 & -y_1Y_1 & -y_1Z_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -x_nX_n & -x_nY_n & -x_nZ_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -y_nX_n & -y_nY_n & -y_nZ_n \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_{10} \\ L_{11} \end{bmatrix} \qquad (3.6)$$

$$U = A \cdot L \Rightarrow L = [A^T A]^{-1} A^T \cdot U \qquad (3.7)$$

The mayor benefit of the DLT algorithm is the simplicity and convenience it offers. However, its accuracy is severely restricted by the need for an ideal camera setting and

optimal control points. Several improvements have been suggested over the years to account for these DLT limitations (Hatze, 1988; Heikkila & Silven, 1997; Zhao et al., 2016).

As an improvement to the DLT method, Tsai's method considers the lens distortions as a radial distortion factor. This method is separated into two steps; and thus it is also referred to as the two-step method. In the first step, the extrinsic parameters are solved by linear equations, while in the second step, the intrinsic parameters are solved by non-linear optimization. This calibration method assumes the camera model to adhere to the framework presented in Figure 4.

There are two different cases for which Tsai's method can be used, one for planar control points and the other for non-planar control points. Assuming that a 2D planar pattern is used as a calibration object, an important prerequisite for this method is that the camera scale factor m is known, usually supplied by the manufacturer. Moreover, the principal point O' is assumed to be on the center of the camera's digital sensor. The first step starts with the estimation of the rotation matrix **R**, Equation (3.8), and part of the translation matrix **T**, Equation (3.9), by solving Equation (3.10).

$$\boldsymbol{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \qquad (3.8)$$

$$\boldsymbol{T} = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \qquad (3.9)$$

$$[y_i X_i \quad y_i Y_i \quad y_i \quad -x_i X_i \quad x_i Y_i] \begin{bmatrix} {T_Y}^{-1} r_1 \\ {T_Y}^{-1} r_2 \\ {T_Y}^{-1} T_X \\ {T_Y}^{-1} r_4 \\ {T_Y}^{-1} r_5 \end{bmatrix} = x_i \qquad (3.10)$$

This linear system consists of 5 unknowns, so it needs more than five known control points from the calibration pattern in order to be solved. The factor $|T_x|$ is arbitrarily scaled

to ±1, so as to calculate the other factors. Remark that the sign of $T_x$ is determined by choosing a control point that is far from the image center, calculating the projection coordinates, and checking that the result is consistent with the observed image coordinates. The other factors are then calculated by solving the Equations (3.11).

$$\begin{aligned} r_1 &= {T_Y}^{-1} r_1, \quad & r_2 &= \left({T_Y}^{-1} r_2\right) T_Y, \quad & r_4 &= \left({T_Y}^{-1} r_4\right) T_Y, \\ r_5 &= \left({T_Y}^{-1} r_5\right) T_Y, \quad & T_X &= \left({T_Y}^{-1} T_X\right) T_Y \end{aligned} \tag{3.11}$$

The remaining part of the rotation matrix can be solved by Equation (3.12).

$$\boldsymbol{R} = \begin{bmatrix} r_1 & r_2 & \sqrt{(1 - {r_1}^2 - {r_2}^2)} \\ r_4 & r_5 & s\sqrt{(1 - {r_4}^2 - {r_5}^2)} \\ r_7 & r_8 & r_9 \end{bmatrix} \tag{3.12}$$

where $s = -\mathrm{sgn}(r_1 r_4 + r_2 r_5)$

Having calculated most of the extrinsic parameters, the second step of Tsai's method consists of estimating the intrinsic parameters, including the radial distortions. The principal distance and the factor $T_z$ are calculated by solving the linear system in Equation (3.13).

$$\begin{bmatrix} u_i & -\Delta_y y_i \end{bmatrix} \begin{bmatrix} f \\ T_Z \end{bmatrix} = w_i \Delta_y y_i \tag{3.13}$$

where $u_i = r_4 X_i + r_5 Y_i + r_6 \cdot 0 + T_Y$ and $w_i = r_7 X_i + r_8 Y_i + r_9 \cdot 0$

For the estimation of the radial distortion coefficients and the correct principal point O', a non-linear optimization method has to be used to minimize Equation (3.14). This step is referred to as bundle adjustment.

$$\sum_{i=1}^{N} (x_i - x_{pi})^2 + \sum_{i=1}^{N} (y_i - y_{pi})^2 \tag{3.14}$$

where $x_{pi}$ and $y_{pi}$ are the predicted image coordinates based on the estimated values.

This process can also be performed for non-planar 3D calibration points and for multiple images in order to achieve better precision with small changes to the algorithm.

The results of this method are very accurate and the implementation is simple, therefore it has been very popular for various photogrammetric applications (Choi & Seo, 2011; Gee et al., 2015). However, it has also been criticized for being very sensitive to errors (Tiscareño et al., 2019), which is detrimental for low-quality images or inexperienced users.

## 3.2 Multi-Image Approach

Zhang's method utilizes a planar pattern, usually a chessboard pattern, in order to streamline the detection of feature points. This method requires a set of at least three images where the planar pattern is placed at different positions and orientations relative to the camera. The relations of the feature points are used for the linear calculation of the projection matrix, and subsequently, the camera parameters are extracted using a closed-form solution. Next, the radial distortion coefficients are calculated by solving the linear least-squares, and, finally, the maximum likelihood criterion is applied to refine the results.

The target plane is assumed to be at Z=0 of the world coordinates. The planar projection of that plane can be defined by the homography **H** of Equation (3.15).

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \boldsymbol{K}[r_1 \quad r_2 \quad r_3 \quad t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \qquad (3.15)$$

$$[h_1 \quad h_2 \quad h_3] = \lambda \boldsymbol{K}[r_1 \quad r_2 \quad t] \qquad (3.16)$$

where λ is a scale factor.

Since $r_1$ and $r_2$ are orthonormal, the Equations (3.17) and (3.18) can be derived. These equations are related to the image of the absolute conic, which is explained in Section 5.

$$h_1{}^T \boldsymbol{K}^{-T} \boldsymbol{K}^{-1} h_2 = 0 \qquad (3.17)$$

$$h_1{}^T \boldsymbol{K}^{-T} \boldsymbol{K}^{-1} h_1 = h_2{}^T \boldsymbol{K}^{-T} \boldsymbol{K}^{-1} h_2 \qquad (3.18)$$

To solve the linear equations, a symmetric 3x3 matrix **B** is defined by Equation (3.19).

$$\boldsymbol{B} = \boldsymbol{K}^{-T} \boldsymbol{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \qquad (3.19)$$

$$\boldsymbol{B} = \begin{bmatrix} \frac{1}{c_x{}^2} & -\frac{s}{c_x{}^2 c_y} & \frac{y_0 s - x_0 c_y}{c_x{}^2 c_y} \\ -\frac{s}{c_x{}^2 c_y} & \frac{s^2}{c_x{}^2 c_y{}^2} + \frac{y_0}{c_y{}^2} & -\frac{s(y_0 s - x_0 c_y)}{c_x{}^2 c_y{}^2} - \frac{y_0}{c_y{}^2} \\ \frac{y_0 s - x_0 c_y}{c_x{}^2 c_y} & -\frac{s(y_0 s - x_0 c_y)}{c_x{}^2 c_y{}^2} - \frac{y_0}{c_y{}^2} & \frac{(y_0 s - x_0 c_y)^2}{c_x{}^2 c_y{}^2} + \frac{y_0{}^2}{c_y{}^2} + 1 \end{bmatrix} \qquad (3.20)$$

where $c_x = c * m_x$ and $c_y = c * m_y$

The matrix **B** can be expressed as a vector **b** in equation (3.21).

$$\boldsymbol{b} = \begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix} \qquad (3.21)$$

From the above, the Equation (3.22) is derived.

$$h_i{}^T \boldsymbol{B} h_j = v_{ij}{}^T \boldsymbol{b} \qquad (3.22)$$

where $h_i$ is the vector for column i=1,2,3 of homography **H**.

$$h_i = \begin{bmatrix} h_{i1} \\ h_{i2} \\ h_{i3} \end{bmatrix} \qquad (3.23)$$

$$v_i = \begin{bmatrix} h_{i1}h_{j1} \\ h_{i1}h_{j2} + h_{i2}h_{j1} \\ h_{i2}h_{j2} \\ h_{i3}h_{j1} + h_{i1}h_{j3} \\ h_{i3}h_{j2} + h_{i2}h_{j3} \\ h_{i3}h_{j3} \end{bmatrix} \qquad (3.24)$$

By associating Equations (3.17) and (3.18) with Equation (3.24), the Equation (3.25) is inferred.

$$\begin{bmatrix} {v_{12}}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \Rightarrow Vb = 0 \qquad (3.25)$$

where **V** contains an array for all observed values and therefore is a 2nx6 matrix for n equal to the number of images obtained for the calibration.

Given that the homography matrix **H** can be solved by the DLT method in subsection 3.1, the linear system of Equations (3.25) can be solved for **b**. Since the system has 6 unknowns, there needs to be at least 3 images to arrive at a unique solution for **b**. The intrinsic parameters can be extracted from vector **b** by solving Equations (3.26).

$$\begin{aligned} y_0 &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - {B_{12}}^2} \\ \lambda &= B_{33} - \frac{{B_{13}}^2 + y_0(B_{12}B_{13} - B_{11}B_{23})}{B_{11}} \\ c\, m_x &= \sqrt{\frac{\lambda}{B_{11}}} \\ c\, m_y &= \sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - {B_{12}}^2}} \\ \gamma &= \frac{-B_{12}(c\, m_x)^2 c\, m_y}{\lambda} \\ x_0 &= \frac{\gamma y_0}{c\, m_y} - \frac{B_{13}(c\, m_x)^2}{\lambda} \end{aligned} \qquad (3.26)$$

Considering that the camera matrix **K** has been estimated, the extrinsic parameters can be subsequently calculated by Equations (3.27).

$$\begin{aligned} r_1 &= \lambda \boldsymbol{K}^{-1} h_1 \\ r_2 &= \lambda \boldsymbol{K}^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda \boldsymbol{K}^{-1} h_3 \\ with\ \lambda &= \frac{1}{\|\boldsymbol{K}^{-1} h_1\|} = \frac{1}{\|\boldsymbol{K}^{-1} h_2\|} \end{aligned} \qquad (3.27)$$

To estimate the radial distortion coefficients, Equation (3.28) is used to associate the first two distortion coefficients with the ideal coordinates ($x_i$,$y_i$), the observed coordinates (x,y) and the observed normalized image coordinates (u,v).

$$\begin{bmatrix} (x_i - x_0)(u^2 + v^2) & (x_i - x_0)(u^2 + v^2)^2 \\ (y_i - y_0)(u^2 + v^2) & (y_i - y_0)(u^2 + v^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} x - x_i \\ y - y_i \end{bmatrix} \qquad (3.28)$$

By combining all control points for all images, the resulting Equation (3.29) is solved by linear least-squares, as given by Equation (3.30).

$$\boldsymbol{D}\boldsymbol{k} = \boldsymbol{d} \qquad (3.29)$$

where $k = [k_1, k_2]^T$

$$\boldsymbol{k} = (\boldsymbol{D}^T \boldsymbol{D})^{-1} \boldsymbol{D}^T \boldsymbol{d} \qquad (3.30)$$

The results can be refined by applying a non-linear optimization method to minimize the distance between estimated projected image points and observed image points, similar to Tsai's method in subsection 3.1. Through the process, the properties of the rotation matrix **R** have to be considered to estimate the best rotation matrix from a general 3×3 matrix.

This method is much more complicated than the ones mentioned above but the results are very accurate and the approach is highly flexible. This has made it one of the

most popular calibration algorithms, getting implemented in various open-source and commercial software, like MATLAB (Mathworks Inc., 2016) and OpenCV (Eser, 2021).

## 3.3 Comparisons between traditional methods

Table 1 summarizes in the author's opinion the advantages and disadvantages of the most popular traditional calibration methods. The flexibility of each method is considered as its ability to be adapted for different photogrammetric applications, while the complexity is assessed by the user knowledge required in conjunction with the computational cost.

The image quantity required for each method is important because methods that can work with only a single image offer versatility for applications with limited views. On the contrary, single-image methods tend to result in lower precision and noise sensitivity. The DLT algorithm has the capacity to get high accuracy for the results but only if the lens distortions are insignificant or have been pre-corrected by other means. In terms of flexibility, the DLT algorithm has the advantage of having a very low computational cost and only requiring one image, which means that the complexity is very low and therefore can be used for a variety of simple applications. However, the lack of accurate lens distortion calibration and the need for appropriate accurate 3D control points reduces the flexibility of the DLT algorithm. Tsai's method is more flexible due to the ability to calibrate radial distortions but counterbalances by having a slightly more complex calibration, due to the non-linear step. The accuracy of Tsai's method is also very sensitive to noise which can make it undesirable for various applications where very accurate control points are not present. Zhang's method has a substantial increase in complexity due to more images being required and the calculations having higher computation costs. Regardless of the complexity, Zhang's method is highly flexible and accurate through a more robust calibration framework that is not sensitive to noise, considers radial distortions, and can easily be adapted for a variety of applications. Moreover, the chessboard calibration pattern that is used for Zhang's method is easy to be acquired, since it is a common and simple pattern.

It is apparent that these methods provide high accuracy but are not very flexible and require a good understanding of the method by the user. The practical advancement for

these methods is the automation of the calibration process, which is discussed in the next section.

Table 1: Traditional Calibration Methods Comparisons

| Method | Image Quantity | Accuracy | Flexibility | Complexity |
|---|---|---|---|---|
| DLT (Abdel-Aziz & Karara, 2015) | 1 | High, for no dis-tortions | Medium | Low |
| Tsai (Tsai, 1987) | 1 or more | High, but sensitive to errors | Medium | Medium |
| Zhang (Zhang, 2000) | At least 3 | High | High | High |

# Section 4: Automated Calibration Methods

The automation of the traditional camera calibration methods can be achieved by the development of algorithms that automatically detect planar patterns within the image. The procedure of the photogrammetric application still needs to ensure that such an object exists within the images, but once the images are obtained the calibration process is fully automatic. The most common calibration target for such methods has been the 2D chessboard pattern because of how effortless its creation is; nevertheless, many other designs have also been proposed.

## 4.1 State-of-the-art Automated Methods

An approach for automatic calibration using 2D chessboard patterns has been presented by Douskos et al. (2007). This approach uses corner extraction and point ordering to detect the chessboard pattern in multiple images, the vanishing points for an initial estimation of the camera parameters, and bundle adjustment for the final parameter estimation. It cannot not calculate the actual extrinsic orientation of the camera because of the chessboard's symmetry. The results are as accurate as other planar pattern methods, considering that enough good-quality images are used. It is a very simple and effective approach that has also been developed into an open-source software called FAUCCAL (Fully Automatic Camera Calibration) (Douskos et al., 2009).

A very similar but more sophisticated approach was developed Wohlfeil et al. (2019). The proposed approach uses a chessboard pattern that covers almost the entire image plane and contains specific markers in order to calculate the camera orientation. The advantages of this approach are the low computational cost and the very accurate results. A concern about this method is the construction of a sufficiently large pattern, which can be unreasonable for applications where the camera is located far away from the target. The authors suggest using the markers to configure multiple chessboard patterns instead of one.

The employment of multiple chessboard patterns within each image was examined by Grammatikopoulos et al. (2019). The proposed approach uses multiple unordered coplanar chessboards instead of a single one. The identical small chessboard patterns are

placed in a semi-structured composition on a bigger plane, and the pattern detection and parameter estimations follow a similar methodology as the previously mentioned approaches. Figure 7 shows examples for the images used for the implementation of this method. The results are as accurate as other approaches with the added benefit of an easier-to-implement configuration for specific applications. The application that motivated this approach is the Unmanned Aerial Vehicle (UAV) photogrammetry, where the camera is situated very far from the calibration target and therefore the construction of a sufficiently large pattern is impractical.

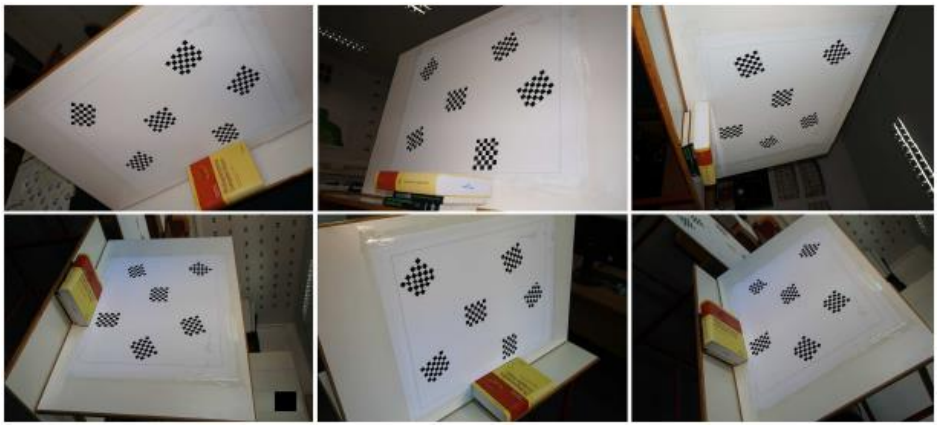

*Figure 7: Image dataset for calibration using multiple unordered coplanar chessboards by* (Grammatikopoulos et al., 2019)

Another approach to improve upon the chessboard pattern method was suggested by Xu et al. (2016). The proposition of this approach is to calibrate based on the 2D perpendicular lines extracted from a 2D chessboard pattern, instead of feature points. The objective of this approach is to reduce the instability of the traditional calibration methods by utilizing the geometrical information of perpendicular lines. An example to illustrate the line recognition of this method is presented in Figure 8. The results are indeed more accurate and resistant to noise when compared to Zhang's and Tsai's methods. The approach is also able to handle larger distances of the pattern from the camera in a more reliable way.

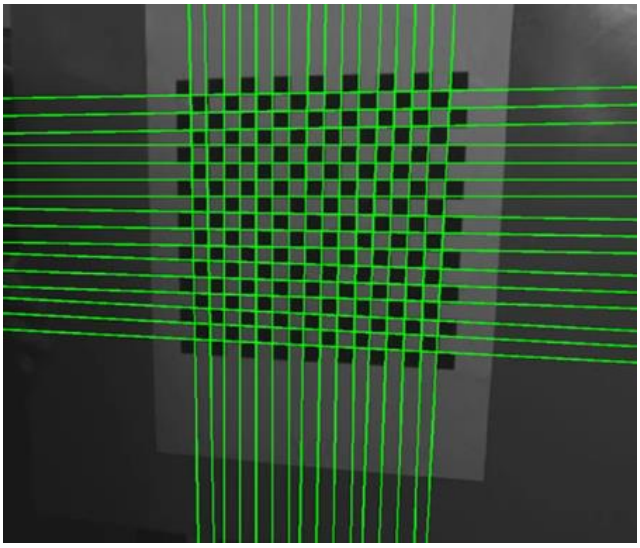

*Figure 8: Image example for the recognition of 2D lines from a chessboard pattern by* ( Xu et al., 2016)

Instead of the traditional chessboard pattern, Chen & Pan (2020) suggested the use of a synthetic random speckle pattern, illustrated in Figure 9. For this approach, the feature points of the calibration target are detected using the digital image correlation (DIC) algorithm (Pan et al., 2013), and subsequently the camera parameters are calculated using Zhang's method. Compared to chessboard patterns, this method offers better automatic feature point extraction and improved lens distortion correction. While the results are very accurate, this approach is limited by the speckle pattern resolution, which needs to be large enough to be correctly perceived by the camera sensor. This can be a significant drawback for low-quality digital cameras.

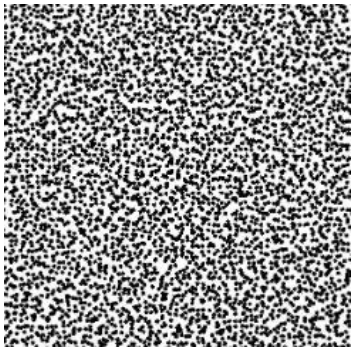

*Figure 9: Example of random speckle pattern*

As an alternative to the 2D calibration targets presented so far, Forbes et al. (2002) examined automatic camera calibration using 3D target objects. The 3D structure is marked with coded target patterns so that feature points can be detected without knowing the object geometry beforehand. The calibration process combines Tsai's method with a pose estimation algorithm for an initial estimation and then refines the parameters using bundle adjustment. The precision of this approach has been reported as very high, and the configuration is relatively flexible. However, the construction of the 3D calibration target is much more complex than that of a 2D plane.

*Figure 10: Example of 3D calibration object by* (Forbes et al., 2002)

## 4.3 Comparisons between automated methods

Our summary and comparison of the featured methods is presented in Table 2. All methods presented here have high adequate accuracy to be comparable to previous traditional calibration methods. The method that utilizes 2D perpendicular lines and that using a synthetic random speckle pattern can even surpass the precision of traditional methods under specific conditions. In terms of flexibility, certain methods are constrained by impractical requirements. In particular, the full-image chessboard is infeasible for long-range photography or any other application where the chessboard pattern has to cover a large surface. The synthetic random speckle pattern is also limited by the camera resolution and the availability of such a pattern in a realistic setting. The 2D chessboard pattern

approach of the FAUCCAL software relies upon the variety in images to cover the different perspectives in order to produce good results. Contrary to the restricted character of these methods, the approaches of multiple unordered coplanar chessboards and the 2D perpendicular lines attempt to rectify difficult applications by offering better flexibility.

The complexity of these methods is another important factor that has to be evaluated, as the purpose of automating the camera calibration process is to simplify the procedure. On that account, the methods that only require a chessboard pattern photographed from different perspectives are quite simple to implement and have low computational cost, depending only on the optimization of the pattern detection algorithms. The full-image chessboard method is a slight exception because it needs special markers on top of the chessboard pattern. The method of multiple unordered coplanar chessboards is considered more complex than other methods because the setup of multiple chessboards is more elaborate. Similarly, the synthetic random speckle pattern is more difficult to produce than its chessboard counterpart. The 3D calibration target is, of course, also more difficult to create. Overall, the presented methods have low computational costs and in principle do not require any advanced expertise on the subject to be conducted.

Each of these methods presents certain advantages over the others that can make it stand out for specific applications. The full image chessboard method achieves very high precision for very low computation cost for close-range photography. The multiple unordered coplanar chessboards method presents a practical solution for UAVs or long-range photography. The 2D perpendicular lines method is very resilient to noisy or low-resolution images. The synthetic random speckle pattern method has the ability to rectify images with significant lens distortions. Finally, the 3D calibration target method can be adapted for a very wide variety of applications where the camera is moving around the world space without moving the calibration target for each image.

All these approaches are accessible, low-cost, and automated enough to be usable in real applications, provided that each method can be more effective for different applications. In the author’s view, the methods presenting a promising solution for common photogrammetric problems are the multiple unordered coplanar chessboards and the 2D perpendicular lines. These two methods offer a cost-effective solution for calibrations

where the camera may be quite far away from the target. Further research on this topic would be valuable for the photogrammetric community.

Table 2: Automated Calibration Methods Comparisons

| Method | Accuracy | Flexibility | Complexity | Advantage |
|---|---|---|---|---|
| FAUCCAL (Douskos et al., 2009) | High, but no extrinsic orientation | Medium | Low | Fully implemented on open-source software |
| Full Image Chessboard (Wohlfeil et al., 2019) | High | Low, because of the target size | Low | Very low computation cost for very high accuracy |
| Multiple Unordered Coplanar Chessboards (Grammatikopoulos et al., 2019) | High | High | Medium | Very practical for applications where the camera is far from target |
| 2D Perpendicular Lines (G. Xu et al., 2016) | Very high | High | Low | Resilience to noise |
| Synthetic Random Speckle Pattern (Chen & Pan, 2020) | Very high | Medium | Medium | High accuracy for cameras with high lens distortions |
| 3D Calibration Target (Forbes et al., 2002) | High | High | Medium | Very flexible method |

# Section 5: Self-Calibration Methods

As explained in the previous section, the traditional calibration methods require some known geometry within the image, which in some cases could be impossible. That is where the need for self-calibration methods emerges. The premise of camera self-calibration is that, by establishing specific constraints, the camera parameters can be recovered simply from a series of images. The constraints can be applied to either the intrinsic parameters, the scene, or the camera motion.

## 5.1 Theoretical background for self-calibration

An important concept for self-calibration is the absolute conic Ω, which is a conic at the infinity plane Π∞, as seen in Figure 11. The absolute conic's perspective projection is invariant to the camera position and only depends on the camera's intrinsic parameters. That attribute is useful because the intrinsic parameters can be derived from the motion between multiple images of the same camera.

A practical representation of the absolute conic is constituted by the Dual Absolute Quadric Ω*, defined in Equation (5.1), whose image projection ω* is calculated by Equation (5.2). The absolute quadric is a quadric at the plane at infinity, whose planes are tangent to the absolute conic.

$$\boldsymbol{\Omega}^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (5.1)$$

$$\boldsymbol{\omega}_i^* \approx \boldsymbol{P}_i \boldsymbol{\Omega}^* \boldsymbol{P}_i{}^T \qquad (5.2)$$

where i is an index for the camera viewpoint.

Solving Equation (5.2) using (2.1) and (5.1), the image of the absolute quadric is defined by Equation (5.3):

$$\boldsymbol{\omega}_i^* \approx \boldsymbol{K}_i \boldsymbol{K}_i^T \qquad (5.3)$$

The image of the absolute conic can be translated between two camera viewpoints by using the homography **H** of the plane at infinity, where the conic is situated, as in Equation (5.4).

$$\boldsymbol{\omega}_i \approx \boldsymbol{H}_{ij}^{\infty\,-T} \boldsymbol{\omega}_j \boldsymbol{H}_{ij}^{\infty\,-1} \qquad (5.4)$$

By considering the absolute conic present in all images and applying the epipolar transformation, we obtain the Equation (5.5) which is known as the Kruppa equation (Faugeras et al., 1992):

$$[\boldsymbol{e}]_\times^T \boldsymbol{K}\boldsymbol{K}^T [\boldsymbol{e}]_\times \cong \boldsymbol{F}^T \boldsymbol{K}\boldsymbol{K}^T \boldsymbol{F} \qquad (5.5)$$

where **F** is the fundamental matrix and $\mathbf{e}_i$ are the image epipoles

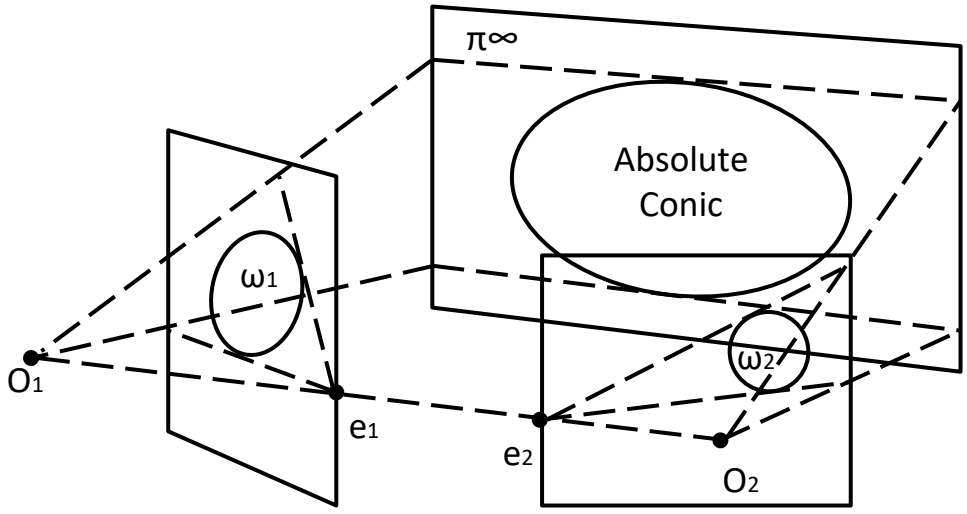


*Figure 11: The absolute conic **Ω** (located in the infinity plane π∞) and its projection ω in the images*

## 5.2 Conventional self-calibration methods

The first self-calibration method presented by Faugeras et al. (1992) proposed that by solving the Kruppa equations for at least three images the camera parameters will be obtained. For each pair of images, the first step is to find points matched between the

images and compute their epipoles. Then, a Kruppa equation (5.5) can be computed for each image of the pair. Once at least 6 Kruppa equations have been solved, the intrinsic parameters can be extracted from the image of the absolute conic ω.

The reason that three images are required is because each pair of images produces two Kruppa equations and 5 intrinsic parameters are unknown. It is obviously important for this method that the intrinsic parameters are constant for all images.

While reviewing this method, it is important to acknowledge that Kruppa equations have problematic specific cases that cannot be resolved, as analyzed by Kahl (1999) and Sturm (2000a). Figure 12 displays such a case, where the geometry of the scene is problematic because one pair of viewpoints has an absolute quadric that is identical with another absolute quadric of a different pair of viewpoints.

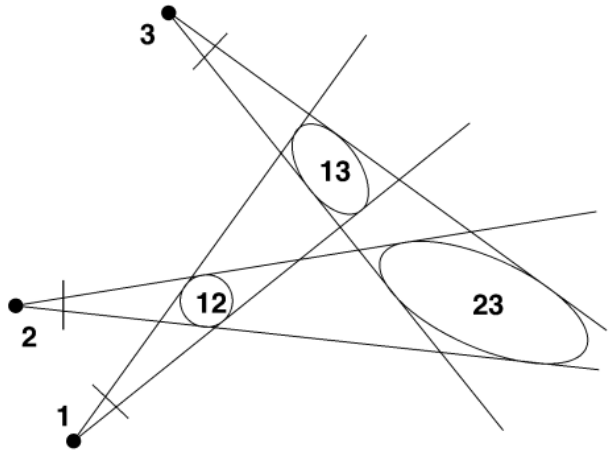


*Figure 12: Problematic pair of viewpoints where Kruppa equations fail, by* (Sturm, 2000a)

An application of this method was examined by Luong & Faugeras (1997) for 3D reconstruction. The application produced acceptable results with the only input being the point correspondences between the images. The precision was lower than traditional calibration methods and the computational cost much higher, but it proved that camera calibration can be performed without using calibration targets. Several other approaches based on this method were proposed and produced similar results (Hartley, 1994a; Pollefeys & Van Gool, 1999; Pollefeys & van Goot, 1997).

As a substitution for the absolute conic method, Triggs (1997) introduced the concept of absolute quadric. The method starts with an initial projective reconstruction and then

the absolute quadric and the image conic are found at the same time by using nonlinear (sequential quadratic programming) or quasi-linear techniques. The camera calibration and Euclidean structure are recovered by correcting the initial estimated projections. Finally, the bundle adjustment method is used to refine the results. Like in the absolute conic case, 3 images from different viewpoints and constant intrinsic parameters are enough for the calibration. The procedure still has a high computational cost but is simpler than the absolute conic method.

The requirement of constant intrinsic parameters for the implementation of these self-calibration methods poses a fundamental problem for practical applications. Especially in modern digital photography where the feature of automatic focus to adapt to the scene is very common. Pollefeys et al. (1999) proved that self-calibration using the Kruppa equations is possible for varying intrinsic camera parameters as long as the image rows and columns are orthogonal, meaning that there is no skewness in the image plane. Similarly, Heyden & Åström (1997) proved that knowing even one intrinsic parameter is sufficient for the absolute quadric self-calibration method. A more robust and flexible algorithm was proposed by Seo & Heyden (2000), where the orthogonality constraints are used after an initial projective reconstruction to iteratively calculate the varying intrinsic parameters. The important restriction of this algorithm is that the images must have a constant aspect ratio and no skewness.

A method for self-calibration for a camera with a fixed position and variable rotation has been proposed by Hartley (1994b). In this case, the camera calibration is accomplished by finding matching points and calculating the projective transformations between the images. This method has lower computational cost and provides better accuracy than most self-calibration methods but addresses a limited number of practical cases. A similar method was proposed by Moons et al. (1996), with the difference that the relative object-camera translation remains constant between the images. Further analysis on rotating camera self-calibration has been done by Zheng & Li (2014).

Another approach for self-calibration is performed by utilizing the planar geometry in the world. Triggs (1998) described a method of self-calibration by observing a plane shape located in the images. This method establishes a numerical algorithm that iteratively optimizes the constraints of the scene plane to projected plane collineations. The

first iteration of this method was efficient but unstable and unoptimized, so Malis & Cipolla (2000a, 2000b) made several improvements. Those and other planar self-calibration techniques were evaluated by Gurdjos & Sturm (2003) and were found to produce good results.

## 5.3 Comparisons between conventional self-calibration methods

The characteristics of the various self-calibration methods, as concentrated by the author, are presented in Table 3. A considerable disadvantage of these applications is that they tend to produce low-accuracy results, unlike traditional calibration methods where high accuracy is expected. The planar self-calibration techniques tend to perform more accurately than other self-calibration methods. The rotating camera method is one of the more accurate self-calibration methods, which however is not that important because the constraint of the camera position significantly limits the flexibility of this method. On the contrary, all other presented self-calibration methods are highly flexible and can be utilized for a great variety of applications where the precision threshold is low.

The complexity of these methods is difficult to measure because it often depends on the application where they are being applied. In general, the absolute conic and absolute quadric methods have very high computational costs due to the amount of non-linear equations that have to be solved. At the same time, these methods require a good understanding from the perspective of the user in order to be implemented. The orthogonality constraint method and the planar self-calibration method have the advantage of requiring fewer computations than the preceding self-calibration methods. The rotating cameras method is the exception because the fixed camera potion simplifies the computations by a considerable amount.

Table 3: Self-Calibration Methods Comparisons

| Method | Image Quantity | Accuracy | Flexibility | Complexity |
|---|---|---|---|---|
| Absolute Conic (Faugeras et al., 1992) | 3 or more | Low | High | High |
| Absolute Quadric (Triggs, 1997) | 3 or more | Low | High | High, but lower than absolute conic |
| Orthogonality Constraint (Seo & Heyden, 2000) | 3 or more | Low | High, assuming camera with no skewness | Medium |
| Rotating Camera (Hartley, 1994b) | 3 or more | Medium | Low, because the camera is stationary | Low |
| Planar Self-calibration (Malis & Cipolla, 2000a) | 5 or more | Medium | High | Medium |

# Section 6: State-of-the-art Self-calibration Methods

As observed in the previous section, the domain of camera self-calibration is intricate and constantly improving. There is no established ideal approach because each method is meant to accommodate different applications with different constraints. However, a common aspect between the most conventional self-calibration methods is that the calibration precision is generally lower than that of traditional methods. This section examines the state-of-the-art self-calibration approaches that attempt to overcome the typical self-calibration limitations.

## 6.1 State-of-the-art self-calibration methods

A novel approach to self-calibration was introduced by Wong et al. (2003), which takes advantage of the contours of objects of revolution. An example of the geometry of such an object is illustrated by Figure 13. The assumption is that objects of revolution, like bowls and vases, are very common in photographs, so their symmetric features can be utilized. Specifically, the silhouette of the object of revolution is detected, its revolution axis is defined based on the silhouette's symmetry, and by relating the vanishing points with the revolution axis an initial parameter estimation is done. The initial estimation is then optimized by repeating for multiple images. This method was also implemented and evaluated by Hödlmoser et al. (2010) for low-cost cameras. The results were accurate, but not as precise as traditional methods.

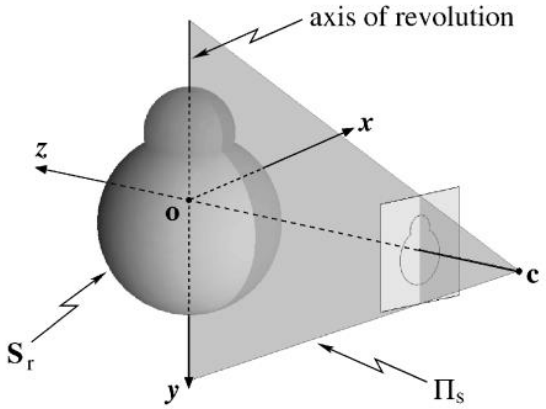


*Figure 13: Object of revolution projected to a camera, by* (Wong et al., 2003)

A camera calibration algorithm based on vanishing points was proposed by Grammatikopoulos et al. (2007). The algorithm starts by automatically extracting three finite vanishing points from at least one image and performs line-fitting combined with lens distortions estimation. The intrinsic camera parameters are then estimated from the vanishing point pairs. Lastly, the calibration adjustment procedure is implemented to refine the estimated parameters. The progress for each step of this method is illustrated by Figure 14. This approach relies on the orthogonality and parallelism that exist in man-made environments. The algorithm is fully automatic and produces accurate enough results.

*Figure 14: Left: Extracted edges, Middle: Vanishing lines, Right: Vanishing lines after lens distortion correction, by* (Grammatikopoulos et al., 2007)

A practical planar self-calibration approach was presented by Herrera et al. (2016). The framework of this approach starts with feature matching, continues with projective reconstruction and bundle adjustment, follows it up with homography-based calibration, and ends with a metric reconstruction and bundle adjustment. The results are surprisingly accurate for a self-calibration method and can even rival the accuracy of traditional methods.

Adam et al. (2013) utilized unstructured planar objects to implement an automated plane-based self-calibration method that is accomplished directly by bundle adjustment. The method requires multiple images of geometry with a distinct texture so that the Scale-Invariant Feature Transform (SIFT) operator can be used to find matching points. Using the matching point, the images can be connected with inter-image homographies via a

Random Sample Consensus (RANSAC) method. After the initial estimations, the unsuitable images are discarded so the self-calibration bundle adjustment process is sufficient for the parameter calibration. This method is also fully automatic, requiring only an appropriate use case.

A more specialized camera self-calibration method for multiple cameras was introduced by Wang et al. (2019) for use in motion capture systems. For the purposes of this method, a moving small object, like a tennis ball, is considered as a single point on the image plane. That point is moved in a specific motion and through projective reconstruction and reprojection the full trajectory of the point is calculated. The projection matrix for each camera is then estimated based on that trajectory. The resulting precision is efficient for the purposes of motion capture.

## 6.2 Comparisons between state-of-the-art self-calibration methods

The state-of-the-art self-calibration methods are compared in Table 4. The objects of revolution method appears to achieve better accuracy than most conventional self-calibration methods but not by a significant margin. This method also has lowered flexibility because the existence of an appropriate shape in the image is not granted. The vanishing points method is similar in terms of accuracy and also depends on the scene geometry, even though the existence of orthogonal objects in the scene is much more likely. The improved planar self-calibration method is more robust than other methods and even surpasses the traditional calibration methods in terms of accuracy. The primary advantage of planar self-calibration is that planar objects are very common in most photogrammetric applications. Herrera et al. (2016) suggest that this method can even replace traditional chessboard patterns. The unstructured planar objects method presents similar results with the only difference that a distinct texture is used to define the object, which is slightly less common and therefore reduces the flexibility of the system. All of these methods are accompanied by very high computational costs due to the advanced non-linear computations involved with the process. An exception to this is the method for the unmarked human motion capture system, which is more specialized and less complicated for an inexperienced user.

Most of the presented methods appear to have less flexibility than the conventional self-calibration methods while gaining significantly on the accuracy scale. This phenomenon indicates that the camera calibration process has to adapt to the relevant application it is used for in order to achieve better results.

Table 4: State-of-the-art Self-Calibration Methods Comparisons

| Method | Accuracy | Flexibility | Complexity |
|---|---|---|---|
| Objects of Revolution (Wong et al., 2003) | Medium | Medium | High |
| Vanishing Points (Grammatikopoulos et al., 2007) | Medium | Medium | High |
| Improved Planar Self-calibration (Herrera et al., 2016) | High | High | High |
| Unstructured Planar Objects (Adam et al., 2013) | High | Medium | High |
| Unmarked Human Motion Capture System (J. Wang et al., 2019) | High, considering the application and not general use | Low | Low |

# Section 7: Conclusions

The camera calibration process has been a subject of research for decades and through this time it has matured and advanced into more sophisticated algorithms. The discussion around the subject has shifted from just increasing accuracy to decreasing complexity and reinforcing convenience. Because of the amount of photogrammetric and computer vision applications where camera calibration is needed, an analogous amount of research is being conducted. As this thesis has exemplified, there is no universal camera calibration method that can be integrated into every possible application.

The general trend in camera calibration has been to automate the process as much as possible in order to accommodate the influx of inexperienced users who have easy access to photogrammetric applications. While self-calibration methods have come a long way in terms of accuracy, traditional methods are still very popular among researchers because of the low computational costs. In theory, self-calibration should be more desirable because of the lack of prerequisites, but in practice the cost of the calibration target for automated traditional methods is negligible. There are still many applications where calibration targets are impossible or problematic, like UAV photogrammetry, virtual reality, and augmented reality.

In this thesis, the fundamentals of camera calibration, and specifically the camera intrinsic and extrinsic parameters, were described and analyzed. Some of the most popular traditional camera calibration methods were explained and compared in terms of performance, expressed by accuracy, flexibility and complexity. For the category of automated calibration methods, a variety of state-of-the-art calibration methods, that use calibration target but are fully automatic, were presented and compared with the same metrics of performance. Then, the basis for self-calibration methods was described and the most important self-calibration methods were outlined and compared. Lastly, a variety of state-of-the-art self-calibration methods were analyzed and compared by the performance metrics explained above.

A lot of the state-of-the-art methods reviewed in this thesis look very promising for further research, on how they can apply to different photogrammetric applications or what algorithmic optimizations can be done to improve the performance. A branch that has great potential and was not included in this thesis is that of camera calibration using deep

learning, which is being steadily adopted by more researchers. This review can hopefully be a guide to reinforce future studies.

# Bibliography


Abdel-Aziz, Y. I., & Karara, H. M. (2015). Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, *81*(2), 103–107. https://doi.org/10.14358/PERS.81.2.103

Adam, K., Kalisperakis, I., Grammatikopoulos, L., Karras, G., & Petsa, E. (2013). AUTOMATIC CAMERA CALIBRATION FOR CULTURAL HERITAGE APPLICATIONS USING UNSTRUCTURED PLANAR OBJECTS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-5/W2*, 1–6. https://doi.org/10.5194/isprsarchives-xl-5-w2-1-2013

Baratoff, G., Neubeck, A., & Regenbrecht, H. (2002). Interactive multi-marker calibration for augmented reality applications. *Proceedings - International Symposium on Mixed and Augmented Reality, ISMAR 2002*, 107–116. https://doi.org/10.1109/ISMAR.2002.1115079

Chen, B., & Pan, B. (2020). Camera calibration using synthetic random speckle pattern and digital image correlation. *Optics and Lasers in Engineering*, *126*, 105919. https://doi.org/10.1016/j.optlaseng.2019.105919

Choi, K., & Seo, Y. (2011). Automatic initialization for 3D soccer player tracking. *Pattern Recognition Letters*, *32*(9), 1274–1282. https://doi.org/10.1016/j.patrec.2011.03.009

Davis, W. A., Paquette, L., Stampfler, R., & Caelli, T. M. (2017). *A new camera calibration method for robotic vision*. 84. https://doi.org/10.1117/12.2294327

Douskos, V., Grammatikopoulos, L., Kalisperakis, I., Karras, G., & Petsa, E. (2009). FAUCCAL: An open source toolbox for fully automatic camera calibration. *XXII CIPA Symposium on Digital Documentation, Interpretation & Presentation of Cultural Heritage, Kyoto, Japan*. https://www.researchgate.net/publication/228951969_FAUCCAL_an_open_source_toolbox_for_fully_automatic_camera_calibration

Douskos, V., Kalisperakis, I., & Karras, G. (2007). Automatic calibration of digital cameras using planar chess-board patterns. *Proceedings of the 8th Conference on Optical 3-D Measurement Techniques*, 9--12. https://www.researchgate.net/publication/228345254_Automatic_calibration_of_digital_cameras_using_planar_chess-board_patterns

Eser, A. Y. (2021). *OpenCV Camera Calibration*. https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html

Fathi, H., & Brilakis, I. (2014). Multistep Explicit Stereo Camera Calibration Approach to Improve Euclidean Accuracy of Large-Scale 3D Reconstruction. *Journal of Computing in Civil Engineering*,

*30*(1), 04014120. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000454

Faugeras, O. D., Luong, Q. T., & Maybank, S. J. (1992). Camera self-calibration: Theory and experiments. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *588 LNCS*, 321–334. https://doi.org/10.1007/3-540-55426-2_37/COVER

Forbes, K., Voigt, A., & Bodika, N. (2002). An inexpensive, automatic and accurate camera calibration method. *In Proceedings of the Thirteenth Annual South African Workshop on Pattern Recognition (PRASA)*, 1–6. http://www.vision.caltech.edu/bouguetj/

Forlani, G., Roncella, R., & Nardinocchi, C. (2015). Where is photogrammetry heading to? State of the art and trends. *Rendiconti Lincei*, *26*(1), 85–96. https://doi.org/10.1007/s12210-015-0381-x

Fraser, C. S. (2013). Automatic camera calibration in close range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, *79*(4), 381–388. https://doi.org/10.14358/PERS.79.4.381

Garcia, M. V. Y., & de Oliveira, H. C. (2021). The influence of flight configuration, camera calibration, and ground control points for digital terrain model and orthomosaic generation using unmanned aerial vehicles imagery. *Boletim de Ciencias Geodesicas*, *27*(2), e2021015. https://doi.org/10.1590/S1982-21702021000200015

Gee, T., Delmas, P., Stones-Havas, N., Sinclair, C., Van Der Mark, W., Li, W., Friedrich, H., & Gimel'Farb, G. (2015). Tsai camera calibration enhanced. *Proceedings of the 14th IAPR International Conference on Machine Vision Applications, MVA 2015*, 435–438. https://doi.org/10.1109/MVA.2015.7153104

Gibson, S., Cook, J., Howard, T., Hubbold, R., & Oram, D. (2002). Accurate camera calibration for off-line, video-based augmented reality. *Proceedings - International Symposium on Mixed and Augmented Reality, ISMAR 2002*, 37–46. https://doi.org/10.1109/ISMAR.2002.1115068

Grammatikopoulos, L, Adam, K., Petsa, E., & Karras, G. (2019). Camera calibration using multiple unordered coplanar chessboards. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(2/W18), 59–66. https://doi.org/10.5194/isprs-archives-XLII-2-W18-59-2019

Grammatikopoulos, Lazaros, Karras, G., & Petsa, E. (2007). An automatic approach for camera calibration from vanishing points. *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*(1), 64–76. https://doi.org/10.1016/j.isprsjprs.2007.02.002

Gurdjos, P., & Sturm, P. (2003). Methods and geometry for plane-based self-calibration. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*. https://doi.org/10.1109/cvpr.2003.1211394

Hartley, R. I. (1994a). Algorithm for self calibration from several views. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 908–912. https://doi.org/10.1109/cvpr.1994.323923

Hartley, R. I. (1994b). Self-calibration from multiple views with a rotating camera. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *800 LNCS*, 471–478. https://doi.org/10.1007/3-540-57956-7_52

Hatze, H. (1988). High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. *Journal of Biomechanics*, *21*(7), 533–538. https://doi.org/10.1016/0021-9290(88)90216-3

Heikkila, J., & Silven, O. (1997). Four-step camera calibration procedure with implicit image correction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1106–1112. https://doi.org/10.1109/cvpr.1997.609468

Herrera, D., Kannala, C. J., & Heikkila, J. (2016). Forget the checkerboard: Practical self-calibration using a planar scene. *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016*. https://doi.org/10.1109/WACV.2016.7477641

Heyden, A., & Åström, K. (1997). Minimal conditions on intrinsic parameters for euclidean reconstruction. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1352*, 169–176. https://doi.org/10.1007/3-540-63931-4_212

Hödlmoser, M., Zollner, H., & Kampel, M. (2010). An Evaluation of Camera Calibration Methods Using Digital Low Cost Cameras. *Computer Vision Winter Workshop 2010, Libor ˇSpaˇcek, Vojtˇech Franc (Eds.)*, *June*, 1–7. http://cvlab.epfl.ch/software/

ITO, M. I. N. O. R. U. (1990). Robot vision modelling—camera modelling and camera calibration. *Advanced Robotics*, *5*(3), 321–335. https://doi.org/10.1163/156855391X00232

Kahl, F. (1999). Critical motions and ambiguous Euclidean reconstructions in auto-calibration. *Proceedings of the IEEE International Conference on Computer Vision*, *1*, 469–475. https://doi.org/10.1109/iccv.1999.791258

Kroeger, O., Huegle, J., & Niebuhr, C. A. (2019). An automatic calibration approach for a multi-camera-robot system. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, *2019-Septe*, 1515–1518. https://doi.org/10.1109/ETFA.2019.8869522

Läbe, T., & Förstner, W. (2004). Geometric stability of low-cost digital consumer cameras. *Proceedings of the 20th ISPRS Congress, Istanbul, Turkey*, 528–535.

Liao, K., Nie, L., Huang, S., Lin, C., Zhang, J., Zhao, Y., Gabbouj, M., & Tao, D. (2023). *Deep Learning for Camera Calibration and Beyond: A Survey*. http://arxiv.org/abs/2303.10559

Long, L., & Dongri, S. (2019). Review of Camera Calibration Algorithms. *Advances in Intelligent Systems and Computing*, *924*, 723–732. https://doi.org/10.1007/978-981-13-6861-5_61

Luong, Q. T., & Faugeras, O. D. (1997). Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *International Journal of Computer Vision*, *22*(3), 261–289. https://doi.org/10.1023/A:1007982716991

Malis, E., & Cipolla, R. (2000a). Self-calibration of zooming cameras observing an unknown planar structure. *Proceedings - International Conference on Pattern Recognition*, *15*(1), 85–88. https://doi.org/10.1109/icpr.2000.905281

Malis, E., & Cipolla, R. (2000b). Multi-view constraints between collineations: Application to self-calibration from unknown planar structures. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1843*, 610–624. https://doi.org/10.1007/3-540-45053-x_39

Martins, H. A., Birk, J. R., & Kelley, R. B. (1981). Camera models based on data from two calibration planes. *Computer Graphics and Image Processing*, *17*(2), 173–180. https://doi.org/10.1016/0146-664X(81)90024-1

Martins, P. F., Costelha, H., Bento, L. C., & Neves, C. (2020). Monocular camera calibration for autonomous driving - A comparative study. *2020 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2020*, 306–311. https://doi.org/10.1109/ICARSC49921.2020.9096104

Mathworks Inc. (2016). *Estimate Geometric Parameters of a Single Camera*. https://www.mathworks.com/help/vision/ref/cameracalibrator-app.html

Moons, T., Van Gool, L., Proesmans, M., & Pauwels, E. (1996). Affine reconstruction from perspective image pairs with a relative object-camera translation in between. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(1), 77–83. https://doi.org/10.1109/34.476015

Pan, B., Li, K., & Tong, W. (2013). Fast, Robust and Accurate Digital Image Correlation Calculation Without Redundant Computations. *Experimental Mechanics*, *53*(7), 1277–1289. https://doi.org/10.1007/s11340-013-9717-6

Pollefeys, M., Koch, R., & Van Gool, L. (1999). Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*, *32*(1), 7–25. https://doi.org/10.1023/A:1008109111715

Pollefeys, M., & Van Gool, L. (1999). Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(8), 707–724. https://doi.org/10.1109/34.784285

Pollefeys, M., & van Goot, L. (1997). Self-calibration from the absolute conic on the plane at infinity. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1296*, 175–182. https://doi.org/10.1007/3-540-63460-6_115

Romero, R. P., & Quintana, I. R. (2023). Photogrammetry workflow for obtaining low-polygon 3D models using free software. *Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*. https://doi.org/10.24132/CSRN.3301.42

Seo, Y., & Heyden, A. (2000). Auto-calibration from the orthogonality constraints. *Proceedings - International Conference on Pattern Recognition*, *15*(1), 67–71. https://doi.org/10.1109/icpr.2000.905277

Setiawan, M., Rutzinger, M., Wichmann, V., Stoetter, J., & Sartohadi, J. (2013). Evaluation of methods for digital elevation model interpolation of tillage systems. *Journal of Natural Resources and Development*. https://doi.org/10.5027/jnrd.v3i0.13

Song, L., Wu, W., Guo, J., & Li, X. (2013). Survey on camera calibration technique. *Proceedings - 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2013*, *2*, 389–392. https://doi.org/10.1109/IHMSC.2013.240

Song, S., Chandraker, M., & Guest, C. C. (2016). High Accuracy Monocular SFM and Scale Correction for Autonomous Driving. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(4), 730–743. https://doi.org/10.1109/TPAMI.2015.2469274

Sturm, P. (2000a). Case against Kruppa's equations for camera self-calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(10), 1199–1204. https://doi.org/10.1109/34.879804

Sturm, P. (2000b). A method for 3D reconstruction of piecewise planar objects from single panoramic images. *Proceedings - IEEE Workshop on Omnidirectional Vision, OMNIVIS 2000*, 119–126. https://doi.org/10.1109/OMNVIS.2000.853818

Tiscareño, J., Albajez, J. A., & Santolaria, J. (2019). Analysis of different camera calibration methods on a cameraprojector measuring system. *Procedia Manufacturing*, *41*, 539–546. https://doi.org/10.1016/j.promfg.2019.09.041

Triggs, B. (1997). Autocalibration and the absolute quadric. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 609–614. https://doi.org/10.1109/cvpr.1997.609388

Triggs, B. (1998). Autocalibration from planar scenes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1406*, 89–105. https://doi.org/10.1007/BFb0055661

Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W. (2000). Bundle adjustment – a modern synthesis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1883*, 298–372. https://doi.org/10.1007/3-540-44480-7_21

Tsai, R. Y. (1987). A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Journal on Robotics and Automation*, *3*(4), 323–344. https://doi.org/10.1109/JRA.1987.1087109

Wang, G., Tsui, H. T., Hu, Z., & Wu, F. (2005). Camera calibration and 3D reconstruction from a single view based on scene constraints. *Image and Vision Computing*, *23*(3), 311–323. https://doi.org/10.1016/J.IMAVIS.2004.07.008

Wang, J., Miao, Z., Liang, Q., & Jiang, Z. (2019). Camera Self-Calibration for Unmarked Human Motion Capture System. *International Conference on Signal Processing Proceedings, ICSP*, *2018-Augus*, 480–484. https://doi.org/10.1109/ICSP.2018.8652333

Wang, Q., Fu, L., & Liu, Z. (2010). Review on camera calibration. *2010 Chinese Control and Decision Conference, CCDC 2010*, 3354–3358. https://doi.org/10.1109/CCDC.2010.5498574

Wei, G. Q., & De Ma, S. (1991). Two plane camera calibration: A unified model. *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 133–138. https://doi.org/10.1109/cvpr.1991.139675

Weng, J., Coher, P., & Herniou, M. (1992). Camera Calibration with Distortion Models and Accuracy Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*(10), 965–980. https://doi.org/10.1109/34.159901

Wilczkowiak, M., Boyer, E., & Sturm, P. (2001). Camera calibration and 3D reconstruction from single images using parallelepipeds. *Proceedings of the IEEE International Conference on Computer Vision*, *1*, 142–148. https://doi.org/10.1109/ICCV.2001.937510

Wohlfeil, J., Grießbach, D., Ernst, I., Baumbach, D., & Dahlke, D. (2019). Automatic camera system calibration with a chessboard enabling full image coverage. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(2/W13), 1715–1722. https://doi.org/10.5194/isprs-archives-XLII-2-W13-1715-2019

Wong, K. Y. K., Mendonça, P. R. S., & Cipolla, R. (2003). Camera calibration from surfaces of revolution.

*IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(2), 147–161. https://doi.org/10.1109/TPAMI.2003.1177148

Wu, W., Xu, M., Liang, Q., Mei, L., & Peng, Y. (2020). Multi-camera 3d ball tracking framework for sports video. *IET Image Processing*, *14*(15), 3751–3761. https://doi.org/10.1049/iet-ipr.2020.0757

Xu, G., Zheng, A., Li, X., & Su, J. (2016). A method to calibrate a camera using perpendicularity of 2D lines in the target observations. *Scientific Reports*, *6*(1), 1–15. https://doi.org/10.1038/srep34951

Xu, H., Lan, G., Wu, S., & Hao, Q. (2019). Online Intelligent Calibration of Cameras and LiDARs for Autonomous Driving Systems. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 3913–3920. https://doi.org/10.1109/ITSC.2019.8916872

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330–1334. https://doi.org/10.1109/34.888718

Zhao, Z., Ye, D., Zhang, X., Chen, G., & Zhang, B. (2016). Improved Direct Linear Transformation for Parameter Decoupling in Camera Calibration. *Algorithms 2016, Vol. 9, Page 31*, *9*(2), 31. https://doi.org/10.3390/A9020031

Zheng, K., & Li, X. (2014). The quantitative analysis of self-calibration based on rotating cameras. *Proceedings - 2014 7th International Congress on Image and Signal Processing, CISP 2014*, 520–524. https://doi.org/10.1109/CISP.2014.7003835