Unbiased Estimator for Distorted Conics in Camera Calibration

Chaehyeon Song, Jaeho Shin, Myung-Hwan Jeon, Jongwoo Lim, Ayoung Kim
Seoul National University
{chaehyeon, leah100, myunghwan.jeon, jongwoo.lim, ayoungk}@snu.ac.kr

Abstract

In the literature, points and conics have been major features for camera geometric calibration. Although conics are more informative features than points, the loss of the conic property under distortion has critically limited the utility of conic features in camera calibration. Many existing approaches addressed conic-based calibration by ignoring distortion or introducing 3D spherical targets to circumvent this limitation. In this paper, we present a novel formulation for conic-based calibration using moments. Our derivation is based on the mathematical finding that the first moment can be estimated without bias even under distortion. This allows us to track moment changes during projection and distortion, ensuring the preservation of the first moment of the distorted conic. With an unbiased estimator, the circular patterns can be accurately detected at the sub-pixel level and can now be fully exploited for an entire calibration pipeline, resulting in significantly improved calibration. The entire code is readily available from https://github.com/ChaehyeonSong/discocal.

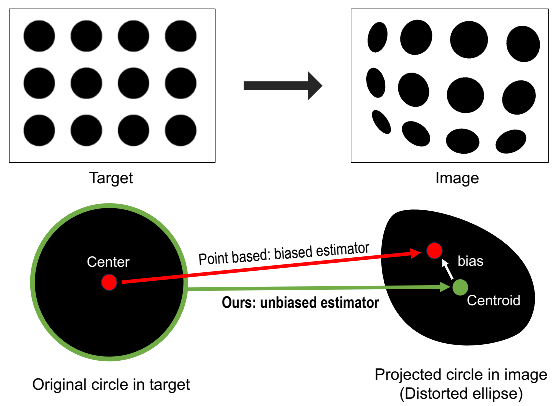


Figure 1. **Centroid estimation in distorted Images.** This figure illustrates the effects of camera projection and lens distortion on circular targets within an image. Due to these distortions, circles lose their conic properties and deform into distorted ellipses, making center point tracking challenging. (Red) failure of conventional control point estimation methods, as indicated by an incorrectly tracked center point. (Green) the proposed unbiased estimator accurately identifies the center point of the transformed ellipse, as derived from closed-form calculations.

1. Introduction

Camera calibration is essential in 3D computer vision and vision-based perception. Significantly, the processes of understanding the geometry from images rely on accurate camera calibration, including 3D dense reconstruction [17, 18], visual simultaneous localization and mapping (SLAM) [12, 19, 22], and depth estimation [9]. This fundamental problem has been tackled in various ways [4, 5, 14, 26, 27] and the most calibration methods exploit planar targets covered with particular patterns such as grid structure of squares [25, 28] or circles [10, 13]. This planar target-based approach requires precise measurements with an unbiased projection model of control points (i.e., corner of the square or centroid of the circle) to achieve accurate calibration results. Much literature has studied the distinct advantages and disadvantages of the two patterns. The checkerboard pattern ensures precise (unbiased) estimation in projective transformation and distortion, yet it is limited to pixel-level detection accuracy. On the other hand, the circular pattern [11] excels in achieving sub-pixel level detection accuracy, while the biased projection model has led to poor calibration results. Compensating the perspective bias using the conic feature is not sufficient, and the more dominant factor is the nonlinear lens distortion [16]. It degrades the geometric properties of the cone, making it difficult to estimate the centroid of the projected circle as described in Fig. 1.

In this paper, we push the envelope on the circular pattern by proposing an unbiased estimator in a closed-form solution to handle the distortion bias. To the best of our knowledge, this is the first analytic solution describing projected conic features under radial polynomial distortion. Inspired by the fact that moment representation is possible to describe the general distribution transformation under any polynomial mapping, we proved that the characteristics of the distorted conic, such as the centroid, could be expressed as a linear combination of n^{th} moments of the undistorted conic. Our work is not only limited to distortion models for pinhole camera calibration. Moreover, we suggest a general and differentiable approach to track the conic under arbitrary nonlinear transformations that can be approximated as polynomial functions.

By conducting both synthetic and real experiments, we have validated that our estimator is unbiased by investigating the reprojection error under various distortions and circle radius with known camera parameters. Our estimator is also applied to the calibration of RGB and thermal infrared (TIR) cameras which have difficulty in detecting control points due to the boundary blur effect (see Appendix 9.1). As a result, our method outperforms existing circular pattern-based methods and the checkerboard method in reprojection error and 6D pose manner. The main contribution of our work is summarized as follows.

- We pioneer the unbiased estimator for distorted conic to fully exploit the virtue of circular patterns inheriting a simple, robust, and accurate detector. In a blend of mathematical elegance and practical ingenuity, our work completes the missing piece in the conic-based calibration pipeline.
- We leverage the probabilistic concept of a moment, which had not been previously attempted in the calibration, and provide general moments of conic as an analytic form with thorough mathematical derivation and proof. This approach enables us to design an unbiased estimator for circular patterns.
- Our unbiased estimator improves the overall calibration performance when tested on both synthetic and real images. Especially, we showcase that our method yields substantially improved calibration results for TIR images, which often include high levels of blur, noise, and significant distortion. We open our algorithm to support the community using the conic features in their calibration.

2. Related Works

Planar pattern and control point. Zhang [28] and Sturm and Maybank [25] introduced a calibration method utilizing a planar target with some specific pattern printed on it. The planar target should include control points that can be easily and accurately extracted from the specific pattern. The checkerboard pattern comprising black and white squares was initially considered, and then the alternative pattern of circles with grid structure was introduced by Heikkila [10]. The checkerboard pattern uses the corners of squares as control points, and the circular pattern uses the centroid of projected circles in the image with subpixel accuracy [11]. Since the exact position of control points is crucial, an interactive way to refine the location of control points was also considered [7].

Unbiased estimator. To achieve accurate calibration results, the quality of measurement and estimation value of the control point are both vital. In contrast to a checkerboard pattern, which is proven to have an unbiased estimator in most cases, applying the same estimator to a circular pattern results in bias [16]. Specifically, the center point of circles in the target is not projected to the center point of the projected circle in the image. Several works [10, 14] challenged this problem by adopting conic-based transformation. These methods achieved unbiased estimators under linear transformation, such as perspective transformation. However, bias resulting from distortion is unresolved since conic characteristics are not preserved under nonlinear transformation. The bias originating from distortion is a dominant factor in most cameras [16]. To resolve this issue, Kannala and Brandt [13] introduced a generalized concept of the unbiased estimator for the circular pattern but could not derive an analytic solution for the given integral equation.

Methods without control point. The approach to directly use the conic characteristic has been developed simultaneously. By utilizing concentric circles, early works [4, 14] could obtain the image of absolute conic from the conic equations. Some works [26, 27] focused on the sphere, whose projected shape only depends on the configuration between the center point and the camera. Unfortunately, these methods also suffered from distortion bias, which destroys all crucial geometry of the conic or sphere. To deal with the nonlinear distortion function, Devernay and Faugeras [8] leveraged the line feature, which is always straight in undistorted images. Although this work did not explicitly estimate the intrinsic parameter, recent work [5] introduced the closed-form solution of the intrinsic parameter using line features with the undistorted images.

General model approach. The existing camera model assumes an ideal lens and approximates the nonlinear mapping with few parameters. Alleviating this assumption, a more general ray tracing model has been introduced. Schops et al. [23] suggested a general un-projection model using B-spline interpolation with nearest points. Another works [20] proposed pixel-wise focal length to consider general nonlinear mapping at the radial direction. These alternative approaches possess potential; however, existing applications of 3D vision still need geometric camera models. Therefore, in this paper, we focused on geometric model-based calibration, especially the pin-hole model.

3. Preliminary

3.1. Notations

A vector $\boldsymbol{p} \in \Re^n$ and a matrix $\boldsymbol{Q} \in \Re^{n\times n}$ are denoted by lowercase and uppercase in bold with the coordinate in the subscript. We set the target plane to be the same as the xy-plane of the world coordinate. Thus, $\boldsymbol{p}_w = (x_w, y_w)^\top$ is a 2D point in the target plane written in the target coordinate with $\boldsymbol{p}_i$ being a corresponding point in the image. This target point $\boldsymbol{p}_w = (x_w, y_w)^\top$ is projected to a point $\boldsymbol{p}_n$ in the normalized plane via perspective projection, and then $\boldsymbol{p}_n$ is mapped to a point $\boldsymbol{p}_i$ in image plane under distortion and intrinsic matrix.

Also, $\boldsymbol{Q}_w$ is a matrix representation of an ellipse in the target plane, and $\boldsymbol{Q}_n$ is an ellipse in the normalized plane. More details of these notations and the geometric relationship between the coordinates are described in Fig. 2. We use $\sim$ for a homogeneous vector representation. Also, $\tilde{\boldsymbol{p}} \simeq \tilde{\boldsymbol{q}}$ means that two vectors are identical up-to-scale.

3.2. Matrix representation of conic

A conic is a set of points (x, y) that satisfy the following equation:

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \tag{1}$$

which can be also written in matrix form as:

$$\boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x} = 0, \quad \boldsymbol{Q} = \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix}, \ \boldsymbol{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{2}$$

The symmetric matrix $\boldsymbol{Q}$ is the characteristic matrix of the conic. In general, the conic comprises an ellipse, parabola, and hyperbola, but we only concentrate on the ellipse in this paper since the images of circles are ellipses except for extreme cases. Given characteristic matrix $\boldsymbol{Q}$, the center of an ellipse is calculated as:

$$\tilde{\boldsymbol{p}} = \boldsymbol{Q}^{-1} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top. \tag{3}$$

Detailed derivation for geometric features of the ellipse is explained in Appendix 8.1.

3.3. Camera model

Adopting a pinhole camera model, a 2D point $\tilde{\boldsymbol{p}}_w$ in the target plane is projected to the image by:

$$\begin{aligned}
\boldsymbol{K} &= \begin{bmatrix} f_x & \eta & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, && (4) \\
\boldsymbol{T}_{cw} &= \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^\top & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \boldsymbol{r}_3 & \boldsymbol{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}, && (5) \\
\tilde{\boldsymbol{p}}_i &= \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \boldsymbol{K} \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \boldsymbol{r}_3 & \boldsymbol{t} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} && (6) \\
&\simeq \boldsymbol{K} \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \boldsymbol{t} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \quad (\because z_w = 0) && (7) \\
&\simeq \boldsymbol{K} \boldsymbol{E} \tilde{\boldsymbol{p}}_w \quad (\tilde{\boldsymbol{p}}_w \triangleq [x_w, y_w, 1]^\top) && (8) \\
&\simeq \boldsymbol{K} \tilde{\boldsymbol{p}}_n \simeq \boldsymbol{H} \tilde{\boldsymbol{p}}_w. && (9)
\end{aligned}$$

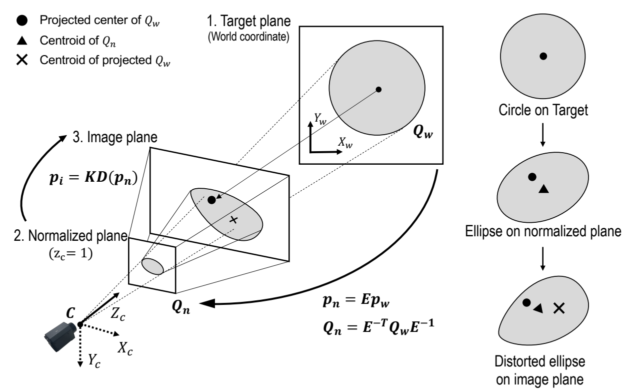


Figure 2. **Image projection geometry.** Due to projection and lens distortion, there exists some mismatch between the projected center of the circle on the target plane (circled dot), the projected center of the ellipse on the normalized plane(triangle sign), and the actual centroid of the shape on the image plane (crossed sign). The transformed shape, caused by non-linear distortion in the normalized plane, cannot be analytically described as the original conic. However, despite these distortions, our algorithm successfully locates the true centroid through moment tracking.

Here, $\boldsymbol{K}$ is the intrinsic matrix of the camera, $\boldsymbol{E}$ is the extrinsic parameter between the target coordinate and the camera coordinate, and $\boldsymbol{H}$ is the homography matrix. The z_w is zero since the p_w is a point located on the xy-plane. The above equation fully determines the relation between the target point and the projected point ($\tilde{\boldsymbol{p}}_n$) in the normalized plane. Unfortunately, this projected point ($\tilde{\boldsymbol{p}}_n$) further undergoes a nonlinear transformation (lens distortion) when projected onto the image plane. Between the two types of distortion, radial and tangential, we only consider radial distortion, which is known to be sufficient in most cameras [15, 20]. The radial distortion is typically modeled with a polynomial function [3, 6] as:

$$\begin{aligned}
s_n &= x_n^2 + y_n^2, && (10) \\
k &= \sum_{i=0}^{n_d} d_i s_n^i \qquad (d_0 = 1), && (11) \\
\tilde{\boldsymbol{p}}_d &= D(\tilde{\boldsymbol{p}}_n) = [x_d, y_d, 1]^\top = [kx_n, ky_n, 1]^\top, && (12) \\
\tilde{\boldsymbol{p}}_i &\simeq \boldsymbol{K} \tilde{\boldsymbol{p}}_d = \boldsymbol{K} D(\tilde{\boldsymbol{p}}_n) = \boldsymbol{K} D(\boldsymbol{E} \tilde{\boldsymbol{p}}_w), && (13)
\end{aligned}$$

where d_i are the distortion parameters and D is the distortion function. The value n_d is the maximum order of distortion parameters, typically less than three in existing calibration methods [2].

3.4. Measurement model (Control point)

The centroid of a black dot in an image is obtained as

$$\bar{\boldsymbol{p}}_i = \frac{\int w \boldsymbol{p}_i dA_i}{\int w dA_i} \approx \frac{\sum_j w_j \boldsymbol{p}_i(j)}{\sum_j w_j} \tag{14}$$

In the ideal case, w is $(255 - \text{intensity})$. In the real world, other light sources frequently distort the color value of the projected circle so that w is set to one for robustness.

4. Unbaised Estimator for Circular Patterns

4.1. Intuition: moment approach

The main objective of this section is to provide a thorough mathematical derivation of the unbiased estimator of the control point, which is the center point of the projected circle in the image originating from the target pattern. Tracking the control point under homography transformation is straightforward. An ellipse is projected to another ellipse under homography transformation $\boldsymbol{H}$ as:

$$\tilde{\boldsymbol{p}}_f \simeq \boldsymbol{H}\tilde{\boldsymbol{p}}_i, \tag{15}$$

$$\boldsymbol{Q}_f \simeq \boldsymbol{H}^{-\top}\boldsymbol{Q}_i\boldsymbol{H}^{-1}. \tag{16}$$

The subscript indicates the original frame i to transformed frame f. By combining Eqs. (3) and (16), the center of the transformed ellipse is calculated as $\boldsymbol{H}\boldsymbol{Q}_i^{-1}\boldsymbol{H}^\top(0,0,1)^\top$ and it is not the same as the transformed center point of the original ellipse, $\boldsymbol{H}\boldsymbol{Q}_i^{-1}(0,0,1)^\top$. Before adapting our approach, we define some concepts of geometry as preliminary.

Definition 1 (Shape)**.** *A shape A_{xy} is a set of points enclosed by the closed curve in xy-plane. $|A_{xy}|$ is the inner area of the closed curve. A_k denotes $A_{x_k y_k}$.*

The conic feature suffers from losing its characteristic under nonlinear transformations like distortion. To build an unbiased estimator under distortion, we utilize moment theory. In probability theory, $(i+j)^{\text{th}}$ moment of random variables X and Y is defined as $E[X^iY^j]$. By assuming points of the given shape lie on uniform distribution, the spatial average corresponds to the expectation of a random variable. We define the $(i+j)^{\text{th}}$ moment of a 2D shape in xy coordinate.

Definition 2 (Moment)**.** *For any $m, n \in \mathbb{Z}^*$(nonnegative integer set), $M_{xy}^{m,n}$ is $(m+n)^{th}$ moment of A_{xy} such that*

$$M_{xy}^{m,n} \triangleq \frac{1}{|A_{xy}|}\int x^m y^n dA_{xy}. \tag{17}$$

The first-moment factors ($M^{1,0}$ and $M^{0,1}$) are the center point of the shape, and the second-moment factors ($M^{2,0}$, $M^{1,1}$, and $M^{0,2}$) are related to the covariance of the shape. Therefore, an ellipse can be described by only using the first and second-moment factors. The matrix form of ellipse is $\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x} \leq 0$, and the moment form is $[M^{1,0}, M^{0,1}, M^{2,0}, M^{1,1}, M^{0,1}]$. Each representation has the same five-degree of freedom (DoF).

Theorem 1. *Given a polynomial function $D : (x,y) \longrightarrow (x',y')$ which is invertible in domain X, let $A_{xy} \subset X$ is transformed to $A_{x'y'}$ by D. For any $m, n \in \mathbb{Z}^*$, there exist $c_{ij}(D) \in \mathbb{R}$ and $p, q \in \mathbb{Z}^*$ such that*

$$|A_{x'y'}|M_{x'y'}^{m,n} = |A_{xy}|\sum_{i=0}^{p}\sum_{j=0}^{q} c_{ij}M_{xy}^{i,j}.$$

Proof. See Appendix 7.1. □

Theorem. 1 implies that for any polynomial mapping, the moment of the transformed shape can always be expressed by the linear combination of the moment of the original shape. The required number of p and q depends on the order of the polynomial function.

4.2. Tracking control point under distortion

Even though it is impossible to describe a distorted ellipse in any analytic form, its moment representation always exists by Theorem. 1. Despite using high-order moments, we only use the first moments, which is the centroid of shape. The reason is explained in Appendix 8.4.

To examine the shape under distortion mapping, let A_n be a shape in a normalized plane comprising points (x_n, y_n) and A_d be the distorted shape comprising points (x_d, y_d). Using Theorem. 1, the center of the distorted ellipse in the normalized plane can be calculated as follows.

$$w_{0r} = \sum_i (2i+1)d_i d_{r-i} \tag{18}$$

$$w_{1r} = \sum_i \sum_j (2i+1)d_i d_j d_{r-i-j} \tag{19}$$

$$1 = M_d^{0,0} = \frac{|A_n|}{|A_d|}\sum_{r=0}^{2n_d} w_{0r}\left[\frac{1}{|A_n|}\int s_n^r dA_n\right] \tag{20}$$

$$\bar{x}_d = M_d^{1,0} = \frac{|A_n|}{|A_d|}\sum_{r=0}^{3n_d} w_{1r}\left[\frac{1}{|A_n|}\int x_n s_n^r dA_n\right] \tag{21}$$

$$\bar{y}_d = M_d^{0,1} = \frac{|A_n|}{|A_d|}\sum_{r=0}^{3n_d} w_{1r}\left[\frac{1}{|A_n|}\int y_n s_n^r dA_n\right] \tag{22}$$

Refer to Eq. (10) for the symbol s_n defined as $x_n^2 + y_n^2$ and Eq. (20) is need for calculate $\frac{|A_n|}{|A_d|}$. The $2n_d$ and $3n_d$ come from calculating the Riemannian Metric and the details are provided in the Appendix 8.2.

The centroid of the projected shape in the image plane is readily calculated as below with details in Appendix 8.3.

$$\bar{x}_i = f_x\bar{x}_d + \eta\bar{y}_d + c_x \tag{23}$$

$$\bar{y}_i = f_y\bar{y}_d + c_y. \tag{24}$$

Therefore, to build the unbiased estimator of $[\bar{x}_i, \bar{y}_i]$, we only need $\frac{1}{|A_n|}\int s_n^r dA_n$, $\frac{1}{|A_n|}\int x_n s_n^r dA_n$, and

$\frac{1}{|A_n|}\int y_n s_n^r dA_n$ for every integer r from 0 to $3n_d$, where (x_n, y_n) is a point in the set $A_n = \{\boldsymbol{p}_n | \tilde{\boldsymbol{p}}_n^T \boldsymbol{Q}_n \tilde{\boldsymbol{p}}_n \leq 0\}$. We define these values as a vector $\boldsymbol{v}_n^r$:

$$\boldsymbol{v}_n^r \triangleq \begin{bmatrix} \frac{1}{|A_n|}\int x_n s_n^r dA_n \\ \frac{1}{|A_n|}\int y_n s_n^r dA_n \\ \frac{1}{|A_n|}\int s_n^r dA_n \end{bmatrix} \tag{25}$$

Calculating $\boldsymbol{v}_n^r$ directly from an ellipse on a normalized plane is not straightforward; therefore, we divide the process into two steps as shown in Fig. 3.

Theorem 2 (Rotation Equivariant)**.** *For any 2D rotation transformation* $R : (x_s, y_s) \rightarrow (x_n, y_n)$*, there exist* $\alpha \in [0, 2\pi]$ *such that,*

$$\begin{aligned} x_n &= \cos\alpha x_s - \sin\alpha y_s \\ y_n &= \sin\alpha x_s + \cos\alpha y_s \\ \boldsymbol{v}_n^r &= \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{v}_s^r \end{aligned} \tag{26}$$

Proof. See Appendix 7.2 . □

Theorem. 2 implies that it is possible to obtain $\boldsymbol{v}_n^r$ of arbitrary ellipse A_n if we can calculate the vector $\boldsymbol{v}_s^r$ of any unrotated ellipse A_s whose major and minor axis are parallel to the axis of the coordinate system.

Lemma 3. *Consider* $I^{m,n} = \frac{1}{2\pi}\int_0^{2\pi} \cos^m\theta \sin^n\theta d\theta$ *with* $m, n, i, j \in \mathbb{Z}^*$ *then,*

$$I^{m,n} = \begin{cases} \frac{\binom{2i+2j}{i+j}\binom{i+j}{i}}{\binom{2i+2j}{2i} 2^{2i+2j}} & \text{if } m = 2i, n = 2j \\ 0 & otherwise. \end{cases}$$

Proof. See Appendix 7.3. □

Lemma 4. *The* $(m+n)^{th}$ *moment of* $A_0 = \{(x_0, y_0) | (x_0/a)^2 + (y_0/b)^2 \leq 1\}$ *is*

$$M_0^{m,n} = \frac{a^m b^n}{1 + (m+n)/2} I^{m,n}$$

Proof. See Appendix 7.4. □

We can obtain $\boldsymbol{v}_s^r$ from Theorem. 5 using Lemma. 3 and Lemma. 4. $M_0^{m,n}$ denotes analytic solutions for $(m+n)^{\text{th}}$ moments of the unrotated ellipse located at the origin.

Theorem 5 (Solution of $\boldsymbol{v}_s^r$)**.** *When* x_s *and* y_s *satisfy* $((x_s - t_x)/a)^2 + ((y_s - t_y)/b)^2 \leq 1$ *then,*

$$\boldsymbol{v}_s^r[0] = \sum_{i=0}^{r}\sum_{j=0}^{r-i} M_0^{2i,2j} \sum_{k=i}^{r-j} \binom{r}{k}\binom{2k+1}{2i}\binom{2r-2k}{2j} t_x^{2k-2i+1} t_y^{2r-2k-2j}$$

$$\boldsymbol{v}_s^r[1] = \sum_{i=0}^{r}\sum_{j=0}^{r-i} M_0^{2i,2j} \sum_{k=i}^{r-j} \binom{n}{k}\binom{2k}{2i}\binom{2r-2k+1}{2j} t_x^{2k-2i} t_y^{2r-2k-2j+1}$$

$$\boldsymbol{v}_s^r[2] = \sum_{i=0}^{r}\sum_{j=0}^{r-i} M_0^{2i,2j} \sum_{k=i}^{r-j} \binom{r}{k}\binom{2k}{2i}\binom{2r-2k}{2j} t_x^{2k-2i} t_y^{2r-2k-2j}$$

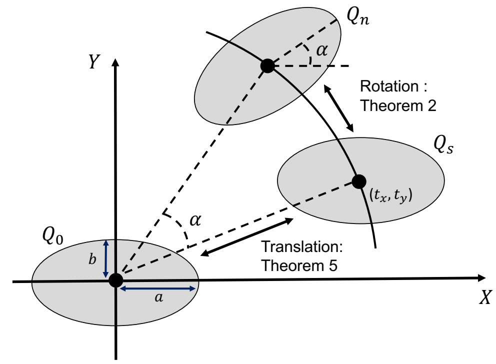


Figure 3. **Moment calculation strategy for a arbitrary ellipse $\boldsymbol{Q}_n$.** The moments of the rotated ellipse $\boldsymbol{Q}_n$ are obtained from the moments of the un-rotated standard ellipse $\boldsymbol{Q}_s$ using Theorem. 2. The moments of $\boldsymbol{Q}_s$ are obtained from moments of another standard ellipse $\boldsymbol{Q}_0$, which is located at the origin, using Theorem. 5.

Proof. See Appendix 7.5. □

We have reduced the computational cost of calculating the $\boldsymbol{v}_s^r$ using dynamic programming. The coefficients comprise products of binomials while $I^{m,n}$ is invariant to the conic shape; therefore, it is possible to calculate and store these values in advance. Hence, $\boldsymbol{v}_s^r$ is obtained in $O(1)$[1]. The $\boldsymbol{v}_n^r$ can be obtained simply by multiplying the rotation matrix to $\boldsymbol{v}_s^r$ as Theorem. 2. Using each of the previous derivations to compute the Eqs. (20) to (22), the overall process of the unbiased estimator is described as Algorithm 1.

Algorithm 1 Unbiased estimator

Input:
- $\boldsymbol{Q}_w$: matrix of a circle in target plane
- $\boldsymbol{E} = [\boldsymbol{r}_1 \boldsymbol{r}_2 \boldsymbol{t}]$: extrinsic parameter
- $\boldsymbol{K}$: intrinsic parameter
- $\boldsymbol{D} = [1, d_1, d_2, d_3, ..., d_n]$: distortion parameter

Output:
- (x_i, y_i) : the center point of projected circle in image

```
Q_n = E^{-T} Q_w E^{-1}
m_x, m_y, m_0 = 0
for r = 0 : 3n_d do
    t'_x, t'_y, a, b, α = GeometryOfEllipse(Q_n)   % Appendix 8.1
    t_x = cos α t'_x + sin α t'_y
    t_y = − sin α t'_x + cos α t'_y
    v_s^r = CalcualteVs(t_x, t_y, a, b, D, r)   % Theorem. 5
    v_n^r = R_z(α) v_n^r   % Theorem. 2
    w_0r, w_1r = GetCoeff(D, r)   % Appendix 8.2
    m_x = m_x + w_1r v_n^r[0]
    m_y = m_y + w_1r v_n^r[1]
    m_0 = m_0 + w_0r v_n^r[2]
end for
x_d = m_x / m_0
y_d = m_y / m_0
x_i = K[0,0] x_d + K[0,2]
y_i = K[1,1] y_d + K[1,2]
return x_i, y_i
```

[1] Indeed, $O(n_d^3)$. However, n_d is a constant number and typically lower than three so that we can treat it as $O(1)$

4.3. Calibration using unbiased estimator

The calibration process is divided into two stages. First, we establish the initial value of the intrinsic parameter in closed form using Zhang [28]'s method, assuming zero distortion. The initial value is unusable as it does not consider the distortion parameter, requiring refinement through optimization. In the optimization process, we minimize the reprojection error, which is the squared difference between the observed position of the control point and the estimated position of the control point using the unbiased estimator as:

$$\boldsymbol{K}, \boldsymbol{D} = \underset{\boldsymbol{K},\boldsymbol{D},\boldsymbol{E}^{1,2,\ldots n}}{\operatorname{argmin}} \sum_{j=1}^{n}\sum_{k=1}^{m} \left\| \bar{\boldsymbol{p}}_i^{jk} - \hat{\boldsymbol{p}}_i^{jk} \right\|^2, \tag{27}$$

$$\hat{\boldsymbol{p}}_i^{jk} = \text{UnbiasedEstimator}(\boldsymbol{K}, \boldsymbol{E}^j, \boldsymbol{Q}_w^{jk}), \tag{28}$$

where the $\bar{\boldsymbol{p}}_i^{jk}$ is the k^{th} control point in the j^{th} image, and the $\hat{\boldsymbol{p}}_i^{jk}$ is the estimated control point corresponding to $\boldsymbol{p}_i^{jk}$. The circle in target plane corresponding to $\boldsymbol{p}_i^{jk}$ is $\boldsymbol{Q_w}^{jk}$. The above optimization problem is solved with the Ceres Solver [1].

5. Results

5.1. Comparison of estimators

We compared our unbiased estimator with the checkerboard method and other existing estimators, such as point-based, conic-based, and numerical method [13], as defined below:

$$\text{Point-based} : \hat{\hat{\boldsymbol{p}}}_i = \boldsymbol{K}D(\boldsymbol{E}\boldsymbol{Q}_w^{-1}[0,0,1]^\top), \tag{29}$$

$$\text{Conic-based} : \hat{\hat{\boldsymbol{p}}}_i = \boldsymbol{K}D(\boldsymbol{E}\boldsymbol{Q}_w^{-1}\boldsymbol{E}^T[0,0,1]^\top), \tag{30}$$

$$\text{Numerical} : \hat{\hat{\boldsymbol{p}}}_i = \text{Eq. (19) in [13]} \tag{31}$$

$$\text{Ours} : \text{Algorithm 1}$$

Numerical(n) means the iteration step is n.

Most of the existing calibration algorithms use the point-based estimator, which disregards the conic geometry and directly matches the center of the shape in the distorted image to the center of the circle in the target. The conic-based estimator compensates for the bias introduced by the perspective transformation. However, the conic-based estimator still obtains the distortion bias. The numerical method, which is theoretically unbiased, also shows considerable error induced by numerical integration. These limitations are well described in Fig. 4. Each data point indicates the average of reprojection errors in fixed 24 scenes, and we illustrated the error graph with standard deviations.

Our moment-based unbiased estimator maintains very small reprojection errors regardless of radius size and distortion. However, the errors of other methods are significant, while the errors of circular pattern methods such as

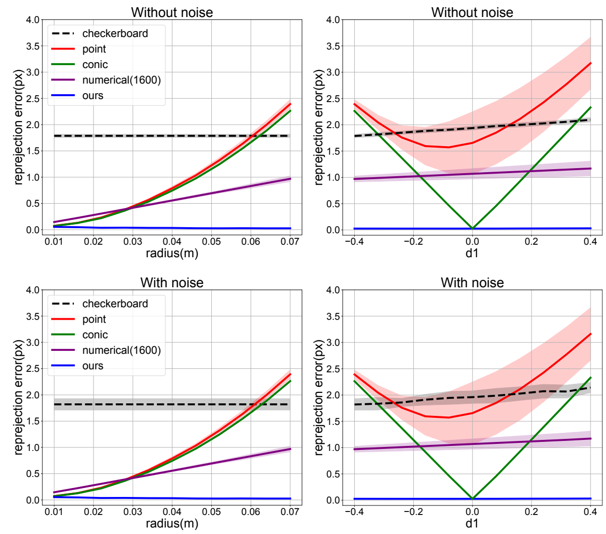


Figure 4. **Reprojection error comparison in synthetic images.** We examined the consistency and magnitude of the reprojection errors of each method as we varied the radius of the circles and the amount of distortion, represented as d_1. In the graphs, each curve represents the mean reprojection error and the envelope represents the corresponding standard deviation boundary. The first row uses raw images and the second row applies the Gaussian blur to the images to check robustness. Unlike other circular pattern methods, our method is unbiased, resulting in near-zero errors regardless of distortion and radius changes.

conic and point gradually increase as the radius or distortion increases. Although the amount of distortion does not greatly influence the result of the checkerboard method, it has the inherent error caused by the inaccurate control point detector compared to the circular pattern. The control point of the checkerboard is the corner of the squares, and due to the discontinuity of the pixel, the corner position should be approximated by interpolation near the pattern boundary. In contrast, the control point of the circular pattern is obtained from the average of thousands of inner points of the circle, resulting in the precise position in decimal places.

Because the conic and point-based estimators have biases resulting from the conic geometry, the error increases as the circle size and distortion increase. The two estimators have slightly different tendencies. The conic-based estimator has a bias due to the distortion, not the perspective transformation. Hence, the reprojection error depends only on the absolute value of the distortion coefficient. In comparison, the point-based estimator has both perspective and distortion biases. The perspective bias causes a high standard deviation and asymmetry in the error plot.

5.2. Calibration accuracy of synthetic images

We conducted experiments on synthetic images to evaluate how our unbiased estimator improves calibration per-

Table 1. **Calibration results of synthetic images.** This table shows the mean and standard deviation of the calibration results. We generated 100 synthetic images and obtained intrinsic parameters using each method from 30 randomly selected images. By repeating this action 30 times, we were able to calculate the mean and standard deviation of the calibration results corresponding to each method. Our method shows the closest result to the ground truth value, and the variance is dramatically lower than other methods. Meanwhile, other circular pattern methods, such as point-based and conic-based methods, converge to inappropriate values despite using the same control points as ours. More iterations allow the numerical method to behave unbiased while the computational time significantly grows. The checkerboard method, which is also unbiased, converges to ground truth values; however, its variance is significant due to measurement noise. Significantly, the measurement accuracy deteriorates further, and almost half of the images are not detected when the image quality is low due to Gaussian blur.

	Low distortion ($d_1 = -0.2$)							High distortion ($d_1 = -0.4$)						
Methods	f_x	f_y	c_x	c_y	d_1	#fail	runtime(s)	f_x	f_y	c_x	c_y	d_1	#fail	runtime(s)
GT	600	600	600	450	-0.2			600	600	600	450	-0.4		
checkerboard	599.8±0.37	599.8±0.38	600.3±0.21	449.9±0.41	-0.20±0.001	0.0	0.3	600.1±0.22	600.2±0.21	599.9±0.15	450.1±0.12	-0.40± 0.001	0.0	0.4
point-based	598.9±0.28	598.9±0.28	600.0±0.29	450.0± 0.43	-0.20±0.001	0.0	0.3	603.1±0.82	603.1±0.79	599.9±0.53	450.3±0.4	-0.41±0.003	0.0	0.4
conic-based	603.5±0.41	603.5±0.38	600.1±0.24	449.8±0.25	-0.21±0.002	0.0	0.7	606.9±0.80	606.9±0.79	599.7±0.74	450.4±0.51	-0.41±0.005	0.0	0.9
numerical(25)	600.6±0.28	600.4±0.30	600.0±0.3	449.7±0.22	-0.20±0.001	0.0	2.0	598.4±0.71	597.9±0.75	600.9±0.72	449.7±0.36	-0.40±0.002	0.0	2.7
numerical(1600)	600.1±0.07	600.0±0.07	**600.0±0.05**	449.9±0.06	**-0.20±0.000**	0.0	75.5	599.9±0.1	599.7±0.14	600.2±0.18	450.0±0.06	**-0.40±0.001**	0.0	86.6
ours	**600.0± 0.06**	**600.0±0.06**	**600.0±0.05**	**450.0±0.05**	**-0.20±0.000**	0.0	2.0	**599.9±0.09**	**599.9±0.10**	**600.0±0.03**	**450.0±0.03**	**-0.40±0.001**	0.0	2.0
	Gaussian blur ($\sigma = 2$)													
GT	600	600	600	450	-0.2			600	600	600	450	-0.4		
checkerboard	598.7±1.99	599.1±1.84	599.3±2.00	450.5±1.52	-0.20±0.013	14.2	0.2	596.7±2.83	597.1±2.70	599.6±1.98	451.6±2.02	-0.36±0.026	15.8	0.2
point-based	598.7±0.29	598.7±0.28	600.0±0.32	449.9±0.25	-0.20±0.003	0.0	0.4	603.3±0.65	603.3±0.63	600.4±0.64	450.4±0.31	-0.41±0.002	0.0	0.4
conic-based	603.6±0.47	603.6±0.46	600.0±0.28	449.8±0.18	-0.20±0.002	0.0	0.7	607.6±1.57	607.7±1.57	599.9±0.60	450.6±0.43	-0.41±0.004	0.0	0.7
numerical(25)	600.5±0.26	600.3±0.26	600.1±0.27	449.6±0.18	**-0.20±0.000**	0.0	2.0	598.4±0.90	598.0±0.92	600.8±0.65	449.6±0.35	**-0.40±0.001**	0.0	2.8
numerical(1600)	600.1±0.09	600.0±0.08	**600.0±0.06**	449.9±0.07	**-0.20±0.000**	0.0	75.5	599.9±0.15	599.7±0.15	600.2±0.18	450.1±0.06	**-0.40±0.001**	0.0	84.7
ours	**600.0±0.07**	**600.0±0.07**	**600.0±0.06**	**450.0±0.05**	**-0.20±0.000**	0.0	2.0	**599.9±0.07**	**599.9±0.08**	**600.0±0.04**	**450.0±0.03**	**-0.40±0.001**	0.0	2.0

formance. We set the Ground Truth (GT) values for f_x, f_y, c_x, and c_y to make the field of view (FOV) approximately 90 degrees. We prepared two scenarios for the distortion coefficients from (11): one with low distortion (d_1 = -0.2) and one with high distortion (d_1 = -0.4). In the latter case, the distortion function is not invertible within the given FOV range, so we slightly adjusted d_2 to make it invertible, and we only evaluated d_1 in our analysis since d_2 is as small as negligible. We prepared 100 images and performed calibration on a set of 30 randomly selected images. We repeated this process 30 times to calculate the averages and standard deviations shown in Tab. 1.

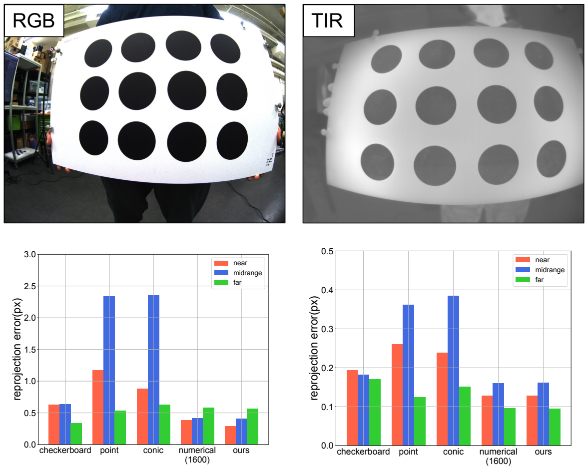
As a result, our method consistently produces the best results in all cases. Our approach benefits from an accurate mean and is particularly valuable for its low variance. Since the estimator of the checkerboard pattern is unbiased, it converges closely to the ground truth in the absence of noise. However, the calibration results show significant variance from the imprecise control point measurements, as shown in the experiment in Sec. 5.1. This tendency increases when noise is introduced into the images, and control point detection fails in many cases, resulting in worse performance.

In the case of point-based and conic-based methods using a circular pattern, the measurements are accurate and robust. However, due to the biased estimator, they perform worse than the checkerboard pattern. As the distortion increases, the bias becomes more pronounced, leading to poor calibration results. The numerical method has a trade-off between accuracy and runtime, and it requires significant time to achieve meaningful results. Our method, which addresses these limitations, exhibits higher accuracy, efficiency, and robustness than the checkerboard and the others. Note that one of the critical contributions of this paper is to prove the value of circular patterns, which have more informative features but have been underutilized due to algorithmic limitations.

5.3. Calibration accuracy of real images

We further evaluated our method using real images captured from both RGB and TIR cameras. Another evaluation metric is required to extend the experiment to the real world since there is no way to find the GT intrinsic parameter of real cameras. Instead of comparing the intrinsic parameter directly, we evaluated the distribution of reprojection errors and the relative translation and rotation value between each target in different scenes. To highlight clear differences among the methods, we utilized two types of cameras with high distortion. One is an RGB camera, Trition 5.4 MP, whose resolution is 1200×930. The other is a TIR camera, FLIR A65, whose resolution is 640×512. The sample images of these two cameras are shown in Fig. 5. Since the TIR is limited in recognizing colored patterns and only distinguishes objects by infrared energy we utilized a printed circuit board (PCB) composed of squares or circles with different heat conductivity introduced in [24].

Distribution of reprojection errors. We expected calibration results to be better when the total reprojection error is small and the error is uniformly distributed. We collected 24 images from three different distances (i.e., near, midrange, and far) to investigate these two aspects. As summarized in Fig. 5, point-based and conic-based methods result in high reprojection errors and significant differences in error values between different distances, which implies that the calculated intrinsic parameters do not explain the projection model consistently across all distances. On the contrary, the checkerboard method and our method, whose unbiased



(a) RGB (b) TIR

Figure 5. **Real world experiment: Reprojection error.** The first row shows sample images from (a) RGB and (b) TIR cameras. The second row shows the distribution of the reprojection error divided by the distance from the camera to the target. The error distribution of the biased methods, such as point and conic-based methods, varies greatly with distance. The unbiased methods, such as checkerboard and ours, show a low and uniform error distribution. In TIR images, our method significantly outperforms the checkerboard method due to the robust control point of circular patterns.

estimators show low reprojection error at all distances. The checkerboard and our methods show similar performance in the RGB camera. Meanwhile, our methods outperform the checkerboard method in the TIR camera due to accurate control point detection. As shown in §5.2, the accuracy of the numerical method converges to our method while the number of iteration steps is increasing. However, Numerical(1600) takes almost five minutes to converge, while our method takes 30 seconds. More information, such as the distribution of the error vectors on the 2D image, is described in Appendix 9.2.

6D pose errors. To validate the un-projection performance of each method, we investigated the relative 6-DoF pose error. Each camera captured 20 in different scenes, and we conducted the calibration with these images. During the calibration, the pose of the target in each scene is also optimized, and these poses are used to evaluate the calibration accuracy. The GT value of the relative pose is obtained by the motion capture system by OptiTrack. Let $\boldsymbol{T}^c_{t_i}$ be a $SE(3)$ matrix from the i^{th} target coordinates to the camera coordinate, and $\boldsymbol{T}^m_{o_i}$ be a $SE(3)$ matrix from the i^{th} object coordinates to the motion capture system coordinate. Note that the object frame is the frame attached to the target plane. However, the two frames are not identical since the motion capture system defines another coordinate system for the target as Fig. 6. It is impossible to evaluate the relative pose directly without the coordinate transformation matrices $\boldsymbol{T}^o_t$ and $\boldsymbol{T}^m_c$. Therefore, we first calculated the transformation matrices, and the final error is defined as

$$error = \sum_{i=1}^{n} ||\boldsymbol{T}^m_{o_i} \ominus \hat{\boldsymbol{T}^m_c} \boldsymbol{T}^c_{t_i} (\hat{\boldsymbol{T}^o_t})^{-1}|| \tag{32}$$

with the estimated value $\hat{\boldsymbol{T}^m_c}$ and $\hat{\boldsymbol{T}^o_t}$. The detail of estimating $\boldsymbol{T}^o_t$ and $\boldsymbol{T}^m_c$ is described in Appendix 9.3 According to Tab. 2, our method outperforms the entire case, and the checkerboard method shows lower performance at the unprojection task despite the low reprojection error compared to the point and conic-based methods.

Table 2. **Real world experiment: 6D pose error.** The rotation and translation error between the target pose obtained from the calibration results and the motion capture system. Our method outperforms other methods in the rotation part. The checkerboard method shows comparable results in the translation part on the RGB camera; however, it malfunctions on the TIR camera.

	RGB		TIR	
Methods	Rotation	Translation	Rotation	Translation
checkerboard	0.32°	**2.5** mm	1.63°	16.5 mm
point-based	0.36°	4.2 mm	0.50°	3.8 mm
conic-based	0.31°	2.7 mm	0.47°	3.4 mm
numerical(1600)	0.26°	3.8 mm	**0.45°**	3.3 mm
ours	**0.24°**	**2.5** mm	**0.45°**	**3.1** mm

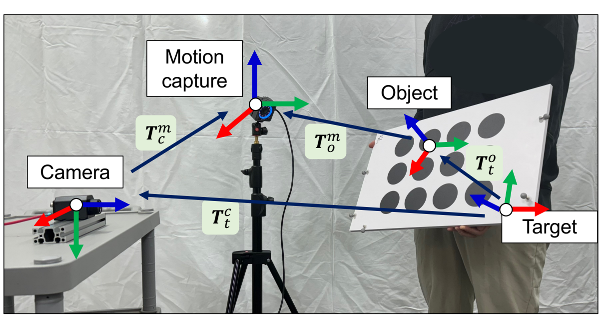


Figure 6. **Experiment setup.** We utilize the motion capture system to develop the ground truth 6D pose of the target. The $\boldsymbol{T}^m_c \boldsymbol{T}^c_t = \boldsymbol{T}^m_o \boldsymbol{T}^o_t$ relationship is satisfied.

6. Conclusion

In this paper, we introduce the closed-form n^{th} moment of conics and prove that the centroid of a distorted ellipse is expressed by a linear combination of the original moments. Based on these results, we developed an unbiased estimator of circular patterns under distortion for camera calibration. Using this estimator, circular pattern-based calibration overcomes its algorithmic limitation and outperforms checkerboard-based calibration. .

Acknowledgments. This work was supported by the NRF grant funded by MSIP (2020R1C1C1006620) and IITP grant funded by MSIT (No.2022-0-00480).

References

[1] Sameer Agarwal, Keir Mierle, and The Ceres Solver Team. Ceres Solver, 2022. 6

[2] J.Y. Bouguet. Matlab camera calibration toolbox. 2000. 3

[3] Duane Brown. Decentering distortion of lenses. *Photogrammetric engineering*, 32(3):444–462, 1996. 3

[4] Qian Chen, Haiyuan Wu, and Toshikazu Wada. Camera calibration with two arbitrary coplanar circles. In *Proceedings of the European Conference on Computer Vision*, pages 521–532. Springer, 2004. 1, 2

[5] Jen-Hui Chuang, Chih-Hui Ho, Ardian Umam, Hsin-Yi Chen, Jenq-Neng Hwang, and Tai-An Chen. Geometry-based camera calibration using closed-form solution of principal line. *IEEE Transactions on Image Processing*, 30: 2599–2610, 2021. 1, 2

[6] Alexander Eugen Conrady. Decentred lens-systems. *Monthly notices of the royal astronomical society*, 79(5): 384–390, 1919. 3

[7] Ankur Datta, Jun-Sik Kim, and Takeo Kanade. Accurate camera calibration using iterative refinement of control points. In *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pages 1201–1208. IEEE, 2009. 2

[8] Frederic Devernay and Olivier Faugeras. Straight lines have to be straight. *Machine vision and applications*, 13:14–24, 2001. 2

[9] Clément Godard, Oisin Mac Aodha, and Gabriel J Brostow. Unsupervised monocular depth estimation with left-right consistency. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 270–279, 2017. 1

[10] Janne Heikkila. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1066–1077, 2000. 1, 2

[11] Janne Heikkila and Olli Silven. Calibration procedure for short focal length off-the-shelf ccd cameras. In *Proceedings of the International Conference Pattern Recognition*, pages 166–170. IEEE, 1996. 1, 2

[12] Myung-Hwan Jeon, Jeongyun Kim, Jee-Hwan Ryu, and Ayoung Kim. Ambiguity-aware multi-object pose optimization for visually-assisted robot manipulation. *IEEE Robotics and Automation Letters*, 8(1):137–144, 2022. 1

[13] Juho Kannala and Sami S Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1335–1340, 2006. 1, 2, 6

[14] Jun-Sik Kim, Pierre Gurdjos, and In-So Kweon. Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):637–642, 2005. 1, 2

[15] Manuel Lopez, Roger Mari, Pau Gargallo, Yubin Kuang, Javier Gonzalez-Jimenez, and Gloria Haro. Deep single image camera calibration with radial distortion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11817–11825, 2019. 3

[16] John Mallon and Paul F Whelan. Which pattern? biasing aspects of planar calibration patterns and detection methods. *Pattern recognition letters*, 28(8):921–930, 2007. 1, 2

[17] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *Proceedings of the European Conference on Computer Vision*, pages 405–421, 2020. 1

[18] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41 (4):1–15, 2022. 1

[19] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015. 1

[20] Linfei Pan, Marc Pollefeys, and Viktor Larsson. Camera pose estimation using implicit distortion models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12819–12828, 2022. 2, 3

[21] Frank C Park and Bryan J Martin. Robot sensor calibration: solving ax= xb on the euclidean group. *IEEE Transactions on Robotics*, 10(5):717–721, 1994. 8

[22] Tong Qin, Peiliang Li, and Shaojie Shen. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, 2018. 1

[23] Thomas Schops, Viktor Larsson, Marc Pollefeys, and Torsten Sattler. Why having 10,000 parameters in your camera model is better than twelve. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2535–2544, 2020. 2

[24] Young-Sik Shin and Ayoung Kim. Sparse depth enhanced direct thermal-infrared slam beyond the visible spectrum. *IEEE Robotics and Automation Letters*, 4(3):2918–2925, 2019. 7, 6

[25] Peter F Sturm and Stephen J Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 432–437. IEEE, 1999. 1, 2

[26] Junhua Sun, Xu Chen, Zheng Gong, Zhen Liu, and Yuntao Zhao. Accurate camera calibration with distortion models using sphere images. *Optics & Laser Technology*, 65:83–87, 2015. 1, 2

[27] Kwan-Yee Kenneth Wong, Guoqiang Zhang, and Zhihu Chen. A stratified approach for camera calibration using spheres. *IEEE Transactions on Image Processing*, 20(2): 305–316, 2010. 1, 2

[28] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000. 1, 2, 6