Camera Calibration Based on Mirror Reflections

Kosuke Takahashi[1,a)] Shohei Nobuhara[1,b)]

Abstract: This paper addresses the use of mirror reflections for camera calibration. Camera calibration is an essential technique for analyzing the geometric and radiometric relationship between a 3D space and a 2D image. Most conventional camera calibration methods are based on a fundamental assumption: *a camera can directly observe a reference object of known geometry*. However, there are cases in which this assumption does not hold in practical scenarios.
One approach to camera calibration in such cases is the use of a "mirror" as a supporting device. A mirror generates a virtual reference object that can be expressed using a small number of parameters. In addition, the 2D projection of the reflection object is equal to that of the known reference object from the virtual viewpoint. This paper utilizes these features and tackles two challenges of the geometric camera calibration; the first challenge is the intrinsic camera calibration when a known reference object is not available and the second challenge is the extrinsic camera calibration when the camera cannot directly observe a known reference object due to a physical constraint on the imaging system. The proposed algorithms introduce novel constraints, *kaleidoscopic projection constraint* and *orthogonality constraint*, which are hold with the mirror reflections for solving these problems. Evaluations with synthesized and real data demonstrates that the proposed algorithms can work properly and report the robustness of it in comparison with conventional methods.

1. Introduction

A ray omitted from a light source or reflected from the surface of an object reaches an image sensor through a lens and is collected as an "image". Understanding this generating process of a image, that is, describing *to where* a point in 3D space is projected with *how much* intensity, is a fundamental and important problem for various tasks in computer vision, such as 3D reconstruction and motion analysis. In order to analyze the generating process, to use an appropriate camera model and to estimate the camera parameters are called "camera calibration."

Camera calibration has been a fundamental research topic in computer vision for many years. While the strict meaning of camera calibration depends on the configuration of the imaging system, camera calibration can be divided into two types: geometric calibration and radiometric calibration.

Geometric camera calibration: With this type, the "to where" a 3D point is projected onto a 2D image plane is estimated. That is, there is a geometric relationship between a 3D space and a 2D image plane. This type of calibration is thus appropriate for analyzing geometric transformation involving a 3D space and a 2D image plane.

Radiometric camera calibration: With this type, the "how much" intensity is estimated. It is appropriate for analyzing radiometric transformation involving a 3D space and a 2D image plane. This calibration includes removing shading caused by the lens [27] or noise on the image sensor [11], estimation of the radiometric response function [18], and so on.

This paper focuses on the analysis of geometric properties involving a 3D space and a 2D image plane, so it mainly discusses geometric camera calibration of a perspective camera model.

[1] Kyoto Uniersity, Yoshida-honmachi, Kyoto, Japan
[a)] takahasi@vision.kyoto-u.ac.jp
[b)] nob@i.kyoto-u.ac.jp

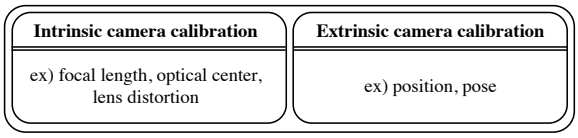


Fig. 1 The overview of the geometric camera calibration.

As illustrated in Figure 1, geometric camera calibration comprises intrinsic camera calibration and extrinsic camera calibration. The former estimates the *intrinsic parameters*, which describe the camera properties: focal length, optical center, lens distortion, and so on. The latter estimates the *extrinsic parameters*, which describe the position and pose of the camera.

Geometric camera calibration has three basic steps: (1) capture a reference object for which the geometric features are known, (2) determine the correspondences between the features and their projections, and (3) estimate each parameter from the correspondences using the camera model.

For example, Tsai [29] estimated the intrinsic and extrinsic parameters from the correspondences between the 3D points for which the geometries were known in a world coordinate system and their 2D projections. Zhang [31] used a known reference point on a single plane as a reference object, *e.g.* , a chessboard, to estimate the camera parameters.

These conventional methods share a fundamental assumption: *the camera can directly observe a reference object for which the geometry is known*. However, this assumption does not always hold in practical cases, so these methods do not always work properly. For example, while intrinsic camera calibration should be done using known reference objects in the camera's field-of-view for high precision estimation, a practical known reference object is not always available when capturing images on a certain scale. Figure 2(a) illustrates the case in which a known reference object is not practical in terms of scale. In addition, Figure 2(b) illustrates the case in which a reference object is not observable

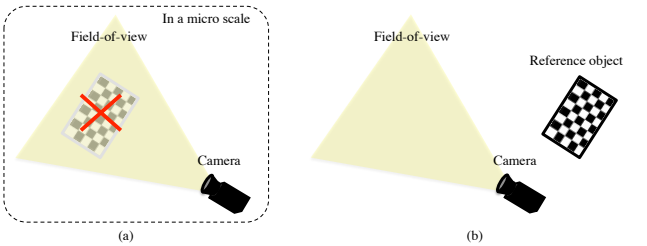


Fig. 2 Example problems with previous calibration methods: (a) using a known reference object is not practical in terms of scale; (b) a reference object is not observable directly from the camera.

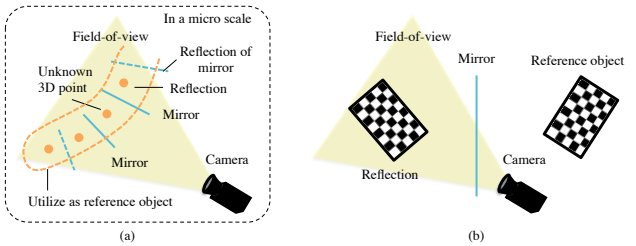


Fig. 3 Mirror-based calibration uses (a) multiple reflections of a 3D point as a reference object of a parametric 3D model or (b) mirrors to enable the camera to observe the reflections of a known reference object.

from the camera. For a vision-based robot [4] with a display-camera system [6, 13, 16], although the extrinsic parameters between the camera and each part of the system (*e.g.* a robot arm and a body) are important, they are often unobservable directly from the camera due to a physical constraint on the imaging system. Thus, the previous methods can be problematic depending on the configuration of the imaging system and the situation.

The methods introduced in this paper for solving these two problems uses algorithms that enable the use of a *mirror* as a supporting device for calibration. Mirrors can generate reflections of a real object, and the reflections can be defined using a small number of parameters. This means that they can be recognized as a parametric 3D model. Furthermore, 2D projections of reflections are equal to those of the known reference object from the virtual viewpoint. In other words, the mirror extends the camera's field-of-view. As illustrated in Figure 3, the problems are overcome by

- setting planar mirrors and generating multiple reflections consisting of a parametric 3D model with an unknown 3D point and its reflections (Figure 3(a)) and
- setting a planar mirror that enables the camera to observe a known reference object (Figure 3(b)).

The potential of using mirrors to solve these problems of camera calibration in practical situations is explored in this paper.

1.1 Problem Statement and Contributions

This paper focuses on geometric camera calibration in cases where the fundamental assumption, *i.e. the camera can directly observe a reference object for which the geometry is known*, does not hold and tackles the two problems; intrinsic camera calibration without a known reference object and extrinsic camera calibration with an unobservable reference object.

Intrinsic Camera Calibration Without Known Reference Object: In cases where a known reference object is not available, we introduce mirrors to generate reflections of a 3D point of unknown geometry and recognize its reflections as a reference object of a parametric 3D model, as shown in Figure 3(a).

In this paper, a novel intrinsic camera calibration algorithm is introduced that utilizes the parametric 3D model by multiple planar mirrors consisting of a kaleidoscopic imaging system. Three problems must be solved in order to realize this method: chamber assignment of the kaleidoscopic projections, estimation of the mirror parameters, and estimation of the intrinsic parameters. The key contribution of this work is the introduction of a novel geometric constraint, the *kaleidoscopic projection constraint*, which is satisfied by projections of high-order reflections. This constraint provides multiple linear equations for the mirror parameters for a single 3D point and solves the three problems.

Extrinsic Camera Calibration with Unobservable Reference Object In cases where the camera cannot directly observe a reference object due to a physical constraint on the imaging system, we introduce mirrors to enable the camera to directly observe the reflections of a known reference object (Figure 3(b)).

In this paper, a novel two types of mirror-based algorithms are introduced that estimates the extrinsic parameters between the camera and a reference object located outside its field-of-view.

The first algorithm utilizes a planar mirror. Since the poses and positions of the reference object and the planar mirror are unknown, the extrinsic parameters cannot be determined uniquely from a single image of the reflection. To overcome this problem, an *orthogonality constraint* that is satisfied among reflections by multiple mirror poses is introduced. This constraint is used to estimate the extrinsic parameters, that is, three reference points and three mirror poses, with the minimal configuration.

The mirror-based methods can be troublesome for preparing a mirror and calibrating it every time in a casual scenario, such as gaze correction in a video conference [16]. For such cases, the second algorithm utilizes the human cornea based on the fact that the surface of the human eye reflects light like a mirror, meaning that the human eye can be modeled as a spherical mirror. The introduction of a geometric model of the human cornea enables estimation of extrinsic parameters with a simple configuration, i.e., one mirror pose and three or five reference points.

The rest of this paper is organized as follows. Section 2 provides fundamental knowledge and related work on camera calibration and mirror geometry. Section 3 presents a novel algorithm for intrinsic camera calibration using multiple planar mirrors. Section 4 and Section 5 present novel methods for planar mirror-based and human-cornea-based extrinsic camera calibration. Section **??** cocludes this paper and outlines future works.

2. Camera Calibration and Mirror Geometry

2.1 Geometric Camera Calibration of Perspective Camera

The perspective camera model is illustrated in Figure 4. Let $\boldsymbol{p}^{\{W\}} = (X^{\{W\}}, Y^{\{W\}}, Z^{\{W\}})^\top$ and $\boldsymbol{p}^{\{C\}} = (X^{\{C\}}, Y^{\{C\}}, Z^{\{C\}})^\top$ denote a 3D point $\boldsymbol{p}$ in a world coordinate system and a camera coordinate system, respectively. They satisfy

$$\boldsymbol{p}^{\{C\}} = R\boldsymbol{p}^{\{W\}} + \boldsymbol{t}, \tag{1}$$

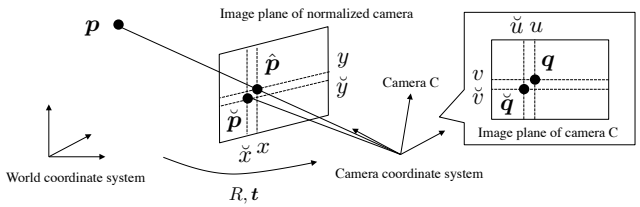


Fig. 4 The perspective camera model. The effect of lens distortion is modeled as the non-linear transformation in the normalized image coordinates.

where R is a rotation matrix and $\boldsymbol{t}$ is a translation vector. Parameters R and $\boldsymbol{t}$ are *extrinsic camera parameters* representing the pose and position of the camera in the world coordinate system. The goal of extrinsic camera calibration is to estimate these parameters. Note that we hereinafter omit the superscript representing the coordinate system for cases in which $\boldsymbol{p}$ is represented in the camera coordinate system.

In addition, let $\boldsymbol{q} = (u, v)$ denote the projection of $\boldsymbol{p}^{\{C\}}$ in a pixel image coordinate system. This $\boldsymbol{q}$ is given by a perspective projection:

$$\lambda\tilde{\boldsymbol{q}} = A\boldsymbol{p}^{\{C\}} = \begin{bmatrix} f & 0 & c_u \\ 0 & f & c_v \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{p}^{\{C\}}, \tag{2}$$

where $\tilde{\boldsymbol{q}}$ denotes the homogeneous coordinate of $\boldsymbol{q}$ and λ is a scale parameter. The f is focal length and (c_u, c_y) are the optical centers expressed in pixels coordinates.

This projection model is extended by taking into account lens distortion. Suppose $\hat{\boldsymbol{p}} = (x, y, 1)$ denotes the normalized image coordinates of $\boldsymbol{p}^{\{C\}}$, and $\breve{\boldsymbol{p}}$ denotes the distorted coordinates of $\hat{\boldsymbol{p}}$. This $\breve{\boldsymbol{p}} = (\breve{x}, \breve{y}, 1)$ is expressed as

$$\begin{aligned} \breve{x} =& x(1 + k_1r^2 + k_2r^4 + k_3r^6) + 2p_1xy + p_2(r^2 + 2x^2) \\ \breve{y} =& y(1 + k_1r^2 + k_2r^4 + k_3r^6) + p_1(r^2 + 2y^2) + 2p_2xy, \end{aligned} \tag{3}$$

where $r^2 = x^2 + y^2$. The $k_i (i = 1, 2, 3)$ and $p_i (i = 1, 2)$ denote the coefficients of the radial and tangential factors of lens distortion, respectively.

A 2D observation $\breve{\boldsymbol{q}}$ including lens distortion is given by the perspective projection of $\breve{\boldsymbol{p}}$; that is, $\lambda\tilde{\breve{\boldsymbol{q}}} = A\breve{\boldsymbol{p}}$. The ideal 2D projection $\boldsymbol{q}$ without lens distortion can be computed from $\breve{\boldsymbol{q}}$ by solving Eq. (3) numerically for $\hat{\boldsymbol{p}}$ with *intrinsic camera parameters* A and $\boldsymbol{d} = (k_1, k_2, k_3, p_1, p_2)$. The goal of intrinsic camera calibration is to determine these parameters.

2.2 Mirror Geometry

The planar mirror geometry is illustrated in Figure 5. Consider a 3D point $\boldsymbol{p}$ and its reflection $\boldsymbol{p}'$ from mirror π. Their ideal projections without lens distortion, $\boldsymbol{q}$ and $\boldsymbol{q}'$, are given by the perspective projection:

$$\lambda\boldsymbol{q} = A\boldsymbol{p}, \quad \lambda'\boldsymbol{q}' = A\boldsymbol{p}', \tag{4}$$

where λ and λ' are scale parameters.

Let $\boldsymbol{n}$ and $d(> 0)$ denote the normal and the distance of the mirror π satisfying

$$\boldsymbol{n}^\top\boldsymbol{x} + d = 0, \tag{5}$$

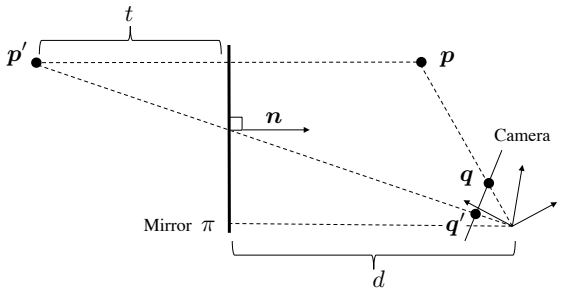


Fig. 5 Planar mirror geometry: mirror π with normal $\boldsymbol{n}$ and distance d reflects 3D point $\boldsymbol{p}$ to $\boldsymbol{p}'$; they are projected to $\boldsymbol{q}$ and $\boldsymbol{q}'$ respectively.

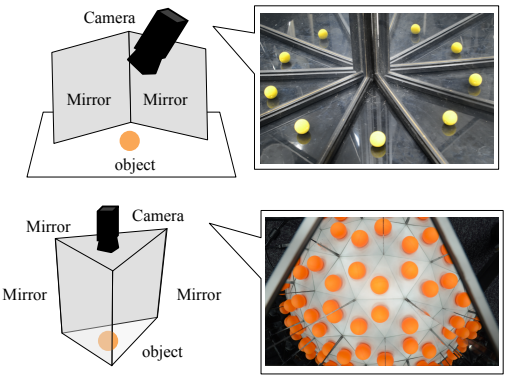


Fig. 6 Kaleidoscopic imaging system consists of multiple planer mirrors. Up: two mirrors. Bottom: three mirrors. The projections of reflected target object in each chamber are recognized as the target observed by the virtual cameras, which are generated by the planer mirrors.

where $\boldsymbol{x}$ is a 3D position in the scene. Here the normal vector is oriented toward the camera center.

As illustrated in Figure 5, the distance t from $\boldsymbol{p}$ and $\boldsymbol{p}'$ to mirror π satisfies $\boldsymbol{p} = \boldsymbol{p}' + 2t\boldsymbol{n}$. The projection of $\boldsymbol{p}'$ to $\boldsymbol{n}$ gives $t + d = -\boldsymbol{n}^\top\boldsymbol{p}'$. By eliminating t from these two equations, we have

$$\boldsymbol{p} = -2(\boldsymbol{n}^\top\boldsymbol{p}' + d)\boldsymbol{n} + \boldsymbol{p}', \tag{6}$$

$$\Leftrightarrow\tilde{\boldsymbol{p}} = S\,\tilde{\boldsymbol{p}}' = \begin{bmatrix} H & -2d\boldsymbol{n} \\ \boldsymbol{0}_{1\times3} & 1 \end{bmatrix} \tilde{\boldsymbol{p}}', \tag{7}$$

where H is a 3×3 Householder matrix given by $H = I_{3\times3} - 2\boldsymbol{n}\boldsymbol{n}^\top$, $\boldsymbol{0}_{m\times n}$ denotes the $m\times n$ zero matrix, and $I_{n\times n}$ denotes the $n\times n$ identity matrix. Note that this S also satisfies inverse transformation; that is $\tilde{\boldsymbol{p}}' = S\,\tilde{\boldsymbol{p}}_0$.

3. Mirror-based Intrinsic Camera Calibration

This section provides a novel intrinsic camera calibration algorithm by introducing multiple planar mirrors and consisting a kaleidoscopic imaging system (Figure 6).

3.1 Kaleidoscopic Imaging System

Suppose the camera observes the target 3D point directly and indirectly via N_π mirrors as shown in Figure 6. Let $\boldsymbol{p}_0$ denote the original 3D point and $\boldsymbol{p}_i$ denote the first reflection of $\boldsymbol{p}_0$ by the mirror π_i $(i = 1, \cdots, N_\pi)$ (Figure 8). The reflection $\boldsymbol{p}_i$ is given by

$$\tilde{\boldsymbol{p}}_0 = S_i\tilde{\boldsymbol{p}}_i = \begin{bmatrix} H_i & -2d_i\boldsymbol{n}_i \\ \boldsymbol{0}_{1\times3} & 1 \end{bmatrix} \tilde{\boldsymbol{p}}_i, \tag{8}$$

where $\boldsymbol{n}_i$ and d_i denote the mirror normal and its distance respectively and H_i is given by $H_i = I_{3\times3} - 2\boldsymbol{n}_i\boldsymbol{n}_i^\top$.

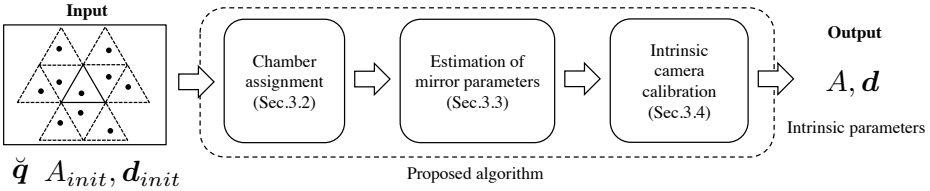


Fig. 7 The outline of the proposed optimization strategy for intrinsic camera calibration.

Furthermore, such mirrors define virtual mirrors as a result of multiple reflections. Let π_{ij} $(i, j = 1, \cdots, N_\pi,\ i \neq j)$ denote the virtual mirror defined as a mirror of π_j by π_i, $\boldsymbol{p}_{ij}$ denote the reflection of $\boldsymbol{p}_i$ by π_{ij}, and L_{ij} denote the chamber where $\boldsymbol{p}_{ij}$ is projected to. Also the matrices S_{ij} and H_{ij} for π_{ij} are given by

$$\begin{aligned} S_{ij} &= S_i S_j, \\ H_{ij} &= H_i H_j. \end{aligned} \tag{9}$$

The third and further reflections, virtual mirrors, and chambers are defined in the same manner:

$$\Pi_{k=1}^{N_k} S_{i_k}\ (i_k = 1, 2, 3,\ i_k \neq i_{k+1}), \tag{10}$$

where N_k is the number of reflections.

Obviously the 3D subspaces where $\boldsymbol{p}_0$ and $\boldsymbol{p}_i$ can exist are mutually exclusive, and the captured image can be subdivided into regions called *chambers* corresponding to such subspaces. Suppose the perspective projections of $\boldsymbol{p}_x$ is denoted by $\boldsymbol{q}_x \in Q$ $(x = 0, 1, \ldots, N_\pi, 12, 13, \ldots)$ in general. In this paper we denote the 2D region where $\boldsymbol{q}_0$ exists as the *base chamber* L_0, and we use L_x to denote the chamber where $\boldsymbol{q}_x$ exists.

3.2 Problem Description

Conventional intrinsic camera calibration techniques have been conducted from a 3D model of "known" geometry and its 2D projections [31]. That is such approaches require a reference object whose surface has several feature points such that their 3D positions are provided a priori and they are uniquely identifiable in 2D images. Hence in cases of microscopic or large-scale environment, it is not trivial task to provide such 3D models in practice.

The key idea to solve this problem is to introduce a parametric 3D model whose 3D feature positions are defined by a small number of parameters. To realize the idea, this research utilizes multiple reflections of 3D points by planar mirrors. That is our calibration estimates the intrinsic parameters as well as the mirror parameters to identify the 3D model structure simultaneously.

The configuration we consider are as follows: (a) each of the detected observation is not assigned to the corresponding chamber, (b) it has two or more planar mirrors as a supporting device whose normals and distances are unknown, and (c) it has one perspective camera and its intrinsic parameters are unknown.

Figure 7 illustrates an outline of the proposed algorithm. In our algorithm, there are three problems to be solved. Consider a 2D point set $R = \{\boldsymbol{r}_i\}$ detected from the captured image as candidates of $\boldsymbol{q}_x$. The problems are:

- to assign the chamber label L_x to $\boldsymbol{r}_i \in R$ to identify to which chamber each of the projections $\boldsymbol{q}_x$ belong (Section 3.4),
- to estimate the parameters of the real mirrors $\pi_i (i = 1, \cdots, N_\pi)$, *i.e.* normals $\boldsymbol{n}_i$ and distances d_i of them, from kaleidoscopic projections $\boldsymbol{q}_x$ (Section 3.5), and
- to estimate the intrinsic parameters, *i.e.* A and $\boldsymbol{d}$, from $\breve{\boldsymbol{q}}_x$ and the mirror parameters as the model parameters (Section 3.6).

For solving these problems, we utilize a kaleidoscopic projection constraint introduced in the next section.

3.3 Kaleidoscopic Projection Constraint

Suppose the camera observes a 3D point of unknown geometry $\boldsymbol{p}$. The mirror π of matrix S defined by the normal $\boldsymbol{n}$ and the distance d reflects $\boldsymbol{p}$ to $\boldsymbol{p}' = S\boldsymbol{p}$ (Eq (7)).

Based on the epipolar geometry [10,30], $\boldsymbol{n}$, $\boldsymbol{p}$ and $\boldsymbol{p}'$ are coplanar and satisfy $(\boldsymbol{n} \times \boldsymbol{p})^\top \boldsymbol{p}' = 0$. By substituting Eq (4), we obtain

$$\boldsymbol{q}^\top A^{-\top} [\boldsymbol{n}]_\times^\top A^{-1} \boldsymbol{q}' = 0, \tag{11}$$

where $[\boldsymbol{n}]_\times$ denotes the 3×3 skew-symmetric matrix representing the cross product by $\boldsymbol{n}$ and this is the essential matrix of this mirror-based binocular geometry [30].

By representing the normalized image coordinates of $\boldsymbol{q}$ and $\boldsymbol{q}'$ by $(x, y, 1)^\top = A^{-1}\boldsymbol{q}$ and $(x', y', 1)^\top = A^{-1}\boldsymbol{q}'$ respectively, Eq (11) can be rewritten as

$$\begin{pmatrix} y - y' & x' - x & xy' - x'y \end{pmatrix} \boldsymbol{n} = 0. \tag{12}$$

We call this Eq (12) as *kaleidoscopic projection constraint* in this paper. This constraint is satisfied by not only single reflections but also high-order reflections as below.

Single reflection: Let $\boldsymbol{p}_0$ denote a 3D point and $\boldsymbol{p}_i$ denote the reflection by mirror π_i. Since $\boldsymbol{p}_i$ is expressed as $\boldsymbol{p}_i = S_i \boldsymbol{p}_0$, the normalized image coordinates of their projections, $\boldsymbol{q}_0$ and $\boldsymbol{q}_i$, obviously satisfy Eq (12) as:

$$\begin{pmatrix} y_0 - y_i & x_i - x_0 & x_0 y_i - x_i y_0 \end{pmatrix} \boldsymbol{n}_i = 0, \tag{13}$$

where $(x_0, y_0, 1)^\top = A^{-1}\boldsymbol{q}_0$ and $(x_i, y_i, 1)^\top = A^{-1}\boldsymbol{q}_i$.

High-order reflections: Let $\boldsymbol{p}_{ij}$ $(i, j = 1, \cdots, N_\pi,\ i \neq j)$ denote the reflection of $\boldsymbol{p}_i$ by π_{ij}. This $\boldsymbol{p}_{ij}$ can be expressed as $\boldsymbol{p}_{ij} = S_{ij}\boldsymbol{p}_i = S_{ij}S_i\boldsymbol{p}_0 = S_i S_j \boldsymbol{p}_0$ based on Eq (9). Here $\boldsymbol{p}_j = S_j \boldsymbol{p}_0$ holds as well, and we obtain $\boldsymbol{p}_{ij} = S_i S_j \boldsymbol{p}_0 \Leftrightarrow \boldsymbol{p}_{ij} = S_i \boldsymbol{p}_j$. This equation means that $\boldsymbol{p}_{ij}$ can be recognized as the first reflection of $\boldsymbol{p}_j$ by π_i, and hence the normalized image coordinates of their projections, $\boldsymbol{q}_{ij}$ and $\boldsymbol{q}_j$, also satisfy Eq (12) as:

$$\begin{pmatrix} y_j - y'_{ij} & x_{ij} - x_j & x_j y_{ij} - x_{ij} y_j \end{pmatrix} \boldsymbol{n}_i = 0, \tag{14}$$

where $(x_{ij}, y_{ij}, 1)^\top = A^{-1}\boldsymbol{q}_{ij}$.

In N_kth reflections, we obtain the kaleidoscopic projection constraint between $\boldsymbol{p}_{i_k} = S_{i_{N_k}} \Pi_{k=1}^{N_k - 1} S_{i_k} \boldsymbol{p}_0$ and $\boldsymbol{p}_{i'_k} = \Pi_{k=1}^{N_k - 1} S_{i_k} \boldsymbol{p}_0$ in the same manner.

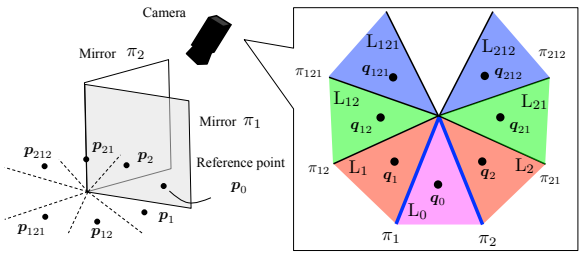


Fig. 8 Chamber assignment. The magenta region indicates the base chamber. The red, green and blue regions indicate the chambers corresponding to the first, second, third reflections respectively.

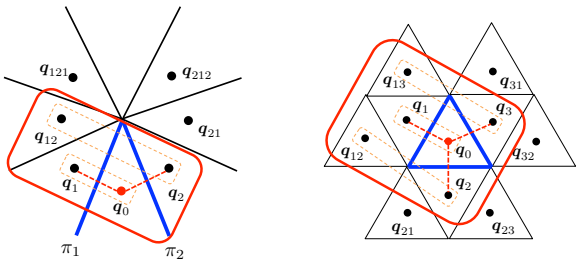


Fig. 9 Red boxes show examples of base structures in case of (a) $N_\pi = 2$ and (b) $N_\pi = 3$. A red point indicates the point assumed as the base chamber and dotted boxes indicate doublets. Red dotted lines indicate reflection pairs and blue lines indicate discovered mirrors.

3.4 Chamber Assignment

Based on the kaleidoscopic projection constraint, we introduce a new algorithm that identifies the chamber label of each projections. Our algorithm utilizes an analysis-by-synthesis approach which iteratively draws a number of projections and evaluates their geometric consistency in terms of the kaleidoscopic projection in order to find the best chamber assignment.

In what follows the concept of *base structure*, *i.e.* minimal configuration for estimating the real mirror parameters using the kaleidoscopic projection constraint, is introduced. Our algorithm hypothesizes a number of base structure candidates from observed points and evaluate each of their consistencies as a kaleidoscopic projection. Notice that we introduce our algorithm in a two mirror case as an example here, and it can be extended to three or more mirror cases easily.

3.4.1 Base Structure

Suppose $2N_\pi$ points of the observed points R are selected and they could be hypothesized as $\boldsymbol{q}_0, \boldsymbol{q}_1, \ldots$ correctly. The mirror normal $\boldsymbol{n}_i$ has two degrees of freedom and can be linearly estimated by collecting two or more linear constraints on it. In case of $N_\pi = 2$, the mirror normal $\boldsymbol{n}_1$ can be estimated as the eigenvector corresponding to the smallest eigenvalue of the coefficient matrix of the following system defined by the kaleidoscopic projection constraint using $\{\langle \boldsymbol{q}_0, \boldsymbol{q}_1\rangle, \langle \boldsymbol{q}_2, \boldsymbol{q}_{12}\rangle\}$ in Figure 9(a):

$$\begin{bmatrix} y_0 - y_1 & x_0 - x_1 & x_0 y_1 - x_1 y_0 \\ y_2 - y_{12} & x_{12} - x_2 & x_0 y_{12} - x_{12} y_0 \end{bmatrix} \boldsymbol{n}_1 = \boldsymbol{0}_{3\times 1}, \qquad (15)$$

where $\langle \boldsymbol{q}, \boldsymbol{q}' \rangle$ denotes a *doublet*, the pair of $\boldsymbol{q}$ and $\boldsymbol{q}'$ for Eq (12).

Using the estimated $\boldsymbol{n}_1$ and assuming $d_1 = 1$ without loss of generality, the 3D point $\boldsymbol{p}_1$ can be described as $\tilde{\boldsymbol{p}}_1 = S_1 \tilde{\boldsymbol{p}}_0$ by Eq (7). By substituting $\boldsymbol{p}_0$ and $\boldsymbol{p}_1$ in this equation by using $\boldsymbol{q}_0$ and $\boldsymbol{q}_1$ as expressed in Eq (4), the 3D point $\boldsymbol{p}_0$ and $\boldsymbol{p}_1$ can be triangulated by solving the following linear system for λ_0 and λ_1:

$$\tilde{\boldsymbol{p}}_0 = S_1 \tilde{\boldsymbol{p}}_1, \qquad (16)$$

$$\Leftrightarrow \begin{bmatrix} H_1 A^{-1} \boldsymbol{q}_0 & -A^{-1} \boldsymbol{q}_1 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \end{bmatrix} = 2\boldsymbol{n}_1. \qquad (17)$$

Similarly the 3D points $\boldsymbol{p}_2$ and $\boldsymbol{p}_{12}$ can be triangulated by solving the linear system for λ_2 and λ_{12}. Because $\boldsymbol{p}_2$ is the reflection of $\boldsymbol{p}_0$ by the mirror π_2, the mirror normal $\boldsymbol{n}_2$ as well as the distance d_2 can be estimated as

$$\boldsymbol{n}_2 = \frac{\boldsymbol{p}_0 - \boldsymbol{p}_2}{|\boldsymbol{p}_0 - \boldsymbol{p}_2|}, \qquad d_2 = -\boldsymbol{n}_2^\top \frac{\boldsymbol{p}_0 + \boldsymbol{p}_2}{2}. \qquad (18)$$

This doublets pair $\{\langle \boldsymbol{q}_0, \boldsymbol{q}_1\rangle, \langle \boldsymbol{q}_2, \boldsymbol{q}_{12}\rangle\}$ is a minimal configuration for linear estimation of the real mirror parameters in $N_\pi = 2$ case and we call such minimal configuration as a *base structure* of our chamber assignment. Notice that the above doublet pair is not the unique base structure. That is, $\{\langle \boldsymbol{q}_0, \boldsymbol{q}_2\rangle, \langle \boldsymbol{q}_1, \boldsymbol{q}_{21}\rangle\}$ is also a base structure for $N_\pi = 2$ case.

3.4.2 Chamber Assignment Algorithm

Given the mirror parameters and the triangulated 3D point from one base structure, the kth reflection and its projection can be synthesized by Eq (7) and Eq (4) with known chamber labels. These labels are assigned by finding correspondences between the synthesized point $\hat{Q}$ and the observed point set $R = \{\boldsymbol{r}_i\}$ as a sort of biparite matching.

Based on this assignment, the proposed method introduces a recall ratio $\mathcal{R}$ which measures how many of the synthesized projections that are supposed to be visible have been assigned detected points:

$$\mathcal{R} = \frac{|R_c|}{|\hat{Q}|}, \qquad (19)$$

where $R_c \subseteq R$ is the set of detected points assigned labels, $|\hat{Q}|$ and $|R_c|$ denote the size of the set $\hat{Q}$ and R_c respectively.

The proposed chamber assignment algorithm examines if each possible base structure satisfies the kaleidoscopic projection constraint expressed by Eq (15) and other geometric constraints, which are detailed in [26]. Once the base structure passes these verifications, its recall ratio $\mathcal{R}$ is computed by Eq (19). Finally the best estimate of the chamber assignment is returned by finding the base structure of the highest recall ratio.

3.5 Mirror Parameters Estimation

This section introduces a novel algorithm of mirror parameters estimation given the chamber assignment for kaleidoscopic projections of single 3D point.

While the algorithm in Section 3.4.1 can also estimate the mirror parameters, it is a per-mirror estimation and it is not guaranteed to estimate mirror parameters consistent with projections of higher order reflections. Instead of such mirror-wise estimations, this section proposes a new linear algorithm which calibrates the kaleidoscopic mirror parameters simultaneously by observing a single 3D point in the scene. Notice that the algorithm is first introduced by utilizing up to the second reflections, but they can be extended to third or further reflections.

3.5.1 Mirror Normals

As illustrated in Figure 10 (a), suppose a 3D point $\boldsymbol{p}_0$ is projected to $\boldsymbol{q}_0$ in the base chamber, and its mirror $\boldsymbol{p}_i$ by π_i is pro-

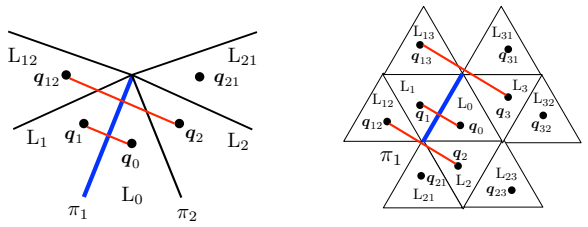


Fig. 10 Corresponding points in (a) $N_\pi = 2$ and (b) $N_\pi = 3$ case. (a) Two pairs $\langle \boldsymbol{q}_0, \boldsymbol{q}_1 \rangle$ and $\langle \boldsymbol{q}_2, \boldsymbol{q}_{12} \rangle$ (red) are available on mirror π_1 (blue). (b) Three pairs $\langle \boldsymbol{q}_0, \boldsymbol{q}_1 \rangle$, $\langle \boldsymbol{q}_2, \boldsymbol{q}_{12} \rangle$ and $\langle \boldsymbol{q}_3, \boldsymbol{q}_{13} \rangle$(red) are available on mirror π_1 (blue)

jected to $\boldsymbol{q}_i$ in the chamber L_i. Likewise, the second mirror $\boldsymbol{p}_{ij}$ by π_{ij} is projected to $\boldsymbol{q}_{ij}$ in the chamber L_{ij}, and so forth. Here, kaleidoscopic projection constraints are satisfied by two pairs of projections on each mirror π_1 and π_2. From these constraints, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ can be estimated by solving

$$\begin{bmatrix} y_0 - y_1 & x_1 - x_0 & x_0 y_1 - x_1 y_0 \\ y_2 - y_{12} & x_{12} - x_2 & x_2 y_{12} - x_{12} y_2 \end{bmatrix} \boldsymbol{n}_1 = \boldsymbol{0}_{3\times 1}. \tag{20}$$

and

$$\begin{bmatrix} y_0 - y_2 & x_2 - x_0 & x_0 y_2 - x_2 y_0 \\ y_2 - y_{21} & x_{21} - x_1 & x_1 y_{21} - x_{21} y_1 \end{bmatrix} \boldsymbol{n}_2 = \boldsymbol{0}_{3\times 1}. \tag{21}$$

An important observation in this simple algorithm is the fact that (1) this is a linear algorithm while it utilizes multiple reflections, and (2) the estimated normals $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are enforced to be consistent with each other while they are computed on a per-mirror basis apparently.

The first point is realized by using not the multiple reflections of a 3D position but their 2D projections. Intuitively a reasonable formalization of kaleidoscopic projection is to define a real 3D point in the scene, and then to express each of the projections of its reflections by Eq (7) coincides with the observed 2D position as introduced in Section 3.5.3 later. This expression, however, is nonlinear in the normals $\boldsymbol{n}_i$ $(i = 1, 2)$ (*e.g.* $\boldsymbol{p}_{12} = S_1 S_2 \boldsymbol{p}_0$). On the other hand, projections of such multiple reflections can be associated as a result of single reflection by Eq (14) directly (*e.g.* $\boldsymbol{n}_1$ with $\boldsymbol{q}_{12}$ and $\boldsymbol{q}_2$ as the projections of $\boldsymbol{p}_{12}$ and $S_2 \boldsymbol{p}_0$ respectively). As a result, we can utilize 2D projections of multiple reflections in the linear systems above.

This explains the second point as well. The above constraint on $\boldsymbol{q}_{12}$, $\boldsymbol{q}_2$ and $\boldsymbol{n}_1$ in Eq (20) assumes $\boldsymbol{p}_2 = g S_2 \boldsymbol{p}_0$ being satisfied, and it is enforced by $(A^{-1}\boldsymbol{q}_2 \times A^{-1}\boldsymbol{q}_0)^\top \boldsymbol{n}_2 = 0$ in the first row of Eq (21). Inversely, on estimating $\boldsymbol{n}_1$ by Eq (20), it enforces $\boldsymbol{p}_1 = S_1 \boldsymbol{p}_0$ for Eq (21).

Note that this algorithm can be extended to third or further reflections intuitively. For example, if $\boldsymbol{p}_{21}$ and its reflection by π_1 are observable as $\lambda_{121}\boldsymbol{q}_{121} = A\boldsymbol{p}_{121} = AS_1\boldsymbol{p}_{21}$, then they provide

$$(y_{21} - y_{121}, x_{121} - x_{21}, x_{21} y_{121} - x_{121} y_{21})\, \boldsymbol{n}_1 = 0, \tag{22}$$

and can be integrated with Eq (20).

Also, this algorithm can be extended to $N_\pi \geq 3$ cases. In case of $N_\pi = 3$, for example, we solve

$$\begin{bmatrix} y_0 - y_1 & x_1 - x_0 & x_0 y_1 - x_1 y_0 \\ y_2 - y_{12} & x_{12} - x_2 & x_2 y_{12} - x_{12} y_2 \\ y_3 - y_{13} & x_{13} - x_3 & x_3 y_{13} - x_{13} y_3 \end{bmatrix} \boldsymbol{n}_1 = \boldsymbol{0}_{3\times 1}, \tag{23}$$

instead of Eqs (20) from point correspondences in Figure 10 (b).

3.5.2 Mirror Distances

Once the mirror normals $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are given linearly, the mirror distances d_1 and d_2 can also be estimated linearly as follows.

Kaleidoscopic Re-projection Constraint: The perspective projection Eq (4) indicates that a 3D point $\boldsymbol{p}_i$ and its projection $\boldsymbol{q}_i$ should satisfy the colinearity constraint:

$$(A^{-1}\boldsymbol{q}_i) \times \boldsymbol{p}_i = \boldsymbol{x}_i \times \boldsymbol{p}_i = \boldsymbol{0}_{3\times 1}, \tag{24}$$

where $\boldsymbol{x}_i = \begin{pmatrix} x_i & y_i & 1 \end{pmatrix}^\top$ is the normalized camera coordinate of $\boldsymbol{q}_i$. Since the mirrored points $\boldsymbol{p}_i$ $(i = 1, 2)$ are given by Eq (7) as $\boldsymbol{p}_i = H_i \boldsymbol{p}_0 - 2d_i \boldsymbol{n}_i$, we obtain

$$\boldsymbol{x}_i \times \boldsymbol{p}_i = [\boldsymbol{x}_i]_\times \begin{bmatrix} H_i & -2\boldsymbol{n}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_0 \\ d_i \end{bmatrix} = \boldsymbol{0}_{3\times 1}. \tag{25}$$

Similarly, $\boldsymbol{p}_{ij}$ is also collinear with its projection $\boldsymbol{q}_{ij}$:

$$(A^{-1}\boldsymbol{q}_{ij}) \times \boldsymbol{p}_{ij} = [\boldsymbol{x}_{ij}]_\times \begin{bmatrix} H_i H_j & -2\boldsymbol{n}_i & -2H_i\boldsymbol{n}_j \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_0 \\ d_i \\ d_j \end{bmatrix} = \boldsymbol{0}_{3\times 1}. \tag{26}$$

Based on them, we obtain a linear system of $\boldsymbol{p}_0$, d_1 and d_2:

$$\begin{bmatrix} [\boldsymbol{x}_0]_\times & \boldsymbol{0}_{3\times 1} & \boldsymbol{0}_{3\times 1} \\ h_1 & -2[\boldsymbol{x}_1]_\times \boldsymbol{n}_1 & \boldsymbol{0}_{3\times 1} \\ h_2 & \boldsymbol{0}_{3\times 1} & -2[\boldsymbol{x}_2]_\times \boldsymbol{n}_2 \\ h'_{1,2} & -2[\boldsymbol{x}_{12}]_\times \boldsymbol{n}_1 & -2h''_{1,2} \\ h'_{2,1} & -2h''_{2,1} & -2[\boldsymbol{x}_{21}]_\times \boldsymbol{n}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_0 \\ d_1 \\ d_2 \end{bmatrix} = K \begin{bmatrix} \boldsymbol{p}_0 \\ d_1 \\ d_2 \end{bmatrix} = \boldsymbol{0}_{15\times 1}, \tag{27}$$

where $h_i = [\boldsymbol{x}_i]_\times H_i$, $h'_{i,j} = [\boldsymbol{x}_{ij}]_\times H_i H_j$, $h''_{i,j} = [\boldsymbol{x}_{ij}]_\times H_i \boldsymbol{n}_j$. By computing the eigenvector corresponding to the smallest eigenvalue of $K^\top K$, $(\boldsymbol{p}_0, d_1, d_2)^\top$ can be determined up to a scale factor. In this paper, we choose the scale that normalizes $d_1 = 1$.

Notice that Eq (27) apparently has 15 equations, but only 10 of them are linearly independent. This is simply because each of the cross products by Eqs (24) and (26) has only two independent constraints by definition.

Also, as discussed in Section 3.5.1, the above algorithm can be extended to third or further reflections and $N_\pi \geq 3$ cases as well. In $N_\pi = 3$, considering the reflection of $\boldsymbol{p}_{23}$ by π_1 as $\lambda_{123}\boldsymbol{q}_{123} = A\boldsymbol{p}_{123} = AS_1\boldsymbol{p}_{23}$, we have

$$[\boldsymbol{x}_{123}]_\times \begin{bmatrix} (H_1 H_2 H_3)^\top \\ -2\boldsymbol{n}_1^\top \\ -2(H_1\boldsymbol{n}_2)^\top \\ -2(H_1 H_2 \boldsymbol{n}_3)^\top \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{p}_0 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \boldsymbol{0}_{3\times 1}. \tag{28}$$

3.5.3 Kaleidoscopic Bundleadjustment

Once estimated the mirror normals $\boldsymbol{n}_i$ and the distances d_i $(i = 1, 2)$ linearly, the triangulation from kaleidoscopic projections of a single 3D point can be given in a DLT manner by solving:

$$K'\boldsymbol{p}_0 = -K''\boldsymbol{\delta}, \tag{29}$$

as $\boldsymbol{p}_0^* = -(K'^\top K')^{-1} K'^\top K''\boldsymbol{\delta}$, where $\boldsymbol{\delta} = (d_1, d_2)^\top$, K' is the 15×3 matrix corresponding to the first three columns of K:

$$K' = \left[[\boldsymbol{x}_0]_\times^\top, h_1^\top, h_2^\top, h_{1,2}'^\top, h_{2,1}'^\top \right]^\top, \tag{30}$$

and K'' is the 15×3 matrix corresponding to the 4th and 5th columns of K:

$$K'' = \begin{bmatrix} \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} \\ -2[\boldsymbol{x}_1]_\times \boldsymbol{n}_1 & \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times1} & -2[\boldsymbol{x}_2]_\times \boldsymbol{n}_2 \\ -2[\boldsymbol{x}_{12}]_\times \boldsymbol{n}_1 & -2h''_{1,2} \\ -2h''_{2,1} & -2[\boldsymbol{x}_{21}]_\times \boldsymbol{n}_2 \end{bmatrix}. \tag{31}$$

By reprojecting this $\boldsymbol{p}_0^*$ to each of the chambers as

$$\begin{aligned} \lambda\hat{\boldsymbol{q}}_0 &= A\boldsymbol{p}_0^*, \\ \lambda\hat{\boldsymbol{q}}_i &= AS_i\boldsymbol{p}_0^* \ (i = 1, 2), \\ \lambda\hat{\boldsymbol{q}}_{i,j} &= AS_iS_j\boldsymbol{p}_0^* \ (i, j = 1, 2,\ i \neq j), \end{aligned} \tag{32}$$

we obtain a reprojection error as

$$\boldsymbol{E}(\boldsymbol{n}_1, \boldsymbol{n}_2, d_1, d_2) = \left[\boldsymbol{q}_0 - \hat{\boldsymbol{q}}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}'_{1,2}, \boldsymbol{e}'_{2,1},\right]^\top, \tag{33}$$

where $\boldsymbol{e}_i = \boldsymbol{q}_i - \hat{\boldsymbol{q}}_i$ and $\boldsymbol{e}'_{i,j} = \boldsymbol{q}'_{i,j} - \hat{\boldsymbol{q}}'_{i,j}$. By minimizing $\|\boldsymbol{E}(\cdot)\|^2$ nonlinearly over $\boldsymbol{n}_1, \boldsymbol{n}_2, d_1, d_2$, we obtain a best estimate of the mirror normals and the distances.

3.6 Intrinsic Parameters Estimation

As described in Section 3.5, the mirror parameters can be computed from the kaleidoscopic observations undistorted with the intrinsic parameters. Finally, the proposed method optimizes these parameters by minimizing the reprojection error as follow.

Suppose $\breve{\boldsymbol{q}}^*(A, \boldsymbol{d}, \boldsymbol{n}, \delta)$ denotes the reprojected point from A, $\boldsymbol{d}$, $\boldsymbol{n}$ and δ, the proposed method defines the reprojection error as,

$$\boldsymbol{E}(A, \boldsymbol{d}, \boldsymbol{n}, \delta) = \Sigma_x \|\breve{\boldsymbol{q}}_x - \breve{\boldsymbol{q}}_x^*(A, \boldsymbol{d}, \boldsymbol{n}, \delta)\|^2. \tag{34}$$

By minimizing $\boldsymbol{E}(\cdot)$ nonlinearly over A, $\boldsymbol{d}$, $\boldsymbol{n}$ and δ, we obtain a best estimate of the intrinsic parameters.

3.7 Evaluations of Intrinsic Camera Calibration

3.7.1 Quantitative Evaluations with Synthesized Data

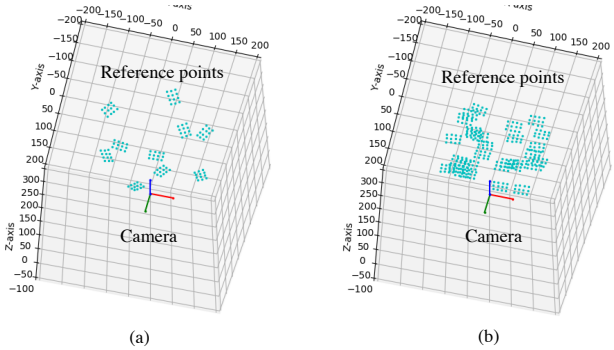
Experimental Environment The performance of intrinsic camera calibration with synthesized data is evaluated in the following configuration.

The kaleidoscopic imaging system in this evaluation consists of three mirrors $\pi_i (i = 0, 1, 2)$ whose normal vectors $\boldsymbol{n}_i$ of each mirror are set to $(\cos\theta\cos\phi, \sin\theta, \cos\theta\sin\phi)$ with $(\theta, \phi) = (-8, 0)$ for $\boldsymbol{n}_0$, $(\theta, \phi) = (186, 60)$ for $\boldsymbol{n}_1$, and $(\theta, \phi) = (190, -60)$ for $\boldsymbol{n}_2$. The distance of them d_i are set to $d_0 = 50$mm, $d_1 = 53$mm and $d_2 = 54$mm. The image size is set to 1920×1080. The ground truth of intrinsic parameters (f, u_0, v_0) and $\boldsymbol{d} = (k_1, k_2, k_3, p_1, p_2)$ are set to $(2700, 960, 540)$ and $(0.001, -0.001, 0.001, 0.002, -0.002)$. Notice that the chamber labels are assigned correctly in this valuation.

The evaluations in this section utilize the absolute error of each parameters as error metrics, that is $E_p = |p - p_g|$ where p denotes each intrinsic camera parameters and p_g denotes its ground truth.

The performance of our method is compared with the most common intrinsic camera calibration algorithm of Zhang [31].

Results In this evaluation, the performance of our proposed method is compared with the most common intrinsic camera calibration algorithm of Zhang [31] (Baseline 1). For comparison, we set a 3×4 chessboard whose distance of each chess corner is 10mm in the base chamber whose distance from the camera is 160mm in the kaleidoscopic imaging system. As illustrated in Figure 11(1), we utilize the projections of its reflections in 10 chambers, that is until second reflection, as input for each method.



(a) (b)

Fig. 11 Configurations of intrinsic camera calibration. (a) utilizes the 3×4 chessboards and its reflections in 10 chambers. (b) utilizes 30 random poses of the chessboards.

In addition, we compare the results by [31] with the ideal configuration as a reference data, that is the input data consists of 30 observations of above chessboard scattered covering the camera's field-of-view in the 3D space randomly as illustrated in Figure 11(b) (Baseline 2).

Figure 12 shows average estimation errors of each intrinsic camera parameters over 30 trials at different noise σ_q. The σ_q denotes the standard deviation of zero-mean Gaussian pixel noise injected to the observations $\boldsymbol{q}$. In each trial, the initial values of A and $\boldsymbol{d}$ are generated by adding random noise whose level is less than 5% to the ground truth of each parameters. The red, and blue are results by the proposed ethod and Baseline 1 and green line is the result by Baseline 2.

In Figure 12, we can see that the proposed method outperforms Baseline 1 and is comparable with the Baseline 2. Especially, distortion parameters estimated by the proposed method have higher precision that those by the other methods. We consider that this improvement is caused by the properties of the kaleidoscopic imaging system, *i.e.* the reference objects are scattered in an isotropic manner by mirror reflections and the reflections are strictly constrained each other with mirror parameters.

From these results, we can conclude that our method can esitmate intrinsic parameters in an ideal and a noised environment robustly compared with baseline methods.

3.7.2 Qualitative Evaluations with Real Data

Experimental Environment As illustrated in Figure 13, the capture setup consists of one camera (Nikon D7000, 4948×3280 resolution) with a single-focus lens of 28mm and three planar mirrors. In this evaluation, we compare our method with Zhang [31] (Baseline 1) as with 3.7.1. The reference object for Baseline 1 is a 5×8 chessboard whose distance of each chess corner is 4.45mm. As an ideal configuration for Zhang [31], we utilize a larger reference object which has 7×10 chessboard with 20.5mm corners (Baseline 2).

Results Table 1 reports the each parameters estimated by each method. Notice that proposed method and [31] with the kaleidoscopic imaging system utilize 10 projections of the reference

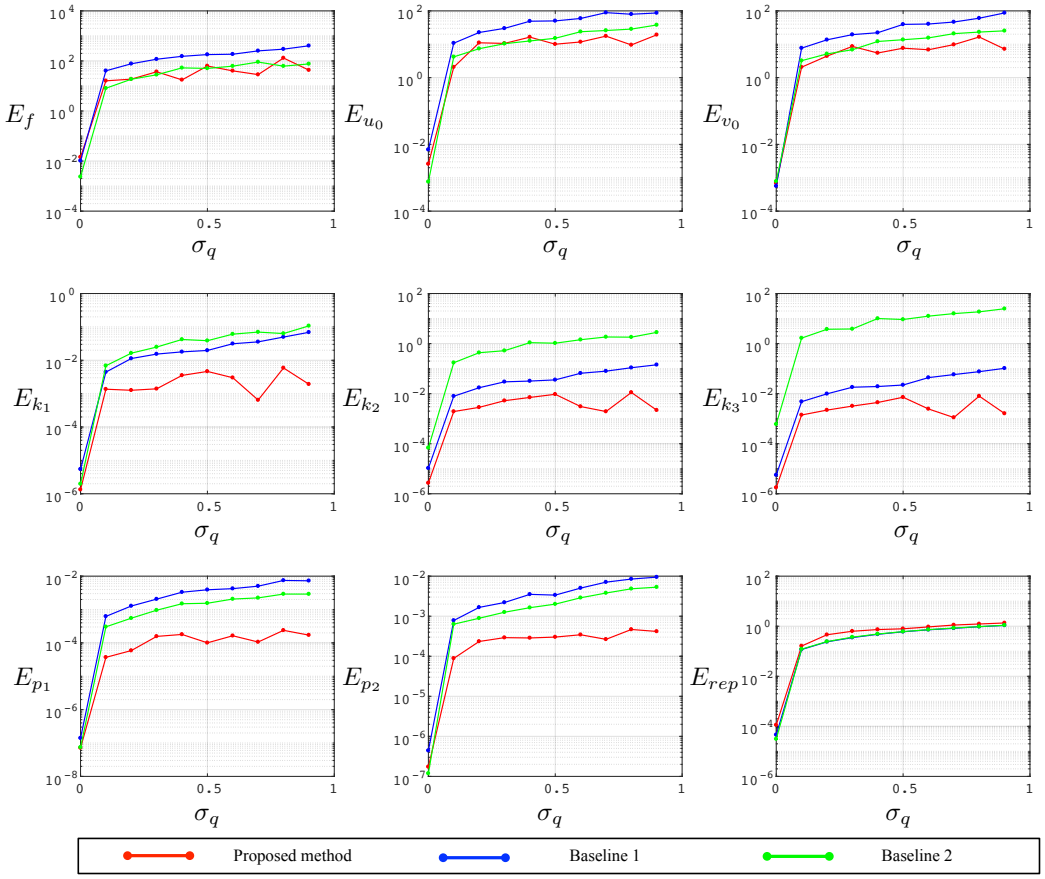
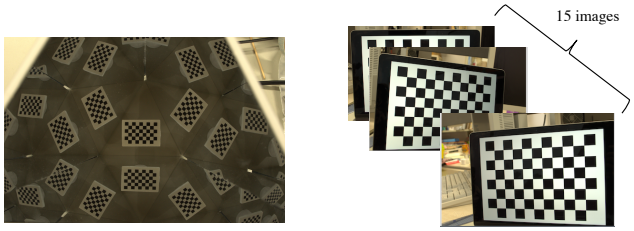


Fig. 12 Estimation errors of each intrinsic parameter and reprojection error at different noise levels σ_q.



(a) (b)

Fig. 13 Configurations of intrinsic camera calibration with real data. (a) utilizes the 5×8 chessboards and its reflections in 10 chambers. (b) utilizes 15 random poses of the larger 7×10 chessboards.

Table 1 Intrinsic parameters estimated by each method.

Parameters	Proposed	Baseline 1	Baseline 2
f	5882.2	12318.3	6283.4
u_0	2474.4	2462.7	2427.6
v_0	1639.4	1622.9	1686.2
k_1	-0.0103	-0.3381	-0.1059
k_2	0.0143	-86.1886	0.3035
k_3	-0.0020	2007.4	-0.9600
p_1	0.0056	0.0740	0.00109
p_2	0.0036	0.0625	0.0006

Table 2 Reprojection error by each method for evaluating robustness.

Data	Proposed	Baseline 1	Baseline 2
1	0.526	0.487	0.524
2	1.735	13.014	0.933

object, that is projections in each chambers (Figure 13(a)), and citezhang2000flexible with the ideal configuration utilizes 15 images of the larger reference object (Figure 13(b)). From these results, while the intrinsic parameters estimated by Baseline 1 are different greatly compared with those by Baseline 2 which is considered as the ground truth, the parameters estimated by proposed method are close to those by Baseline 2.

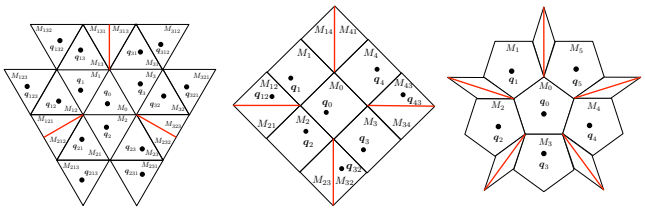
Table 2 shows the average of reprojection errors by each method for evaluating their robustness. Data 1 consists of 10 observations of 5×8 chessboard used for estimating intrinsic parameters by proposed method and Baseline 1. Data 2 consists of 5 images of large 7×10 chessboard used for estimating intrinsic parameters by Baseline 2. As to the Baseline 1, based on the fact that the reprojection errors with Data 2 get worse apparently, the intrinsic parameters estimated by Baseline 1 does not have robustness. On the other hand, the reprojection errors by our proposed method with both Data 1 and Data 2 are small enough and this shows that it estimates robust intrinsic parameters.

From above results, we can state that our proposed method works properly in the real configuration.

3.8 Discussion

3.8.1 Ambiguity of chamber assignment

In the case of the 8-point algorithm for the regular two-view extrinsic calibration [10], the linear algorithm return four possible combinations of the rotation and the translation, and we can choose the right combination by examining if triangulated 3D points appear in front of the cameras. The mirror normal estimation in Section 3.4.1 is a special case of the 8-point algorithm, and this has such sign ambiguity on the mirror normal as described in



(a) Three mirrors (b) Four mirrors (c) Five mirrors

Fig. 14 Kaleidoscopic imaging system using (a) three, (b) four, and (c) five mirrors. Discontinuities (red lines) appear on the boundaries of overlapping chambers.

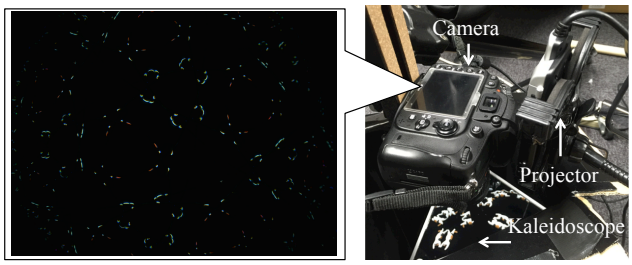


Fig. 15 A kaleidoscopic capture setup for 3D reconstruction. It consists of three first surface mirrors, a camera and a laser projector.

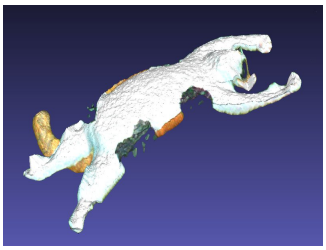


Fig. 16 Reconstructed 3D shape of *cat* object.

Section 3.4.1. This ambiguity is also solved by considering the result of triangulation. In other words, estimating the essential matrix is identical to estimating mirror normal.

In addition, the normal estimation for kaleidoscopic system has another family of ambiguity due to multiple reflections. As introduced in Section 3.4.1, particular combinations of kaleidoscopic projections can return physically infeasible solutions, and they can be rejected by additional geometric constraints as done for the 8-point algorithm. However, there exists another class of solutions due to a *sparse sampling* of the observations.

Consider a base structure by the pairs $\langle \boldsymbol{q}_0, \boldsymbol{q}_{12} \rangle$ and $\langle \boldsymbol{q}_2, \boldsymbol{q}_{121} \rangle$ in Figure 8. This configuration can estimate the mirror parameters successfully, one between $\boldsymbol{p}_0$ and $\boldsymbol{p}_2$, and the other between $\boldsymbol{p}_0$ and $\boldsymbol{p}_{12}$. While the latter is a virtual mirror, this interpretation satisfies all the constraints in Section 3.4.1. In other words, we can assemble a mirror system of this configuration in practice.

To solve this problem, Section 3.4 utilizes the recall ratio (Eq (19)) so that our algorithm returns the solution which reproduces as many as possible candidate points $\boldsymbol{r}$ observed in the image.

3.8.2 RANSAC or PROSAC for Chamber Assignment

While the proposed algorithm examines all possible base structures as introduced in Section 3.4.2 to evaluate the performance thoroughly, we can also consider a RANSAC or PROSAC approach [3]. For example, we can first hypothesize the base chamber from N_r candidates, and then can consider only $N_\pi + 1$ nearest points around it for estimating the mirror parameters. Designing and evaluating such approach is one of our future work.

3.8.3 Degenerate Cases

The both proposed algorithms of chamber assignment and mirror parameters estimation are based on the kaleidoscopic projection constraint (Eq (12)) satisfied by more than or equal to second reflections. Therefore these algorithms do not work in the following two cases. (1) If the two mirror are parallel, the mirror normals are not computable by solving Eq (15), Eq (20) and Eq (21) because the constraints are linearly dependent. (2) If the second reflections are not observable due to the angle of view or discontinuities, the mirror normals are not computable. Especially, in case of using more than three mirrors, discontinuities are more likely to happen in general, and finding the second reflections itself become difficult (Figure 14).

3.8.4 Application: 3D Reconstruction

The kaleidoscopic system can be recognized as a virtual multi-view capture system. Here we evaluated the feasibility of 3D reconstruction with estimated mirror parameters.

Figure 15 shows our kaleidoscopic capture setup. In this evaluation, the intrinsic parameter A of the camera (Nikon D600, 6016×4016 resolution) is calibrated beforehand [31]. As a target object, we utilized a *cat* (about 4×5×1 cm) with three planar first surface mirrors. The projector (MicroVision SHOWWX+ Laser Pico Projector, 848×480 resolution) is used to cast line patterns to the object for simplifying the correspondence search problem in a light-sectioning fashion (Figure 15 left), and the projector itself is not involved in the calibration *w.r.t.*the camera and the mirrors. In this evaluations, the mirror parameters are estimated by the algorithm introduced in Section 3.5.

Figure 16 shows a 3D rendering of the estimated 3D shape using the mirror parameters calibrated by the proposed method, while the residual reprojection error indicates the parameters can be further improved for example through the 3D shape reconstruction process itself [7]. From these results, we can conclude that the proposed provides a sufficiently accurate calibration for 3D shape reconstruction.

3.9 Summary

This section proposes a novel algorithm of full-automatic intrinsic camera calibration algorithm introducing multiple mirrors and consisting a kaleidoscopic imaging system. To realize this algorithm, there are hree challenges to be solved, *i.e.* (1) chamber assignment, (2) mirror parameters estimation and (3) intrinsic parameters estimation. For solving them, the proposed method introduces a *kaleidoscopic projection constraint* which is satisfied by projections of high-order reflections. Evaluations with synthesized data and real data prove that the proposed method estimates intrinsic parameters with reasonable precision robustly.

This research mainly focuses on the perspective camera model with simple lens distortion model. Along with the development of imaging systems, a wide variety of camera models [9, 15, 21] and lens models [2, 25] have been proposed. The adaptation to such configuration should be investigated as a future work.

4. Mirror-based Extrinsic Camera Calibration

This section provides a novel planar-mirror based extrinsic camera calibration algorithm in case where a camera cannot observe a reference object directly as illustrated in Figure **??**.

4.1 Measurement Model

As illustrated by Figure 17, we denote a camera by C and a mirror by π_j ($j = 1, \cdots, N_\pi$). We use $\{C\}$ to describe the camera C coordinate system which is used as the world coordinate system in this chapter. A vector $\boldsymbol{p}$ in the Y coordinate system is expressed as $\boldsymbol{p}^{\{Y\}}$, while we may omit Y if it is clear from the context.

Let $\boldsymbol{p}^{i\{X\}}$ ($i = 1, \cdots, N_p$) denote the positions of the reference points given a priori in its local coordinate system X. These positions are modeled as

$$\boldsymbol{p}^{i\{C\}} = R \cdot \boldsymbol{p}^{i\{X\}} + \boldsymbol{t} \quad (i = 1, \cdots, N_p), \tag{35}$$

in $\{C\}$ with a rotation matrix R and a translation vector $\boldsymbol{t}$. The reflection of the ith reference point $\boldsymbol{p}^{i\{C\}}$ mirrored by π_j appears as $\boldsymbol{p}_j^{i\{C\}}$ in $\{C\}$. These mirrored reference points are projected to the image screen of the real camera C as $\boldsymbol{q}_j^{i\{I\}}$. We model each mirror π_j by its normal vector $\boldsymbol{n}_j$ and its distance d_j from the camera C. The distance t_j^i from the mirror π_j to $\boldsymbol{p}_j^{i\{C\}}$ is equal to the distance from π_j to $\boldsymbol{p}^{i\{C\}}$ by definition. The goal of the extrinsic calibration is to estimate R and $\boldsymbol{t}$ from projected reference points $\boldsymbol{q}_j^{i\{I\}}$.

Based on the planar mirror geometry introduced in Section 2.2, the reference point $\boldsymbol{p}^{i\{C\}}$ and its reflection $\boldsymbol{p}_j^{i\{C\}}$ satisfy,

$$\boldsymbol{p}^{i\{C\}} = -2(\boldsymbol{n}_j^\top \boldsymbol{p}_j^{i\{C\}} + d_j)\boldsymbol{n}_j + \boldsymbol{p}_j^{i\{C\}}. \tag{36}$$

By removing $\boldsymbol{p}^{i\{C\}}$ from Eq (35) and Eq (36), we obtain

$$R \cdot \boldsymbol{p}^{i\{W\}} + \boldsymbol{t} = -2(\boldsymbol{n}_j^\top \cdot \boldsymbol{p}_j^{i\{C\}} + d_j)\boldsymbol{n}_j + \boldsymbol{p}_j^{i\{C\}}. \tag{37}$$

This is the fundamental equation which describes our measurement model.

4.2 Orthogonality Constraint on Mirror Reflections

Consider a reference point $\boldsymbol{p}^i$ and its two mirrored points $\boldsymbol{p}_j^i$, $\boldsymbol{p}_{j'}^i$ by two different mirror planes π_j and $\pi_{j'}$ respectively. The axis vector $\boldsymbol{m}_{jj'}$ lying along the intersection of the two mirror planes is expressed as the cross product of each mirror normals, $\boldsymbol{m}_{jj'} = \boldsymbol{n}_j \times \boldsymbol{n}_{j'}$. This axis vector $\boldsymbol{m}_{jj'}$ satisfies the following *orthogonality constraint* [24] (Figure 18),

$$(\boldsymbol{p}_j - \boldsymbol{p}_{j'})^\top \cdot \boldsymbol{m}_{jj'} = 0. \tag{38}$$

This is the key constraint of this paper. The next section provides our algorithm to estimate the extrinsic parameters which utilizes this constraint.

4.3 Extrinsic Camera Calibration Using Orthogonality Constraint

This section introduces the proposed algorithm which analytically determines the camera extrinsic parameter from the projections of N_p reference points observed via N_π different mirror poses based on the orthogonality constraint.

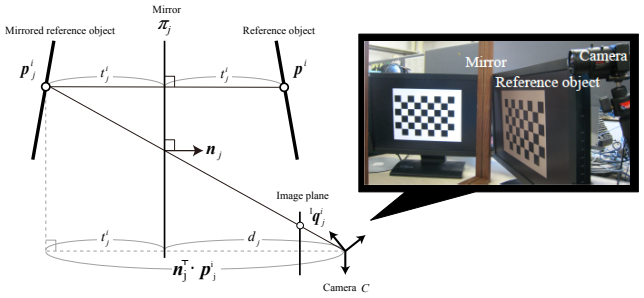


Fig. 17 The measurement model

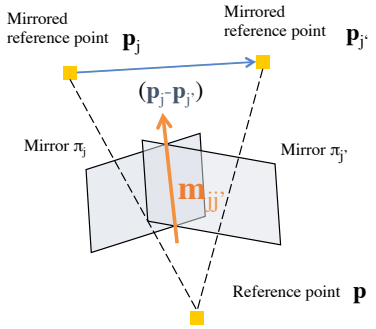


Fig. 18 An illustration of orthogonality constraint.

Algorithm 1 shows an overview of our calibration algorithm. Firstly, we solve the PnP problem from mirrored reference points projected to the image plane, ${}^I\boldsymbol{q}_j^i$, and obtain their 3D positions $\boldsymbol{p}_j^i$ by EPnP [17] ($N_p > 3$ case). Notice that the "handedness" of the extrinsic parameters obtained by solving PnP with mirrored reference points are flipped. However this does not affect the 3D position $\boldsymbol{p}_j^i$, and hence we ignore such flipped extrinsic parameters. Secondly, we estimate the axis vectors of each pair of mirror planes and obtain the mirror normals from them based on the othogonality constraint. Finally, we compute R and $\boldsymbol{t}$ by solving a large system of linear equations.

4.3.1 Computing the axis vector from mirror planes

As described in Sec 4.2, the axis vector $\boldsymbol{m}_{jj'}(j, j' = 1, \cdots, N_\pi, \quad j \neq j')$ and two mirrored points $\boldsymbol{p}_j^i$, $\boldsymbol{p}_{j'}^i(i = 1, 2, \cdots, N_p)$ satisfy the orthogonality constraint (Eq (38)).

By applying this orthogonality constraint to N_p mirrored reference points $\boldsymbol{p}_j^i$, we obtain:

$$\begin{pmatrix} (\boldsymbol{p}_j^1 - \boldsymbol{p}_{j'}^1)^\top \\ (\boldsymbol{p}_j^2 - \boldsymbol{p}_{j'}^2)^\top \\ \vdots \\ (\boldsymbol{p}_j^{N_p} - \boldsymbol{p}_{j'}^{N_p})^\top \end{pmatrix} \boldsymbol{m}_{jj'} = \boldsymbol{Q}_{jj'}\boldsymbol{m}_{jj'} = 0. \tag{39}$$

An axis vector $\boldsymbol{m}_{jj'}$ can be computed as the right-singular vector corresponding to the smallest singular value of $\boldsymbol{Q}_{jj'}$.

4.3.2 Computing the normal vector of a mirror plane

The axis vector $\boldsymbol{m}_{jj'}$ is perpendicular to the normal vectors $\boldsymbol{n}_j$ and $\boldsymbol{n}_{j'}$ of each mirror planes π_j and $\pi_{j'}$ respectively. That is,

$$\begin{aligned} \boldsymbol{n}_j^\top \cdot \boldsymbol{m}_{jj'} &= 0, \\ \boldsymbol{n}_{j'}^\top \cdot \boldsymbol{m}_{jj'} &= 0. \end{aligned} \tag{40}$$

When using N_π mirror poses, we obtain $(N_\pi - 1)$ equations of Eq

Algorithm 1 An overview of extrinsic camera calibration using N_π mirror poses and N_p reference points

Require: ${}^l\boldsymbol{q}^i_j$ $(i = 1, \cdots, N_p, \quad j = 1, \cdots, N_\pi)$
Ensure: $R, \boldsymbol{t}$
  for all N_π mirror poses $\pi_j, j = 1, \cdots, N_\pi$ **do**
    Solve the PnP and obtain the mirrored points $\boldsymbol{p}^i_j$.
  end for
  for all N_π mirror poses $\pi_j, \pi_{j'}, (j, j' = 1, \cdots, N_\pi, \quad j \neq j')$ **do**
    Compute the axis vector $\boldsymbol{m}_{jj'}$ as the right-singular vector corresponding to the smallest singular value of $\boldsymbol{Q}_{jj'}$ in Eq (39).
  end for
  for all N_π mirror poses $\pi_j, (j = 1, \cdots, N_\pi)$ **do**
    Compute the normal vector $\boldsymbol{n}_j$ of the mirror plane π_j by Eq (41).
  end for
  Compute the each column of rotation matrix $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ and the translation vector by solving Eq (42).
  Refine $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ by solving the orthogonal Procrustes problem [8].
  Refine $R, \boldsymbol{t}$ by applying non-linear optimization [28]

(40) for one normal vector $\boldsymbol{n}_j$. By collecting these equations, we have

$$\begin{aligned} \boldsymbol{S}_j \boldsymbol{n}_j &= 0, \\ \boldsymbol{S}_j &= (\boldsymbol{m}_{j1}\boldsymbol{m}_{j2}\cdots\boldsymbol{m}_{jj-1}\boldsymbol{m}_{jj+1}\cdots\boldsymbol{m}_{jN_\pi})^\top \end{aligned} \tag{41}$$

where $\boldsymbol{S}_j$ is a $(N_\pi - 1)\times 3$ matrix. A normal vector $\boldsymbol{n}_j$ can be computed as the right-singular vector corresponding to the smallest singular value of $\boldsymbol{S}_j$.

This equation also indicates that we have to provide $N_\pi \geq 3$ mirror poses in order to estimate $\boldsymbol{n}_j$, because the degree of freedom of $\boldsymbol{n}_j$ is 2.

4.3.3 Computing Extrinsic Parameters

Up to this point, we obtain the 3D positions of mirrored reference points $\boldsymbol{p}^i_j (i = 1, \cdots, N_p, j = 1, \cdots, N_\pi)$ and mirror normals $\boldsymbol{n}_j$. The 3D positions of reference points $\boldsymbol{p}^{i\{W\}} = (x_i, y_i, z_i)$ are supposed to be given a priori in its local coordinate system X. By substituting these known parameters into Eq (37), we can derive a large system of linear equations:

$$A\boldsymbol{Z} = \boldsymbol{B}. \tag{42}$$

where

$$A = \begin{bmatrix} \boldsymbol{I}_3 & 2\boldsymbol{n}_1 & \boldsymbol{0}_{3\times1} & \cdots & \boldsymbol{0}_{3\times1} & x_1\boldsymbol{I}_3 & y_1\boldsymbol{I}_3 & z_1\boldsymbol{I}_3 \\ \boldsymbol{I}_3 & 2\boldsymbol{n}_1 & \boldsymbol{0}_{3\times1} & \cdots & \boldsymbol{0}_{3\times1} & x_2\boldsymbol{I}_3 & y_2\boldsymbol{I}_3 & z_2\boldsymbol{I}_3 \\ & & & \vdots & & & & \\ \boldsymbol{I}_3 & 2\boldsymbol{n}_1 & \boldsymbol{0}_{3\times1} & \cdots & \boldsymbol{0}_{3\times1} & x_{N_p}\boldsymbol{I}_3 & y_{N_p}\boldsymbol{I}_3 & z_{N_p}\boldsymbol{I}_3 \\ \boldsymbol{I}_3 & \boldsymbol{0}_{3\times1} & 2\boldsymbol{n}_2 & \cdots & \boldsymbol{0}_{3\times1} & x_1\boldsymbol{I}_3 & y_1\boldsymbol{I}_3 & z_1\boldsymbol{I}_3 \\ & & & \vdots & & & & \\ \boldsymbol{I}_3 & \boldsymbol{0}_{3\times1} & 2\boldsymbol{n}_2 & \cdots & \boldsymbol{0}_{3\times1} & x_{N_p}\boldsymbol{I}_3 & y_{N_p}\boldsymbol{I}_3 & z_{N_p}\boldsymbol{I}_3 \\ & & & \vdots & & & & \\ \boldsymbol{I}_3 & \boldsymbol{0}_{3\times1} & \boldsymbol{0}_{3\times1} & \cdots & 2\boldsymbol{n}_{N_\pi} & x_1\boldsymbol{I}_3 & y_1\boldsymbol{I}_3 & z_1\boldsymbol{I}_3 \\ & & & \vdots & & & & \\ \boldsymbol{I}_3 & \boldsymbol{0}_{3\times1} & \boldsymbol{0}_{3\times1} & \cdots & 2\boldsymbol{n}_{N_\pi} & x_{N_p}\boldsymbol{I}_3 & y_{N_p}\boldsymbol{I}_3 & z_{N_p}\boldsymbol{I}_3 \end{bmatrix}, \tag{43}$$

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{t}^\top & d_1 & d_2 & \cdots & d_{N_\pi} & \boldsymbol{r}_1^\top & \boldsymbol{r}_2^\top & \boldsymbol{r}_3^\top \end{bmatrix}^\top, \tag{44}$$

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_{N_\pi} \end{bmatrix}^\top, \tag{45}$$

$$\boldsymbol{B}_j = \begin{bmatrix} \boldsymbol{b}^1_j & \boldsymbol{b}^2_j & \cdots & \boldsymbol{b}^{N_p}_j \end{bmatrix}, \tag{46}$$

$$\boldsymbol{b}^i_j = (-2\boldsymbol{n}_j^\top \boldsymbol{p}^i_j \boldsymbol{n}_j + \boldsymbol{p}^i_j)^\top. \tag{47}$$

The vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ denote the first, second and third column of the rotation matrix R. From N_π mirror poses and N_p reference points, we have $12+N_\pi$ unknown parameters and $3\times N_\pi \times N_p$ equations. Hence, when $N_\pi \geq 3$ and $3\times N_\pi \times N_p > 12+N_\pi$, we can solve the Eq (42) by $\boldsymbol{Z} = \boldsymbol{A}^*\boldsymbol{B}$, where $\boldsymbol{A}^*$ is the pseudo-inverse matrix of $\boldsymbol{A}$.

In case of that reference points are on a single plane, the 3D position of reference points in its local coordinate system can be expressed as $\boldsymbol{p}^{i\{W\}} = (x_i, y_i, 0)^\top$ and we cannot compute the third column vector $\boldsymbol{r}_3$ of the rotation matrix. In this case, we compute $\boldsymbol{r}_3$ as the cross product of first and second column vector $\boldsymbol{r}_1, \boldsymbol{r}_2$, that is $\boldsymbol{r}_3 = \boldsymbol{r}_1 \times \boldsymbol{r}_2$.

4.3.4 Linear Refinement of Rotation Matrix by Solving the Orthogonal Procrustes Problem

Now we obtain columns of rotation matrix $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ and translation vector $\boldsymbol{t}$ linearly, but $\boldsymbol{r}_1, \boldsymbol{r}_2$ and $\boldsymbol{r}_3$ do not necessarily satisfy the following constraints as a rotation matrix due to noise:

$$\begin{aligned} |\ \boldsymbol{r}_1\ | = |\ \boldsymbol{r}_2\ | = |\ \boldsymbol{r}_3\ | &= 1, \\ \boldsymbol{r}_1^\top \boldsymbol{r}_2 = \boldsymbol{r}_2^\top \boldsymbol{r}_3 = \boldsymbol{r}_3^\top \boldsymbol{r}_1 &= 0. \end{aligned} \tag{48}$$

Here, we solve the orthogonal Procrustes problem [8] and obtain a rotation matrix which satisfies Eq (48) and is closest to the original linear solution as proposed in Zhang's method [31]. That is $R = \boldsymbol{U}\boldsymbol{V}^\top$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are given by as the SVD of the original matrix $(\boldsymbol{r}_1 \quad \boldsymbol{r}_2 \quad \boldsymbol{r}_3) = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$.

4.4 Non-Linear Refinement of Extrinsic Parameters

In general, obtained extrinsic parameters can be refined by non-linear optimization [28]. Here, we minimize following reprojection error function,

$$\boldsymbol{E}_{opt} = \Sigma_{j=1}^{N_\pi}\Sigma_{i=1}^{N_p} |\boldsymbol{q}^i_j - \breve{\boldsymbol{q}}^i_j(R, \boldsymbol{t}, \boldsymbol{n}_j, d_j)|, \tag{49}$$

where $\breve{\boldsymbol{q}}^i_j(R, \boldsymbol{t}, \boldsymbol{n}_j, d_j)$ denote the reprojected point calculated from estimated parameters. We solved this non-linear optimization problem of Eq (49) with Levenberg-Marquardt algorithm.

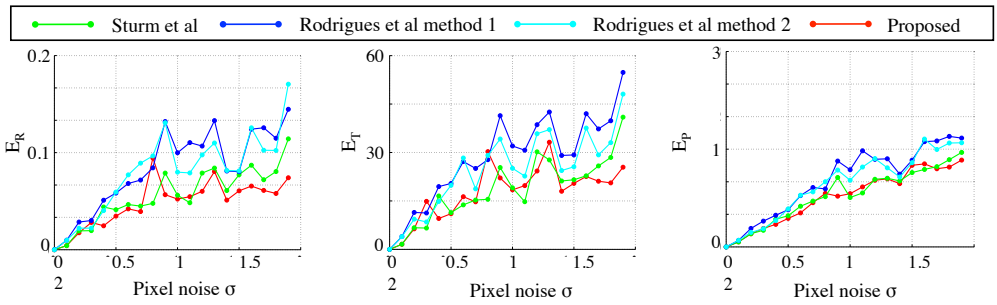


Fig. 19 Estimation error of each parameter in changing the standard deviation σ of pixel noise added to the input.

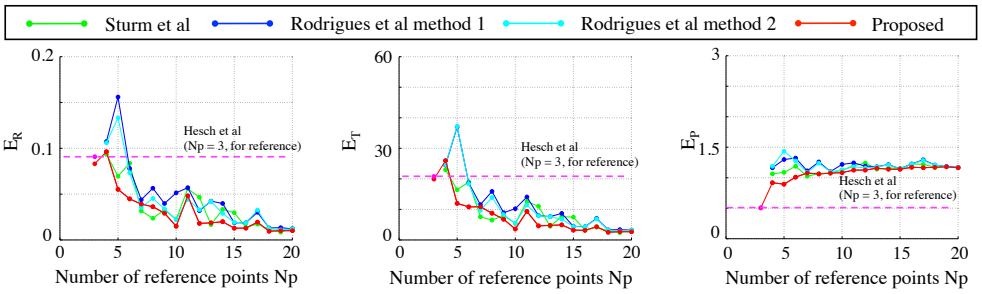


Fig. 20 Estimation error of each parameter in changing the number of reference points N_p.

4.5 Performance Evaluations

This section provides experimental evaluations using synthesized and real data in $N_p > 3$ case. In these evaluations, we compare our method with state-of-the-arts proposed by Sturm *et al*. [24] and by Rodrigues *et al*. [23] with non-linear refinement.

4.5.1 Quantitative Evaluations with Synthesized Data

Experiment Environment: To synthesize data, we used the following experiment setup by default. The matrix of intrinsic parameters, K, consists of (fx, fy, cx, cy); fx and fy represents the focal length in pixels, and cx and cy represent the 2D coordinates of the principle point. We set them to $(500, 500, 300, 250)$ in this evaluation respectively.

The normal vectors $\boldsymbol{n}_j (j = 1, \cdots, N_\pi)$ of mirror poses π_j are set to $(\sin\theta_z \sin\theta_x + \cos\theta_x \cos\theta_z \sin\theta_y, \sin\theta_x \cos\theta_z + \cos\theta_x \sin\theta_z \sin\theta_y, \cos\theta_x \cos\theta_y)$ where $\theta_k (k = x, y, z)$ is the angle respect to each axis, and drawn randomly within the ranges of $(-20 \leq \theta_x \leq 20, 160 \leq \theta_y \leq 200, -20 \leq \theta_z \leq 20)$. The distance between each mirror plane and camera center was set to 300 mm.

The reference object consists of N_p reference points forming a grid pattern and the distance between each reference point is 50mm. The center of X is located at the centroid of these points.

We represent the ground truth of rotation matrix as a product of three elemental rotation matrices, that is $R = R_1(\theta_1)R_2(\theta_2)R_3(\theta_3)$, and we set random values to each angles θ_1, θ_2 and θ_3 within $[-10 : 10]$ respectively. The position $\boldsymbol{t}$ is generated of each trial by assigning a random value within $[-5 : 5]$ to each x, y and z element of $\boldsymbol{t}$.

In this experiment, we evaluate the performance of each method under various conditions of the following parameters.

(a) σ: the standard deviation of Gaussian pixel noise of zero-mean.
(b) N_p: the number of reference points.
(c) N_π: the number of mirror poses.

Table 3 describes the min, max, increment step and default value of the parameters. We computed the average of the estimation errors of 100 trials for each of combinations. While changing these parameters respectively, the other parameters are set to values in Default column, that is the minimum setup for Sturm *et al*. [24] and Rodrigues *et al*. [23].

4.5.2 Error metrics

Throughout this evaluation, we used the following metrics to measure the performance of the calibration methods.

The estimation error of R is defined as the Riemannian distance [19]:

$$E_R = \frac{1}{N_\pi} \sum_{j=1}^{N_\pi} \frac{1}{\sqrt{2}} \| \mathrm{Log}(R_j^\top R_{g,j}) \|_F \tag{50}$$

$$\mathrm{Log}R' = \begin{cases} 0 & (\theta = 0), \\ \frac{\theta}{2\sin\theta}(R' - R'^\top) & (\theta \neq 0), \end{cases} \tag{51}$$

where $\theta = \cos^{-1}(\frac{\mathrm{Tr}(R')-1}{2})$.

The estimation error of $\boldsymbol{t}$ is defined as the root mean square error:

$$E_t = \frac{1}{N_\pi} \sum_{j=1}^{N_\pi} \sqrt{| \boldsymbol{t}_j - \boldsymbol{t}_{g,j} |^2 / 3}. \tag{52}$$

The estimation error of $\boldsymbol{n}$ is defined as the angle against the ground truth $\boldsymbol{n}_g$:

$$E_n = \frac{1}{N_\pi} \sum_{j=1}^{N_\pi} \arccos(\boldsymbol{n}_j^\top \boldsymbol{n}_{g,j}). \tag{53}$$

The reprojection error is defined as follows:

$$E_P = \frac{1}{N_\pi \times N_p} \sum_{j=1}^{N_\pi} \left(\sum_{i=1}^{N_p} (\breve{\boldsymbol{q}}_j^i - \boldsymbol{q}_j^i) \right), \tag{54}$$

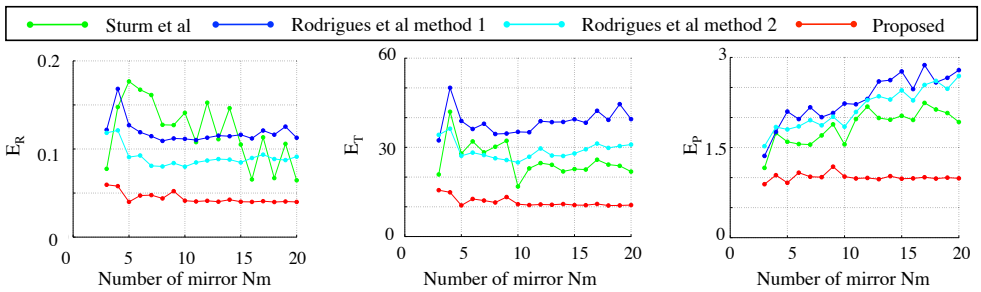


Fig. 21 Estimation error of each parameter in changing the number of mirror poses N_π.

Parameter	Min	Max	Step	Default
σ	0	2	0.1	1
N_p	4	20	1	4
N_π	3	20	1	3

Table 3 The range of changing parameters.

where $\boldsymbol{q}_j^i$ is the observation and $\breve{\boldsymbol{q}}_j^i$ is the reprojected point calculated from estimated parameters.

Results: Figure 19 shows results for different standard deviations σ of pixel noise added to the observations. The averages of each estimation error of proposed method are smaller than those of Sturm *et al*. [24] and Rodrigues *et al*. [23]. This fact indicates that our method can estimate better initial values from same inputs, because of the larger number of constraints involved in the estimation. Notice that there exist some trials in each of which all the methods result in a same optimal value regardless of the differences between initial values returned by their linear methods. Besides, there exists some trials where all the methods fall in local minima. This is the reason for the spikes.

Figure 20 shows results for difference number of reference points. We added the results of Hesch *et al*. [12] designed for $N_p = 3$ case as a reference. From these results, we can observe that the number of reference points affects the performance of each method drastically. These improvements are thought to be due to the improvement of estimation of mirrored reference point by PnP. In fact, Lepetit *et al*. [17] shows that the result of PnP in increasing number of points follows similar pattern of this experiments in their paper.

Figure 21 shows results in changing the number of mirror poses N_π. While Sturm's method [24] does not improve with increasing number of mirror pose, we can see that our method and Rodrigues's method [23] improves. This is considered to be due to the scalability of formulation for mirror pose, that is the number of equations for estimating extrinsic parameters by our method and [23] changes depends on the number of mirror poses.

These results prove that proposed method works robustly with observation noise and has the scalability for the number of reference points and mirror poses.

4.5.3 Qualitative Evaluations with Real Data

Experiment Environment: We evaluated the performance of our proposed method with real data assuming calibration of a display-camera system, such as digital-signage, laptop computer and so on. We utilized the same configuration introduced in **??** (Figure **??**). We used two cameras (Pointgrey Flea3) C_1 and C_2, a 20-inch flat panel display and a sputtering mirror. The goal

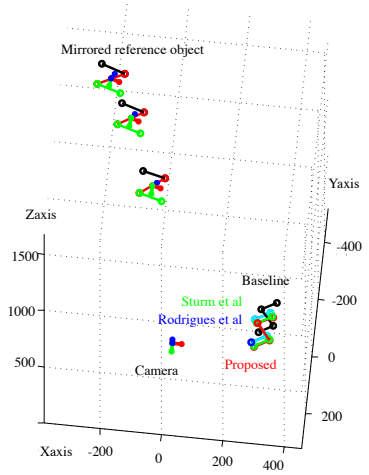


Fig. 22 Estimated positions of the reference object by the proposed method (red), by [23](blue and cyan), by [24](green) and by Eq (**??**) (black). Notice that this figure renders estimate positions of $\boldsymbol{p}$ and X_1 in $\{C\}$, and therefore C_1 is located at $(0,0,0)^\top$.

of this experiment is to calibrate the extrinsic parameters of C_1 against a 7×10 chess pattern X_1 rendered in the display. Notice that we used only N_p reference points for calibration. The length between each reference point is 82.5 mm. C_1 is located where it cannot observe X_1 directly. It captures N_π UXGA images of different mirror poses $\pi_j (j = 1, \cdots, N_\pi)$ for calibration.

Results: Figure 22 renders the estimated positions of the reference object by each method with $N_p = 4$ and $N_\pi = 3$ configuration. We can see that the reference objects estimated by each method are located near the baseline result. This precision is acceptable for applications using display-camera system such as gaze detection for dibgital-signage or gaze correction [16] in a video conference scenario.

Figure 23 and 24 shows results in changing the number of reference points N_p and the number of mirror poses N_π respectively. Notice that the estimated parameters by each method are almost identical, and therefore we observe only one line in Figure 24.

From these results, we can observe that our method performs better than conventional methods [12, 23, 24] in real situation qualitatively and quantitatively.

4.6 Discussion

4.6.1 Degenerate Case

Our algorithm does not work if it cannot compute enough axis vectors $\boldsymbol{m}_{jj'}$ for estimating mirror normals. This happens in the following three cases. (1) If two mirrors are parallel, then the in-

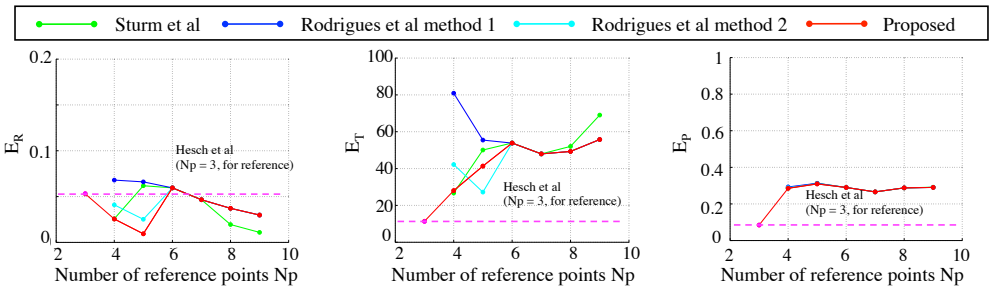


Fig. 23 Estimation error of each parameter with real data in changing the number of reference points N_p.

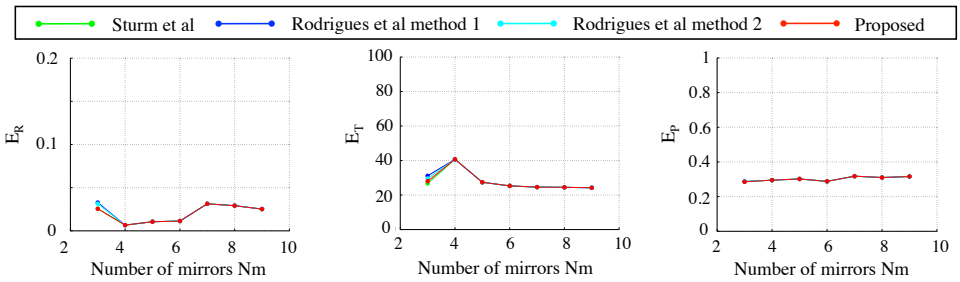


Fig. 24 Estimation error of each parameter with real data in changing the number of mirror poses N_π.

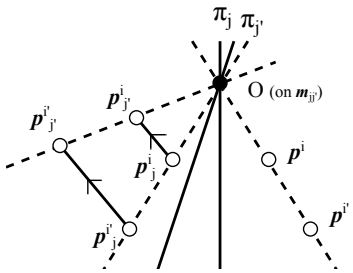


Fig. 25 Degenerate case

tersection of them does not exist and therefore not be computable. (2) If all the mirror planes intersect at single axis in 3D, the mirror normals cannot be computable by solving Eq (41). These (1) and (2) cases has been originally observed by Sturm *et al.* [24]. (3) If reference points and the intersection of two mirrors π_j and $\pi_{j'}$ are on a same plane, the axis vector $\boldsymbol{m}_{jj'}$ is not be computable by solving Eq (39) though $\boldsymbol{m}_{jj'}$ does exist physically because of the following reason which makes the two rows of $M'_{jj'}$ corresponding to π_j and $\pi_{j'}$ be linearly dependent.

Proposition 1. *If two reference points $\boldsymbol{p}^i$ and $\boldsymbol{p}^{i'}$ and the intersection of two mirrors π_j and $\pi_{j'}$ are on a same plane, the two lines connecting results of different Householder transformations of $\boldsymbol{p}^i$ and $\boldsymbol{p}^{i'}$, i.e. , the lines connecting $\boldsymbol{p}^i_j$ to $\boldsymbol{p}^i_{j'}$ and $\boldsymbol{p}^{i'}_j$ to $\boldsymbol{p}^{i'}_{j'}$, are parallel (Figure 25).*

Proof. Suppose the line connecting $\boldsymbol{p}^i$ and $\boldsymbol{p}^{i'}$ intersects with the intersection of the two mirrors at $\boldsymbol{O}$ as shown in Figure 25. By definition of the reflection, the distance from $\boldsymbol{p}^i$ to $\boldsymbol{O}$ is equal to the one from $\boldsymbol{p}^i_j$ to $\boldsymbol{O}$. Similarly, the distance from $\boldsymbol{p}^{i'}$ to $\boldsymbol{O}$ is equal to the one from $\boldsymbol{p}^{i'}_j$ to $\boldsymbol{O}$. Also, these distances are equal to the ones from $\boldsymbol{p}^i_{j'}$ to $\boldsymbol{O}$, and from $\boldsymbol{p}^{i'}_{j'}$ to $\boldsymbol{O}$ respectively. Here $\triangle \boldsymbol{O}\boldsymbol{p}^i_j\boldsymbol{p}^i_{j'}$ and $\triangle \boldsymbol{O}\boldsymbol{p}^{i'}_j\boldsymbol{p}^{i'}_{j'}$ are isosceles triangles sharing the apex $\angle \boldsymbol{p}^i_j\boldsymbol{O}\boldsymbol{p}^i_{j'}$. Therefore, the two lines $\boldsymbol{p}^i_j$ to $\boldsymbol{p}^i_{j'}$ and $\boldsymbol{p}^{i'}_j$ to $\boldsymbol{p}^{i'}_{j'}$ are parallel. □

These three degenerate cases can be detected by observing the rank of $M_{jj'}$ in Eq (39). If the rank is less than 2, we can discard the mirror pair and try with more mirrored images in practice.

4.6.2 Sufficiency of the Orthogonality Constraint

The orthogonality constraint holds for two reflections of a single reference point and the axis vector, as a necessary condition. Obviously, this does not constrain the position of the mirror (the parameter d of Eq (37)). That is a mirror of another distance satisfies Eq (38) as long as it has the same intersection direction.

This fact indicates that the orthogonality constraint itself does not serve as the sufficient condition to determine all of the mirror parameters. Instead, given three mirrors, it becomes the sufficient condition to obtain the mirror normals uniquely as described in Sections 4.3. By using the ted normals, we can define the linear equations (Eq (42)) based on the measurement model (Eq (37)).

4.7 Summary

In this section, we proposed a new algorithm to extrinsically calibrate a camera to a 3D reference object that is not directly visible from the camera. We introduced an orthogonality constraint which should be satisfied by all families of reflections of a same reference object and established a mirrored-points-based formulation. This formulation allow us to obtain a larger number of constraints which contribute to make the calibration more robust even with a simple configuration, that is using fewer reference points and fewer mirror poses. The evaluations of the extrinsic calibration by synthesized and real data showed our improvement on the accuracy and robustness against state-of-the-arts quantitatively and qualitatively.

The proposed method utilized a single reflection of a reference object. However, in case of calibrating widely scattered cameras or an omnidirectional cameras, the reference object cannot be observable via single reflection. For such cases, the multiple reflections should be considered and the extrinsic camera calibration with them is the future work.

5. Extrinsic Camera Calibration using Human Cornea Reflections

This section provides a human cornea reflection based extrinsic calibration algorithm of a camera and a reference object located out of the camera's filed-of-view.

5.1 Measurement Model

As illustrated in Figure 26, we denote a center of the spherical mirror, its radius, and a reflection point of a reference point $\boldsymbol{p}_i$ on the cornea sphere by $\boldsymbol{S}$, r and $\boldsymbol{m}_i$ respectively. Let us assume a unit vector from $\boldsymbol{m}_i$ to $\boldsymbol{p}_i$ expressed as $\boldsymbol{u}_i$, $\boldsymbol{p}_i$ is expressed as $\boldsymbol{p}_i = k_i\boldsymbol{u}_i + \boldsymbol{m}_i$, where k_i is the distance between $\boldsymbol{m}_i$ and $\boldsymbol{p}_i$. This $\boldsymbol{p}_i$ also satisfies $\boldsymbol{p}_i = R\boldsymbol{p}_i^{\{X\}} + \boldsymbol{t}$, where R and $\boldsymbol{t}$ are the extrinsic paraemters between the camera and the refefence object. From these equations, we obtain the following equation:

$$R\boldsymbol{p}_i^{\{X\}} + \boldsymbol{t} = k_i\boldsymbol{u}_i + \boldsymbol{m}_i. \tag{55}$$

This Eq.(55) has 10 unknown parameters, that is R, $\boldsymbol{t}$, $\boldsymbol{S}$ and r, and it is defined as the measurement model in this configuration.

5.2 Approach: Solving Absolute Orientation Problem

Determining extrinsic parameters between two coordinate systems, such as $\{C\}$ and $\{X\}$, through the use of a set of corresponding points in each coordinate system is known as the *Absolute Orientation Problem*. By solving this problem, we can obtain the extrinsic parameters from at least three point correspondences. Since the 3D positions of reference points $\boldsymbol{p}_i^{\{X\}}$ in $\{X\}$ are supposed to be given a priori, we estimate the 3D positions of reference point $\boldsymbol{p}_i$ in $\{C\}$ by estimating $k_{m_i p_i}$, $\boldsymbol{S}$ and r. In order to estimate them, we introduce the following two constraints, a geometric model of the cornea sphere and the equidistance constraint.

5.2.1 Cornea Sphere Parameters Estimation Based on its Geometric Model

Based on the geometric model [20], the radius of the cornea is recognize as the average of it based on [22], and the limbus projection is modeled as an ellipse represented by five parameters: the center, $\boldsymbol{i}_L$, the major and minor radii, r_{max} and r_{min}, respectively, and rotation angle ϕ. Since the depth of a tilted limbus is much smaller than the distance between camera and the cornea sphere, we assume weakly perspective projection. Under this assumption, the 3D position of the center of limbus $\boldsymbol{L}$ is expressed as $\boldsymbol{L} = dK^{-1}\boldsymbol{i}_L$, where d denotes the distance between the center of camera, $\boldsymbol{O}$, and the center of limbus $\boldsymbol{L}$, and is expressed as $d = f \cdot r_L/r_{max}$. f and K represent the focal length in pixels and intrinsic parameters, respectively. Gaze direction $\boldsymbol{g}$ is approximated by the optical axis of the eye, and is theoretically determined by $\boldsymbol{g} = [\sin\tau\sin\phi, -\sin\tau\cos\phi, -\cos\tau]^\top$, where $\tau = \pm\arccos(r_{min}/r_{max})$; τ corresponds to the tilt of the limbus plane with respect to the image plane. Since the center of cornea sphere, $\boldsymbol{S}$, is located at distance d_{LS} $(= \sqrt{r^2 - r_L^2} = \sqrt{7.7^2 - 5.6^2} \approx 5.3$ mm), the radius of the cornea sphere from $\boldsymbol{L}$, we compute $\boldsymbol{S}$ as $\boldsymbol{S} = \boldsymbol{L} - d_{LS}\boldsymbol{g}$. In this way, we estimate $\boldsymbol{S}$ from the ellipse parameters of the limbus projected onto the image plane.

5.2.2 Equidistance Constraint

To obtain k_i, we introduce the *Equidistance Constraint*. The

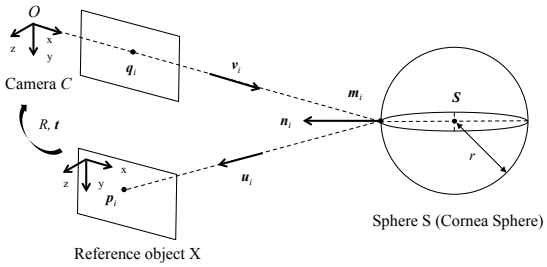


Fig. 26 Reflection model of spherical mirror.

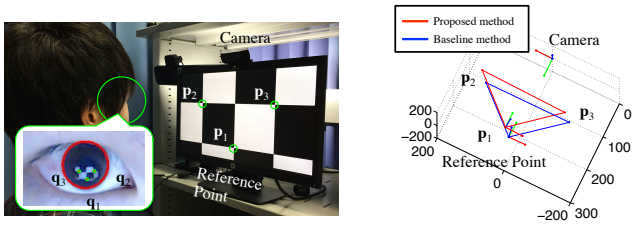


Fig. 27 Left: Configuration for experiments with real data, Right: positions of reference point computed from estimated extrinsic parameters.

Equidistance Constraint states that the distance from reference point $\boldsymbol{p}_i$ to the center of cornea sphere S is equal to the distance from the center of the camera, $\boldsymbol{O}$, to $\boldsymbol{S}$. If this equidistance constraint is satisfied, when user sets his/her center of cornea sphere such that the equidistance constraint does hold, triangle $\triangle\boldsymbol{O}\boldsymbol{m}_i\boldsymbol{p}_i$ should be an isosceles triangle that satisfies $|\boldsymbol{p}_i - \boldsymbol{m}_i| = |\boldsymbol{O} - \boldsymbol{m}_i|$, that is $k_i = k_i'$ and this k_i' is a known parameter.

By introducing above constraints, we compute 3D positions of reference point $\boldsymbol{p}_i$ and obtain extrinsic parameters R and $\boldsymbol{t}$ by solving the Absolute Orientation Problem. This algorithm works with the minimal configuration where three reference points and one mirror pose.

5.3 Performance Evaluations

Figure **??** (a) illustrates the geometric relation of camera and reference points. The camera is set at the top of display which has three reference points. Figure **??** (b) renders the estimated positions of the reference points. It is difficult to obtain the ground truth of extrinsic parameters in any real configuration, so we used the baseline method as the reference parameters. From this result, we can see that the positions estimated by the proposed method are almost identical to those of baseline. Notice that the difference in these rotation matrices for x-axis, y-axis and z-axis are 2.44, 6.88, and 0.49 degrees, respectively ($E_R = 0.1278$), and E_t is 27.4989 mm in Figure **??**.

6. Summary

In this chapter, we proposed a new algorithm that calibrates a camera to a 3D reference object via cornea reflection with minimal configuration. The key features of our method are its introduction of two constraints: cornea reflection model and equidistance constraint. In evaluations, our method outperformed a state-of-the-art of plane mirror based method with both synthesized and real data.

The proposed method introduced in this chapter mainly focuses on the geometric relationship between the 3D reference

point and its 2D projection. While the human eye reflects ray and can be utilized as a spherical mirror, the human eye also reflects a cue of mental condition of humans, such as the gaze direction [13]. These cues are sometimes available for analyzing the relationship of each parts of a imaging system. The fusion of reflections of a ray and a mental condition is one of the future works.

References

[1] Agrawal, A.: Extrinsic Camera Calibration without a Direct View Using Spherical Mirror, *Proceedings of International Conference on Computer Vision* (2013).

[2] Basu, A. and Licardie, S.: Alternative models for fish-eye lenses, *Pattern recognition letters*, Vol. 16, No. 4, pp. 433–441 (1995).

[3] Chum, O. and Matas, J.: Matching with PROSAC - progressive sample consensus, *Proceedings of IEEE Computer Vision and Pattern Recognition*, Vol. 1, pp. 220–226 (2005).

[4] DeSouza, G. N. and Kak, A. C.: Vision for mobile robot navigation: a survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 2, pp. 237 –267 (2002).

[5] Fitzgibbon, A. W., Fisher, R. B. et al.: A buyer's guide to conic fitting, *DAI Research paper* (1996).

[6] Funk, N. and Yang, Y.-H.: Using a raster display for photometric stereo, *Computer and Robot Vision, 2007. CRV'07. Fourth Canadian Conference on*, IEEE, pp. 201–207 (2007).

[7] Furukawa, Y. and Ponce, J.: Accurate Camera Calibration from Multi-View Stereo and Bundle Adjustment, *International Journal of Computer Vision*, Vol. 84, No. 3, pp. 257–268 (2009).

[8] Golub, G. H. and Van Loan, C. F.: *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, third edition (1996).

[9] Grossberg, M. D. and Nayar, S. K.: The raxel imaging model and ray-based calibration, *International Journal of Computer Vision*, Vol. 61, No. 2, pp. 119–137 (2005).

[10] Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press (2000).

[11] Healey, G. E. and Kondepudy, R.: Radiometric CCD camera calibration and noise estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 3, pp. 267–276 (1994).

[12] Hesch, J. A., Mourikis, A. I. and Roumeliotis, S. I.: Mirror-based extrinsic camera calibration, *Algorithmic Foundation of Robotics VIII*, Springer, pp. 285–299 (2009).

[13] Hirayama, T., Dodane, J.-B., Kawashima, H. and Matsuyama, T.: Estimates of user interest using timing structures between proactive content-display updates and eye movements, *IEICE TRANSACTIONS on Information and Systems*, Vol. 93, No. 6, pp. 1470–1478 (2010).

[14] Iyamu, E. and Osuobeni, E.: Age, gender, corneal diameter, corneal curvature and central corneal thickness in Nigerians with normal intra ocular pressure, *Journal of optometry*, Vol. 5, No. 2, pp. 87–97 (2012).

[15] Kannala, J. and Brandt, S. S.: A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 8, pp. 1335–1340 (2006).

[16] Kuster, C., Popa, T., Bazin, J.-C., Gotsman, C. and Gross, M.: Gaze Correction for Home Video Conferencing, *ACM Transactionson Graphics (In Proceedings of ACM SIGGRAPH ASIA)*, Vol. 31, No. 6, p. to appear (2012).

[17] Lepetit, V., Moreno-Noguer, F. and Fua, P.: Epnp: An accurate o (n) solution to the pnp problem, *International Journal of Computer Vision*, Vol. 81, No. 2, pp. 155–166 (2009).

[18] Mitsunaga, T. and Nayar, S. K.: Radiometric self calibration, *Proceedings of IEEE Computer Vision and Pattern Recognition*, Vol. 1, IEEE (1999).

[19] Moakher, M.: Means and Averaging in the Group of Rotations, *SIAM J. Matrix Anal. Appl.*, Vol. 24, pp. 1–16 (2002).

[20] Nakazawa, A. and Nitschke, C.: Point of gaze estimation through corneal surface reflection in an active illumination environment, *Proceedings of European Conference on Computer Vision* (2012).

[21] Nishimura, M., Nobuhara, S., Matsuyama, T., Shimizu, S. and Fujii, K.: A Linear Generalized Camera Calibration from Three Intersecting Reference Planes, *Proceedings of International Conference on Computer Vision*, pp. 2354–2362 (2015).

[22] Richard S.Snell, M. A. L.: *Clinical Anatomy of the Eye*, Wiley-Blackwell, second edition (1997).

[23] Rodrigues, R., Barreto, P. and Nunes, U.: Camera Pose Estimation Using Images of Planar Mirror Reflections, *Proceedings of European Conference on Computer Vision*, pp. 382–395 (2010).

[24] Sturm, P. and Bonfort, T.: How to Compute the Pose of an Object without a Direct View, *Proceedings of Asian Conference on Computer Vision*, pp. 21–31 (2006).

[25] Swaminathan, R. and Nayar, S. K.: Nonmetric calibration of wide-angle lenses and polycameras, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1172–1178 (2000).

[26] Takahashi, K.: Camera Calibration Based on Mirror Reflections, PhD Thesis, Kyoto University (2018).

[27] Tomaževič, D., Likar, B. and Pernuš, F.: Comparative evaluation of retrospective shading correction methods, *Journal of Microscopy*, Vol. 208, No. 3, pp. 212–223 (2002).

[28] Triggs, B., McLauchlan, P., Hartley, R. and Fitzgibbon, A.: Bundle Adjustment — A Modern Synthesis, *Vision Algorithms: Theory and Practice* (Triggs, B., Zisserman, A. and Szeliski, R., eds.), Lecture Notes in Computer Science, Vol. 1883, Springer Berlin Heidelberg, pp. 298–372 (2000).

[29] Tsai, R. Y.: An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, *Proceedings of IEEE Computer Vision and Pattern Recognition*, Vol. 1, IEEE (1986).

[30] Ying, X., Peng, K., Hou, Y., Guan, S., Kong, J. and Zha, H.: Self-Calibration of Catadioptric Camera with Two Planar Mirrors from Silhouettes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 5, pp. 1206–1220 (2013).

[31] Zhang, Z.: A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330–1334 (2000).