Ray-Space Projection Model for Light Field Camera*

Qi Zhang[1], Jinbo Ling[1], Qing Wang[1], and Jingyi Yu[2]

[1]School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, P.R. China,
qwang@nwpu.edu.cn

[2]ShanghaiTech University, Shanghai 200031, P.R. China

Abstract

Light field essentially represents the collection of rays in space. The rays captured by multiple light field cameras form subsets of full rays in 3D space and can be transformed to each other. However, most previous approaches model the projection from an arbitrary point in 3D space to corresponding pixel on the sensor. There are few models on describing the ray sampling and transformation among multiple light field cameras. In the paper, we propose a novel ray-space projection model to transform sets of rays captured by multiple light field cameras in term of the Plücker coordinates. We first derive a 6×6 *ray-space intrinsic matrix based on multi-projection-center (MPC) model. A homogeneous ray-space projection matrix and a fundamental matrix are then proposed to establish ray-ray correspondences among multiple light fields. Finally, based on the ray-space projection matrix, a novel camera calibration method is proposed to verify the proposed model. A linear constraint and a ray-ray cost function are established for linear initial solution and non-linear optimization respectively. Experimental results on both synthetic and real light field data have verified the effectiveness and robustness of the proposed model.*

1. Introduction

Light field cameras [18, 23] can record spatial and angular information of light rays in 3D space. Based on angular sampling of light rays, sophisticated post-processing techniques [20, 27, 13, 26, 34, 35, 33] ranging from digital refocusing to depth estimation have been introduced in decades. However, the major disadvantages of such handheld systems are the spatio-angular trade-off and narrow baseline. Applications of registration [15, 31, 32] and light field stitching [1, 8, 24, 7] are proposed to overcome these limitations. In general, the performance of these applications can be enhanced by accurate geometric information of multiple light field cameras. But, there are fewer works on establishing a generalized model for defining the ray sampling and transformation among multiple light fields.

Existing light field camera models [3, 30] mostly define the projection from an arbitrary point in 3D space (passing through micro-lens) to corresponding pixel on the sensor. Nevertheless, light field essentially represents the collection of rays in space. The Plücker line coordinates explicitly provide a homogeneous parameterization for rays to effectively formulate ray-ray correspondence, which has already verified in generalized epipolar geometry [22]. Dansereau *et al.* [4] describe pixel-ray correspondences in 3D space and present a 4D intrinsic matrix. However, they only focus on the transformation of rays in monocular light field camera. Meanwhile, their model has redundancy and dependency, which results in irregular ray sampling and makes parameterization impossible in term of the Plücker line coordinates. In order to explore the ray sampling and transformation among multiple light field cameras, multi-projection-center model (MPC) [30] which provides independent and effective intrinsic parameters for the Plücker representation is used.

In the paper, we exploit the transformation of light fields captured by multiple cameras and propose a novel ray-space projection model in term of the Plücker coordinates. We first propose a so-called *ray-space intrinsic matrix* (RSIM) to relate the recorded indices to corresponding physical ray. We show that the RSIM is a 6×6 matrix consisted of 6-parameter intrinsics, which is analogous to traditional camera matrix. Furthermore, we derive a homogeneous ray-space projection matrix and a fundamental matrix using the RSIM and ray-space extrinsic matrix to describe ray-ray correspondences. Secondly, we propose a novel light field camera calibration method to certify the proposed model. According to the ray-sapce projection matrix, a linear constraint between 3D point and the rays captured by light field

*The work was supported by NSFC under Grant 61531014.

cameras is established. Subsequently, an effective linear initial solution for both intrinsic and extrinsic parameters is computed for optimization. We also define an effective ray-ray cost function to minimize the distance among rays in the Plücker coordinates. Finally, we exhibit empirical performances in calibrating synthetic light field cameras and commercial light field cameras (Lytro and Lytro Illum [18]). Quantitative and qualitative comparisons demonstrate the effectiveness and robustness of the proposed ray-space projection model.

Our main contributions are:

1) The ray-space projection model and fundamental matrix among multiple light field cameras are exploited.

2) A 6×6 RSIM is deduced to relate the recorded pixel to physical ray in the Plücker coordinates.

3) An effective calibration algorithm is proposed to verify the proposed model, including a linear constraint for initial solution and a novel ray-ray cost function for optimization.

2. Related Work

Light field model. Since the hand-held light field camera is put forwarded by Ng [21], many research groups [4, 14, 2, 9, 28, 3, 30, 29] have extensively explored various projection models for light field cameras. Johannsen *et al.* [14] exhibit a model for reconstructing 3D points from the parallax in adjacent micro-lens images. Heinze *et al.* [9] present a similar model with [14] and deduce a linear initialization for the focused light field camera.

Dansereau *et al.* [4] propose a 12-free-parameter light field camera model, corresponding the recorded pixels to the rays outside the camera. They derive a 4D decoding matrix from a conventional pinhole lenslet and thin-lens model. Nevertheless, the calibration method is initialized by a traditional camera calibration algorithm which is not effective for light field camera. More importantly, the 4D intrinsic matrix has redundancy and dependency, which results in irregular rays sampling during the calibration and rectification. Different from the calibration based on corner features of sub-aperture images, Bok *i.e.* [2, 3] utilize line features which are directly extracted from micro-lens images of raw data to calibrate light field camera. They formulate a 6-parameter geometric projection model for light field camera to estimate initial solution of intrinsic parameters. But, this method confronts a significant challenge to obtain line features accurately (in practice, the checkerboard should be shot under an unfocused status in order to make the measurements detectable).

More recently, Zhang *et al.* [28] exhibit a 4-parameter projective distortion model to estimate the parameters using a parallel bi-planar board. However, these parameters are insufficient to model the light field camera geometry. Zhang *et al.* [30] propose a 6-parameter multi-projection-center (MPC) model for light field cameras, including traditional and focused light field cameras. A 3D projective transformation is deduced to describe the relationship between geometric structure and light field camera coordinates. Then, a light field camera calibration method is presented to verify the effectiveness of MPC model.

Generalized epipolar geometry. The generalized epipolar geometry is the intrinsic projective geometry in the ray-space. It only depends on essential parameters of ray and relative pose instead of scene structure. Techniques for estimating camera motion and structure from multiple images have been improved and perfected over the decades [11]. Grossberg and Nayar [6] define the image pixel as the light from a cone around a ray and propose a generalized camera model for calibration. For the geometric analysis of multiple images, Pless [22] simplifies this calibration so that it only includes the definition of ray that the pixel samples. They then propose a general linear framework to describe any cameras as an unordered collection of rays which are obtained by sensor elements. The correspondences between rays need to be established with the assumption that the rays intersect at a single scene point. Then, the generalized epipolar constraint is proposed in the Plücker coordinates. Sturm [25] introduces a hierarchy of general camera model. In this framework, 17 corresponding rays are sufficient to solve linearly for pose estimation. Li *et al.* [17] carry out a pose estimation based on the generalized epipolar constraint. This also can be applied to estimate the motion of light field camera. Guo *et al.* [8] propose a ray-space motion matrix to establish ray-ray correspondences for motion estimation.

Moreover, all above mentioned methods first reconstruct the 3D points, and then formulate the Plücker coordinates based on the 3D points. In other words, the performance of these techniques are influenced by accurate reconstruction. Johannsen *et al.* [15] directly generate the Plücker coordinates based on light field captured by camera. A linear mathematical framework is built from the relationships between scene geometry and light rays for motion estimation. However, little attention has been paid to model the sampling and transformation of rays captured by multiple light field cameras in the Plücker coordinates. In the work, we explore the relationship between a ray in 3D space and the corresponding ray in light field camera and propose a ray-space projection model for light field camera.

3. Ray-Space Projection Model

3.1. The MPC Model

Light field cameras, especially micro-lens array assembled inside, which are innovated from traditional 2D camera, record 3D world in different but similar ways. With the shifted views, light field camera maps 3D space to many

sub-aperture images, which produces 4D light field. The ray in 4D light field is parameterized in a *relative* two-parallel-plane coordinates [16], where $Z = 0$ denotes the view plane and $Z = f$ for the image plane. In this parametrization, the decoded physical ray is described by $\boldsymbol{r} = (s, t, x, y)$ in term of specific (*e.g.*, meter) dimension. The ray $\boldsymbol{r}$ intersects with the view plane at projection center $(s, t, 0)$. The pair (x, y) is the intersection of the ray $\boldsymbol{r}$ with the image plane, but it is relative to (s, t, f) which is the origin of image plane. The (x, y, f) describes the direction of ray. Then, according to the MPC model [30], a 3D point $\boldsymbol{X} = (X, Y, Z)^\top$ is mapped to the pixel (x, y) in the image plane,

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & -fs \\ 0 & f & 0 & -ft \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{1}$$

This is analogous to classical projective camera model with projection center at $(s, t, 0)$ and principal axis parallelling to the Z-axis.

The light field $L(i, j, u, v)$ captured by a light field camera is transformed into a normalized undistorted physical light field $L(s, t, x, y)$ by a homogeneous decoding matrix $\boldsymbol{D} \in \mathbb{R}^{5\times5}$ [30],

$$\begin{bmatrix} s \\ t \\ x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} k_i & 0 & 0 & 0 & 0 \\ 0 & k_j & 0 & 0 & 0 \\ 0 & 0 & k_u & 0 & u_0 \\ 0 & 0 & 0 & k_v & v_0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \\ u \\ v \\ 1 \end{bmatrix}, \tag{2}$$

where $(k_i, k_j, k_u, k_v, u_0, v_0)$ are intrinsic parameters of a light field camera. (k_i, k_j) are scale factors for s and t axes in the view plane and (k_u, k_v) for x and y axes in the image plane respectively. In addition, $(-u_0/k_u, -v_0/k_v)$ represent the coordinates of principal point in the sub-aperture image.

3.2. Ray-Space Intrinsic Matrix

According to the MPC model, light field camera is assumed as a pinhole camera array. In this framework, a light field camera is described by the set of rays sampled by a collection of perspective cameras. In order to simplify the discussion of geometric analysis in multiple light fields, the pixel captured by the camera is generalized and simplified to a ray [22, 21]. The light field essentially represents a set of rays. Consequently, we need a new mechanism to describe arbitrary rays in free space. The Plücker coordinates provide convenience to mathematically formulate concise and efficient correspondence equations (*e.g.*, rotation and translation). In addition, the Plücker coordinates representation is also a homogeneous parameterization to unambiguously represent a ray in 3D projective geometry. We will briefly review the core theory leading the equations for ray-space projection model.

In the Plücker coordinates, the ray is mathematically represented by a pair of vectors $(\boldsymbol{m}, \boldsymbol{q})$, named moment and direction vector respectively. Moreover, the moment vector $\boldsymbol{m} = \boldsymbol{X} \times \boldsymbol{q}$, for an arbitrary point $\boldsymbol{X}$ on the ray. Further, the physical ray $\boldsymbol{r} = (s, t, x, y)$ in 3D space contains directional information (x, y) and positional information (s, t). Therefore, the moment vector and direction vector of $\boldsymbol{r}$ are defined as,

$$\begin{cases} \boldsymbol{m} = (s, t, 0)^\top \times (x, y, 1)^\top = (t, -s, sy - tx)^\top \\ \boldsymbol{q} = (x, y, 1)^\top \end{cases}, \tag{3}$$

where $(\boldsymbol{m}^\top, \boldsymbol{q}^\top)^\top$ are the Plücker line coordinates.

Substituting Eq. (2) into Eq. (3), there is a transformation caused by the intrinsic parameters $(k_i, k_j, k_u, k_v, u_0, v_0)$. Then the RSIM $\boldsymbol{K} \in \mathbb{R}^{6\times6}$ is established to describe the relationship between the ray $\mathcal{L} = (\boldsymbol{n}^\top, \boldsymbol{p}^\top)^\top$ captured by light field camera and the normalized undistorted physical ray $\mathcal{L}_c = (\boldsymbol{m}^\top, \boldsymbol{q}^\top)^\top$ in the Plücker coordinates, *i.e.*,

$$\begin{bmatrix} \boldsymbol{m} \\ \boldsymbol{q} \end{bmatrix} = \underbrace{\begin{bmatrix} k_j & 0 & 0 & 0 & 0 & 0 \\ 0 & k_i & 0 & 0 & 0 & 0 \\ -k_j u_0 & -k_i v_0 & k_i k_v & 0 & 0 & 0 \\ 0 & 0 & 0 & k_u & 0 & u_0 \\ 0 & 0 & 0 & 0 & k_v & v_0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{=:\boldsymbol{K}} \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{p} \end{bmatrix}, \tag{4}$$

which needs to satisfy the condition $k_u/k_v = k_i/k_j$. (u, v) are pixel coordinates extracted from sub-aperture image at the view (i, j). Then, $\boldsymbol{p} = (u, v, 1)^\top$ represents the direction of ray in the sub-aperture image coordinates. $\boldsymbol{n} = (i, j, 0)^\top \times (u, v, 1)^\top$ denotes the moment of ray.

3.3. Ray-Space Projection Matrix

In general, considering $\boldsymbol{X}_w$ is a point in the world coordinates, the transformation between the world and camera coordinates is described by a rotation matrix $\boldsymbol{R} \in SO(3)$ and a translation vector $\boldsymbol{t} = (t_x, t_y, t_z)^\top \in \mathbb{R}^3$, formulated as $\boldsymbol{X} = \boldsymbol{R}\boldsymbol{X}_w + \boldsymbol{t}$. Consequently, the Plücker transformation can be formulated according to generalized epipolar geometry [22]

$$\mathcal{L}_w = \begin{bmatrix} \boldsymbol{R}^\top & \boldsymbol{E}^\top \\ \boldsymbol{0}_{3\times3} & \boldsymbol{R}^\top \end{bmatrix} \mathcal{L}_c, \tag{5}$$

where $\boldsymbol{E} = [\boldsymbol{t}]_\times \boldsymbol{R}$ is the essential matrix and $[\cdot]_\times$ refers to the vector cross product [11]. The rays $\mathcal{L}_c = (\boldsymbol{m}^\top, \boldsymbol{q}^\top)^\top$ and $\mathcal{L}_w = (\boldsymbol{m}_w^\top, \boldsymbol{q}_w^\top)^\top$ are expressed by the camera coordinates and world coordinates respectively. Subsequently, according to Eqs. (3) and (4), the homogeneous ray-space projection matrix $\boldsymbol{P}$ can be written as,

$$\mathcal{L}_w = \underbrace{\begin{bmatrix} \boldsymbol{R}^\top & \boldsymbol{E}^\top \\ \boldsymbol{0}_{3\times3} & \boldsymbol{R}^\top \end{bmatrix} \boldsymbol{K}}_{=:\boldsymbol{P}} \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{p} \end{bmatrix}, \tag{6}$$

which implies the relationship between $\mathcal{L} = (\boldsymbol{n}^\top, \boldsymbol{p}^\top)^\top$ in the camera coordinates and $\mathcal{L}_w$ in the world coordinates.

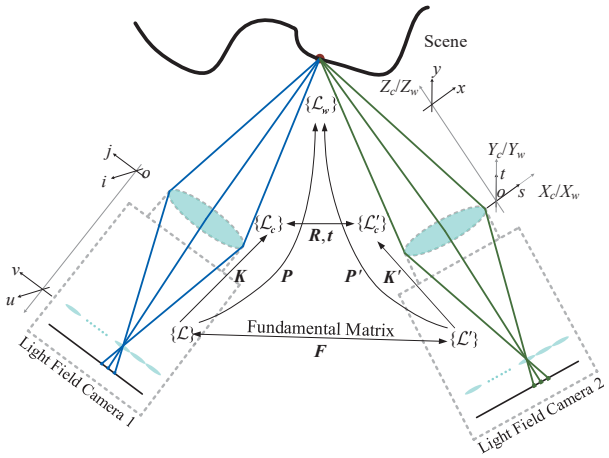


Figure 1. Ray-space projection model and ray-ray transformation among two light field cameras.

3.4. Fundamental Matrix

In order to establish the ray-ray transformation among two light field cameras, the camera coordinates of the second light field camera are assumed as the world coordinates, as shown in Fig. 1. Based on the ray-space projection model Eq. (6), the fundamental matrix $\boldsymbol{F}$ between two light field cameras is then derived,

$$\mathcal{L}'^{\top} \underbrace{\boldsymbol{K}'^{\top} \begin{bmatrix} \boldsymbol{0}_{3\times3} & \boldsymbol{R}^{\top} \\ \boldsymbol{R}^{\top} & \boldsymbol{E}^{\top} \end{bmatrix} \boldsymbol{K}}_{=:\boldsymbol{F}} \mathcal{L} = 0, \tag{7}$$

which represents the ray-ray correspondence $\{\mathcal{L}'\} \leftrightarrow \{\mathcal{L}\}$. $\boldsymbol{R}$, $\boldsymbol{t}$ denote the rotation and translation between two light field cameras' coordinates. For a valid correspondence, all rays in both light fields must come from the same scene point, as shown in Fig. 1. The detailed derivation of Eq. (7) is presented in the supplemental material.

In order to verify the effectiveness of ray-space projection model, we propose a light field camera calibration algorithm.

4. Light Field Camera Calibration

4.1. Constraint of Ray-Space Projection Matrix

In 3D projective geometry, a point $\boldsymbol{X}_w$ in the world coordinates can be described as the intersection of $\mathcal{L}_w = (\boldsymbol{m}_w^{\top}, \boldsymbol{q}_w^{\top})^{\top}$ with the plane $Z = Z_w$. The plane is expressed by a homogeneous vector $(\boldsymbol{\pi}^{\top}, d)^{\top}$, $\boldsymbol{\pi} \in \mathbb{R}^3$, $d \in \mathbb{R}$. Therefore, according to the theorem $\boldsymbol{X} = (\boldsymbol{\pi} \times \boldsymbol{m} - d\,\boldsymbol{q}) / \boldsymbol{\pi}^{\top}\boldsymbol{q}$ [12], we find the constraint between $\boldsymbol{X}_w$ and $\mathcal{L}_w$ in the Plücker coordinates,

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & Z_w & -Y_w \\ 0 & 1 & 0 & -Z_w & 0 & X_w \end{bmatrix}}_{=:\boldsymbol{M}(\boldsymbol{X}_w)} \begin{bmatrix} \boldsymbol{m}_w \\ \boldsymbol{q}_w \end{bmatrix} = 0. \tag{8}$$

where $\boldsymbol{M}$ is the measurement matrix. The detailed derivation of Eq. (8) is presented in the supplemental material.

According to the ray-space projection model of Eq. (6), the relationship among the RSIM $\boldsymbol{K}$, extrinsic parameters $[\boldsymbol{R}|\boldsymbol{t}]$, and the points $\boldsymbol{X}_w$ in the world coordinates is extended by Eqs. (4), (5) and (8),

$$\boldsymbol{M}(\boldsymbol{X}_w) \begin{bmatrix} \boldsymbol{R}^{\top} & \boldsymbol{E}^{\top} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{R}^{\top} \end{bmatrix} \boldsymbol{K} \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{p} \end{bmatrix} = 0. \tag{9}$$

In summary, the solution space of Eq. (9) is the ray set intersecting at a point $\boldsymbol{X}_w$. The set of ray sampled by the light field camera is a subspace of the whole-solution space of Eq. (9).

4.2. Initialization

Without a loss of generality, there is an assumption that the checkerboard is on the plane $Z_w = 0$ in the world coordinates, which leads to a simplified form of Eq. (9),

$$\underbrace{\begin{bmatrix} 1 & 0 & -Y_w \\ 0 & 1 & X_w \end{bmatrix}}_{=:\boldsymbol{M}_s} \otimes \begin{bmatrix} \boldsymbol{n}^{\top} & \boldsymbol{p}^{\top} \end{bmatrix} \vec{\boldsymbol{H}}_s = 0, \tag{10}$$

where $\vec{\boldsymbol{H}}_s$ is an 18×1 matrix stretched on row from the simplified homogeneous ray-space projection matrix $\boldsymbol{H}_s$. $\boldsymbol{M}_s$ is the simplified measurement matrix of checkerboard corners $\boldsymbol{X}_w$. $\otimes$ is a direct product operator. The RSIM is abbreviated to a lower triangle matrix $\boldsymbol{K}_{ij}$ and the upper triangle matrix $\boldsymbol{K}_{uv}$. Subsequently, $\boldsymbol{H}_s$ denotes a 3×6 matrix only using intrinsic and extrinsic parameters,

$$\boldsymbol{H}_s = \begin{bmatrix} \boldsymbol{r}_1^{\top} & -\boldsymbol{r}_1^{\top}[\boldsymbol{t}]_{\times} \\ \boldsymbol{r}_2^{\top} & -\boldsymbol{r}_2^{\top}[\boldsymbol{t}]_{\times} \\ \boldsymbol{0}_{1\times3} & \boldsymbol{r}_3^{\top} \end{bmatrix} \begin{bmatrix} \boldsymbol{K}_{ij} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{K}_{uv} \end{bmatrix}, \tag{11}$$

where $\boldsymbol{r}_i$ is the i-th column vector of rotation matrix $\boldsymbol{R}$.

In order to derive intrinsic parameters, we abbreviate with $[\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{0}_{1\times3}]^{\top}$ the first three and with $[\boldsymbol{h}_3, \boldsymbol{h}_4, \boldsymbol{h}_5]^{\top}$ the second three columns of $\boldsymbol{H}_s$ respectively. $\boldsymbol{h}_i$ denotes the row vector (h_{i1}, h_{i2}, h_{i3}). Utilizing the orthogonality of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, we have

$$\begin{aligned} \boldsymbol{h}_1 \boldsymbol{K}_{ij}^{-1} \boldsymbol{K}_{ij}^{-\top} \boldsymbol{h}_2^{\top} &= 0 \\ \boldsymbol{h}_1 \boldsymbol{K}_{ij}^{-1} \boldsymbol{K}_{ij}^{-\top} \boldsymbol{h}_1^{\top} &= \boldsymbol{h}_2 \boldsymbol{K}_{ij}^{-1} \boldsymbol{K}_{ij}^{-\top} \boldsymbol{h}_2^{\top}, \end{aligned} \tag{12}$$

where $\boldsymbol{K}_{ij}^{-1} = \begin{bmatrix} 1/k_j & 0 & 0 \\ 0 & 1/k_i & 0 \\ u_0/k_j k_u & v_0/k_i k_v & 1/k_i k_v \end{bmatrix}$.

Let a symmetric matrix $\boldsymbol{B}$ denote $\boldsymbol{K}_{ij}^{-1} \boldsymbol{K}_{ij}^{-\top}$,

$$\boldsymbol{B} = \begin{bmatrix} 1/k_j^2 & 0 & u_0/k_u k_j^2 \\ 0 & 1/k_i^2 & v_0/k_v k_i^2 \\ u_0/k_u k_j^2 & v_0/k_v k_i^2 & (u_0^2 + v_0^2 + 1)/k_i^2 k_j^2 \end{bmatrix}. \tag{13}$$

Note that $\boldsymbol{B}$ has only 5 distinct non-zero elements, expressed by $\boldsymbol{b} = (b_{11}, b_{13}, b_{22}, b_{23}, b_{33})^\top$. In order to get the solution of $\boldsymbol{B}$, Eq. (12) is rewritten as $\boldsymbol{V}\boldsymbol{b} = \boldsymbol{0}$, *i.e.*,

$$\begin{bmatrix} h_{11}h_{21} & h_{11}^2 - h_{21}^2 \\ h_{11}h_{23}+h_{13}h_{21} & 2(h_{11}h_{13}-h_{21}h_{23}) \\ h_{12}h_{22} & h_{12}^2 - h_{22}^2 \\ h_{12}h_{23}+h_{13}h_{22} & 2(h_{12}h_{13}-h_{22}h_{23}) \\ h_{13}h_{23} & h_{13}^2 - h_{23}^2 \end{bmatrix}^\top \begin{bmatrix} b_{11} \\ b_{13} \\ b_{22} \\ b_{23} \\ b_{33} \end{bmatrix} = \boldsymbol{0}, \quad (14)$$

where $\boldsymbol{V}$ is a $2n \times 5$ matrix. What is more, a general non-zero solution $\boldsymbol{b}$ is computed, only if there are at least two such equations (from two positions) as Eq. (14). $\boldsymbol{b}$ is determined up to an unknown scale factor.

Once $\boldsymbol{b}$ is estimated, it is easy to solve $\boldsymbol{K}_{ij}^{-1}$ by Cholesky factorization [10]. The effect of the scale factor is eliminated by calculating the ratio of elements. Therefore the intrinsic parameters except (k_i, k_j) are obtained,

$$\begin{aligned} k_u &= \hat{k}_{11}/\hat{k}_{33}, \quad k_v = \hat{k}_{22}/\hat{k}_{33}, \\ u_0 &= \hat{k}_{31}/\hat{k}_{33}, \quad v_0 = \hat{k}_{32}/\hat{k}_{33}, \end{aligned} \quad (15)$$

where $\hat{k}_{mn}$ is the m-th row and n-th column of the estimated intrinsic matrix $\hat{\boldsymbol{K}}_{ij}^{-1}$. Furthermore, the rest intrinsic parameters and extrinsic parameters of different poses can be obtained as follows,

$$\begin{aligned} &\lambda = \frac{1}{2}\left(\left\|\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_1^\top\right\| + \left\|\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_2^\top\right\|\right), \\ &\tau = 1/\left\|\hat{\boldsymbol{K}}_{uv}^{-\top}\boldsymbol{h}_5^\top\right\|, \\ &\boldsymbol{r}_1 = \frac{\alpha}{\lambda}\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_1^\top,\ \boldsymbol{r}_2 = \frac{\alpha}{\lambda}\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_2^\top,\ \boldsymbol{r}_3 = \boldsymbol{r}_1 \times \boldsymbol{r}_2, \\ &\boldsymbol{t} = (\boldsymbol{G}^\top\boldsymbol{G})^{-1}(\boldsymbol{G}^\top\boldsymbol{g}), \\ &\boldsymbol{G} = (-[\boldsymbol{r}_1]_\times, -[\boldsymbol{r}_2]_\times)^\top,\ \boldsymbol{g} = (\tau\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_3^\top, \tau\hat{\boldsymbol{K}}_{ij}^{-\top}\boldsymbol{h}_4^\top)^\top, \\ &\alpha = \begin{cases} 1 & t_z > 0 \\ -1 & t_z < 0 \end{cases}, \\ &k_i = \frac{\lambda\tau}{\hat{k}_{22}},\ k_j = \frac{\lambda\tau}{\hat{k}_{11}}, \end{aligned} \quad (16)$$

where $\|\cdot\|$ denotes L_2 norm, $\hat{\boldsymbol{K}}_{uv}$ is formed by the intrinsic parameters which are obtained from Eq. (15). α is determined by t_z because it must be positive (*i.e.* the checkerboard is put in front of light field camera). The details of derivation are given in the supplemental material.

4.3. Distortion Model

Due to the special imaging design of two-parallel-plane in light field camera, there exists a radial distortion on the image plane and a sampling distortion on the view plane simultaneously. With the assumption that the angular sampling is ideal without distortion, only radial distortion is considered in the paper. The undistorted coordinates $(\widetilde{x}, \widetilde{y})$ is rectified by the distorted coordinates (x, y) under the view (s, t),

$$\begin{cases} \widetilde{x} = x + (k_1 r_{xy}^2 + k_2 r_{xy}^4)(x - b_1) + k_3 s \\ \widetilde{y} = y + (k_1 r_{xy}^2 + k_2 r_{xy}^4)(y - b_2) + k_4 t \end{cases}, \quad (17)$$

where $r_{xy}^2 = (x - b_1)^2 + (y - b_2)^2$ and $\boldsymbol{r} = (s, t, x, y)$ is transformed from the measurement $\boldsymbol{l} = (i, j, u, v)$ by the intrinsic parameters $\mathcal{P}$ according to Eq. (2). Compared with existing radial distortion of light field camera, we add k_3 and k_4 to represent the distortion affected by the shifted view. k_1, k_2, b_1, b_2 regulate conventional radial distortion on the image plane. In summary, we use $\boldsymbol{k}^d = (k_1, k_2, k_3, k_4, b_1, b_2)$ to denote distortion vector.

4.4. Nonlinear Optimization

The initial solution computed by the linear method is refined via nonlinear optimization. Instead of minimizing the distance between checkerboard corners and rays [4] and the re-projection error in traditional multiview geometry [11], we define a ray-ray cost function to acquire the nonlinear solution. The ray-ray cost function is the minimization of the distance between the rays and the line on the checkerboard,

$$d_x = \frac{|\boldsymbol{m}^\top\boldsymbol{q}_x + \boldsymbol{q}^\top\boldsymbol{m}_x|}{\|\boldsymbol{q}\times\boldsymbol{q}_x\|},\ d_y = \frac{|\boldsymbol{m}^\top\boldsymbol{q}_y + \boldsymbol{q}^\top\boldsymbol{m}_y|}{\|\boldsymbol{q}\times\boldsymbol{q}_y\|}, \quad (18)$$

where $\mathcal{L} = (\boldsymbol{m}^\top, \boldsymbol{q}^\top)^\top$ is the ray in the Plücker coordinates. $\mathcal{L}_x = (\boldsymbol{m}_x^\top, \boldsymbol{q}_x^\top)^\top$ and $\mathcal{L}_y = (\boldsymbol{m}_y^\top, \boldsymbol{q}_y^\top)^\top$ are the lines parallel to the X_w-axis and the Y_w-axis respectively. Both lines cross the checkerboard corners $\boldsymbol{X}_w$. The derivation of the distance between two lines is given in the supplemental material. Further, we formulate the following cost function according Eq. (18),

$$\sum_{p=1}^{\#pose}\sum_{n=1}^{\#point}\sum_{i=1}^{\#view} \|d(\widetilde{\mathcal{L}}'_w(\mathcal{P}, \boldsymbol{k}^d, \boldsymbol{R}_p, \boldsymbol{t}_p), \mathcal{L}_w(\boldsymbol{X}_w))\|, \quad (19)$$

where $\widetilde{\mathcal{L}}'_w$ is expressed as the Plücker coordinates after computing by Eq. (2), followed by the distortion according to Eq. (17). $\mathcal{L}'_w$ denotes the line in Eq. (18). $\mathcal{P}$ represents intrinsic parameters, $\boldsymbol{k}^d$ is distortion vector and $\boldsymbol{R}_p$, $\boldsymbol{t}_p$ are extrinsic parameter at each position, $1 \le p \le P$.

Eq. (19) are nonlinear objective functions which can be solved using Levenberg-Marquardt algorithm based on the trust region reflective method [19]. In addition, $\boldsymbol{R}$ is parameterized by Rodrigues formula [5]. MATLAB's lsqnonlin function is utilized to implement the optimization.

5. Experiments

In this section, we evaluate the performance of our algorithm on calibrating synthetic light field cameras as well as commercial light field cameras. We compare the proposed method in ray re-projection error and re-projection

Algorithm 1 Light Field Camera Calibration.

Input: Checkerboard corners $\boldsymbol{X}_w$,
Corresponding rays $\boldsymbol{l}=(i,j,u,v)$.
Output: Intrinsic parameter $\mathcal{P}=(k_i,k_j,k_u,k_v,u_0,v_0)$,
Distortion vector $\boldsymbol{k}^d=(k_1,k_2,k_3,k_4,b_1,b_2)$,
Extrinsic parameters $\boldsymbol{R}_p,\boldsymbol{t}_p$, $(1\leqslant p\leqslant P)$.

for $p=1$ to P **do**
  $\boldsymbol{H}_s = \mathit{EstimateProjectionMatrix}(\boldsymbol{X}_w,\boldsymbol{l})$ ▷ Eq. (10)
end for
$\boldsymbol{B} = \mathit{EstimateMatrix}(\boldsymbol{H}_s)$ ▷ Eqs. (13), (14)
$(k_u,k_v,u_0,v_0) = \mathit{CalculateK}_{uv}(\boldsymbol{B})$ ▷ Eq. (15)
for $p=1$ to P **do**
  $(\boldsymbol{R}_p,\boldsymbol{t}_p) = \mathit{CalculateRT}(\boldsymbol{H}_s,k_u,k_v,u_0,v_0)$ ▷ Eq. (16)
end for
$(k_i,k_j) = \mathit{CalculateK}_{ij}(\hat{\boldsymbol{K}}_{ij})$ ▷ Eq. (16)
$\mathit{Optimization}(\mathcal{P},\boldsymbol{k}^d,\bigcup_{p=1}^{P}(\boldsymbol{R}_p,\boldsymbol{t}_p))$ ▷ Eq. (19)

error with state-of-the-arts, including DPW by Dansereau *et al.* [4], BJW by Bok *et al.* [3] and MPC by Zhang *et al.* [30].

5.1. Simulated Data

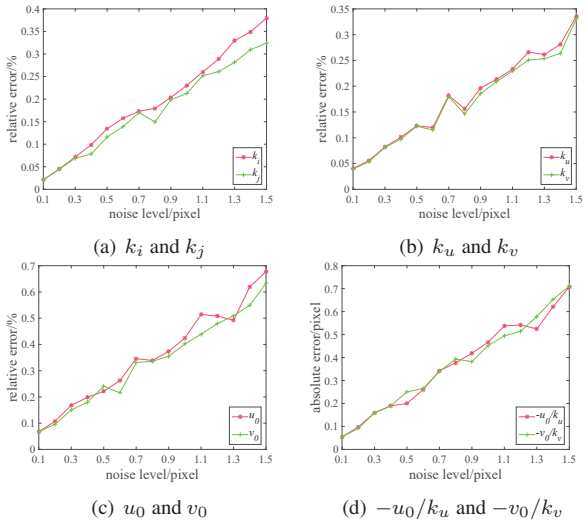
In order to evaluate the performance of our algorithm, we simulate a light field camera, whose intrinsic parameters are listed in Tab. 1. These parameters are close to the setting of a Lytro camera so that we obtain plausible input close to real-world scenarios. The checkerboard is a pattern with a 12×12 grid of $3.51mm$ cells.

Table 1. Intrinsic parameters of the simulated light field camera.

k_i	k_j	k_u	k_v	u_0	v_0
2.4e-4	2.5e-4	2.0e-3	1.9e-3	-0.32	-0.33

Performance w.r.t. the noise level. In this experiment, we employ the measurements of 3 poses and 7×7 views to verify the robustness of calibration algorithm. The rotation angles of 3 poses are $(6°,28°,-8°)$, $(12°,-10°,15°)$ and $(-5°,5°,-27°)$ respectively. Gaussian noise with zero mean and a standard deviation σ is added to the projected image points. We vary σ from 0.1 to 1.5 pixels with a 0.1 pixel step. For each noise level, we perform 150 independent trials. The estimated intrinsic parameters are evaluated by the average of relative errors with ground truth. As shown in Fig. 2, the errors almost linearly increase with noise level. For $\sigma=0.5$ pixels which is larger than normal noise in practical calibration, the errors of (k_i,k_j) and (k_u,k_v) are less than 0.14%. Although the relative error of (u_0,v_0) is 0.25%, the absolute error of $(-u_0/k_u,-v_0/k_v)$ is less than 0.24 pixel (In Eq. (2), $u=(x-u_0)/k_u$ and $v=(y-v_0)/k_v$, where $(-u_0/k_u,-v_0/k_v)$ is the principal point of a sub-aperture image), which demonstrates the robustness of the proposed method to high noise level.

Performance w.r.t. the number of poses and views. This experiment investigates the performance with respect



(a) k_i and k_j (b) k_u and k_v (c) u_0 and v_0 (d) $-u_0/k_u$ and $-v_0/k_v$

Figure 2. Performance evaluation of intrinsic parameters on the simulated data with different levels of noise σ.

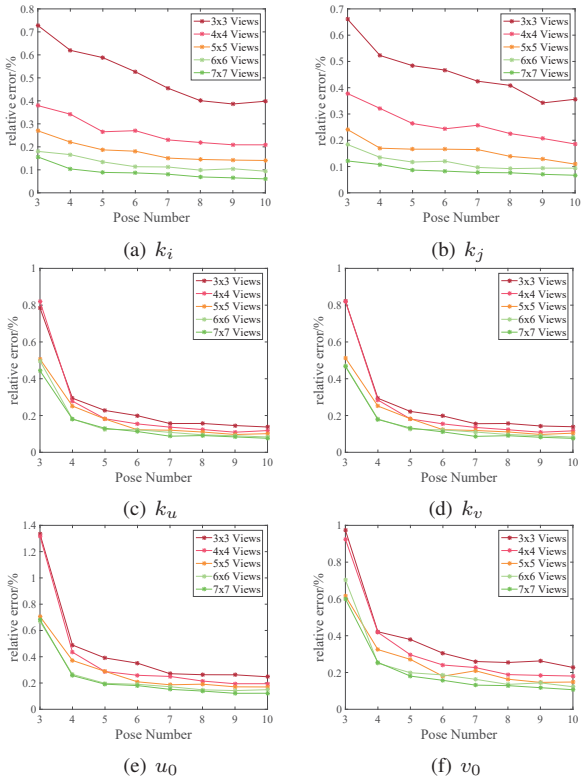
to the number of poses and views. We vary the number of poses from 3 to 10 and the number of views from 3×3 to 7×7. For each combination of pose and view, by adding a Gaussian noise with zero mean and a standard deviation of 0.5 pixel, 200 trails with independent checkerboard poses are conducted. The rotation angles are randomly generated from $-30°$ to $30°$. The average relative errors of calibration results with increasing measurements are shown in Fig. 3. The relative errors decrease with the number of poses. Meanwhile, when the number of poses is fixed, the errors reduce with the number of views. In particular, when $\#pose\geq4$ and $\#view\geq4\times4$, all relative errors are less than 0.5%, which further exhibits the effectiveness of the proposed calibration method.

5.2. Real Data

We also perform experiments on real scene light fields, including light field datasets (Lytro) released by DPW and light field datasets[1] (Lytro and Illum) released by MPC.

The sub-aperture images are easy to be decoded by raw data. We improve the preprocessing of raw data described in [4] to obtain sub-aperture images. First, normalized cross-correlation (NCC) of the white images is used to locate the centers of micro-lens images and estimate the average size of micro-lens images. In addition, due to a slight rotation between micro-lens array and the image sensor, a line fitting is utilized to estimate the rotation and refine the misalignment of micro-lens array. It is easy to extract 4D light field from 2D raw data once the centers and size of micro-lens images are obtained. The preprocess of raw data begins with demosaicing after alignment of micro-lens array. Then, the vignetting raw data is refine by divided by

[1] http://www.npu-cvpg.org/opensource



(a) k_i (b) k_j (c) k_u (d) k_v (e) u_0 (f) v_0

Figure 3. Relative errors of intrinsic parameters on simulated data with different numbers of poses and views.

Table 2. RMS ray re-projection errors (unit: mm).

	A	B	C	D	E
DPW [4]	0.0835	0.0628	**0.1060**	0.1050	0.3630
MPC [30]	0.0810	0.0572	0.1123	0.1046	0.5390
Ours	**0.0705**	**0.0438**	0.1199	**0.0740**	**0.2907**

white image. Finally, a resampling is utilize to extract sub-aperture images.

We firstly conduct calibration on the datasets collected with DPW [4]. For a fair comparison, the middle 7×7 sub-apertures are utilized. Tab. 2 summarizes the root mean square (RMS) ray re-projection error. Compared with DPW which employs 12 intrinsic parameters, the proposed ray-space projection model only employs a half of parameters but achieves smaller ray re-projection error except on dataset C. Considering the fact that the errors exhibited in DPW are minimized in its own optimization (*i.e.*, ray re-projection error), we additionally evaluate the performance in mean re-projection error with DPW and BJW. As exhibited in Tab. 3, the errors of the proposed method are obviously smaller than those of DPW and BJW, which further verifies the effectiveness of nonlinear optimization (*i.e.* the cost function in Eq. (19)).

Unlike the core idea of DPW, BJW conducts the calibration on the raw data directly instead of sub-aperture images.

Table 3. Mean re-projection errors (unit: *pixel*). The errors of DPW are calculated by their least released code. The errors of BJW are obtained from their extended work [3].

	A	B	C	D	E
DPW [4]	0.2284	0.1582	0.1948	0.1674	0.3360
BJW [3]	0.3736	0.2589	-	-	0.2742
MPC [30]	0.2200	0.1568	0.1752	0.1475	0.2731
Ours	**0.1843**	**0.1245**	**0.1678**	**0.1069**	**0.1383**

Table 4. RMS ray re-projection errors of optimizations without and with distortion rectification (unit: mm). The datasets are captured with Lytro and Illum cameras.

		Illum-1	Illum-2	Lytro-1	Lytro-2
Optimized without Rectification	DPW [4]	0.5909	0.4866	0.1711	0.1287
	BJW [3]	-	-	-	-
	MPC [30]	0.5654	0.4139	0.1703	0.1316
	Ours	**0.5641**	**0.4132**	**0.1572**	**0.1237**
Optimized with Rectification	DPW [4]	0.2461	0.2497	0.1459	0.1228
	BJW [3]	0.3966	0.3199	0.4411	0.2673
	MPC [30]	0.1404	0.0936	0.1400	0.1124
	Ours	**0.1294**	**0.0837**	**0.1142**	**0.0980**

It poses a significant challenge to obtain line feature accurately which is extracted from raw data to estimate an initial solution of intrinsic parameters. The light field data for calibration must be out of focus to make the measurement detectable. Therefore, as shown in Tab. 3, several datasets, *i.e.* C and D by [4], can not be estimated by BJW.

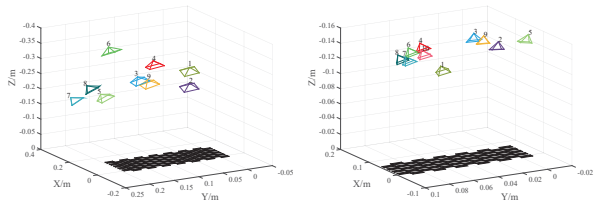
In order to comprehensively compare with DPW, BJW and MPC, we also carry out calibration on the datasets captured by MPC [30]. Tab. 4 lists the RMS ray re-projection errors compared with DPW, BJW and MPC at two calibration stages. As exhibited in Tab. 4, the proposed method obtains smaller ray re-projection errors on the item of optimization without rectification compared with DPW and MPC. Furthermore, it is more important we achieve small errors once the distortion is introduced in the optimization. According to the item of optimization with rectification, the proposed method outperforms DPW, BJW and MPC. Consequently, such optimization results substantiate that our 6-parameter ray-space projection model is effective to describe light field camera.

Fig. 4 demonstrates the results of pose estimation on datasets of MPC. In addition, we carry out light field stitching to examine the accuracy of intrinsic and extrinsic parameters estimation. 4-th and 7-th poses are center poses of datasets Illum-1 and Illum-2 respectively (see Fig. 4(a-b)). They are regarded as the reference views for stitching. Fig. 5 illustrates stitching results of light fields (actually these light fields are shot at a calibration board), from which we can see all light fields are registered and stitched very well.

It is vital to accurately calibrate light field camera and reconstruct 3D geometry. In order to verify the effective-

Table 5. Quantitative comparison of different calibration methods (unit: mm). The relative error is indicated in parentheses.

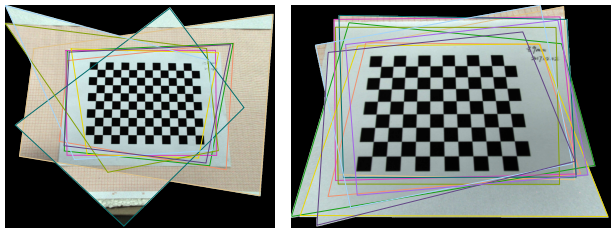
	'C'	'V'	'P'	'R'	'2'	'0'	'1'	'9'
Ruler	128.0	97.5	99.0	78.5	147.5	102.0	84.5	158.0
DPW [4]	124.3 (2.89%)	100.2 (2.77%)	97.6 (1.41%)	80.2 (2.17%)	145.7 (1.22%)	106.6 (4.51%)	86.1 (1.89%)	157.4 (0.69%)
BJW [3]	120.6 (5.78%)	106.8 (9.54%)	101.5 (2.53%)	80.7 (2.80%)	151.9 (2.98%)	103.4 (**1.37**%)	86.1 (1.89%)	160.6 (1.20%)
MPC [30]	127.5 (**0.39**%)	97.0 (**0.51**%)	98.3 (0.71%)	77.4 (1.40%)	144.9 (1.76%)	103.6 (1.57%)	85.7 (1.42%)	159.3 (0.82%)
Ours	127.4 (0.47%)	97.0 (**0.51**%)	98.9 (**0.10**%)	77.7 (**1.02**%)	145.9 (**1.08**%)	103.6 (1.57%)	85.3 (**0.95**%)	158.6 (**0.38**%)



(a) Illum-1 (b) Illum-2

(c) Lytro-1 (d) Lytro-2

Figure 4. Pose estimation results of our datasets.



(a) Illum-1 (b) Illum-2

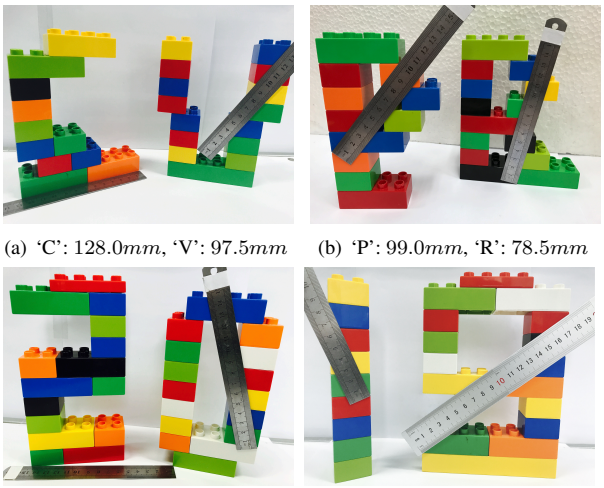
Figure 5. The stitching results of Illum-1 and Illum-2 datasets.

ness of geometric reconstruction of the proposed method compared with state-of-the-art methods, we capture four real scene light fields and reconstruct several specific corner points and estimate the distances between them. As shown in Tab. 5, the estimated distances between the reconstructed points are nearly equal to those measured lengths from real objects by rulers (see Fig. 6). For these four measurement examples, the relative errors of distance between reconstructed points demonstrate the performance of the proposed model compared with state-of-the-art methods.

6. Conclusion

In the paper, we explore the linear projection relationship between the rays in 3D space and the corresponding rays captured by multiple light field cameras. We first deduce a 6×6 RSIM composed of 6-parameter under the Plücker coordinates. A linear ray-space projection matrix and a fundamental matrix are then formulated to describe the ray sampling and transformation among multiple view light fields. Consequently, a linear constraint between 3D point and rays captured by light field camera is established to actuate an effective initial solution for both intrinsic and extrinsic parameters. Finally, a novel ray-ray cost function is established to nonlinearly optimize the 12-parameter model (6 for intrinsics and 6 for distortion). Extensive experiments are conducted on synthetic and real light field data. Qualitative and quantitative comparisons verify the effectiveness and robustness of the proposed ray-space projection model.

In the future, we intend to focus on exploring the invariance for multiple light field cameras based on the ray-space projection model. The future work also includes improving the effectiveness of initial solution and reconstructing large-scale complex scenes using multiple light field cameras.



(a) 'C': $128.0mm$, 'V': $97.5mm$ (b) 'P': $99.0mm$, 'R': $78.5mm$

(c) '2': $147.5mm$, '0': $102.0mm$ (d) '1': $84.5mm$, '9': $158.0mm$

Figure 6. Measurements between specific points by rulers.

Acknowledgement

Qi Zhang is also sponsored by Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University under CX201919. The authors thank Dr. Xue Wang for helpful discussions and Dahe Ma for the decoding program of light field.

References

[1] Clemens Birklbauer and Oliver Bimber. Panorama light-field imaging. In *Computer Graphics Forum*, volume 33, pages 43–52. Wiley Online Library, 2014.

[2] Yunsu Bok, Hae-Gon Jeon, and In So Kweon. Geometric calibration of micro-lens-based light-field cameras using line features. In *ECCV*, pages 47–61, 2014.

[3] Yunsu Bok, Hae-Gon Jeon, and In So Kweon. Geometric calibration of micro-lens-based light field cameras using line features. *IEEE T-PAMI*, 39(2):287–300, 2017.

[4] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *IEEE CVPR*, pages 1027–1034, 2013.

[5] Olivier Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 1993.

[6] Michael D Grossberg and Shree K Nayar. A general imaging model and a method for finding its parameters. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 108–115. IEEE, 2001.

[7] Mangtang Guo, Hao Zhu, Guoqing Zhou, and Qing Wang. Dense light field reconstruction from sparse sampling using residual network. In *ACCV*, 2018.

[8] Xinqing Guo, Zhan Yu, Sing Bing Kang, Haiting Lin, and Jingyi Yu. Enhancing light fields through ray-space stitching. *IEEE T-VCG*, 22(7):1852–1861, 2016.

[9] Christopher Hahne, Amar Aggoun, Shyqyri Haxha, Vladan Velisavljevic, and Juan Carlos Jácome Fernández. Light field geometry of a standard plenoptic camera. *Optics Express*, 22(22):26659–26673, 2014.

[10] Richard Hartley. Self-calibration of stationary cameras. *IJCV*, 22(1):5–23, 1997.

[11] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003.

[12] William Vallance Douglas Hodge, WVD Hodge, and Daniel Pedoe. *Methods of algebraic geometry*, volume 2. Cambridge University Press, 1994.

[13] Hae-Gon Jeon, Jaesik Park, Gyeongmin Choe, Jinsun Park, Yunsu Bok, Yu-Wing Tai, and In So Kweon. Accurate depth map estimation from a lenslet light field camera. In *IEEE CVPR*, pages 1547–1555, 2015.

[14] Ole Johannsen, Christian Heinze, Bastian Goldluecke, and Christian Perwaß. On the calibration of focused plenoptic cameras. In *Time-of-Flight and Depth Imaging*, pages 302–317. 2013.

[15] Ole Johannsen, Antonin Sulc, and Bastian Goldluecke. On linear structure from motion for light field cameras. In *IEEE ICCV*, pages 720–728, 2015.

[16] Marc Levoy and Pat Hanrahan. Light field rendering. In *ACM SIGGRAPH*, pages 31–42, 1996.

[17] Hongdong Li, Richard Hartley, and Jae-hak Kim. A linear approach to motion estimation using generalized camera models. In *IEEE CVPR*, pages 1–8, 2008.

[18] Lytro. Lytro redefines photography with light field cameras. `http://www.lytro.com`, 2011.

[19] Kaj Madsen, Hans Bruun Nielsen, and Ole Tingleff. *Methods for non-linear least squares problems, 2nd Edition*. Informatics and Mathematical Modelling, Technical University of Denmark, 2004.

[20] Ren Ng. Fourier slice photography. *ACM TOG*, 24(3):735–744, 2005.

[21] Ren Ng. *Digital light field photography*. PhD thesis, Stanford University, 2006.

[22] Robert Pless. Using many cameras as one. In *IEEE CVPR*, volume 2, pages 587–593, 2003.

[23] Raytrix. 3d light field camera technology. `http://www.raytrix.de`, 2013.

[24] Zhao Ren, Qi Zhang, Hao Zhu, and Qing Wang. Extending the FOV from disparity and color consistencies in multiview light fields. In *Proc. ICIP*, pages 1157–1161, 2017.

[25] Peter Sturm. Multi-view geometry for general camera models. In *IEEE CVPR*, volume 1, pages 206–212, 2005.

[26] Ting-Chun Wang, Alexei A Efros, and Ravi Ramamoorthi. Depth estimation with occlusion modeling using light-field cameras. *IEEE T-PAMI*, 38(11):2170–2181, 2016.

[27] Sven Wanner and Bastian Goldluecke. Globally consistent depth labeling of 4d light fields. In *IEEE CVPR*, pages 41–48, 2012.

[28] Chunping Zhang, Zhe Ji, and Qing Wang. Rectifying projective distortion in 4d light field. In *IEEE ICIP*, 2016.

[29] Qi Zhang and Qing Wang. Common self-polar triangle of concentric conics for light field camera calibration. In *ACCV*, 2018.

[30] Qi Zhang, Chunping Zhang, Jinbo Ling, Qing Wang, and Jingyi Yu. A generic multi-projection-center model and calibration method for light field cameras. *IEEE T-PAMI*, 2018.

[31] Yingliang Zhang, Zhong Li, Wei Yang, Peihong Yu, Halting Lin, and Jingyi Yu. The light field 3d scanner. In *IEEE ICCP*, pages 1–9, 2017.

[32] Yingliang Zhang, Peihong Yu, Wei Yang, Yuanxi Ma, and Jingyi Yu. Ray space features for plenoptic structure-from-motion. In *IEEE ICCV*, pages 4631–4639, 2017.

[33] Hao Zhu, Xiaoming Sun, Qi Zhang, Qing Wang, Antonio Robles-Kelly, Hongdong Li, and Shaodi You. Full view optical flow estimation leveraged from light field superpixel. *IEEE T-CI*, 2019.

[34] Hao Zhu, Qing Wang, and Jingyi Yu. Occlusion-model guided anti-occlusion depth estimation in light field. *IEEE J-STSP*, 11(7):965–978, 2017.

[35] Hao Zhu, Qi Zhang, and Qing Wang. 4D light field superpixel and segmentation. In *IEEE CVPR*, pages 6709–6717, 2017.