Light field planar homography and its application

Qi Zhang, Xue Wang, Qing Wang

Qi Zhang, Xue Wang, Qing Wang, "Light field planar homography and its application," Proc. SPIE 11187, Optoelectronic Imaging and Multimedia Technology VI, 111870S (18 November 2019); doi: 10.1117/12.2538015

SPIE.

Event: SPIE/COS Photonics Asia, 2019, Hangzhou, China

Light Field Planar Homography and Its Application

Qi Zhang, Xue Wang, and Qing Wang

School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China

ABSTRACT

Light field planar homography is essential for light field camera calibration and light field raw data rectification. However, most previous researches assume light field camera as pinhole camera array and deduce the planar homography matrix of sub-aperture image, which is analogous to traditional method. In this paper, we regard light field as a whole and present a novel light field planar homography matrix based on multi-projection-center (MPC) model. The projections of point, line and conic are exploited based on light field planar homography. In addition, a camera calibration method and a homography estimation are proposed to verify the light field planar homography. Experimental results on light field datasets have verified the performance of the proposed methods.

Keywords: Light field camera, planar homography, homography estimation, camera calibration

1. INTRODUCTION

In contrast to conventional camera which captures the 2D projection of a 3D scene, a hand-held light field camera[1,2] records both spatial as well as angular information of rays in a 3D scene. Based on the angular sampling of rays, researches on applications using light field camera have garnered interest, such as refocusing[3] and free-viewpoint roaming. However, the performance of hand-held light field camera is limited by the trade-off between the angular resolution and spatial resolution and narrow baseline in light field. In order to overcome these disadvantages, several applications using multiple light field cameras are proposed. Light field registration methods[4–6] are presented to register multiple light field cameras and reconstruct the 3D scene. In addition, light field camera calibration,[7–11] light field panorama[12] and light field stitching methods[13,14] are also interesting applications.

It is widely known that homography estimation plays an important role in the computer vision. The purpose of homography estimation is to reconstruct the mapping matrix between different coordinates.[15] Consider the fact that the traditional calibration pattern is the planar pattern, it is vital to estimate planar homography matrix between image plane and calibration pattern. In order to recover the planar homography matrix, correspondences are need. In general, there are three traditional types of planar patterns for homography estimation and camera calibration, *i.e.*, points, lines and conics. Previous researches[7–9,11] usually utilize corner points or line features on the calibration pattern to estimate mapping matrix and calibrate light field camera. Bok *et al.*[8] directly detect line features from the raw data and establish a geometric mapping matrix between calibration pattern and micro-lens image for light field camera calibration. In addition, Zhang *et al.* present a multi-projection-center (MPC) model with 6 intrinsic parameters for light field camera, including traditional and focused light field camera. A 3D mapping matrix is derived for describing the relationship between light field coordinates and camera coordinates. And then the light field camera is calibrated by recovering the 3D mapping matrix. What's more, in order to represent the ray sampling and transformation among light field cameras, Zhang *et al.*[11] present a homogeneous 6×6 mapping matrix with 6 intrinsic parameters for light field camera. A linear ray-ray constraint is then deduced to estimate the mapping matrix and then calibrate light field camera. On the other hand, Zhang *et al.*[5] propose a ray-plane geometric mapping matrix between two light fields, and then recover light field poses and conduct bundle adjustment based on line-line correspondences.

Further author information: (Send correspondence to Qing Wang)
Qing Wang: E-mail: qwang@nwpu.edu.cn, Telephone: +86 29 8843 1518
The work was supported by NSFC under Grants No.61531014 and No.61801396. Qi Zhang is also sponsored by Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University under CX201919.

Optoelectronic Imaging and Multimedia Technology VI, edited by Qionghai Dai, Tsutomu Shimura, Zhenrong Zheng,
Proc. of SPIE Vol. 11187, 111870S · © 2019 SPIE · CCC code: 0277-786X/19/$21 · doi: 10.1117/12.2538015

However, several open issues still remain. It is difficult to extract accurate locations of corners due to the effect of noise and the quality of sub-aperture images. On the other hand, in view of the fact that line features are extracted from the low resolution micro-lens images in raw data, the precision of the line features cannot be ensured. Consequently, compared with corner points and line features, conics have two advantages: one is that the conic is well studied in mathematics and can be simply represented by a 3×3 matrix, the other is that conics can be detected and estimated robustly by existing algorithms. Conics have been utilized to calibrate conventional camera over decades. Huang *et al.*[16] propose properties of the common self-polar triangle of concentric circles to estimate the intrinsic parameters of traditional camera. In addition, Huang *et al.*[17] explore the location features of the common self-polar triangle of separate conics for homography estimation. However, a few attentions are drawn to light field planar homography estimation. Zhang *et al.*[10] propose the property and reconstruction of common self-polar triangle with respect to concentric circle and ellipse. Making use of this property, the imaged conic center and vanishing line of sub-aperture image can be recovered simultaneously. This property can also be used to estimate the intrinsic and extrinsic parameters of light field camera.

In this paper, instead of the checkerboard, a concentric conics calibration pattern with known size (*i.e.* a circle and a ellipse with the same center) for light field camera calibration. Based on the MPC model, we first exploit the light field planar homography matrices among light field coordinates, camera coordinates and world coordinates. Then, under a point homography transformation, the line and the conic transformation are deduced. According to the conic homography and the property of common self-polar triangle, a light field camera calibration and a light field homography estimation based on concentric conics are proposed. Finally, we illustrate empirical performances in calibrating light field camera and estimating the homography between two light fields.

Our main contributions are:

1) Light field planar homography matrices among light field coordinates, camera coordinates and world coordinates are exploited.

2) A light field camera calibration algorithm and a homography estimation method between light fields are proposed.

2. PRELIMINARIES

2.1 MPC Model

Light field camera, especially micro-lens array assembled inside, which are innovated from traditional 2D camera, record the 3D world in different but similar rays. With the shifted view, light field camera maps the 3D world to many sub-aperture images. As we all know, light field can be parameterized as a *relative* two-parallel-plane coordinates[18] in general. In prior work, according to the MPC model which is proposed to describe light field camera, a 3D point $\mathbf{X} = (X, Y, Z)^\top$ is mapped to the *relative* pixel (x, y) of the image plane in the view point (s, t) of the view plane,[9]

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -s \\ 0 & 1 & 0 & -t \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \tag{1}$$

where the distance between the two-parallel-plane is normalized as 1 unit to describe the ray (s, t, x, y) in camera coordinates. This is analogous to classical projective camera model with projection center at $(s, t, 0)$ and principal axis paralleling to the Z-axis.

A normalized undistorted ray (s, t, x, y) in camera coordinates can be transformed from a ray captured by light field camera in light field coordinates through a homogeneous matrix $\mathbf{D} \in \mathbb{R}^{5\times5}$,[9]

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

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. $(-u_0/k_u, -v_0/k_v)$ represent the coordinates of principal point in the sub-aperture image.

According to Eq. (1), it is known that at least two rays from one point in space enable 3D reconstruction. The projective transformation between 3D points reconstructed by rays in different coordinates (*i.e.* light field coordinates, camera coordinates) is derived by Eq. (2),

$$\begin{bmatrix} X_d \\ Y_d \\ Z_d \end{bmatrix} = \underbrace{\begin{bmatrix} 1/k_i & 0 & -u_0/k_i \\ 0 & 1/k_j & -v_0/k_j \\ 0 & 0 & k_u/k_i \end{bmatrix}}_{\mathbf{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \tag{3}$$

where $\mathbf{K}$ denotes the intrinsic matrix which describes the transformation between light field coordinates and camera coordinates. $\mathbf{X} = (X, Y, Z)^\top$ is the 3D scene point reconstructed by rays (s, t, x, y) in camera coordinates, $\mathbf{X}_d = (X_d, Y_d, Z_d)^\top$ is the 3D scene point reconstructed by rays (i, j, u, v) in light field coordinates.

In addition, consider $\mathbf{X}_w$ denoting the scene point in world coordinates, the transformation between world and camera coordinates is described by a rotation matrix $\mathbf{R} \in SO(3)$ and a translation vector $\mathbf{t} = (t_x, t_y, t_z)^\top \in \mathbb{R}^3$, formulated as $\mathbf{X} = \mathbf{R}\mathbf{X}_w + \mathbf{t}$. Hence, the projective matrix $\mathbf{P}$ between light field coordinates and world coordinates can be written as,

$$\mathbf{X}_d = \underbrace{\mathbf{K}\begin{bmatrix} \mathbf{R} \mid \mathbf{t} \end{bmatrix}}_{\mathbf{P}} \begin{bmatrix} \mathbf{X}_w \\ 1 \end{bmatrix}. \tag{4}$$

2.2 Common Self-polar Triangle of Concentric Conics

Compared with point and line, conic is well studied in mathematics and easy to be represented by matrix. Consequently, we utilize the conic features for homography estimation between two light fields. We first briefly introduce the notations and property of common self-polar shared by concentric conics. In addition, the reconstruction of common self-polar triangle is proposed for light field homography estimation and light field camera calibration.

According to the notations,[15] a point $\mathbf{x}$ and a conic $\mathbf{C}$ define a line $\mathbf{l} = \mathbf{C}\mathbf{x}$ which is described as *pole-polar relationship*. The line $\mathbf{l}$ is defined as the *polar* of $\mathbf{x}$ with respect to $\mathbf{C}$, and the point $\mathbf{x}$ denotes the *pole* of $\mathbf{l}$ with respect to $\mathbf{C}$. If the poles of a conic form the vertices of a triangle and their respective polars form its opposite sides, it is defined as a *self-polar triangle*. A self-polar which is shared by several conics is what we called *common* self-polar triangle. The property of common self-polar triangle with respect to concentric conics is illustrated,[10]

Property 1. The concentric circle and ellipse have and only have one common self-polar triangle which is right triangle. The three sides of this common self-polar triangle are major axis, minor axis and the line at infinity.

According to the definition of common self-polar triangle, we can reconstruct the common self-polar triangle which is common to the circle $\mathbf{C}_1$ and ellipse $\mathbf{C}_2$. Consider point $\mathbf{x}$ and line $\mathbf{l}$ denoting the common pole-polars of $\mathbf{C}_1$ and $\mathbf{C}_2$. The following relationships should be satisfied:

$$\mathbf{x} = \mathbf{C}_1^{-1}\mathbf{l} \qquad \textit{and} \qquad \mathbf{x} = \lambda\mathbf{C}_2^{-1}\mathbf{l}, \tag{5}$$

where λ is a scale factor. Simplifying these relationships in Eq. (5), we obtain $\left(\lambda\mathbf{C}_2^{-1} - \mathbf{C}_1^{-1}\right)\mathbf{l} = 0$. By multiplying $\mathbf{C}_2$ on both sides, we get the following equation,

$$\left(\lambda\mathbf{I} - \mathbf{C}_2\mathbf{C}_1^{-1}\right)\mathbf{l} = 0. \tag{6}$$

The solutions of Eq. (6) are three eigenvectors with different eigenvalues of $\mathbf{C}_2\mathbf{C}_1^{-1}$. What's more, these eigenvectors are also the common polars of point $\mathbf{x}$ with respect to circle $\mathbf{C}_1$ and ellipse $\mathbf{C}_2$. According to the *Property* 1, these polars are three special sides of common self-polar triangle, *i.e.* major axis, minor axis and the line at infinity. Similarly, we can also derive the vertices fo the common self-polar triangle by computing the eigenvectors of $\mathbf{C}_2^{-1}\mathbf{C}_1$.

3. LIGHT FIELD PLANAR HOMOGRAPHY

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. The light field essentially represents a set of rays come from scene points. Consequently, we consider that Eq. (4) is sufficient to represent light field. Eq. (4) is a map from a point in world coordinates to a point in light field coordinates. We have the freedom to choose the world coordinates. Suppose it is chosen such that the OX_wY_w corresponds to a plane π in the scene, so that points on the scene plane have $Z_w = 0$. Then, the light field planar homography is given by

$$\mathbf{X}_d = \underbrace{\mathbf{K}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix}}_{\mathbf{H}}\begin{bmatrix}X_w\\ Y_w\\ 1\end{bmatrix}, \tag{7}$$

so that the map between points $\mathbf{X}_w = (X_w, Y_w, 1)^\top$ on plane π and their point $\mathbf{X}_d$ recorded by light field camera is a general planar homography.

Eq. (7) describe the planar homography between light field coordinates and world coordinates. It is well-known that light field camera is assumed as a pinhole camera array. Considered the fact that the transform between sub-aperture images within same light field is merely a translation $\mathbf{t}_{ij} = (-i, -j, 0)^\top$ without rotation, according to the planar homography matrix between 2D pinhole cameras[15] and MPC model,[9] we can also derive the planar homography between views in the light field,

$$\begin{bmatrix}u\\ v\\ 1\end{bmatrix} \sim \underbrace{\left(\mathbf{I} + \mathbf{t}_{ij}\mathbf{n}^\top/d\right)}_{\mathbf{H}_{ij}}\mathbf{X}_d, \tag{8}$$

where $\sim$ refers to equality up to a scale, $\mathbf{n}$ and d denote the scene plane which can be computed by the vanishing line of light field. Then, substituting Eq. (8) into Eq. (7), there is a planar homography between scene plane and sub-aperture images,

$$\mathbf{u} \sim \mathbf{H}_{ij}\mathbf{H}\mathbf{X}_w. \tag{9}$$

If a 3D line $\mathbf{l}_w$ on the scene plane π, the line transformation between the projective line $\mathbf{l}$ on the sub-aperture image of view (i, j) and scene line $\mathbf{l}_w$ is,

$$\mathbf{l} \sim \mathbf{H}_{ij}^{-\top}\mathbf{H}^{-\top}\mathbf{l}_w. \tag{10}$$

Similarly, we can also derive the transformation between projective conic $\widetilde{\mathbf{C}}$ on the sub-aperture image of view (i, j) and conic $\mathbf{C}_w$ on the scene plane,

$$\widetilde{\mathbf{C}} \sim \mathbf{H}_{ij}^{-\top}\mathbf{H}^{-\top}\mathbf{C}_w\mathbf{H}^{-1}\mathbf{H}_{ij}^{-1}. \tag{11}$$

We call this constraint light field planar homography and we show in the following section how to use the planar homography of conic for light field camera calibration and homography estimation between two light fields.

4. CAMERA CALIBRATION AND HOMOGRAPHY ESTIMATION

4.1 Light Field Camera Calibration

Let the images of two concentric conics under the conic homography matrix Eq. (11) be $\widetilde{\mathbf{C}_1}$ and $\widetilde{\mathbf{C}_2}$ on the view (i, j), we have

$$\mathbf{H}_{ij}^\top\widetilde{\mathbf{C}}_1\mathbf{H}_{ij} \sim \mathbf{H}^{-\top}\mathbf{C}_{w1}\mathbf{H}^{-1} \quad \textit{and} \quad \mathbf{H}_{ij}^\top\widetilde{\mathbf{C}}_2\mathbf{H}_{ij} \sim \mathbf{H}^{-\top}\mathbf{C}_{w2}\mathbf{H}^{-1}. \tag{12}$$

According to the *Property* 1, line at infinity can be calculated by the product $\mathbf{C}_2\mathbf{C}_1^{-1}$. Therefore, it is easy to obtain the vanishing line of light field and compute the planar homography $\mathbf{H}_{ij}$. In addition, by computing the product $\mathbf{C}_2\mathbf{C}_1^{-1}$, we also have

$$\mathbf{H}_{ij}^\top\left(\widetilde{\mathbf{C}}_2\widetilde{\mathbf{C}}_1^{-1}\right)\mathbf{H}_{ij} \sim \mathbf{H}^{-\top}\left(\mathbf{C}_{w2}\mathbf{C}_{w1}^{-1}\right)\mathbf{H}^\top. \tag{13}$$

As illustrated in Eq. (13), $\widetilde{\mathbf{C}}_2\widetilde{\mathbf{C}}_1^{-1}$ is the similar matrix of $\mathbf{C}_{w2}\mathbf{C}_{w1}^{-1}$, which means they have the same eigenvalues. Consider λ and $\mathbf{l}_w$ denoting the eigenvalue and eigenvector of $\mathbf{C}_{w2}\mathbf{C}_{w1}^{-1}$ respectively, according to the property of similarity transformation and Eq. (10), λ and $\mathbf{H}_{ij}^{-\top}\mathbf{H}^{-\top}\mathbf{l}_w$ are eigenvalue and eigenvector of $\widetilde{\mathbf{C}}_2\widetilde{\mathbf{C}}_1^{-1}$ respectively. Based on this observation, we can find three line-line correspondences by matching the eigenvalues. In order to estimate the light field homography matrix $\mathbf{H}$, at least four line-line correspondences are needed.[15] Then, we add the lines through intersecting points at the conics with the sides of self-polar triangle to estimate $\mathbf{H}$.

In order to derive intrinsic parameters, we abbreviate with $[\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$ of $\mathbf{H}$. Utilizing the orthogonality and identity of $\mathbf{r}_1$ and $\mathbf{r}_2$, we have

$$\begin{aligned} &\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2 = 0, \\ &\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2. \end{aligned} \tag{14}$$

Consider a symmetric matrix $\mathbf{B}$ denoting $\mathbf{K}^{-\top}\mathbf{K}^{-1}$ which has four degree of freedom, Eq. (14) can be rewritten as follows,

$$\begin{bmatrix} h_{11}h_{12} & h_{11}h_{32}+h_{12}h_{31} & h_{21}h_{22} & h_{21}h_{32}+h_{22}h_{31} & h_{31}h_{32} \\ h_{11}^2-h_{12}^2 & 2(h_{11}h_{31}-h_{12}h_{32}) & h_{21}^2-h_{22}^2 & 2(h_{21}h_{31}-h_{22}h_{32}) & h_{31}^2-h_{32}^2 \end{bmatrix} \mathbf{b} = 0, \tag{15}$$

where h_{ij} refers to the element on the $i-th$ row and $j-th$ column of $\mathbf{H}$. $\mathbf{b}$ denotes the vector $(b_{11}, b_{13}, b_{22}, b_{23}, b_{33})^\top$. By stacking at least two such equations (from two poses) as Eq. (15) in view of the degree of freedom, we can reconstruct the $\mathbf{B}$ with a scale factor. Furthermore, it is easy to utilize Cholesky factorization[19] to solve $\mathbf{K}$ and obtain intrinsic parameters (k_u, k_v, u_0, v_0). In addition, extrinsic parameters in different poses can be calculated as follows,

$$\begin{aligned} \alpha &= \frac{1}{2}\left(\left\|\hat{\mathbf{K}}^{-1}\mathbf{h}_1\right\| + \left\|\hat{\mathbf{K}}^{-1}\mathbf{h}_2\right\|\right), \\ \mathbf{r}_1 &= \frac{1}{\alpha}\hat{\mathbf{K}}^{-1}\mathbf{h}_1, \\ \mathbf{r}_2 &= \frac{1}{\alpha}\hat{\mathbf{K}}^{-1}\mathbf{h}_2, \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2, \\ \mathbf{t} &= \frac{1}{\alpha}\hat{\mathbf{K}}^{-1}\mathbf{h}_3, \end{aligned} \tag{16}$$

where $\hat{\mathbf{K}}$ refers to the estimation of $\mathbf{K}$. $\|\cdot\|$ denotes L_2 norm. In order to estimate the rest intrinsic parameters k_i and k_j, we first calculate the center point (x_o, y_o) of concentric circle and ellipse in the camera coordinate according to the *Property* 1. Then, based on Eqs. (2) and (8), we have

$$\begin{bmatrix} i & 0 \\ 0 & j \end{bmatrix} \begin{bmatrix} k_i \\ k_j \end{bmatrix} = \begin{bmatrix} t_x/t_z - x_o \\ t_y/t_z - y_o \end{bmatrix}. \tag{17}$$

Stacking the measurements in different poses, we can obtain a unique non-zeros solution for k_i and k_j.

The most common distortion model of light field camera is the radial distortion shifted with the view (s, t).[9] Compared with radial distortion coefficient (k_1, k_2) of traditional camera, k_3 and k_4 are added to represent the distortion affected by the shifted view. Once we obtain the initial linear solution computed by Eqs. (15), (16) and (17), a cost function, based on Sampson error[15] and Eqs. (9) and (12) is proposed to conduct the non-linear optimization,

$$\sum \frac{\left|\mathbf{u}\left(\mathbf{k}^d\right)^\top \mathbf{C}\left(\mathcal{P}, \mathbf{R}, \mathbf{t}, \mathbf{C}_w\right)\mathbf{u}\left(\mathbf{k}^d\right)\right|}{2\left\|\mathbf{C}\left(\mathcal{P}, \mathbf{R}, \mathbf{t}, \mathbf{C}_w\right)\mathbf{u}\left(\mathbf{k}^d\right)\right\|}, \tag{18}$$

where $\mathbf{u}$ refers to the sub-aperture image point on conic, $\mathcal{P}$ denotes the light field camera parameters, $\mathbf{k}^d$ is the distortion coefficient, $\mathbf{R}, \mathbf{t}$ refer to the extrinsic parameters and $\mathbf{C}_w$ denotes the conic matrix in the world coordinates. This non-linear cost function can be solved using Levenberg-Marquardt algorithm based on trust region method.[20] Matlab's lsqnonlin is utilized to carry out the non-linear optimization. The light field calibration algorithm is summarized in Alg 1.

Algorithm 1 Light Field Camera Calibration Algorithm.

Input: Concentric conics $\mathbf{C}_{w1}$ and $\mathbf{C}_{w2}$ and projected conics $\widetilde{\mathbf{C}}_1$ and $\widetilde{\mathbf{C}}_2$ on each view.
Output: Intrinsic parameters $\mathcal{P} = (k_i, k_j, k_u, k_v, u_0, v_0)$;
Extrinsic parameters $\mathbf{R}_p$, $\mathbf{t}_p (1 \leq p \leq P)$;
Distortion vector $\mathbf{k}^d = (k_1, k_2, k_3, k_4)^\top$.

for $p = 1$ to P **do**
  for each view (i, j) **do**
    Obtain the eigenvectors of $\widetilde{\mathbf{C}}_2 \widetilde{\mathbf{C}}_1^{-1}$ and planar homography matrix $\mathbf{H}_{ij}$. ▷ Eqs. (8) and (11)
  end for
  Find the line-line correspondences by matching the eigenvalues of $\widetilde{\mathbf{C}}_2 \widetilde{\mathbf{C}}_1^{-1}$ and $\mathbf{C}_{w2} \mathbf{C}_{w1}^{-1}$. ▷ Eq. (10)
  Solve light field homography matrix $\mathbf{H}$. ▷ Eq. (13)
end for
Obtain four intrinsic parameters (k_u, k_v, u_0, v_0) by Cholesky factorization. ▷ Eq. (15)
for $p = 1$ to P **do**
  Get extrinsic parameters $\mathbf{R}_p$ and $\mathbf{t}_p$. ▷ Eq. (16)
end for
Obtain other two intrinsic parameters (k_i, k_j). ▷ Eq. (17)
Initialize distortion coefficient. $\mathbf{d} = (0, 0, 0, 0)^\top$
Create the cost function based on Sampson error and obtain optimized results using nonlinear LM algorithm.

4.2 Light Field Homography Estimation

Similarly, consider the sub-aperture images of two concentric conics $\mathbf{C}_{w1}$ and $\mathbf{C}_{w2}$ on the view (i, j) of first light field being $\widetilde{\mathbf{C}}_{11}$ and $\widetilde{\mathbf{C}}_{21}$ and the images on the view (m, n) of second light field being $\widetilde{\mathbf{C}}_{12}$ and $\widetilde{\mathbf{C}}_{22}$. Let the planar transformation from the first light field to the second one is $\mathbf{H}$. According to Eq. 11, we have

$$\mathbf{H}_{mn}^\top \widetilde{\mathbf{C}}_{12} \mathbf{H}_{mn} \sim \mathbf{H}^{-\top} \mathbf{H}_{ij}^\top \widetilde{\mathbf{C}}_{11} \mathbf{H}_{ij} \mathbf{H}^{-1} \quad \textit{and} \quad \mathbf{H}_{mn}^\top \widetilde{\mathbf{C}}_{22} \mathbf{H}_{mn} \sim \mathbf{H}^{-\top} \mathbf{H}_{ij}^\top \widetilde{\mathbf{C}}_{21} \mathbf{H}_{ij} \mathbf{H}^{-1}. \tag{19}$$

Based on the *Property* 1, we also obtain,

$$\mathbf{H}_{mn}^\top \left(\widetilde{\mathbf{C}}_{22} \widetilde{\mathbf{C}}_{12}^{-1} \right) \mathbf{H}_{mn} \sim \mathbf{H}^{-\top} \mathbf{H}_{ij}^\top \left(\widetilde{\mathbf{C}}_{21} \widetilde{\mathbf{C}}_{11}^{-1} \right) \mathbf{H}_{ij} \mathbf{H}^{-1}, \tag{20}$$

which illustrates that $\widetilde{\mathbf{C}}_{22} \widetilde{\mathbf{C}}_{12}^{-1}$ is the similar matrix of $\widetilde{\mathbf{C}}_{21} \widetilde{\mathbf{C}}_{11}^{-1}$. They all have the same eigenvalues. According to the method provided in Sec. 4.1, we can also find the line-line correspondence by matching the eigenvalues of $\widetilde{\mathbf{C}}_{21} \widetilde{\mathbf{C}}_{11}^{-1}$ and $\widetilde{\mathbf{C}}_{22} \widetilde{\mathbf{C}}_{12}^{-1}$ and estimate the homography matrix $\mathbf{H}$ between two light fields by existing algorithms.[15]

In addition, the estimated homography matrix solved by linear method Eq. (20) is refined via non-linear optimization. We define a cost function based on Sampson distance and Eq. (19) to optimize results,

$$\sum \frac{\left| \mathbf{u}_2^\top (m,n) \, \mathbf{H}^{-\top} \widetilde{\mathbf{C}}_1 (i,j) \, \mathbf{H}^{-1} \mathbf{u}_2 (m,n) \right|}{2 \left\| \mathbf{H}^{-\top} \widetilde{\mathbf{C}}_1 (i,j) \, \mathbf{H}^{-1} \mathbf{u}_2 (m,n) \right\|} + \frac{\left| \mathbf{u}_1^\top (i,j) \, \mathbf{H}^{\top} \widetilde{\mathbf{C}}_2 (m,n) \, \mathbf{H} \mathbf{u}_1 (i,j) \right|}{2 \left\| \mathbf{H}^{\top} \widetilde{\mathbf{C}}_2 (m,n) \, \mathbf{H} \mathbf{u}_1 (i,j) \right\|}. \tag{21}$$

The light field calibration algorithm is summarized in Alg 2.

5. EXPERIMENTS AND RESULTS

5.1 Light Field Camera Calibration

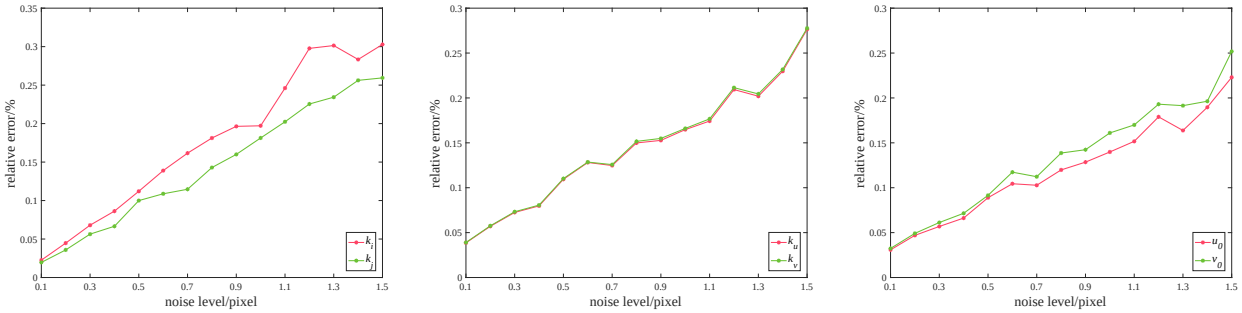
In order to evaluate the performance of the proposed light field calibration algorithm, we carry out the experiments on both simulated data and real scene. We first simulate a light field camera which is same as the baseline method[10] (*i.e.* $k_i = 1.4e-4$, $k_j = 1.5e-4$, $k_u = 2.0e-3$, $k_v = 1.9e-3$, $u_0 = -0.59$, $v_0 = -0.52$). The model pattern is conic pattern containing two concentric circle and ellipse. The radius of circle is $0.05m$. The major and minor axes of ellipse are $0.13m$ and $0.07m$. In this experiment, we employ the measurement of 3 poses and 7×7 views to demonstrate the robustness of our calibration algorithm. We choose 100 sub-aperture image points

Algorithm 2 Light Field Homography Estimation Algorithm.

Input: Projected conics $\widetilde{\mathbf{C}}_{11}$ and $\widetilde{\mathbf{C}}_{21}$ on each view of the first light field;
Projected conics $\widetilde{\mathbf{C}}_{12}$ and $\widetilde{\mathbf{C}}_{22}$ on each view of the second light field.
Output: Homography matrix $\mathbf{H}$ between two light fields.

1: **for** each view (i, j) of first light field and each view (m, n) of second light field **do**
2: Obtain the eigenvectors of $\widetilde{\mathbf{C}}_{21}\widetilde{\mathbf{C}}_{11}^{-1}$, $\widetilde{\mathbf{C}}_{22}\widetilde{\mathbf{C}}_{12}^{-1}$ and planar homography matrix $\mathbf{H}_{ij}$, $\mathbf{H}_{mn}$. ▷ Eq. (8)
3: **end for**
4: Find the line-line correspondences by matching the eigenvalues of $\widetilde{\mathbf{C}}_{21}\widetilde{\mathbf{C}}_{11}^{-1}$ and $\widetilde{\mathbf{C}}_{22}\widetilde{\mathbf{C}}_{12}^{-1}$. ▷ Eq. (10)
5: Solve homography matrix $\mathbf{H}$ between two light fields. ▷ Eq. (20)
6: Create the cost function based on Sampson error and obtain optimized results using nonlinear LM algorithm.

on each conic. Gaussian noise with zero-mean and σ standard deviation is added to these images points. We vary σ from 0.1 to 1.5 pixels with a 0.1 pixel step. For each noise level, 150 independent trials are conducted. The accuracy is evaluated by the average of relative errors with ground truth. As illustrated in Fig. 1, it is demonstrated the errors almost linearly increase with noise level, which exhibits that the proposed algorithm is robust to higher noise level.



(a) k_i and k_j (b) k_u and k_v (c) u_0 and v_0

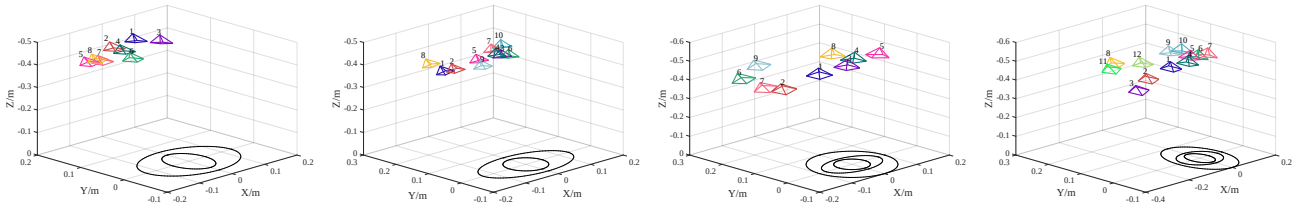
Figure 1. Performance evaluation of intrinsic parameters on the simulated data with different levels of noise σ.

We also verify the calibration method on real scene light fields collected with the baseline method.[10] The sub-aperture images are obtained by the method.[7] The Canny edge detector[21] is utilized to detect the conic from sub-aperture images, and conics are fitted to these images using a least squares conic fitting algorithm. For every different pose, the middle 11×11 sub-aperture images are utilized similar to the compared method.[10] We than use the proposed method (Alg. 1) to calibrate light field camera. Tab. 1 summarizes the root mean square (RMS) Sampson errors of initialization, optimization without and with distortion compared with the method proposed by Zhang *et al.*[10] As illustrated in Tab. 1, the proposed method obtains similar Sampson errors on the item of initial solution compared with the state-of-the-art method, which verified the effectiveness of linear initial solution for both intrinsic and extrinsic parameters. From the results, we find both methods have small different, which further exhibits the proposed method is robustness and effectiveness. In addition, the results of pose estimation on the datasets are illustrated in Fig. 2

Table 1. RMS Sampson errors of initialization, optimizations without and with distortion (unit: *pixel*).

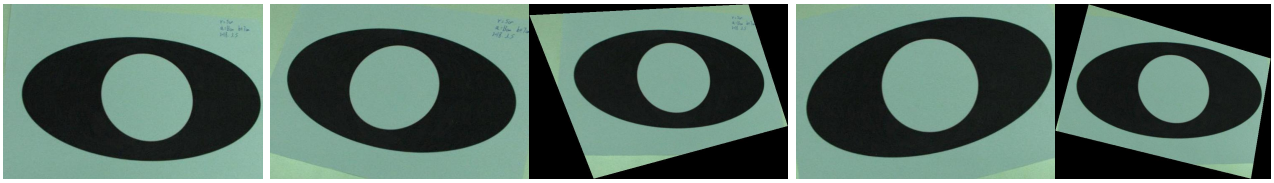
		Illum-1 (8)	Illum-2 (10)	Illum-3 (9)	Illum-4 (12)
Initial	Zhang[10]	5.2401	6.6361	4.1867	**5.0426**
	Ours	**5.0497**	**6.0590**	**3.9557**	5.1443
Opt. w/o distortion	Zhang[10]	1.1173	1.1081	0.7818	0.7838
	Ours	**1.0519**	**1.0428**	**0.7285**	**0.7433**
Opt. with distortion	Zhang[10]	**0.1746**	**0.1778**	0.2879	**0.2889**
	Ours	0.1759	0.1809	**0.2482**	0.2908

The (N) denotes the number of light fields used for calibration.

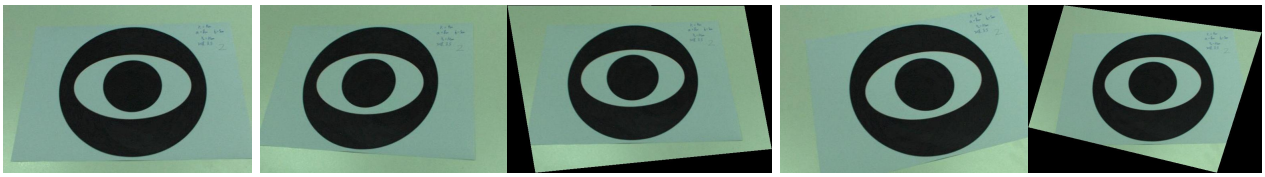


(a) Illum-1 (b) Illum-2 (c) Illum-3 (d) Illum-4

Figure 2. Pose estimation results of the light field datasets.



(a) Reference view (b) Central view of first light filed (c) Central view of second light filed



(d) Reference view (e) Central view of first light filed (f) Central view of second light filed

Figure 3. The homography rectification results. The first column shows the central view of reference light field. The second and third columns show the original and rectified central view of the first light field. The forth and fifth columns show the original and rectified central view of the second light field.

5.2 Light Field Homography Estimation

We conduct two real experiments to certify the performance of the proposed light field homography estimation method. We capture three light fields using the two concentric conics pattern and three concentric conics pattern respectively. It is noticed that there exists distinct sidtortion in the Illum camera. We first utilize the proposed method in Sec. 5.1 for distortion rectification. Similarly, the conic matrices on sub-aperture images are obtained by Canny edge detector. We than utilize our presented approach Alg. 2 to compute the homography for light fields. Finally, we warp the light field back to the reference view. Fig. 3 illustrate the central view of the reference light field and the results.

6. CONCLUSION

In this paper, based on the property of concentric conics, a novel light field camera calibration method and a homography matrix estimation between two light fields are proposed. Compared with point and line pattern, conic pattern is more simply and robustness. We first explore the light field planar homography matrices among light field coordinates, camera coordinates and world coordinates. According to the conic homography matrix and common self-polar triangle, four line-line correspondences can be uniquely determined. Than, a linear method for light field camera calibration and a linear method for homography estimation are presented. Finally, qualitative and quantitative analyses on extensive experiments verify the performance of the proposed methods.

REFERENCES

[1] Lytro, "Lytro redefines photography with light field cameras." http://www.lytro.com (2011).

[2] Raytrix, "3d light field camera technology." http://www.raytrix.de (2013).

[3] Ng, R., "Fourier slice photography," *ACM TOG* **24**(3), 735–744 (2005).

[4] Johannsen, O., Sulc, A., and Goldluecke, B., "On linear structure from motion for light field cameras," in [*IEEE ICCV*], 720–728 (2015).
[5] Zhang, Y., Yu, P., Yang, W., Ma, Y., and Yu, J., "Ray space features for plenoptic structure-from-motion," in [*IEEE CVPR*], 4631–4639 (2017).
[6] Nousias, S., Lourakis, M., and Bergeles, C., "Large-scale, metric structure from motion for unordered light fields," in [*IEEE CVPR*], 3292–3301 (2019).
[7] Dansereau, D. G., Pizarro, O., and Williams, S. B., "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in [*IEEE CVPR*], 1027–1034 (2013).
[8] Bok, Y., Jeon, H.-G., and Kweon, I. S., "Geometric calibration of micro-lens-based light field cameras using line features," *IEEE T-PAMI* **39**(2), 287–300 (2017).
[9] Zhang, Q., Zhang, C., Ling, J., Wang, Q., and Yu, J., "A generic multi-projection-center model and calibration method for light field cameras," *IEEE T-PAMI* (2018).
[10] Zhang, Q. and Wang, Q., "Common self-polar triangle of concentric conics for light field camera calibration," in [*ACCV*], **Part VI**, 18–33 (2018).
[11] Zhang, Q., Ling, J., Wang, Q., and Yu, J., "Ray-space projection model for light field camera," in [*IEEE CVPR*], 10121–10129 (2019).
[12] Birklbauer, C. and Bimber, O., "Panorama light-field imaging," in [*Computer Graphics Forum*], **33**(2), 43–52, Wiley Online Library (2014).
[13] Guo, X., Yu, Z., Kang, S. B., Lin, H., and Yu, J., "Enhancing light fields through ray-space stitching," *IEEE T-VCG* **22**(7), 1852–1861 (2016).
[14] Ren, Z., Zhang, Q., Zhu, H., and Wang, Q., "Extending the FOV from disparity and color consistencies in multiview light fields," in [*Proc. ICIP*], 1157–1161 (2017).
[15] Hartley, R. and Zisserman, A., [*Multiple view geometry in computer vision*], Cambridge University Press (2003).
[16] Huang, H., Zhang, H., and Cheung, Y.-m., "The common self-polar triangle of concentric circles and its application to camera calibration.," in [*IEEE CVPR*], 4065–4072 (2015).
[17] Huang, H., Zhang, H., and Cheung, Y.-m., "Homography estimation from the common self-polar triangle of separate ellipses," in [*IEEE CVPR*], 1737–1744 (2016).
[18] Levoy, M. and Hanrahan, P., "Light field rendering," in [*ACM SIGGRAPH*], 31–42 (1996).
[19] Hartley, R., "Self-calibration of stationary cameras," *IJCV* **22**(1), 5–23 (1997).
[20] Madsen, K., Nielsen, H. B., and Tingleff, O., [*Methods for non-linear least squares problems, 2nd Edition*], Informatics and Mathematical Modelling, Technical University of Denmark (2004).
[21] Canny, J., "A computational approach to edge detection," in [*Readings in Computer Vision*], 184–203, Elsevier (1987).