



# Light field calibration and 3D shape measurement based on epipolar-space

PING ZHOU,\* ZI YANG, WEIJIA CAI, YUNLEI YU, AND GUANGQUAN ZHOU

School of Biological Science & Medical Engineering, Southeast University, Nanjing 210096, China

\*capzhou@163.com

**Abstract:** Light field camera calibration is much more complicated by the fact that a single point in the 3D scene appears many times in the image plane. Compared to the previous geometrical models of light field camera, which describe the relationship between 3D point in the scene and 4D light field, we proposed an epipolar-space (EPS) based geometrical model in this paper, which determines the relationship between 3D point in the scene and 3-parameter vector in the EPS. Moreover, a close-form solution for the 3D shape measurement based on the geometrical model is accomplished. Our calibration method includes an initial linear solution and nonlinear optimization with the Levenberg-Marquardt algorithm. The light field model is validated with the commercially available light field camera Lytro iIllum, and the performance of 3D shape measurement is verified by both real scene data and the data set on the internet.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Unlike traditional 2D imaging that integrates a light beam from a point, the light field (LF) camera that consists a main-lens and a micro-lens array (MLA) allows to capture both spatial and angular information of a light ray from our world simultaneously. The data captured by light field camera is equivalent to that captured by cameras from different viewpoints, so that the light field data contains information about the three-dimensional (3D) shape of a scene in a single photographic exposure. The light field imaging has been recently developed for both scientific researches and industrial applications, such as LF rendering, scene reconstruction, 3D microscopy, 3D endoscopy, *et al.* One of important applications of light field is depth estimation, in which the light field calibration is not necessary completely, making it convenient for data acquisition. However, the 3D shape measurement is much more essential than the depth estimation for many applications. To support 3D shape measurement, it is crucial to perform light field calibration accurately and establish precise relationship between a certain point in the 3D scene and its corresponding parameters in the light field.

Generally, the main-lens is treated as a pinhole model, the micro-lens is regarded as a thin-lens model, and the 3D light field is defined via the so-called two-parallel-plane (TPP) model where a light ray is described by the intersection points of two parallel planes. Recently, some state-of-the-art light field models have been proposed to accomplishment light field calibration. In 2013, Dansereau [1] *et al.* proposed 15-parameter light field camera model (12 free parameters in the intrinsic matrix), where a reference plane outside the light field camera is present as one of the two-parallel-planes. However, the reference plane is lack of specific meaning and there are redundant parameters in the transformation matrix, the calibration model is not easy to be used. In 2018, Wang [2] *et al.* proposed a multi-projection-center model based on two-parallel-planes. In their model, a ray is described by two planes with an alterable distance, which generally parameterizes the light field in focused and defocused formation. Instead of defining two parallel planes with variable positions, most light field models regard the main-lens plane and micro-lens array plane as the two parallel planes. In 2014 and 2017, Bok [3] *et al.* complete their calibration model by utilizing light field raw data directly. In 2018, Peng [4] *et al.* proposed an active calibration method with the aid of an auxiliary camera and a projector. In this method, a series of target points along light field rays

within a measurement volume are used to determine a look-up table (LUT), which describes the relationship between the light field rays and calibration parameters. Thurow [5] *et al.* also proposed a volumetric calibration method based on polynomial mapping function, where the lens distortion and thin-lens assumption are considered to improve the calibration accuracy.

In a word, a single 3D point has only one projected image in traditional camera, while the light field camera is much more complicated by the fact that a single point will appear in the image plane multiple times. As mentioned above, the previous light field camera calibration models often illustrated the relationship between the 3D scene and 4D rays, which are not only complicated but also redundant. In this paper, a light field calibration model is proposed with a one-to-one correspondence between a point in the 3D scene and 3D parameters in a so called epipolar-space (EPS), along with accurate 3D shape measurement based on our calibration model. Both intrinsic and extrinsic matrices are initialized with EPI images, then refined by nonlinear distortion correction and minimizing the squared sum of ray reprojection errors. In addition, the 3D shape measurement could be accomplished by combining the light field camera calibration parameters and the depth estimation results. In 2012, Wanner and Goldluecke [6] estimated the depth information by measuring the local line orientation in the epipolar-plane image (EPI), where the structure tensor algorithm is utilized to calculate the orientation and assess its reliability. Zeller [7] *et al.* proposed a high-order curve model for depth estimation, which was iteratively updated using depth error compensation. Williem [8] *et al.* proposed a framework for occlusion and noise-aware light field depth estimation, where they introduced the constrained angular entropy metric to measure the randomness of pixel color in the angular patch while reducing the effect of the occlude and noise.

The remainder of this paper is organized as follows. Section 2 introduces the background and some related works of light field. Section 3 introduces our EPS-based calibration model and the transformation matrix between the 3D point and 3D parameters in the EPS. In section 4, our calibration method is described in detail, including a nonlinear optimization. The experimental results are presented in section 5 to demonstrate the performance of our method.

## 2. Background and related works

A light field was defined as a 7D plenoptic function  $L(x, y, z, \theta, \Phi, \lambda, t)$  by Adelson and Bergen [9] in 1991. Under the assumption that the scene is static when exposures, the wavelength of a particular light is constant, and the light propagates in a transparent medium, the 7D light field plenoptic function was reduced to a 4D function by Levoy [10]. In 2005, Ng [11] *et al.* integrated a micro-lens array between the image sensor and main-lens to accomplish a compact version of light field camera, where the object points are imaged at different angles via the micro-lenses, as shown in Fig. 1(a). Without loss of generality, the 4D light field is denoted  $L(s, t, x, y)$  in this paper, which has been widely utilized for its conciseness. As shown in Fig. 1(b), the 4D light field  $L(s, t, x, y)$  intersects the angular plane at  $(s, t)$  and the spatial plane at  $(x, y)$ . In this paper, the angular plane and spatial plane are the main-lens plane and micro-lens array plane, respectively.

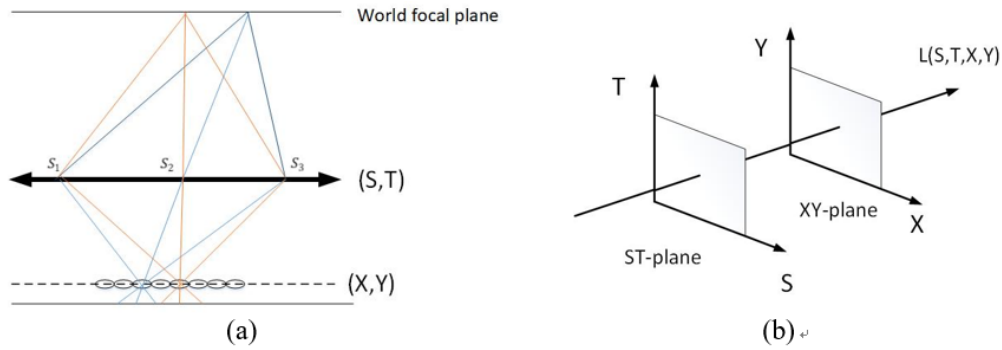


Fig. 1. 4D light field.

As shown in Fig. 2, the image plane, micro-lens array (MLA) plane and main-lens plane are parallel to each other and all perpendicular to the optical axis. To express the light field ray with TPP model, the notation of symbols used in the light field model is given in Table 1.

Table 1. Notation of symbols in the light field model.

$O_i - X_p Y_p$	Image coordinate system
$O_c - X_c Y_c Z_c$	Camera coordinate system
$O_w - X_w Y_w Z_w$	World coordinate system
$O_0 - X_m Y_m$	Micro-lens array coordinate system
$(i, j)$	pixel coordinate in the sub-aperture image
$L(S, T, X, Y)$	Light field in TPP model in the unit of millimeter
$L(u, v, x, y)$	Decoded light field

Without loss of generality, the optical center of the main lens and optical axis are defined as the origin  $O_c$  and the  $z_c$ -axis of camera coordinate system, respectively. All coordinate systems follow the same convention: from the observation view (towards right in Fig. 2(a)), the  $Z$  axis points towards the object ( $Z_c > 0$ ). The  $Y$  axis points downwards and the  $X$  axis points to right. All the coordinates with origins  $O_x$  and in the light field  $L(S, T, X, Y)$  are in the unit of millimeter. To analysis the light field in a certain light field camera, a decoded light field  $L(u, v, x, y)$  is obtained based on Dansereau's method, as shown in Figs. 2(b) and 2(c), where  $(u, v)$  and  $(x, y)$  are indices in element image and micro-lens array respectively.

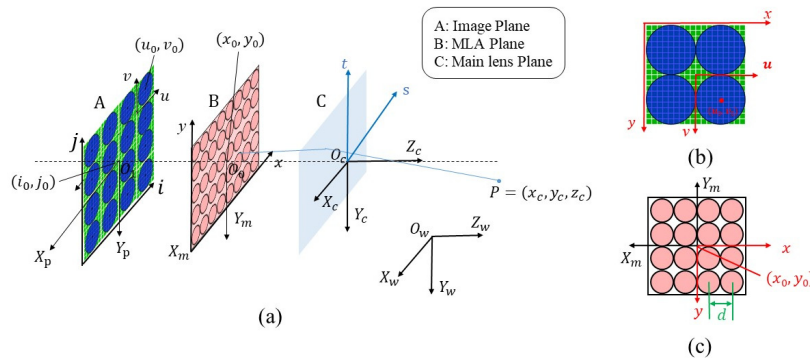


Fig. 2. Geometrical model of light field cameras. (a) Projection model. (b) Image plane with four element images. (c) Micro-lens array plane.

In 1987, Bolles *et al.* proposed the epipolar-plane image (EPI) method, which they used to estimate sparse disparity information by analyzing the slopes of lines. As shown in Fig. 3, the epipolar-plane images are generated by collecting the light field data with a fixed angular coordinate  $u^*$  and a fixed position coordinate  $x^*$  (expressed as  $I_{u^*x^*}(v, y)$ ), or with a fixed angular coordinate  $v^*$  and a fixed position coordinate  $y^*$  (expressed as  $I_{v^*y^*}(u, x)$ ), which contain information in both angular and spatial dimensions in one image simultaneously. Considering a certain point with depth  $Z_w$  in the 3D scene, the coordinate  $u$  changes when the coordinate  $x$  changes, therefore a line is formed on the EPI. The points with different depths are visualized as the lines with corresponding different slopes on the EPI. In other words, the slopes of the lines are indicative of the depths of the different points in the 3D scene, which is the basis for depth estimation. Generally speaking, it is imprecise to calculate slope from EPI straightly because of the limited number of sub-aperture images. In 2012, Wanner [6] *et al* proposed the extending epipolar-plane image method to calculate the slope in EPI, which is more accurate than these methods using sub-aperture images directly.

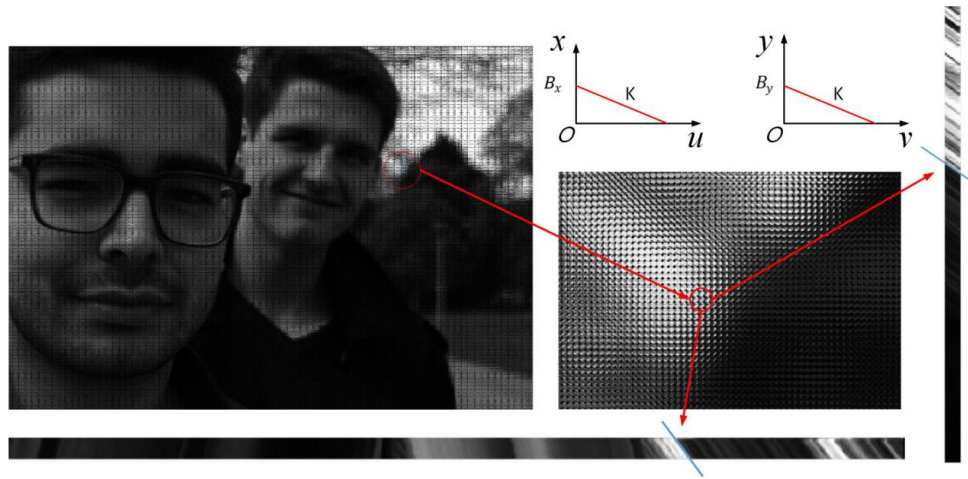


Fig. 3. Epipolar-plane images. Bottom: the EPI  $I_{v^*y^*}(u, x)$  generated by fixing  $v$  and  $y$ , Right: the EPI  $I_{u^*x^*}(v, y)$  generated by fixing  $u$  and  $x$ .

For a certain point in the scene, two corresponding lines are generated in the EPIs  $I_{v^*y^*}(u, x)$  and  $I_{u^*x^*}(v, y)$  respectively, as shown in the bottom and right of Fig. 3.

### 3. EPS-based geometrical model of light field camera

Given a 3D point  $P(x_c, y_c, z_c)$  in the scene, as shown in Fig. 4, its light field imaging process is illustrated in 2D condition for simplification. A light ray from point  $P$  intersects main-lens plane at  $P_m$ , where  $S$  is the distance in the unit of millimeter of the intersection point  $P_m$  with respect to the optical axis. Subsequently, the light ray intersects the micro-lens array plane at  $P_x$  and image plane at  $i$ , where  $x_m$  is the distance in the unit of millimeter of the intersection point  $P_x$  with respect to the optical axis. The corresponding decoded light field is expressed as  $L(u, v, x, y)$ .

The following relationship is derived as the triangles that have equal angles are similar.

$$\frac{S - x_c}{z_c} + \frac{S - x_m}{h'_m} = \frac{S - x_d}{h_m} + \frac{S - x_m}{h'_m} \quad (1)$$

where  $h_m$  is the distance between the focal plane and main-lens plane,  $h'_m$  is the distance between the main-lens plane and micro-lens array plane,  $x_c$  is the distance between the point  $P$

to the optical axis. In addition, the light ray intersects the focal plane at  $P'$ , where  $x_d$  is the distance between the point  $P'$  to the optical axis. All the parameters in Eq. (1) are in the unit of millimeter.

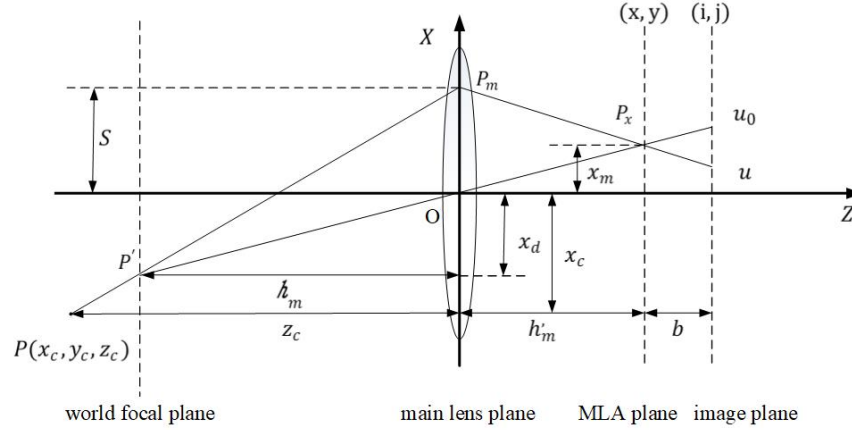


Fig. 4. Geometrical Model of Light Field Camera.

In this paper, the main-lens is considered to be an ideal thin-lens. Therefore, the relationship between  $h_m$  and  $h'_m$  is defined by the thin-lens Gauss theory.

$$\frac{1}{f} = \frac{1}{h_m} + \frac{1}{h'_m} \quad (2)$$

where  $f$  is the main-lens focal length. After rearranging Eqs. (1) and (2), the relationship between the point  $P$  coordinates and the imaging position is given by

$$\frac{S - x_c}{z_c} + \frac{S - x_m}{h'_m} = \frac{S}{f} \quad (3)$$

In addition, according to the similar triangles, the relationship between intersection point  $P_m$  and the imaging pixel in the element image is expressed as

$$S = q \frac{h'_m}{b} (u - u_0) \quad (4)$$

where  $u_0$  is the  $u$ -coordinate center of element image as shown in Fig. 2(b), and  $q$  is the pixel width in the image plane. After rearranging Eqs. (3) and (4), the relationship between the imaging position in the micro-lens array plane and the certain point  $P$  is derived, as expressed in Eq. (5),

$$x = \frac{qh_m^2}{bd} \times \left[ \frac{1}{z_c} - \frac{1}{h_m} \right] (u - u_0) - \frac{x_c h'_m}{z_c d} \quad (5)$$

where  $x_0$  is the  $x$ -coordinate center of the micro-lens array plane as shown in Fig. 2(c). The relationship between  $x$  and  $x_m$  is  $x_m = (x - x_0) \cdot d$ , and  $d$  is pitch of micro-lens. The Eq. (5) describes a line corresponding to the certain point  $P(x_c, y_c, z_c)$  on the epipolar-plane image  $I_{v^*y^*}(u, x)$ . As illustrated in Eq. (5), when the angular coordinate  $u$  changes, the spatial coordinate  $x$  changes according to Eq. (5). The slope of the line in Eq. (5) is related to the depth of the certain point  $P$  and is independent of the coordinate  $x_c$ , which is the basis of depth

estimation described in many works previously. The  $x$ -intercept term of the line in Eq. (5) is determined by  $x_c$  and  $z_c$  of the point  $P$ , but has no relation to the coordinate  $y_c$  completely. Subsequently, the similar linear equation is also deduced in EPI  $I_{u^*x^*}(v, y)$ , as expressed in Eq. (6).

$$y = \frac{qh_m'^2}{bd} \times \left[ \frac{1}{z_c} - \frac{1}{h_m} \right] (v - v_0) - \frac{y_c h_m'}{z_c d} + y_0 \quad (6)$$

where  $v_0$  and  $y_0$  are the centers of element image and micro-lens array plane, respectively, which are similar to those in Eq. (5). Obviously, the slopes in Eqs. (5) and (6) corresponding to the certain point  $P(x_c, y_c, z_c)$  are equal, no matter which EPI is in consideration. On the other hand, although both intercepts are related to the depth  $z_c$  of the certain point, the intercepts in different EPI is related to only one coordinate, *i.e.*,  $x$ -coordinate or  $y$ -coordinate.

When both lines in the epipolar-plane images  $I_{v^*y^*}(u, x)$  and  $I_{u^*x^*}(v, y)$  are taken into consideration, there are three parameters corresponding to a certain point in the 3D scene, a slope and two intercepts, as the slopes in different EPIs are equal. In other words, for a certain 3D point  $P(x_c, y_c, z_c)$ , a 3-parameter vector is determined, which belongs to two lines corresponding to the certain point  $P(x_c, y_c, z_c)$  in two epipolar-plane images. In this paper, the space constituted by the 3-parameter vectors is termed epipolar-space. Therefore, a 3D point  $P(x_c, y_c, z_c)$  in the scene determines a corresponding 3D point in the epipolar-space. The relationship is described in Eq. (7).

$$\begin{bmatrix} K \\ B_x \\ B_y \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} 0 & 0 & -\frac{qh_m'^2}{bdh_m} & \frac{qh_m'^2}{bd} \\ -\frac{h_m'}{d} & 0 & x_0 & 0 \\ 0 & -\frac{h_m'}{d} & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (7)$$

where  $[x_c, y_c, z_c, 1]^T$  is the homogeneous coordinates of the 3D measurement point in the scene,  $K$  is the slope of the lines in epipolar-plane images  $I_{v^*y^*}(u, x)$  and  $I_{u^*x^*}(v, y)$ , which is also considered as the slope in traditional epipolar-plane theory,  $B_x$  is the intercept of the line in  $I_{v^*y^*}(u, x)$  where the horizontal axis  $x$  is 0, and the vertical axis  $u$  is  $u_0$ ,  $B_y$  is the intercept of the line in  $I_{u^*x^*}(v, y)$ , where the horizontal axis  $y$  is 0, and the vertical axis  $v$  is  $v_0$ , as expressed in Eq. (8).

$$\begin{aligned} K &= \frac{qh_m'^2}{bd} \times \left[ \frac{1}{z_c} - \frac{1}{h_m} \right] \\ B_x &= x_0 - \frac{x_c h_m'}{z_c d} \\ B_y &= y_0 - \frac{y_c h_m'}{z_c d} \end{aligned} \quad (8)$$



In this paper,  $K$  is calculated from the extending epipolar-plane image [6,12, 14-17],  $B_x$ ,  $B_y$  are extracted from the center sub-aperture image, which means  $[K, B_x, B_y, 1]^T$  can be obtained from a single light field camera image.

Therefore, the lines in epipolar-plane images  $I_{v^*y^*}(u, x)$  and  $I_{u^*x^*}(v, y)$  are expressed as

$$x = K \cdot (u - u_0) + B_x$$

$$y = K \cdot (v - v_0) + B_y$$

Moreover, the 3D point in the world coordinate system is related to the 3D point in the camera coordinate system by a rigid transformation, with the rotation  $R$  and translation  $t$ , as expressed in Eq. (9).

$$\begin{bmatrix} K \\ B_x \\ B_y \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} 0 & 0 & -\frac{qh_m'^2}{bdh_m} & \frac{qh_m'^2}{bd} \\ -\frac{h_m'}{d} & 0 & x_0 & 0 \\ 0 & -\frac{h_m'}{d} & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ \bar{0} & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (9)$$

#### 4. Calibration method

The details how to effectively solve the light field camera calibration problem is provided in this section. We started with an analytical solution, followed by a nonlinear optimization based on the maximum likelihood criterion, where the lens distortion is taken into account.

##### 4.1 Linear initialization

Without loss of generality, the calibration board is assumed on the plane of  $Z = 0$  in the world coordinate system, and the feature point on the top-left corner is the origin of the coordinate system. Let  $[X, Y]^T$  denote a feature point on the calibration plane since  $Z$  is always equal to 0. From Eq. (9), we have:

$$\begin{bmatrix} K \\ B_x \\ B_y \\ 1 \end{bmatrix} = H_{4 \times 3} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{1}{z_c} M_1 M_2 \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (10)$$

where  $H$  is homographic matrix,  $M_1$  is intrinsic matrix and  $M_2$  is extrinsic matrix, expressed as:

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}$$

$$M_1 = \begin{bmatrix} 0 & 0 & m_{11} & m_{12} \\ m_{21} & 0 & m_{22} & 0 \\ & m_{31} & m_{32} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -\frac{qh_m'^2}{bdh_m} & \frac{qh_m'^2}{bd} \\ -\frac{h_m'}{d} & 0 & x_0 & 0 \\ 0 & -\frac{h_m'}{d} & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} r_1 & r_2 & t \\ 0 & 0 & 1 \end{bmatrix}$$

As  $r_1$  and  $r_2$  are orthonormal, there are:

$$h_1^T M_1^{-T} M_1^{-1} h_1 = h_2^T M_1^{-T} M_1^{-1} h_2 \quad (11)$$

$$h_1^T M_1^{-T} M_1^{-1} h_2 = 0 \quad (12)$$

Let

$$B = M_1^{-T} M_1^{-1} = \begin{bmatrix} B_{11} & 0 & 0 & B_{12} \\ 0 & B_{21} & 0 & B_{22} \\ 0 & 0 & B_{31} & B_{32} \\ B_{12} & B_{22} & B_{32} & B_{41} \end{bmatrix}$$

$$= \begin{bmatrix} (\frac{1}{m_{12}})^2 & 0 & 0 & -\frac{m_{11}}{(m_{12})^2} \\ 0 & (\frac{1}{m_{21}})^2 & 0 & -\frac{m_{22}}{(m_{21})^2} \\ 0 & 0 & (\frac{1}{m_{31}})^2 & -\frac{m_{32}}{(m_{31})^2} \\ -\frac{m_{11}}{(m_{12})^2} & -\frac{m_{22}}{(m_{21})^2} & -\frac{m_{32}}{(m_{31})^2} & 1 + \text{sigma} \end{bmatrix} \quad (13)$$

$$\text{sigma} = (\frac{m_{11}}{m_{12}})^2 + (\frac{m_{22}}{m_{21}})^2 + (\frac{m_{32}}{m_{31}})^2$$

Note that  $B$  is symmetric, defined by a 7D vector:

$$\mathbf{b} = [B_{11} \quad B_{12} \quad B_{21} \quad B_{22} \quad B_{31} \quad B_{32} \quad B_{41}]^T$$

Let the  $i^{\text{th}}$  column vector of  $H$  be  $h_i = [h_{i1}, h_{i2}, h_{i3}, h_{i4}]^T$ , there are:

$$h_i^T B h_j = v_{ij}^T \mathbf{b}$$

where

$$v_{ij} = [h_{i1}h_{j1} \quad h_{i4}h_{j1} + h_{i1}h_{j4} \quad h_{i2}h_{j2} \quad h_{i4}h_{j2} + h_{i2}h_{j4} \quad h_{i3}h_{j3} \quad h_{i4}h_{j3} + h_{i3}h_{j4} \quad h_{i4}h_{j4}]$$



Therefore, Eqs. (12) and (13), from a given homography, can be written as 2 homogeneous equations in  $\mathbf{b}$ :

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} \mathbf{b} = 0 \quad (14)$$

When  $N$  images of the calibration board are observed in  $N$  poses, by stacking  $N$  such equation as Eq. (14), we have:

$$\mathbf{V}\mathbf{b} = 0 \quad (15)$$

where  $\mathbf{V}$  is a  $2N \times 6$  matrix. Once  $\mathbf{b}$  is determined, it is easy to compute camera intrinsic matrix  $M_1$  using Cholesky factorization [13]. Furthermore, the extrinsic parameters are determined by the following Eq. (16)

$$\begin{aligned} z_c &= 1 / \|M_1^{-1}h_1\| \\ [m_1 \quad m_2 \quad m_3]r_1 &= z_c h_1 \\ [m_1 \quad m_2 \quad m_3]r_2 &= z_c h_2 \\ r_3 &= r_1 \times r_2 \\ [m_1 \quad m_2 \quad m_3]t &= z_c h_2 - m_4 \end{aligned} \quad (16)$$

where

$$[m_1 \quad m_2 \quad m_3 \quad m_4] = M_1$$

If  $N \geq 4$ , we will have in general a unique solution  $\mathbf{b}$  defined up to a scale factor, which are the initial value of nonlinear optimization.

#### 4.2 Nonlinear optimization

There are many ways leading to ray distortion in the light field camera simultaneously, such as radial distortion of the main-lens, the mismatching between the imaging sensor and MLA, *et al.* To improve calibration and 3D shape measurement accuracy, the nonlinear optimization is necessary. Generally, the radial distortion of the main-lens is the most common distortion of light field camera, the distortion generated by the MLA is ignored in this paper as the structure of micro-lens is ideal approximately, which is also ignored as described in [1]. The undistorted coordinate  $(\bar{x}, \bar{y})^T$  is computed from the distorted coordinate  $(x, y)^T$  in the central sub-aperture image coordinate.

$$\begin{aligned} \bar{x} &= x + (x - x_0)[k_1(x_m^2 + y_m^2) + k_2(x_m^2 + y_m^2)^2] \\ \bar{y} &= y + (y - y_0)[k_1(x_m^2 + y_m^2) + k_2(x_m^2 + y_m^2)^2] \end{aligned} \quad (17)$$

where  $(k_1, k_2)$  denotes distortion vector. To accomplish the nonlinear optimization, the following cost function is minimized with the initial values solved in Section 4.1 to refine the parameters, including the intrinsic matrix  $M_1$ , distortion vector  $(k_1, k_2)$  and the extrinsic matrix  $R_i, t_i$ ,  $i = 1, 2, \dots, N$ .  $N$  is the number of pose.

$$\arg \min \sum_{u=1}^{N_u} \sum_{v=1}^{N_v} \sum_{i=1}^N \sum_{j=1}^M E_{i,j}^{u,v}(M_1, k_1, k_2, R_i, t_i) \quad (18)$$

where  $E$  is the Euclidean distance from observed to expected projected feature locations,  $N_u$  is the maximum number of sub aperture images in horizontal direction,  $N_v$  is the maximum number of sub aperture images in vertical direction,  $M$  is the number of feature points on calibration board.

In this paper, this nonlinear minimization problem is solved with the Levenberg-Marquardt algorithm, and the “*lsqnonlin*” function in MATLAB is adopted to accomplish the optimization.

## 5. Experimental results

The light field camera, Lytro Illum, is used to verify the proposed calibration and 3D shape measurement methods, as shown in Fig. 5. After light field decoding, the 4D light field contains  $15 \times 15$  array of sub-aperture images with  $625 \times 434$  pixels. In experiment, a calibration board with circular patterns was captured at  $M = 13$  different perspectives, which is about 600mm away from camera, as shown in Fig. 5. Circular centers are considered as feature points, the nominal distance between adjacent circular centers of the calibration board is 30.00 mm with the error of  $\pm 0.005$ mm. The 35 mm equivalent focal length of the main lens is 30 mm, the pixel size  $q$  of the image plane is 0.0014 mm, and the distance  $d$  between adjacent micro-lens after decoding is 0.01732 mm, which are obtained from the metadata provided by Lytro. The calibration results of the light field camera are detailed in Table 2, and the position of the calibration board at  $M$  perspectives are shown in Fig. 6.



Fig. 5. Light field camera and calibration board.

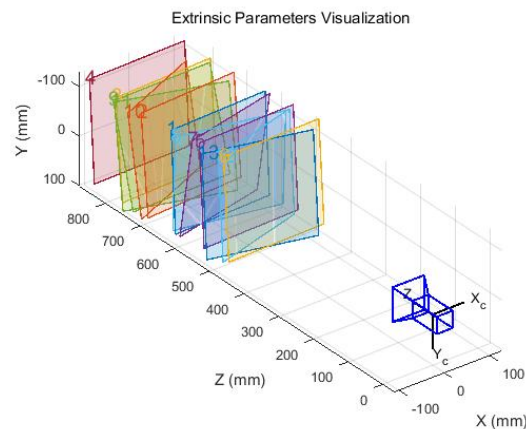


Fig. 6. Calibration results of the light field camera and the calibration board.

Table 2. Parameters of the light field camera before and after optimization.

Parameters	Initial Values	Optimized Values
------------	----------------	------------------

Intrinsic parameters	$m_{11}$	-0.1448	-0.1689
	$m_{12}$	0.3284	0.2792
	$m_{21}$	-759.4334	-763.9315
	$m_{21}$	315.5135	321.7423
	$m_{31}$	-752.6002	-755.8924
	$m_{32}$	200.0482	216.5287
Extrinsic parameters	$\mathbf{R} (\times 10^{-3} \text{ rad})$	(-0.0913, 0.1144, 0.6343)	(-0.0925, 0.1257, 0.6482)
	$\mathbf{t}(\text{mm})$	(2.2242, 2.1757, 0.01727)	(2.2573, 2.1496, 0.01881)
	$(k_1, k_2)$	(0.2271, -16.4822)	(2.7350, -16.3859)

The reprojection error of the proposed geometrical model before and after nonlinear optimization and distortion correction is shown in Fig. 7. Obviously, the reprojection error on the margin of the main lens is larger than that in the middle of the main lens due to main-lens distortion, as shown in the Fig. 7(a). After the nonlinear optimization and correction, the reprojection error is identical approximately, as shown in Fig. 7(b).

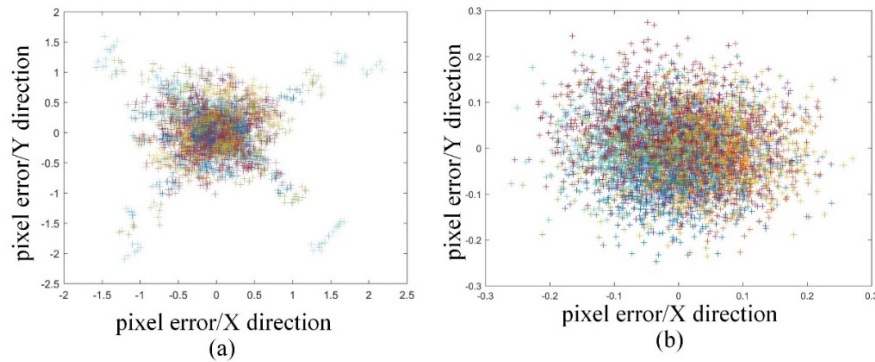


Fig. 7. Reprojection error. (a) without nonlinear optimization and distortion correction; (b) with nonlinear optimization and distortion correction.

To further verify the proposed calibration method, the light field camera is also calibrated with checkerboard images, which are contained in public data sets [20] (CVPR 2013 Plenoptic Calibration Data sets), and the calibration results are compared to that provided by Dansereau [1] *et al.* The experimental results show that the RMS reprojection error of the proposed calibration method is 0.0164mm, while the RMS reprojection error of Dansereau's calibration method is 0.0628mm. The proposed method achieves better light field camera calibration.

To verify the performance of 3D shape measurement, the calibration board was reconstructed based on Eq. (8). The nominal distance between the adjacent feature points is treated as ground truth, by which the 3D shape measurement error is evaluated, as shown in Fig. 8. There are  $8 \times 8$  feature points that are reconstructed, so that there are 112 points totally in Fig. 8. The maximum error is less than 0.9mm, and the RMS is 0.3670mm.

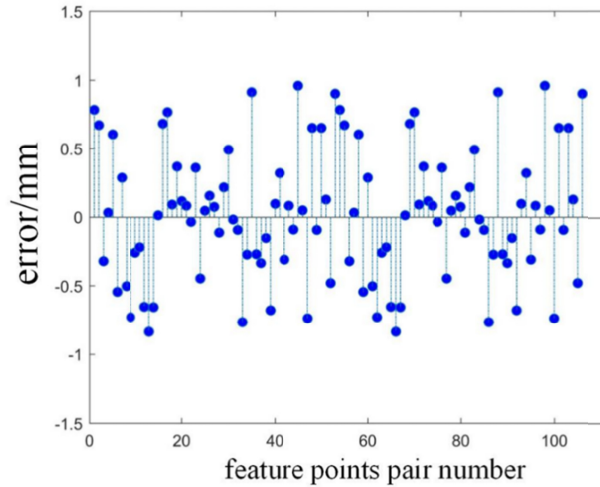


Fig. 8. Distance error between adjacent feature points.

Both reprojection error and 3D shape measurement error demonstrate that the calibration model and 3D shape measurement method based on epipolar-space work well. In addition, some light field camera data set from JPEG Pleno Database [18] was used to verify the 3D shape measurement proposed in this paper. The detailed information about the data set is illustrated in New Light Field Image Data set [19]. One raw data of the light field camera data set is shown in Fig. 9(a), where there are two men in different depth and background. To accomplish 3D shape measurement based on Eq. (8), the slope  $K$  and intercepts  $B_x$  and  $B_y$  that construct the epipolar-space were computed. Therefore, the 3D coordinates of the men and the background in the scene are derived and shown in Fig. 9(d). The objects with different depth are illustrated with different color in Fig. 9(d) to be distinguished easily. In the 3D shape measurement results, the detail of the faces and the distance between the two men are reconstructed accurately, which demonstrate that our method could reconstruct the object well. However, the hair of the front man is reconstructed to a plane nearly, and the hair of the second man is reconstructed to some irregular shapes. The reason is that the region of hair is textureless, which is challenging to all kinds of passive methods and is considered in our future works. Some 3D shape measurement results are also shown in Fig. 9. Two figures in the light field camera data set are shown in Figs. 9(b) and 9(c), and corresponding 3D shape measurement results are shown in Figs. 9(e) and 9(f). The viewpoint in this figure is somewhat deceptive, because the coordinate system is rotated artificially to make the depth clear and easy to sense.

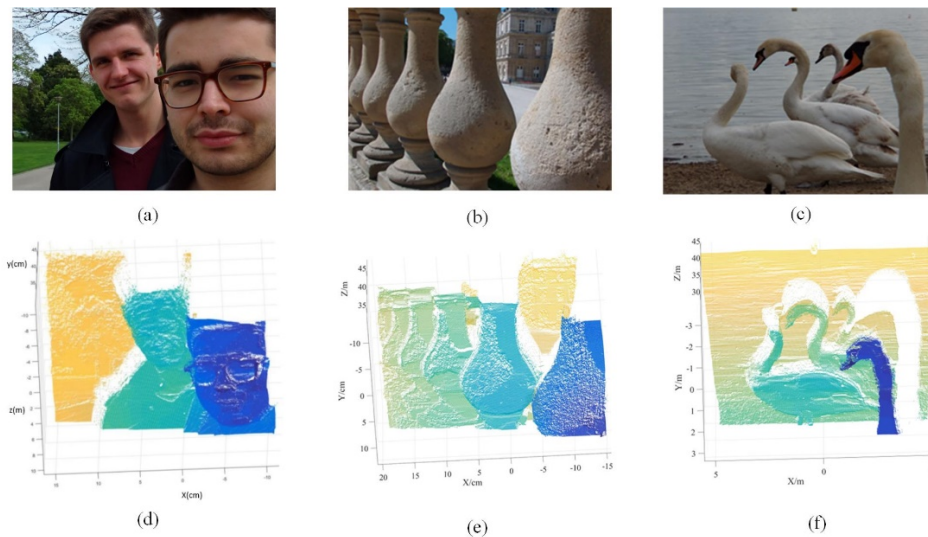


Fig. 9. 3D shape measurement results. Top: raw images from JPEG Pleno Database [18]; bottom: corresponding 3D measurement results of top images.

## 6. Conclusion

Compared to the traditional camera that captures 2D images, the MLA-based light field camera enables a single camera to record 4D light field. For light field applications, the light field camera calibration is necessary for 3D shape measurement. The previous works indicated that the 3D point depth is inversely proportional to the slope of its corresponding line on the EPI. In this paper, we deduced that the slopes of both lines corresponding to a certain point in EPIs are equal and the intercept is dependent on only  $x_c$  or  $y_c$  of the point except for the depth. Therefore, we proposed an epipolar-space based light field geometrical model, which determined the relationship between the 3D point to be measured and the corresponding 3-parameter vector in the epipolar-space, instead of the 4D light field description. Moreover, the coordinates of the 3D point were deduced with a close-form solution based on the 3-parameter vector. Experimental demonstration has verified that the proposed light field geometrical model is suitable for light field camera calibration and has the potential to accomplish 3D shape measurement. Future works will focus on the slope estimation method to improve the accuracy of the light field calibration and 3D shape measurement.

## Funding

National Nature Science Foundation of China (NSFC) (61771130).

## References

1. D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras," in *Proceedings of IEEE International Conference on Computer Vision*, IEEE, 1027–1034, 2013.
2. Q. Zhang, C. Zhang, J. Ling, Q. Wang, and J. Yu, "A Generic Multi-Projection-Center Model and Calibration Method for Light Field Cameras," *IEEE Trans. Pattern Anal. Mach. Intell.* **8430574**, 1 (2018).
3. Y. Bok, H. G. Jeon, and I. S. Kweon, "Geometric Calibration of Micro-Lens-Based Light Field Cameras Using Line Features," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(2), 287–300 (2017).
4. Z. Cai, X. Liu, X. Peng, and B. Z. Gao, "Ray calibration and phase mapping for structured-light-field 3D reconstruction," *Opt. Express* **26**(6), 7598–7613 (2018).
5. E. M. Hall, T. W. Fahringer, D. R. Guildenbecher, and B. S. Thurow, "Volumetric calibration of a plenoptic camera," *Appl. Opt.* **57**(4), 914–923 (2018).
6. S. Wanner and B. Goldluecke, "Globally consistent depth labelling of 4D light fields," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 41–48, (2012).

7. N. Zeller, F. Quint, and U. Stilla, "Depth estimation and camera calibration of a focused plenoptic camera for visual odometry," *ISPRS J. Photogramm. Remote Sens.* **118**, 83–100 (2016).
8. I. K. Williem, I. K. Park, and K. M. Lee, "Robust light field depth estimation using occlusion-noise aware data costs," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(10), 2484–2497 (2018).
9. E. H. Adelson, J. R. Bergen, "The plenoptic function and the elements of early vision" *Computational Models of Visual Processing*, M. Landy and J. A. Movshon, eds. (MIT Press, 1991), 3–20.
10. M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, 31–42, (1996).
11. R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, **2**(11), 1–11 (2005).
12. P. Yang, Z. Wang, Y. Yan, W. Qu, H. Zhao, A. Asundi, and L. Yan, "Close-range photogrammetry with light field camera: from disparity map to absolute distance," *Appl. Opt.* **55**(27), 7477–7486 (2016).
13. Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
14. Y. Zhang, H. Lv, Y. Liu, H. Wang, X. Wang, Q. Huang, X. Xiang, and Q. Dai, "Light-field depth estimation via epipolar plane image analysis and locally linear embedding," *IEEE Trans. Circ. Syst. Video Tech.* **27**(4), 739–747 (2017).
15. G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, and Y. Liu, "Light field image processing: an overview," *IEEE J. Sel. Top. Signal Process.* **11** (7), 926–954 (2017).
16. S. Heber, W. Yu, and T. Pock, "Neural EPI-volume networks for shape from light field," in *Proceeding of IEEE International Conference on Computer Vision*, IEEE, 2271–2279, (2017).
17. L. Su, Q. Yan, J. Cao, and Y. Yuan, "Calibrating the orientation between a microlens array and a sensor based on projective geometry," *Opt. Lasers Eng.* **82**, 22–27 (2016).
18. "JPEG Pleno Database: EPFL Light-field data set," <https://jpeg.org/plenodb/lf/epfl/>
19. M. Řeřábek and T. Ebrahimi, "New light field image dataset," in *Proceedings of the 8th International Workshop on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal, 218363, (2016).
20. <http://marine.acfr.usyd.edu.au/research/plenoptic-imaging/>