

Decoding and calibration method on focused plenoptic camera

Chunping Zhang¹ (✉), Zhe Ji¹, and Qing Wang¹

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The ability of light gathering of plenoptic camera opens up new opportunities for a wide range of computer vision applications. An efficient and accurate method to calibrate plenoptic camera is crucial for its development. This paper describes a 10-intrinsic-parameter model for focused plenoptic camera with misalignment. By exploiting the relationship between the raw image features and the depth–scale information in the scene, we propose to estimate the intrinsic parameters from raw images directly, with a parallel biplanar board which provides depth prior. The proposed method enables an accurate decoding of light field on both angular and positional information, and guarantees a unique solution for the 10 intrinsic parameters in geometry. Experiments on both simulation and real scene data validate the performance of the proposed calibration method.

Keywords calibration; focused plenoptic camera; depth prior; intrinsic parameters

1 Introduction

The light field cameras, including plenoptic camera designed by Ng [1, 2] and focused plenoptic camera designed by Georgiev [3–5], capture both angular and spatial information of rays in space. With the micro-lens array between image sensor and main lens, the rays from the same point in the scene fall on different locations of image sensor. With a particular camera model, the 2D raw image can be decoded into a 4D light field [6, 7], which allows applications on refocusing, multiview imaging, depth estimation, and so on [1, 8–10]. To support the applications, an

accurate calibration method for light field camera is necessary.

Prior work in this area has dealt with the calibration of plenoptic camera and focused plenoptic camera by projecting images into the 3D world, but their camera models are still improvable. These methods make an assumption that the geometric center of micro-lens image lies on the optical axis of its corresponding micro-lens, and do not consider the constraints on the high-dimensional features of light fields. In this paper, we concentrate on the focused plenoptic camera and analyze the variance and invariance between the distribution of rays inside the camera and in real world scene, namely the relationship between the raw image features and the depth–scale information. We fully take into account the misalignment of the micro-lens array, and propose a 10-intrinsic-parameter light field camera model to relate the raw image and 4D light fields by ray tracing. Furthermore, to improve calibration accuracy, instead of a single-planar board, we design a parallel biplanar board to provide depth and scale priors. The method is verified on simulated data and a physical focused plenoptic camera. The effects of rendered images on different intrinsic parameters are compared.

In summary, our main contributions are listed as follows:

- (1) A full light field camera model taking into account the geometric relationship between the center of micro-lens image and the optical center of micro-lens, which is ignored in most literature.
- (2) A loop-locked algorithm which is capable of exploiting the 3D scene prior for estimating the intrinsic parameters in one shoot with good stability and low computational complexity.

The remainder of this paper is organized as follows. Section 2 summarizes related work on

¹ School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China. E-mail: C. Zhang, 724495506@qq.com (✉); Z. Ji, 1277440141@qq.com; Q. Wang, qwang@nwpu.edu.cn.

Manuscript received: 2015-12-01; accepted: 2016-01-13

light field camera models, decoding and calibration methods. Section 3 describes the ideal model for a traditional camera or a focused plenoptic camera, and presents three theorems we utilize for intrinsic parameter estimation. In Section 4, we propose a more complete model for a focused plenoptic camera. Section 5 presents our calibration algorithm. In Section 6, we evaluate our method on both simulation and real data. Finally, Section 7 concludes with summary and future work.

2 Related work

A light field camera captures light field in a single exposure. The 4D light field data is rearranged on a 2D image sensor in accordance with the optical design. Moreover, the distribution of raw image depends on the relative position of the focused point inside the camera and the optical center of the micro-lens, as shown in Fig. 1. Figure 1(a) shows the design of Ng's plenoptic camera, where the micro-lens array is on the image plane of the main lens and the rays from the focused point almost fall on the same micro-lens image. Figure 1(b) and Fig. 1(c) show the design of Georgiev's focused plenoptic camera with a micro-lens array focused on the image plane of main lens,

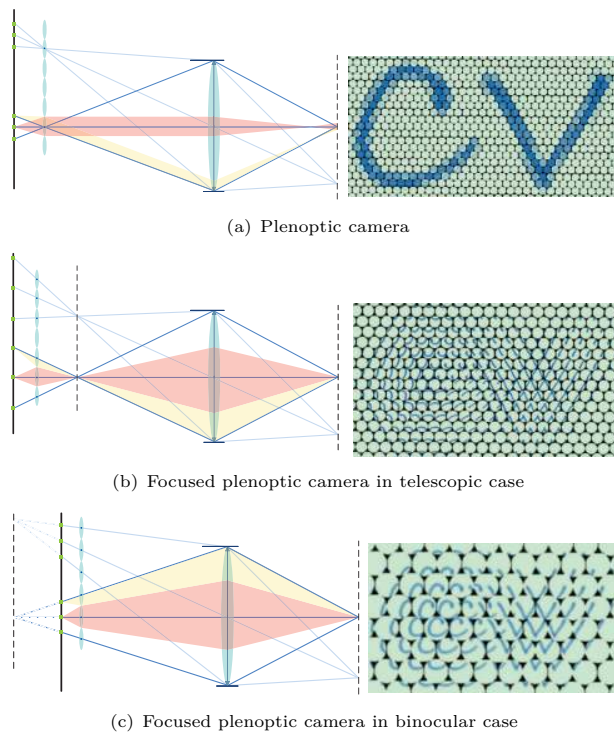


Fig. 1 Different designs of light field camera and raw images that consist of many micro-lens images closely packed.

and the rays from the focused point fall on different micro-lenses.

Decoding light field is equivalent to computing multiview images in two perpendicular directions. Multiview images are reorganized by selecting a contiguous set of pixels from each micro-lens image, for example, one pixel for plenoptic camera [2] and a patch for focused plenoptic camera [3,10]. However, for a focused plenoptic camera, the patch size influences the focus depth of the rendered image. Such decoding method causes discontinuity on out-of-focus area and results in artifact of aliasing.

For decoding a 2D raw image to a 4D light field representation, a common assumption is made that the center of each micro-lens image lies on the optical axis of its corresponding micro-lens [7, 11, 12] in ideal circumstances. Perwaß et al. [7] synthesized refocused images on different depths by searching pixels from multiple micro-lens images. Georgiev et al. [13] decoded into light field using ray transfer matrix analysis. Based on this assumption, the deviation in the ray's original direction has little effect on rendering a traditional image. However, the directions of decoded rays are crucial for an accurate estimation of camera intrinsic parameters, which is particularly important for absolute depth estimation [14] or light field reparameterization for cameras in different poses [15].

The calibration of a physical light field camera aims to decode rays more accurately. Several methods are proposed for the plenoptic camera. Dansereau et al. [6] presented a 15-parameter plenoptic camera model to relate pixels to rays in 3D space, which provides theoretical support for light field panorama [15]. The parameters are initialized using traditional camera calibration techniques. Bok et al. [16] formulated a geometric projection model to estimate intrinsic and extrinsic parameters by utilizing raw images directly, including analytical solution and non-linear optimization. Thomason et al. [17] concentrated on the misalignment of the micro-lens array and estimated its position and orientation. In this work, the directions of rays may deviate due to an inaccurate solution of the installation distances among main lens, micro-lens array, and image sensor. On the other hand, Johannsen et al. [12] estimated the intrinsic and extrinsic parameters

for a focused plenoptic camera by reconstructing a grid pattern from the raw image directly. The depth distortion caused by main lens was taken into account in their method. More importantly, expect for Ref. [17], these methods do not consider the deviation of the image center or the optical center for each micro-lens, which tends to cause inaccuracy in decoded light field.

3 The world in camera

The distribution of rays refracted by a camera lens is different from the original light field. In this section, we first discuss the corresponding relationship between the points in the scene and inside the camera modelled as a thin lens. Then we analyze the invariance in an ideal focused plenoptic camera, based on a thin lens and a pinhole model for the main lens and micro-lens respectively. Finally we conclude the relationship between the raw image features and the depth-scale information in the scene. Our analysis is conducted in the non-homogeneous coordinate system.

3.1 Thin lens model

As shown in Fig. 2, the rays emitted from the scene point $(x_{\text{obj}}, y_{\text{obj}}, z_{\text{obj}})^T$ in different directions are refracted through the lens aperture and brought to a single convergence point $(x_{\text{in}}, y_{\text{in}}, z_{\text{in}})^T$ if $z_{\text{obj}} > F$, where F denotes the focal length of the thin lens. The relationship between the two points is described as follows:

$$\frac{1}{|z_{\text{obj}}|} + \frac{1}{|z_{\text{in}}|} = \frac{1}{F} \quad (1)$$

$$\begin{pmatrix} x_{\text{in}} \\ y_{\text{in}} \\ z_{\text{in}} \end{pmatrix} = \frac{F}{F - z_{\text{obj}}} \begin{pmatrix} x_{\text{obj}} \\ y_{\text{obj}} \\ z_{\text{obj}} \end{pmatrix} \quad (2)$$

Equation (2) shows that the ratio on the

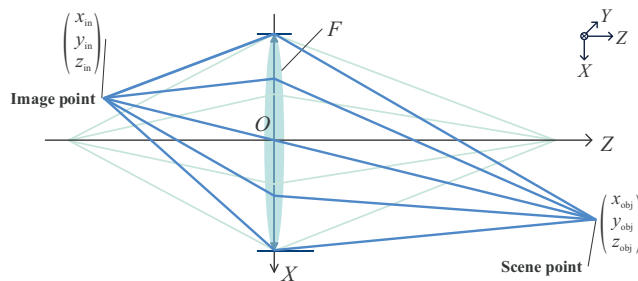


Fig. 2 Thin lens model.

coordinates of the two points changes with z_{obj} . Furthermore, there is a projective relationship between the coordinates inside and outside the camera. For example, as shown in Fig. 3, the objects with the same size in different depths in the scene correspond to the objects with different sizes inside the camera. The relationship can be described as

$$\frac{T^2}{S_1 S_2} = \frac{a_1 - a_2}{b_1 - b_2} \quad (3)$$

where the focal length F satisfies:

$$F = \frac{b_2 S_1 - b_1 S_2}{S_1 - S_2} \quad (4)$$

3.2 Ideal focused plenoptic camera model

As shown in Fig. 1, there are two optical designs of the focused cameras. In this paper, we only consider the design in Fig. 1(b). The case in Fig. 1(c) is similar to the former, only with the difference in the relative position of the focus point and the optical center of the micro-lens.

In this section, the main lens and the micro-lens array are described by a thin lens and a pinhole model respectively. As shown in Fig. 4, the main lens, the micro-lens array, and the image sensor are parallel to each other and all perpendicular to the optical axis. The optical center of the main lens lies on the optical axis.

Let d_{img} and d_{lens} be the distance between two geometric centers of arbitrary adjacent micro-lens images and the diameter of the micro-lens respectively, as shown in Fig. 4(a). The ratio between them is

$$\frac{d_{\text{lens}}}{d_{\text{img}}} = \frac{L}{L + l} \quad (5)$$

where L and l are the distances among the main lens, the micro-lens array, and the image sensor respectively. We can find that the ratio L/l is

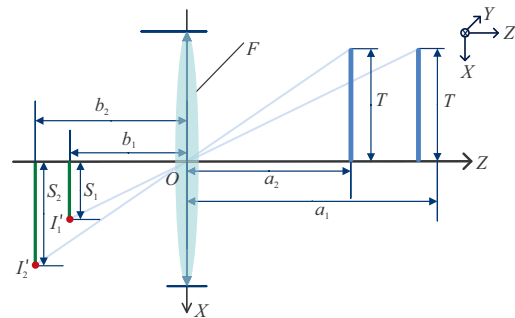


Fig. 3 Two objects with the same size of T in the scene at different depths focus inside a camera with focal length F .

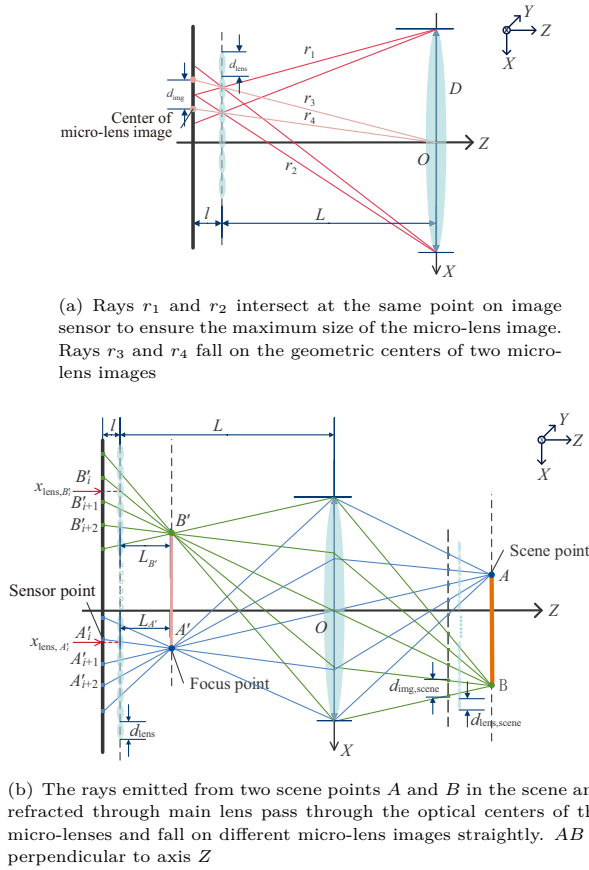


Fig. 4 Ideal focused plenoptic camera in telescopic case.

dependent on the raw image and the diameter of micro-lens, which is useful for our calibration model in Section 5. Moreover, there is a deviation between the optical center of micro-lens and the geometric center of its image, and d_{img} is constant in the same plenoptic camera.

Let $d_{\text{lens, scene}}$ and $d_{\text{img, scene}}$ be the size of micro-lens and its image refracted through the main lens into the scene respectively (Fig. 4(b)), combining Eqs. (2) and (5), the ratio between them satisfies:

$$\alpha = \frac{d_{\text{lens, scene}}}{d_{\text{img, scene}}} = \frac{L^2 + lL - LF}{L^2 + lL - LF - lF} \neq 1 \quad (6)$$

Equation (6) shows that though the rays are refracted through the main lens, the deviation between the geometric center of micro-lens image and the optical center of micro-lens still can not be ignored. The effect of deviations on the rendered images will be demonstrated and discussed in experiment.

In Fig. 4(b), A' and B' are the focus points of two scene points A and B respectively. The rays emitted from every focus point fall on multiple micro-lens

and focus on the image sensor, resulting in multiple images A'_i and B'_i . The distance between sensor points A'_i and A'_{i+1} is computed:

$$d_{A'} = |x'_{A_i} - x'_{A_{i+1}}| = d_{\text{lens}} \frac{L_{A'} + l}{L_{A'}} \quad (7)$$

where $L_{A'}$ is the distance between focus point A' and the micro-lens array, and $|\cdot|$ denotes the absolute operator. Equation (7) indicates that the distance between arbitrary two adjacent sensor points of the same focus point inside the camera is only dependent on intrinsic parameters. Once the raw image is shot (thus $d_{A'}$ is determined), $L_{A'}$ is only dependent on l and d_{lens} . According to triangle similarity, we can get the coordinate of the focus point:

$$x_{A'} = x_{\text{lens}, A'_i} - \frac{L_{A'}}{l} (x_{A'_i} - x_{\text{lens}, A'_i}) \quad (8)$$

Based on Eq. (7), we can simplify Eq. (8) as

$$x_{A'} = \frac{d_{A'}}{d_{A'} - d_{\text{lens}}} x_{\text{lens}, A'_i} - \frac{d_{\text{lens}}}{d_{A'} - d_{\text{lens}}} x_{A'_i} \quad (9)$$

According to Eq. (9), once a raw image is shot (thus $d_{A'}$ and $x_{A'}$ are determined) and d_{lens} is given, $x_{A'}$ and $y_{A'}$ can be calculated and they are independent on other intrinsic parameters. Furthermore, the length of AB can be calculated using only the raw image and d_{lens} .

Imaging that there are two objects with equal size in the scene, as shown in Fig. 3, the distance between the focus point and the micro-lens array can be calculated via Eq. (7). Replacing b_1 and b_2 in Eq. (4) and simplifying via Eqs. (5) and (7), we get the relationship:

$$F = L - \frac{S_1 L_{I'_2} - S_2 L_{I'_1}}{S_1 - S_2} \quad (10)$$

where S_1 , S_2 , $L_{I'_1}$, and $L_{I'_2}$ are dependent on only three factors, including the raw image, d_{lens} , and l . Equation (10) shows that the value of F can be calculated uniquely once the other intrinsic parameters are determined.

In the same manner, Eq. (3) can be simplified as

$$T = \frac{b_2 S_1 - b_1 S_2}{b_1 - b_2} \quad (11)$$

From Eq. (11), the size of an object in the scene is independent on l . The size of an object which we reconstruct from the raw image can not be taken as a cost function to constrain l .

In summary, given the coordinates of micro-lens and the raw image, three theorems can be concluded as follows:

(1) The size of a reconstructed object inside the camera and its distance to the micro-lens array are constant (Eq. (9)).

(2) The unique F can be determined by the prior of the scene (Eq. (10)).

(3) The size of the reconstructed object in the scene is constant with changing L (Eq. (11)).

4 Micro-lens-based camera model

In this section we present a more complete model for a focused plenoptic camera with misalignment of the micro-lens array [17], which is capable of decoding more accurate light field. There are 10 intrinsic parameters totally to be presented in this section, including the distance between the main lens and the micro-lens array, L , the distance between the micro-lens array and the image sensor, l , the misalignment of micro-lens array, $x_m, y_m, (\theta, \beta, \gamma)$, the focal length of the main lens, F , and the shift of image coordinate, (u_0, v_0) .

4.1 Distribution of micro-lens image

As shown in Fig. 5(a), every micro-lens with its unique coordinate $(x_i, y_i, 0)^T$ is tangent with each other. In addition, $(x_i, y_i, 0)^T$ is only dependent on d_{lens} . To simplify the discussion, we assume the layout of the micro-lens array is square-like. For hexagon-like configuration, it is easy to partition the whole array into two square-like ones. With the transformation shown in Fig. 5(b), the coordinate of the optical center of the micro-lens is represented as

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \mathbf{R} \begin{pmatrix} x_i \\ y_i \\ 0 \end{pmatrix} + \mathbf{t} \quad (12)$$

where $\mathbf{t} = (x_m, y_m, L)^T$ and \mathbf{R} is the rotation matrix with three degrees of freedom, i.e., the rotations (θ, β, γ) about three coordinate axes, which are similar to the traditional camera calibration model [18].

Although the main lens and the image sensor are parallel, the case between the micro-lens array and the image sensor is not similar (Fig. 5(c)). Each geometric center of the micro-lens image is represented as

$$\begin{pmatrix} x'_c \\ y'_c \\ L+l \end{pmatrix} = \frac{L+l}{z_c} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \quad (13)$$

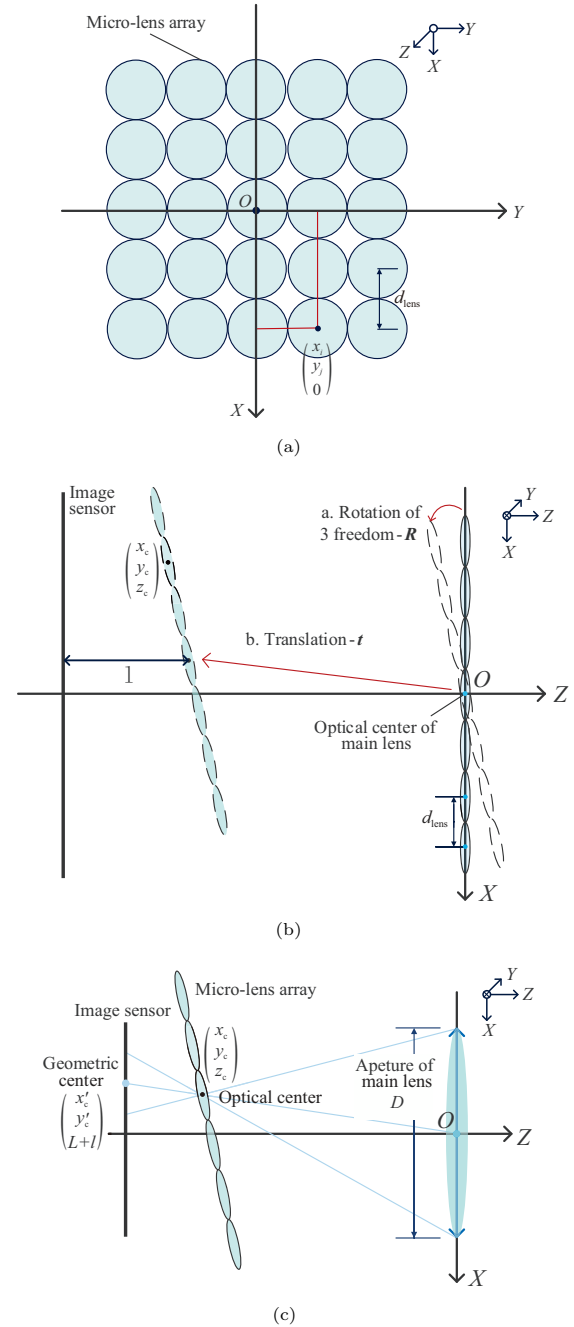


Fig. 5 The coordinate system of a focused plenoptic camera.

4.2 Projections from the raw image

Once the coordinate of a micro-lens's optical center $(x_c, y_c, z_c)^T$ and its image point $(x_{\text{img}}, y_{\text{img}}, L+l)^T$ are calculated, we can get a unique ray \mathbf{r}_i represented as

$$\mathbf{r}_i = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} + t_i \left(\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} - \begin{pmatrix} x_{\text{img}} \\ y_{\text{img}} \\ L+l \end{pmatrix} \right), \quad t_i \in \mathbf{R} \quad (14)$$

As shown in Fig. 4(b), the multiple images $\{(x_{A'_i}, y_{A'_i})^T | i = 1, \dots, n\}$ on the image sensor from the same focus point A' can be located if a proper pattern is shot, such as a grid-array pattern [12]. Thus the multiple rays emitted from point A' through different optical centers of the micro-lenses are collected to calculate the coordinate of point A' :

$$\hat{\mathbf{t}} = \underset{\mathbf{t}}{\operatorname{argmin}} \sum_{j=1}^n \left\| \mathbf{r}_j(t_j) - \frac{1}{n} \sum_{k=1}^n \mathbf{r}_k(t_k) \right\|_2 \quad (15)$$

$$\hat{A}' = \frac{1}{n} \sum_{i=1}^n \mathbf{r}_i(\hat{t}_i) \quad (16)$$

where $\|\cdot\|_2$ represents L_2 norm. Till now, we have accomplished the decoding process of light field inside the camera. To obtain the light field data in the scene, combining the depth-dependent scaling ratio described in Eq. (2), the representation of the focused points \hat{A}' can be transformed using the focal lens F easily.

5 Calibration

Compared to the ideal focused plenoptic camera model, the shift caused by the rotations of related micro-lenses is far less than l and the difference in the numerical calculation is trivial, therefore the three theorems concluded for an ideal focused plenoptic camera still hold for our proposed model with misalignment. More importantly, when there is zero machining error, the diameter of the micro-lens d_{lens} is set, and does not need to be estimated during the calibration. Consequently, the unique solution of the intrinsic parameters $\mathcal{P} = (\theta, \beta, \gamma, x_m, y_m, L, l, u_0, v_0)^T$ and F can be estimated using the two steps described in the following.

5.1 Decoding by micro-lens optical center

To locate the centers of the micro-lens images, we shoot a white scene [19, 20]. Then a template of proper size is cut out from the white image and its similarity with the original white image is calculated via normalized cross-correlation (NCC). To find the locations with subpixel accuracy, a threshold is placed on the similarity map such that all values less than 50% of the maximum intensity are set to zero. Then we take the filtered similarity map as weight and calculate the weighted coordinate of every small region. The results are shown in Fig. 6.

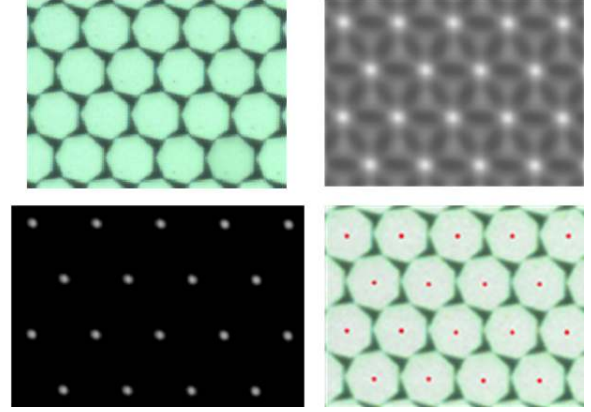


Fig. 6 The template (top-left), the crops of similarity map (top-right), the filtered similarity map (bottom-left), and the final location of the micro-lens image centers (bottom-right).

To estimate parameters \mathcal{P} , we minimize the cost function:

$$\hat{\mathcal{P}} = \underset{\mathcal{P}}{\operatorname{argmin}} \sum_{i,j} \left\| \begin{pmatrix} x'_{c,i} \\ y'_{c,i} \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} - \begin{pmatrix} \hat{x}'_{c,i} \\ \hat{y}'_{c,i} \end{pmatrix} \right\|_2 \quad (17)$$

where (u_0, v_0) is the offset between the camera coordinate and the image coordinate. After this optimization, \mathcal{P} is used to calculate micro-lens optical centers and reconstruct the calibration points. Then the rays are obtained via Eq. (14).

According to Eq. (5), the solution of Eq. (17), changing with the initial value of L , is not unique. Moreover, the ratio L/l is almost constant with changing initial value of \mathcal{P} . Although there are differences between the models described in Section 3.2 and Section 4, the theorems still hold since the shift caused by the rotations can be ignored. This observation will be verified in experiment later.

In addition, the value of l influences the direction of decoded rays. Due to the coupling relationship of angle and depth, either of them can be used as the prior to be introduced to estimate the unique \mathcal{P} .

5.2 Reconstruction of calibration points

To reconstruct a plane in the scene, we may shoot a certain pattern in order to recognize multiple images from different scene points. A crop of the calibration board and its raw image we shoot are shown in Fig. 7. To locate the multiple images of every point on the calibration board, we preprocess the grid image by adding the inverse color of the white image to the grid image (Fig. 7). Then one of the sensor points corresponding to the focus point A' , denoted by

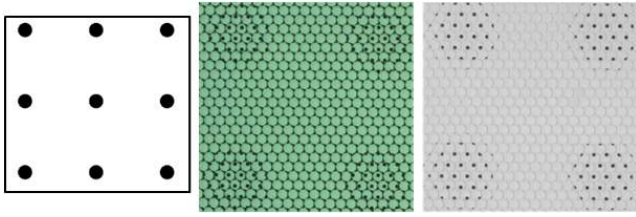


Fig. 7 A crop of calibration board, its raw image, and the preprocessed image by white image.

$(\hat{x}_{A'_i}, \hat{y}_{A'_i})^T$, is located by the same method described in Section 5.1. Consequently, the plane we shoot in the scene, denoted by $\hat{\Pi} = \{A_i | i = 1, \dots, n\}$, is easy to be reconstructed using Eqs. (2) and (14).

As shown in Fig. 8, we design a parallel biplanar board with known distance between the two parallel planes and the distance between adjacent grids, which can provide depth prior Pr_{dp} and scale prior Pr_{sc} . Equivalently, we can shoot a single-plane board twice while we move the camera on a guide rail to a fixed distance.

After the sensor point $(\hat{x}_{A'_i}, \hat{y}_{A'_i})^T$ of arbitrary scene point A is located and the intrinsic parameters \mathcal{P} are determined, we can reconstruct the grid-array plane $\hat{\Pi}_1$ and $\hat{\Pi}_2$ in the scene. Then the minimum distance of arbitrary point on the two calibration board planes can be calculated, referred as \hat{T}_1 and \hat{T}_2 respectively. Finally, we can minimize the cost function to estimate the focal length F of main lens:

$$\hat{F} = \argmin_F \left\| \hat{T}_1(F) - \hat{T}_2(F) \right\|_2, 0 < F < \max(z) \quad (18)$$

where \hat{T}_1 and \hat{T}_2 are only dependent on F in this step. According to Eq. (10), there is an optimal solution for Eq. (18) if \mathcal{P} is determined.

Note that if the values of L or l is incorrect, the distance between plane $\hat{\Pi}_1$ and $\hat{\Pi}_2$ is not equal to the prior distance. Therefore we take the distance

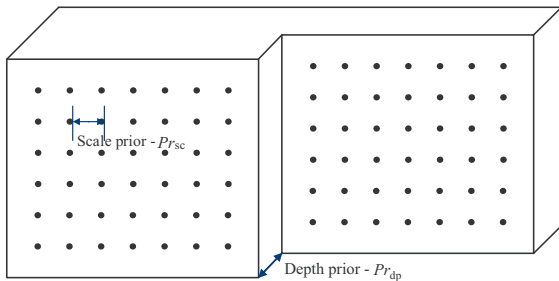


Fig. 8 The parallel biplanar board we designed to provide depth prior for calibration.

between plane $\hat{\Pi}_1$ and $\hat{\Pi}_2$ as the last cost function:

$$\hat{L} = \argmin_L \left\| dis(\hat{\Pi}_1, \hat{\Pi}_2) - Pr_{dp} \right\|_2, L > 0 \quad (19)$$

where $dis(\cdot, \cdot)$ represents the distance between two parallel planes. In practice, we take the mean distance of reconstructed points on $\hat{\Pi}_1$ to plane $\hat{\Pi}_2$ as the value of dis .

Moreover, \hat{T}_1 and \hat{T}_2 may not be equal to Pr_{sc} due to possible calculation error, so we must refine the value of depth prior to ensure the correct ratio of scale and depth.

5.3 Algorithm summary

The complete algorithm is summarized in Algorithm 1.

To make the algorithm more efficiently, the search

Algorithm 1: Calibration method for a focused camera with a parallel calibration board

Input:

Micro-lens images' centers

$\{(\hat{x}'_{c,i}, \hat{y}'_{c,j})^T | i = 1, \dots, p, j = 1, \dots, q\}$ extracted from a white image;

The diameter of micro-lens d_{lens} ;

Sensor points of P_1 extracted from grid image:

$\{(\hat{x}'_{1,i}, \hat{y}'_{1,i}, \hat{z}'_{1,i})^T | i = 1, \dots, m_1\}$;

Sensor points of P_2 extracted from grid image:

$\{(\hat{x}'_{2,i}, \hat{y}'_{2,i}, \hat{z}'_{2,i})^T | i = 1, \dots, m_2\}$;

The resolution of the image sensor $H * W$;

The installation parameter L_0, l_0 ;

Pr_{dp}, Pr_{sc} ;

Output:

$\mathcal{P} = (\theta, \beta, \gamma, x_m, y_m, L, l, u_0, v_0)^T$;

F .

Initialize:

$\mathcal{P}_0 = (0, 0, 0, 0, 0, L_0, l_0, H/2, W/2)^T$;

$cnt = 1$;

for $L_0 - searchRange$ to $L_0 + searchRange$ **do**

Optimize \mathcal{P} using Eq. (17);

Reconstruct focus points on $\hat{\Pi}'_1$ using Eqs. (15) and

(16): $\{(\hat{x}'_{1,i}, \hat{y}'_{1,i}, \hat{z}'_{1,i})^T | i = 1, \dots, n_1\}$;

Reconstruct focus points on $\hat{\Pi}'_2$ using Eqs. (15) and

(16): $\{(\hat{x}'_{2,i}, \hat{y}'_{2,i}, \hat{z}'_{2,i})^T | i = 1, \dots, n_2\}$;

Optimize \hat{F} using Eq. (18);

Reconstruct $\hat{\Pi}_1$ and $\hat{\Pi}_2$ using Eq. (2);

$distance(cnt) = dis(\hat{\Pi}_1, \hat{\Pi}_2)$;

$PSet(cnt) = \mathcal{P}$;

$FSet(cnt) = \hat{F}$;

$cnt = cnt + 1$;

end for

ind = the index of the closest value to repaired Pr_{dp} in $distance$;

$\mathcal{P} = PSet(ind)$;

$F = FSet(ind)$.

step of the loop of L should be changed with the value of $\|dis(\hat{\Pi}_1, \hat{\Pi}_2) - Pr_{dp}\|_2$ in Eq. (19). The same principle is applied to the search step of F . In addition, because of the monotonicity of $\|dis(\hat{\Pi}_1, \hat{\Pi}_2) - Pr_{dp}\|_2$ with L , and F with $\|\hat{T}_1(F) - \hat{T}_2(F)\|_2$, we can use dichotomy to search an accurate value more efficiently.

6 Experimental results

In experiments, we apply our calibration method on simulated and real world scene data. We capture three datasets of white images and grid images using a self-assembly focused plenoptic camera (Fig. 9). The camera includes a GigE camera with a CCD image sensor whose resolution is 4008×2672 pixels that are $9 \mu\text{m}$ wide, F-mount Nikon lens with 50 mm focal length, and a micro-lens array whose diameter is $300 \mu\text{m}$ with negligible error in hexagon layout.

We use the function “fminunc” in MATLAB to complete the non-linear optimization in Eqs. (15), (17), and (18). The initial parameters are set as the installation parameters, and $\theta, \beta, \gamma, x_m, y_m$ are set to zero.

6.1 Simulated data

First we verify the calibration method on simulated images rendered in MATLAB, as shown in Fig. 1. The ground truth and the calibrated parameters are shown in Table 1. We compare the estimated angle of the ray passing through each optical center of micro-



Fig. 9 The focused plenoptic camera we installed and its micro-lens array inside the camera.

Table 1 The parameters we estimated and the ground truth

Parameter	Ground truth	Calibration
θ ($^\circ$)	0.3000	0.2998
β ($^\circ$)	0.1500	0.1493
γ ($^\circ$)	0.1000	0.0997
(x_m, y_m) (mm)	0.7200, 0.6300	1.5030, -4.5009
(L, l) (mm)	67.3168, 3.3162	67.1861, 3.3096
(u_0, v_0) (pixel)	-1326.0, -2000.0	-1337.4, -1999.7
F (mm)	50.0000	50.02558

lens and the one of the main lens to the ground truth, which is shown in Fig. 10. The differences are less than 1.992×10^{-3} rad.

We compare the geometric centers of the micro-lens images we locate and the ones with optimization. The error maps of 84×107 geometric centers optimized with different L are shown in Fig. 11(a). From Fig. 11(b), we find that there are 96.53% of the centers whose error is less than 0.1 pixel, which is the input for the following projection step.

The comparison of the locations of optical centers of micro-lenses with different L is illustrated in Fig. 12. The difference in x -coordinate and y -coordinate of the optical center is trivial with changing L . The maximal difference is 4.2282×10^{-6} mm when L changes from 55 to 84 mm, which proves our observation mentioned in Section 5.1.

The values of F , $dis(\hat{\Pi}_1, \hat{\Pi}_2)$, \hat{S}_1 , \hat{S}_2 , \hat{T}_1 , and \hat{T}_2 are shown in Fig. 13. It is obvious that \hat{S}_1 and \hat{S}_2 are almost constant when L changes, proving the correctness made in Eqs. (9) and (11). In addition, the values of $dis(\hat{\Pi}_1, \hat{\Pi}_2)$ correlate linearly with L , which testifies the reasonability of the cost function described in Eq. (18). The relationship among \hat{T}_1 , \hat{T}_2 , and F is shown in Fig. 14, which proves the

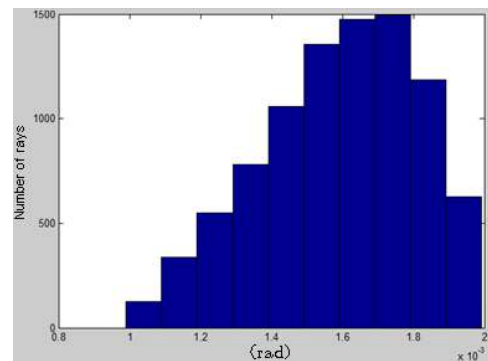


Fig. 10 The histogram of the deviation between the estimated angles of the rays and the ground truth.

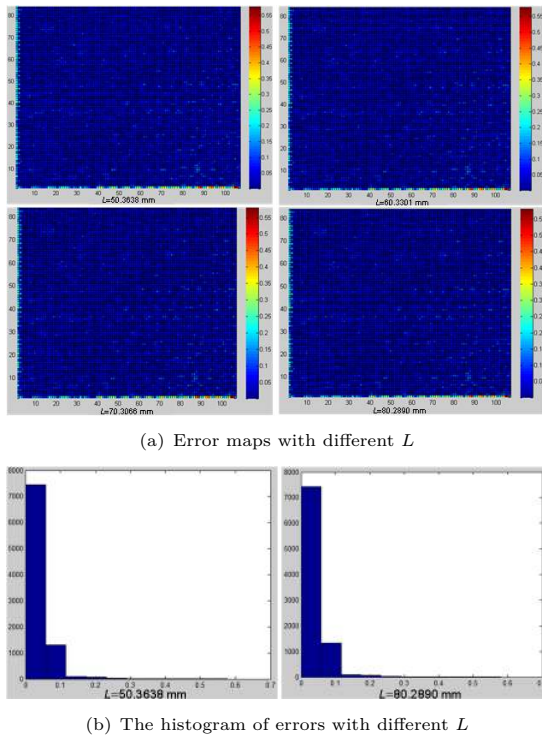


Fig. 11 The results of optimization on geometric centers of the micro-lens image on simulated data.

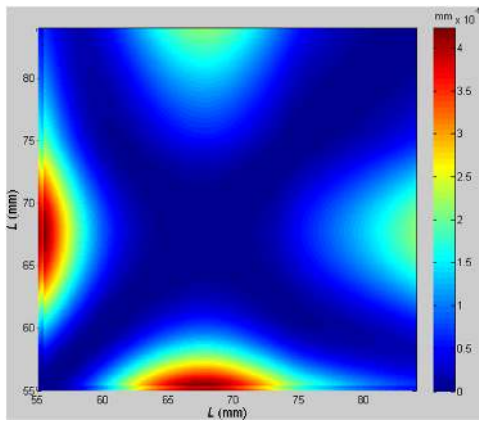


Fig. 12 The comparison of the locations of optical centers of micro-lenses with different L from 55 to 84 mm on simulated data. The value in row i and column j represents the difference in x -coordinate and y -coordinate between the results optimized in L_i and L_j .

analysis about Eq. (10).

6.2 Physical camera

Then we verify the calibration method on the physical focused plenoptic camera. To obtain the equivalent data of parallel biplanar board, we shoot a single-plane board twice while we move the camera on a guide rail to an accurate fixed distance, as shown in Fig. 9. The depth prior Pr_{dp} is precisely controlled to be 80.80 mm and the scale prior Pr_{sc}

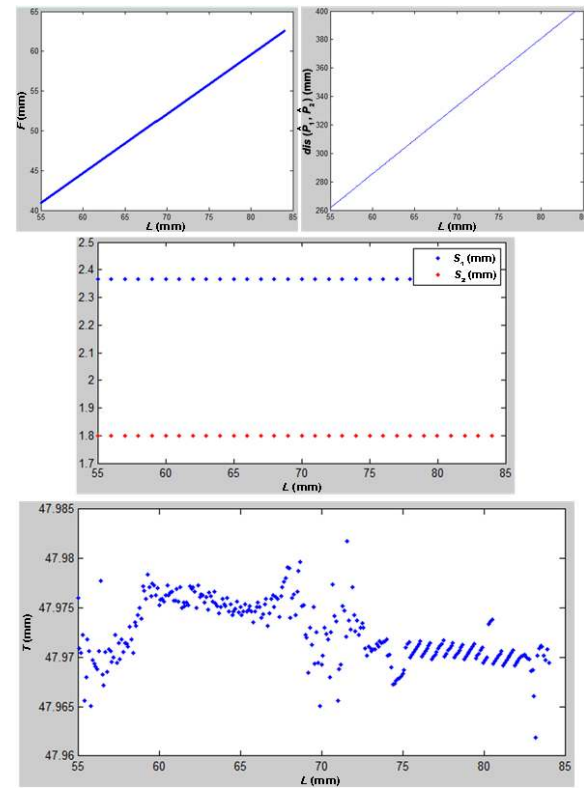


Fig. 13 The values of F , $dis(\hat{\Pi}_1, \hat{\Pi}_2)$, \hat{S}_1 , \hat{S}_2 , and \hat{T} ($\hat{T}_2 = \hat{T}_1$) with different L on simulated data.

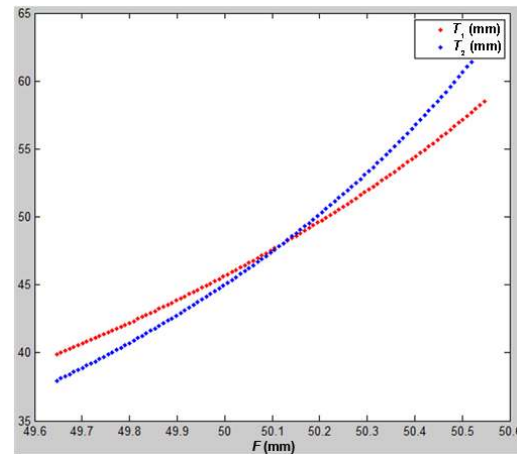
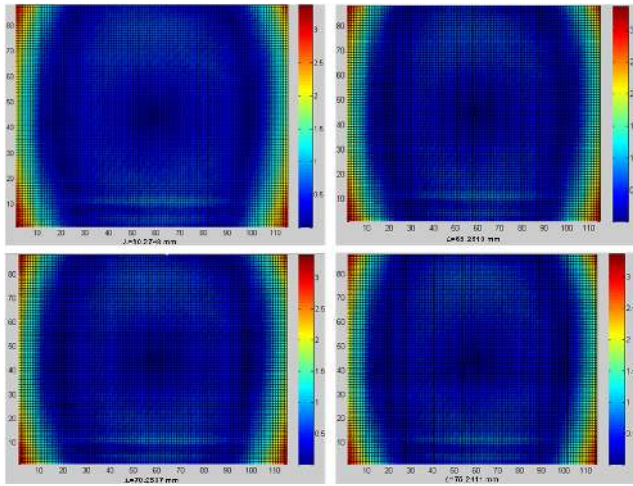
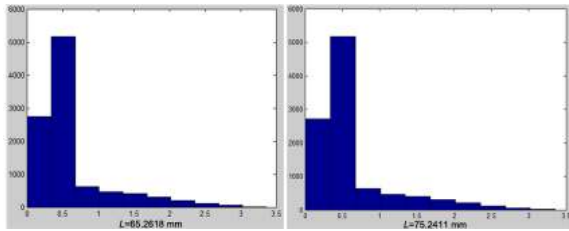


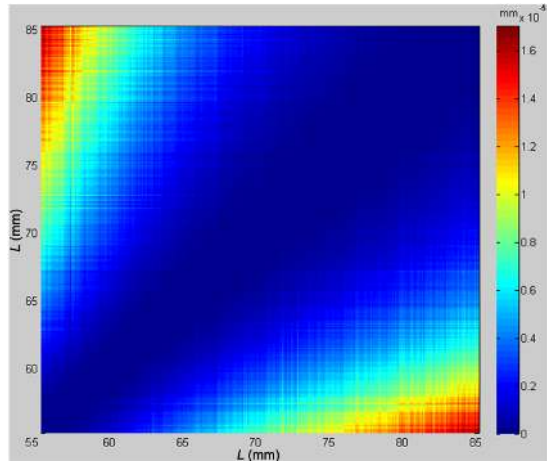
Fig. 14 The relationship of \hat{T}_1 , \hat{T}_2 , and F when $L = 67.3129$ mm.

is 28.57 mm. The calibration results are shown in Fig. 15.

As shown in Fig. 15(a), there is an obvious error between the computed geometric centers and the located centers on the edge of the error map, which may result from the distortion of lenses or the machining error of micro-lens. However, we find that there are 73.00% of the centers whose error is less than 0.6 pixel, as shown in Fig. 15(b). The

(a) Error maps with different L 

(b) The histogram of errors corresponding to (a)

(c) The comparison of the locations of optical centers of micro-lenses with different L from 55 to 85 mm on the physical camera. The value in row i and column j represents the difference in x -coordinate and y -coordinate between the results optimized in L_i and L_j .**Fig. 15** The results of optimization on geometric centers of micro-lens image on physical data.

mean difference of geometric centers of micro-lens images optimized with different L is 1.89×10^{-4} pixel (Fig. 15(c)). The results of F , $dis(\hat{\Pi}_1, \hat{\Pi}_2)$, \hat{S}_1 , \hat{S}_2 , \hat{T} ($\hat{T}_1 = \hat{T}_2$) with different L are similar to the results on simulated data.

Finally, to verify the stability of our algorithm, we calibrate intrinsic parameters with different poses of

calibration board. Corresponding results are shown in Table 2.

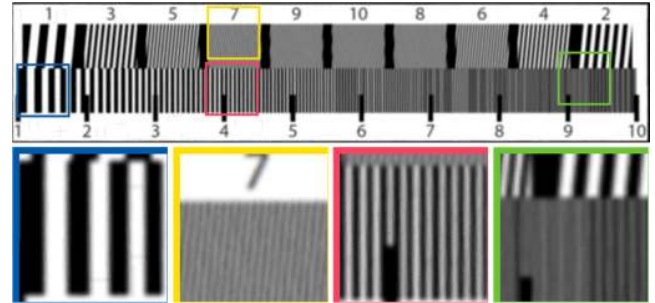
6.3 Rendering

We render the focused image with deviations between the optical center of micro-lens and the geometric center of micro-lens image.

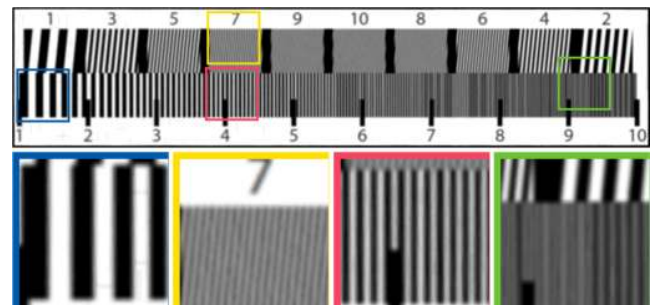
We shoot a resolution test chart on the same depth for simulated data (Fig. 16), which indicates that the deviation surely effects the accuracy of decoded light

Table 2 Parameters estimated with calibration board with different poses. The third parameter is the angle between the calibration board and the optical axis

Parameter	Dataset 1	Dataset 2	Dataset 3
Pr_{dp} (mm)	80.80	80.80	80.80
Pr_{sc} (mm)	28.57	28.57	28.57
$Angle$ ($^\circ$)	175.8386	151.9994	139.5982
θ ($^\circ$)	0.3978	0.3978	0.3978
β ($^\circ$)	-0.0616	-0.0615	-0.0616
γ ($^\circ$)	0.1377	0.1377	0.1378
x_m (mm)	0.0300	0.0299	0.0302
y_m (mm)	-0.0341	-0.0344	-0.0343
L (mm)	67.8059	67.7860	67.8109
l (mm)	2.1215	2.1209	2.1217
u_0 (pixel)	-12.0622	-12.0623	-12.0620
v_0 (pixel)	-17.9686	-17.9689	-17.9688
F (mm)	54.1801	53.9759	54.1841



(a) Rendering without deviation of the micro-lens' optical center and its image center



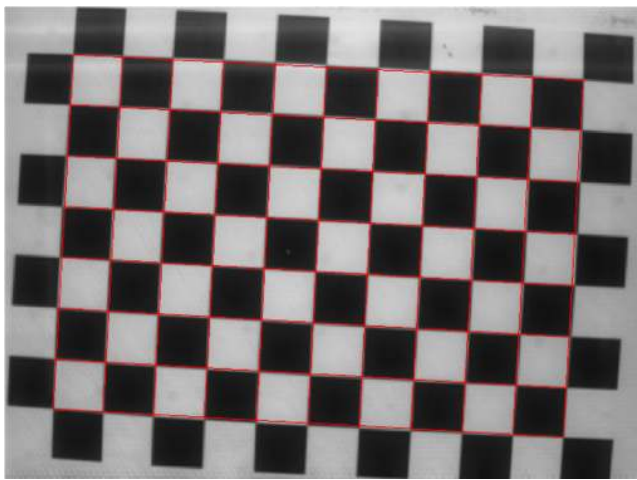
(b) Rendering with the calibrated result

Fig. 16 The rendered images from simulated data.

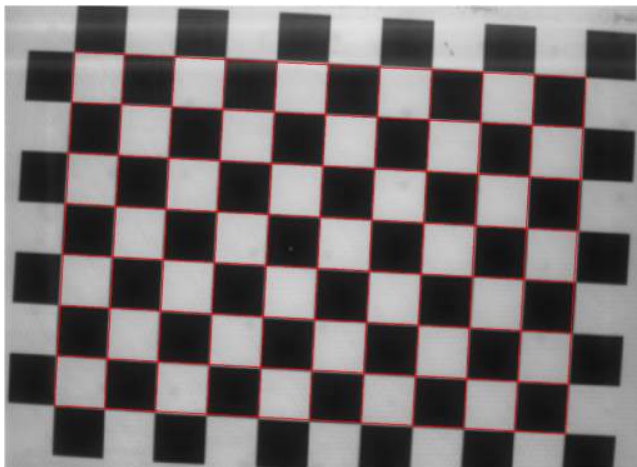
field. Then we shoot a chess board for simulated data to evaluate the width of every grid in the rendered images. We resize the images by setting the mean width of the grids to be 100 pixels. Then we calculate the range and the standard deviation of the grid width. The results are shown in Table 3, which indicates that the calibration contributes to the uniform scale in the same depth and reduces the distortion caused by incorrect deviations. The results on physical camera are shown in Table 4 and Fig. 17. The decoded light field with the estimated intrinsic parameters leads to more accurate refocus

Table 3 The range and variance of rendered chess board on simulated data

State	Std (pixel)	Range (pixel)
No deviation	0.053281	0.38697
Calibrated	0.040079	0.25109



(a) Rendering without deviation of the micro-lens' optical center and its image center



(b) Rendering with the calibrated result

Fig. 17 The image rendered from physical camera.

Table 4 The range and variance of rendered chess board on physical camera

State	Std (pixel)	Range (pixel)
No deviation	1.1145	0.042755
Calibrated	0.87151	0.034322

distance [14], which is equivalent to a correct ratio of scale and depth.

7 Conclusions and future work

In the paper we present a 10-intrinsic-parameter model to describe a focused plenoptic camera with misalignment. To estimate the intrinsic parameters, we propose a calibration method based on the relationship between the raw image features and the depth-scale information in the real world scene. To provide depth and scale priors to constrain the intrinsic parameters, we design a parallel biplanar board with grids. The calibration approach is evaluated on simulation as well as real data. Experimental results show that our proposed method is capable of decoding more accurate light field for the focused plenoptic camera.

Future work includes modelling the distortion caused by the micro-lens and main lens, optimization of extrinsic parameters, and the reparameterization of multiple and re-sampling light field data from cameras with different poses.

Acknowledgements

The work is supported by the National Natural Science Foundation of China (Nos. 61272287 and 61531014) and the research grant of State Key Laboratory of Virtual Reality Technology and Systems (No. BUAAVR-15KF-10).

References

- [1] Ng, R. Digital light field photography. Ph.D. Thesis. Stanford University, 2006.
- [2] Ng, R.; Levoy, M.; Bredif, M.; Duval, G.; Horowitz, M.; Hanrahan, P. Light field photography with a hand-held plenoptic camera. Stanford University Computer Science Tech Report CSTR 2005-02, 2005.
- [3] Georgiev, T. G.; Lumsdaine, A. Focused plenoptic camera and rendering. *Journal of Electronic Imaging* Vol. 19, No. 2, 021106, 2010.
- [4] Lumsdaine, A.; Georgiev, T. Full resolution lightfield rendering. Technical Report. Indiana University and Adobe Systems, 2008.

- [5] Lumsdaine, A.; Georgiev, T. The focused plenoptic camera. In: *Proceedings of IEEE International Conference on Computational Photography*, 1–8, 2009.
- [6] Dansereau, D. G.; Pizarro, O.; Williams, S. B. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1027–1034, 2013.
- [7] Perwaß, C.; Wietzke, L. Single lens 3D-camera with extended depth-of-field. In: *Proceedings of SPIE 8291, Human Vision and Electronic Imaging XVII*, 829108, 2012.
- [8] Bishop, T. E.; Favaro, P. Plenoptic depth estimation from multiple aliased views. In: *Proceedings of IEEE 12th International Conference on Computer Vision Workshops*, 1622–1629, 2009.
- [9] Levoy, M.; Hanrahan, P. Light field rendering. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, 31–42, 1996.
- [10] Wanner, S.; Fehr, J.; Jähne, B. Generating EPI representations of 4D light fields with a single lens focused plenoptic camera. In: *Lecture Notes in Computer Science, Vol. 6938*. Bebis, G.; Boyle, R.; Parvin, B. et al. Eds. Springer Berlin Heidelberg, 90–101, 2011.
- [11] Hahne, C.; Aggoun, A.; Haxha, S.; Velisavljevic, V.; Fernández, J. C. J. Light field geometry of a standard plenoptic camera. *Optics Express* Vol. 22, No. 22, 26659–26673, 2014.
- [12] Johannsen, O.; Heinze, C.; Goldluecke, B.; Perwaß, C. On the calibration of focused plenoptic cameras. In: *Lecture Notes in Computer Science, Vol. 8200*. Grzegorzec, M.; Theobalt, C.; Koch, R.; Kolb, A. Eds. Springer Berlin Heidelberg, 302–317, 2013.
- [13] Georgiev, T.; Lumsdaine, A.; Goma, S. Plenoptic principal planes. *Imaging Systems and Applications*, OSA Technical Digest (CD), paper JTuD3, 2011.
- [14] Hahne, C.; Aggoun, A.; Velisavljevic, V. The refocusing distance of a standard plenoptic photograph. In: *Proceedings of 3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video*, 1–4, 2015.
- [15] Birklbauer, C.; Bimber, O. Panorama light-field imaging. *Computer Graphics Forum* Vol. 33, No. 2, 43–52, 2014.
- [16] Bok, Y.; Jeon, H.-G.; Kweon, I. S. Geometric calibration of micro-lens-based light-field cameras using line features. In: *Lecture Notes in Computer Science, Vol. 8694*. Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T. Eds. Springer International Publishing, 47–61, 2014.
- [17] Thomason, C. M.; Thurow, B. S.; Fahringer, T. W. Calibration of a microlens array for a plenoptic camera. In: *Proceedings of the 52nd Aerospace Sciences Meeting, AIAA SciTech, AIAA 2014-0396*, 2014.
- [18] Zhang, Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 22, No. 11, 1330–1334, 2000.
- [19] Cho, D.; Lee, M.; Kim, S.; Tai, Y.-W. Modeling the calibration pipeline of the Lytro camera for high quality light-field image reconstruction. In: *Proceedings of IEEE International Conference on Computer Vision*, 3280–3287, 2013.
- [20] Sabater, N.; Drazic, V.; Seifi, M.; Sandri, G.; Perez, P. Light-field demultiplexing and disparity estimation. 2014. Available at <https://hal.archives-ouvertes.fr/hal-00925652/document>.



and light field computing theory and application.



field computing theory and application.



obtained his master and Ph.D. degrees from the Department of Computer Science and Engineering, Northwestern Polytechnical University, respectively. In 2006, he was awarded the Program for New Century Excellent Talents in University of Ministry of Education, China. He is now a member of IEEE and ACM. He is also a senior member of China Computer Federation (CCF).

He worked as research assistant and research scientist in the Department of Electronic and Information Engineering, the Hong Kong Polytechnic University from 1999 to 2002. He also worked as a visiting scholar at the School of Information Engineering, the University of Sydney, Australia, in 2003 and 2004. In 2009 and 2012, he visited the Human Computer Interaction Institute, Carnegie Mellon

University, for six months and the Department of Computer Science, University of Delaware, for one month, respectively.

Professor Wang's research interests include computer vision and computational photography, such as 3D structure and shape reconstruction, object detection, tracking and recognition in dynamic environment, and light field imaging and processing. He has published more than 100 papers in the international journals and conferences.

Open Access The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License ([http://](http://creativecommons.org/licenses/by/4.0/)

creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Other papers from this open access journal are available free of charge from <http://www.springer.com/journal/41095>. To submit a manuscript, please go to <https://www.editorialmanager.com/cvmj>.