Light field endoscope calibration based on virtual objective lens and virtual feature points

Ping Zhou,[a,*] **Cheng Gu,**[a] **Weizhong Zhang,**[b] **Zi Yang,**[a] **Yuting Zhang,**[a] **Weijia Cai,**[a] **and Guangquan Zhou**[a]

[a]Southeast University, School of Biological Science and Medical Engineering, Nanjing, China
[b]The First Affiliated Hospital of Nanjing Medical University, Department of Ophthalmology, Nanjing, China

Abstract. Three-dimensional light field imaging in laparoscopic surgery is an emerging technology, which has the potential to enable three-dimensional imaging. Calibration of the three-dimensional light field endoscope (3D LFE) is essential but challenging as the disparity is much smaller than that of the conventional light field camera. The geometrical model for 3D LFE is established, and a calibration method based on virtual objective lens and virtual feature points is proposed. First, the virtual objective lens is introduced and the parameters about it are calibrated using corner features in center subaperture images. Second, two types of virtual feature points are proposed to calibrate the parameters about the microlens array, one is on the black-and-white board line and the other is selectively determined but can be anywhere on the checkerboard. Moreover, the relationship between the virtual feature points mapping in the microlens image and the virtual feature points mapping in the central subaperture image is deduced to overcome tiny light field disparity. Experimental results verify the performance of our calibration method. © 2020 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.OE.59.10.104101]

Keywords: three-dimensional light field endoscope; calibration; virtual objective lens; virtual feature points.

1 Introduction

Imaging in minimally invasive surgery is usually performed via endoscope or laparoscope, which is an important imaging mode in the medical imaging area, such as computed tomography, optical coherent tomography, and ultrasound imaging.[1–3] Although traditional two-dimensional (2D) endoscope has advantages of small scars and quick recovery, it is restricted in finite surgical operations as three-dimensional (3D) information of organs is missing. Therefore, 3D endoscope is an interesting research area.

Existing 3D endoscopes are mainly based on stereo vision principles using polarized glasses.[4] However, the depth information measured by the stereo endoscope is not precise enough for many applications, such as 3D display and surgical instrument navigation, as the baseline of stereo endoscope is too short. In recent years, a great deal of attention has been given to light field imaging, which is a technique that records four-dimensional slices including 2D spatial information and 2D angular information. Light field imaging records volumetric information in single shot, therefore, has potential applications in many areas, such as light field camera (LFC),[5,6] light field microscopy, and light field endoscope (LFE).[7]

In 3D light field imaging, the imaging model and calibration approach play essential roles. In general, the light field imaging models are based on a two-parallel-plane (TPP) model,[8–14] where the main lens (or objective lens) is regarded as an angular plane and the microlens array (MLA) is regarded as a spatial plane. Accordingly, many calibration approaches have been proposed for the conventional LFC in the last decades. The main differences between the calibration

---

approaches exist in light field imaging models, appropriate features, and corresponding parameters. Duan et al.[8] presented a homogenous intrinsic matrix to describe the relationship between the pixel on the image plane and its corresponding light field. Bok et al.[9] proposed a similar relationship while the raw light field images and line features were applied. Zhang et al.[13] proposed a generic multiprojection-center model, which is similar to Dansereau's imaging model while less parameters are needed.[10] Hall et al.[11] presented a correction method due to lens distortion and thin-lens assumptions, where the volumetric calibration based on polynomial mapping function was applied. In general, the light field images are equivalent to images of traditional cameras with coplanar projection centers and a short baseline. The accuracy of calibration and correction depends on the light field disparity and how to use it. However, the variation of the light field disparity of 3D LFE is much smaller than that of the conventional LFC due to the two nonlinear transformations from 3D coordinates to the light field disparity in the LFE imaging.[12] Therefore, the current calibration methods for the conventional LFC are not suitable for 3D LFE. To the best of our knowledge, little work has been performed on the 3D LFE calibration so far.

In 2013, Dansereau et al.[13] proposed a TPP model to describe the relationship between a ray and its image, where an outside virtual plane was treated as the angular plane and a 12-free-parameter transformation matrix was achieved. Although the parameters in the imaging model lose physical meanings, the imaging model makes the calibration flexible and inspires us. In this paper, the geometrical model for 3D LFE is established and a valid calibration method for 3D LFE based on virtual objective lens and virtual feature points is proposed. The virtual objective lens is introduced as the angular plane in the TPP model, which is neither the objective lens of the 2D endoscope nor the objective lens of 3D LFE, but an outside virtual plane, just like Dansereau's work.[13] The parameters about the virtual objective lens are calibrated using the corner features of the checkboard in center subaperture images. Instead of the widely used corner features, pairs of virtual feature points are utilized to accomplish the calibration of MLA. Furthermore, the relationship between the virtual feature points mapping in the microlens image and the virtual feature points mapping in the central subaperture image is deduced to overcome tiny light field disparity.

2 Setup of 3D LFE

In this section, an experimental LFE system is first established and its simplified optical schematic is described accordingly. The 3D LFE proposed in this paper is based on an industrial 2D endoscope and consists of an objective lens (objective lens II), an MLA, a 1:1 relay lens, and a picture sensor, as shown in Fig. 1(a). To avoid optical misalignment, which will reduce the achievable resolution, the objective lens II, the MLA, and the relay lens are mounted inside kinematic mounts (Thorlabs), allowing fine adjustment in respect to the optical axis. The relay lens is utilized to facilitate the 3D LFE imaging system, as the focal length of the MLA is normally <1 mm and it is difficult to place the picture sensor on the focal plane of the MLA directly.

To simplify the imaging model of the 3D LFE, the 2D endoscope is treated as an equipment composed of an objective lens and an eyepiece lens. The optical schematic of 3D LFE is shown in Fig. 1(b). F_o, F_o' refer to the focal points of objective lens I, and F_e, F_e' refer to the focal points of eyepiece. The imaging process of 3D LFE could be simplified into the following steps. For an object, the objective lens I produces its real image, then the eyepiece produces a magnified virtual image. Subsequently, the objective lens II is used to capture the light passing through the 2D endoscope, then the MLA captures different perspectives of the objective lens II, and the light through MLA is further imaged on the picture sensor. After being fully recorded, the raw light field images are adjusted by locating the microlens with respect to the pixels and then they are cropped to microlens images for calibration.

3 Virtual Objective Lens and Its Calibration

The geometrical model of 3D LFE is proposed based on TPP principles, as shown in Fig. 2, where the relay lens is ignored as the imaging procedure is not affected by the relay lens. Inspired
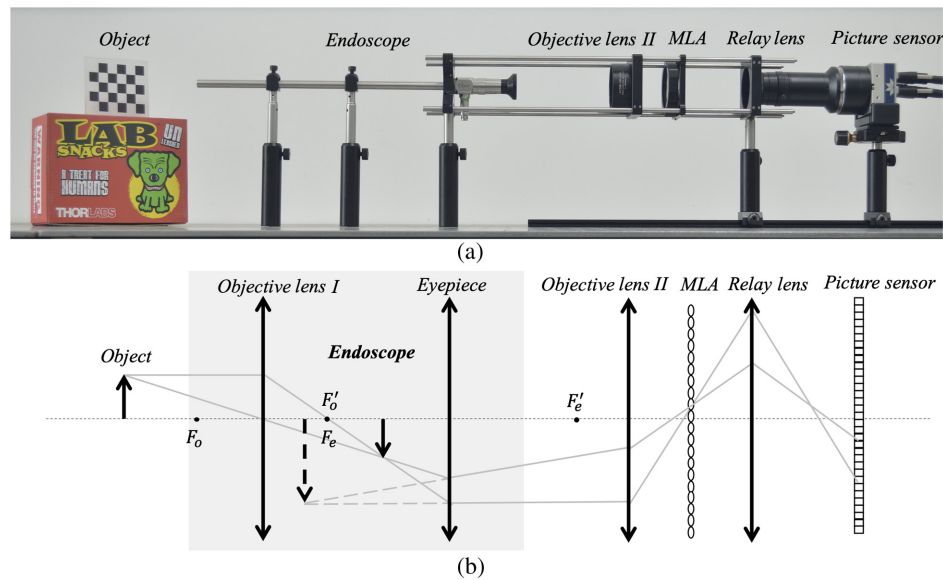
Fig. 1 3D LFE: (a) setup of the 3D LFE and (b) simplified optical schematic of the 3D LFE.
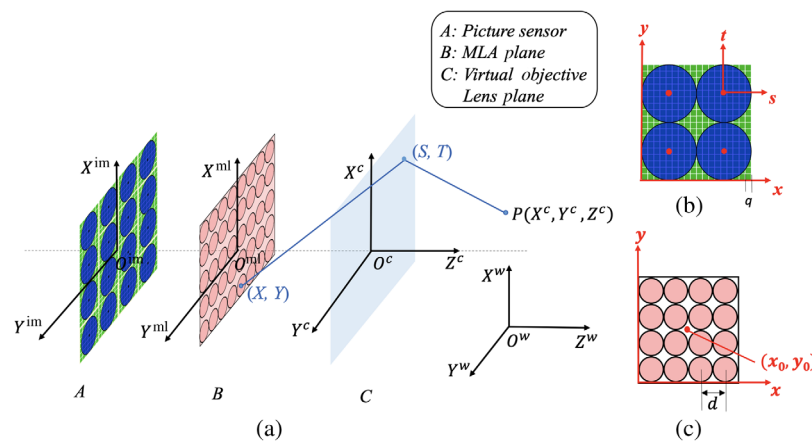

Fig. 2 Geometrical model of 3D LFE and the coordinate system: (a) the simplified geometrical model and related coordinates, (b) main view of partial picture sensor, and (c) main view of partial MLA plane.

by Dansereau's work, a virtual objective lens is introduced to simplify the geometrical model in our work, which is an abstraction of the 2D endoscope and the objective lens II together. The concept of virtual objective lens refers to a single virtual thin lens that is equivalently abstracted from multiple lenses. In the 3D LFE TPP model, two parallel planes are the virtual objective lens plane (S, T) and the MLA plane (X, Y), which describe the angular and spatial information, respectively. Based on the TPP principles, the virtual objective lens plane and the MLA plane are parallel to each other. Therefore, the light emitted from the 2D endoscope is defined by connecting a point in the virtual objective lens plane to a point in the MLA plane. In other words, the light field is expressed as $L(S, T, X, Y)$ in TPP model.

To express the imaging procedure of 3D LFE, five coordinate systems are defined as follows and the notations of these coordinate system symbols are listed in Table 1. All the coordinate systems in Fig. 2 follow the same convention: the Z axis points toward the object space ($Z > 0$), the X axis points upward and the Y axis points out of this figure. Moreover, the coordinates in capital letters, such as $P(X^c, Y^c, Z^c)$, are in units of millimeter, and the coordinates in lowercase letters are in units of pixel.

Table 1 Notation of symbols in the light field model.

$O^{im} - X^{im} Y^{im}$	$O^c - X^c Y^c Z^c$	$O^w - X^w Y^w Z^w$
Image coordinate system	Camera coordinate system	World coordinate system
$O^{ml} - X^{ml} Y^{ml}$	$L(S, T, X, Y)$	$L(s, t, x, y)$
MLA coordinate system	Light field in TPP model in units of millimeter	Decoded light field

Without loss of generality, the optical center and the optical axis of the virtual objective lens are defined as origin O^c and Z^c axis of the camera coordinate system, respectively. Based on the TPP model and LFE model, the virtual objective lens plane, the MLA plane, and the picture sensor are perpendicular to Z^c axis. In the image plane, the raw light field image can be first decoded as $L(s, t, x, y)$, where (s, t) are the pixel coordinates of microlens images, and (x, y) are both the indices of microlens images and pixel coordinates of subaperture image.

Based on the light field imaging principles,[14] for a certain point $P(X^w, Y^w, Z^w)$ in the scene, it may correspond to a certain point in the MLA plane and appear multiple times in the image plane, as shown in Fig. 3. $P(X^c, Y^c, Z^c)$ is the corresponding point of $P(X^w, Y^w, Z^w)$ in the camera coordinate system and the relationship between $P(X^w, Y^w, Z^w)$ and its corresponding image point $P(x, y)$ in the MLA image plane is expressed as

$$Z^c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{d_{ob-ml}}{d} & 0 & x_0 - \frac{S \cdot d_{ob-ml}}{d \cdot h_m} & -\frac{S \cdot d_{ob-m}}{d} \\ 0 & \frac{d_{ob-ml}}{d} & y_0 - \frac{T \cdot d_{ob-ml}}{d \cdot h_m} & -\frac{T \cdot d_{ob-ml}}{d} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix}, \quad (1)$$

where R and t are the rotation and translation matrices, respectively, which are named as extrinsic parameters in general. d_{ob-ml} is the distance between the virtual objective lens plane and the MLA plane, and h_m is the distance between the world focal plane and the virtual objective lens plane. (x_0, y_0) is the principle point of the MLA plane. S and T are the distance from optical center O^c to a certain subaperture of the virtual objective lens that corresponding rays pass through. The "pixel size" of the virtual image sensor is the diameter d of the microlens rather than the realistic pixel size of the picture sensor. Note that d_{ob-ml}, h_m, and d are in units of millimeter.

By setting S and T to zero, Eq. (1) is rewritten as

$$Z^c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{d_{ob-ml}}{d} & 0 & x_0 & 0 \\ 0 & \frac{d_{ob-ml}}{d} & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix}, \quad (2)$$
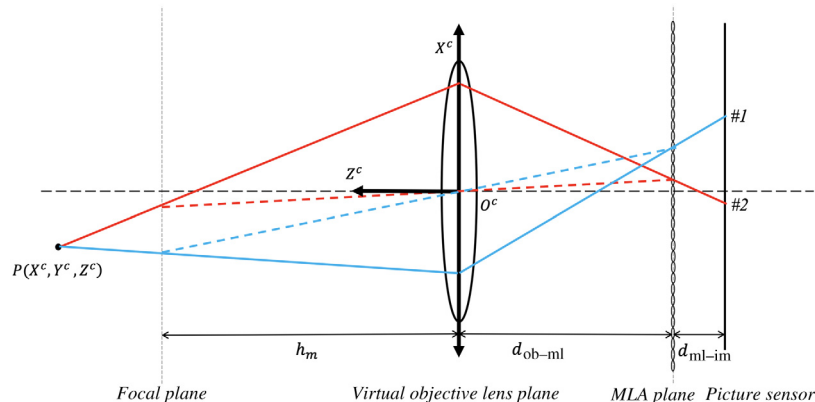
Fig. 3 Two-dimensional geometrical model of LFE.

which means that the MLA plane is treated as a virtual image plane and calibration of the parameters about the virtual objective lens is accomplished first. Although the conventional camera is substituted by the central subaperture image, the calibration principles are still identical. Therefore, the intrinsic and extrinsic parameters about the virtual objective lens in Eq. (2) are derived based on Zhang's calibration method and our previous work[15,16] with a series of corner feature points in central subaperture images recorded from different angles.

4 MLA Calibration Based on Virtual Feature Points

Although the traditional calibration methods for MLA with estimation of light field disparity are available in conventional LFC, they are almost not possible for the MLA in 3D LFE. The tiny disparity in light field images requires the calibration board to move a long distance, which cannot be achieved for 3D LFE as its working distance is limited. A calibration method for MLA is presented by introducing the concept of virtual feature points. Two different types of virtual feature points are utilized, one is on the black-and-white board line and the other one is selectively determined but can be anywhere on the checkerboard.

The imaging process of two typical virtual feature points $P_1^c(X_1^c, Z_1^c)$ and $P_2^c(X_2^c, Z_2^c)$ is depicted in Fig. 4, the scale of which is deceptively magnified to see the microlens clearly. All rays emitted by point $P_1^c(X_1^c, Z_1^c)$ pass through the virtual objective lens and converge to an imaging point $P_1^e(X_1^e, Z_1^e)$ in the image space of the virtual objective lens, and then the light that passes through the center subaperture of the virtual objective lens heads to $P_1^{ml}(X_1^{ml}, Z_1^{ml})$ on the MLA plane and $P_1^{im}(X_1^{im}, Z_1^{im})$ on the picture sensor. The mapping procedure about $P_1^c(X_1^c, Z_1^c)$ is shown as blue lines in Fig. 4. It is important to note that P_1^{ml} is the central point of a microlens image, which is named MI. Similarly, the mapping procedure about $P_2^c(X_2^c, Z_2^c)$ comes to $P_2^e(X_2^e, Z_2^e)$ and $P_2^{ml}(X_2^{ml}, Z_2^{ml})$ is shown as red lines in Fig. 4. Particularly, among the multiple imaging points of $P_2^c(X_2^c, Z_2^c)$ on the picture sensor, which are emitted from different subapertures of the virtual objective lens, the one $P_2^{im}(X_2^{im}, Z_2^{im})$ in the same microlens image MI with $P_1^{im}(X_1^{im}, Z_1^{im})$ is chosen, where P_2^{im} is mapped by the light ray emitted from $P_2^c(X_2^c, Z_2^c)$ and passing through the specific subaperture S of the virtual objective lens (shown as the yellow solid line in Fig. 4), and d_{s-ob} is the distance between the specific subaperture S and the optical center of the virtual objective lens. The imaging point P_1^{im} of the virtual feature point P_1^c is easily determined in our calibration method as it is the center of the microlens image MI, therefore, P_1^c can be any point on the checkerboard in the scene. How to extract the imaging point P_2^{im} of the other type of virtual feature points will be explained in the next section.
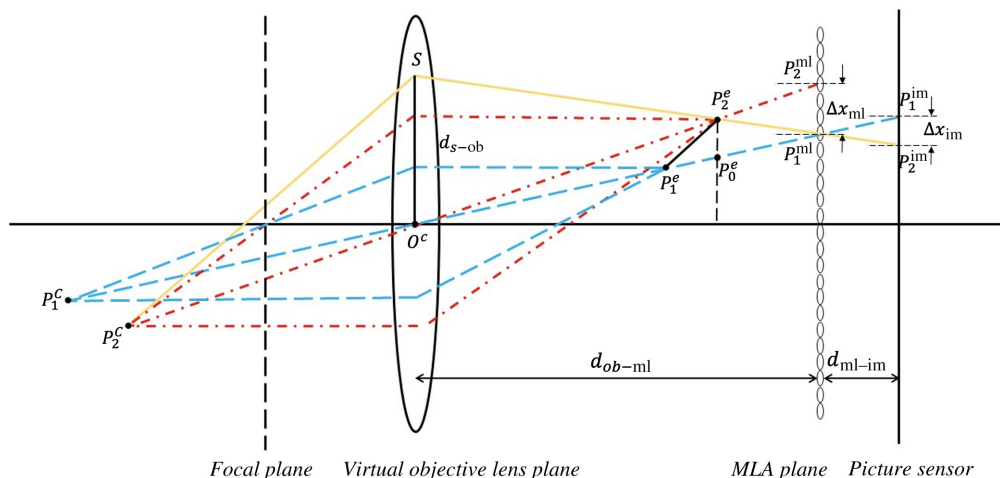

Fig. 4 Calibration model about MLA.

In order to implement the calibration of MLA, the relationship between two distances is deduced, which are the distance from P_1^{im} to P_2^{im} in the microlens image and the distance from P_1^{ml} to P_2^{ml} in the central subaperture image. The distance from P_1^{im} to P_2^{im} in the microlens image is defined as ΔX_{im}, and the distance from P_1^{ml} to P_2^{ml} in the central subaperture image is defined as ΔX_{ml}. The two distances are in units of millimeter.

In the derivation of the calibration formula about the parameter $d_{\mathrm{ml-im}}$, an auxiliary point P_0^e is introduced as the intersection of the chief light ray of P_1^c and the perpendicular line from P_2^e to the optical axis. As shown in Fig. 4, the triangles $\Delta OSP_1^{\mathrm{ml}}$ and $\Delta P_0^e P_2^e P_1^{\mathrm{ml}}$ are similar, and the triangles $\Delta OSP_1^{\mathrm{ml}}$ and $\Delta P_1^{\mathrm{im}} P_2^{\mathrm{im}} P_1^{\mathrm{ml}}$ are also similar. Therefore, there is

$$\frac{d_{s-\mathrm{ob}}}{d_{\mathrm{ob-ml}}} = \frac{X_2^e - X_0^e}{d_{\mathrm{ob-ml}} - Z_0^e} = \frac{\Delta X_{\mathrm{im}}}{d_{\mathrm{ml-im}}}. \tag{3}$$

Given the constraint that P_1^c, P_0^e and the optical center O^c are collinear, there is

$$\frac{X_1^c}{Z_1^c} = \frac{X_0^e}{Z_0^e}. \tag{4}$$

Similarly, P_2^c, P_2^e, and the optical center O^c are collinear, there is

$$\frac{X_2^c}{Z_2^c} = \frac{X_2^e}{Z_2^e}. \tag{5}$$

By combining Eqs. (3)–(5) and the relationship that Z_0^e and Z_2^e are equal, Eq. (3) is rewritten as

$$\frac{\Delta X_{\mathrm{im}}}{d_{\mathrm{ml-im}}} = \frac{Z_2^e}{d_{\mathrm{ob-ml}} - Z_2^e} \left(\frac{X_2^c}{Z_2^c} - \frac{X_1^c}{Z_1^c} \right). \tag{6}$$

Furthermore, as points P_2^c, P_2^{ml}, and O^c are collinear, and points P_1^c, P_1^{ml}, and O^c are collinear, Eq. (6) is rewritten as

$$\frac{\Delta X_{\mathrm{im}}}{d_{\mathrm{ml-im}}} = \frac{Z_2^e}{d_{\mathrm{ob-ml}} - Z_2^e} \frac{1}{d_{\mathrm{ob-ml}}} \cdot \Delta X_{\mathrm{ml}} \tag{7}$$

considering that $d_{\mathrm{ob-ml}}$, Z_1^{ml} and Z_2^{ml} are equal. The distances Δx_{im} and Δx_{ml} (in units of pixel) are corresponded to ΔX_{im} and ΔX_{ml} (in units of millimeter), respectively. There are $q \cdot \Delta x_{\mathrm{im}} = \Delta X_{\mathrm{im}}$ and $d \cdot \Delta x_{\mathrm{ml}} = \Delta X_{\mathrm{ml}}$, where d is the pitch of the microlens and q is the pixel size of the picture sensor, which could be obtained from the manufacturing information of the MLA and the picture sensor. Therefore, Eq. (7) is rearranged as follows to yield a simple expression for the parameter $d_{\mathrm{ml-im}}$ about MLA:

$$d_{\mathrm{ml-im}} = \frac{q}{d} \frac{(d_{\mathrm{ob-ml}} - Z_2^e) \cdot d_{\mathrm{ob-ml}}}{Z_2^e} \frac{\Delta x_{\mathrm{im}}}{\Delta x_{\mathrm{ml}}}. \tag{8}$$

Eq. (8) shows that $d_{\mathrm{ml-im}}$ is a linear function of a ratio, which is determined by four imaging points P_1^{ml}, P_2^{ml}, P_1^{im} and P_2^{im} on the MLA plane and picture sensor of two virtual feature points $P_1^c(X_1^c, Z_1^c)$ and $P_2^c(X_2^c, Z_2^c)$. Therefore, it is important to determine these imaging points of virtual feature points to finish the calibration of the parameter about the MLA.

The microlens image MI, which contains a board line (the light blue line) of black-and-white regions is determined first of all, as shown in Fig. 5(a). Then the center [the black dot in Fig. 5(a)] of the MI is considered as P_1^{im}, which has two properties as follows. (1) The pixel indices of P_1^{im} in the determined MI are both zero as it is the center of MI. (2) P_1^{ml} is located in the central subaperture image, and its coordinates are equal to the indices of the microlens in the MLA. Furthermore, a line in MI is determined, which passes through the center P_1^{im} and is perpendicular to the black-and-white board line. The perpendicular foot [the dark blue dot in Fig. 5(a)] is considered as P_2^{im}, which is also the imaging point in the MI of another virtual feature point P_2^c. As the microlens image is a picture of the virtual objective lens viewed from
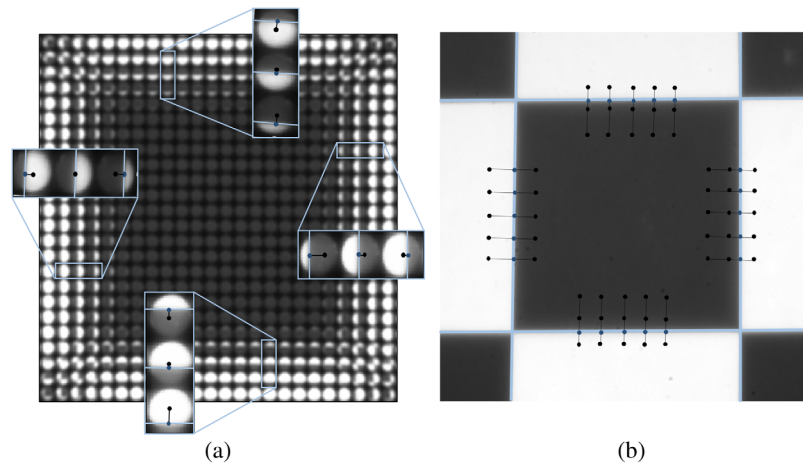
Fig. 5 Extraction of the virtual feature points. (a) 4D light field raw data and the partial magnified microlens images with marked line features and virtual feature points mapping in the MIs. (b) Partial mapping points of the virtual feature points in the central subaperture image.

specific position, the board line in the microlens image is part of the board line in the central subaperture image [the light blue line in Fig. 5(b)]. Therefore, P_2^{ml} in the central subaperture image [the dark blue dot in Fig. 5(b)] could be also determined by the perpendicular foot from P_1^{ml} to the board line. For every microlens image that contains a board line, the distance Δx_{im} and the corresponding distance Δx_{ml} in the central subaperture image could be derived from these four imaging points of two virtual feature points. The calibration of the parameter about MLA expressed in Eq. (8) is consequently accomplished when numerous such microlens images exist.
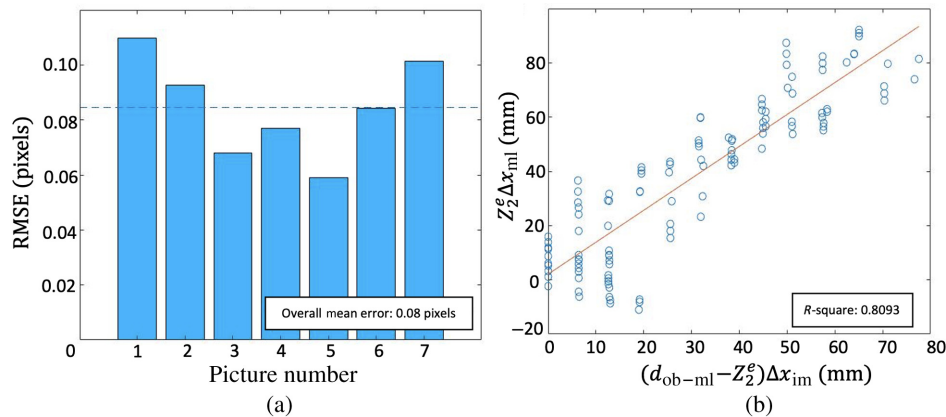
5 Experimental Results

The 3D LFE system proposed in this paper is equipped mainly with the Shenda 2D conventional endoscope (ZG-3A), the Thorlabs objective lens (N-BK7 LA1509-ML, focal length is 50 mm), the AUS MLA (APO-Q-P136-R0.94), the Myutron 1:1 relay lens (LS010F) and the Teledyne DALSA industrial camera (G3-CC10-G5105). To extract central subaperture images and microlens images, the 4D light field is recovered from the original 2D image recorded on the picture sensor of the industrial camera, where the MATLAB LFToolbox (V 0.4) designed by Dansereau is used. The setup of 3D LFE system follows the description in Sec. 2 and is shown in Fig. 1(a).

We captured seven checkerboard images to accomplish the LFE calibration. The size of raw images is 5112×5112 pixels, the size of central subaperture image is 176×176 pixels, and the size of microlens image is 29×29 pixels. The pixel size q is 0.0045 mm, and the microlens pitch d is 0.136 mm, which are obtained from the metadata provided by product supplier. The checkerboard is a pattern of 5×7 cells, whose grid size is 12 mm $\times$ 12 mm. In our experiments, the checkerboard is located at the distance of 100 to 200 mm from the objective lens I of 2D endoscope. To accomplish the calibration, 140 features are utilized for calibration about the virtual objective lens, and 140 pairs of virtual feature points are determined for calibration about MLA. Zhang's calibration method and the least square method are performed, respectively. Note that the radial distortion of the 3D LFE is mainly caused by the 2D endoscope because its objective lens has short focal length but wide field of view, whereas the MLA is ideal approximately. Therefore, the undistorted parameters are derived in calibration about the virtual objective lens only. The nonlinear optimization is performed with the Levenberg–Marquardt algorithm. The calibration results are detailed in Table 2.

The reprojection error of the corner points in the central subaperture images is shown in Fig. 6(a) to evaluate the performance of calibration about the virtual objective lens, and the line fitting result is shown in Fig. 6(b) to illustrate the performance of calibration about MLA. As

Table 2 Parameters of 3D LFE.

(a) Intrinsic parameters		
$d_{\mathrm{ml-im}}$ (mm)	$d_{\mathrm{ob-ml}}$ (mm)	(x_o, y_0)
0.907	11.276	(89.364, 87.497)
(b) Extrinsic parameters		
R ($\times 10^{-3}$ rad)	t (mm)	(k_1, k_2)
$(-0.220, -0.185, -3.080)$	(22.627, 20.700, 104.632)	$(-0.595, 0.365)$

Fig. 6 Experimental results of the proposed calibration method for 3D LFE. (a) Reprojection error of the calibration about the virtual objective lens. (b) Line fitting results of the calibration about MLA.

shown in Fig. 6(a), the typical root mean square (RMS) reprojection error is about 0.08 pixels. According to Dansereau's results,[13] the calibration accuracy increases with decreasing the checkerboard grids; a too small grid will make the corner extraction inaccurate. Therefore, three checkerboards are used in our calibration method, and typical RMS reprojection errors are 0.10, 0.08, and 0.11 pixels for 15, 12, and 9 mm checkerboard grids, respectively. Consequently, the 12-mm checkerboard is used in this paper. Although the calibration about the virtual objective lens is precise enough for the 3D LFE, it can be improved compared to the conventional cameras and LFC calibration.[13,15,17,18] For the calibration result of MLA, as shown in Fig. 6(b), the distance $d_{\mathrm{ml-im}}$ between the MLA and picture sensor is 0.907 mm and the root mean square error (RMSE) is 0.008 mm only. To demonstrate the performance of our 3D LFE calibration method, three other calibration methods for conventional LFC were applied to the 3D LFE for comparison, but the results were not satisfied. For the calibration method proposed in Ref. 15, the RMSE of $d_{\mathrm{ml-im}}$ is 3.422 mm due to tiny light field disparity in 3D LFE. Furthermore, other calibration methods proposed in Refs. 12 and 19 even cannot be achieved.

To further evaluate the validity of our 3D LFE calibration method, another eight checkerboard pictures are captured and their central subaperture images are shown in Fig. 7(a). The checkboard grid is 12.000 ± 0.005 mm. All the corner feature points corresponding to the eight checkerboard pictures are reconstructed and shown in Fig. 7(b), respectively, where the EPI-based 3D reconstruction principles are used.[12] It is obvious that distortion was well corrected, although it was apparent in Fig. 7(a). The Euclidean distances between adjacent reconstructed feature points are computed, and the RMSE is 0.280 mm. Moreover, the standard deviation (STD) of the planes fitted by the reconstructed feature points shown in Fig. 7(b) is calculated. The mean STD of the 8 fitted planes is 0.123 mm. Both RMSE and STD demonstrate the good performance of the calibration method proposed in this paper.
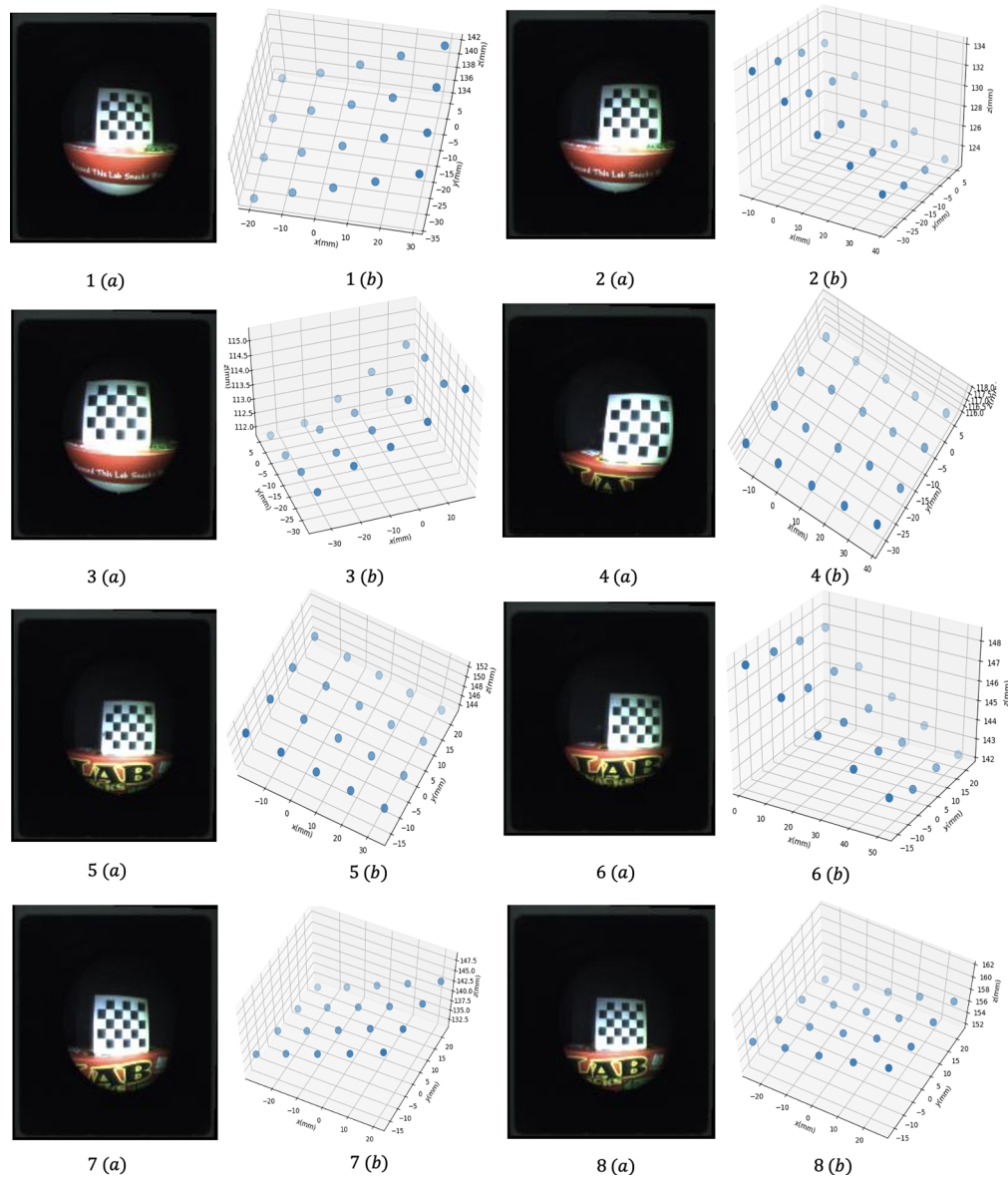
Fig. 7 EPI-based 3D reconstruction results using calculated parameters: (a) the central subaperture images and (b) reconstructed coordinates of the corner features in the camera coordinate system.

6 Conclusion and Discussion

In summary, a new calibration method based on virtual objective lens and virtual feature points for 3D LFE is introduced and two main contributions are made in this paper. On the one hand, inspired by Dansereau's work, a virtual objective lens is introduced in 3D LFE imaging model, which decomposes the 3D LFE calibration to the calibrations about the angular and spatial planes respectively and makes the calibration similar to that of the conventional camera. On the other hand, the concept "virtual feature points" is proposed in this paper. To calibrate the MLA, the relationship between pairs of virtual feature points mapping in the microlens image and mapping in the central subaperture image is deduced to overcome tiny light field disparity. Our calibration method's performance has been verified by the RMS reprojection error, RMS line fitting error, and EPI-based 3D reconstruction results with the calibrated parameters.

Although light field imaging is capable of recording volumetric information in single shot, it has not been used for many medical imaging applications, as it is acquired at the cost of reduced spatial resolution and inherent trade-offs between spatial and angular resolution. The reduced

spatial resolution that results from ray-space multiplexing may be overcome by future improvement such as high-resolution picture sensors, and the trade-off between angular and spatial resolution may be solved by the super-resolution method.[20,21] Our future work may include a more appropriate 3D LFE imaging model to overcome the limitations of pinhole and thin lens, as well as a more complicated lens distortion model to improve the precision of 3D LFE calibration.

Acknowledgments

References

1. Q. Huang, B. Wu, J. Lan, and X. Li, "Fully automatic three-dimensional ultrasound imaging based on conventional B-scan," *IEEE Trans. Biomed. Circuits Syst.* **12**(2), 426–436 (2018).
2. S. Gim et al., "Parallel guidance endoscopic optical coherence tomography system for internal diagnosis through active cannulas," *Opt. Eng.* **53**(8) 084105 (2014).
3. G. Q. Zhou et al., "Automatic measurement of spine curvature on 3-D ultrasound volume projection image with phase features," *IEEE Trans. Med. Imaging* **36**(6), 1250–1262 (2017).
4. S. Y. Bae et al., "Development of a miniature single lens dual-aperture stereo imaging system towards stereo endoscopic imaging application," *Opt. Eng.* **51**(10), 103202 (2012).
5. T. Iwane and M. Shoda, "Light-field camera and display as information retrieval systems for three-dimensional images," *Opt. Eng.* **57**(6), 061616 (2018).
6. K. Inoue and M. Cho, "Enhanced depth estimation of integral imaging using pixel blink rate," *Opt. Lasers Eng.* **115**, 1–6 (2019).
7. J. D. Liu et al., "Light field endoscopy and its parametric description," *Opt. Lett.* **42**(9), 1804–1807 (2017).
8. H. X. Duan et al., "A new imaging model of Lytro light field camera and its calibration," *Neurocomputing*, **328**, 189–194 (2017).
9. Y. Bok et al., "Calibration of micro-lens-based light field cameras using line features," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(2), 287–300 (2017).
10. Q. Zhang et al., "A generic multi-projection-center model and calibration method for light field cameras," *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(11), 2539–2552 (2018).
11. E. M. Hall et al., "Volumetric calibration of a plenoptic camera," *Appl. Opt.* **57**(4), 914–923 (2018).
12. P. Zhou et al., "Light field calibration and 3D shape measurement based on epipolar-space," *Opt. Express* **27**(7), 10171–10174 (2019).
13. D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding calibration and rectification for lenslet-based plenoptic cameras," *IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1027–1034 (2013).
14. R. Ng et al., "Light field photography with a hand-held plenoptic camera," *Comput. Sci. Tech. Rep.* **2**(11), 1–11 (2005).
15. P. Zhou et al., "A two-step calibration method of lenslet-based light field camera," *Opt. Lasers Eng.* **115**, 190–196 (2019).
16. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
17. S. Dong et al., "Extrinsic calibration of a non-overlapping camera network based on close-range photogrammetry," *Appl. Opt.* **55**(23), 6363–6370 (2016).
18. Y. L. Xiao, X. Y. Su, and W. J. Chen, "Flexible geometrical calibration for fringe-reflection 3D measurement," *Opt. Lett.* **37**(4), 620–622 (2012).
19. Z. W. Cai et al., "Unfocused plenoptic metric modeling and calibration," *Opt. Express* **27**, 20177–20198 (2019).

20. M. S. K. Gul and B. K. Gunturk, "Spatial and angular resolution enhancement of light fields using convolutional neural networks," *IEEE Trans. Image Process.* **27**(5), 2146–2159 (2018).
21. H. W. F. Yeung et al., "Light field spatial super-resolution using deep efficient spatial-angular separable convolution," *IEEE Trans. Image Process.* **28**(5), 2319–2330 (2018).

Ping Zhou received his BSc degree in electronic engineering from the University of Science and Technology of China (USTC) in 2002 and his PhD in biomedical engineering from USTC in 2007. He is currently an associate professor at the School of Biological Science and Medical Engineering of the Southeast University. His research interests include three-dimensional image processing, light field imaging, and three-dimensional medical imaging.

Cheng Gu received her BSc degree in biomedical engineering from the Southeast University, Nanjing, China, in 2019 and is currently working toward her MS degree from the School of Biomedical Science and Medical Engineering of the Southeast University. Her current research area includes light field technology and active three-dimensional reconstruction algorithms.

Weizhong Zhang received his PhD in medicine from China Medical University in 2001 and underwent postdoctoral training at Harvard Medical School. He is currently an ophthalmologist and the vice chief of the Ophthalmology Department at Jiangsu Province Hospital. His research interests include the diabetic retinopathy, applying of ocular endoscopes in vitreous-retina surgeries, and three-dimensional image processing.

Zi Yang received his BSc degree from the School of Biomedical Science and Medical Engineering of the Southeast University in 2018 and is currently pursuing his MS degree. His current research focuses on optical systems and calibration algorithms.

Yuting Zhang received her BSc degree in biomedical engineering from the School of Biomedical Science and Medical Engineering of the Southeast University in 2018 and is pursuing her MS degree. Her current research focuses on light field technology and its passive depth estimation.

Weijia Cai received his MS and BSc degrees in biomedical engineering from the School of Biomedical Science and Medical Engineering of the Southeast University in 2020 and 2017, respectively. His current research focuses on light field technology and its passive depth estimation.

Guangquan Zhou received his BSc and MS degrees in biomedical engineering from the Southeast University, Nanjing, China, in 2000 and 2003, respectively, and his PhD in biomedical engineering (BME) from The Hong Kong Polytechnic University, Hong Kong, in 2015. He is currently an associate professor at the School of Biological Sciences and Medical Engineering of the Southeast University. His research interests include ultrasound imaging, optical imaging, image analysis, and computer-aided diagnosis.