A Corner Detection Method for Conventional Light Field Camera by Jointly Using Line-Features

DONGYANG JIN, SAIPING ZHANG, (Student Member, IEEE), XIAO HUO, AND FUZHENG YANG, (Member, IEEE)
State Key Laboratory of Integrated Services Networks, Xidian University, Xi’an 710071, China
Corresponding author: Saiping Zhang (spzhang@stu.xidian.edu.cn)

ABSTRACT Micro-lens array (MLA) based light field cameras have been commercialized and had a wide range of applications in various fields in recent years. To promote these applications, especially for 3D reconstruction, it is crucial to calibrate light field cameras. Among the calibration steps, it is of great importance to detect corners in checkerboard images and establish the point-to-point correspondence between the checkerboard and its images. However, almost all existing algorithms detect corners in the sub-aperture images instead of raw images for conventional light field cameras. In this paper, a corner detection algorithm that can recognize corners in the raw images of conventional light field cameras was proposed. Firstly, template matching was used to detect the 3D lines in the image space of the main lens. Then, the locations of the corners were obtained by calculating the intersection of 3D line segments and reprojecting it to the raw image again. Experimental results have demonstrated that the algorithm can still achieve good performance in actual datasets with blurred and low resolution micro lens image.

INDEX TERMS Conventional light field cameras, raw images, corner detection, intersection of lines.

I. INTRODUCTION

Light field photography technology has achieved large developments in recent years since it can capture both the spatial and angular information of light rays at the same time. Various kinds of light field cameras have been proposed gradually [1]–[5]. Among them, micro-lens array (MLA) based light field cameras are the most widely used [4], [5]. According to the functionality of MLA in the whole imaging path, MLA based light field cameras can be divided into two types, i.e., conventional light field cameras and focused light field cameras, and both have been commercialized [6], [7].

Light field technology has wide applications in multiple fields, for instance, depth estimation [8]–[10], light field super-resolution [11], digital refocusing [12], synthetic aperture imaging [13], [14] and visual simultaneous localization and mapping (VSLAM) [15], [16]. To promote these applications, especially for 3D reconstruction, it is fundamental to calibrate light field cameras. Camera calibration, which aims to establish the correspondence between pixels in the sensor and the rays out of the camera, relies on a series of photos of the checkerboard with different poses. Corner detection of checkerboard images is the basis of the whole calibration algorithm. The performance of detection algorithm has direct influence on the performance of the calibration algorithm. It is the key of the whole calibration process.

In this paper, a corner detection algorithm is proposed aiming at checkerboard images of conventional light field cameras. Taking raw images as the input, this algorithm find the locations of the 3D corner points of the checkerboard image in the image space of main lens by calculating intersections of two 3D line segments. After projecting these 3D corner points on micro-lens images, the coordinates of 2D corner points in raw images can be acquired. Our code is publicly available.1

Our main contributions are as follows:

1) We propose a joint detection algorithm to recognize 2D line features in micro-lens images with the help of 3D line segments in the image space of main lens. Due to the joint use of multiple micro-lens images, this line detection algorithm achieves a result that is more accurate and robust than independently line detection in each micro lens images. Besides, the parameters of 3D line segments can be obtained simultaneously.

1https://gitee.com/HxRecons/LightFieldCornerDetection
We propose a corner detection algorithm that is suitable for raw images of conventional light field cameras. The locations of corners in the raw images is acquired by calculating the intersection of 3D line segments and reprojecting the intersection to the raw image. By doing so, the difficulty of directly detecting corners in micro-lens images which is in extreme low resolution is avoided. Experimental results demonstrate that the algorithm is accurate and works well in blurred actual dataset.

II. RELATED WORKS
Dansereau et al. [17] was the first to deliver an end-to-end geometry calibration method for conventional light field cameras. Sub-aperture images were first extracted from raw images and corners in all sub-aperture images were detected as the input of the subsequent calibration process. Reference [18] introduced generic multi-projection-center model to the calibration of light field camera, and their algorithm can be applied to both conventional light field cameras and focused light field cameras. However, for conventional light field cameras, sub-aperture images should still be extracted before the implementation of corner detection. Reference [19] present a two-step calibration scheme for light field cameras. Intrinsic parameters with different properties were established in each step. Intrinsic parameters relating to the main lens and extrinsic parameters were first determined by using the sub-aperture image at the center view. Then disparities between different sub-apertures were determined by using corner detection results of other sub-aperture images and were further utilized to estimate other intrinsic parameters. All of the above algorithms recognize corners from sub-aperture images, however, the resolution of sub-aperture images is far less than that of the raw image. Taking Lytro Illum as an example, which is a high end hand-held conventional light field camera. The resolution of its raw image is 7728 × 5368, while the resolution of sub-aperture images extracted from raw images is only 625 × 434, which has a negative influence on the accuracy of corner detection algorithm. Moreover, the final detection result cannot strictly guarantee the detected corners of different sub-aperture images are on a single parallax line in the epipolar plane image (EPI), since corners in different sub-aperture images are detected separately. Then, when calculating the disparity of a corner, a line fitting method [19] would be adopted.

Besides, there are some algorithm detecting corners directly in raw images. Noury et al. [20] proposed a corner detection algorithm consist of two steps. They first classified micro lens images according to their characteristics. Then they detected subpixelic corners locations in those micro lens images which contain corners using a pattern registration method. Nousias et al. [21] estimated the location of corners in micro lens images by calculating the intersection point of two potential saddle axes. However, the two methods mentioned above are designed only for focused light field cameras which has clear micro lens images. As for conventional light field cameras, the resolution of micro lens images is far lower than that of focused light field cameras. For instance, the resolution of micro lens images of Lytro Illum, a typical conventional light field camera, is only 15 × 15, while that of Raytrix R5, a typical focused light field camera, achieves 30 × 30. Moreover, for the conventional light field cameras, each micro lens plays a role of splitting light, rather than focusing light. This makes each micro lens image of conventional light field cameras become more blurred than that of focused light field cameras. Thus, low resolution and blurred micro lens images, which are the inherent characteristic of conventional light field cameras, become the largest challenge for detecting corners from the raw images of conventional light field cameras. To avoid directly detecting corner, [22] showed a calibration algorithm using “line features”. They took high resolution raw images as input, rather than sub-aperture images with relatively low resolution. Instead of detecting micro lens images with corners, they detect micro lens images with “line features” which is the 2D image of the 3D line segment connecting two adjacent corners. Besides, this method does not directly export corner locations in the raw image and it does not use corners correspondence to calibrate cameras either. Instead, they solved intrinsic parameters of light field cameras by establishing the relationship between line features and points on 3D line segments. Inspired by the line feature detecting method in raw images in [22], we further improve their work. And our algorithm can not only boost the accuracy of the line feature detection results, but also export the corner locations in raw image.

III. PREPARATIONS
In the optical path model of this algorithm, main lens and micro lens array are regarded as thin lens and pinhole array, respectively. To better express the algorithm, three coordinate systems are defined in this paper.

The first is “world coordinate system”, the origin of which is set at the center of micro lens array. Z axis is perpendicular to micro lens array and points out of the camera. X axis and Y axis are located on the plane of micro lens array (any two perpendicular lines on the plane of micro lens array). As shown in Figure 1, the distance between micro lens array and the sensor is d. The distance between the main lens and the sensor is D. X_0, Y_0, Z_0 represents the origin of the coordinate system. X_corner, Y_corner, Z_corner represents the image point of a checkerboard corner in the image space of the main lens. X_{mla}^{im}, Y_{mla}^{im}, Z_{mla}^{im} represents the center of the i-th micro lens.

The second is the “image coordinate system of raw images”. For each pixel, its image coordinate denotes its pixel index in the raw image. The origin of the image coordinate system is at the top left corner of the raw image, the coordinate of which in the world coordinate system is denoted as X^{im}, Y^{im}, Z^{im}.

The third is “micro lens image coordinate system”. Every micro lens image correspond to a “micro lens image coor-
coordinate system”, the origin of which is set at the center of this micro lens image. The directions of coordinate axes are the same as that of the world coordinate system. For the i-th micro lens image, suppose the coordinate of its center in “image coordinate system of raw images”, i.e. the “image coordinate” of $X_{mla}^i$, $Y_{mla}^i$, $Z_{mla}^i$, is $U_c^i$, $V_c^i$, then

$$
U_c^i = (X_{mla}^i - X_{im}^0) \frac{D}{(D-d)} \frac{1}{d_{pix}},
$$
$$
V_c^i = (Y_{mla}^i - Y_{im}^0) \frac{D}{(D-d)} \frac{1}{d_{pix}},
$$

(1)

where $d_{pix}$ is the distance between two adjacent pixels. After imaging of the i-th micro lens, suppose the 2D projected point of the 3d image point $X_{corner}$, $Y_{corner}$, $Z_{corner}$ in the i-th micro lens image is $U_s^i$, $V_s^i$, which denotes the coordinate of the “micro lens image coordinate system” of this 2D projected point. And $U_s^i$, $V_s^i$ can be calculated as follows,

$$
\begin{bmatrix}
U_s^i \\
V_s^i \\
1
\end{bmatrix}
= K_i \ast [R_i \ast T_i] \ast
\begin{bmatrix}
X_{corner} \\
Y_{corner} \\
Z_{corner} \\
1
\end{bmatrix}
$$

(2)

which can be further expanded as

$$
\begin{bmatrix}
U_s^i \\
V_s^i \\
1
\end{bmatrix}
= \begin{bmatrix}
\frac{d}{d_{pix}} & \frac{X_{mla}^i}{D-d} & \frac{d}{d_{pix}} \\
\frac{d}{d_{pix}} & \frac{Y_{mla}^i}{D-d} & \frac{d}{d_{pix}} \\
1 & -X_{mla}^i & 1 \\
1 & -Y_{mla}^i & 1 \\
1 & -Z_{mla}^i & 1
\end{bmatrix}
\begin{bmatrix}
X_{corner} \\
Y_{corner} \\
Z_{corner} \\
1
\end{bmatrix}
$$

(3)

where $K_i$ is the intrinsic matrix of the pinhole camera corresponding to the i-th micro lens. $[R_i \ast T_i]$ is the transformation
matrix from the world coordinate system to this pinhole camera coordinate system, where $R$ is a unit matrix, i.e. $R = I$. For expressing brevity, let $P_i = K_i^{-1} [R | T_i]$ denotes the projection matrix from the world coordinate system to the “micro lens image coordinate system” of the i-th micro lens.

For the sake of clarity, some nouns used in the following sections are explained in advance.

- **Checkerboard image**: The 3D image behind the main lens, which is re-converged by the light rays emitted by the checkerboard outside the camera, i.e. the 3D image of the checkerboard in the image space of the main lens, as shown in Figure 1.
- **3D line segment**: The line segment connecting two adjacent corners in the checkerboard image, which is represented by $L$, as shown in Figure 1.
- **3D corner point**: Corners in the checkerboard image.
- **2D line feature**: The 2D image of a 3D line segment in the micro lens image, which is represented by $l$. The corresponding equation for $l$ is $ax + by + c = 0$.
- **2D corner point**: The projection of the 3D corner point in the micro lens image of the raw image.
- $L^H$ and $L^V$: Two categories of $L$ divided according to the angle between its direction and the direction of X axis of the world coordinate system. $L^H$ represents horizontal lines whose angle is less than 45 degrees, while $L^V$ represents vertical lines whose angle is more than 45 degrees.
- **HorArea and VerArea**: Two areas composed of micro lens images with different line features, as shown in Figure 2(a). For a 3D corner point in the checkerboard image, there are two 3D line segments ($L^H$ and $L^V$) intersecting at it. HorArea and VerArea represent the areas composed of micro lenses images with 2D line features corresponding to $L^H$ and $L^V$, respectively.

Without specification, in the following sections, when the word “image” appears alone, it means an image in the camera sensor. And “image point” means the projected point of a scene point by a lens.

**IV. PROPOSED ALGORITHMS**

Inspired by [22], the first procedure of our algorithm is line detection. Different from [22], it is not to detect the 2D line feature in a single micro lens image, but to detect 3D line segments in the checkerboard image. Then the coordinates of the 3D corner point in the checkerboard image are obtained by calculating the intersection of two 3D line segments. Finally, the 3D corner point is re-projected onto the raw image to acquire the coordinates of the 2D corner point in micro lens images.

In this section, two steps are taken to find the optimal 3D line segments. The first step is to calculate a rough initial solution of the 3D line segment. Then starting form the initial solution, we get a series of new 3D line segments by changing the parameters of the line. After re-projecting every 3D line segment to HorArea or VerArea and matching them with raw data, the one with the max similarity to raw data is selected as the final refined solution.

**A. INITIAL SOLUTION**

Taking horizontal 3D line segment ($L^H$) as an example, the method in [22] is utilized to establish HorArea and to detect 2D line feature in each micro lens image in the HorArea respectively. The detection result which is denoted as $l_0 \sim l_k$ is considered as the input of the subsequent steps to calculate the 3D line segments corresponding to these 2D line features.

According to [23], the set of spatial points mapped into a line by a pinhole camera is a plane. Specially, for a 2D line feature $l_i$ in the i-th micro lens image, its corresponding plane equation can be determined by

$$\Pi_i = P_i^T l_i$$

(4)

where $\Pi_i = [A_i, B_i, C_i, D_i]^T$ represents parameters of the plane equation, i.e., $A_iX + B_iY + C_iZ + D_i = 0$. $P_i$ is the projection matrix from the world coordinate system to the i-th micro lens image. Simplifying (4), we can obtain

$$[A_i, B_i, C_i, D_i] = \left[ a_i, b_i, c_i, d_{pix} \right] - \left( a_i U_c^i + b_i V_c^i \right)$$

(5)

where $a_i, b_i, c_i$ represent the parameters of the line equation represented by $l_i$, and $U_c^i, V_c^i$ represent the coordinates of the
center of this micro lens image in the raw image. Suppose $L$ represents the 3D line segment corresponding to $l_0 \sim l_k$, then the plane $\Pi^t$ must contain $L$ since $l_i$ is the image of $L$ in the micro lens image. For every 2D line segment $l_0 \sim l_k$, we can determine a set of $\Pi_0 \sim \Pi_k$. By calculating the intersection of these planes, we can get an initial solution of $L$. This can be achieved by solving the following linear equations.

$$
\begin{bmatrix}
A_0 & B_0 & C_0 & D_0 \\
A_1 & B_1 & C_1 & D_1 \\
\vdots & \vdots & \vdots & \vdots \\
A_k & B_k & C_k & D_k
\end{bmatrix} \begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix} = 0,
$$

(6)

The solution set of the system $(X, Y, Z, 1)$ represents the intersection of planes $\Pi_0 \sim \Pi_k$, i.e., the line $L$. Therefore, the dimension of the null space of the linear equations is 2.

We make singular valuable decomposition (SVD) for the coefficient matrix in (6) and obtain $U \Sigma V^\top$. And the two column vectors $V_3$ and $V_4$ of $V^\top$ which corresponds to the smallest and second-smallest eigenvalue of $\Sigma$ are just a set of basis of the solution space of (6). After normalizing the last term of $V_3$ and $V_4$ to $1$, $P_1 = (X_1, Y_1, Z_1, 1)$ and $P_2 = (X_1, Y_1, Z_1, 1)$ are acquired which are just two points on the line $L$.

**B. REFINED SOLUTION**

The $L$ calculated by the above method is not accurate due to the influence of the raw images’ blur. We further optimize $L$ as follows. First, we change the parameters of $L$ to generate a series of new 3D line segments. Then, every 3D line segment is re-projected to the sensor plane, and the similarity between the projection and the actual raw data in HorArea or VerArea is calculated. The most similar one is regarded as the final solution of $L$.

First of all, we show the parameterization of a 3D line segment. Taking the horizontal line $L^H$ as an example, suppose $P^H_1 = (X^H_1, Y^H_1, Z^H_1, 1)$ and $P^H_2 = (X^H_2, Y^H_2, Z^H_2, 1)$ represent two points on $L^H$ solved in the section IV-A, then $(Y^H_1, Z^H_1, Y^H_2, Z^H_2)$ can be regarded as a two-plane-parameterization (TPP) [24], [25]. These two planes are located at $X = X^H_1$ and $X = X^H_2$ respectively. In this section, $(Y^H_1, Z^H_1, Y^H_2, Z^H_2)$ is regarded as the parameterization of $L^H$.

Then we can obtain the refined solution for $L^H$ with the following steps:

1) Generating new 3D line segments.

Starting from the initial solution of $L^H$ solved in Section IV-A, by changing the four parameters $(Y^H_1, Z^H_1, Y^H_2, Z^H_2)$, a series of new 3D line segments are generated. Since every new generated 3D lines need to be projected to the raw image to calculate its similarity, rules of changing parameters of 3D lines should be determined by its 2D projections on the raw image.

Specifically, take $(Y^H_1, Z^H_1)$ which belongs to $P^H_1$ as an example. Then, rules of changing $(Y^H_1, Z^H_1)$ depend on 2D projection of $P^H_1$. Details of how to change $(Y^H_1, Z^H_1)$ are described as follows:

**Step1**

select a micro lens whose center is $\left(\tilde{X}_{mla}, \tilde{Y}_{mla}\right)$, and the micro lens image coordinate of $\tilde{p}_1$ is $\tilde{p}_1 = (U_{\tilde{p}_1}, V_{\tilde{p}_1})$. Then, the strategy of how to change $(Y^H_1, Z^H_1)$ in this paper is detailed as follows:

For a 3D point $P^H_1$, it makes its 2D projection $\tilde{p}_1$ change 0.25 pixel along y axis (i.e. $V_{\tilde{p}_1}$ change 0.25 pixel), to add a change to its $Z^H_{\tilde{p}_1}$ value or $Y^H_{\tilde{p}_1}$ value at every step.

According to this, the specific method of assigning value to $Z^H_{\tilde{p}_1}$ and $Y^H_{\tilde{p}_1}$ could be derived from Equation 3. First, the relationship between $(Y^H_1, Z^H_1)$ and the $V_{\tilde{p}_1}$ value of $\tilde{p}_1$ can be derived from Equation 3

$$
V_{\tilde{p}_1} = \frac{d}{d_{pix}} \cdot \frac{(Y^H_1 - Y^H_{mla})}{Z^H_1},
$$

(8)

which can be also expressed as

$$
Z^H_1 = \frac{(Y^H_1 - Y^H_{mla})}{V_{\tilde{p}_1} d_{pix}},
$$

(9)

or

$$
Y^H_1 = \frac{V_{\tilde{p}_1} Z^H_1 \cdot d_{pix}}{d} + Y^H_{mla},
$$

(10)

Thus, according to Equation 9 and Equation 10, the generation formulas of new $(\tilde{Y}^H_1, \tilde{Z}^H_1)$ can be expressed as,

$$
\tilde{Z}^H_1 = \frac{(Y^H_1 - Y^H_{mla})}{(V_{\tilde{p}_1} + \Delta V) d_{pix}},
$$

$$
\tilde{Y}^H_1 = \frac{(V_{\tilde{p}_1} + \Delta V) Z^H_1 \cdot d_{pix}}{d} + Y^H_{mla},
$$

$\Delta V = \pm 1 \times 0.25$Pixel, $\pm 2 \times 0.25$Pixel $\cdots$ (11)
Then, do the same operation to $(Y^H, Z^H)$, and a series of new 3D line segments are generated.

2) Projection to HorArea.

Suppose one of the newly generated 3D line segments is $L^*$, and $P_{1}^* = (X_1^*, Y_1^*, Z_1^*)$ and $P_{2}^* = (X_2^*, Y_2^*, Z_2^*)$ are two points on $L^*$. The 2D projections of $P_{1}^*$ and $P_{2}^*$ on the micro lens image whose center is $U_1^*$, $V_1^*$ in the HorArea, is calculated by the following equations.

\[
\begin{align*}
\tilde{p}_1 &= \frac{1}{Z_1^*} P_1^* = \frac{1}{Z_1^*} K_i^T [I | T^i] P_1^* \\
\tilde{p}_2 &= \frac{1}{Z_2^*} P_2^* = \frac{1}{Z_2^*} K_i^T [I | T^i] P_2^*
\end{align*}
\] (12)

where $P_i$ is the projection matrix of the i-th micro lens in the HorArea.

After acquiring $\tilde{p}_1^*$, $\tilde{p}_2^*$, (14) can be used to calculate the equations of 2D line features which is the projection of $L^*$ in this micro lens image.

\[
\begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{bmatrix} x_{\tilde{P}_1^*} \\ y_{\tilde{P}_1^*} \end{bmatrix} \times \begin{bmatrix} x_{\tilde{P}_2^*} \\ y_{\tilde{P}_2^*} \end{bmatrix}
\] (14)

where $(a^*, b^*, c^*)$ is the equation coefficient of the 2D line feature in this micro lens image.

3) Calculating similarity.

We use the normalized cross-correlation (NCC) to calculate the similarity between the 2D line feature $l^*$ in the micro lens image and the actual raw data. The sum of NCCs of the 2D line features of $L^*$ in all the micro lens images of the HorArea is considered as the similarity between the 3D line segment $L^*$ and the actual raw data, which is represented by NCC ($L^*$), i.e.

\[
NCC \left( L^* \right) = \sum_{j \in \text{HorArea}} NCC \left( l^*_j \right)
\] (15)

The $L^*$ with the largest NCC ($L^*$), i.e. the one most similar to the actual raw data, is chose as the final refined solution of the 3D line segment.

V. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed algorithm, we compared the results of the proposed algorithm on four different datasets. The first two datasets are from dataset A and dataset E in [17]. They are captured by Lytro camera, and the checkerboard sizes of them are 19 × 19 grid of 3.61 mm cells and 8 × 6 grid of 35.1 × 35.0 mm cells separately. These two datasets are termed Dataset A and Dataset B in this paper. The third is come from [22], which is captured by Lytro Illum camera and with a checkerboard size of 9 × 6 grid of 26.25 mm cells. This dataset is denoted as Dataset C in this paper. The last dataset is captured by ourselves using a Lytro Illum camera, with a checkerboard size of 7 × 11 grid of 29.92 mm cells. After photographing, we use MATLAB toolbox LFToolbox V0.4 designed by Dansereau et al. [17] to generate raw images, then use white images to remove vignetting. And this dataset is denoted as Dataset D in this paper.

Images in different datasets brings different degrees of difficulties for feature detection. As shown in Figure 3(a) and Figure 3(b), there are two images from Dataset A and Dataset D respectively. It can be seen that, comparing with the image in Dataset A, the image in Dataset D has a longer span of a single 3D line segment, and there are more 2D line features in micro lens images in the HorArea or VerArea. And it can be noticed that the image in Dataset A is more blurrier and has micro images in lower resolution compared to Dataset D, which brings a greater challenge to the detection of 3D line segments.

To visibly demonstrate the performance improvement of the refined solution relative to the initial solution, before comparing the final corner detection results, the 2D line features and 3D line segments calculated by the initial solution and refined solution on two datasets (Dataset A and Dataset D) are compared in Figure 5. Figure 5(a) and Figure 5(c) shows a chunk of raw image from Dataset A which is blurrier and in lower resolution, and it is termed chunk 1. Figure 5(b) and Figure 5(d) shows a chunk of raw image from Dataset D.
which is clearer and in higher resolution, and it is termed chunk 2.

From these two figures, it can be found that, the overall distribution of 2D line features of the initial solution is a little irregular, even a little messy in Dataset A (see Figure 5(a)), since they are detected independently in each micro lens image. But the overall distribution of 2D line features of the refined solution is regular since they are projections of the same 3D line segment in different micro lens images. Moreover, although the 2D line features of the initial solution and the refined solution both match the raw data very well, the 3D line segments calculated by them differ greatly, as shown in Figure 6(a), and Figure 6(c). This is because the initial solution locally matches each micro-lens best, while the refined solution globally matches the whole chunk best.

Then, we compare the final corner detection results of the proposed algorithm (including initial solution and refined solution) to other competitors. To our knowledge, for conventional light field camera, there’re no other algorithms which directly detect corners in raw image, so we choose some classic corner detecting algorithms—“Harris” and “MinEigen” algorithm—as the competitors which are widely used for images of traditional camera. Besides, methods based on sub-aperture images are not included in this section, since they could not directly give the locations of the corners in raw image.

In the four datasets, 291 3D corners are totally used to demonstrate the result of each algorithm. Among them, 217 3D corners are from Dataset A which correspond to 3105 2D corners in 3105 micro lens images, 24 3D corners are from Dataset B which correspond to 313 2D corners in 313 micro lens images, 11 3D corners are from Dataset D which correspond to 475 2D corners in 475 micro lens images and 39 3D corners are from Dataset D which correspond to 1436 2D corners in 1436 micro lens images. Since real corner locations in raw image are not known for actual data, we invite eight people to mark the corner locations manually, and the average of their result is selected as the ground truth. So, with the previously mentioned four methods (Harris, MinEigen, the initial solution of the proposed algorithm, the refined solution of the proposed algorithm) and the ground truth, there are totally five detection results in the experiment.

For Harris and MinEigen algorithm, it would bring some additional trouble to directly apply them in the raw image of light field camera. The first is “false corner”— areas between two adjacent micro lens images is often falsely identified as a corner, as shown in Figure 4. Another problem is “missed detection”— corners whose location is near the edge of a micro lens image or whose intensity gradient is not large enough are often failed to be detected, as shown in Figure 4. And we count the number of totally detected corners, which is shown in Table 1.

**TABLE 1. Number of detected corners.**

| Algorithm    | Detected Corners /Ground Truth | Detection Ratio |
|--------------|--------------------------------|-----------------|
| Dataset A    |                                |                 |
| Harris       | 2108/3105                      | 67.89%          |
| MinEigen     | 2045/3105                      | 77.46%          |
| Initial      | 2954/3105                      | 95.14%          |
| Refined      | 2952/3105                      | 95.07%          |
| Dataset B    |                                |                 |
| Harris       | 176/313                        | 56.23%          |
| MinEigen     | 252/313                        | 80.51%          |
| Initial      | 312/313                        | 99.68%          |
| Refined      | 288/313                        | 92.01%          |
| Dataset C    |                                |                 |
| Harris       | 443/475                        | 93.68%          |
| MinEigen     | 443/475                        | 93.26%          |
| Initial      | 475/475                        | 100.00%         |
| Refined      | 474/475                        | 99.78%          |
| Dataset D    |                                |                 |
| Harris       | 932/1436                       | 64.90%          |
| MinEigen     | 1085/1436                      | 75.56%          |
| Initial      | 1403/1436                      | 97.70%          |
| Refined      | 1364/1436                      | 94.99%          |
2D corners of each algorithm, which is termed “detected corners” in Table 1 and we also calculate the ratio of the “detected corners” of each algorithm to the number of ground truth, which is termed “detection ratio” in Table 1.

Finally, to measure the quantitative “detection error” for each algorithm, the deviation of every detected corner from its ground truth is calculated. And the mean value of the deviations of all corners, the maximum deviations and the minimum deviations are all shown in Table 2. Besides, in order to make a fair comparison of the detection error of each algorithm, the “false corner” and “missed detection” are both eliminated manually for Harris and MinEigen algorithm.

From Table 2, it can be noticed that, even after eliminated the “false corner and “missed detection”, the detection error of Harris and MinEigen algorithm are still large. More specifically, for dataset D which are clearer, the difference between Harris or MinEigen and the proposed algorithm are not so large. And for Dataset A and Dataset B which are blurrier, the detection errors of Harris and MinEigen become larger, while the proposed algorithm still achieves a good performance, which indicates the robustness of the proposed algorithm. And due to the further optimization, the refined solution gets a even better result than the initial solution.

**TABLE 2. Detection error (in pixels).**

| Algorithm | Mean ± Std | Maximum | Minimum |
|-----------|------------|---------|---------|
| **Dataset A** |            |         |         |
| Harris    | 1.5731 ± 0.9944 | 7.4064  | 0.0450  |
| MinEigen  | 1.2241 ± 0.7982 | 7.1946  | 0.0134  |
| Initial   | 0.4844 ± 0.2794 | 1.8027  | 0.0096  |
| Refined   | 0.3363 ± 0.1796 | 1.2615  | 0.0021  |
| **Dataset B** |            |         |         |
| Harris    | 2.9064 ± 1.8844 | 8.8054  | 0.5705  |
| MinEigen  | 2.6849 ± 1.1958 | 6.9473  | 0.4284  |
| Initial   | 0.7239 ± 0.1321 | 1.7908  | 0.0119  |
| Refined   | 0.3887 ± 0.0308 | 0.8275  | 0.0108  |
| **Dataset C** |            |         |         |
| Harris    | 1.0057 ± 1.4239 | 5.3379  | 0.0387  |
| MinEigen  | 1.0390 ± 1.4351 | 5.4106  | 0.0106  |
| Initial   | 0.4922 ± 0.0611 | 1.2363  | 0.0476  |
| Refined   | 0.3353 ± 0.0467 | 1.0995  | 0.0042  |
| **Dataset D** |            |         |         |
| Harris    | 0.8166 ± 0.6836 | 4.6485  | 0.0351  |
| MinEigen  | 0.6617 ± 0.7112 | 6.0998  | 0.0081  |
| Initial   | 0.5348 ± 0.4317 | 2.8507  | 0.0112  |
| Refined   | 0.2503 ± 0.1502 | 1.3367  | 0.0133  |
In this paper, 3D line segments of the checkerboard image in the image space of the main lens are detected by the initial solution first and then further optimized. The coordinates of 2D corners on the raw image are obtained by calculating the intersections of 3D line segments and re-projecting them to the raw data. Experimental results demonstrate that the corner detection algorithm is accurate and reliable in actual datasets with blurred and low resolution micro lens images. Our future work will focus on how to extend the corner detection algorithm from checkerboard images to natural images.

VI. CONCLUSION AND FUTURE WORK

In this paper, 3D line segments of the checkerboard image in the image space of the main lens are detected by the initial solution first and then further optimized. The coordinates of 2D corners on the raw image are obtained by calculating the intersections of 3D line segments and re-projecting them to the raw data. Experimental results demonstrate that the corner detection algorithm is accurate and reliable in actual datasets with blurred and low resolution micro lens images. Our future work will focus on how to extend the corner detection algorithm from checkerboard images to natural images.

REFERENCES

[1] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, “Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing,” *ACM Trans. Graph.*, vol. 26, no. 3, p. 69, Jul. 2007.

[2] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen, “Programmable aperture photography: Multiplexed light field acquisition,” *ACM Trans. Graph.*, vol. 27, no. 3, p. 55, Aug. 2008.

[3] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar, “Compressive light field photography using overcomplete dictionaries and optimized projections,” *ACM Trans. Graph.*, vol. 32, no. 4, p. 46, Jul. 2013.

[4] R. Ng, *Digital Light Field Photography*. Stanford, CA, USA: Stanford Univ., 2006.

[5] A. Lumsdaine and T. Georgiev, “The focused plenoptic camera,” in *Proc. IEEE Int. Conf. Comput. Photography (ICCP)*, Apr. 2009, pp. 1–8.

[6] Lytro. (2016). *The Lytro Camera*. Available: http://www.lytro.com

[7] Raytrix. (2016). *3D Light Field Camera Technology*. [Online]. Available: http://www.raytrix.de/

[8] H.-G. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y.-W. Tai, and I. S. Kweon, “Accurate depth map estimation from a lenslet light field camera,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (CVPR), Jun. 2015, pp. 1547–1555.

[9] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi, “Depth from combining defocus and correspondence using light-field cameras,” in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 673–680.

[10] S. Wanner and B. Goldluecke, “Globally consistent depth labeling of 4D light fields,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 41–48.

[11] S. Wanner and B. Goldluecke, “Variational light field analysis for disparity estimation and super-resolution,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 606–619, Mar. 2014.

[12] R. Ng, “Fourier slice photography,” *ACM Trans. Graph.*, vol. 24, no. 3, pp. 735–744, Jul. 2005.

[13] V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. Bing Kang, “Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 2331–2338.

[14] T. Yang, Y. Zhang, J. Yu, J. Li, W. Ma, X. Tong, R. Yu, and L. Ran, “All-in-focus synthetic aperture imaging,” in *Proc. Eur. Conf. Comput. Vis.*, Cham, Switzerland: Springer, 2014, pp. 1–15.

[15] F. Dong, S.-H. Ieng, X. Savatier, R. Etienne-Cummings, and R. Benosman, “Plenoptic cameras in real-time robotics,” *Int. J. Robot. Res.*, vol. 32, no. 2, pp. 206–217, Feb. 2013.

[16] N. Zeller, F. Quint, and U. Stilla, “From the calibration of a light-field camera to direct plenoptic odometry,” *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 1004–1019, Oct. 2017.

[17] D. G. Dansereau, O. Pizarro, and S. B. Williams, “Decoding, calibration and rectification for lenslet-based plenoptic cameras,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1027–1034.

[18] 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.*, vol. 41, no. 11, pp. 2539–2552, Nov. 2019.

[19] P. Zhou, W. Cai, Y. Yu, Y. Zhang, and G. Zhou, “A two-step calibration method of lenslet-based light field cameras,” *Opt. Lasers Eng.*, vol. 115, pp. 190–196, Apr. 2019.

[20] C.-A. Noury, C. Teuliere, and M. Dhome, “Light-field camera calibration from raw images,” in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Oct. 2017, pp. 1–8.

[21] S. Nousias, F. Chadebecq, J. Pichat, P. Keane, S. Ourselin, and C. Bergeles, “Corner-based geometric calibration of multi-focus plenoptic cameras,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 957–965.

[22] 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.*, vol. 39, no. 2, pp. 287–300, Feb. 2017.

[23] R. Hartley and A. Zisserman, *Multiple View Geometry Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[24] M. Levoy and P. Hanrahan, “Light field rendering,” in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn.*, 1996, pp. 31–42.

[25] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, “The lumigraph,” in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn.*, 1996, pp. 43–54.

DONGYANG JIN received the B.E. and M.E. degrees from Xidian University, Xi’an, China, in 2009 and 2013, respectively, where he is currently pursuing the Ph.D. degree with the State Key Laboratory of Integrated Services Networks. His research interests include computational photography, light field imaging, visual SLAM, and 3D reconstruction.
SAIPING ZHANG (Student Member, IEEE) received the B.E. degree from Xidian University, Xi’an, China, in 2016, where she is currently pursuing the Ph.D. degree with the State Key Laboratory of Integrated Services Networks. Her research interests include video coding and multimedia communication.

XIAO HUO received the B.E. degree from Xidian University, Xi’an, China, in 2019, where he is currently pursuing the master’s degree with the State Key Laboratory of Integrated Services Networks. His research interests include light field camera calibration, light field imaging, and 3D reconstruction.

FUZHENG YANG (Member, IEEE) received the B.E. degree in telecommunication engineering, the M.E. degree, and the Ph.D. degree in communication and information system from Xidian University, Xi’an, China, in 2000, 2003, and 2005, respectively. He was a Lecturer and an Associate Professor with Xidian University, in 2005 and 2006, respectively. From 2006 to 2007, he was a Visiting Scholar and a Postdoctoral Researcher with the Department of Electronic Engineering, Queen Mary University of London. He has been a Professor of communications engineering with Xidian University, since 2012. He is currently an Adjunct Professor with the School of Engineering, RMIT University. His research interests include video quality assessment, video coding, and multimedia communication.

* * *