

Blur Aware Calibration of Multi-Focus Plenoptic Camera

Mathieu Labussière¹, Céline Teulière¹, Frédéric Bernardin², Omar Ait-Aider¹

¹ Université Clermont Auvergne, CNRS, SIGMA Clermont,
 Institut Pascal, F-63000 Clermont-Ferrand, France

² Cerema, Équipe-projet STI, 10 rue Bernard Palissy, F-63017 Clermont-Ferrand, France

mathieu.labu@gmail.com, firstname.name@{uca,cerema}.fr

Abstract

This paper presents a novel calibration algorithm for Multi-Focus Plenoptic Cameras (MFPCs) using raw images only. The design of such cameras is usually complex and relies on precise placement of optic elements. Several calibration procedures have been proposed to retrieve the camera parameters but relying on simplified models, reconstructed images to extract features, or multiple calibrations when several types of micro-lens are used. Considering blur information, we propose a new Blur Aware Plenoptic (BAP) feature. It is first exploited in a pre-calibration step that retrieves initial camera parameters, and secondly to express a new cost function for our single optimization process. The effectiveness of our calibration method is validated by quantitative and qualitative experiments.

1. Introduction

The purpose of an imaging system is to map incoming light rays from the scene onto pixels of photo-sensitive detector. The radiance of a light ray is given by the *plenoptic function* $\mathcal{L}(\mathbf{x}, \theta, \lambda, \tau)$, introduced by Adelson *et al.* [1], where $\mathbf{x} \in \mathbb{R}^3$ is the *spatial* position of observation, $\theta \in \mathbb{R}^2$ is the *angular* direction of observation, λ is the wavelength of the light and τ is the time. Conventional cameras capture only one point of view. A *plenoptic camera* is a device that allows to retrieve spatial as well as angular information.

From *Lumigraph* [17] to commercial *plenoptic cameras* [19, 23], several designs have been proposed. This paper focuses on plenoptic cameras based on a Micro-Lenses Array (MLA) placed between the main lens and the photo-sensitive sensor (see Fig. 2). The specific design of such a camera allows to multiplex both types of information onto the sensor in the form of a Micro-Images Array (MIA) (see Fig. 1 (b)), but implies a trade-off between the angular and spatial resolutions [10, 16, 8]. It is balanced according to the MLA position with respect to the main lens focal plane (*i.e.*, *focused* [23, 9] and *unfocused* [19] configurations).

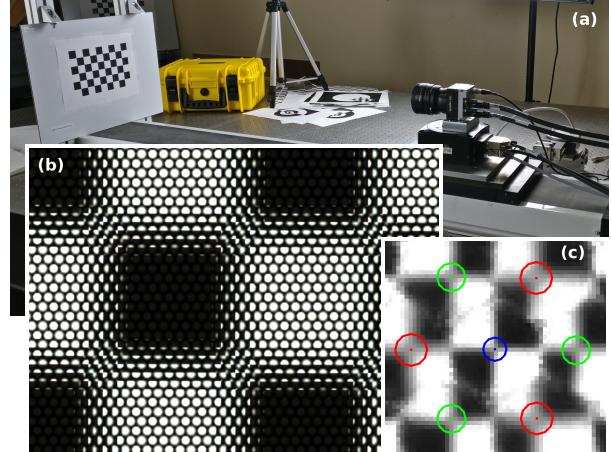


Figure 1: The Raytrix R12 multi-focus plenoptic camera used in our experimental setup (a), along with a raw image of a checkerboard calibration target (b). The image is composed of several micro-images with different blurred levels and arranged in an hexagonal grid. In each micro-image, our new Blur Aware Plenoptic (BAP) feature is illustrated by its center and its blur radius (c).

The mapping of incoming light rays from the scene onto pixels can be expressed as a function of the camera model. Classical cameras are usually modeled as pinhole or thin lens. Due to the complexity of plenoptic camera's design, the used models are usually high dimensional. Specific calibration methods have to be developed to retrieve the intrinsic parameters of these models.

1.1. Related Work

Unfocused plenoptic camera calibration. In this configuration, light rays are focused by the MLA on the sensor plane. The calibration of unfocused plenoptic camera [19] has been widely studied in the literature [6, 2, 25, 24, 11, 36]. Most approaches rely on a thin-lens model for the main lens and an array of pinholes for the micro-lenses. Most of them require reconstructed images to extract features, and limit their model to the unfocused configuration, *i.e.*, set-

ting the micro-lens focal length at the distance MLA-sensor. Therefore those models cannot be directly extended to the focused or multi-focus plenoptic camera.

Focused plenoptic camera calibration. With the arrival of commercial focused plenoptic cameras [18, 23], new calibration methods have been proposed. Based on [14], Heinze *et al.* [13] have developed a new projection model and a metric calibration procedure which is incorporated in the RxLive software of Raytrix GmbH. Other calibration methods and models have been proposed, either to overcome the fragility of the initialization [26], or to model more finely the camera parameters using depth information [33, 32]. O’Brien *et al.* [22] introduced a new 3D feature called plenoptic disc and defined by its center and its radius. Nevertheless, all previous methods rely on reconstructed images meaning that they introduce error in the reconstruction step as well as in the calibration process.

To overcome this problem, several calibration methods [35, 34, 29, 20, 3, 21, 31] have been proposed using only raw plenoptic images. In particular, features extraction in raw micro-images has been studied in [3, 21, 20] achieving improved performance through automation and accurate identification of feature correspondences. However, most of the methods rely on simplified models for optic elements: the MLA is modeled as a pinholes array making it impossible to retrieve the focal lengths, or the MLA misalignment is not considered. Some do not consider distortions [35, 21, 31] or restrict themselves to the focused case [35, 34, 20].

Finally, few have considered the multi-focus case [13, 3, 21, 31] but dealt with it in separate processes, leading to different intrinsic and extrinsic parameters according to the type of micro-lenses.

1.2. Contributions

This paper focuses on the calibration of micro-lenses-based Multi-Focus Plenoptic Camera (MFPC). To the best of our knowledge, this is the first method proposing a single optimization process that retrieves intrinsic and extrinsic parameters of a MFPC directly from raw plenoptic images. The main contributions are the following:

- We present a new Blur Aware Plenoptic (BAP) feature defined in raw image space that enables us to handle the multi-focus case.
- We introduce a new pre-calibration step using BAP features from white images to provide a robust initial estimate of internal parameters.
- We propose a new reprojection error function exploiting BAP features to refine a more complete model, including in particular the multiple micro-lenses focal lengths. Our checkerboard-based calibration is conducted in a single optimization process.

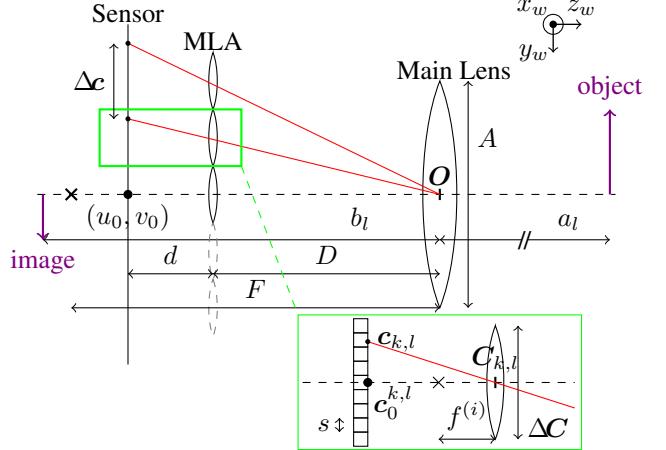


Figure 2: Focused Plenoptic Camera model in Galilean configuration (*i.e.*, the main lens focuses behind the sensor) with the notations used in this paper.

A visual overview of our method is given in Fig. 3. The remainder of this paper is organized as follows: first, the camera model and BAP feature are presented in Section 2. The proposed pre-calibration step is explained in Section 3. Then, the feature detection is detailed in Section 4 and the calibration process in Section 5. Finally, our results are presented and discussed in Section 6. The notations used in this paper are shown in Fig. 2. Pixel counterparts of metric values are denoted in lower-case Greek letters.

2. Camera model and BAP feature

2.1. Multi-focus Plenoptic Camera

We consider multi-focus plenoptic cameras as described in [9, 23]. The main lens, modeled as a thin-lens, maps object point to virtual point behind (*resp.*, in front of) the image sensor in Galilean (*resp.*, Keplerian) configuration. Therefore, the MLA consists of I different lens types with focal lengths $f^{(i)}, i \in \{1, \dots, I\}$ which are focused on I different planes behind the image sensor. This multi-focus setup corresponds to the Raytrix camera system described in [23] when $I = 3$. The micro-lenses are modeled as thin-lenses allowing to take into account blur in the micro-image. Our model takes into account the MLA misalignment, freeing all six degrees of freedom.

The tilt of the main lens is included in the distortion model and we make the hypothesis that the main lens plane Π_l is parallel to the sensor plane Π_s . Furthermore, we choose the main lens frame as our camera reference frame, with O being the origin, the z -axis coinciding with the optical axis and pointing outside the camera, and the y -axis pointing downwards. Only distortions of the main lens are considered. We use the model of Brown-Conradry [4, 5] with three coefficients for the radial component and two for the tangential.

Furthermore, we take into account the deviation of the image center and the optical center for each micro-lens because it tends to cause inaccuracy in decoded light field. Therefore, the principal point $\mathbf{c}_0^{k,l}$ of the micro-lens indexed by (k, l) is given by

$$\mathbf{c}_0^{k,l} = \begin{bmatrix} u_0^{k,l} \\ v_0^{k,l} \end{bmatrix} = \frac{d}{D+d} \left(\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} - \mathbf{c}_{k,l} \right) + \mathbf{c}_{k,l}, \quad (1)$$

where $\mathbf{c}_{k,l}$ is the center in pixels of the micro-image (k, l) , $[u_0 \ v_0]^\top$ is the main lens principal point, d is the distance MLA-sensor and D is the distance main lens-MLA, as illustrated in Fig. 2.

Finally, each micro-lens produces a micro-image onto the sensor. The set of these micro-images has the same structural organization as the MLA, *i.e.*, in our case an hexagonal grid, alternating between each type of micro-lens. The data can therefore be interpreted as an array of micro-images, called by analogy the MIA. The MIA coordinates are expressed in image space. Let $\Delta\mathbf{c}$ be the pixel distance between two arbitrary consecutive micro-images centers $\mathbf{c}_{k,l}$. With s the metric size of a pixel, let $\Delta C = s \cdot \Delta\mathbf{c}$ be its metric value, and ΔC be the metric distance between the two corresponding micro-lenses centers $C_{k,l}$. From similar triangles, the ratio between them is given by

$$\frac{\Delta C}{\Delta\mathbf{c}} = \frac{D}{d+D} \iff \Delta C = \Delta\mathbf{c} \cdot \frac{D}{d+D}. \quad (2)$$

We make the hypothesis that ΔC is equal to the micro-lens diameter. Since $d \ll D$, we can make the following approximation:

$$\frac{D}{D+d} = \lambda \approx 1 \implies \Delta C = \Delta\mathbf{c} \cdot \frac{D}{D+d} \approx \Delta\mathbf{c}. \quad (3)$$

This approximation will be validated in the experiments.

2.2. BAP feature and projection model

Using a camera with a circular aperture, the blurred image of a point on the image detector is circular in shape and is called the *blur circle*. From similar triangles and from the thin-lens equation, the signed blur radius of a point in an image can be expressed as

$$\rho = \frac{1}{s} \cdot \frac{A}{2} d \left(\frac{1}{f} - \frac{1}{a} - \frac{1}{d} \right), \quad (4)$$

with s the size of a pixel, d the distance between the considered lens and the sensor, A the aperture of this lens, f its focal length, and a the distance of the object from the lens.

This radius appears at different levels in the camera projection: in the blur introduced by the thin-lens model of the micro-lenses and during the formation of the micro-image while taking a white image. To leverage blur information,

we introduce a new Blur Aware Plenoptic (BAP) feature characterized by its center and its radius, *i.e.*, $\mathbf{p} = (u, v, \rho)$.

Therefore, our complete plenoptic camera model allows us to link a scene point \mathbf{p}_w to our new BAP feature \mathbf{p} through each micro-lens (k, l)

$$\begin{bmatrix} u \\ v \\ \rho \\ 1 \end{bmatrix} \propto \mathcal{P}(i, k, l) \cdot \mathbf{T}_\mu(k, l) \cdot \varphi(\mathbf{K}(F) \cdot \mathbf{T}_c \cdot \mathbf{p}_w), \quad (5)$$

where $\mathcal{P}(i, k, l)$ is the blur aware plenoptic projection matrix through the micro-lens (k, l) of type i , and computed as

$$\begin{aligned} \mathcal{P}(i, k, l) &= \mathbf{P}(k, l) \cdot \mathbf{K}(f^{(i)}) \\ &= \begin{bmatrix} d/s & 0 & u_0^{k,l} & 0 \\ 0 & d/s & v_0^{k,l} & 0 \\ 0 & 0 & s \frac{\Delta C}{2} & -s \frac{\Delta C}{2} d \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f^{(i)} & 1 \end{bmatrix}. \end{aligned} \quad (6)$$

$\mathbf{P}(k, l)$ is a matrix that projects the 3D point onto the sensor. $\mathbf{K}(f)$ is the thin-lens projection matrix for the given focal length. \mathbf{T}_c is the pose of the main lens with respect to the world frame and $\mathbf{T}_\mu(k, l)$ is the pose of the micro-lens (k, l) expressed in the camera frame. The function $\varphi(\cdot)$ models the lateral distortion.

Finally, the projection model defined in Eq. (5) consists of a set Ξ of $(16 + I)$ intrinsic parameters to be optimized, including the main lens focal length F and its 5 lateral distortion parameters, the sensor translation, encoded in (u_0, v_0) and d , the MLA misalignment, *i.e.*, 3 rotations $(\theta_x, \theta_y, \theta_z)$ and 3 translations (t_x, t_y, D) , the micro-lens inter-distance ΔC , and the I micro-lens focal lengths $f^{(i)}$.

3. Pre-calibration using raw white images

Drawing inspiration from *depth from defocus* theory [27], we leverage blur information to estimate parameters (*e.g.*, here our blur radius) by varying some other parameters (*e.g.*, the focal length, the aperture, etc.) in combination with known (*i.e.*, fixed or measured) parameters. For instance, when taking a white picture with a controlled aperture, each type of micro-lens produces a micro-image (MI) with a specific size and intensity, providing a way to distinguish between them. In the following all distances are given with reference to the MLA plane. Distances are signed according to the following convention: f is positive when the lens is convergent; distances are positive when the point is real, and negative when virtual.

3.1. Micro-image radius derivation

Taking a white image is equivalent for the micro-lenses to image a white uniform object of diameter A at a distance

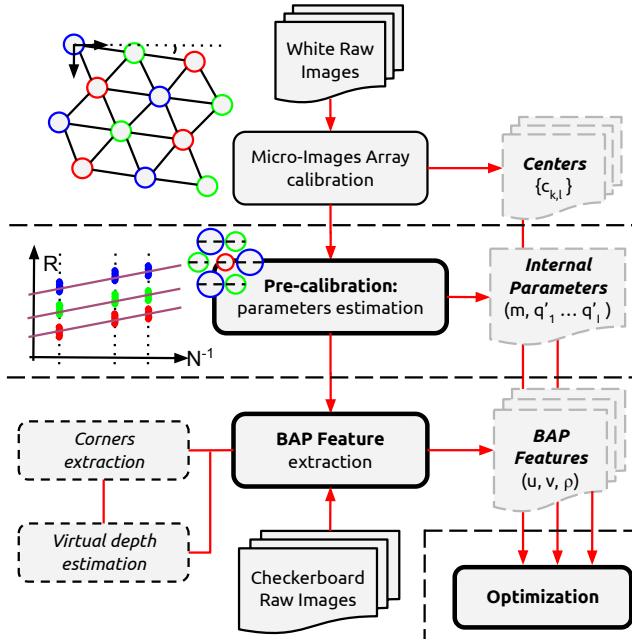


Figure 3: Overview of our proposed method with the pre-calibration step and the detection of BAP features that are used in the non-linear optimization process.

D . This is illustrated in Fig. 4. We relate the micro-image (MI) radius to the plenoptic camera parameters. From optics geometry, the image of this object, *i.e.* the resulting MI, is equivalent to the image of an imaginary point constructed as the vertex of the cone passing through the main lens and the considered micro-lens (noted V in Fig. 4). Let a' be the distance of this point from the MLA plane, given from similar triangles and Eq. (2) by

$$a' = -D \frac{\Delta C}{A - \Delta C} = -D \left(A \left(\frac{d + D}{\Delta c D} \right) - 1 \right)^{-1}, \quad (7)$$

with A the main lens aperture. Note the minus sign is due to the fact that the imaginary point is always formed behind the MLA plane at a distance a' , and thus considered as a virtual object for the micro-lenses. Conceptually, the MI formed can be seen as the *blur circle* of this imaginary point. Therefore, using Eq. (4), the metric MI radius R is given by

$$\begin{aligned} R &= \frac{\Delta C}{2} d \left(\frac{1}{f} - \frac{1}{a'} - \frac{1}{d} \right) \\ &= A \cdot \frac{d}{2D} + \left(\frac{\Delta c D}{d + D} \right) \cdot \frac{d}{2} \cdot \left(\frac{1}{f} - \frac{1}{D} - \frac{1}{d} \right). \end{aligned} \quad (8)$$

From the above equation, we see that the radius depends linearly on the aperture of the main lens. However, the main lens aperture cannot be computed directly whereas we have access to the *f*-number value. The *f*-number of an optical system is the ratio of the system's focal length F to the diameter of the entrance pupil, A , given by $N = F/A$.

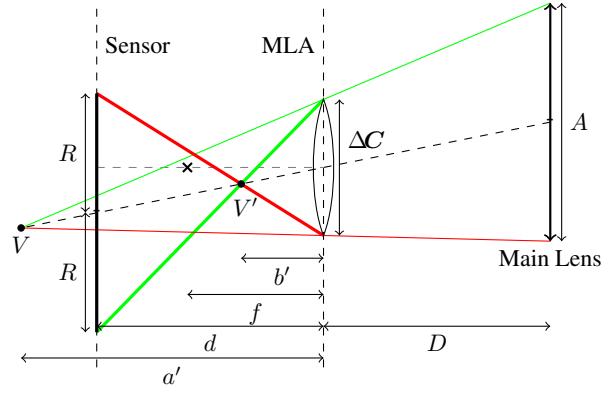


Figure 4: Formation of a micro-image with its radius R through a micro-lens while taking a white image at an aperture A . The point V is the vertex of the cone passing by the main lens and the considered micro-lens. V' is the image of V by the micro-lens and R is the radius of its blur circle.

Finally, we can express the MI radius for each micro-lens focal length type i as

$$R(N^{-1}) = m \cdot N^{-1} + q_i \quad (9)$$

with

$$m = \frac{dF}{2D} \quad \text{and} \quad q_i = \frac{1}{f^{(i)}} \cdot \left(\frac{\Delta c D}{d + D} \right) \cdot \frac{d}{2} - \frac{\Delta c}{2}. \quad (10)$$

Let q'_i be the value obtained by $q'_i = q_i + \Delta c/2$.

3.2. Internal parameters estimation

The *internal parameters* $\Omega = \{m, q'_1, \dots, q'_l\}$ are used to compute the radius part of the BAP feature and to initialize the parameters of the calibration process. Given several raw white images taken at different apertures, we estimate the coefficients of Eq. (9) for each type of micro-image. The standard full-stop *f*-number conventionally indicated on the lens differs from the real *f*-number calculated with the aperture value AV as $N = \sqrt{2^{AV}}$.

From raw white images, we are able to measure each micro-image (MI) radius $\varrho = |R|/s$ in pixels for each distinct focal length $f^{(i)}$ at a given aperture. Due to vignetting effect, the estimation is only conducted on center micro-images which are less sensitive to this effect. Our method is based on image moments fitting. It is robust to noise, works under asymmetrical distribution and is easy to use, but needs a parameter α to convert the standard deviation σ into a pixel radius $\varrho = \alpha \cdot \sigma$. We use the second order central moments of the micro-image to construct a covariance matrix. Finally, we choose σ as the square root of the greater eigenvalue of the covariance matrix. The parameter α is determined such that at least 98% of the distribution is taken into account. According to the standard normal distribution *Z*-score table, α is picked up in [2.33, 2.37]. In our experiments, we set $\alpha = 2.357$.

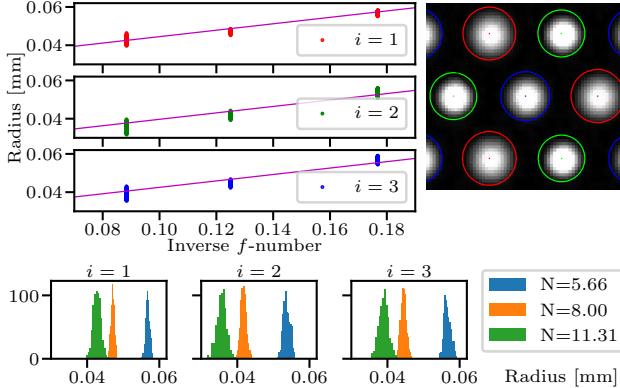


Figure 5: Micro-image radii as function of the inverse f -number with the estimated lines (in magenta). Each type of micro-lens is identified by its color (type (1) in red, type (2) in green and type (3) in blue) with its computed radius. For each type an histogram of the radii distribution is given.

Finally, from radii measurements at different f -numbers, we estimate the coefficients of Eq. (9), $\mathbf{X} = \{m, q_1, \dots, q_I\}$, with a least-square estimation. Fig. 5 shows the radii computed in a white image taken with an f -number $N = 8$, the histogram of these radii distributions for each type, and the estimated linear functions.

4. BAP feature detection in raw images

At this point, the internal parameters Ω , used to express our blur radius ρ , are available. The process (illustrated in Fig. 3) is divided into three phases: 1) using a white raw image, the MIA is calibrated and micro-image centers are extracted; 2) checkerboard images are processed to extract corners at position (u, v) ; and 3) with the internal parameters and the associated virtual depth estimate for each corner, the corresponding BAP feature is computed.

4.1. Blur radius derivation through micro-lens

To respect the f -number matching principle [23], we configure the main lens f -number such that the micro-images fully tile the sensor without overlap. In this configuration the working f -number of the main imaging system and the micro-lens imaging system should match. The following relation is then verified for at least one type of micro-lens:

$$\frac{\Delta C}{d} = \frac{A}{D} \iff \frac{1}{d} \cdot \left(\frac{\Delta c D}{d + D} \right) = \frac{F}{ND}. \quad (11)$$

We consider the general case of measuring an object p at a distance a_l from the main lens. First, p is projected through the main lens according to the thin lens equation, $1/F = 1/a_l + 1/b_l$, resulting in a point p' at a distance b_l behind the main lens, *i.e.* at a distance $a = D - b_l$ from the MLA. The metric radius of the blur circle r formed on the sensor

for a given point p' at distance a through a micro-lens of type i is expressed as

$$\begin{aligned} r &= \left(\frac{\Delta c D}{d + D} \right) \cdot \frac{d}{2} \cdot \left(\frac{1}{f^{(i)}} - \frac{1}{a} - \frac{1}{d} \right) \\ &= \underbrace{\frac{\Delta c D}{d + D} \cdot \frac{d}{2} \cdot \frac{1}{f^{(i)}}}_{=q'_i \text{ (10)}} - \underbrace{\frac{\Delta c D}{d + D} \cdot \frac{d}{2} \cdot \frac{1}{d}}_{=\Delta C/2 \text{ (2)}} - \underbrace{\frac{\Delta c D}{d + D} \cdot \frac{d}{2} \cdot \frac{1}{a}}_{=\Delta C \text{ (2)}} \\ &= \left(-\Delta C \cdot \frac{d}{2} \right) \cdot \frac{1}{a} + \left(q'_i - \frac{\Delta C}{2} \right). \end{aligned} \quad (12)$$

In practice, we do not have access to the value of ΔC but we can use the approximation from Eq. (3). Moreover, a and d cannot be measured in the raw image space, but the virtual depth can. Virtual depth refers to relative depth value obtained from disparity. It is defined as the ratio between the object distance a and the sensor distance d :

$$\nu = \frac{a}{d}. \quad (13)$$

The sign convention is reversed for virtual depth computation, *i.e.* distances are negative in front of the MLA plane. If we re-inject the latter in Eq. (12), taking caution of the sign, we can derive the radius of the *blur circle* of a point p' at a distance a from the MLA by

$$r = \frac{\lambda \Delta c}{2} \cdot \nu^{-1} + \left(q'_i - \frac{\lambda \Delta c}{2} \right). \quad (14)$$

This equation allows to express the pixel radius of the blur circle $\rho = r/s$ associated to each point having a virtual depth without explicitly evaluating A, D, d, F and $f^{(i)}$.

4.2. Features extraction

First, the Micro-Images Array has to be calibrated. We compute the micro-image centers observations $\{c_{k,l}\}$ by intensity centroid method with sub-pixel accuracy [30, 20, 28]. The distance between two micro-image centers Δc is then computed as the optimized edge-length of a fitted 2D grid mesh with a non-linear optimization. The pixel translation offset in the image coordinates, (τ_x, τ_y) , and the rotation around the $(-z)$ -axis, ϑ_z , are also determined during the optimization process. From the computed distance and the given size of a pixel, the parameter $\lambda \Delta c$ is computed.

Secondly, as we based our method on checkerboard calibration pattern, we detect corners in raw images using the detector introduced by Noury *et al.* [20]. The raw images are devignetted by dividing them by a white raw image taken at the same aperture.

With a plenoptic camera, contrarily to a classical camera, a point is projected into more than one observation onto the sensor. Due to the nature of the spatial distribution of the data, we used the DBSCAN algorithm [7] to identify the

clusters. We then associate each point with its cluster of observations.

Once each cluster is identified, we can compute the virtual depth ν from disparity. Given the distance ΔC_{1-2} between the centers of the micro-lenses C_1 and C_2 , *i.e.* the baseline, and the Euclidean distance $\Delta i = |\mathbf{i}_1 - \mathbf{i}_2|$ between images of the same point in corresponding micro-images, the virtual depth ν can be calculated with the intercept theorem:

$$\nu = \frac{\Delta C_{1-2}}{\Delta C_{1-2} - \Delta i} = \frac{\eta \Delta C}{\eta \Delta C - \Delta i} = \frac{\eta \lambda \Delta c}{\eta \lambda \Delta c - \Delta i}, \quad (15)$$

where $\Delta C = \lambda \Delta c$ is the distance between two consecutive micro-lenses and $\eta \geq 1.0$. Due to noise in corner detection, we use a median estimator to compute the virtual depth of the cluster taking into account all combinations of point pairs in the disparity estimation.

Finally, from internal parameters Ω and with the available virtual depth ν we can compute the BAP features using Eq. (14). In each frame n , for each micro-image (k, l) of type i containing a corner at position (u, v) in the image, the feature $p_{k,l}^n$ is given by

$$p_{k,l}^n = (u, v, \rho), \text{ with } \rho = r/s. \quad (16)$$

In the end, our observations are composed of a set of micro-image centers $\{c_{k,l}\}$ and a set of BAP features $\{p_{k,l}^n\}$ allowing us to introduce two reprojection error functions corresponding to each set of features.

5. Calibration process

To retrieve the parameters of our camera model, we use a calibration process based on non-linear minimization of a reprojection error. The calibration process is divided into three phases: 1) the intrinsics are initialized using the internal parameters Ω ; 2) the initial extrinsics are estimated from the raw checkerboard images; and 3) the parameters are refined with a non-linear optimization leveraging our new BAP features.

5.1. Parameters initialization

Optimization processes are sensitive to initial parameters. To avoid falling into local minima during the optimization process, the parameters have to be carefully initialized not too far from the solution.

First, the camera is initialized in *Keplerian* or *Galilean* configuration. First, given the camera configuration, the internal parameters Ω , the focus distance h , and from the Eq. (17) of [23], the following parameters are set as

$$d \leftarrow \frac{2mH}{F + 4m} \quad \text{and} \quad D \leftarrow H - 2d, \quad (17)$$

where H is given by

$$H = \left| \frac{h}{2} \left(1 \pm \sqrt{1 \pm 4 \frac{F}{h}} \right) \right|, \quad (18)$$

with positive (*resp.*, negative) sign in *Galilean* (*resp.*, *Keplerian*) configuration.

The focal length, and the pixel size s are set according to the manufacturer value. All distortion coefficients are set to zero. The principal point is set as the center of the image. The sensor plane is thus set parallel to the main lens plane, with no rotation, at a distance $- (D + d)$.

Seemingly, the MLA plane is set parallel to the main lens plane at a distance $-D$. From the pre-computed MIA parameters the translation takes into account the offsets $(-s\tau_x, -s\tau_y)$ and the rotation around the z -axis is initialized with $-\vartheta_z$. The micro-lenses inter-distance ΔC is set according to Eq. (2). Finally, from internal parameters Ω , the focal lengths are computed as follows:

$$f^{(i)} \leftarrow \frac{d}{2 \cdot q_i'} \cdot \Delta C. \quad (19)$$

5.2. Initial poses estimation

The camera poses $\{T_c^n\}$, *i.e.*, the extrinsic parameters are initialized using the same method as in [20]. For each cluster of observation, the barycenter is computed. Those barycenters can be seen as the projections of the checkerboard corners through the main lens using a standard pin-hole model. For each frame, the pose is then estimated using the Perspective-n-Point (PnP) algorithm [15].

5.3. Non-linear optimization

The proposed model allows us to optimize all the parameters into one single optimization process. We propose a new cost function Θ taking into account the blur information using our new BAP feature. The cost is composed of two main terms both expressing errors in the image space: 1) the blur aware plenoptic reprojection error, where, for each frame n , each checkerboard corner p_w^n is reprojected into the image space through each micro-lens (k, l) of type i according to the projection model of Eq. (5) and compared to its observations $p_{k,l}^n$; and 2) the micro-lens center reprojection error, where, the main lens center O is reprojected according to a pinhole model in the image space through each micro-lens (k, l) and compared to the detected micro-image center $c_{k,l}$.

Let $\mathcal{S} = \{\Xi, \{T_c^n\}\}$ be the set of intrinsic Ξ and extrinsic $\{T_c^n\}$ parameters to be optimized. The cost function $\Theta(\mathcal{S})$ is expressed as

$$\sum \|p_{k,l}^n - \Pi_{k,i}(p_w^n)\|^2 + \sum \|c_{k,l} - \Pi_{k,l}(O)\|^2. \quad (20)$$

The optimization is conducted using the Levenberg-Marquardt algorithm.

R12-A ($h = 450\text{mm}$)				R12-B ($h = 1000\text{mm}$)				R12-C ($h = \infty$)		
Unit	Initial	Ours	[20]	Initial	Ours	[20]	Initial	Ours	[20]	
F	[mm]	50	49.720	54.888	50	50.047	51.262	50	50.011	53.322
D	[mm]	56.66	56.696	62.425	52.11	52.125	53.296	49.38	49.384	52.379
ΔC	[μm]	127.51	127.45	127.38	127.47	127.45	127.40	127.54	127.50	127.42
$f^{(1)}$	[μm]	578.15	577.97	-	581.10	580.48	-	554.35	556.09	-
$f^{(2)}$	[μm]	504.46	505.21	-	503.96	504.33	-	475.98	479.03	-
$f^{(3)}$	[μm]	551.67	551.79	-	546.39	546.37	-	518.98	521.33	-
u_0	[pix]	2039	2042.55	2289.83	2039	1790.94	1759.29	2039	1661.95	1487.2
v_0	[pix]	1533	1556.29	1528.24	1533	1900.19	1934.87	1533	1726.91	1913.81
d	[μm]	318.63	325.24	402.32	336.84	336.26	363.17	307.93	312.62	367.40

Table 1: Initial intrinsic parameters for each dataset along with the optimized parameters obtained by our method and with the method of [20]. Some parameters are omitted for compactness.

6. Experiments and Results

We evaluate our calibration model quantitatively in a controlled environment and qualitatively when ground truth is not available.

6.1. Experimental setup

For all experiments we used a Raytrix R12 color 3D-light-field-camera, with a MLA of F/2.4 aperture. The mounted lens is a Nikon AF Nikkor F/1.8D with a 50mm focal length. The MLA organization is hexagonal, and composed of 176×152 (width \times height) micro-lenses with $I = 3$ different types. The sensor is a Basler beA4000-62KC with a pixel size of $s = 0.0055\text{mm}$. The raw image resolution is 4080×3068 .

Datasets. We calibrate our camera for three different focus distance configurations hand build three corresponding datasets: R12-A for $h = 450\text{ mm}$, R12-B for $h = 1000\text{ mm}$, and R12-C for $h = \infty$. Each dataset is composed of:

- white raw plenoptic images acquired at different apertures ($N \in \{4, 5.66, 8, 11.31, 16\}$) with augmented gain to ease circle detection in the pre-calibration step,
- free-hand calibration targets acquired at various poses (in distance and orientation), separated into two subsets, one for the calibration process and the other for the qualitative evaluation,
- a white raw plenoptic image acquired in the same luminosity condition and with the same aperture as in the calibration targets acquisition,
- and calibration targets acquired with a controlled translation motion for quantitative evaluation, along with the depth maps computed by the Raytrix software (RxLive v4.0.50.2).

We use a 9×5 of 10mm side checkerboard for R12-A, a 8×5 of 20mm for R12-B, and a 7×5 of 30mm for R12-C. Datasets and our source code are publicly available¹.

¹ <https://github.com/comsee-research>

Free-hand camera calibration. The white raw plenoptic image is used for devignetting other raw images and for computing micro-images centers. From the set of calibration targets images, BAP features are extracted, and camera intrinsic and extrinsic parameters are then computed using our non-linear optimization process.

Controlled environment evaluation. In our experimental setup (see Fig. 1), the camera is mounted on a linear motion table with micro-metric precision. We acquired several images with known relative motion between each frame. Therefore, we are able to quantitatively evaluate the estimated displacement from the extrinsic parameters with respect to the ground truth. The extrinsics are computed with the intrinsics estimated from the free-hand calibration. We compared our computed relative depth to those obtained by the RxLive software.

Qualitative evaluation. When no ground truth is available, we evaluate qualitatively our parameters on the evaluation subset by estimating the reprojection error using the previously computed intrinsics. We use the Root-Mean-Square Error (RMSE) as our metric to evaluate the reprojection, individually on the corner reprojection and the blur radius reprojection.

Comparison. Since our model is close to [20], we compare our intrinsics with the ones obtained under their pin-hole assumption using only corner reprojection error and with the same initial parameters. We also provide the calibration parameters obtained from the RxLive software.

6.2. Results

Internal parameters. Note that the pixel MI radius is given by $\varrho = |R|/s$, and R is either positive if formed after the rays inversion (as in Fig. 4), or negative if before. With our camera $f^{(i)} > d$ [12], so $R < 0$ implying that m and c_i are also negative. In practice, it means we use the value $-\varrho$ in the estimation process.

In our experiments, we set $\lambda = 0.9931$. We verified that the error introduced by this approximation is less than the metric size of a pixel (*i.e.*, less than 1 % of its optimized value). Fig. 5 shows the estimated lines (see Eq. (9)), with a cropped white image taken at $N = 8$, where each type of micro-lens is identified with its radius. An histogram of the radii distribution is also given for each type. From the radii measurements, the computed internal parameters are estimated and given for each dataset in Tab. 2.

	R12-A	R12-B	R12-C
Δc	128.222	128.293	128.333
m	-140.596	-159.562	-155.975
q'_1	35.135	36.489	35.443
q'_2	40.268	42.075	41.278
q'_3	36.822	38.807	37.858
m^*	-142.611	-161.428	-158.292
ϵ_m	1.41%	1.16%	1.52%

Table 2: Internal parameters (in μm) computed during the pre-calibration step for each dataset. The expected value for m is given by m^* and the relative error $\epsilon_m = (m^* - m)/m^*$ is computed.

As expected, the internal parameters Δc and m are different for each dataset, as D and Δc vary with the focus distance h , whereas the q'_i values are close for each dataset. Using intrinsics, we compare the expected coefficient m^* (from the closed form in Eq. (10) with the optimized parameters) with its calculated m value. The mean error over all datasets is $\bar{\epsilon}_m = 1.36\%$ which is smaller than a pixel.

Free-hand camera calibration. The initial parameters for each dataset are given in Tab. 1 along with the optimized parameters obtained from our calibration and from the method in [20]. Some parameters are omitted for compactness (*i.e.*, distortions coefficient and MLA rotations which are negligible). The main lens focal lengths obtained with the proprietary RxLive software are: $F_{h=450} = 47.709$ mm, $F_{h=1000} = 50.8942$ mm, and $F_{h=\infty} = 51.5635$ mm.

With our method and [20], the optimized parameters are close to their initial value, showing that our method provides a good initialization for our optimization process. The F , d and ΔC are consistent across the datasets with our method. In contrast, the F and d obtained with [20] show a larger discrepancy. This is also the case for the focal lengths obtained by RxLive.

Poses evaluation. Tab. 3 presents the relative translations and their errors with respect to the ground truth for the controlled environment experiment. Even if our relative errors are similar with [20], we are able to retrieve more parameters. With our method, absolute errors are of the order of the mm (R12-A: 0.37 ± 0.15 mm, R12-B: 1.66 ± 0.58 mm and R12-C: 1.38 ± 0.85 mm) showing that the retrieved scale is coherent. Averaging over all the datasets,

our method presents the smallest relative error, with low discrepancy between datasets, outperforming the estimations of the RxLive software.

Error [%]	R12-A		R12-B		R12-C		All
	$\bar{\epsilon}_z$	σ_z	$\bar{\epsilon}_z$	σ_z	$\bar{\epsilon}_z$	σ_z	$\bar{\epsilon}_z$
Ours	3.73	1.48	3.32	1.17	2.95	1.35	3.33
[20]	6.83	1.17	1.16	1.06	2.70	0.86	3.56
RxLive	4.63	2.51	4.26	5.79	11.52	3.22	6.80

Table 3: Relative translation error along the z -axis with respect to the ground truth displacement. For each dataset, the mean error $\bar{\epsilon}_z$ and its standard deviation σ_z are given. Results are given for our method and compared with [20] and the proprietary software RxLive.

Reprojection error evaluation. For each evaluation dataset, the total squared pixel error is reported with its computed RMSE in Tab. 4. The error is less than 1pix per feature for each dataset demonstrating that the computed intrinsics are valid and can be generalized to images different from the calibration set.

	R12-A (#11424)		R12-B (#3200)		R12-C (#9568)	
	Total	RMSE	Total	RMSE	Total	RMSE
$\bar{\epsilon}_{all}$	8972.91	0.886	1444.98	0.672	5065.33	0.728
$\bar{\epsilon}_{u,v}$	8908.65	0.883	1345.20	0.648	5046.68	0.726
$\bar{\epsilon}_p$	64.257	0.075	99.780	0.177	18.659	0.044

Table 4: Reprojection error for each evaluation dataset with their number of observations. For each component of the feature, the total squared pixel error is reported with its computed RMSE.

7. Conclusion

To calibrate the Multi-Focus Plenoptic Camera, state-of-the-art methods rely on simplifying hypotheses, on reconstructed data or require separate calibration processes to take into account the multi-focal aspect. This paper introduces a new pre-calibration step which allows us to compute our new BAP feature directly in the raw image space. We then derive a new projection model and a new reprojection error using this feature. We propose a single calibration process based on non-linear optimization that enables us to retrieve camera parameters, in particular the micro-lenses focal lengths. Our calibration method is validated by qualitative experiments and quantitative evaluations. In the future, we plan to exploit this new feature to improve metric depth estimation.

Acknowledgments. This work has been supported by the AURA Region and the European Union (FEDER) through the MMII project of CPER 2015-2020 MMASyF challenge. We thank Charles-Antoine Noury and Adrien Coly for their insightful discussions and their help during the acquisitions.

References

- [1] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, pages 3–20, 1991. 1
- [2] Yunsu Bok, Hae-Gon Jeon, and In So Kweon. Geometric Calibration of Micro-Lens-Based Light-Field Cameras Using Line Features. In *Computer Vision – ECCV 2014*, pages 47–61. Springer International Publishing, 2014. 1
- [3] Yunsu Bok, Hae-Gon Jeon, and In So Kweon. Geometric Calibration of Micro-Lens-Based Light Field Cameras Using Line Features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2):287–300, 2017. 2
- [4] Duane Brown. Decentering Distortion of Lenses - The Prism Effect Encountered in Metric Cameras can be Overcome Through Analytic Calibration. *Photometric Engineering*, 32(3):444–462, 1966. 2, 11
- [5] Ae Conrady. Decentered Lens-Systems. *Monthly Notices of the Royal Astronomical Society*, 79:384–390, 1919. 2, 11
- [6] Donald G. Dansereau, Oscar Pizarro, and Stefan B. Williams. Decoding, calibration and rectification for lenslet-based plenoptic cameras. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1027–1034, 2013. 1
- [7] Martin Ester, Hans-Peter Kriegel, Jürg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *KDD*, 1996. 5
- [8] Todor Georgiev and Andrew Lumsdaine. Resolution in Plenoptic Cameras. *Frontiers in Optics 2009/Laser Science XXV/Fall 2009 OSA Optics & Photonics Technical Digest*, page CTuB3, 2009. 1
- [9] Todor Georgiev and Andrew Lumsdaine. The multifocus plenoptic camera. In *Digital Photography VIII*, pages 69–79. International Society for Optics and Photonics, SPIE, 2012. 1, 2
- [10] Todor Georgiev, Ke Colin Zheng, Brian Curless, David Salesin, Shree K Nayar, and Chintan Intwala. Spatio-Angular Resolution Tradeoff in Integral Photography. *Rendering Techniques*, 2006(263-272):21, 2006. 1
- [11] Christopher Hahne, Andrew Lumsdaine, Amar Aggoun, and Vladan Velisavljevic. Real-time refocusing using an FPGA-Based standard plenoptic camera. *IEEE Transactions on Industrial Electronics*, 65(12):9757–9766, 2018. 1
- [12] Christian Heinze. *Design and test of a calibration method for the calculation of metrical range values for 3D light field cameras*. Master’s thesis, Hamburg University of Applied Sciences - Faculty of Engineering and Computer Science, 2014. 7
- [13] Christian Heinze, Stefano Spyropoulos, Stephan Hussmann, and Christian Perwass. Automated Robust Metric Calibration Algorithm for Multifocus Plenoptic Cameras. *IEEE Transactions on Instrumentation and Measurement*, 65(5):1197–1205, 2016. 2
- [14] Ole Johannsen, Christian Heinze, Bastian Goldluecke, and Christian Perwass. On the calibration of focused plenoptic cameras. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8200 LNCS:302–317, 2013. 2
- [15] Laurent Kneip, Davide Scaramuzza, and Roland Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2969–2976, 2011. 6
- [16] Anat Levin, William T. Freeman, and Frédo Durand. Understanding camera trade-offs through a Bayesian analysis of light field projections. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5305 LNCS(PART 4):88–101, 2008. 1
- [17] Gabriel Lippmann. Integral Photography. *Academy of the Sciences*, 1911. 1
- [18] Andrew Lumsdaine and Todor Georgiev. The focused plenoptic camera. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8, apr 2009. 2
- [19] Ren Ng, Marc Levoy, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light Field Photography with a Hand-held Plenoptic Camera. Technical report, Stanford University, 2005. 1, 12
- [20] Charles Antoine Noury, Céline Teulière, and Michel Dhome. Light-Field Camera Calibration from Raw Images. *DICTA 2017 – International Conference on Digital Image Computing: Techniques and Applications*, pages 1–8, 2017. 2, 5, 6, 7, 8, 13, 14
- [21] Sotiris Nousias, Francois Chadebecq, Jonas Pichat, Pearse Keane, Sébastien Ourselin, and Christos Bergeles. Corner-Based Geometric Calibration of Multi-focus Plenoptic Cameras. *Proceedings of the IEEE International Conference on Computer Vision*, pages 957–965, 2017. 2
- [22] Sean O’Brien, Jochen Trumpf, Viorela Ila, and Robert Mahony. Calibrating light-field cameras using plenoptic disc features. In *2018 International Conference on 3D Vision (3DV)*, pages 286–294. IEEE, 2018. 2
- [23] Christian Perwass and Lennart Wietzke. Single Lens 3D-Camera with Extended Depth-of-Field. In *Human Vision and Electronic Imaging XVII*, volume 49, page 829108. SPIE, feb 2012. 1, 2, 5, 6, 12
- [24] Shengxian Shi, Junfei Ding, T. H. New, You Liu, and Hanmo Zhang. Volumetric calibration enhancements for single-camera light-field PIV. *Experiments in Fluids*, 60(1):21, jan 2019. 1
- [25] Shengxian Shi, Jianhua Wang, Junfei Ding, Zhou Zhao, and T. H. New. Parametric study on light field volumetric particle image velocimetry. *Flow Measurement and Instrumentation*, 49:70–88, 2016. 1
- [26] Klaus H. Strobl and Martin Lingenauber. Stepwise calibration of focused plenoptic cameras. *Computer Vision and Image Understanding*, 145:140–147, 2016. 2
- [27] Murali Subbarao and Gopal Surya. Depth from defocus: A spatial domain approach. *International Journal of Computer Vision*, 13(3):271–294, 1994. 3
- [28] Piotr Suliga and Tomasz Wrona. Microlens array calibration method for a light field camera. *Proceedings of the 19th International Carpathian Control Conference (ICCC)*, pages 19–22, 2018. 5

- [29] Jun Sun, Chuanlong Xu, Biao Zhang, Shimin Wang, Md Moinul Hossain, Hong Qi, and Heping Tan. Geometric calibration of focused light field camera for 3-D flame temperature measurement. In *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, July 2016. 2
- [30] Chelsea M. Thomason, Brian S. Thurow, and Timothy W. Fahringer. Calibration of a Microlens Array for a Plenoptic Camera. *52nd Aerospace Sciences Meeting*, (January):1–18, 2014. 5
- [31] Yuan Wang, Jun Qiu, Chang Liu, Di He, Xinkai Kang, Jian Li, and Ligen Shi. Virtual Image Points Based Geometrical Parameters' Calibration for Focused Light Field Camera. *IEEE Access*, 6(c):71317–71326, 2018. 2
- [32] Niclas Zeller, Charles Antoine Noury, Franz Quint, Céline Teulière, Uwe Stilla, and Michel Dhome. Metric Calibration of a Focused Plenoptic Camera based on a 3D Calibration Target. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3(July):449–456, jun 2016. 2
- [33] Niclas Zeller, Franz Quint, and Uwe Stilla. Calibration and accuracy analysis of a focused plenoptic camera. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3(September):205–212, 2014. 2
- [34] Chunping Zhang, Zhe Ji, and Qing Wang. Decoding and calibration method on focused plenoptic camera. *Computational Visual Media*, 2(1):57–69, 2016. 2
- [35] Chunping Zhang, Zhe Ji, and Qing Wang. Unconstrained Two-parallel-plane Model for Focused Plenoptic Cameras Calibration. pages 1–20, 2016. 2
- [36] Ping Zhou, Weijia Cai, Yunlei Yu, Yuting Zhang, and Guangquan Zhou. A two-step calibration method of lenslet-based light field cameras. *Optics and Lasers in Engineering*, 115:190–196, 2019. 1

(Supplementary Material)

Blur Aware Calibration of Multi-Focus Plenoptic Camera

Mathieu Labussière¹, Céline Teulière¹, Frédéric Bernardin², Omar Ait-Aider¹

¹ Université Clermont Auvergne, CNRS, SIGMA Clermont,
Institut Pascal, F-63000 Clermont-Ferrand, France

² Cerema, Équipe-projet STI, 10 rue Bernard Palissy, F-63017 Clermont-Ferrand, France

mathieu.labu@gmail.com, firstname.name@{uca,cerema}.fr

As supplementary material, we first provide details about the camera model, second we detail in a more complete fashion the equations for the MI radii computation, third we complete the calibration results with the previously omitted parameters, and finally we give insight about the dataset images and poses.

A. Camera model

For completeness, the notations used in this paper are summarized in Fig. 6.

A.1. Main lens distortions

Recall that in our model, only distortions of the main lens are considered. Distortions represent deviations from the theoretical thin lens projection model. To correct those errors, we can undistort a distorted point $\mathbf{p} = [x \ y \ z]^\top$ by applying a function φ to it, such as $\mathbf{p} \xrightarrow{\varphi} \mathbf{p}_u$. To model the radial $\varphi^{(r)}$ and tangential $\varphi^{(t)}$ components of the lateral distortion model, we use the model of Brown-Conradry [4, 5]. The radial component is thus expressed as

$$\begin{cases} x_u^{(r)} = x \cdot (1 + A_0\kappa^2 + A_1\kappa^4 + A_2\kappa^6) \\ y_u^{(r)} = y \cdot (1 + A_0\kappa^2 + A_1\kappa^4 + A_2\kappa^6), \end{cases} \quad (21)$$

and the tangential component as

$$\begin{cases} x_u^{(t)} = B_0(\kappa^2 + 2x^2) + 2B_1xy \\ y_u^{(t)} = B_1(\kappa^2 + 2y^2) + 2B_0xy, \end{cases} \quad (22)$$

where $\kappa = \sqrt{x^2 + y^2}$.

Finally, our set of intrinsic Ξ includes 5 lateral distortion parameters: three coefficients for the radial component, $\mathbf{k}^{(r)} = \{A_0, A_1, A_2\}$, and two for the tangential, $\mathbf{k}^{(t)} = \{B_0, B_1\}$.

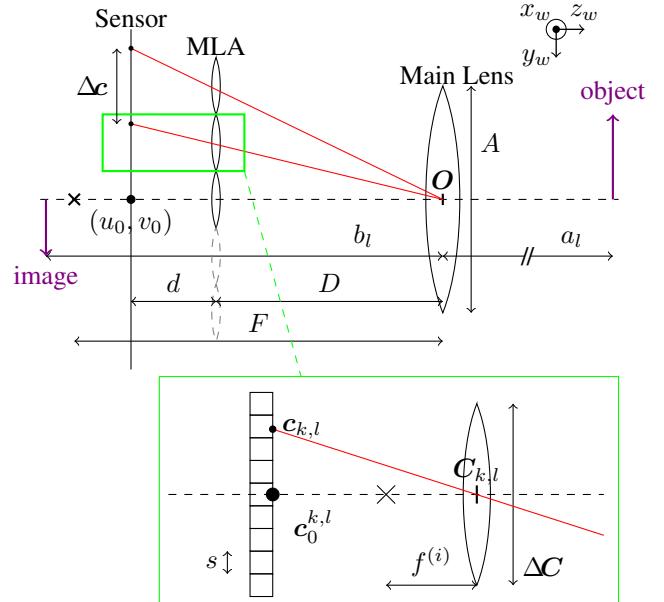


Figure 6: Focused Plenoptic Camera model in Galilean configuration (*i.e.*, the main lens focuses behind the sensor) with the notations used in this paper. Pixel counterparts of metric values are denoted in lower-case Greek letters.

A.2. F-number matching principle

The f -number of an optical system is the ratio between the system's focal length F and the diameter of the entrance pupil, A , given by

$$N = \frac{F}{A}. \quad (23)$$

The f -number accurately describes the light-gathering ability of a lens only for objects at infinite distance away. In optical design, an alternative is often needed for systems where the object is not far from the lens. In these cases

the working f -number is used. The working f -number is defined as

$$N^* \approx \frac{1}{2\text{NA}} \approx (1 + |\gamma|) N, \quad (24)$$

where NA is the numerical aperture, *i.e.*, the number that characterizes the range of angles over which the system can accept or emit light, and γ is the magnification of the current focus setting. Let a be the quantity that verifies the following thin lens equation:

$$\frac{1}{F} = \frac{1}{a} + \frac{1}{D}, \quad (25)$$

where F is the focal length and D is the distance between the sensor and the lens. Then we can express the magnification as

$$\gamma = \frac{D}{a}, \quad (26)$$

and thus we can rearrange Eq. (23) and Eq. (24) to obtain the working f -number expressed as:

$$N^* = \left(1 + \frac{D}{a}\right) \frac{F}{A} = D \left(\frac{1}{D} + \frac{1}{a}\right) \frac{F}{A} = \frac{D}{A}. \quad (27)$$

The fundamental design principle for light-field imaging is that the working f -numbers of the micro-lenses and the main lens are matched. This condition maximizes the fill factor of the sensor while avoiding overlap between micro-images [19]. Both unfocused and focused plenoptic camera designs follow the f -number matching principle. As highlighted in [23], the micro images generated by the micro-lenses in a plenoptic camera should just touch to make the best use of the image sensor, meaning

$$\frac{d}{\Delta C} = \frac{D_s - d}{A} \iff N_\mu^* = N_l^* - \frac{d}{A}, \quad (28)$$

where d is the distance between the MLA and the sensor, $D_s = D + d$ is the distance between the main lens and the sensor, ΔC and A are respectively the diameters of the micro-lenses and the main lens, and, N_μ^* and N_l^* are respectively the working f -numbers of the micro-lenses and the main lens.

Since typically $d \ll A$, we have $N_\mu^* \approx N_l^*$. So, the working f -numbers of the main imaging system and the micro lens imaging system should match. This also implies that the design of the micro lenses fixes the f -number of the main lens that is used with the plenoptic camera.

A.3. Projection model

The blur aware plenoptic projection matrix $\mathcal{P}(i, k, l)$ through the micro-lens (k, l) of type i is computed as

$$\begin{aligned} \mathcal{P}(i, k, l) &= \mathcal{P}(k, l) \cdot \mathbf{K}\left(f^{(i)}\right) \\ &= \begin{bmatrix} d/s & 0 & u_0^{k,l} & 0 \\ 0 & d/s & v_0^{k,l} & 0 \\ 0 & 0 & s \frac{\Delta C}{2} & -s \frac{\Delta C}{2} d \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f^{(i)} & 1 \end{bmatrix} \\ &= \begin{bmatrix} d/s & 0 & u_0^{k,l} & 0 \\ 0 & d/s & v_0^{k,l} & 0 \\ 0 & 0 & s \frac{\Delta C}{2} d \left(\frac{1}{d} - \frac{1}{f^{(i)}}\right) & -s \frac{\Delta C}{2} d \\ 0 & 0 & -1 & 0 \end{bmatrix}, \end{aligned} \quad (29)$$

where d is the distance between the MLA and the sensor, s is the size of a pixel, $c_0^{k,l} = [u_0^{k,l} \ v_0^{k,l}]^\top$ is the principal point of the micro-lens (k, l) , ΔC is the diameter of the micro-lens, and $f^{(i)}$ is the micro-lens focal length.

B. Micro-image radii computation

B.1. Image moments fitting

From raw white images, we measure each micro-image (MI) radius $\varrho = |R|/s$ in pixel based on image moments fitting. We use the second order central moments of the micro-image to construct a covariance matrix. Raw moments and centroid are given by

$$M_{ij} = \sum_{x,y} x^i y^j \mathcal{I}(x, y) \quad \text{and} \quad \{\bar{x}, \bar{y}\} = \left\{ \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right\},$$

and the central moments are given by

$$\mu_{pq} = \sum_{x,y} (x - \bar{x})^p (y - \bar{y})^q \mathcal{I}(x, y). \quad (30)$$

The covariance matrix is then given by

$$\text{cov}[\mathcal{I}(x, y)] = \frac{1}{\mu_{00}} \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix}. \quad (31)$$

We choose σ as the square root of the greater eigenvalue of the covariance matrix, *i.e.*,

$$\sigma^2 = \lambda_{\max} = \frac{\sigma_{xx} + \sigma_{yy}}{2} + \frac{\sqrt{4\sigma_{xy}^2 + (\sigma_{xx} - \sigma_{yy})^2}}{2}. \quad (32)$$

Finally, the radius ϱ is proportional to the computed standard deviation σ .

Type	$N = 5.66$	$N = 8$	$N = 11.31$
$i = 1$	56.79 ± 0.56	47.08 ± 0.53	42.76 ± 1.04
$i = 2$	53.68 ± 0.95	41.71 ± 0.78	35.68 ± 1.36
$i = 3$	56.58 ± 0.91	44.46 ± 0.73	38.98 ± 1.32

Table 5: Statistics (mean \pm std) over radii measurements (in μm) for each type of micro-image at different apertures.

B.2. Radii distribution analysis

An analysis of the micro-image radii distribution is given for three apertures $N \in \{5.66, 8, 11.31\}$ in Tab. 5 for the dataset R12-A.

As expected, the radius decreases whilst the f -number N increases. The standard deviation is less than one-fifth of a pixel, meaning that our method provides precise results.

B.3. Coefficients estimation

From radii measurements at different f -numbers, we want to estimate the coefficients $\mathbf{X} = [m \quad q_1 \quad \dots \quad q_I]^\top$ of

$$R(N^{-1}) = m \cdot N^{-1} + q_i$$

with

$$m = \frac{dF}{2D} \quad \text{and} \quad q_i = \frac{1}{f^{(i)}} \cdot \left(\frac{\Delta c D}{d+D} \right) \cdot \frac{d}{2} - \frac{\Delta c}{2}.$$

Note that m is a function of fixed physical parameters independent of the micro-lenses focal lengths and the main lens aperture. Therefore, we obtain a set of linear equations, sharing the same slope, but with different y -intercepts. This set of equations can be rewritten as

$$\mathbf{A}\mathbf{X} = \mathbf{B}, \text{ and then } \mathbf{X} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{B}$$

where the matrix \mathbf{A} (containing the f -numbers and a selector of the corresponding y -intercept coefficient) and \mathbf{B} (containing the radii measurements) are constructed by arranging the terms given the focal length at which they have been calculated. Finally, we compute \mathbf{X} with a least-square estimation.

C. Supplements on calibration results

Some parameters have been omitted in the main paper for compactness. Tab. 6 presents the complete set of intrinsic parameters with their initial and optimized values for our method and compared to [20]. As stated, the distortions coefficients are really low, and the MLA rotations are negligible. Indeed, in the case of industrial plenoptic cameras, such as Raytrix ones, a careful attention is payed to the co-planarity of the MLA and the sensor plane.

D. Insight about the datasets

We have presented three datasets in the submitted work: R12-A, R12-B, and R12-C. The devignetted images of the calibration targets from the dataset R12-A (Fig. 7), R12-B (Fig. 9), and R12-C (Fig. 11) taken at various angles and distances are presented below along with the poses at which they have been taken (Fig. 8, Fig. 10, and Fig. 12).

D.1. Software and setup

All images has been acquired using the free software MultiCam Studio (v6.15.1.3573) of the company Euresys. The shutter speed has been set to 5 ms. While taking white images for the pre-calibration step, the gain has been set to its maximum value. For Raytrix data, we use their proprietary software RxLive (v4.0.50.2) to calibrate the camera, and compute the depth maps used in the evaluation.

D.2. Dataset R12-A

The dataset has been taken at short focus distance, $h = 450$ mm. Therefore, the checkerboard squares size had to be decreased to 10 mm so we can observe the corner in image space. All the poses have been acquired at distances between 400 and 175 mm from the checkerboard.

Controlled evaluation. The dataset is composed of 11 poses taken with a relative step of 10 mm between each pose along the z -axis direction, at distances between 385 and 265 mm.

D.3. Dataset R12-B

The dataset has been taken at middle focus distance, $h = 1000$ mm. Therefore, the checkerboard squares size is set to 20 mm so we can observe the corner in image space. All the poses have been acquired at distances between 775 and 400 mm from the checkerboard.

Controlled evaluation. The dataset is composed of 10 poses taken with a relative step of 50 mm between each pose along the z -axis direction, at distances between 900 and 450 mm.

D.4. Dataset R12-C

The dataset has been taken at long focus distance, $h = \infty$. Therefore, the checkerboard squares size had to be increased to 30mm so we can observe the corner in image space. All the poses have been acquired at distances between 2500 and 500 mm from the checkerboard.

Controlled evaluation. The dataset is composed of 18 poses taken with a relative step of 50 mm between each pose along the z -axis direction, at distances between 1250 and 400 mm.

R12-A ($h = 450\text{mm}$)				R12-B ($h = 1000\text{mm}$)				R12-C ($h = \infty$)		
	Unit	Initial	Ours	[20]	Initial	Ours	[20]	Initial	Ours	[20]
F	[mm]	50	49.720	54.888	50	50.047	51.262	50	50.011	53.322
A_0	[$\times 10^{-5}$]	0	23.145	6.099	0	1.686	0.023	0	18.647	1.393
$-A_1$	[$\times 10^{-6}$]	0	2.934	0.925	0	0.177	0.093	0	2.652	0.382
A_2	[$\times 10^{-8}$]	0	1.078	0.303	0	0.023	0.007	0	1.036	0.104
B_0	[$\times 10^{-5}$]	0	-9.870	-15.028	0	14.373	12.143	0	18.971	27.720
$-B_1$	[$\times 10^{-5}$]	0	5.387	5.020	0	18.947	18.164	0	7.600	4.461
D	[mm]	56.658	56.696	62.425	52.113	52.125	53.296	49.384	49.384	52.379
$-t_x$	[mm]	11.293	11.129	9.771	11.299	12.503	12.672	11.302	13.213	14.159
$-t_y$	[mm]	8.411	8.186	8.334	8.416	6.305	6.114	8.418	7.256	6.231
$-\theta_x$	[μrad]	0	875.200	468.600	0	648.400	576.200	0	575.100	487.700
θ_y	[μrad]	0	669.700	321.800	0	417.400	350.400	0	445.600	366.100
θ_z	[μrad]	0.553	30.100	25.300	16.989	36.600	35.300	27.961	44.500	49.100
ΔC	[μm]	127.51	127.45	127.38	127.47	127.45	127.40	127.54	127.50	127.42
$f^{(1)}$	[μm]	578.15	577.97	-	581.10	580.48	-	554.35	556.09	-
$f^{(2)}$	[μm]	504.46	505.21	-	503.96	504.33	-	475.98	479.03	-
$f^{(3)}$	[μm]	551.67	551.79	-	546.39	546.37	-	518.98	521.33	-
u_0	[pix]	2039	2042.55	2289.83	2039	1790.94	1759.29	2039	1661.95	1487.2
v_0	[pix]	1533	1556.29	1528.24	1533	1900.19	1934.87	1533	1726.91	1913.81
d	[μm]	318.63	325.24	402.32	336.84	336.26	363.17	307.93	312.62	367.40

Table 6: Initial intrinsic parameters for each dataset along with the optimized parameters obtained by our method and with the method of [20].

E. Source code and datasets redistribution

The datasets and the source code are publicly available to the community. Datasets can be downloaded from <https://github.com/comsee-research/plenoptic-datasets>.

We also developed an open-source C++ library for plenoptic camera named libpleno which is available at <https://github.com/comsee-research/libpleno>.

Along with this library, we developed a set of tools to pre-calibrate and calibrate a multifocus plenoptic camera, named COMPOTE (standing for Calibration Of Multi-focus PlenOpTic camEra) which is available at <https://github.com/comsee-research/compote>.

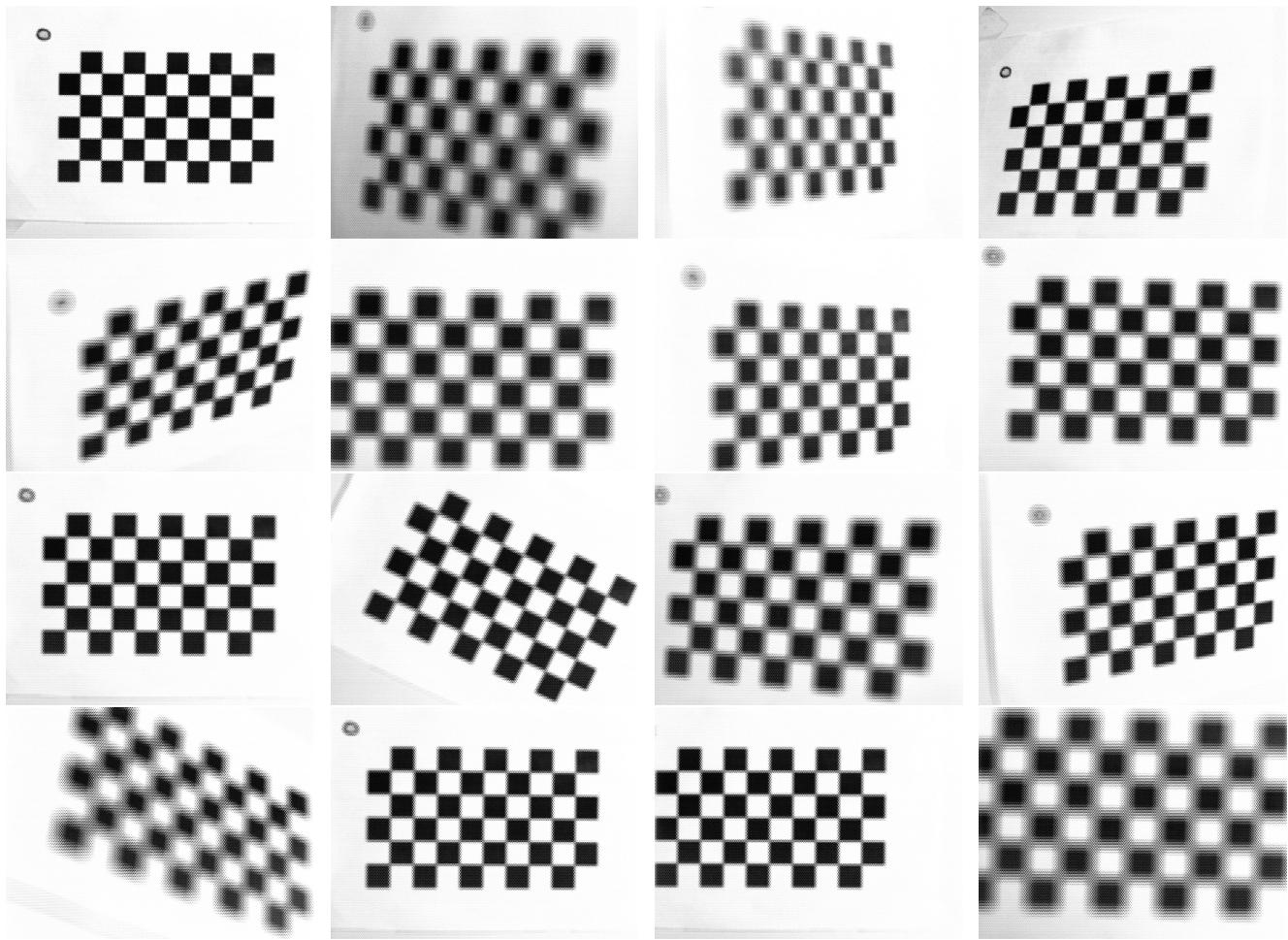


Figure 7: Devignetted images of the calibration targets (9×5 of 10mm side checkerboard) from the dataset R12-A taken at various angles and distances.

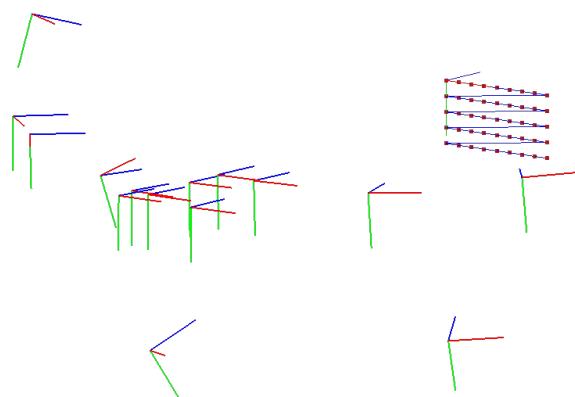


Figure 8: Poses of the camera while capturing the calibration targets from dataset R12-A.

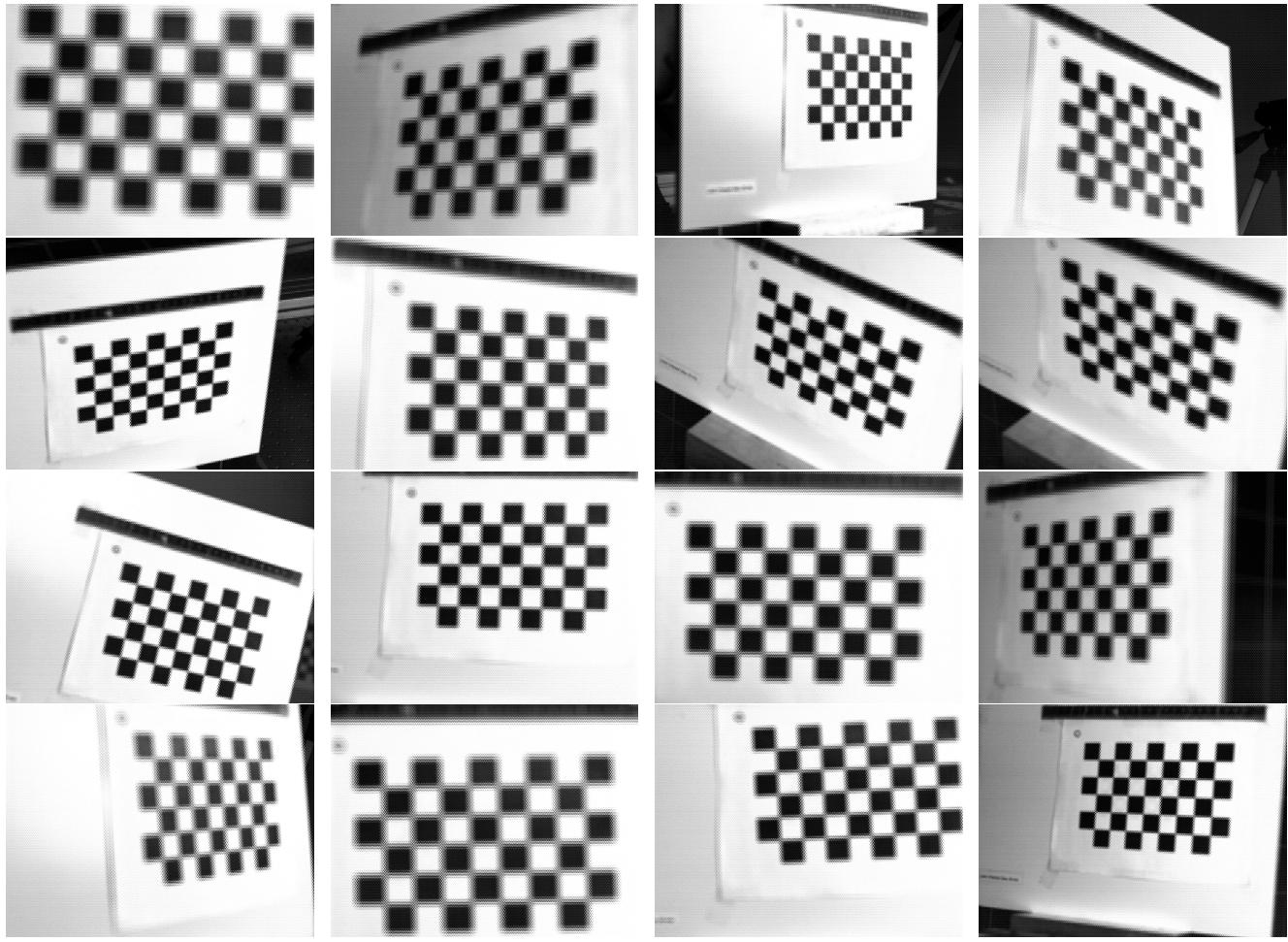


Figure 9: Devignetted images of the calibration targets (8×5 of 20mm side checkerboard) from the dataset R12-B taken at various angles and distances.

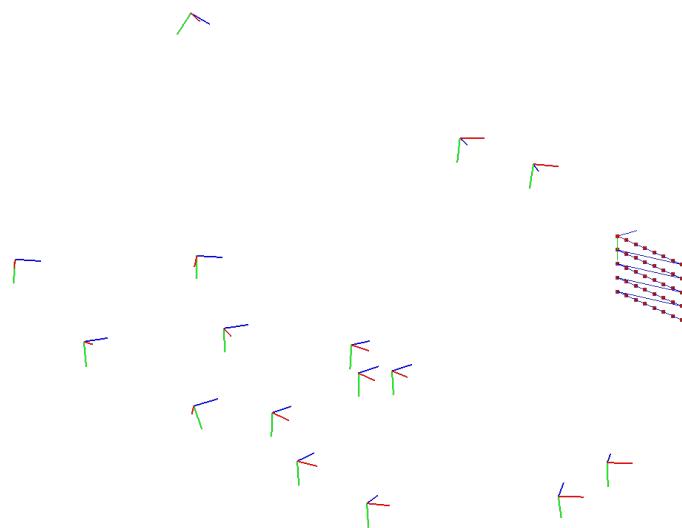


Figure 10: Poses of the camera while capturing the calibration targets from dataset R12-B.

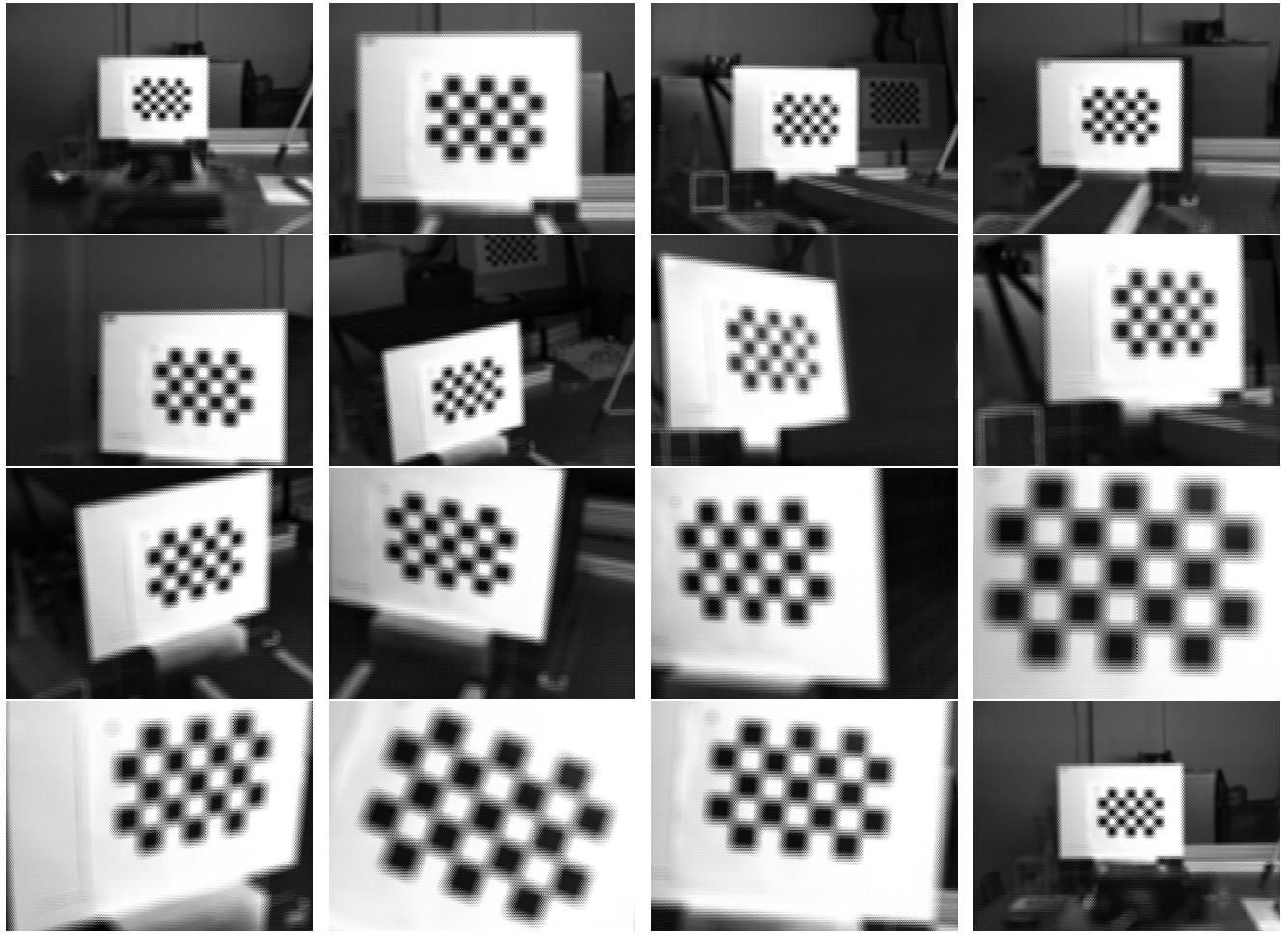


Figure 11: Devignetted images of the calibration targets (7×5 of 30mm side checkerboard) from the dataset R12-C taken at various angles and distances.

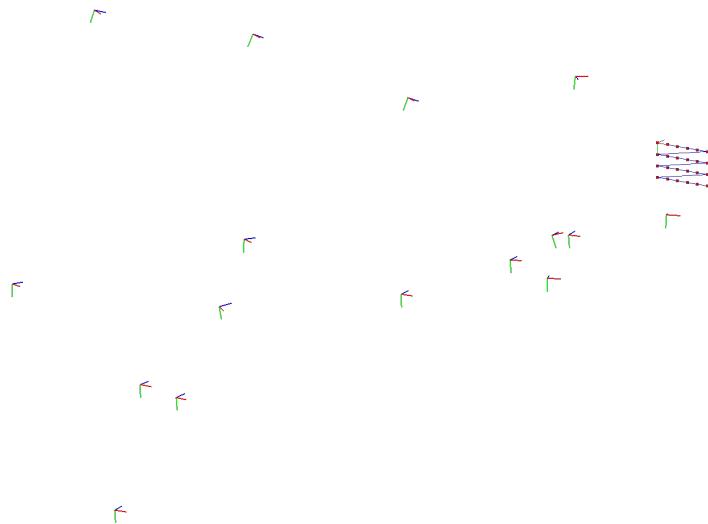


Figure 12: Poses of the camera while capturing the calibration targets from dataset R12-C.