Applying a Traditional Calibration Method to a Focused Plenoptic Camera

Niclas Zeller, Franz Quint

Faculty of Electrical Engineering and Information Technology
Karlsruhe University of Applied Sciences
76133 Karlsruhe, Germany
Email: niclas.zeller@hs-karlsruhe.de,
franz.quint@hs-karlsruhe.de

Uwe Stilla

Department of Photogrammetry and Remote Sensing
Technische Universität München
80290 Munich, Germany
Email: stilla@tum.de

Abstract—This article presents a method to calibrate the optical imaging process of a focused plenoptic camera. At first, the concept of a focused plenoptic camera is presented, where the synthesis of images from the recorded light-field is described. It is shown that the synthesized image imitates the image of a real camera and thus traditional methods can be used for calibration. In this article a method approved for traditional cameras is applied to the recordings of a light-field camera. Based on experiments, the quality of the calibration method is evaluated. Amongst others it is shown that a pinhole camera with a conventional lens distortion model also holds true for a focused plenoptic camera.

I. Introduction

A plenoptic camera or light-field camera is an optical camera which records, different from a traditional camera, not only the intensity of incident light on the image sensor but an image of the light-field of a scene.

During the last years plenoptic cameras gained more and more attention in the research fields of computer vision and photogrammetry. The main reason for revisiting this concept, which was already investigated more than hundred years ago [1], [2], is today's availability of fast graphic processor units (GPUs). Today's GPUs are capable to evaluate recorded light-field sequences with high frame rates ($\geq 30\,\text{fps}$). In [3] and [4] the first developed prototypes of plenoptic cameras are described.

Today, there exist different concepts of light-field cameras. While [4], [5] and [6] use a micro lens array (MLA) in front of the image sensor, the concept described in [7] relies on a 4×4 micro camera array. There are basically two different MLA-based concepts of a plenoptic camera. The "unfocsued" plenoptic camera as described in [4] and the focused plenoptic camera (or plenoptic camera 2.0) [8], [6]. Compared to the "unfocused" plenoptic camera the focused plenoptic camera has a higher spatial resolution. This results in a higher resolution of the synthesized image. In contrast, the "unfocused" plenoptic camera has a higher angular resolution.

For photogrammetric applications it is important to accurately know the relationship between a point in object space and the corresponding image point. To define this relationship, the intrinsic parameters of the camera have to be known precisely and have to be determined by a camera calibration process. In the last years some calibration methods for plenoptic cameras were proposed already. While [9] describes the calibration of a Lytro camera [4], [10] and [11] present calibration methods for a Raytrix camera [6]. Paper [11] mainly focuses on the calibration of the supplied depth information up to a range of about $10\,\text{m}$ and investigates the depth resolution in this range, whereas [10] presents depth and image calibration of a Raytrix camera in a very short range. Therefore a more complex calibration setup is required.

In this article we apply a very simple calibration method [12], which is commonly used for traditional cameras, to the recordings of a Raytrix camera. Thus, we want to investigate the suitability of such traditional methods for plenoptic cameras.

This article is organized as follows. Section II briefly presents the concept of a focused plenoptic camera. In Section III we describe the calibration method [12] which was used in the experiment presented in Section IV. Section V illustrates the results of the experiments and Section VI draws conclusion.

II. Concept of a Focused Plenoptic Camera

A plenoptic camera records the light-field of a scene as a four dimensional (4D) function, by one single shot. Thus, a plenoptic camera gathers much more information about a scene than a traditional camera. In [13] Gortler et al. show that in free space it is sufficient to define the light-field as a 4D function since here the intensity along a ray does not change. Thus, the constant intensity along a ray can be defined by two position and two angle coordinates. Based on the 4D representation, the light-field of a convex object emitted in one direction can be described.

Since this article discusses calibration methods applied to a Raytrix camera, only the concept of this camera will be presented here. From a schematic point of view the only difference between a traditional camera and a focused plenoptic camera is the MLA in front of the image sensor. Thus, we will derive the concept of the focused plenoptic camera from a common thin lens projection.

Figure 1 shows the projection of an object which is in the distance a_L in front of the main lens to a focused image in the distance b_L behind the main lens. Here, the relationship
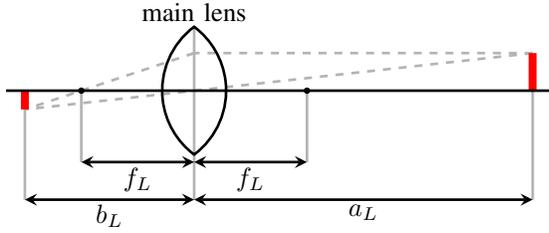
Fig. 1. Thin lens projection. A thin lens projects an object in distance a_L in front of the lens to a focused image in distance b_L behind the lens. The relationship between a_L and b_L dependents on the focal length f_L of the lens and is defined by the thin lens equation.
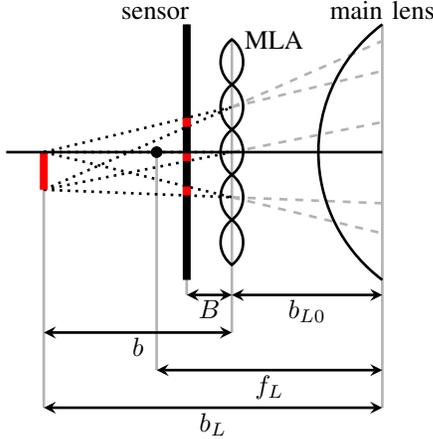

Fig. 2. Image projection inside a focused plenoptic camera. The virtual image, which would be formed in distance b_L behind the main lens, is focused on the image sensor by several micro lenses, which are placed in the distance b_{L0} behind the main lens. Out of the micro images of a point the distance b between MLA and the corresponding virtual main lens image can be estimated.

between the object distance a_L and the image distance b_L is defined by the thin lens equation given in eq. (1). Here f_L represents the focal length of the thin lens.

$$\frac{1}{f_L} = \frac{1}{a_L} + \frac{1}{b_L} \tag{1}$$

For a traditional camera the sensor would be placed in the image distance b_L behind the sensor and thus all object within the depth of field (DOF) around the object distance a_L occur focused in the recorded image.

In a Raytrix camera the sensor is placed closer than the image distance b_L to the sensor. Furthermore, a MLA is placed in distance B ahead of the sensor. Instead of placing the sensor in front of the image plane, a focused plenoptic camera can also be realized by placing the sensor behind the image plane as described in [5]. Figure 2 shows a schematic cross view of the interior of a Raytrix camera. The micro lenses of the MLA focus the virtual main lens image, which would occur behind the sensor, on the sensor. Thus, each micro lens forms a micro image on the sensor. One distinct feature of Raytrix cameras is that they have MLAs which consist of micro lenses with three different focal lengths. Each type of micro lenses focuses a different image distance b_L on the sensor. Thus, the DOF of the synthesized image is increased by a factor of three.

Within its DOF each micro lens can be considered as a pinhole. Thus, each pixel of a micro image represents one

light ray (central ray of the corresponding micro lens). Since the optical center of each micro lens is known, for each ray additionally to the position coordinates (pixel position (x_I, y_I)) two angle coordinates (α_x and α_y) can be calculated.

A. Image Synthesis

One feature of a plenoptic camera is, that after capturing a light-field, images for different focus distances can be synthesized. This is done by calculating the image which would be formed on a sensor placed in a certain image distance b_L by the recorded light-field rays. To calculate the intensity of a pixel in the synthesized image, at first all rays of the 4D light-field which intersect the synthesized image plane at the corresponding position are searched. Out of the intensities of the selected rays a weighted average value is calculated. This average value represents the intensity of the synthesized pixel.

Because of the low angular resolution of a focused plenoptic camera there will occur artifacts in the synthesized image for regions which are not in focus on the selected image plane. Nevertheless, it is also possible to focus each pixel on a different image plane. Thus, if the image distance for each pixel is known, a totally focused image can be synthesized.

For an image point with a long image distance b_L (short object distance a_L) more rays are sampled than for an image point with a short image distance because it occurs focused in more micro images. Thus, in the synthesized totally focused image close objects have less spatial resolution than objects which are further away from the camera.

Here we will not go further into details of the image synthesis. For a detailed mathematical description we refer to [6].

B. Depth Estimation

It was already mentioned that if the image distance for each virtual image point is known, a totally focused image can be calculated. If a virtual image point is projected to at least two micro images, its distance to the MLA b can be estimated by triangulation of the corresponding rays. Thus, the depth map needed to synthesize a totally focused image also can be estimated from the recorded light-field function. Since the distance between MLA and sensor B usually is not known, a standardized value of the distance b, called virtual depth $v = \frac{b}{B}$, is estimated. For further descriptions on the depth estimation for plenoptic cameras we refer to the following papers: [6], [14], [15].

III. IMAGE CALIBRATION METHOD

Section II presented the concept of a focused plenoptic camera. Here it was shown that one can synthesize the image of a traditional camera from the recorded light field of a focused plenoptic camera.

In photogrammetry as well as in some fields of computer vision it is important to have a precisely defined relation between a pixel in the image and the corresponding point in the object space. Since the synthesized image imitates the recording of a traditional camera it is obvious to apply traditional camera calibration methods to the synthesized image.

An approved method for calibration of traditional cameras is the method described by Zhang [12]. In [12] the imaging process of the camera is simplified by using a pinhole camera model, like it is done in most of the common calibration methods. For our experiments we define the following intrinsic parameters for the pinhole camera model:

- f_x - focal length of the pinhole camera in x-direction (in pixels)

- f_y - focal length of the pinhole camera in y-direction (in pixels)

- (c_x, c_y) - image coordinates of the camera's principal point (in pixels)

By the definition of different focal lengths in x- and y-direction we consider the case that the pixels on the image sensor are not square but rectangular.

For the following definitions we use x_I as notation for an image point in Cartesian coordinates and $\tilde{x}_I$ for the corresponding homogeneous coordinates. x_C defines a three dimensional (3D) point in camera coordinates and x_W in world coordinates.

$$x_I = \begin{pmatrix} x_I & y_I \end{pmatrix}^T \tag{2}$$

$$\tilde{x}_I = \begin{pmatrix} k \cdot x_I & k \cdot y_I & k \end{pmatrix}^T = k \cdot \begin{pmatrix} x_I^T & 1 \end{pmatrix}^T \tag{3}$$

$$x_C = \begin{pmatrix} x_C & y_C & z_C \end{pmatrix}^T \tag{4}$$

$$x_W = \begin{pmatrix} x_W & y_W & z_W \end{pmatrix}^T \tag{5}$$

Based on the intrinsic parameters, the intrinsic matrix M can be defined, which describes the transformation from 3D camera coordinates x_C to image coordinates x_I, as given in eq. (6).

$$\tilde{x}_I = M \cdot x_C = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot x_C \tag{6}$$

Since real optical lenses are never perfect, they add distortion to the projected image. In the presented calibration method we use the distortion model presented by Brown [16], as it is implemented in the OpenCV calibration method [17]. Eq. (7) to (10) define the distortion model, where x_I and y_I are the undistorted and x'_I and y'_I ($x'_I = \begin{pmatrix} x'_I & y'_I \end{pmatrix}^T$) the distorted image coordinates in pixels. This distortion model considers radial as well as tangential distortion.

$$x_I^* = \frac{x_I - c_x}{f_x} \quad \text{and} \quad y_I^* = \frac{y_I - c_y}{f_y} \tag{7}$$

$$r = \sqrt{(x_I^*)^2 + (y_I^*)^2} \tag{8}$$

$$x'_I = \left[x_I^* \cdot \left(1 + k_0 \cdot r^2 + k_1 \cdot r^4 + k_2 \cdot r^6 \right) + \right.$$
$$\left. 2 \cdot p_0 \cdot x_I^* \cdot y_I^* + p_1 \cdot \left(r^2 + 2 \cdot (x_I^*)^2 \right) \right] \cdot f_x + c_x \tag{9}$$

$$y'_I = \left[y_I^* \cdot \left(1 + k_0 \cdot r^2 + k_1 \cdot r^4 + k_2 \cdot r^6 \right) + \right.$$
$$\left. p_0 \cdot \left(r^2 + 2 \cdot (y_I^*)^2 \right) + 2 \cdot p_1 \cdot x_I^* \cdot y_I^* \right] \cdot f_y + c_y \tag{10}$$

Radial distortion is a radial symmetric distortion component with its origin in the principal point (c_x, c_y). Thus, it can be defined by a function of the distance r to the principal point. In Brown's distortion model radial distortion is defined by a polynomial of r, where the coefficients of odd exponents are

zero. The nonzero coefficients are k_0, k_1 and k_2, as given in eq. (9) and (10). Tangential distortion is a radial asymmetric distortion which comes from decentralized lenses within the lens system. Brown defines the tangential distortion by the coefficients p_0 and p_1 as given in eq. (9) and (10).

The transform from world coordinates x_W to camera coordinates x_C can be defined by a 3D rigid transform which combines a 3D rotation and translation. The rigid transform is defined in eq. (11), where R is the rotation matrix and t the translation vector.

$$x_C = (R \quad t) \cdot \begin{pmatrix} x_W \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ z_W \\ 1 \end{pmatrix} \tag{11}$$

The presented calibration method considers only planar objects. The world coordinate system for a recorded object is defined such that the x-y-plane is equivalent to the plane of the object. Thus, the z-component of the world coordinate system is always zero ($z_W = 0$) and the transform becomes independent of the third colum of R, as given in eq. (12).

$$x_C = \begin{pmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix} \tag{12}$$

If we now combine eq. (6) and (12), the transform from a point on the planar object to a point on the image plane is defined as given in eq. (13).

$$\begin{pmatrix} k \cdot x_I \\ k \cdot y_I \\ k \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix} \tag{13}$$

After dividing eq. (13) by h_{33} eq. (14) results.

$$\begin{pmatrix} k^* \cdot x_I \\ k^* \cdot y_I \\ k^* \end{pmatrix} = \begin{pmatrix} h_{11}^* & h_{12}^* & h_{13}^* \\ h_{21}^* & h_{22}^* & h_{23}^* \\ h_{31}^* & h_{32}^* & 1 \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix}$$
$$= H \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix} \tag{14}$$

The planar homography between world and image coordinates, as given in eq. (14), is defined by eight linear independent coefficients h_{ij}^*. For each perspective from which the planar object is recorded the rotation matrix R and the translation vector t change. Thus each perspective results in a new transformation matrix H. The rotation matrix R is defined by three independent angles (α, β and γ). Besides, the translation vector t relies on three independent coefficients (t_x, t_y and t_z). Since the estimated planar homography, defined by H, gives us eight linear independent conditions, the six coefficients of the rigid transform between world and camera coordinates can be calculated. Thus, for each perspective two conditions are left to estimate the intrinsic parameters. The intrinsic matrix mostly is defined by four independent coefficients as given in eq. (6). Hence, the planar object has to be recorded from at

least two perspective to receive a unique solution. In reality recordings from much more than two perspectives are taken to average measurement errors.

After the estimation of the intrinsic and extrinsic parameters the distortion coefficients are estimated. This is done based on the calculated undistorted projection of an object point x_W to an image point x_I, as defined in eq. (13), and the corresponding recorded and distorted image point x_I'. Thus, the distortion coefficients (k_0, k_1, k_2, p_0 and p_1) can be estimated from eq. (7) to (10) by linear regression.

Based on the calculated undistorted points the intrinsic and extrinsic coefficients are updated and the distortion coefficients are calculated again. This procedure is repeated until consistency is reached.

For a more detailed description on how the intrinsic, extrinsic and distortion coefficients are estimated we refer to [12] and [18].

IV. EXPERIMENTS

This section presents experiments which were performed to evaluate the calibration method described in Section III when applying it to a focused plenoptic camera. In the presented experiments the OpenCV [17] implementation of the calibration method was used. Here calibration points are recorded by using a planar chessboard pattern. Each corner point between four adjacent chessboard fields is detected in the recorded image. Since for all those points the corresponding world coordinates on the pattern are known, the points can be used for calibration. For the experiment a pattern with 10×7 fields was used.

In the presented experiments we used a Raytrix R5 camera with the following two different lenses mounted to it:

1) 4 mm–12 mm zoom lens, set to 12 mm focal length
2) 35 mm fixed focal length

The main goal of the performed experiments was to distinguish if the synthesized imaging of a focused plenoptic camera can also be defined by a traditional camera model. Besides, we wanted to evaluate how the two different lenses, with different focal lengths and levels of distortion, are effecting the calibration results.

For the experiments, for both lenses three measurement series were recorded. In each series the chessboard pattern was recorded from 50 as different as possible perspectives. To each series the calibration method presented in Section III was applied multiple times. For each calibration the camera model was slightly changed as will be described in Section IV-A to IV-D. The calibration method was performed to each of the three measurement series to evaluate the consistency of the results.

A. Complete Calibration Model

In the first experiment for both lenses the calibration was performed using the complete calibration model with four intrinsic and five distortion coefficients.

B. Constant Aspect Ration ($f_x = f_y$)

For the second experiment the aspect ration was set to one ($f_x / f_y = 1$). Thus, during the calibration one intrinsic parameter less has to be estimated. If this assumption conforms the real camera model a more consistent result can be expected.

C. Constant Aspect Ration and Fixed Principal Point

One further assumption is made in the third experiment. Beside the constant aspect ration the principal point of the camera model is considered to be in the image center and is set constantly to that point. This reduces again the number of coefficients to be estimated. Furthermore, the principal point certainly lies somewhere around the image center.

D. Constant Aspect Ration, Fixed Principal Point and Only One Distortion Coefficient

The fourth experiment is only performed to the 35 mm lens since for the 12 mm lens no improvement is expected. Here the assumptions from the second and third experiment are still considered to hold true. Besides, the distortion model is defined by only the first radial distortion coefficient k_0. All other distortion coefficients are set to zero ($k_1 = k_2 = p_0 = p_1 = 0$). This experiment is performed since the distortion for the 35 mm lens seems to be very weak. Thus, by reducing again the degrees of freedom the remaining coefficients should be estimated more consistent if the assumption holds true.

V. RESULTS

In this section the results for the performed experiments are presented. For evaluation at first some statistics are defined.

The root mean square (RMS) of the reprojection error σ_{Rep} gives a measure how good the estimated model represents the measured values. The reprojection error $e_{Rep}^{(i,j)}$ is the distance between an object point $x_W^{(i)}$ projected on the image plane, based on the projection model, and the corresponding recorded image point $x_I'^{(i,j)}$. Eq. (15) gives the definition of σ_{Rep}.

$$\sigma_{Rep} = \sqrt{\frac{1}{N \cdot M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left(e_{Rep}^{(i,j)} \right)^2} \qquad (15)$$

Here, $e_{Rep}^{(i,j)}$ represents the reprojection error of the i-th chessboard corner in the j-th image. N is the number of chessboard corners and M the number of recorded images.

Besides, we define the root mean square error (RMSE) of the estimated focal length σ_f and the principal point σ_c, as given in eq. (16) and (17).

$$\sigma_f = \sqrt{\frac{1}{2 \cdot S} \sum_{i=0}^{S-1} \left(f_x^{(i)} - \overline{f}_x \right)^2 + \left(f_y^{(i)} - \overline{f}_y \right)^2} \qquad (16)$$

$$\sigma_c = \sqrt{\frac{1}{S} \sum_{i=0}^{S-1} \left(c^{(i)} - \overline{c} \right) \cdot \left(c^{(i)} - \overline{c} \right)^T} \qquad (17)$$

In eq. (16) and (17) S represents the number of measurement series which were recorded ($S = 3$). Of course, from a series of three no real statistics can be calculated. Nevertheless, based

TABLE I.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT A USING A 12 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	1058.8	1053.8	1048.1
f_y [Pixel]	1064.3	1060.9	1053.9
c_x [Pixel]	494.0	488.4	475.8
c_y [Pixel]	496.9	528.4	502.4

TABLE II.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT B USING A 12 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	1065.9	1068.7	1059.7
f_y [Pixel]	1065.9	1068.7	1059.7
c_x [Pixel]	495.3	492.6	471.0
c_y [Pixel]	494.2	525.6	500.7

TABLE III.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT C USING A 12 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	1064.1	1068.6	1058.8
f_y [Pixel]	1064.1	1068.6	1058.8
c_x [Pixel]	511.5	511.5	511.5
c_y [Pixel]	511.5	511.5	511.5

TABLE IV.   CALCULATED STATISTICS FOR THE 12 mm FOCAL LENGTH

experiment no.	1	2	3
σ_{Rep} [Pixel]	0.804	0.807	0.808
σ_f [Pixel]	4.333	3.758	4.032
σ_c [Pixel]	15.727	17.368	–

TABLE V.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT A USING A 35 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	3259.1	3261.6	3284.0
f_y [Pixel]	3262.3	3261.0	3291.0
c_x [Pixel]	520.7	430.0	426.4
c_y [Pixel]	346.3	389.5	323.7

TABLE VI.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT B USING A 35 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	3262.1	3261.1	3285.0
f_y [Pixel]	3262.1	3261.1	3285.0
c_x [Pixel]	521.1	430.0	423.5
c_y [Pixel]	347.0	389.4	327.0

TABLE VII.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT C USING A 35 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	3257.5	3247.8	3250.5
f_y [Pixel]	3257.5	3247.8	3250.5
c_x [Pixel]	511.5	511.5	511.5
c_y [Pixel]	511.5	511.5	511.5

TABLE VIII.   ESTIMATED INTRINSIC PARAMETERS FOR THE
EXPERIMENT D USING A 35 mm FOCAL LENGTH

series no.	1	2	3
f_x [Pixel]	3255.7	3252.7	3275.4
f_y [Pixel]	3255.7	3252.7	3275.4
c_x [Pixel]	511.5	511.5	511.5
c_y [Pixel]	511.5	511.5	511.5

TABLE IX.   CALCULATED STATISTICS FOR THE 35 mm FOCAL LENGTH

experiment no.	1	2	3	4
σ_{Rep} [Pixel]	0.336	0.336	0.352	0.358
σ_f [Pixel]	12.583	11.006	9.709	10.053
σ_c [Pixel]	51.461	51.609	–	–

on σ_f and σ_c the quality of different calibration results can be compared.

In the following two subsections the calibration results for the 12 mm and 35 mm lens will be presented separately.

A. Calibration Results for $f_L = 12$ mm

Table I to III present the estimated intrinsic parameters which resulted from the three experiments performed for the 12 mm lens. Besides, Table IV shows the corresponding calculated statistics. Since for the third experiment the principal point was set to the image center prior to the calibration, σ_c is not meaningful.

If we compare σ_{Rep}, one can see that the models of all three experiments conform quite well to the measured data. For the model of the first experiment which uses all degrees of freedom, the RMS of the reprojection errors of course is minimum. This case uses the most complex model and thus, the measured data can be adapted best.

Nevertheless, fixing the aspect ratio to $f_x/f_y = 1$ or setting the principal point to a certain image coordinate improves the estimation of the other parameters, at least as long as the assumption conforms more or less to the real model.

For the third experiment σ_f is worse than for the second experiment. This indicates, that the selected principal point differs from the real one. By adjusting the principal point it should be possible to achieve a more consistent estimation of the focal length.

B. Calibration Results for $f_L = 35$ mm

Table V to VIII shows the intrinsic parameters which were estimated for the 35 mm lens and Table IX gives the corresponding calculated statistics. For the 35 mm lens the same behavior as for the 12 mm lens can be observed. Here, the most consistent estimation of the focal length was achieved when fixing the aspect ratio as well as the principal point. One reason therefore could be that for this lens the principal point conforms quite well to the image center. Another reason is that for long focal lengths a shift of the principal point has not as much effect as for short focal lengths. This can also be seen when comparing σ_c for the 35 mm and the 12 mm lens. For the 35 mm lens σ_c is much higher than for the 12 mm lens. Besides, one can see that the results for experiment C and D are almost the same. This means, for the 35 mm lens the distortion model can be reduced to only one parameter with only a small rise of σ_{Rep} and σ_f.

VI. CONCLUSION

In conclusion it can be said that traditional camera calibration methods, like the method of Zhang [12] can be used

to calibrate the synthesized images of a plenoptic camera. This can be seen by the small reprojection errors for all experiments. Nevertheless, for a focal length of the main lens which is long with respect to the image size (small field of view (FOV)) Zhang's calibration method seems to be inappropriate. Especially for long focal lengths, the intrinsic parameters are strongly correlated to the extrinsic orientation. Thus, errors in the extrinsic orientation will also affect the intrinsic parameters. One way to improve the calibration would be to used a 3D calibration object instead of a planer object. A 3D object brings more perspective distortion to the recorded image. Thus, extrinsic and intrinsic parameters can be separated more accurate. Another problem is that Zhang's method does not really minimize the squared error between recorded and calculated image points like it is done for a bundle adjustment. Instead is uses several optimization steps. Here it could be investigated, how the errors of estimation are propagated form one step to the next, to see how the error of any parameter effects the reprojection error.

References

[1] F. E. Ives, "Parallax stereogram and process of making same," USA Patent US725 567, 04 14, 1903.

[2] G. Lippmann, "Epreuves reversibles. photographies integrales," *Comptes Rendus De l'Academie Des Sciences De Paris*, vol. 146, pp. 446–451, 1908.

[3] E. H. Adelson and J. Y. A. Wang, "Single lens stereo with a plenoptic camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 99–106, February 1992.

[4] R. Ng, "Digital light field photography," Ph.D. dissertation, Stanford University, Stanford, USA, July 2006.

[5] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *IEEE International Conference on Computational Photography (ICCP)*, San Francisco, CA, April 2009, pp. 1–8.

[6] C. Perwaß and L. Wietzke, "Single lens 3D-camera with extended depth-of-field," in *Human Vision and Electronic Imaging XVII*, Burlingame, California, USA, January 2012.

[7] K. Venkataraman, D. Lelescu, J. Duparre, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar, "Picam: An ultra-thin high performance monolithic camera array," *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2013*, vol. 32, no. 6, pp. 1–13, 11 2013.

[8] A. Lunsdaine and T. Georgiev, "Full resolution lightfield rendering," Adobe Systems, Inc., Tech. Rep., 2008.

[9] D. Dansereau, O. Pizarro, and S. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 1027–1034.

[10] O. Johannsen, C. Heinze, B. Goldlücke, and C. Perwaß, "On the calibration of focused plenoptic cameras," in *GCPR Workshop on Imaging New Modalities*, 2013.

[11] N. Zeller, F. Quint, and U. Stilla, "Kalibrierung und Genauigkeituntersuchung einer fokussierten plenoptischen Kamera," in *34. Wissenschaftlich-Technische Jahrestagung der DGPF (DGPF Tagungsband 23 / 2014)*, vol. 23, 3 2014.

[12] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 1, 1999, pp. 666–673.

[13] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proceedings of the 23rd annual conference on computer graphics and interactive techniques, SIGGRAPH*. New York, NY, USA: ACM, 1996, pp. 43–54.

[14] N. Zeller, F. Quint, and U. Stilla, "Calibration and accuracy analysis of a focused plenoptic camera," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3, pp. 205–212, 09 2014.

[15] S. Wanner and B. Goldlücke, "Globally consistent depth labeling of 4d lightfields," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[16] D. C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, no. 8, pp. 855–866, 1971.

[17] G. Bradski, "The opencv library," *Dr. Dobb's Journal of Software Tools*, 2000.

[18] G. Bradski and A. Kaehler, *Learning OpenCV - Computer Vision in C++ with the OpenCV Library*. O'Reilly Media, 09 2008.