

CALIBRATION AND ACCURACY ANALYSIS OF A FOCUSED PLENOPTIC CAMERA

N. Zeller^{a, c, *}, F. Quint^b, U. Stilla^c

^a Institute of Applied Research (IAF), Karlsruhe University of Applied Sciences,
76133 Karlsruhe, Germany - niclas.zeller@hs-karlsruhe.de

^b Faculty of Electrical Engineering and Information Technology, Karlsruhe University of Applied Sciences,
76133 Karlsruhe, Germany - franz.quint@hs-karlsruhe.de

^c Department of Photogrammetry and Remote Sensing, Technische Universitaet Muenchen, 80290 Munich, Germany - stilla@tum.de

Commission III

KEY WORDS: Accuracy, calibration, depth map, imaging model, plenoptic camera, Raytrix, virtual depth

ABSTRACT:

In this article we introduce new methods for the calibration of depth images from focused plenoptic cameras and validate the results. We start with a brief description of the concept of a focused plenoptic camera and how from the recorded raw image a depth map can be estimated. For this camera, an analytical expression of the depth accuracy is derived for the first time. In the main part of the paper, methods to calibrate a focused plenoptic camera are developed and evaluated. The optical imaging process is calibrated by using a method which is already known from the calibration of traditional cameras. For the calibration of the depth map two new model based methods, which make use of the projection concept of the camera are developed. These new methods are compared to a common curve fitting approach, which is based on Taylor-series-approximation. Both model based methods show significant advantages compared to the curve fitting method. They need less reference points for calibration than the curve fitting method and moreover, supply a function which is valid in excess of the range of calibration. In addition the depth map accuracy of the plenoptic camera was experimentally investigated for different focal lengths of the main lens and is compared to the analytical evaluation.

1. INTRODUCTION

The concept of a plenoptic camera already has been developed more than hundred years ago (Ives, 1903, Lippmann, 1908). Nevertheless, only for the last few years the existing graphic processor units (GPUs) are capable to evaluate the recordings of a plenoptic camera with acceptable frame rates (≥ 30 fps).

Today, there exist basically two concepts of a plenoptic camera which use a micro lens array (MLA) in front of the sensor. Those two concepts are the "unfocused" plenoptic camera developed by Ng (Ng, 2006) and the focused plenoptic camera, which was described for the first time by Lunsdaine and Georgiev (Lunsdaine and Georgiev, 2008). The focused plenoptic camera can also be found as plenoptic camera 2.0 in the literature. One big advantage of it compared to the "unfocused" plenoptic camera is the high resolution of the synthesized image. This is beneficial for estimating a depth map out of the recorded raw image (Perwass and Wietzke, 2012).

A camera system which is supposed to be use for a photogrammetric purpose has to be calibrated to transform the recorded images to the metric object space. While in (Dansereau et al., 2013) the calibration of a Lytro camera (Ng, 2006) is described, (Johannsen et al., 2013) presents the calibration of a Raytrix camera (Perwass and Wietzke, 2012) for object distances up to about 50 cm. The methods presented here were also developed to calibrate a Raytrix camera. Nevertheless, we are focused on developing calibration method for farther object distances then the method described in (Johannsen et al., 2013).

This paper presents an analytical analysis of the depth accuracy which cannot be found in the literature until now. Afterwards, the developed calibration methods are presented. The presented

methods are separated into the calibration of the optical imaging process while disregarding the depth information and into the calibration of the depth map supplied by the camera. For the calibration of the depth map we present two newly developed approaches and compare them to an already known method. All three methods are evaluated for different camera setups in a range of approx. 0.7 m to 5 m.

Different from (Johannsen et al., 2013) we did not investigate the distortion of the depth map by the main lens. However, for the focal lengths used in this article and for the large object distances we are operating with, the depth map distortion can be neglected compared to the stochastic noise of the depth information.

All experiments presented in this paper were performed using the camera Raytrix R5. In these experiments the impact of different focal lengths of the main lens on the depth information was analyzed and the different calibration methods were compared to each other.

This article is organized as follows. Section 2 presents the concept of a focused plenoptic camera. Here, we also derive the analytical expression for the accuracy of the depth map. In Section 3 the calibration of the image projection as well as the three depth calibration methods are presented. Section 4 presents experiments which were performed to evaluate the developed calibration methods and Section 5 shows the corresponding results.

2. CONCEPT OF THE CAMERA

Different from traditional cameras, which only record the intensity of incident light on the image sensor, a plenoptic camera records the light-field inside the camera as a four dimensional (4D) function. By retracing the path of rays through the main lens the light-field outside the camera can be calculated. In (Gortler et

*Corresponding author.

point and the MLA b can be described as a function of the base line length d , the distance between MLA and sensor B , and the estimated parallax p_x , as given in eq. (5).

$$b = \frac{d \cdot B}{p_x} \quad (5)$$

A point occurs in more or less micro images depending on the distance of its virtual image to the MLA. Thus, dependent on this distance, the length of the base line d , which is used for triangulation, changes. If the triangulation would be performed by using two neighbored micro images, the based line would be equivalent to the micro lens aperture ($d = D$). Since in a Raytrix camera two neighbored micro lenses have different focal lengths, they never focus the same point on the sensor and thus, the baseline is always greater than the micro lens aperture ($d > D$).

The distance B between sensor and MLA is not known exactly, thus, the depth map which is supplied by the plenoptic camera is the distance b divided by B . This relative depth value is called virtual depth and is denoted by v . From eq. (5) the virtual depth v as a function of the estimated parallax p_x and the base line distance d can be derived, as given in eq. (6).

$$v = \frac{b}{B} = \frac{d}{p_x} \quad (6)$$

The virtual depth can only be calculated for a point which occurs focused in at least two micro images. Thus, caused by the hexagonal arrangement of the MLA with three different focal lengths, as it is in a Raytrix camera, a minimum measurable virtual depth of $v_{min} = 2$ results (Perwass and Wietzke, 2012). Since one point usually occurs focused in more than two micro images, its parallax can be estimated by using more than two images.

2.2 Depth accuracy

Based on the rules know from the theory of propagation of uncertainty one can see how an error of the estimated parallax will effect the depth accuracy. From the derivative of v with respect to the measured parallax p_x the standard deviation of the virtual depth σ_v can be approximated as given in eq. (7).

$$\sigma_v \approx \left| \frac{\partial v}{\partial p_x} \right| \cdot \sigma_{p_x} = \frac{d}{p_x^2} \cdot \sigma_{p_x} = \frac{v^2}{d} \cdot \sigma_{p_x} \quad (7)$$

From eq. (7) it is obtained that the accuracy of the virtual depth decays proportional to v^2 . The base line distance d is a discontinuous function of the virtual depth v , because depending on the virtual depth v of a point the maximum base line length of micro lenses which see this point changes. This finally leads to a discontinuous dependency of the depth accuracy as function of the object distance a_L . However, on average the base line distance d is proportional to the virtual depth v . So in total the accuracy of the virtual depth decays approximately proportional to v .

The relationship between the image distance b_L and the virtual depth v is defined by the linear function given in eq. (8).

$$b_L = b + b_{L0} = v \cdot B + b_{L0} \quad (8)$$

Here b_{L0} is the unknown but constant distance between main lens and MLA. Using the thin lens equation (1) one can finally express the object distance a_L as function of the virtual depth v . If the derivative of a_L with respect to b_L is calculated, the standard deviation of the object distance σ_{a_L} can be approximated as given

in eq. (9).

$$\begin{aligned} \sigma_{a_L} &\approx \left| \frac{\partial a_L}{\partial b_L} \right| \cdot \sigma_{b_L} = \frac{f_L^2}{(b_L - f_L)^2} \cdot \sigma_{b_L} \\ &= \frac{(a_L - f_L)^2}{f_L^2} \cdot \sigma_{b_L} = \frac{(a_L - f_L)^2}{f_L^2} \cdot B \cdot \sigma_v \end{aligned} \quad (9)$$

For object distances which are much higher than the focal length of the main lens f_L the approximation in eq. (9) can be further simplified as given in eq. (10). From eq. (10) one can see, that for a constant object distance a_L the depth accuracy increases proportional to f_L^2 .

$$\begin{aligned} \sigma_{a_L} &= \frac{(a_L - f_L)^2}{f_L^2} \cdot B \cdot \sigma_v \\ &\approx \frac{a_L^2}{f_L^2} \cdot B \cdot \sigma_v \quad \text{for } a_L \gg f_L \end{aligned} \quad (10)$$

For a constant focal length of the main lens the depths accuracy decays proportional with a_L^2 . But, since large object distances are equivalent to a small virtual depth, the depth accuracy as a function of the object distance a_L is of better nature than given in eq. (10).

3. METHODS OF CALIBRATION

This section presents the developed methods to calibrate the focused plenoptic camera. We divide the calibration into two separate parts. In the first part we calibrate the optical imaging process disregarding the estimated depth map. The second part addresses the calibration of the depth map.

3.1 Calibration of the optical imaging

For the calibration of optical imaging process the synthesized image is considered to result from a central projection, as it is also done for regular cameras. The intrinsic parameters of the pinhole camera and the distortion parameters are estimated by the calibration method supplied by OpenCV. This method is mainly based on the approach described in (Zhang, 1999). Only the lens distortion model is the one described in (Brown, 1971).

This calibration method uses a number of object points on a plane which are recorded from many different perspectives. The projection from an object point on the plane to the corresponding image point can be described by a planar homography as given in eq. (11). Here, the object point coordinates are denoted by x_W and y_W while $z_W = 0$ and the image point coordinates are denoted by x_I and y_I .

$$\begin{pmatrix} w \cdot x_I \\ w \cdot y_I \\ w \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix} \quad (11)$$

From the planar homography, which has to be estimated for each perspective, eight independent equations are received. From those eight equations six extrinsic parameters have to be calculated. Thus, from each perspective two equations are left to estimate the intrinsic parameters. Since OpenCV defines a pinhole camera model with four intrinsic parameters, the object points have to be recorded from at least two different perspectives. Those four intrinsic parameters are the focal lengths of the pinhole camera in x - and y -coordinates f_x and f_y and the coordinates of the principal point c_x and c_y . In reality much more than two perspectives are used to average measurement errors.

After the determination of the intrinsic parameters, the distortion parameters can be estimated out of the difference between the projection of the object points on the image plane using the estimated extrinsic and intrinsic parameters and the corresponding measured image points.

Based on the corrected image points, again the extrinsic and intrinsic parameters are estimated. This procedure is repeated until consistency is reached.

3.2 Calibration of the depth map

Purpose of the depth map calibration is to define the relationship between the virtual depth v supplied by the Raytrix camera and a metric dimension o which describes the distance to an object.

As described in Section 2 the relationship between the virtual depth v and the object distance a_L relies on the thin lens equation, which is given in eq. (1). Nevertheless, a real lens usually can not be described properly by an ideal thin lens. Thus, for instance the position of the lens' principal plane is not known exactly or the lens conforms to a thick lens with two principal planes.

From Figure 2 one can see, that the image distance b_L is linearly dependent on the virtual depth v . This dependency is defined as given in eq. (12).

$$b_L = v \cdot B + b_{L0} \quad (12)$$

Since the position of the main lens' principal plane cannot be determined, the object distance a_L cannot be measured explicitly. Hence, the object distance a_L is defined as the sum of a measurable distance o and a constant but unknown offset a_{L0} as given in eq. (13).

$$a_L = o + a_{L0} \quad (13)$$

The definitions for the image distance b_L and the object distance a_L (eq. (12) and (13)) are inserted in the thin lens equation (eq. (1)). After rearranging the thin lens equation with the inserted terms, the measurable object distance o is described as a function of the virtual depth v as given in eq. (14).

$$o = \left(\frac{1}{f_L} - \frac{1}{v \cdot B + b_{L0}} \right)^{-1} - a_{L0} \quad (14)$$

This function depends on four unknown but constant parameters (f_L , B , b_{L0} , and a_{L0}). The defined function between the virtual depth v and the measurable object distance o has to be estimated. This estimation can be performed from a bunch of measured calibration points for which the object distance o is known. In this paper we present two novel model based calibration methods. For comparison the function will also be approximated by a traditional curve fitting approach.

3.2.1 Method 1 - Physical model: The first model based approach estimates the unknown parameters of eq. (14) explicitly. However, one additional condition is missing to receive a unique solution for eq. (14) depending on all four unknown parameters. Thus, to solve this equation the focal length of the main lens f_L is specified first by an assumed value. Then the other three unknown parameters are estimated iteratively.

Firstly, the estimated value of the object distance offset \hat{a}_{L0} is set to some initial value. Based on this initial value and the focal length f_L , for each measured object distance $o^{\{i\}}$ the corresponding image distance $b_L^{\{i\}}$ is calculated.

Since the image distance b_L is linearly dependent on the virtual depth v , the calculated image distances $b_L^{\{i\}}$ and the corresponding virtual depths $v^{\{i\}}$ are used to estimate the Parameters B and b_{L0} . Eq. (15) to (17) show the least squares estimation of the parameters.

$$\begin{pmatrix} \hat{B} \\ \hat{b}_{L0} \end{pmatrix} = \left(\mathbf{X}_{Ph}^T \cdot \mathbf{X}_{Ph} \right)^{-1} \cdot \mathbf{X}_{Ph}^T \cdot \mathbf{y}_{Ph} \quad (15)$$

$$\mathbf{y}_{Ph} = \left(b_L^{\{0\}} \quad b_L^{\{1\}} \quad b_L^{\{2\}} \quad \dots \quad b_L^{\{N\}} \right)^T \quad (16)$$

$$\mathbf{X}_{Ph} = \begin{pmatrix} v^{\{0\}} & v^{\{1\}} & v^{\{2\}} & \dots & v^{\{N\}} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}^T \quad (17)$$

Based on the estimated parameters \hat{B} and \hat{b}_{L0} , for each virtual depth $v^{\{i\}}$ the corresponding object distance $\hat{a}_L^{\{i\}}$ is calculated. From the difference between the calculated object distances $\hat{a}_L^{\{i\}}$ and the measured object distances $o^{\{i\}}$ the estimated object distance offset \hat{a}_{L0} is updated as given in eq. (18).

$$\hat{a}_{L0} = \frac{1}{N+1} \cdot \sum_{i=0}^N \hat{a}_L^{\{i\}} - o^{\{i\}} \quad (18)$$

By using the updated value \hat{a}_{L0} , again the image distances $b_L^{\{i\}}$ are calculated and the parameters \hat{B} and \hat{b}_{L0} are updated. The estimation procedure is continued until the variation of \hat{a}_{L0} between two iteration steps is negligible small.

For this method the focal length of the main lens f_L does not have to be known precisely. There exists an optimum solution for any focal length greater than zero. However, the estimated parameters change when the assumed focal length is changed to a different value. Though, if the specified focal length does not comply with the real one, the estimated parameters differ from the real physical dimensions. Nevertheless, the estimated set of values is consistent and will not affect the limits of accuracy of the object distance o .

3.2.2 Method 2 - Behavioral model: The second model based approach relays also on the function defined in eq. (14). However, this method does not estimate the physical parameters explicitly as done in method 1, but a function which behaves similar to the physical model.

Eq. (14) can be rearranged to the term given in eq. (19). Since the virtual depth v and the object distance o both are measurable dimensions, a third measurable variable $u = o \cdot v$ can be defined.

$$o = o \cdot v \cdot \frac{B}{f_L - b_{L0}} + v \cdot \frac{B \cdot a_{L0} - B \cdot f_L}{f_L - b_{L0}} + \frac{b_{L0} \cdot a_{L0} - a_{L0} \cdot f_L - b_{L0} \cdot f_L}{f_L - b_{L0}} \quad (19)$$

Thus, from eq. (19) the term given in eq. (20) results. Here, the object distance o is defined as a linear combination of the measurable variables u and v .

$$o = u \cdot c_0 + v \cdot c_1 + c_2 \quad (20)$$

The coefficients c_0 , c_1 , and c_2 are defined as given in eq. (21) to

(23).

$$c_0 = \frac{B}{f_L - b_{L0}} \quad (21)$$

$$c_1 = \frac{B \cdot a_{L0} - B \cdot f_L}{f_L - b_{L0}} \quad (22)$$

$$c_2 = \frac{b_{L0} \cdot a_{L0} - a_{L0} \cdot f_L - b_{L0} \cdot f_L}{f_L - b_{L0}} \quad (23)$$

Since for eq. (20) all three variables o , v , and u are measurable dimensions, the coefficients c_0 , c_1 , and c_2 can be estimated based on a number of calibration points. For the experiments presented in Section 4 the coefficients c_0 , c_1 , and c_2 are estimated by using the least squares method as given in eq. (24) to (26).

$$\begin{pmatrix} \hat{c}_0 \\ \hat{c}_1 \\ \hat{c}_2 \end{pmatrix} = \left(\mathbf{X}_{Be}^T \cdot \mathbf{X}_{Be} \right)^{-1} \cdot \mathbf{X}_{Be}^T \cdot \mathbf{y}_{Be} \quad (24)$$

$$\mathbf{y}_{Be} = \left(o^{\{0\}} \quad o^{\{1\}} \quad \dots \quad o^{\{N\}} \right)^T \quad (25)$$

$$\mathbf{X}_{Be} = \begin{pmatrix} o^{\{0\}}v^{\{0\}} & o^{\{1\}}v^{\{1\}} & \dots & o^{\{N\}}v^{\{N\}} \\ v^{\{0\}} & v^{\{1\}} & \dots & v^{\{N\}} \\ 1 & 1 & \dots & 1 \end{pmatrix}^T \quad (26)$$

After rearranging eq. (20) the measurable object distance o can be described as a function of the virtual depth v and the estimated parameters c_0 , c_1 , and c_2 as given in eq. (27).

$$o = \frac{v \cdot c_1 + c_2}{1 - v \cdot c_0} \quad (27)$$

3.2.3 Method 3 - Curve fitting: The third method presented in this paper is a common curve fitting approach. This approach approximates the function between the virtual depth v and the measurable object distance o without paying attention to the function defined in eq. (14).

It is known that any differentiable function can be represented by a Taylor-series and thus, by a polynomial of infinite order. Hence, in the approach presented here the functions which describes the object distance o depending on the virtual depth v will be defined as a polynomial as well. A general definition of this polynomial is given in eq. (28).

$$o \approx \sum_{k=0}^M p_k \cdot (v)^k \quad (28)$$

Similar to the second method the polynomial coefficients p_0 to p_M are estimated based on a bunch of calibration points. In the experiments presented in Section 4 a least squares estimator as given in eq. (29) to (31) is used.

$$\begin{pmatrix} \hat{p}_0 \\ \hat{p}_1 \\ \vdots \\ \hat{p}_M \end{pmatrix} = \left(\mathbf{X}_{Pol}^T \cdot \mathbf{X}_{Pol} \right)^{-1} \cdot \mathbf{X}_{Pol}^T \cdot \mathbf{y}_{Pol} \quad (29)$$

$$\mathbf{y}_{Pol} = \left(o^{\{0\}} \quad o^{\{1\}} \quad \dots \quad o^{\{N\}} \right)^T \quad (30)$$

$$\mathbf{X}_{Pol} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ v^{\{0\}} & v^{\{1\}} & \dots & v^{\{N\}} \\ v^{\{0\}^2} & v^{\{1\}^2} & \dots & v^{\{N\}^2} \\ \vdots & \vdots & \ddots & \vdots \\ v^{\{0\}^M} & v^{\{1\}^M} & \dots & v^{\{N\}^M} \end{pmatrix}^T \quad (31)$$

For this method a trade-off between the accuracy of the approxi-

mated function and the order of the polynomial has to be found. A high order of the polynomial results in more effort for calculating an object distance from the virtual depth. Besides, for high orders the matrix inversion as defined in eq. (29) results in numerical inaccuracy. For such cases a different method for solving the least squares problem has to be used (e.g. Cholesky decomposition).

4. EXPERIMENTS

To evaluate the calibration methods presented in Section 3 different experiments were performed. Firstly, Section 4.1 presents the experiments which were performed to evaluate the calibration of the optical imaging process. Secondly, the experiments performed for the evaluation of the depth calibration methods are described in Section 4.2. The results corresponding to the experiments are presented in Section 5.

4.1 Calibration of the optical path

The calibration method for the optical imaging process was performed for a main lens with focal length $f_L = 35$ mm and one with $f_L = 12$ mm.

To calibrate the optical path we used for both focal lengths a planar chessboard pattern with seven times ten fields. This results in a total number of 54 reference points per recorded pattern (one reference point is the connection point of four adjacent fields). One square of the chessboard has a size of 50 mm × 50 mm.

For calibration the pattern was recorded from 50 as different as possible views and it was tried to cover the whole image with calibration points. For both lenses three series of measurements were recorded to determine the variance of the calibration method. The results corresponding to this experiment are presented in Section 5.1.

4.2 Calibration of the depth map

The experiments to investigate the properties of the three presented depth map calibration methods were also performed for the main lens focal length $f_L = 35$ mm and $f_L = 12$ mm.

Based on the chessboard pattern, which was already used to calibrate the optical imaging process, for both focal lengths a series of measurements was recorded. This time, for the 35 mm lens a chessboard with squares of the size 25 mm × 25 mm was used. For the 12 mm lens again the one with square size 50 mm × 50 mm was used. For each object distance, the distance between a defined point close to the camera and a certain point on the chessboard was measured with a laser rangefinder (LRF). Here it can not be guaranteed that the chessboard stands parallel to the image plane of the camera. Thus, the object distance for each of the 54 reference points can be slightly different from the distance measured by the LRF. It is known that the object distance a_L of a square on chessboard is inversely proportional to the size of its image. Thus, for each square on the chessboard a relative object distance can be calculated out of the size of its image. Since it is known that all squares lay on the same plane, this plane is estimated based on the relative distances. Thus, the orientation of the chessboard with respect to the camera is received. Based on the object distance which was measured for one certain point accurately by the LRF and the estimated orientation of the chessboard, for each of the 54 corners in the chessboard pattern a very precise object distance is calculated.

For the 35 mm lens the chessboard pattern was recorded at 50 different object distances o in the range from approximately 0.7 m

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
σ_{Rep} [Pixel]	0.703	0.870	0.830

Table 1: Optical path calibration results for $f_L = 12$ mm. The optical path calibration was performed three times for the 12 mm focal length. For each calibration series a chessboard pattern was recorded from 50 different perspectives.

to 5.2 m. For the 12 mm lens the pattern was recorded at 50 different object distances o in the range from approximately 0.7 m to 3.9 m. Since each pattern has 54 reference points, 54 measurement points are received for each object distance o .

To evaluate the three calibration methods different experiments were performed based on the two series of measurements. Beside the physical model and the behavioral model based calibration method, the curve fitting approach was performed by using a third and a sixth order polynomial.

In the first experiment only the values of five measured object distances o were used for calibration. This experiment was performed to evaluate if a low number of calibration points is sufficient to receive reliable calibration results.

In a second experiment only the 50 % of the measurement points with the lowest object distances (for $f_L = 35$ mm object distances up to approx. 2.6 m and for $f_L = 12$ mm object distances up to approx. 2.1 m) were used for calibration. In this experiment it was supposed to investigate how strong the estimated functions are drifting off from the real model outside the range of calibration.

To evaluate the accuracy of the depth map, based on all measured points the root mean square error (RMSE) with respect to the object distance was calculated for both focal lengths. In this experiment the object distances which were calculated from the virtual depth were converted to metric object distance by using the behavioral model presented in Section 3.2.2.

Section 5.2 presents the results corresponding to the experiments performed for the depth map calibration.

5. RESULTS

This section presents the results of the performed experiments. In Section 5.1 the results of the calibration of the optical imaging process are presented, while Section 5.2 shows the results corresponding to the calibration of the depth map.

5.1 Calibration of the optical path

As already mentioned in Section 3.1 the used calibration method described in (Zhang, 1999) calculates four intrinsic as well as distortion parameters. The intrinsic parameters are the focal lengths of the underlying pinhole camera (f_x and f_y) and the corresponding principal point (c_x and c_y). Table 1 and 2 show the calibration results for the 12 mm and 35 mm focal length respectively. For both focal lengths the results of the three calibration series are presented in the tables. Besides, for each calibration series the root mean square (RMS) of the reprojection error σ_{Rep} was calculated. The reprojection error e_{Rep} is the Euclidean distance between a measured two dimensional (2D) image point and the image point which results from the projection of the corresponding

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	426.3
c_y [Pixel]	346.6	389.5	323.6
σ_{Rep} [Pixel]	0.332	0.333	0.342

Table 2: Optical path calibration results for $f_L = 35$ mm. The optical path calibration was performed three times for the 35 mm focal length. For each calibration series a chessboard pattern was recorded from 50 different perspectives.

three dimensional (3D) object point, based on the camera model. Eq. (32) gives the definition of the RMS of the reprojection error.

$$\sigma_{Rep} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (e_{Rep}^i)^2} \quad (32)$$

As one can see, the reprojection error for the 12 mm focal length is about twice to three times as high as the one for the 35 mm focal length. This can be explained since the effective image resolution depends on the virtual depth. For a low virtual depth the effective resolution is higher than for a high virtual depth. The 12 mm focal length results in a larger field of view (FOV) than the 35 mm focal length. Thus, for the 12 mm focal length the pattern had to be recorded from a much closer distance than for the 35 mm focal length. This likely resulted in a higher virtual depth. Nevertheless, since the calibration results showed a very high variance between the single calibrations we did not investigate this phenomena further.

As one can see from Table 1 and 2, the estimated intrinsic parameter are varying a lot from one calibration series to the next, especially for $f_L = 35$ mm. The reason therefore is that especially for large focal lengths the extrinsic parameter, which have to be estimated for each recorded pattern and the intrinsic parameters are strongly correlated and thus cannot be separated accurately. The calibration could be improved by using a 3D calibration object. From a 3D object three conditions more than from a planar object are received per perspective.

5.2 Calibration of the depth map

In this section only the results for the calibration of the 35 mm lens are presented. For $f_L = 12$ mm comparable results were achieved.

Figure 4 shows the results corresponding to the first experiment. The red dots represent the calibration points for the five object distances. The green dots are the remaining measured points which were not used for calibration. As one can see, the physical model as well as the behavioral model are almost congruent. Both curves match the measured distances very well over the whole range from approx. 0.7 m to 5.2 m. For the polynomials of order three and six instead, five object distances are not sufficient to approximate the function between virtual depth v and object distance o accurately. Both functions fit to the points used for calibration but do not define the underlying model properly in between the calibration points.

Figure 5 shows the results of the second experiment. Again, the points used for calibration are represented as red dots and the green dots represent the remaining points. Both model based calibration approaches again are almost congruent and describe very well the measured distances in the range of calibration up to 2.6 m. In this range also the estimated polynomials fit the

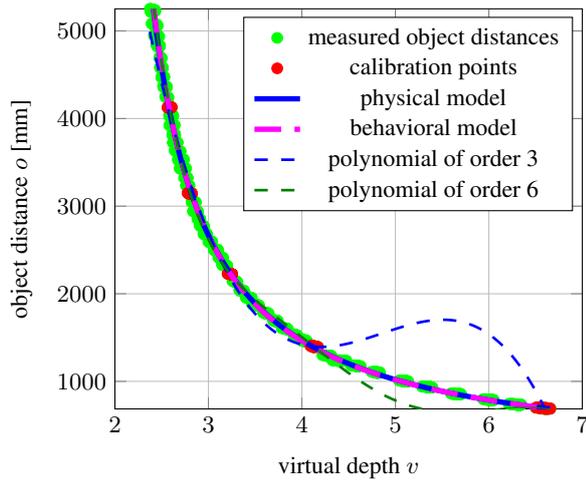


Figure 4: Results of the first experiment for the depth map calibration. Five object distances were used for calibration. The calibration was performed based on the physical model, the behavioral model and the curve fitting approach.

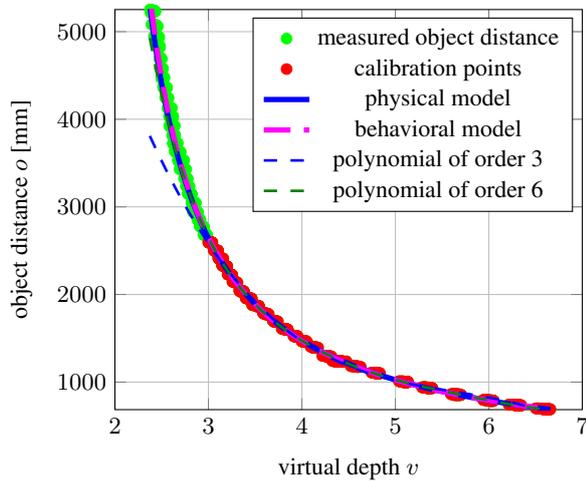


Figure 5: Results of the second experiment for the depth map calibration. Object distances up to approx. 2.6 m were used for calibration. The calibration was performed based on the physical model, the behavioral model and the curve fitting approach.

measured distances very well. Nevertheless, for object distances larger than 2.6 m especially the third order but also the sixth order polynomial are drifting away from the series of measurements. The functions of both model based approaches still match the measured values very well up to an object distance of 5.2 m. Thus, the functions underlying the two model based calibration methods conform with the reality. Both functions are able to convert the virtual depth to an object distance even outside the range of calibration points with good reliability.

As mentioned in Section 4.2, to evaluate the depth resolution of the plenoptic camera the complete series of measurement points was used. For each of the 50 object distances the RMSE is calculated. The RMSE is calculated from the 54 values o_{Meas} measured for each object distance and the corresponding values $o(v)$ calculated as a function of the virtual depth v . The virtual depth values v are transformed into object distance based on the function received from the behavioral model. The RMSE is defined

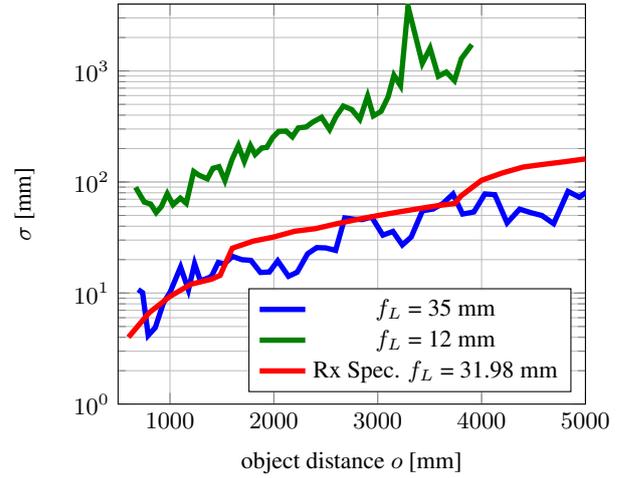


Figure 6: Depth accuracy of a Raytrix R5 camera with two different main lens focal lengths ($f_L = 35$ mm and $f_L = 12$ mm)

as given in eq. (33).

$$\sigma = \sqrt{\frac{1}{54} \cdot \sum_{i=0}^{53} (o_{Meas}^{\{i\}} - o(v^{\{i\}}))^2} \quad (33)$$

Figure 6 shows the RMSE of the depth for the 12 mm and the 35 mm focal length respectively. Besides, the figure shows the simulated accuracy of the camera which is given by Raytrix. For the simulation a maximum focus distance of 10 m and a focal length of $f_L = 31.98$ mm was defined. Since the RMSE for the 12 mm focal length increases strongly with rising object distance, the axis of ordinates is scaled logarithmically.

The results for the focal length $f_L = 35$ mm conform very well with the specification given by Raytrix. However, it cannot be guaranteed that the object distances measured during the experiments are equivalent to the distance given by Raytrix. The distance given by Raytrix is the distance between an object and the sensor plane. Thus, the Raytrix specification actually might be shifted by a small offset along the axis of abscissas.

As one can see, the accuracy measured for $f_L = 12$ mm is worse by a factor of approx. 8 to 10 compared to $f_L = 35$ mm. This conforms to the evaluation made in Section 2.2, according to which $\frac{35^2}{12^2} = 8.5$ holds true.

6. CONCLUSION

In this article we developed an analytical expression for the depth map accuracy supplied by a focused plenoptic camera. The expression describes the dependency of the depth accuracy from camera specific parameters like the focal length of the main lens f_L .

Besides, it was shown that the image synthesized from the camera's raw image can be described by the projection model of a pinhole camera. Hence, the intrinsic and distortion parameters of the synthesized image can be estimated by traditional calibration methods.

The most important contribution of the article is the development of two novel, model based methods for the depth map calibration which proved to be superior to traditional methods. The performed experiments showed that both methods correspond to the behavior of the camera very well. Thus, only a small number of

measurements is needed for calibration. Besides, it was shown that the estimated functions are valid in excess of the calibration range.

In future development the method for calibrating the optical imaging process has to be improved since the estimated intrinsic and distortion parameters showed high variances. A reliable estimation of the camera parameters is needed to calculate an accurate 3D point cloud from the depth map supplied by the camera.

In near future the plenoptic camera is supposed to operate in navigation applications for visually impaired people. The camera seems to be suited for such an applications because of its small dimensions. Furthermore, the depth map accuracy is sufficient for this purposes.

ACKNOWLEDGEMENT

This research is funded by the Federal Ministry of Education and Research of Germany in its program "IKT 2020 – Research for Innovation".

REFERENCES

- Brown, D. C., 1971. Close-range camera calibration. *Photogrammetric Engineering* 37(8), pp. 855–866.
- Dansereau, D., Pizarro, O. and Williams, S., 2013. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1027–1034.
- Gortler, S. J., Grzeszczuk, R., Szeliski, R. and Cohen, M. F., 1996. The lumigraph. In: *Proceedings of the 23rd annual conference on computer graphics and interactive techniques, SIGGRAPH '96*, ACM, New York, NY, USA, pp. 43–54.
- Ives, F. E., 1903. Parallax stereogram and process of making same.
- Johannsen, O., Heinze, C., Goldluecke, B. and Perwa, C., 2013. On the calibration of focused plenoptic cameras. In: *G CPR Workshop on Imaging New Modalities*.
- Lippmann, G., 1908. Epreuves rversibles. photographies integrales. *Comptes Rendus De l'Acadmie Des Sciences De Paris* 146, pp. 446–451.
- Lunsdaine, A. and Georgiev, T., 2008. Full resolution lightfield rendering. Technical report, Adobe Systems, Inc.
- Ng, R., 2006. Digital light field photography. PhD thesis, Stanford University, Stanford, USA.
- Perwass, C. and Wietzke, L., 2012. Single lens 3d-camera with extended depth-of-field. In: *Human Vision and Electronic Imaging XVII*, Burlingame, California, USA.
- Zhang, Z., 1999. Flexible camera calibration by viewing a plane from unknown orientations. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Vol. 1, pp. 666–673.