A Practical Method for Fully Automatic Intrinsic Camera Calibration Using Directionally Encoded Light

Mahdi Abbaspour Tehrani
University of California Irvine
mabbaspo@ics.uci.edu

Thabo Beeler
Disney Research
thabo@disneyresearch.com

Anselm Grundhöfer
Disney Research
anselm@disneyresearch.com

Abstract

Calibrating the intrinsic properties of a camera is one of the fundamental tasks required for a variety of computer vision and image processing tasks. The precise measurement of focal length, location of the principal point as well as distortion parameters of the lens is crucial, for example, for 3D reconstruction [27]. Although a variety of methods exist to achieve this goal, they are often cumbersome to carry out, require substantial manual interaction, expert knowledge, and a significant operating volume. We propose a novel calibration method based on the usage of directionally encoded light rays for estimating the intrinsic parameters. It enables a fully automatic calibration with a small device mounted close to the front lens element and still enables an accuracy comparable to standard methods even when the lens is focused up to infinity.

Our method overcomes the mentioned limitations since it guarantees an accurate calibration without any human intervention while requiring only a limited amount of space. Besides that, the approach also allows to estimate the distance of the focal plane as well as the size of the aperture. We demonstrate the advantages of the proposed method by evaluating several camera/lens configurations using prototypical devices.

1. Introduction

Accurate intrinsic camera calibration is one of the core requirements for a multitude of computer vision tasks. Although reliable methods exist to do so, most of them require substantial human interaction and a carefully configured setup to successfully compute accurate calibration parameters. Especially the accuracy of the estimated lens distortion parameters may suffer if the user is not sufficiently covering the complete area of the image sensor. This is particularly challenging for wide angle lenses where a large volume needs to be covered, or for tele-lenses where the calibration volume is at a distance.

Overcoming these issues is the primary goal of this work and therefore we propose a simple, fully automatic calibration method that estimates the intrinsic parameters of a camera reliably without any human intervention. We present a practical calibration device that only requires a minimal working volume directly in front of the lens and the overall process can be carried out within few minutes with repeatingly reliable accuracy at the push of a button. Unlike most existing calibration methods we employ directionally encoded light which allows to calibrate additional intrinsic parameters, such as focus distance or aperture size. The proposed algorithm is not exclusively aimed for being used by scientists or engineers, but to simplify calibration for maintenance personnel, or amateurs who want to use a specific software application requiring a calibrated camera. From practical experience, achieving an acceptable camera calibration is a seriously difficult task for such non-experts.

In summary, our proposed method includes the following novel contributions:

- A reliable and fully automatic intrinsic camera calibration algorithm.
- A prototypical hardware setup which requires only a minimal working volume directly in front of the lens.
- An algorithm not requiring any system pre-calibration.
- The use of directionally encoded light to calibrate additional lens properties, such as focus distance and aperture size.
- An evaluation of the system based on various different lenses and cameras including a comparison with Zhang's method [31].

The remainder of the paper will start with a short discussion of the related work (2), followed by a description of the proposed setup (3) and calibration method (4). In section 5 we will introduce our hardware prototype used to evaluate the system in section 6.

2. Background and Related Work

In the following we will give a short introduction about the basic camera model and the calibration method we are

using, and will give an overview of the most relevant related work with respect to camera calibration as well as the proposed display configuration.

Pinhole Camera Model   Our work is based on the widely used camera calibration introduced by Zhang et. al. [31] using a pinhole camera model. Assume (X, Y, Z) is the coordinate of a point in 3D space and (x, y) is the coordinate of the projection of this point on the 2D image plane of the camera sensor. Then the relation between these two coordinates is given by:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K \begin{bmatrix} R|t \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (1)$$

Where s is an arbitrary scale factor. R and t are the rotation and translation matrices of the camera to the world coordinate system. K is the camera's intrinsic matrix and is defined as $K = \begin{bmatrix} f_x & \alpha & x_p \\ 0 & f_y & y_p \\ 0 & 0 & 1 \end{bmatrix}$ Where f_x and f_y are the focal lengths of the camera in horizontal and vertical direction. (x_p, y_p) is the coordinate of the camera's principal point on the image plane. α is the skew between two image axes which, in our work, is assumed to be equal to zero which is a reasonable assumption for the vast majority of the existing modern cameras.

To also account for radial and tangential lens distortion, one of the most common models is used [6, 5, 9], where several coefficients k describe radial and tangential distortion depending on the distance to the system's optical axis.

Multi-layer Displays   Adding light-attenuating or -directing layers onto emissive displays has been widely explored for the generation of autostereoscopic displays for multi-view stereo [28] or near-eye displays [19]. These approaches are based on the principle of generating a spatially varying light field such that for varying observer positions a different content is presented. We apply such displays in the context of intrinsic calibration: directional light is generated by light-attenuating layers in front of a high-resolution screen, captured by the camera, and the information is used for calibration. The Bokode system [20] uses a similar optical design approach but within a different context.

Camera Calibration Methods   There is a large body of research papers focusing on the task of accurately calibrating the intrinsic properties of cameras. They differ from each other by employing different optimization methods, camera models, calibration targets or user interfaces. Many

of the existing methods use planar patterns for calibration and offer widely available toolboxes [4, 22] which require expert knowledge to reliably generate an accurate calibration. The method described in [31] is one of the most commonly used methods which uses multiple images of a plane captures from unknown orientations. Many of the existing works use this method as a reference to compare the accuracy of their results. Another approach was presented in [8] which uses parallel lines in the image to find vanishing points to calibrate the camera.

The performance of these methods strongly depends on the accuracy of the extraction of the correspondence points. [21] proposed a highly accurate feature detection method for camera calibration using checker board patterns. [29] uses circular control marks to accurately find the location of features. [10] proposed a calibration algorithm for planar square, circle, and ring patterns. Here an iterative refinement approach to undistort and unproject the calibration pattern to the canonical plane is used to increase the accuracy of the localization.

Most existing works focus on the development of novel camera models and improved calibration accuracy, but the calibration process can be time consuming and partially even a manually guided process for identifying control points. There is a body of work focusing on automating these steps. Since most of these methods use planar markers, such as checkerboards, its robust detection is highly important: [3] proposed a method for automated checkerboard detection. In [11], [12], [1] and [24] self-identifying patterns for plane-based calibration are used which can also be partially occluded. The usage of random dot based markers instead of 2D tags or checkerboards is proposed in [23] which is based on the advantage that those can be detected more reliable in out of focus image regions. [18] presented an automatic technique for grid corner extraction. The paper of Ha et.al. [13] proposes a method which handles out of focused input images by displaying a series of structured light patterns on a small display close to the lens to generate display-plane-to-camera correspondences from a variety of different views.This method requires many different views to gather a sufficient number of correspondences and fails for systems with large sensor sizes and apertures.

All of these methods use images of differently oriented patterns, captured by the user. As a result, the calibration can vary for different image sets, and a bad collection of images can result in a poor quality. Especially for estimating accurate distortion coefficients it is crucial to have a set which covers all areas of the image plane. So an accurate calibration usually needs an expert user. [24] proposed an iterative approach that uses the current calibration state to suggest the next most helpful pattern position to the user. They show that their system yields better calibration results than existing tools using a pre-captured set of images.

In [2], a single image calibration was proposed, but since it is focused on the calibration of endoscope cameras, it is not clear whether the claimed quality can be achieved for other camera and lens configurations as well. Another single image approach using lenticular arrays [26] was proposed recently. Although the method is able to estimate the focal length of a camera, it is not accurate enough to estimate distortion parameters and can only be applied using a centered principal point as well as with color cameras.

Figure 1. Sketches of the proposed setups. *Left:* Minimal configuration using two layers: the first layer is an emissive planar display panel and the second layer, a thin opaque sheet with a grid of holes, is placed slightly angled in front. *Center:* Using two tilted hole planes enables a better coverage for wide angle lenses and an additional calibration plane. *Right:* Three layered hole planes consisting of yellow, cyan and magenta filters allow to capture correspondences to three planes with one configuration.

3. Device Setup

The main goal of the proposed calibration method is to simplify and fully automate intrinsic camera calibration using a small device mounted directly in front of a lens. The proposed hardware can be assembled from low-cost off-the-shelf components and requires no pre-calibration. Unlike most existing calibration methods, we make use of directionally encoded light rays which allows us to calibrate additional intrinsic properties, such as aperture diameter and focal distance of the lens.

To generate such directional light rays we devise a multi-layer device in which one, the background layer, is an emissive high resolution display. The other layers are one or more thin planar sheets ($d < 0.2mm$), which are opaque for at least parts of the visible spectrum, but contain a grid of small holes ($r < 0.1mm$). With this configuration, light originating from the display can be directionally encoded while passing through the masking layers (cf. Fig. 1).

Such a setup leads to some unique effects which have to be considered during the calibration process: each of the holes in a masking layer approximates a pinhole projection, and thus the pixels of the background layer will be imaged in focus by the camera, irrespectively of the actual focus plane of the lens. This is beneficial, since it allows to place the calibration device very close to the lens during calibration. The holes themselves, however, will be subject

to *defocus blur* depending on the size of the camera aperture. Furthermore, the pixels of the background layer may appear shifted on the camera sensor. In order to achieve accurate calibration results, these effects have to be accounted for as we show in the following.
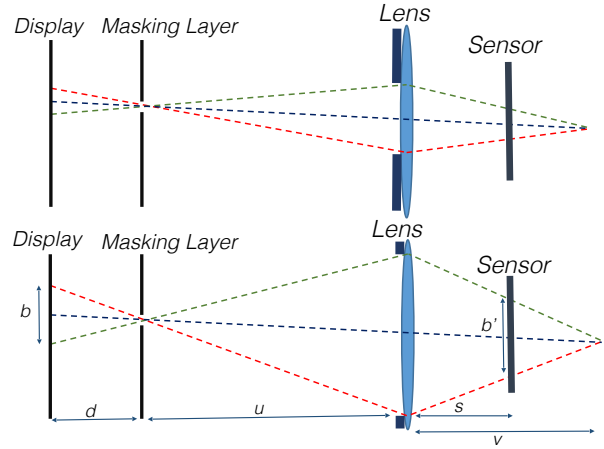

Figure 2. Defocus blur of a point source. The size and shape of the CoC depends on the lens' aperture size and shape. *Top:* A closed physical aperture creates a smaller image of the hole. *Bottom:* A large aperture size creates bigger images of the hole and a larger area of the display panel is visible through the hole.

3.1. Defocus Blur

If the camera was focused on the masking layer, then the image of each hole of the layer would amount to a single illuminated point on the camera sensor. However, since the device is meant to be mounted closely in front of the lens, which is likely focused at a farther distance, the projection of the holes will amount to circular disks on the sensor instead of points, known as circles of confusion (CoC) [16]. Size and shape of the disks depend on the one hand on the distance of the holes to the sensor and the focal plane and on the other hand on the shape and size of the camera aperture as illustrated in Fig. 2 and can be observed in Fig. 3.
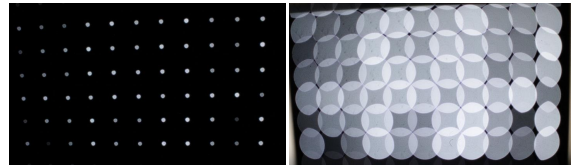

Figure 3. Captured images with closed (*Left*) and open aperture (*right*). The CoC of the holes is influenced by the aperture but also by its shape as well as the spatial location on the image plane.

3.2. Pixel Shift

As described above, the use of pinholes within the optical path leads to circular disks on the camera sensor. The content of these disks is a sharp image of a number of pixels of the background layer, since the holes are essentially
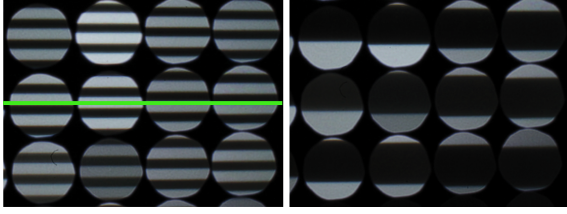
Figure 4. Captured images of a masking layer. Gray code patterns of two varying frequencies (*Left,Right*) are displayed on the background layer. Note how the pixel shift described in Sec. 3.2 leads to slightly shifted lines when observed through different holes, best seen in the left image at the green line.
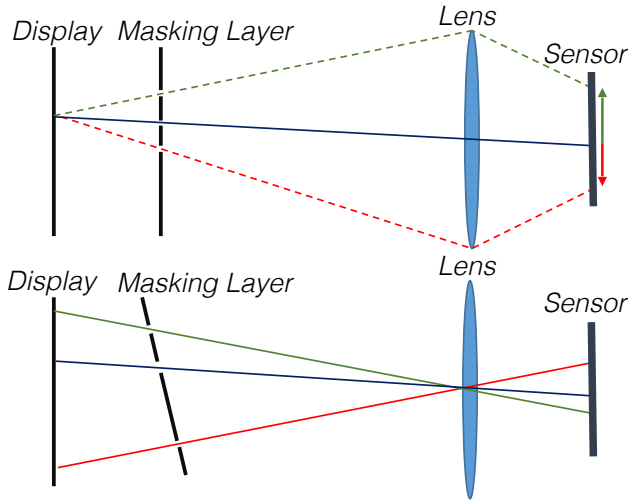

Figure 5. *Top*: Different light rays from a (defocused) pixel on the display will intersect with camera's sensor at different position. Only the ray which passes through the optical center of the lens (blue ray) resides on the same line. *Bottom*: We only consider the rays passing the optical center of the lens for the calibration.

pinhole projections, as can be seen in Fig. 4. Each pixel corresponds to a directional ray travelling through the optics of the camera. Rays which are not passing through the optical center of the lens will be distorted, which amounts in a translational shift on the sensor, which we call pixel shift. This phenomenon is illustrated in Fig. 5 and can be observed in Fig. 4, where the lines in neighbouring disks are slightly shifted. For accurate calibration, we therefore rely solely on the rays that pass through the optical center of the lens and are hence distortion free. How to extract these will be described in Sec. 4.3.

4. Proposed Calibration Method

With the mentioned optical effects in mind, the proposed device can be used to estimate not only the standard intrinsic properties such as focal length and distortion parameters, but also the focal distance and aperture size. Having mounted the device in front of the camera, the correspondences between the holes of the masking layers and the

camera pixels are computed and a mapping between camera and display pixels is generated. Since both background layer and masking layers are planar, we can use the computed correspondences to establish homographies between the planes and employ the method of [31] to calibrate the intrinsic matrix and distortion properties of the camera. Other lens properties such as focal plane or aperture size can then be estimated by leveraging the directionally encoded light rays. Algorithm 1 shows an overview of our method. In the following sections we will explain each step in detail.

Algorithm 1 Calibration Method

1: Choose the smallest aperture for the camera (If possible)
2: Show white for all pixels of the display layer
3: Take an image of the hole layer
4: Find correspondences of masking layer with camera's pixels
5: Increase the aperture size such as to avoid CoC overlap (If possible)
6: Show a sequence of gray code patterns on display layer and capture the images by the camera
7: Generate the mapping betwen camera and display pixels by decoding the images of the gray codes.
8: Calibrate the intrinsics of the camera and position and orientation of layers using the method of [31]
9: Estimage the focal length and aperture size using the directionaly encoded light rays.

4.1. Correspondence to Masking Layer

To establish correspondence between the holes in the masking layers and their projection onto the imaging sensor we need to determine their projected location onto the sensor and identify their coordinate on the masking layer.

Hole pixel locations: To compute the projected location of the holes on the image sensor the background display is turned on completely, yielding homogeneously filled circles of confusion on the sensor. As shown in Fig. 3 the size of the aperture directly impacts the size of the CoC. The smaller the aperture the better but at the very least the disks should not overlap. If the aperture cannot be controlled or closed sufficiently, then the spacing of the holes on the masking layer has to be adjusted to prevent overlapping CoC.

Even with a small aperture, field curvature, optical aberration artifacts, the aperture shape as well as lens distortion might influence the shape of the CoC such that it deviates from a perfect circle (cf. Fig. 6). To account for that in the most generic way, the projected location of the holes is estimated by fitting a 2D Gaussian function to the image of each hole. Fig. 6 shows a sample of a fitted Gaussian function and the estimated hole center. Since the hole can be assumed to have infinitesimally small size, this method does

not suffer from foreshortening effects known to pose challenges for location estimation using spatial fiducials such as circular feature points[17].
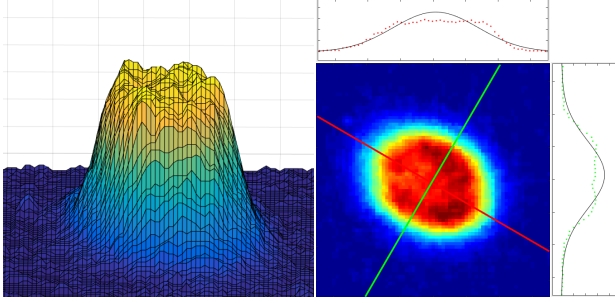

Figure 6. Color coded height field plot of a sample hole of one of the captured images of the blocking layers (*Left*) and the center found by fitting a 2D Gaussian function (*Right*).

Hole layer coordinates: The masking layers contain a regular grid of holes and we need to determine the local coordinates of all detected holes on the layers. Therefore, one of the holes is chosen as the origin of the coordinate system, and by detecting the grid we can find the coordinates of the other holes based on this origin. Other methods, such as automatically detected random patterns [30] would be possible as well, but the grid arrangement guarantees a regular pattern which is easier to process and more robust with respect to potentially overlapping circles.

If the lens contains strong distortions, a straightforward grid search is likely to fail. However, as illustrated in the bottom row of Fig. 5 there is a linear projection between 2D coordinates of holes on the masking layer and the coordinates of corresponding pixels on the display panel. This describes a relationship between two planes independent of the camera and is free of any lens distortion. Fig. 7 shows the position of visible pixels through the holes on the display panel. Since there is only a relatively small angle between these planar layers and no non-linear distortion, only a small amount of keystoning occurs, which can be neutralized using an affine transformation. As a result we can easily detect the grid once we know the correspondences between the holes and display pixels. How this is achieved will be described in the following Section.

4.2. Correspondence to Background Layer

Having established correspondence between the camera and masking layers, the ones to the background layer still have to be estimated. Despite the fact that the camera is out of focus, a sharp and magnified image of the display's pixels is observed through each hole, since they are so small that they act like pinhole projections. This magnified image of the display helps finding accurate pixel correspondences to the camera. For this purpose complementary black and
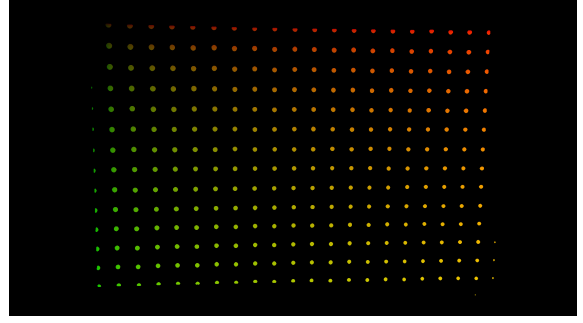

Figure 7. Color coded pixel coordinates sensed by the camera through the holes on the display. Note that the projection of the holes on the display is free of distortions and only contains a small amount of keystoning.

white Gray code patterns [7] are displayed and captured [1]. Depending on the display's resolution, this process requires around 40-50 images to be captured to achieve reliable pixel correspondences. At this step, images are captured with larger aperture sizes, if possible, since this increases light throughput and allows to observe more display pixels through each hole. To avoid distortions resulting from a noticeable amount of light diffracting at the boundary of the hole, pixels close to the edge of a hole are masked and disabled for further processing. Next, by decoding the images of the Gray code patterns, pixel correspondences between the camera and the display are generated. To compute the plane homographies we select only those pixels on the background display whose rays pass through the optical center of the lens to avoid issues with pixel shift as described in Sec. 3.2. These can be easily determined, as they correspond to the pixels whose projection onto the image sensor coincide with the detected pixel locations of the holes. The other correspondences will be used later on to calibrate the focal plane of the lens.

4.3. Camera and Lens Parameter Estimation

Having computed the correspondences from camera pixels to the masking layers as well as the background layer, the standard homography based calibration method proposed by Zhang et al. [31] can be used to calculate the intrinsic calibration matrix as well as distortion parameters. Furthermore, after having estimated this calibration, the distance to the focal plane as well as the current aperture size can be estimated by our system as well.

Intrinsic Matrix and Distortion Parameters The last Sections described how to gather a set of point correspondences between the display and the masking layers to the camera's image plane. These correspondences are directly used for camera calibration using the method introduced by

---

[1]The interested reader is referred to [25] for more information about structured light scanning
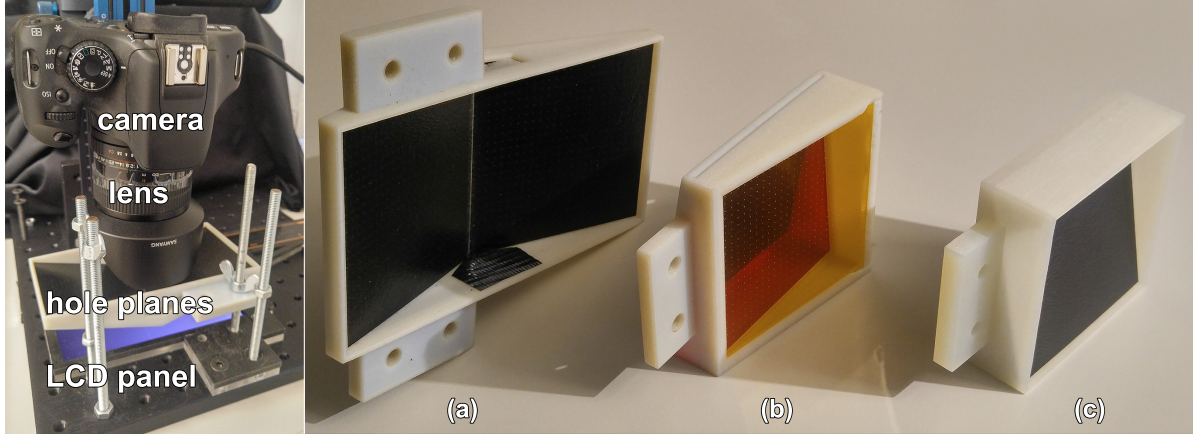
Figure 8. *Left*: Prototypical setup with two masking layers. The background layer is an emmissive LCD panel and the other layers consist of thin opaque sheets perforated by a grid of holes. *Right*: Three different hole layer configurations: (a) v-shaped planes to get better coverage when calibrating wide angle lenses as shown on the left, (b) cyan, magenta and yellow layers to generate superimposed correspondences using color sensors, and (c) a minimal, single plane configuration.

Zhang et al. [31]. Since parallel planes do not provide additional constraints for this calibration method and thus should be avoided, the masking layers are rotated by several degrees (cf. Fig. 1). First an initial guess of the camera parameters is estimated by solving the closed-form solution using the homographies between the planes and the camera. This result is further refined in a non-linear optimization step as described in [14]. Although this method is used for intrinsic calibration, the extrinsic information, i.e. rotation and translation between the camera and the individual planes is computed as well and can be used to estimate the lens aperture diameter as well as the focus distance as described in the following.

Aperture Size As discussed in Sec. 3.1 because of the defocus blur, the image of each hole creates a disk on the camera's image plane. As a result, several pixels of the display panel are visible through each hole. The geometrical relationships are depicted in Fig. 2 and can be described as follows: lets denote the distance of a hole from the background layer as d and from the camera sensor as u and the camera's aperture size as a. Based on similarity of the triangles [20] the size of the visible part of the display panel (b) through the hole is

$$b = \frac{ad}{u}.\tag{2}$$

This relationship can be used to estimate the current lens's aperture setting. The effect of varying the camera aperture is illustrated in Fig. 3.

Mohan et al. [20] show that the size of the CoC for each hole is

$$b' = \frac{(v-s)a}{v},\tag{3}$$

where s is the distance of the sensor from the lens, b' is the size of the CoC on the camera's image plane, and v is the

distance of the image of the hole to the lens. Based on the thin lens model equation [15] we have

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}.\tag{4}$$

Using above equation and removing v from 3 we have

$$b' = \frac{(fu - su + fs)a}{fu}.\tag{5}$$

The magnification scaling of the lens is $M = \frac{b'}{b}$. So using 2 and 5 we have

$$M = \frac{b'}{b} = \frac{fu - su + fs}{fd}.\tag{6}$$

The above equation can be used to estimate s, which describes the distance of the sensor from the lens, which also allows us to compute the distance of the focal plane from the camera.

In Equ. 2 it is shown that the size of the visible part of the display depends on the camera's aperture size and the distance of the hole from the lens and display panel. Having already calibrated the intrinsic and extrinsic parameters allows to calculate the distance of each hole from the lens and display panel. By having estimated the pixel correspondences between the camera's image plane and the display, the size of the visible part of the display panel through each hole can be calculated. With this information, Equ. 2 can be used to compute the lens' aperture size.

Focal Plane Moreover, using our method we can find the distance of the focal plane from the camera. As we discussed in the last section, the magnification scaling in Equ. 6 can be used to find the distance of the sensor from the lens which can be used by Equ. 4 to find the distance of
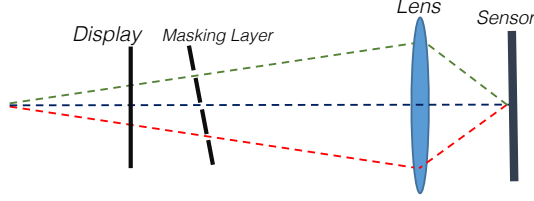
Figure 9. Extending the rays which intersect on the camera's sensor allows to estimate the focus distance.

the focal plane. However, because of the error in calculating b (size of the visible part of display through each hole) and b' (size of CoC) the final result is not accurate enough. For this purpose we found an alternative solution which is more robust.

As mentioned already, the image of each hole creates a disk known as CoC. Assume that images of these holes have overlap with each other as it is shown in Fig. 3. In this case multiple rays from different holes are intersecting at the same position on the camera's image plane (Fig. 9). By extending these light rays we can find their intersection point which is on the focal plane of the camera. For generating these light rays it is required to know the position of the holes and the pixels on the display. But it is not possible to directly look up the correspondences between the camera's image plane and the display in the overlapping area, since this overlapping effect had to be avoided beforehand when displaying the Gray code patterns because they could not be decoded uniquely within these overlapping areas. To solve this problem first it is assumed that the circle of confusion for the holes do not overlap with each other. By displaying the Gray code patterns (cf. Sec. 4.2) the position of the visible pixels through each hole on the image plane is detected. Instead of using only the center correspondence per hole, all generated correspondences between the camera and the display for each individual hole are used to estimate a homography between display and camera pixels for each specific hole. This homography can now be used to virtually extend the image of each individual hole. As a result, virtual overlaps between the images of the holes can be created. Since rotation and translation of the planes are already known during the calibration step, all the required light rays for focal plane estimation can be generated (cf. Table 2).

5. Hardware Protoype

The algorithm was evaluated using three hardware prototypes based on the design presented in Section 3. The three prototypes are shown in Fig. 8). The emissive background layer was realized by a 6", 2K (1440×2560 pixels) LCD IPS display controlled via HDMI[2]. The size of each pixel is approximately $0.051mm$. The blocking layers were realized using black paper sheets which were perforated with a

grid of small holes using a laser cutter device[3]. To ensure that they are planar, they were integrated into customized 3D printed[4] frames and placed $1 - 3cm$ above the display.

We also realized a color multiplexed setup in which three hole planes are stacked tilted on top of each other. Each plane consists out of one cyan, magenta or yellow color filter[5], such that the camera captures red, green and blue colored dots from different layers. This has the advantage that it is able to generate superimposed planes with a fixed, unchanged setup.

6. Evaluation and Discussion

The accuracy of the proposed calibration was evaluated by comparing it to the method of [31]. 30 images of the checker board were captured with different orientations and the toolbox implemented by Bouget [4] was used for calibration. The proposed method was tested with prototype (a) (cf. Fig. 8) using multiple lenses and cameras (cf. Table 1): A Canon 1100D with three different lenses ranging from 14mm up to 40mm, an action camera (GoPro Hero 2), and a smartphone camera (LG V10). The results of both methods were comparable in all setups. A 100mm DSLR lens (Canon EF 100mm f/2.8) was evaluated as well. For the latter, the resulting focal lengths were also comparable to the reference (fx 22252.54 vs. 21802.86, fy: 22262.68 vs. 21801.98), however, with such a combination the limits of the current prototype are reached since it is only able to generate a small number of correspondences because of the relatively wide hole spacing required due to the large CoC even with a minimum apertures size. To account for that, a larger back panel would be required to calibrate such lenses more reliably.

To estimate the accuracy of the aperture size calculation described in Sec. 4.3, the 24mm lens mounted on a Canon DSLR was used to compare the manually set aperture diameter to the estimated one for a range of f-stops. As shown in Fig. 10, all evaluated apertures diameters closely match the ground truth value. The same hardware was used to estimate the accuracy of the focal plane distance estimation. A comparison of the results with the measured ground truth values is given in Table 2). In the following we will explain how we obtained the ground truth data for the evaluation of the calculated aperture size and focal plane by our method.

Aperture Size First we calculate the focal length of the camera adequately accurate using a conventional checkerboard calibration or, if available, obtained the f-number from the camera and lens information. As a result we can use the following equation to find the aperture size of the camera: *aperture size = focal length / f-number*. The result

---

[2]Topfoison TF60010A

[3]Epilog HELIX 24x18, 60W
[4]Stratasys Connex 350, material: digital ABS
[5]Kodak Wratten Color Filters #44A, #32, #12

Table 1. Comparison of the calibration results using the prototypes of the proposed method (cf. Fig. 8) and the algorithm proposed by Zhang: f_x and f_y are focal length of the camera in horizontal and vertical direction. (c_x, c_y) is the coordinate of the principal point of the camera, and $k_1 ... k_4$ are the four estimated distortion parameters

| | Canon EF-S 24mm f/2.8 | | Canon EF 40mm f/2.8 | | Samyang 14mm f/2.8 | | LG V10 ~4.4mm f/1.8 | | GoPro Hero 2 Narrow | |
	Prototype (a)	Checker board	Prototype (c)	Checker board	Prototype (b)	Checker board	Prototype (b)	Checker board	Prototype (b)	Checker board
fx	4871.86	4766.92	7654.27	7521.66	2821.82	2839.23	4027.87	3996.92	1964.38	1903.76
fy	4875.97	4748.65	7654.69	7498.10	2822.46	2828.43	4026.39	3977.37	1980.31	1924.59
cx	2172.70	2161.98	2073.93	2147.74	2147.65	2144.77	2689.62	2626.58	975.03	945.47
cy	1396.46	1419.37	1393.93	1429.39	1411.67	1404.27	1480.87	1502.02	583.19	588.38
k_1	-0.16321	-0.12474	-0.05305	-0.03691	-0.13445	-0.11572	-0.00618	0.02290	-0.38539	-0.3713
k_2	0.06961	0.12403	-0.02518	-0.08091	0.04971	0.03280	0.07026	-0.06092	0.19380	0.2103
k_3	0.00137	-0.00016	0.00078	0.00157	-0.00038	-0.00043	-0.00052	-0.00102	0.00091	0.00085
k_4	-0.00114	0.00060	0.00106	0.00112	0.00174	-0.00132	0.00092	-0.00218	-0.00232	-0.0017

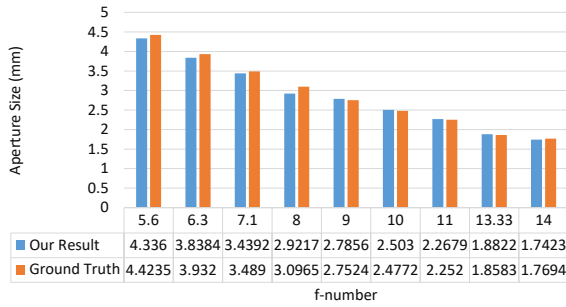|	5.6	6.3	7.1	8	9	10	11	13.33	14
Our Result	4.336	3.8384	3.4392	2.9217	2.7856	2.503	2.2679	1.8822	1.7423
Ground Truth	4.4235	3.932	3.489	3.0965	2.7524	2.4772	2.252	1.8583	1.7694

f-number

Figure 10. Estimated aperture sizes compared to ground truth values.

Table 2. Estimated distances of the camera's focal plane in comparison with the actual, i.e. measured, distance.

Proposed method (cm)	31.8	41.41	50.76	57.02	65.55
Ground truth (cm)	30	40	50	60	70

of this equation is compared to the calculated aperture size derived from our method.

Focal Plane First we put a checker board at an arbitrary distance from the camera and adjust the camera to focus onto this checker board and take a picture of it. We also take several other images of the checker board from different orientations allowing us to use the conventional checker board calibration method to calibrate the camera and also find the position of the planes with respect to the camera. Then we can find the distance of the plane from the camera in the first image. Then, without changing the focus of the camera, we use our calibration method to estimate the distance of the focal plane.

Reprojection Errors The reprojection errors of our method were also compared to Zhang's method. For this purpose we used the result of our method for reprojecting the corners of a checkerboard in 15 different images. The average errors are shown in Table 3.

7. Conclusion

In this paper we proposed a novel, fully automatic method for intrinsic camera calibration using a small device mounted directly in front of a lens. This greatly simplifies calibration, which otherwise requires expert knowledge

Table 3. Reprojection error comparison [pixels]

	proposed method	checkerboard method
24mm f/2.8	[0.85979 0.74322]	[0.40055 0.46826]
40mm f/2.8	[0.51382 0.58675]	[0.36577 0.49497]
14mm f/2.8	[0.36540 0.34291]	[0.21734 0.25747]
4.4mm f/1.8	[2.06272 1.42654]	[1.81404 1.40415]
GoPro Hero	[0.35966 0.37716]	[0.20676 0.22861]

and tedious and lengthy manual calibration procedures. We demonstrate accurate and repeatable calibrations of a variety of different configurations ranging from wide angle lenses to tele optics, which are on par with expert checkerboard calibration. In addition to the intrinsics that can be calibrated with previous methods, we can also estimate the focal plane as well as the size of the aperture since we are employing directionally encoded light for calibration.
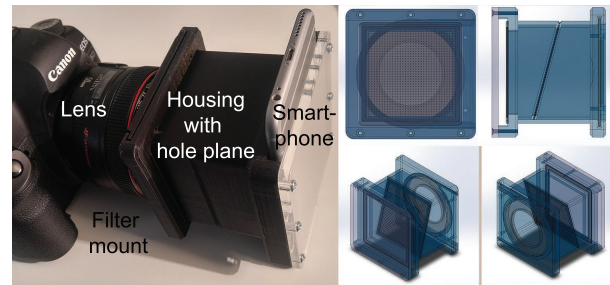

Figure 11. *Left*: 3D printed sample all-in-one prototype with lens mount and smartphone as display panel. *Right*: According CAD images.

Since the presented method does not require any specialized hardware except a high-resolution display which is a standard component of state-of-the-art smartphones, it can be realized by using the latter embedded into a frame holding the hole layers. We 3D printed such a sample calibration device with a smartphone as well as a lens mount using the filter thread as shown in Fig. 11. We envision that our proposed method could become a standard calibration tool.

In the future we are planning to investigate how further information can be extracted using this method such as lens aberration artifacts, the exact aperture shape, as well as spectral wavelength dependent issues.

References

[1] B. Atcheson, F. Heide, and W. Heidrich. Caltag: High precision fiducial markers for camera calibration. In R. Koch, A. Kolb, and C. Rezk-Salama, editors, *VMV*, pages 41–48. Eurographics Association, 2010. 2

[2] J. P. Barreto, J. Roquette, P. Sturm, and F. Fonseca. Automatic Camera Calibration Applied to Medical Endoscopy. *Procedings of the British Machine Vision Conference 2009*, (20091000069650):52.1–52.10, 2009. 3

[3] Y. Bok, H. Ha, and I. S. Kweon. Automated checkerboard detection and indexing using circular boundaries. *Pattern Recogn. Lett.*, 71(C):66–72, Feb. 2016. 2

[4] J. Y. Bouguet. Camera calibration toolbox for Matlab, 2008. 2, 7

[5] D. C. Brown. Decentering distortion and the definitive calibration of metric cameras. *Annual Meeting of the Americamn Society of Photogrammtric Engineering*, 1965. 2

[6] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971. 2

[7] O. Choi, H. Lim, and S. C. Ahn. Robust binarization of graycoded pattern images for smart projectors. In *2016 International Conference on Electronics, Information, and Communications (ICEIC)*, pages 1–4, Jan 2016. 5

[8] R. Cipolla, T. Drummond, and D. Robertson. Camera calibration from vanishing points in images of architectural scenes, 1999. 2

[9] A. E. Conrady. Decentered lens-systems. *Monthly Notices of the Royal Astronomical Society, Vol. 79, p.384-390*, 1919. 2

[10] A. Datta, J. Kim, and T. Kanade. Accurate camera calibration using iterative renement of control points. In *Workshop on Visual Surveillance (VS), 2009 (held in conjunction with ICCV).*, October 2009. 2

[11] M. Fiala and C. Shu. Self-identifying patterns for plane-based camera calibration. *Machine Vision and Applications*, 19(4):209–216, 2008. 2

[12] S. Garrido-Jurado, R. M. noz Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280 – 2292, 2014. 2

[13] H. Ha, Y. Bok, K. Joo, J. Jung, and I. S. Kweon. Accurate Camera Calibration Robust to Defocus Using a Smartphone. *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 828–836, 2015. 2

[14] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, 2003. Choix de documents en appendice. 6

[15] E. Hecht and A. Zajac. *Optics*. Addison-Wesley world student series. Addison-Wesley, 1987. 6

[16] R. Jacobson, S. Ray, G. Attridge, and N. Axford. *Manual of Photography*. Taylor & Francis, 2000. 3

[17] K. Kanatani, Y. Sugaya, and Y. Kanazawa. *Ellipse Fitting for Computer Vision: Implementation and Applications*. Synthesis Lectures on Computer Vision. Morgan & Claypool Publishers, 2016. 5

[18] V. A. V. Lixia Yang, Chao Tian and A. R. Reibman. An automatic grid corner extraction technique for camera calibration. *2012 19th IEEE International Conference on Image Processing*, pages 349 – 352, 2012. 2

[19] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs. Pinlight displays: Wide field of view augmented reality eyeglasses using defocused point light sources. *ACM Trans. Graph.*, 33(4):89:1–89:11, July 2014. 2

[20] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar. Bokode: Imperceptible visual tags for camera based interaction from a distance. *ACM Trans. Graph.*, 28(3):98:1–98:8, July 2009. 2, 6

[21] M. Mühlich and T. Aach. High accuracy feature detection for camera calibration: A multi-steerable approach. In F. A. Hamprecht, C. Schnörr, and B. Jähne, editors, *Pattern Recognition: 29th DAGM Symposium, Heidelberg, Germany, September 12-14, 2007. Proceedings*, pages 284–293, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. 2

[22] OpenCV. Open source computer vision library. https://github.com/itseez/opencv, 2015. 2

[23] Y. Oyamada, P. Fallavollita, and N. Navab. Single camera calibration using partially visible calibration objects based on random dots marker tracking algorithm. In *ISMAR*, Atlanta, USA, November 2012. 2

[24] A. Richardson, J. Strom, and E. Olson. AprilCal: Assisted and repeatable camera calibration. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, November 2013. 2

[25] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recogn.*, 43(8):2666–2680, Aug. 2010. 5

[26] I. Schillebeeckx, S. Louis, R. Pless, and S. Louis. Single Image Camera Calibration with Lenticular Arrays for Augmented Reality. 2016. 3

[27] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from Internet photo collections. *International Journal of Computer Vision*, 80(2):189–210, 2008. 1

[28] H. Urey, K. V. Chellappan, E. Erden, and P. Surman. State of the art in stereoscopic and autostereoscopic displays. *Proceedings of the IEEE*, 99(4):540–555, April 2011. 2

[29] Y. Xiao and B. Fisher. *Accurate Feature Extraction and Control Point Correction for Camera Calibration with a Mono-Plane Target*. 2010. 2

[30] L. Yang, J. M. Normand, and G. Moreau. Local geometric consensus: A general purpose point pattern-based tracking algorithm. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1299–1308, Nov 2015. 5

[31] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *ICCV*, pages 666–673, 1999. 1, 2, 4, 5, 6, 7