David Uhlig* and Michael Heizmann

Model-independent light field reconstruction using a generic camera calibration

Modellunabhängige Lichtfeld-Rekonstruktion mithilfe einer generischen Kamerakalibrierung

https://doi.org/10.1515/teme-2021-0020
Received February 15, 2021; accepted April 19, 2021

Abstract: Sophisticated and highly specialized optical measuring devices are becoming increasingly important for high-precision manufacturing and environment perception. In particular, light field cameras are experiencing an ever-increasing interest in research and industry as they enable a variety of new measurement methods. Unfortunately, due to their complex structure, their calibration is very difficult and usually precisely tailored to the particular type of light field camera. To overcome these difficulties, we present a method that decodes a light field from the raw data of any light field imaging system without knowing and modeling the internal optical elements. We calibrate the camera using a precise generic calibration method and transform the obtained ray set into an equivalent light field representation. Finally, we reconstruct a rectified light field from the irregularly sampled data and in addition we derive the geometric ray properties as intrinsic camera parameters. Experimental results validate the method by showing that both the information of the observed scene and the geometric structure of the light field are preserved by an adequate rectification and calibration.

Keywords: Light field, decoding, rectification, generic camera calibration.

Zusammenfassung: Anspruchsvolle und hochspezialisierte optische Messgeräte werden für die hochpräzise Fertigung und Umfelderkennung immer wichtiger. Insbesondere Lichtfeldkameras erfahren ein immer größeres Interesse in Forschung und Industrie, da sie eine Vielzahl von neuen Messmethoden ermöglichen. Leider ist ihre Kalibrierung aufgrund ihres komplexen Aufbaus sehr schwierig und meist genau auf den jeweiligen Lichtfeldkamera-Typ zugeschnitten. Um diese Schwierigkeiten zu überwinden, stellen wir eine Methode vor, die ein Lichtfeld aus den Rohdaten eines beliebigen Lichtfeldaufnahmesystem decodiert, ohne die internen optischen Elemente zu kennen und zu modellieren. Wir kalibrieren die Kamera mit einer präzisen generischen Kalibrierungsmethode und transformieren das erhaltene Strahlenset in eine äquivalente Lichtfelddarstellung. Schließlich rekonstruieren wir ein rektifiziertes Lichtfeld aus den unregelmäßig abgetasteten Daten und leiten darüber hinaus die geometrischen Eigenschaften der Strahlen als intrinsische Kameraparameter ab. Experimentelle Ergebnisse validieren die Methode, indem sie zeigen, dass sowohl die Informationen der beobachteten Szene als auch die geometrische Struktur des Lichtfeldes durch eine adäquate Rektifizierung und Kalibrierung erhalten bleiben.

Schlagwörter: Lichtfeld, Dekodierung, Rektifizierung, generische Kamerakalibrierung.

1 Introduction

In recent years, research on light fields and light field cameras (plenoptic cameras) has become more and more important. In contrast to traditional cameras, light field cameras are able to capture both angular and spatial information of the light rays that are propagated through space. They are thus able to obtain multiple views of the same scene in a single photographic image exposure, to estimate the depth of the scene or to shift the focus of the image after capturing the image [11]. These advantages have led to light field cameras becoming an important tool in image processing and optical metrology. As a result, a precise calibration of these cameras becomes increasingly important.

The first commercially available light field camera was presented by Ng [11]. He proposed a hand-held camera that consisted of an additional microlens array (MLA) that is placed in a small distance in front of the sensor, see Fig. 1. This array additionally allows to detect the directional dependencies of the rays and thus a light field can be extracted. Since the design of microlens-based cameras is not trivial, the light field has to be decoded from

***Corresponding author: David Uhlig,** Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstraße 16, 76187 Karlsruhe, Germany, e-mail: david.uhlig@kit.edu, ORCID: https://orcid.org/0000-0003-1996-4419
Michael Heizmann, Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstraße 16, 76187 Karlsruhe, Germany

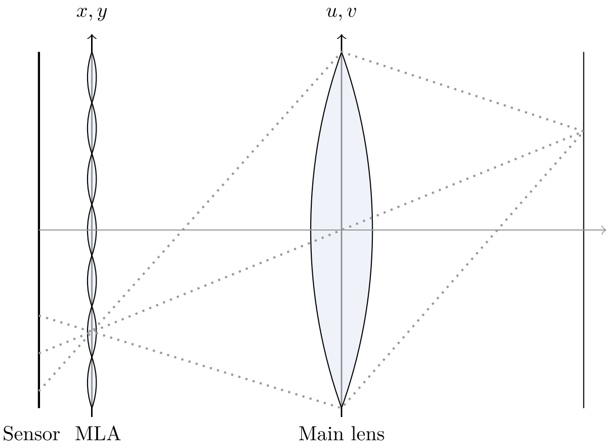


Figure 1: Schematic structure of a light field camera.

the raw sensor image using sophisticated algorithms. Furthermore, each lens (main lens and micro lens) is affected by the usual lens aberrations, *i. e.* a subsequent rectification of the light field is necessary to obtain correct geometric information relevant for image processing and applications in optical metrology. Dansereau *et al.* [4] presented a method that first extracts a light field from the raw sensor data by several reshaping and interpolation operations and then rectifies it by estimating the values of a 12-parameter camera model. Bok *et al.* [3], in contrast, presented a method that could extract the rectified light field directly from the raw sensor data by also using a low-dimensional camera model. In order to be able to extract any light field information from the raw data, both methods must initially detect the centers of the microlenses very precisely. But even with a subpixel accurate detection, the camera rays at the boundary of the microlenses are very difficult to model in both methods, and therefore these pixels are mostly discarded.

Another disadvantage of these methods is the model-based calibration in general. It can't describe highly local errors such as the strong distortions at the boundaries of the microlenses using only a low-dimensional model. As a consequence, in recent years, new camera models were proposed that describe the camera as a generic imaging system. They are able to model the ray of each pixel individually and thus allow for a high-precision calibration [6, 13]. However, the biggest disadvantage of the common light field reconstruction methods is that they are only applicable for a single type of camera, *e. g.* microlens-based light field cameras whose microlenses are exactly focused onto the sensor. To our knowledge, there is no single method yet that can reconstruct a light field from any type of light field camera.

In this work we present a method to reconstruct a light field that was captured by an arbitrary light field imaging system, without knowing the actually used configuration of optical elements inside the camera. We propose to use a generic camera calibration procedure to optimally calibrate each individual pixel of the camera, where all distortions of the optical elements are contained in the unconstrained bundle of sight rays, and thus are modeled very accurately. Further, we propose to use this bundle of rays to obtain an irregularly sampled presentation of the light field, we present a simple reconstruction method to interpolate a rectified light field from the irregularly spaced camera rays, and finally, we demonstrate how to calculate the intrinsic camera parameters. We use the presented method to calibrate and reconstruct light fields from two commercially available light field cameras which are based on different optical designs, a Lytro Illum and a Raytrix R5.

The paper is organized as follows: Section 2 provides the background about light fields and light field cameras. It gives an introduction to the concept of generic camera calibration and motivates our approach. Section 3.1 and 3.2 derive the 4D light field parameters from the unconstrained ray bundle obtained in the generic calibration. Section 3.3 describes the algorithm for the reconstruction of the light field from the rays' intensity values, Section 3.4 derives the intrinsic camera parameters, and finally, Section 4 experimentally validates the proposed method by analyzing real light field images. At last, Section 5 draws conclusions and presents directions for future work.

2 Background

2.1 Light field acquisition

The light propagating through space contains a variety of information. In a scene of interest, light can be described by the plenoptic function. Within the field of geometrical optics, this function can be described with seven variables: three spatial coordinates, two angular coordinates, one spectral value and time. A conventional camera, however, usually only captures a subspace of this function: two spatial coordinates with a color/intensity value and time in case of video cameras.

More information can be extracted when using light field acquisition devices. The 4D light field parameterizes a light ray with four coordinates (x, y, u, v) and to each ray a corresponding spectral value or color λ can be assigned, resulting in a 5D function $L(x, y, u, v, \lambda)$. However, for sake of simplicity, we omit the spectral value λ for the rest of this paper. In most cases, the two-plane parameterization is used, consisting of two parallel planes. The x, y-plane usually represents the spatial dependency of the light field and the u, v-plane symbolizes the angular properties. With this, each ray can be defined by the intersection points with these two planes.

The most commonly used designs for light field acquisition devices are microlens-based light field cameras [7]. Their layout is similar to that of a conventional camera with the essential difference that an array of microscopically sized lenses is placed in front of the sensor, see Fig. 1. By adding this microlens array it is possible to capture a section of the 4D light field $L(x, y, u, v)$ of a scene and to encode it onto the 2D sensor.

When the distance between the MLA and the sensor corresponds to the focal length of the microlenses, the camera is an *unfocused plenoptic camera* [11]. The coordinates of the light field's two-plane parameterization are represented here by the x, y- and u, v-coordinates, whereby x, y define the position of a microlens in front of the sensor and thus, they encode the spatial dimension of the light field. u, v describe the coordinates within the microlens relative to its center and in this way, they implicitly provide information on where a light ray has passed through the main lens. They represent the angular information of the light field. Each u, v coordinate therefore represents a virtual subcamera, which observes only a small part of the main lens, meaning that a light field camera can also be interpreted as a multi-camera array, whereby each subcamera, often referred to as subaperture image, has a slightly different view onto the scene, see Fig. 2. When the distance between the MLA and the sensor differs from the focal length of the microlenses, the camera is a *focused plenoptic camera* [8]. The relation between light field coordinates and the optical components of the camera is no longer as intuitive as it was before. Now, each microlens contains spatial information by imaging a microimage. Further, the position of a microlens contains angular and spatial information, due to a slightly different view of the scene in each microimage. For a comparison of the focused and unfocused plenoptic camera see Fig. 3.

The additional information compared to a standard camera allows to change the perspective of the scene after the exposure, which allows to extract depth information. Moreover, even after capturing a dynamically active scene, it becomes possible to shift the focus plane by rendering new images from the 4D light field data.

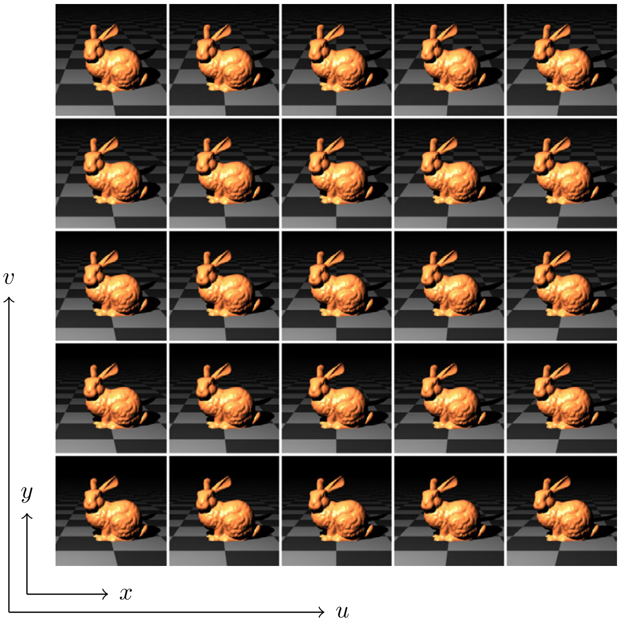


Figure 2: Interpretation of the light field camera as an array of subcameras which have slightly different views of the scene.

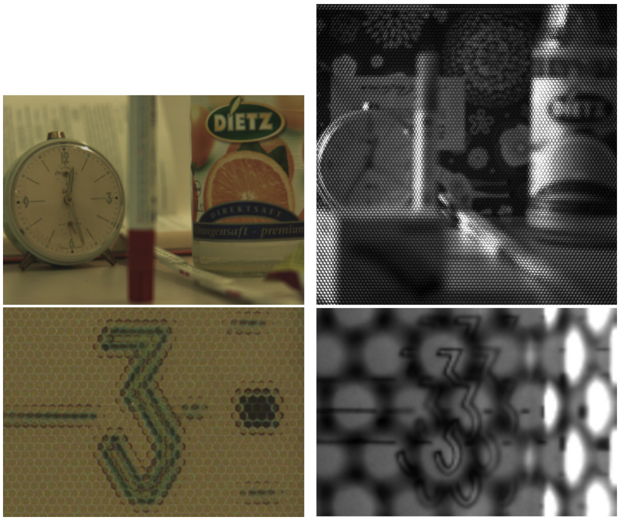


Figure 3: Camera raw data. Left: Unfocused plenoptic camera (Lytro Illum). The pixels under each microlens contain only angular information. Right: Focused plenoptic camera (Raytrix R5). The pixels under each microlens contain spatial information.

In particular, there are different configurations, *e. g.*, the distance of the microlens array to the sensor can be varied or microlenses with multiple focus lengths can be used [5, 8, 11]. Furthermore, there exist a variety of more exotic designs, *e. g.*, coded aperture based light field cameras, multispectral light field cameras, kaleidoscope-like configurations and of course camera arrays [7, 10, 16].

However, all have in common that decoding the light field from the sensor data and calibrating the camera is generally difficult. For example, to reconstruct the light field of microlens-based cameras, the centers of the microlenses, which are often arranged in a hexagonal grid, must be detected very accurately [9]. The 4D light field can then be extracted by shifting the pixels onto a rectangular grid and reshaping the 2D microlens images into a 4D array. The resulting light field, however, generally still contains all the distortions of the main lens and the microlenses, which is why an additional rectification is necessary [3, 4].

2.2 Generic camera calibration

A camera maps the three-dimensional world onto a two-dimensional image. The calibration of a camera generally refers to the determination of the parameters of this mapping operation. A good calibration procedure is based on a careful modeling of the optical elements inside the camera, which of course is strongly influenced by the camera type. Typically, low-dimensional models are used to model the entire camera. Standard cameras, for example, often use the model of Zhang [18], which represents the camera as a pinhole model with additional distortion parameters. However, these simple models have the disadvantage that they have only a limited descriptive capability. Thus, for modern cameras, or rather for more advanced optical systems, not all pixels can be perfectly described by these few model parameters. As the complexity of an optical system increases, it becomes more and more difficult to model it with a low-dimensional representation. Hence, the lack of flexibility and precision has led to the development of new camera models. Cameras are now described as generic imaging systems, which are independent of the specific camera type and allow high-precision calibration [2, 6, 13]. An imaging system is modeled as a set of photosensitive pixels, with all other optical elements not explicitly modeled but represented by a black box. Each pixel collects light from a bundle of rays, referred to as *raxel*, entering the imaging system. The set of all *raxels* with their associated geometric parameters then forms the complete generic imaging model.

The geometric parameters can be described for each pixel i by a single camera ray running through the center of the *raxel* along the direction of light propagation. There are a multitude of options for describing rays. One mathematically easy-to-handle choice are, for example, *Plücker*-line coordinates [1, 14]. Here, a ray $\mathbf{l}_i = (\mathbf{d}_i^{\mathrm{T}}, \mathbf{m}_i^{\mathrm{T}})^{\mathrm{T}}$ consists of a direction vector $\mathbf{d}_i$ and a moment vector $\mathbf{m}_i$. And due to rays having four degrees of freedom in 3D-space, two constraints must be considered: $\mathbf{d}_i^{\mathrm{T}}\mathbf{m}_i = 0\,,\ \|\mathbf{d}_i\| = 1\,.$

The calibration of the entire bundle of rays belonging to the camera is then usually realized with the help of a minimization of the ray projection error ε_i. Meaning, the sum of the Euclidean distances of the rays $\mathbf{l}_i$ to known reference points $\mathbf{p}_{ik}$ in space is minimized [13]:

$$\arg\min_{\mathbf{l}_i} \sum_i \varepsilon_i^2 = \arg\min_{\mathbf{l}_i} \sum_{i,k} d_{\mathrm{e}}(\mathbf{l}_i, \mathbf{p}_{ik})^2\,, \qquad (1)$$

with $d_{\mathrm{e}}(\mathbf{l}_i, \mathbf{p}_{ik}) = \|\mathbf{p}_{ik} \times \mathbf{d}_i - \mathbf{m}_i\|$. A minimization of the commonly used ray reprojection error on pixel level is often not possible, because most generic models do not support a direct projection of points in space onto the pixel plane. See [13] for details.

The advantage of this type of calibration is that there is no longer one global model that has to describe the camera over the entire pixel plane. Instead, with the generic model even high-frequency distortions in the optical imaging system can be modeled equally accurate both locally and globally, resulting in a highly accurately calibrated camera. This is specifically important for light field cameras where it becomes very difficult to model distortions of the microlenses with a global model. In the end, however, one does not obtain an image but rather a set of rays with corresponding intensities. This does not interfere with many applications in optical metrology, *e. g.*, profilometry or deflectometry, where only the geometric ray properties are relevant [15, 17]. Yet it may complicate other tasks since the spatial correlations between pixels and their associated rays are lost. The classical image processing algorithms cannot be applied without further effort. In the specific case of the light field camera, algorithms such as a subsequent refocusing of the image or a simple depth estimation can no longer be performed using standard methods. Our approach and proposal is therefore to use the generic camera model to obtain all the geometric ray parameters of an arbitrary light field camera as accurately as possible. Subsequently, we use this information to reconstruct the light field from the set of all rays. And consequently, with that we obtain a generic algorithm to extract a light field from an arbitrary optical imaging system, neglecting the actual design of the used light field acquisition device.

3 Light field reconstruction

3.1 Normalizing the ray bundle

In order to decode a light field from the raw sensor data, the camera must first be calibrated by using, *e. g.*, a generic calibration method as described in the previous section. As a consequence, all the preprocessing steps of the classical light field calibration are not needed at all. Therefore, it is not necessary to detect the centers of the microlenses and no hexagonal sampling of the MLA has to be compensated. However, due to the black box character of the generic calibration, it is initially not possible to define a consistent camera coordinate system for every calibrated camera. Even when using the same calibration algorithm for the same camera, the outcome could vary. Hence, the result of a generic calibration is (in many cases) not unique, *i. e.* the calibrated camera rays are represented in an arbitrary coordinate system, which usually depends on the starting configuration of the generic calibration procedure or on the used calibration reference target. Therefore, to transform this arbitrary coordinate system into one that is fixed to the individual camera, a few steps are necessary.

First, we define the origin of this coordinate system to be the optical center of the camera, and for a light field camera this corresponds approximately to the center of the exit pupil. The location of it can be understood as the point $\mathbf{p}_\mathrm{o}$ that has the smallest distance to all rays, *i. e.* it can be calculated by minimizing the weighted mean of the Euclidean distances to all rays:

$$\mathbf{p}_\mathrm{o} = \arg\min_{\mathbf{p}} \sum_i w_i \,\|\mathbf{p} \times \mathbf{d}_i - \mathbf{m}_i\|^2 \tag{2}$$

$$= \left(\sum_i w_i [\mathbf{d}_i]_\times [\mathbf{d}_i]_\times^T \right)^{-1} \sum_i w_i [\mathbf{d}_i]_\times \mathbf{m}_i \,, \tag{3}$$

where $[\mathbf{a}]_\times$ is the matrix equivalent of the cross product with $[\mathbf{a}]_\times \mathbf{b} = \mathbf{a} \times \mathbf{b}$. The weighting factor can be chosen to suppress poorly calibrated rays and to remove outliers. For instance, a simple choice is to use the inverse of the ray projection error $w_i = 1/\varepsilon_i$ that is calculated during the generic calibration procedure.

As a next step, we define the z-axis of our camera-fixed coordinate system, *i. e.* the view axis, as the average ray direction which can be found by solving the constrained optimization problem

$$\mathbf{d}_\mathrm{z} = \arg\max_{\mathbf{d}} \sum_i w_i \,\langle \mathbf{d}, \mathbf{d}_i \rangle^2 \;, \text{ s.t. } \|\mathbf{d}\| = 1\,. \tag{4}$$

Using the Lagrange multiplier formalism and solving for $\mathbf{d}$ produces an eigenvalue problem:

$$\mathbf{d}_\mathrm{z} = \arg\max_{\mathbf{d}} \sum_i w_i \,\langle \mathbf{d}, \mathbf{d}_i \rangle^2 - \mu \left(\mathbf{d}^\mathrm{T} \mathbf{d} \right) \,, \tag{5}$$

$$\Rightarrow \left(\sum_i w_i \mathbf{d}_i \mathbf{d}_i^\mathrm{T} \right) \mathbf{d}_\mathrm{z} = \mu \mathbf{d}_\mathrm{z} \,, \tag{6}$$

where the eigenvector $\mathbf{d}_\mathrm{z}$ with largest absolute eigenvalue μ results in the average ray direction. A corresponding rotation matrix $\mathbf{R}_{\angle \mathbf{d}_\mathrm{z}}$, that rotates the bundle of rays into the new z-direction, can then directly be calculated.

The last remaining degree of freedom is the rotation around this new z-axis. Since light field cameras project the light onto a rectangular sensor, we wish to align the coordinate system's x- and y-axis with the corresponding sensor's x'- and y'-axis, respectively. Furthermore, due to the almost perspective projection, the change of ray direction with respect to the x- and y-axis should correspond to the change with respect to the x'- and y'-axis. Thus, using $\mathbf{d}_i = (d_{x,i}, d_{y,i}, d_{z,i})^\mathrm{T}$, the rotation angle that aligns both coordinate systems can be found by calculating the mean image gradients with respect to $\mathbf{x}' = (x', y')^\mathrm{T}$:

$$\begin{pmatrix} d_{xx'} \\ d_{xy'} \end{pmatrix} = \frac{\sum_i w_i \nabla_{\mathbf{x}'} d_{x,i}}{\sum_i w_i} \,, \tag{7}$$

$$\begin{pmatrix} d_{yx'} \\ d_{yy'} \end{pmatrix} = \frac{\sum_i w_i \nabla_{\mathbf{x}'} d_{y,i}}{\sum_i w_i} \,. \tag{8}$$

By estimating the orientation angle of the gradients with respect to the sensor axes, a rotation matrix $\mathbf{R}_\alpha$ can be found that rotates the coordinate system around the z-axis by an angle α:

$$\alpha_x = \operatorname{arctan2}\left(d_{xx'}, d_{xy'}\right) \,, \tag{9}$$

$$\alpha_y = \operatorname{arctan2}\left(d_{yx'}, d_{yy'}\right) + \frac{\pi}{2} \,, \tag{10}$$

$$\alpha = \operatorname{arctan2}(\sin\alpha_x + \sin\alpha_y, \cos\alpha_x + \cos\alpha_y) \,. \tag{11}$$

As final action, we transform the *Plücker*-ray parameters into light field coordinates. For this, we first transform the rays into the camera-fixed coordinate system, by shifting the origin and appropriately rotating the axes, and then calculate the intersections of the rays with the two-plane representation of the light field. The u, v-plane is placed orthogonal to the z-axis into the origin of the coordinate system. The x, y-axis is placed parallel to this at an arbitrary distance f, see Fig. 4. And thus, each ray $\mathbf{l}_i = (\mathbf{d}_i, \mathbf{m}_i)^\mathrm{T}$ can be described by four light field coordinates $\bar{x}_i, \bar{y}_i, \bar{u}_i, \bar{v}_i$:

$$\lambda\,(\bar{x}_i, \bar{y}_i, \bar{u}_i, \bar{v}_i, 1)^\mathrm{T} = \mathbf{P}\mathbf{T}\,\mathbf{l}_i \,, \tag{12}$$

with the coordinate transformation matrix $\mathbf{T}$ and the projection operator $\mathbf{P}$ (see [1, 12] for details[1]):

$$\mathbf{T} = \begin{pmatrix} \mathbf{R}_\alpha \mathbf{R}_{\angle \mathbf{d}_\mathrm{z}} & \mathbf{0} \\ \mathbf{R}_\alpha \mathbf{R}_{\angle \mathbf{d}_\mathrm{z}} [-\mathbf{p}_\mathrm{o}]_\times & \mathbf{R}_\alpha \mathbf{R}_{\angle \mathbf{d}_\mathrm{z}} \end{pmatrix} \,, \tag{13}$$

$$\mathbf{P} = \begin{pmatrix} f & 0 & 0 & 0 & -1 & 0 \\ 0 & f & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \,. \tag{14}$$

[1] The projection operator **P** of Johannsen *et al.* [12] may have a typo because it is slightly different from ours.

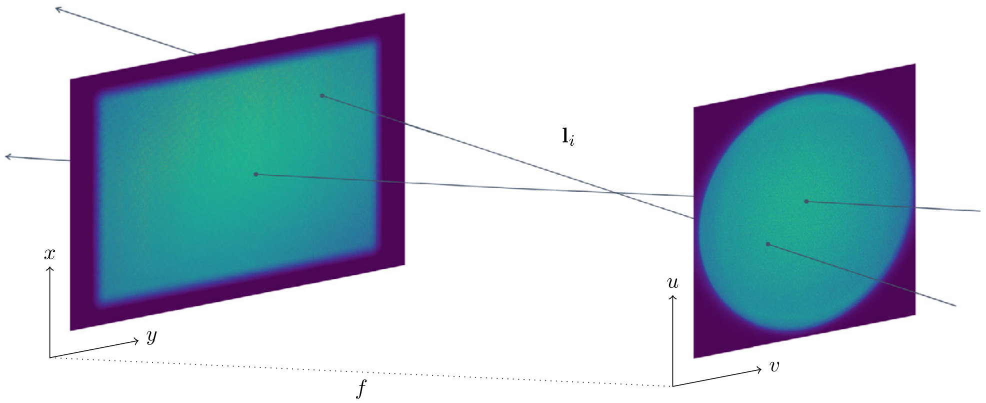


Figure 4: Two-plane parameterization of the light field. The ray $\mathbf{l}_i$ intersects the u, v- and the x, y-plane in (u_i, v_i, x_i, y_i). The intensities in the planes visualize the spatial distribution of the intersection points as a 2D histogram for the Lytro Illum data set. The u, v-plane lies in the plane of the camera's main lens. The x, y-plane corresponds to a projection of the rectangular sensor into space.

3.2 Regular light field grid

To reconstruct a light field from the bundle of rays belonging to the camera, the calibrated ray coordinates must first be transformed to a standardized grid. Afterwards, the observed ray intensities can be interpolated to a discretized light field, which we parameterize in the same two-plane representation as before. The complete set of real camera rays described as a set of 4D-points is arranged in an irregular 4D-grid. Still, the classical light field algorithms require a regular grid with uniform spacing. Therefore, this irregular grid of continuous rays has to be interpolated to a discrete light field described by a regular grid.

Hence, we need to define a regular grid with integer grid points $(x, y, u, v) \in [0, N_x - 1] \times [0, N_y - 1] \times [0, N_u - 1] \times [0, N_v - 1]$ with a fixed number of samples N_x, N_y, N_u, N_v in the respective dimensions. After the discrete target light field has been defined, we need to transform the set of real camera rays for which we need to estimate the parameter space of the actual ray geometry. First, the domains of the real light field dimensions are determined by analyzing the 2D histogram of the intersection points of rays with both planes of the light field representation, see Fig. 4. In order to place the regular grid structure into the irregular data, we define the grid extension by using a threshold value on the histogram data. A threshold of, *e. g.*, 10 % ensures that most of the camera's rays are within the range defined by the grid.

Since the real light field parameters are specified in physical units, *e.g.* mm, they have to be transformed to the previously defined discrete 4D-pixel grid by shifting the minimal value x_o, y_o, u_o, v_o, by normalizing the width of the histogram $\Delta x, \Delta y, \Delta u, \Delta v$ and by considering the sampling rate. The normalized coordinates are defined by:

$$\begin{aligned} x_i &= \frac{N_x - 1}{\Delta x}(\bar{x}_i - x_o)\,, \quad & y_i &= \frac{N_y - 1}{\Delta y}(\bar{y}_i - y_o)\,, \\ u_i &= \frac{N_u - 1}{\Delta u}(\bar{u}_i - u_o)\,, \quad & v_i &= \frac{N_v - 1}{\Delta v}(\bar{v}_i - v_o)\,. \end{aligned} \tag{15}$$

This still results in irregularly spaced data, which however can now be interpolated more easily to the desired regularly sampled light field.

The number of 4D cubes in each direction and the length of their edges could in principle be defined arbitrarily but it is advisable to incorporate knowledge about the physical camera. For example, our light field camera (Lytro Illum) has microlenses with a radius of about 7 pixels. Thus, this sampling can be used directly as a basis for the discretization of the u, v-plane, where $N_u = N_v$. The sampling of the x, y-plane can be determined in the same way by, *e. g.*, the number of microlenses in front of the sensor, whereby it is advisable to choose $N_y = \frac{\Delta y}{\Delta x} N_x$ to obtain approximately square spatial pixels.

3.3 Reconstructing radiometric properties

After the parameters of the light field have been defined, each corresponding light field pixel can be determined for every ray by finding the discrete grid point that is closest to the ray's light field representation. Since the rays and the grid are normalized to the same scale, the set of rays $\mathcal{N}_{x,y,u,v}$ that has an effect on a pixel (x, y, u, v) can easily be found by a fast and simple rounding operation to the closest integer $[\,\cdot\,]$. As a result, each light field pixel is only influenced by rays that lie in the corresponding 4D cube:

$$\mathcal{N}_{x,y,u,v} := \left\{ i \, : \, \frac{m}{2} \geq \left\| \begin{pmatrix} x \\ y \\ u \\ v \end{pmatrix} - \begin{pmatrix} [x_i] \\ [y_i] \\ [u_i] \\ [v_i] \end{pmatrix} \right\|_\infty \right\} , \tag{16}$$

where each individual ray is assigned to a nearest pixel by using $m = 1$. To allow a ray to influence more than the nearest pixel, higher order neighbors can be utilized with $m > 1\,, m \in \mathbb{N}^+$. The intensity of a discrete pixel can then be calculated from the intensity values of the corresponding rays as a weighted average:

$$L(x, y, u, v) = \frac{\sum_{i \in \mathcal{N}_{x,y,u,v}} w_i \, L(x_i, y_i, u_i, v_i)}{\sum_{i \in \mathcal{N}_{x,y,u,v}} w_i} . \tag{17}$$

For the weighting factor we calculate the distance between the ray's light field parameters and its correspondence in the grid. In order to consider larger deviations less, the error is squared and exponentially weighted.

$$w_i = \frac{1}{\varepsilon_i} \exp \left(- \left\| (x, y, u, v)^\mathrm{T} - (x_i, y_i, u_i, v_i)^\mathrm{T} \right\|_2^2 \right) . \tag{18}$$

An additional weighing of the different light field coordinates is not required, since these have already been brought to a unified basis by the normalization of Section 3.2. To additionally benefit from the results of the generic camera calibration, an error measure ε_i is taken into account, *e. g.* the pixelwise ray projection error [13]. This suppresses badly calibrated camera rays, which often do not have good optical properties, *e. g.* dead pixels or pixels at the edges of microlenses, which can be strongly distorted.

3.4 Reconstructing geometric properties

Apart from the radiometric reconstruction of the light field, the geometric ray properties are relevant for many applications. For optical metrology, 3D reconstruction, or other areas of computer vision, a mapping is needed to transform pixel coordinates into world coordinates, and vice versa, to project points from world coordinates onto the pixel plane. Unlike the classic camera model, where each world point is mapped to only a 2D pixel pair, the same point can be mapped to more than one 4D light field pixel. Illustratively, this can be understood by the observation that a light field camera can also be interpreted as an array of individual virtual subcameras, where an observed point is mapped to a 2D pixel pair in each individual camera's virtual sensor plane. We therefore need for every angular coordinate a projection equation from world points to spatial pixels. The intrinsic camera parameters required for this are described for each angular coordinate by a projection matrix (comparable to the standard camera model [18]):

$$\mathbf{K}(u, v) = \begin{pmatrix} f_x & 0 & c_x(u) \\ 0 & f_y & c_y(v) \\ 0 & 0 & 1 \end{pmatrix} . \tag{19}$$

The parameters of this matrix can directly be determined from the two-plane parameterization of the light field:

$$f_x = f \frac{N_x - 1}{\Delta x} , \tag{20}$$

$$f_y = f \frac{N_y - 1}{\Delta y} , \tag{21}$$

$$c_x(u) = u \frac{\Delta u}{\Delta x} \frac{N_x - 1}{N_u - 1} + (N_x - 1) \frac{u_0 + x_0}{\Delta x} , \tag{22}$$

$$c_y(v) = v \frac{\Delta v}{\Delta y} \frac{N_y - 1}{N_v - 1} + (N_y - 1) \frac{v_0 + y_0}{\Delta y} . \tag{23}$$

Since the optical centers of the individual subcameras are slightly displaced with respect to each other in the u, v-plane, a corresponding translation vector is required to represent the relative offset with respect to the central subcamera:

$$\mathbf{t}(u, v) = \begin{pmatrix} t_x(u) \\ t_y(v) \\ 0 \end{pmatrix} = \begin{pmatrix} u \frac{\Delta u}{N_u - 1} + u_0 \\ v \frac{\Delta v}{N_v - 1} + v_0 \\ 0 \end{pmatrix} . \tag{24}$$

The forward projection of a point $\mathbf{p}$ (measured in the coordinate system fixed to the central subcamera) onto the light field pixel (x, y, u, v) can therefore be found by:

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{K}(u, v) \, (\mathbf{p} + \mathbf{t}(u, v)) . \tag{25}$$

The backward projection of light field pixels (x, y, u, v) to points $\mathbf{p}(\lambda)$ along the associated ray is given by:

$$\mathbf{p}(\lambda) = \lambda \, \mathbf{K}(u, v)^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} - \mathbf{t}(u, v) . \tag{26}$$

3.5 Spherical parameterization of angular coordinates

As can be seen in Fig. 4, the parameterization of the u, v-plane using Cartesian coordinates is not always ideal. If the grid is defined to enclose the entire circle, then the light field is reconstructed for areas where no rays pass

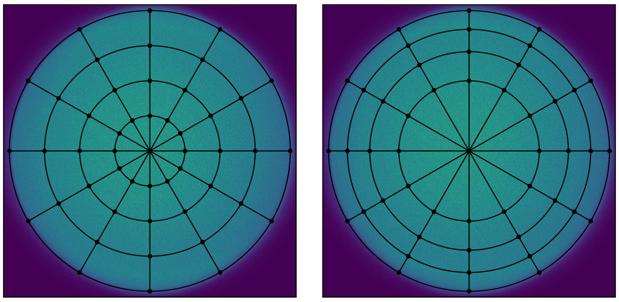


Figure 5: Differences in the sampling pattern of spherical coordinates. Left: Equidistant radius spacing, using $r_n = \text{const}$. Right: Equal pixel area, using $r_n = \sqrt{n}r_1$.

through the u, v-plane. If the grid is placed inside the circle, sufficient rays will pass through each light field pixel. However, information is discarded at the edges. Hence, it would be better to directly use a spherical parameterization of the angular coordinates, which would allow the entire information to be captured without sampling unnecessary areas.

We therefore define the angular coordinates by spherical pixels r and ϕ. And to obtain a resolution comparable to the Cartesian sampling, the number of samples is chosen to be $N_r = N_\phi = N_u$ and the coordinates of the samples are in the domain $r \in \left[-\frac{N_r-1}{2}, \frac{N_r-1}{2}\right]$ and $\phi \in \left[0, \frac{N_\phi-1}{N_\phi}\pi\right]$.

While the advantage of a spherical parametrization is a more efficient sampling, there are also disadvantages. When sampling the angle and the radius in equidistant steps, the effective pixel size grows with increasing radius, see Fig. 5. As a result, fewer rays pass through smaller pixels, which would result in a lower signal-to-noise ratio for these pixels. A possible workaround here is to use more neighbors for the reconstruction of the pixels and to interpolate missing information, which could be achieved with $m > 1$ in equation (16). Another possibility is to define the radius in such a way that each pixel has the same area. This can easily be achieved by transforming the domain of the radius coordinates with $r_n \sim \sqrt{n}$.

The reconstruction of the light field using spherical coordinates is then performed in the same way as presented in the sections above. The only distinction is the different sampling grid in the angular plane, which needs a transformation of the Cartesian coordinates u_i, v_i into spherical coordinates r_i, φ_i in eq. (16), with:

$$r_i = \operatorname{sign}(v_i)\sqrt{u_i^2 + v_i^2}\,, \tag{27}$$

$$\phi_i = \operatorname{arctan2}(v_i - \frac{N_v - 1}{2}, u_i - \frac{N_u - 1}{2}) \mod \pi\,, \tag{28}$$

and a reverse transformation from r, φ back to u, v in eq. (18), with

$$u = r\cos\phi + \frac{N_u - 1}{2}\,, \tag{29}$$

$$v = r\sin\phi + \frac{N_v - 1}{2}\,. \tag{30}$$

The calculation of the camera parameters is equivalent to the one already described in Section 3.4 with merely the u, v variables having to be replaced by the corresponding transformations of r and φ.

4 Results

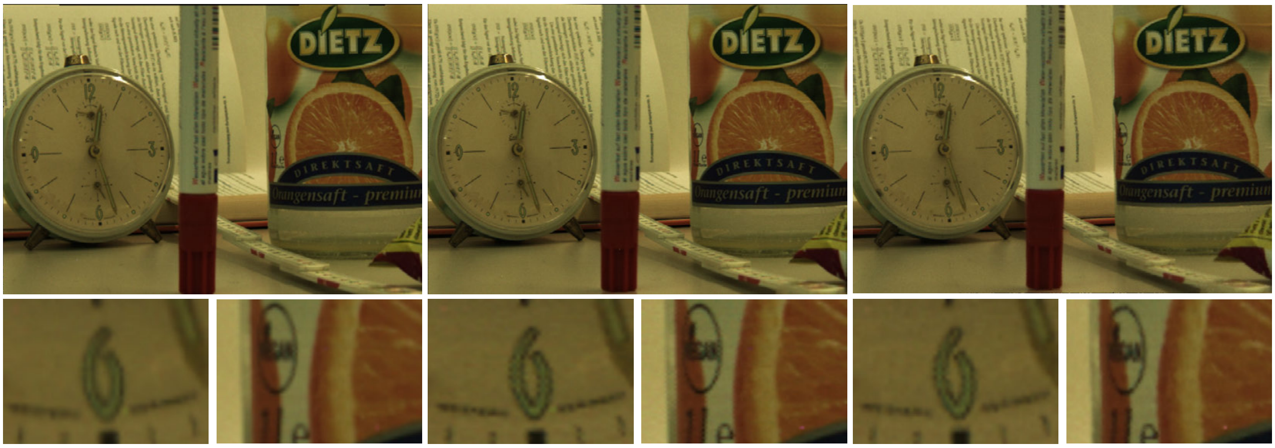
In order to be able to evaluate the method, we recorded light field data experimentally, see Fig. 3. We used a Lytro Illum and a monochromatic Raytrix R5. Both are microlens-based light field cameras. The former is an unfocused plenoptic camera [11] and the latter is a focused plenoptic camera [8], see Section 2.1. Therefore, both cameras are based on a different camera model and for a classical camera calibration both would need a different calibration procedure. However, a generic calibration is independent of the camera. To achieve this independence, the ray geometry of the sight rays of the cameras were estimated using a generic camera calibration [13]. Subsequently, a test scene was captured to be used as a basis for comparison of the proposed light field reconstruction.

To allow a meaningful discussion of the proposed light field reconstruction relative to other methods in the literature, we evaluate the Lytro Illum data by choosing the resolution of the light field grid to be $(N_x, N_y, N_u, N_v) = (625, 434, 15, 15)$, which can be found by following the remark at the end of Section 3.2. The resolution for the reconstruction of the Raytrix R5 was set to $(N_x, N_y, N_u, N_v) = (1000, 1000, 5, 5)$. Since the reconstruction of each pixel can be done independently from all others, it is convenient to parallize the equations (16), (17), (18) using a GPU. The reconstruction of the complete light field then takes only a few seconds (Intel Core i7-6700, Nvidia GTX 1080 Ti, 16 GB RAM).

4.1 Evaluation of the reconstruction

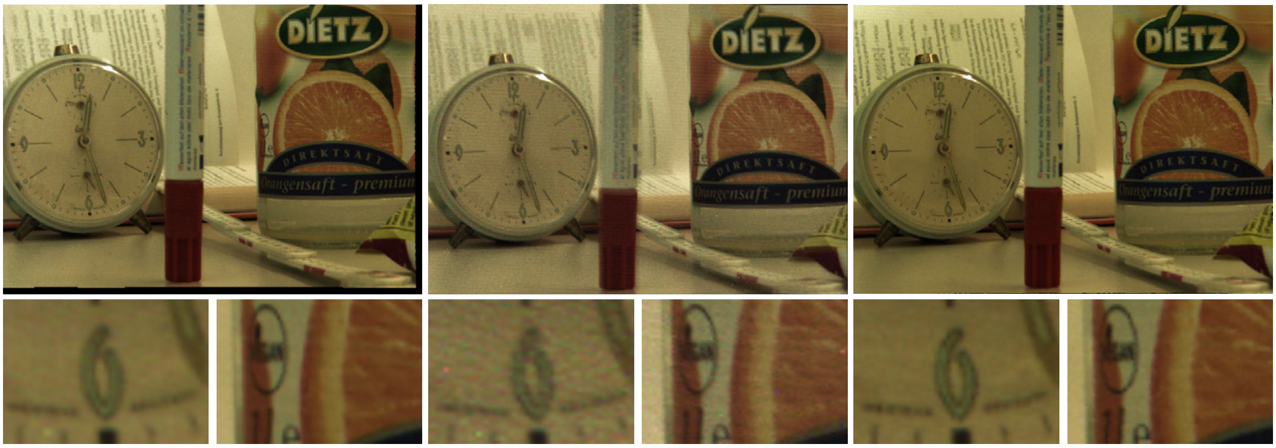
4.1.1 Unfocused plenoptic camera

For a comparison of the proposed method to the state-of-the-art, we also evaluated the light field reconstruction methods of Dansereau *et al.* [4] and Bok *et al.* [3]. Both



(a) Bok *et al.* **(b)** Dansereau *et al.* **(c)** Proposed method

Figure 6: Top: Subaperture images from the center of the u, v-plane (Lytro Illum). Bottom: Zoomed in.



(a) Bok *et al.* **(b)** Dansereau *et al.* **(c)** Proposed method

Figure 7: Top: Subaperture images from the edge of the u, v-plane (Lytro Illum). Bottom: Zoomed in.

methods only work with unfocused plenoptic cameras and can thus only be tested on the Lytro Illum data set.

The reconstruction of the central subaperture image of the test scene captured with the Lytro Illum is shown in Fig. 6. Here, only rays from the center of the u, v-plane were used in the reconstruction. It can be seen that the proposed method can reconstruct the scene correctly, although there were absolutely no presumptions about the internal optical structure of the camera and no information of the connection between rays and pixels on the sensor was used. The reconstruction results of Dansereau *et al.* and the proposed method are relatively similar and show a sharper result compared to the method of Bok *et al.* In detail it can be seen that the proposed method can reconstruct the light field even near object edges very well. The minimally blurrier appearance compared to Dansereau *et al.* is due to the relatively freely chosen sampling of the light field. A better optimized choice of the light field dimensions should result in less rays being summed up, thus reducing the blur. However, moving away from the center and looking at the subaperture images at the edge, one sees that the quality of the images for Dansereau *et al.* decreases significantly, while the proposed method and Bok *et al.*'s method only become slightly blurrier, see Fig. 7. The image of Bok *et al.* shows black borders, *i. e.* invalidated pixels, at the bottom and on the right. The proposed method, shows a similar effect, which, depending on how tight the dimension of the x, y-plane is chosen using the histogram in Section 3.2, could also be stronger.

Regardless of the reconstruction of the subaperture images, the advantage of the proposed method becomes apparent in another area. Apart from the central view that only incorporates spatial information, the light field contains much more, *i. e.* angular information. If one fixes an angular and a spatial coordinate in the 4D light field pointing in the same direction, *e. g.* u and x, one gets a 2D-slice of the light field, a so-called epipolar plane image (EPI) [11]. Lines of different slopes can be seen, whose orientation represents the depth of the observed object point [16]. Depth estimation in light fields is thus reduced to a simple local orientation estimation in these EPIs, whereby the quality of the estimation is significantly influenced by the calibration. The better the quality of the lines, the better the result of the depth estimation. Fig. 8 shows examples of horizontal and vertical EPIs generated by fixing u or v to its center coordinates and by selecting pixel lines for the x (red) or y (green) coordinate, respectively. The EPI of Dansereau *et al.* shows strong deviations from the epipolar geometry, visible through the curvy epipolar lines. This is caused by the poor generalizability of the method which was developed for the old Lytro camera and works only moderately well for the newer Lytro Illum. The EPI of Bok *et al.* on the other hand is much straighter. However, there are errors at the top and the bottom. These areas correspond to pixels which are located at the boundary of the microlenses, where the imaging is more strongly distorted. For the proposed method, it can be seen that the epipolar geometry is maintained much better, visualized by the straight lines in the EPIs. Also, the distortions of the lenses are compensated, resulting in a rectified light field.

Another advantage of the proposed method is the free choice of sampling. Therefore, a more suitable sampling grid can be used. The spherical sampling of the u, v-plane presented in Section 3.5 is better adapted to the data of the Lytro Illum light field camera and can therefore better represent the light field. No unnecessary information is sampled and the result is more compact, or rather, more information is contained in the same amount of data. In contrast, if the u, v-grid in Fig. 5 encloses the shown circle, Cartesian sampling would provide no data for the peripheral images, because there are simply no rays that could measure any information. If the grid is placed inside the circle, however, rays located at the edges of the circle would be discarded. With the same resolution and thus the same size of the reconstructed light field, spherical sampling effectively removes less information while representing the important information more accurately than Cartesian sampling. Fig. 9 shows the comparison, whereby the light field is illustrated as an array of subaperture images.

However, in detail it is important how the spherical sampling is implemented. As already described in Section 3.5, two options for the choice of radial sampling are considered. For the first choice the radius is set in equidistant steps. This has the advantage that all subaperture images for $\phi = 0$ and for $\phi = \frac{\pi}{2}$ correspond to the result of the Cartesian sampling. Further the corresponding

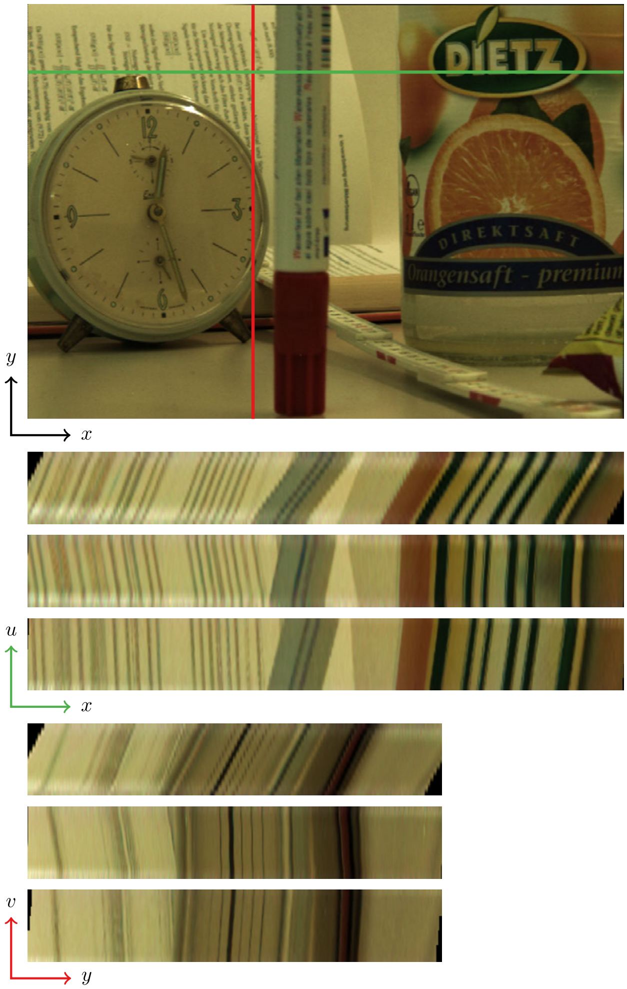


Figure 8: EPIs in comparison: Center image (top). Horizontal EPI (middle) and vertical EPI (bottom), respectively from top to bottom in the order Bok *et al.*, Dansereau *et al.*, proposed method.

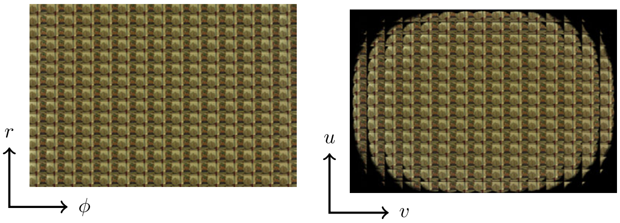


Figure 9: The light field as an array of subaperture images. Left: Spherical sampling results in a more efficient representation of the data. Right: Cartesian sampling reconstructs unnecessary peripheral areas of the circle.

EPIs in these directions also correspond to the EPIs of the Cartesian reconstruction, *i. e.* $\mathrm{EPI}(r,x)_{|\phi=0} = \mathrm{EPI}(u,x)$ and $\mathrm{EPI}(r,y)_{|\phi=\frac{\pi}{2}} = \mathrm{EPI}(v,y)$. The disadvantage is that the effective pixel size of the r,ϕ-plane is no longer the same for each pixel. This means that for small r there might be pixels that are not hit by any ray (u_i, v_i, x_i, y_i). However, those pixels can easily be interpolated from the neighboring rays by selecting $m > 1$ in eq. (16).

Another way to solve this problem is to directly define the pixel area of all pixels of the sampling grid to an equal size, as described in Section 3.5. This has the advantage that the signal-to-noise ratio remains the same for each pixel. Still, a minor disadvantage becomes apparent when analyzing the EPIs. Since the radius now scales with a square root, the lines in the EPIs are no longer straight but curved. The classical light field depth estimation, which analyzes the slope of the lines, can therefore no longer be applied here without further consideration, as it would provide incorrect results or would necessitate corresponding corrections, *e. g.* a local rescaling of the estimated slope of the lines. The comparison of the EPIs is shown in Fig. 10. As a conclusion, it is therefore recommended to use spherical sampling with equivalent pixel area if the light field camera is only used as a multi-view camera array. For use in the field of depth estimation, where the EPIs are used, sampling the radius in equidistant steps is preferable.

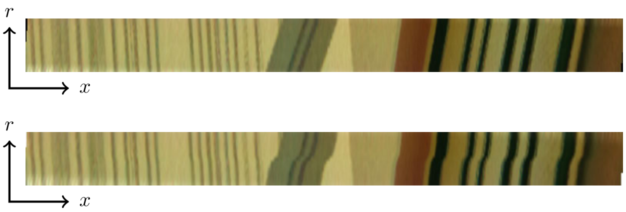


Figure 10: Spherical EPIs for $\phi = 0$: Radial sampling in equidistant steps (top). Radial sampling with $r_n \sim \sqrt{n}$ steps (bottom).

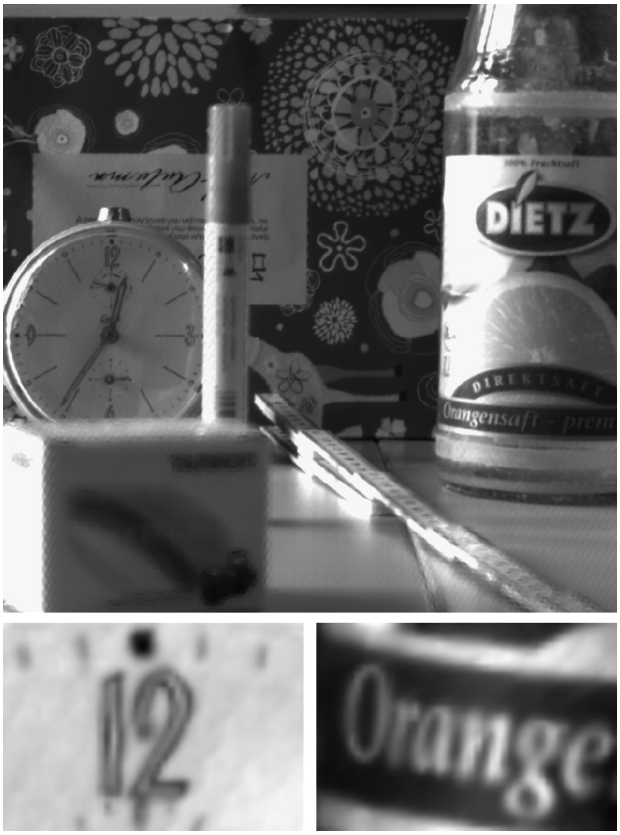


Figure 11: Top: Subaperture image from the center of the u, v-plane (Raytrix R5). Bottom: Zoomed in.

4.1.2 Focused plenoptic camera

While the proposed method can already reconstruct light fields very well from the raw data of the Lytro Illum camera, another advantage of the generic method is that the procedure works with other light field cameras without further adaptation. To show this, the light field of a Raytrix R5 was reconstructed. The reconstruction of the central subaperture view is shown in Fig. 11. One can see that the scene is reconstructed correctly and that even details are recognizable. Since this light field camera is built differently than the Lytro, not everything in the reconstructed image is in focus. With this camera, the depth of field and the focus distance are now determined by the main lens and the main lens setting. Because our lens is not optimally selected for the Raytrix R5, strong vignetting effects are visible at the edges of the microlenses, as can be seen in Fig. 3. While with the Lytro microlens-vignetting reduces the quality of the edge subaperture views, with the Raytrix the effect can also be seen in the central view. Very dark pixels at the edge of the microlenses cause reconstruction artifacts in the image due to a devignetting operation. However, this unwanted effect could be resolved by using a suitable lens with a hexagonal aperture and by manually adjusting the aperture's opening to the correct size.

4.2 Evaluation of the calibration

Apart from the reconstruction of the light field and the qualitative analysis of the result, an exact characterization of the ray geometry is essential for optical metrology and as well for many other areas of computer vision. Since the proposed method is based on a generic camera calibration and in order to be comparable with the very same, we need to investigate the ray projection error. This error equals the distance between a geometric camera ray and an observed point on a reference target, see Section 2.2.

In order to be able to evaluate the error experimentally, a commercially available monitor was used as a reference target, whose pixels serve as reference coordinates. The monitor was captured from different poses using the Lytro Illum. Where in each pose pattern sequences were displayed on the screen encoding the reference coordinates with subpixel accuracy. The corresponding 3D coordinates of these points were then determined using the procedure presented in [13].

The camera raw data with the measured point references were then converted to light fields using the method of Bok *et al.* [3] and the proposed one, respectively. Further, with the help of the respective camera parameters, the camera rays could be determined and the ray projection error as an average value over all rays could be calculated. The method of Danserau *et al.* [4] could unfortunately not be evaluated, as the rectification algorithm and thus the determination of the camera parameters only works for the older Lytro camera, but does not provide any meaningful results for the newer Lytro Illum.

The comparison of the different methods is shown in Tab. 1. As expected, the generic calibration has the lowest calibration error, since each pixel can be calibrated individually and hence with a high precision. However, this result cannot be compared directly to the other methods, since the correlations of the rays and the light field information are lost or can't be used immediately with this camera model. It is therefore only used to represent a lower limit of the calibration error. More importantly, it can be seen that the proposed method has a much smaller mean error and root mean squared error (RMSE) than the method of Bok *et al.*, resulting in a better calibration with less outliers. And thus, the ray geometry is estimated much better although the light field reconstruction results of both methods were very similar. This is due to the fact that the ray calibration of the proposed light field reconstruction itself could be carried out very precisely, starting from the generic calibration. The ray projection error is only slightly worsened during the interpolation and rounding operations of Section 3.3. In the end, a better calibration of the cameras geometrical properties leads to better results when used in the field of optical metrology, depth estimation or environment perception.

Table 1: Comparison of the ray projection errors.

	ε in µm	
	Mean	**RMSE**
Generic calibration [13]	49.4	105.8
Bok *et al.* [3]	375.6	758.8
Proposed method	97.1	155.0

5 Conclusions

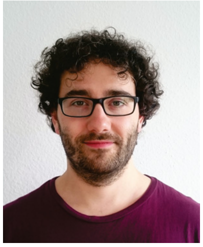
In this paper, we presented a method to calibrate any light field camera (*e. g.*, microlens-based, mirror-based, camera arrays) without having to model the exact optical properties. By means of a generic calibration, we were able to precisely calibrate the individual camera rays. Further, we normalized the result to subsequently transform it into an equivalent light field representation. Since classical algorithms require a regular sampling, we fit a regular 4D grid onto the irregular camera rays. The summation of the weighted intensity values of the rays finally led to the interpolation and reconstruction of a rectified light field. Apart from the usual Cartesian sampling of the angular coordinates, we presented two possibilities to sample them by means of spherical coordinates. This proved to be advantageous, since the light field information can now be represented more compactly. Besides the pure reconstruction of the light fields radiometric quantities, we also presented a derivation of the intrinsic camera parameters, *i. e.* the geometric quantities. The reconstructed light field can therefore also easily be used in the field of optical metrology and computer vision. Eventually, experiments showed that the proposed method can provide good reconstructions and that it provides rectified light fields. The epipolar geometry between the virtual subcameras is preserved and even shows better results than the conventional methods. In addition, an analysis of the geometric parameters by means of the ray projection error showed that the proposed method has a smaller calibration error than the state-of-the-art methods from the literature and thus, it achieves a better calibration. Still there is room for improvement. Hence, further work is dedicated to the improvement of the light field sampling, whereby both the desired resolution and the position of the grid points are to be optimized and adapted to the used camera.

References

1. A. Bartoli and P. Sturm. The 3D Line Motion Matrix and Alignment of Line Reconstructions. *International Journal of Computer Vision*, 57(3):159–178, 2004.

2. F. Bergamasco, A. Albarelli, E. Rodola, and A. Torsello. Can a Fully Unconstrained Imaging Model Be Applied Effectively to Central Cameras? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013*, pages 1391–1398, Piscataway, NJ, 2013. IEEE.
3. Y. Bok, H.-G. Jeon, and I. S. Kweon. Geometric Calibration of Micro-Lens-Based Light Field Cameras Using Line Features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2):287–300, 2017.
4. D. G. Dansereau, O. Pizarro, and S. B. Williams. Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1027–1034, 2013.
5. T. Georgiev and A. Lumsdaine, editors. *The multifocus plenoptic camera*. International Society for Optics and Photonics, 2012.
6. M. D. Grossberg and S. K. Nayar. The Raxel Imaging Model and Ray-Based Calibration. *International Journal of Computer Vision*, 61(2):119–137, 2005.
7. I. Ihrke, J. Restrepo, and L. Mignard-Debise. Principles of Light Field Imaging: Briefly revisiting 25 years of research. *IEEE Signal Processing Magazine*, 33(5):59–69, 2016.
8. A. Lumsdaine and T. Georgiev, editors. *The focused plenoptic camera*. IEEE, 2009.
9. M. Schambach and F. P. León. Microlens Array Grid Estimation, Light Field Decoding, and Calibration. *IEEE Transactions on Computational Imaging*, 6:591–603, 2020.
10. M. Schambach and M. Heizmann. A Multispectral Light Field Dataset and Framework for Light Field Deep Learning. *IEEE Access*, 8:193492–193502, 2020.
11. R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11):1–11, 2005.
12. O. Johannsen, A. Sulc, and B. Goldluecke. On Linear Structure from Motion for Light Field Cameras. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 720–728, 2015.
13. D. Uhlig and M. Heizmann. A Calibration Method for the Generalized Imaging Model with Uncertain Calibration Target Coordinates. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, 2020.
14. W. v. d. Hodge and D. Pedoe. *Methods of Algebraic Geometry*. Cambridge University Press, Cambridge, 1994.
15. S. van der Jeught and J. J. Dirckx. Real-time structured light profilometry: a review. *Optics and Lasers in Engineering*, 87:18–31, 2016.
16. S. Wanner and B. Goldluecke. Variational light field analysis for disparity estimation and super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(3):606–619, 2014.
17. S. Werling, M. Mai, M. Heizmann, and J. Beyerer. Inspection of specular and partially specular surfaces. *Metrology and Measurement Systems*, 16(3):415–431, 2009.
18. Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

Bionotes



David Uhlig
Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstraße 16, 76187 Karlsruhe, Germany
david.uhlig@kit.edu

David Uhlig is a research assistant at the Institute of Industrial Information Technology (IIIT) at the Karlsruhe Institute of Technology (KIT) in the research group of Prof. Heizmann. His main research includes light field cameras, deflectometry and 3D reconstruction.

Michael Heizmann
Institute of Industrial Information Technology, Karlsruhe Institute of Technology, Hertzstraße 16, 76187 Karlsruhe, Germany
michael.heizmann@kit.edu

Michael Heizmann is Professor of Mechatronic Measurement Systems at the Institute of Industrial Information Technology (IIIT) at the Karlsruhe Institute of Technology (KIT). His research areas include machine vision, image processing, image and information fusion, measurement technology, machine learning, artificial intelligence and their applications in industry.