



A Unifying First-Order Model for Light-Field Cameras: The Equivalent Camera Array

Lois Mignard-Debise, John Restrepo, Ivo Ihrke

► To cite this version:

Lois Mignard-Debise, John Restrepo, Ivo Ihrke. A Unifying First-Order Model for Light-Field Cameras: The Equivalent Camera Array. *IEEE Transactions on Computational Imaging*, 2017, 3 (4), pp.798 - 810. 10.1109/TCI.2017.2699427 . hal-01692127

HAL Id: hal-01692127

<https://inria.hal.science/hal-01692127>

Submitted on 24 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Unifying First-Order Model for Light-Field Cameras: the Equivalent Camera Array.

Loïs Mignard-Debise, *Manao, Bordeaux, France,*

John Restrepo, *independent scientist*

and Ivo Ihrke, *independent scientist*

Abstract—Light-field photography is an extension of traditional photography that enables among other effects refocusing, viewpoint change, and aperture synthesis of still images by digital post-processing. It achieves this capability by recording 4-dimensional radiance information rather than 2-dimensional integrated sensor irradiance. Consequently, optical design tools need to change in order to design these new devices.

In this article, we propose an optical first-order model that abstracts the architecture of any light-field camera as an Equivalent Camera Array (ECA). This model enables a comparison between different designs and allows for a simulation of the effects of parameter modifications to a design. We present equations for optical properties such as the depth of field, the angle of view, as well as important parameters for algorithmic performance such as the triangulation baseline. We provide an experimental validation of our model by measuring the properties of a real light-field camera. We are able to extract unknown physical parameters of the system such as the focal length of the main lens.

Index Terms—light-field imaging, computational optics

I. INTRODUCTION

IN the field of computational imaging, light-field cameras have been shown to be a great improvement on traditional cameras. They offer new possibilities to the end user such as synthetic refocusing, view shifting or depth estimation on a picture after only one single shot was taken. Instead of projecting the usual 2D image of the world, a light-field camera records 4D light-field data that require the application of post-processing algorithms for generating images that are understandable by human beings.

A few commercial products exist already, both for professional and consumer use. The technology combines new optical designs together with complex algorithms in order to efficiently measure and treat the light-field information from the scene. The role of the optics is to guide the light to the sensor, whereas the role of the algorithms is to interpret the value of the pixels. On the one hand, the algorithms need an accurate model of the imaging properties of the optical system and the sampling pattern of the sensor to correctly synthesize new images. On the other hand, the optics need to be conceived with the limits of the algorithms on the reconstruction in mind. Designing a light-field camera is, therefore, a co-design task involving both optical design and computer vision knowledge and aiming at optimizing the spatial-angular sampling trade-off.

Common knowledge suggests that light-field cameras are equivalent to an array of cameras since views with different

parallax can be synthesized from the light-field data. However, the details of how such an Equivalent Camera Array may be constructed for a given optical design have not been elaborated so far. We propose to model any light-field camera as an Equivalent Camera Array (ECA) based on first-order imaging. We describe and analyze the construction of the ECA from the components of a light-field camera. The ECA is quite similar to a real camera array and as such, we present the optical properties of its individual cameras such as the field of view and the depth of field. We also define and quantify important field properties such as the baseline, and the spatial and the depth accuracy by using information from multiple cameras. We validate the ECA model with a fit of the properties extracted from experimental data produced by a real light-field camera. We show that parameters such as the focal length of the main lens can be retrieved from our model. We compare the properties of different light-field cameras from the literature.

In sum, the contributions of this paper are the following:

- The unification of light-field cameras under the Equivalent Camera Array model based on first-order optics,
- The derivation of first-order optical properties of the ECA,
- The validation of the model with measurements from a real system, and
- The simulation and comparison of several micro-lens array based designs.

II. RELATED WORK

Light-field camera design: Several light-field camera designs have been studied and they all implement the integral imaging principles described by [1]. The use of an array of pinholes or an array of micro-lenses behind a single camera was first devised in [2] as an alternative to large camera arrays [3], [4] to acquire the light-field of a 3D scene. The idea is to separate light rays by the angle at which they hit the sensor and thus to capture a continuum of viewpoints of the scene.

Two variations of this design have been studied. The first one is the afocal camera [5] where the separation between the micro-lens array and the sensor is equal to the focal length of the micro-lenses. This design allows to directly extract a viewpoint by selecting the same pixel behind each micro-lens. The second design is the focused camera [6] where the separation between the micro-lens array and the sensor is greater or lower than the micro-lens focal length. The image

formed on the sensor is a grid of different points of view, each looking at a small part of the scene. A full view is generated by identifying recurrent elements between neighboring views and selecting the proper pixels.

A generalized light-field camera in [7] extends the afocal design where the distance between the sensor and the lenslet array is tunable. The study concludes that it is best to have this distance be lower than the focal length of the micro-lenses to allow a control over the spatio-angular trade-off. Apart from this work, these two ways of building a light-field camera are treated separately in terms of light-field sampling, calibration, analysis and rendering in the literature. Our goal is to propose a common model that can simultaneously abstract these two architectures and others.

More designs to capture light-fields with a single camera exist. An integrated camera array such as [8] where multiple cameras share the same sensor is a way to simplify the calibration of multiple independent cameras and to miniaturize this multi-view acquisition device. Other devices based on a spherical mirror array [9], [10], [11] or a large lens array [12], [13] placed in front of a perspective camera have been built, calibrated and analyzed. A design using light pipes [14] enables the utilization of a standard sensor and main lens. Fundamentally, all of these designs are sampling the light-field function by multiplexing several views onto the same sensor.

Calibration: The calibration of a light-field camera is a necessary task to extract the light-field data from the sensor. The first goal of the calibration is to compensate for misalignment between the sensor and the micro-lens array [15] and to parameterize and extract the light-field data. The second part is to find the intrinsic and extrinsic parameters of the camera and potentially its views to get an accurate description of the acquired light-field. The actual calibration methods differ between the afocal [16] and the focused case [17], [18]. Our common model predicts the location of the perspective cameras given by the calibration. The calibration would still be necessary as it accounts for variations of the parameters of the real system as well as the distortions due to aberrations. We experimentally validate the ECA model by recovering the parameters of the main lens of the camera from the fit of the measured properties predicted by our model.

Light-field Rendering and Analysis: The purpose of light-field rendering is to extract as much data as possible from an acquired light-field in order to compute refocus stack images, interpolate viewpoints or estimate scene depth. Naive light-field rendering produces low spatial resolution images which is considered as a downside in traditional photography. Super-resolution techniques [19] in conjunction with better depth estimation algorithms [20] have been studied using filtering in the phase space or the Fourier domain [21] to improve the spatial resolution and quality of the output images. A recent overview article [22] summarizes these works and more.

In order to improve light-field rendering techniques, the performance of the light-field camera has been analyzed geometrically. The depth of field of the system is an important

aspect of the rendering process as it sets physical limits to the volume in which images of objects can be synthesized [23]. Finding the best sampling scheme for a scene in a certain depth range [24] allows setting guidelines for the design of a complete system. The f-number matching rule [2], [5] between the main lens and the micro-lens array is one of these guidelines but more is needed. As an example, light-field camera design includes the choice of a suitable lenslet size, sensor specifications and main lens parameters for a specific application. The bases of a global model have been studied in [25] where the position and the baseline of virtual viewpoint cameras were first studied, but the model is incomplete. We position our work in continuation of this model.

Aberrations: Little work has been done on analyzing the effects of aberrations in light-field cameras. Ray-tracing through an afocal system was introduced by [26]. It was shown that digital correction of the aberrations can improve the quality of rendered images. Moreover, a small number of directional samples is already sufficient to significantly improve the effective resolution. The effect of irregularities in the micro-lens array and main lens spherical aberration were studied by [27]. They have shown that these variations are beneficial to the sampling of light and image reconstruction at any depth. However, there is not yet a complete theory of aberrations for light-field systems on par with classical optics. Our work proposes a first-order model for light-field cameras that points in the direction of a theory of light-field aberrations.

III. BACKGROUND

In this section, we recall basic optical concepts such as paraxial optical imaging and the properties of a standard camera. We introduce imaging matrices that map 3D points from one side of the lens (the object side) to the other side (the image side).

A. Lens Imaging

In paraxial optics, real lenses are considered to be idealized thin or thick lenses. The imaging of an object point $\mathbf{A} = (x_{\mathbf{A}}, y_{\mathbf{A}}, z_{\mathbf{A}})$ to its corresponding image $\mathbf{A}' = (x_{\mathbf{A}'}, y_{\mathbf{A}'}, z_{\mathbf{A}'})$ through a thin lens with an optical axis z and focal distance f , centered in the origin is described by the thin lens equation:

$$\frac{1}{z_{\mathbf{A}'}} - \frac{1}{z_{\mathbf{A}}} = \frac{1}{f}. \quad (1)$$

The magnification $M_{\mathbf{AA}'}$ between these two points is defined as :

$$M_{\mathbf{AA}'} = \frac{z_{\mathbf{A}'}}{z_{\mathbf{A}}} = \frac{x_{\mathbf{A}'}}{x_{\mathbf{A}}} = \frac{y_{\mathbf{A}'}}{y_{\mathbf{A}}}. \quad (2)$$

The magnification is constant between two conjugate planes orthogonal to the optical axis whether we use the thin lens or the thick lens model. However, it varies between different pairs of planes. The closer an object approaches the focal point of a lens, the further from it moves its image. At the same time, the magnification increases linearly. This implies that objects are three-dimensionally deformed through the lens following a

perspective transformation of the space. Consider Equation 1. It can be rewritten as:

$$z_{\mathbf{A}'} = \frac{z_{\mathbf{A}} f}{z_{\mathbf{A}} + f}. \quad (3)$$

By using Equation 2 and 1, we obtain:

$$x_{\mathbf{A}'} = \frac{x_{\mathbf{A}} f}{z_{\mathbf{A}} + f} \quad (4)$$

and similarly for $y_{\mathbf{A}'}$. Using homogeneous coordinates for \mathbf{A} and \mathbf{A}' , we can derive the following relation:

$$\begin{pmatrix} \hat{x}_{\mathbf{A}'} \\ \hat{y}_{\mathbf{A}'} \\ \hat{z}_{\mathbf{A}'} \\ \hat{w}_{\mathbf{A}'} \end{pmatrix} = L \begin{pmatrix} x_{\mathbf{A}} \\ y_{\mathbf{A}} \\ z_{\mathbf{A}} \\ 1 \end{pmatrix} \quad (5)$$

with

$$L = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \\ 0 & 0 & 1 & f \end{pmatrix}. \quad (6)$$

The coordinates of \mathbf{A}' can be retrieved as follows: $x_{\mathbf{A}'} = \frac{\hat{x}_{\mathbf{A}'}}{\hat{w}_{\mathbf{A}'}}$, $y_{\mathbf{A}'} = \frac{\hat{y}_{\mathbf{A}'}}{\hat{w}_{\mathbf{A}'}}$ and $z_{\mathbf{A}'} = \frac{\hat{z}_{\mathbf{A}'}}{\hat{w}_{\mathbf{A}'}}$. This imaging matrix L is similar to the perspective projection matrix used in computer graphics to project the coordinates of points in the world space to the pixels of a virtual camera.

Considering a lens that is not situated in the origin, but centered at a point $\mathbf{H} = (x_{\mathbf{H}}, y_{\mathbf{H}}, z_{\mathbf{H}})$ with an optical axis parallel to the z -axis, its imaging matrix $\mathcal{L}_{\mathbf{H}}$ is given by the composition of the imaging matrix L and the translation matrix of the point \mathbf{H} , $T_{\mathbf{H}}$, such that:

$$\mathcal{L}_{\mathbf{H}} = T_{\mathbf{H}} L T_{\mathbf{H}}^{-1} \quad (7)$$

When a lens is too thick, it cannot be considered as a thin lens anymore. Instead of having the refraction occurring on the same lens plane, it is split into two principal planes: $\mathcal{P}_{\mathbf{H}}$ and $\mathcal{P}_{\mathbf{H}'}$. The distance B between these two planes accounts for the effect of the thickness of the lens. The front and back principal planes are perpendicular to the optical axis and their intersections with the optical axis are, respectively, the front and back principal points, \mathbf{H} and \mathbf{H}' . The imaging matrix of a thick lens is given by:

$$\mathcal{L}_{\mathbf{H}} = T_{\mathbf{H}'} L T_{\mathbf{H}}^{-1} \quad (8)$$

with $T_{\mathbf{H}'} = T_{\mathbf{H}} T_B$.

B. Camera Properties

As they are, the previous equations do not give a clear interpretation of what is seen on a camera sensor. The performance of the camera also depends on the parameters of the sensor itself in sampling the light. For simplicity, we consider that the camera is made of an optical system which can be approximated as a single thick lens and a sensor orthogonal to the optical axis of the lens. Most of the common notions that are characteristic of a camera refer to the object side of the camera such as the view direction, the angle of view and the

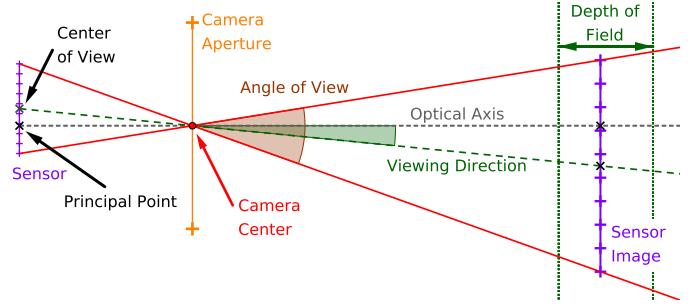


Fig. 1: Our camera model and its properties. These properties characterize the object space of the camera. When imaging off-axis, the center of view and the principal point are not at the same position resulting in the viewing direction being different from the optical axis. The field of view is the whole space that can be projected onto the sensor through the camera center. The angle of view is the angular extent of the field of view and the view direction is its bisector.

depth of field. The definitions of these properties are described in Fig. 1. The same notions will be used in Section IV for the individual cameras of the Equivalent Camera Array.

IV. THE EQUIVALENT CAMERA ARRAY (ECA)

A light-field camera spatially multiplexes directional light information onto its sensor, thereby realizing a 4-dimensional light-field sensor with a 2-dimensional pixel array. In order to do so, it uses a component that directs the light rays that are incident at a particular spatial position from different directions to different pixels. We call this component the *directional multiplexing unit (DMU)*. The DMU is often implemented by a lenslet array. However, since multiple designs exist, we prefer the more general terminology.

Our goal in the following is to abstract a real light-field camera by a virtual camera array (ECA) that is observing the object space. This abstraction is possible for most existing light-field camera designs. We then describe the properties of the virtual cameras as introduced in Section III, as well as additional properties that are derived from coupled information between different virtual cameras in order to interpret and compare two light-field camera designs and configurations in Section V. More examples can be found in the supplementary materials.

A. Principle

Individually, each element of the DMU and the pixels associated with it act as a small camera looking at the world. In a light-field camera, a main lens is often added in front of this array of small cameras. From the point of view of a small camera, instead of looking directly at the object space, it is looking at the in-lens space. The in-lens space is a perspective mapping of the object space following the equations of Section III. In order to retrieve the properties of an equivalent camera array, we would like to transform this in-camera array of small cameras into the object space, where it could be treated much like a normal camera array

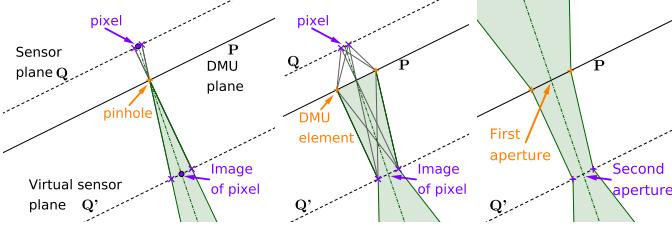


Fig. 2: (Left) pinhole model. (Middle) pixel+DMU element model. Every ray contained in the green area will be integrated by the pixel. (Right) two-aperture model.

consisting of physical cameras. However, mapping a camera through the main lens has no obvious physical solution due to the distortion of space affected by the lens. To arrive at a solution, we need to consider an abstraction that describes the effect of the combined small camera-main lens system. In the process, we will lose some physical properties of the system, in particular, its image side properties will only be defined up to a one-parameter family of solutions.

We resort to the two-aperture model introduced in the pioneering work [28] for the analysis of light-field sampling properties for real camera arrays. In light-field cameras, every pixel is assigned to the main ray passing through the center of that pixel and the center of an attributed element of the directional multiplexing unit.

Let us consider a two-plane parameterization of the light-field in the case of a small camera array with the sensor plane and the DMU plane as the support planes. We respectively label them plane Q and plane P . We consider at first that the DMU element is replaced by a pinhole as illustrated in Fig. 2 (left).

Any light ray hitting a certain position on the sensor also passes through the image of this particular position in the object space of the pinhole. However, a pinhole is a focus free imaging element so there is an infinite number of image planes Q' for the sensor plane Q . In terms of parameterization of the light-field, since the hit position of the light ray with the planes Q and Q' is a relative distance to the optical axis, these planes are equivalent. The pair of planes (Q, P) of the parameterization can be replaced by the pair (Q', P) .

This replacement allows to abstract the effect of the DMU element but still conserves the relation between the parameterization and the sampling of the scene that is implemented by the camera. This model is often used for the calibration of light-field cameras [16] but it neglects the physical focusing aspect of light-field cameras.

In a more realistic system, as shown in Fig. 2 (middle), the DMU element is now a first-order optical element and as such it has focusing properties. A finite-sized pixel on the sensor integrates all light that is passing through the surface of the associated DMU element and that is hitting its finite surface.

Observe that this pencil of light is also passing through the complete surface of the image of the pixel outside the camera, i.e. the unique plane Q' which is the optically conjugate plane of Q , and intersects the same surface area on the DMU element.

Therefore, it suffices to know the positions and the surfaces of the pixel image and the DMU element to predict the light rays that are integrated by the corresponding sensor pixel. Note that these two positions are now located in the object space of the small camera, i.e. that the optical effect of the DMU can now be ignored.

As illustrated in Fig. 2 (right), we abstract a pixel/DMU element combination by a two-aperture system with the positions and extents as described above. This system can be consistently imaged through an optical component and it preserves the information on the focusing properties of the light-field subviews. As mentioned previously, the disadvantage of this abstraction is that the apertures loose their physical properties. We discuss the consequence of this loss on the ECA in Section VII.

This procedure was applied to several setups from the literature in Fig. 3 showing a variety of light-field camera designs that can be analyzed with our model. Note that the main lens may be missing and that the directional multiplexing unit may have additional relay optics. In order to be more specific, we use the relatively complex KaleidoCamera design [14] as an illustrative example. The system is made of a main lens and a sensor with an in-between directional multiplexing unit that consists of two lenses, a field lens and a pickup lens that are at the entrance and the exit of a mirroring light pipe. In Fig. 3(d), the pixel area of the sensor serves as the pixel aperture, whereas the pickup lens generates the aperture of the DMU. The light pipe generates a virtual DMU lens array through mirroring. The field lens is a relay system that images the plane of the DMU into the exit pupil of the main lens.

Rays starting from the pixel aperture pass through the pickup lens and are imaged and reflected through the different system components. Finally, they pass through the images of the pixel and pickup lens aperture in the object space.

The general procedure is illustrated in Fig. 4 and is detailed as follows. Every combination of pixel/DMU elements is decomposed into two-aperture elements. They are imaged through the sequence of the optical components of the light-field camera, DMU element included, to the object space of the camera. From the equations of Section III, a pixel image is given by: $\hat{\mathbf{X}}_{Q_W} = \mathcal{L}_{Main}\mathcal{L}_{DMU}\mathcal{L}_{Relay}\hat{\mathbf{X}}_q$, and a DMU image is given by: $\hat{\mathbf{X}}_{P_W} = \mathcal{L}_{Main}\hat{\mathbf{X}}_P$. The two-aperture elements that share the same aperture imaged from the same DMU element are selected to form a virtual camera. There is one virtual camera per DMU element resulting in a virtual camera array alias the ECA that is equivalent to the in-camera array in the sense that it integrates the same ray bundle as the physical light-field camera.

We investigate more thoroughly the micro-lens based light-field camera designs illustrated in Fig. 3(f) and (e) as well as the properties of their ECA in Section V.

B. Properties

The abstraction of the pixels and DMU elements as apertures is sufficient to define similar properties as those described in Section III for a standard camera since most of them characterize the object side of the camera. Given an

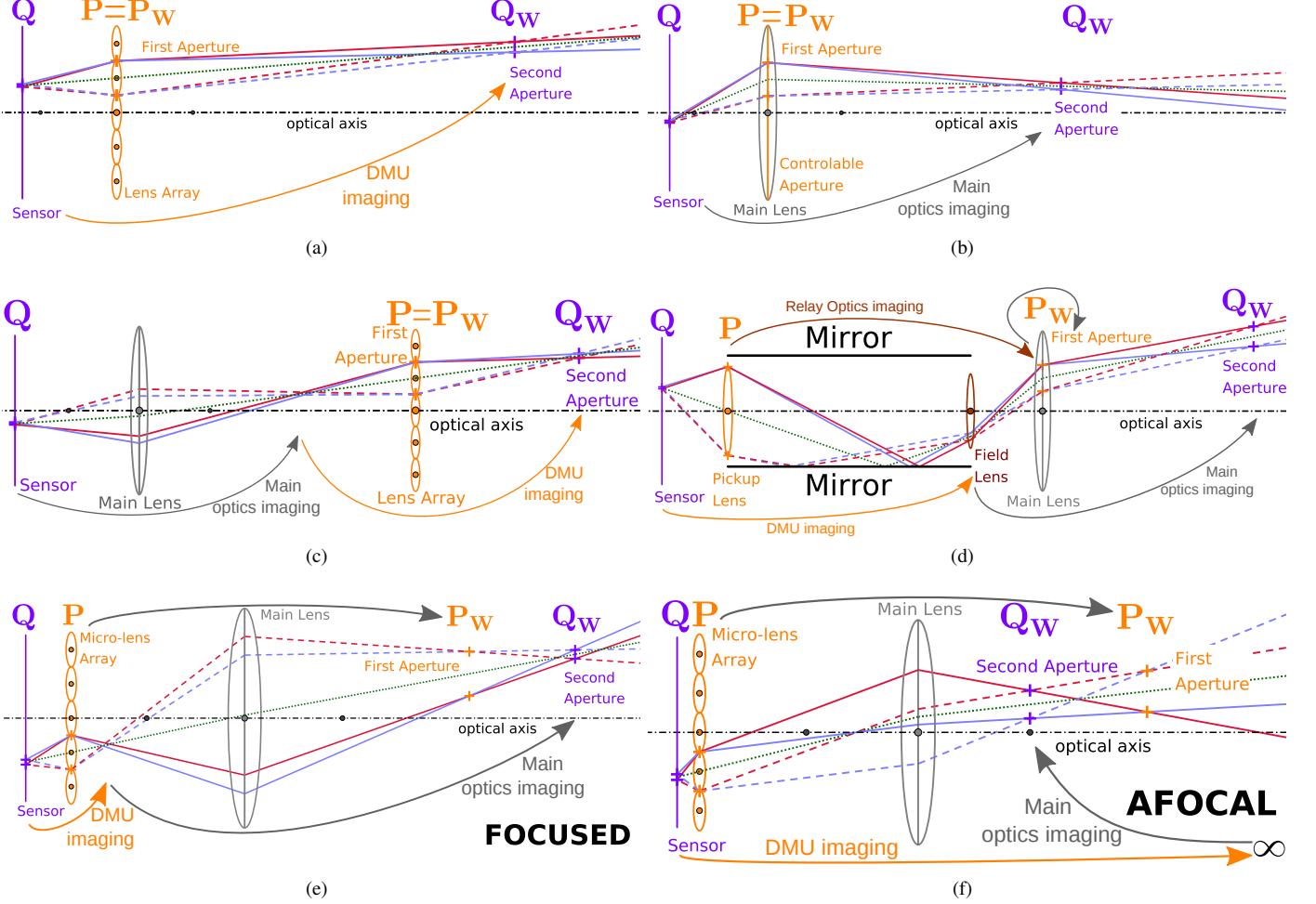


Fig. 3: Light-field camera designs from the literature. (a) The monolithic camera array from [8]. (b) The programmable aperture from [29]. (c) The external lens array from [12]. (d) The KaleidoCamera from [14]. (e) The focused light-field camera from [6]. (f) The afocal light-field camera from [5].

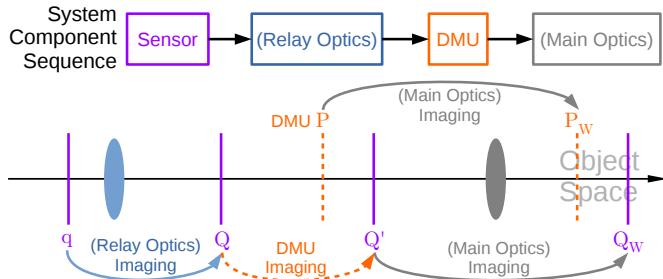


Fig. 4: General model for the mapping of the sensor and the directional multiplexing unit to the object space. The main optics and the optics introduced between the sensor and the DMU planes, such as a relay system, for example, are optional. In this case, the imaging matrix L of the optics is replaced by the identity matrix.

equivalent camera array, we can compute the view direction, the field of view, the depth of field and the resolution that are presented in Section III-B for each of the individual cameras. Since the ECA is made of several cameras, additional

information of two or more cameras can be used to derive new properties of the system. The following properties are only valid for the sharp region of the scene space. This region is delimited by the limits of the depth of field of the cameras of the ECA. Objects outside of this region are out of focus so they appear blurry and become indistinguishable.

Disparity: In stereo vision, the disparity is the difference in pixels of the position of the image of the same feature on two different cameras. The disparity is zero at the plane where the view directions intersect. For a scene point behind the no-parallax plane, the disparity is positive and it is negative when the point is closer than this plane.

Baseline: The baseline is the distance between the centers of projection of two cameras looking at the scene. It is an indicator of the ability of the system to measure the disparity of a point in the scene. In the case of a camera array, multiple cameras can see the same point so the interesting value is the maximum baseline of all pairs of cameras. A baseline map of the whole object space can be computed by

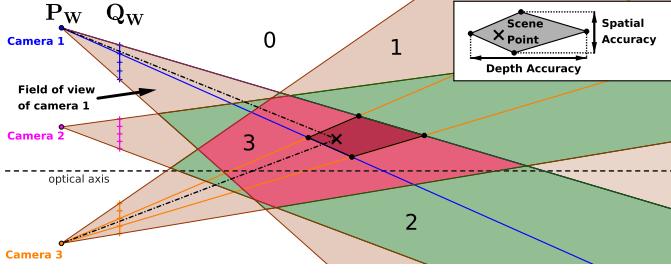


Fig. 5: Definition of the properties for the field. The whole field can be separated into different regions depending on the number of cameras observing it. Any point in the same diamond shaped region would project onto the same pixels as the defined scene point. The largest extent of this region along and orthogonal to the optical axis of the two cameras respectively define the spatial and depth accuracy on the position of a point in this region. Every pair of pixels, one per camera, defines a unique region of space.

intersecting the fields of view of every possible pair of virtual cameras as shown in Fig. 5. It is possible that the baseline is null for scene points that can be seen by only one virtual camera or undefined for scene points outside of the field of view of every camera of the array.

Accuracy: A point in space can be projected onto the virtual sensor of a camera (plane Q_W). This projected point is defined as the intersection between the line formed by the point and the center of the camera, and the plane of the virtual sensor. When a point in object space is projected onto the virtual sensor of two cameras, it will fall onto one pixel in each camera. Similarly to the field of view of a camera, the field of view of a certain pixel is the cone defined by the camera center and the edges of the pixel. As can be seen in Fig. 5, the intersection of the field of view of two pixels from two cameras results in a region in space. The points belonging to this region cannot be differentiated by only utilizing the two cameras. This region is reduced as more pixels from different cameras image a scene point. The spatial and depth accuracy of a system of two cameras is respectively the largest transversal and longitudinal dimension of this region.

C. Pixel/DMU element pairing

The pairing between the pixels and the DMU elements has to obey a few rules in order to build a consistent camera array. The main rule is that there should be no more than one DMU element paired up to each pixel. In micro-lens based light-field cameras, this condition is known as the f-number matching rule. In order to prevent the overlapping or gaps between the images made by two neighboring DMU elements, the working f-number of the DMU elements and of the main lens should match. In the KaleidoCamera, it is the field lens aperture projected onto the sensor through the DMU element center after zero, one or many reflections that is used to select the pairings appropriately. In this case, the main lens also plays a role as it limits the extent of the ray bundle entering the camera, thus cutting down the number of possible reflections.

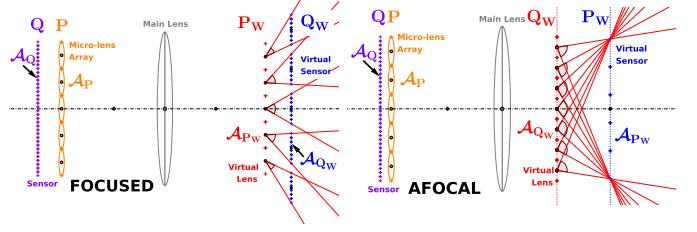


Fig. 6: ECA of the focused and afocal light-field camera. The fields of view of the individual cameras of the ECA are shown in red. The layout of this figure is the same for the following figures plotting the different optical properties. The parameters of the different components have been chosen to have a compact figure and do not represent a realistic imaging system. There would usually be an overlap of the virtual sensors for neighboring virtual cameras in the focused configuration.

V. STUDY OF MICRO-LENS ARRAY BASED LIGHT-FIELD CAMERAS

In order to illustrate the construction of the ECA as explained in Section IV, we study the specific design of light-field cameras based on the use of a micro-lens array as the directional multiplexing unit.

A. Construction of the ECA

We consider that the micro-lenses are thin lenses. The two-aperture system we need to image to the out-camera space is made of a pixel and a micro-lens. We respectively denote the position of the center of these apertures by C_Q and C_P . Their positions in the out-camera space C_{Q_W} and C_{P_W} are obtained by applying the equations from Section III as follows:

$$\hat{C}_{Q_W} = \mathcal{L}_{Main} \mathcal{L}_{DMU} \hat{C}_Q \quad (9)$$

$$\hat{C}_{P_W} = \mathcal{L}_{Main} \hat{C}_P \quad (10)$$

The edges of the apertures \mathcal{A}_{Q_W} and \mathcal{A}_{P_W} in the out-camera space can be imaged in the same way from \mathcal{A}_Q and \mathcal{A}_P , the apertures of the pixel and micro-lens. Fig. 6 illustrates the position and the field of view of the virtual cameras of the ECA.

B. Afocal case

There exists a specific case for which the role of the apertures as the virtual sensor or the virtual camera can be switched. This occurs when the distance between Q and P is equal to the focal length of the micro-lens array. In this case, the image of the pixels are sent to infinity by the micro-lenses and then the main lens images the pixels in its front focal plane as can be seen in Fig. 3(f). The pixels with the same relative position to the center of their assigned micro-lens integrate light rays of the in-camera light-field with the same direction.

The difference with the focused configuration is better explained in a phase space diagram as shown in Fig. 7. A phase space diagram records the height u and the direction s of light rays at a specific plane [22]. Positioning the evaluation

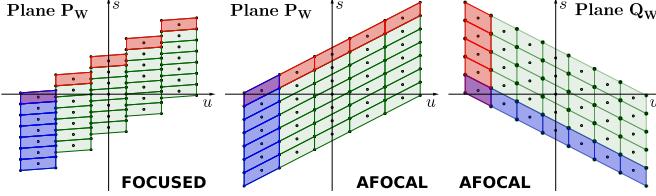


Fig. 7: Phase space of the ECA. A camera view is obtained by summing all cells from one column. In the focused case (left), the evaluation plane is the P_W plane, a column of cells in blue corresponds to different pixels from the same DMU element. For the afocal configuration, there are two ways to obtain a column of cells: when the evaluation plane is P_W as for the focused case (middle) or when the evaluation plane is Q_W (right). In the latter case, a column of cells in red corresponds to the same relative pixel position from different DMU elements.

plane of the phase space at a different location is equivalent to shearing the phase space plot along the u -axis.

In this space, a virtual camera is represented by a vertical column of contiguous pixels. The only position for which this condition is satisfied in the focused case is at the P_W plane (Fig. 7, left).

In the afocal configuration, though, there are two configurations that yield virtual cameras, 1. for the evaluation plane positioned at the P_W plane and 2. for the evaluation plane positioned at Q_W (Fig. 7, middle and right). Choosing one plane or the other allows for creating two different ECAs, that, however, describe the same set of rays. It may be noted that the two apertures of the ECA model determine the boundaries of the phase space parallelograms.

C. Simulation

Since the ECA model is a tool to evaluate the properties of any light-field camera, we simulate different systems taken from the literature and from existing commercial products. The supplementary material contains the exhaustive list of the results of the simulation as well as the values of the parameters for each design. The micro-lens array plane is fixed at the origin, so, the effect of moving the object plane or (physically) refocusing with the main lens is simulated by only changing the position of the main lens along the optical axis and computing the camera array properties. The value of the refocus is the distance between the micro-lens array and the back principal plane of the lens. We also investigate the effect of a varying distance between the sensor and the micro-lens array.

The main lens is approximated by a thin lens. Vignetting effects between the micro-lenses and the main lens are not taken into account as it would change the pairing between pixels and micro-lenses and make the following discussions more difficult.

In the following study, we present the results for the light-field camera from [20] as an illustrative example since its properties show the effects of micro-lens based light-field camera designs most clearly. This camera has been designed

TABLE I: Values for the parameters for the setup described in [20] used in our comparison study.

Parameter	Pixel	Micro-lens	Main Lens
Pitch/Diameter (mm)	0.009	0.135	40.0
Focal length (mm)	-	0.35	80.0
Number	4095	273	1

TABLE II: Position of the two aperture planes P_W and Q_W in out-camera space. The letters R, F, V respectively indicate that the plane is real, in the front focal plane of the main lens or virtual.

Lens Pos	Img P		Img Q'		Neither		
	Plane Pos	P_W	Q_W	P_W	Q_W	P_W	Q_W
Galilean	∞	R	V	∞		R/V	R/V
Afocal	∞	F	-	F		R/V	F
Keplerian	∞	V	R	∞		R/V	R/V

to be used in the focused configuration and as such it is representative of similar existing systems. However, it was not particularly designed to be used in the afocal configuration. In order to compare the focused and afocal configurations on a common basis, we created an afocal version of [20].

The parameters of the components are summarized in Table I. The only difference between the afocal and focused setups is the distance between the sensor and the micro-lens array. It is equal to the micro-lens focal length for the afocal setup and it is 1.2 times this value for the focused setup.

D. Properties

The properties are evaluated and plotted for 2D systems.

Plane positions: The first interesting property is the position of the planes Q_W and P_W along the optical axis. Table II shows a summary of the possible positions of these planes depending on the position of the main lens (either at one focal length away from Q' or P or neither) and the configuration of the sensor and micro-lens array. The distance separating the sensor and the micro-lens array can either be lower, equal, or greater than the focal length of the micro-lenses and corresponds to configurations called respectively galilean, afocal, and keplerian. The virtual camera array can be made of perspective or directional cameras, looking at a real or virtual plane, at a finite or infinite distance. In the afocal setup, Q_W is always located in the front focal plane of the main lens whatever its position is. This also means that, for this case, the pitch and height of the pixel aperture in object space do not depend on the position of the main lens.

View direction: The view direction is the angle between the line connecting C_{P_W} and the center of \mathcal{A}_{Q_W} , and the optical axis. In the focused case, both the center and aperture planes move and the view direction is constant, Fig. 8 (left). However, in the afocal case, since the plane P_W is static but not Q_W , the view direction changes, see Fig. 8 (right).

Angle of view: The field of view is delimited by the cone of rays centered in C_{P_W} and bounded by \mathcal{A}_{Q_W} . Generally, in the same ECA, the magnification for planes P_W and Q_W is different causing the relative position of C_{P_W} and \mathcal{A}_{Q_W} of two

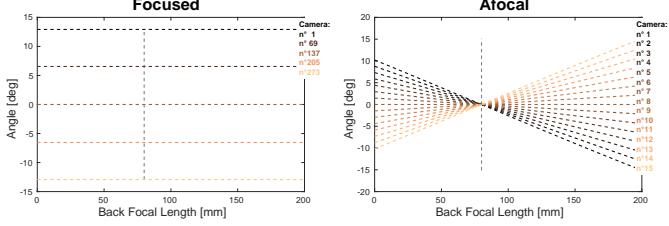


Fig. 8: Viewing direction of each of the virtual cameras of the ECA. There are as many virtual cameras as micro-lens in the focused case (left), so 273. For the afocal case (right), the number of virtual cameras is the number of pixels behind a micro-lens or equivalently the ratio of the number of pixels to the number of micro-lenses, so 15. The vertical gray dashed line indicates the focal length of the main lens.

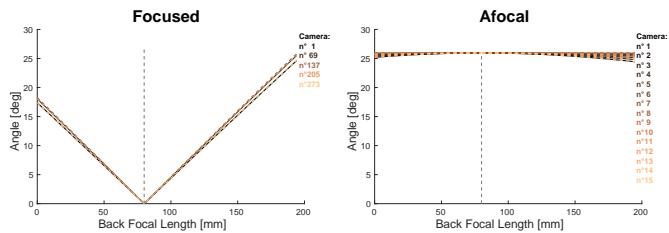


Fig. 9: Angle of view of each of the virtual cameras of the ECA.

neighboring cameras to be different. Consequently, the angle of view of two neighboring cameras is slightly different, Fig. 9. In the focused case, when the plane of the virtual cameras of the ECA, P_W , is imaged to infinity, the angle of view becomes zero when the back focal length is equal to the focal length of the main lens. In a classic configuration where the distance between the main lens and the MLA a bit larger than the focal length of the main lens, the angle of view remains low since the number of pixels of the virtual sensor is small. In the afocal case, the plane of the virtual cameras Q_W never goes to infinity. So, the angle of view hits the maximum value when the plane P_W goes to infinity. Moreover, the number of virtual pixels per camera is high and so is the angle of view.

Depth of Field: In a classical camera, the depth of field is by definition located around the Q_W plane, which is also the plane of best focus. The same is true for the virtual cameras of the ECA. In addition, for light-field cameras, the depth of field of the ECA cameras determines the range of synthetic refocusing. The ECA cameras all have the same depth of field as it depends solely on the pitch of the virtual apertures and the distance between their planes. As the virtual focus plane moves further away from the virtual lens plane, the depth of field grows larger till becoming infinite. This effect can be observed in Fig. 10 when the back focal length approaches the main lens focal length. The asymptote position determines the hyperfocal distance of the system where the sharpness range in the image is the largest.

Baseline: The baseline is a step function that only takes values that are integer multiples of the distance between two

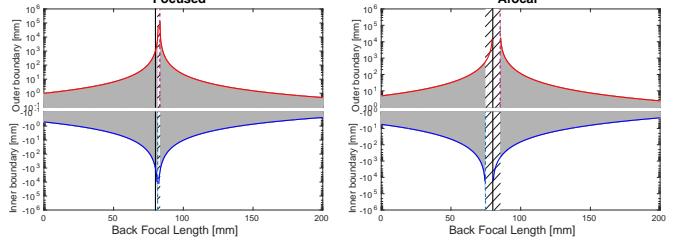


Fig. 10: Depth of field. The blue and red curve represent the distance of the boundaries of the depth of field from the virtual sensor plane. The depth of field is common for all the virtual cameras. The dashed blue and red vertical lines indicate the hyperfocal distances of the camera. The hashed area between these positions is the area where the depth of field is infinite.

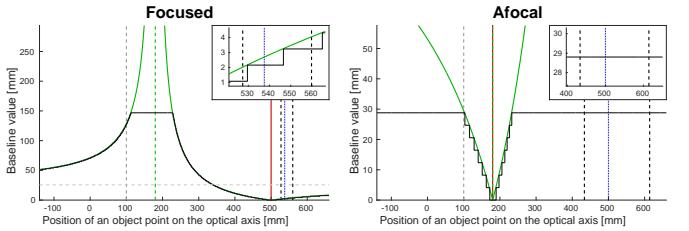


Fig. 11: Baseline for an evaluation point on the optical axis for a Back Focal Length of 100mm. The theoretical baseline (green curve) assumes a continuity of infinitely many cameras in the ECA. The real baseline (black curve) is a step curve computed from the actual position of the ECA cameras. Since the possible values for the baseline (the distances between the centers of the virtual cameras) are discrete, this curve is a step curve that has a maximum equaling the distance between the two extreme cameras of the ECA. The horizontal gray dashed line indicates the theoretical value of the baseline for an evaluation point at infinity. The vertical lines indicate positions of interest such as the main lens plane (in gray), its front focal plane (in green), virtual camera plane (in red), the virtual sensor plane (in blue) and the depth of field boundaries (in black).

neighboring cameras. It is a positive function that is bounded by the maximum distance between the cameras of the ECA. The results are shown in Fig. 11. For both the focused and afocal cases, the baseline is minimum at the virtual camera center plane position. However, the baseline is maximum at the position of the front focal plane of the main lens in the focused case and at the virtual sensor plane position in the afocal case. The important region is the one situated between the depth of field limits. In the focused case, the baseline per camera is low because the overlap between neighboring cameras is reduced as the zero-disparity plane is behind the cameras. In the afocal case, the baseline is at its largest on the full depth of field range since the zero-disparity plane is at the virtual sensor plane.

Accuracy: The results are shown in Fig. 12. The transversal measure of the accuracy is linear with the absolute distance of the evaluation point to the plane of the virtual camera center.

The longitudinal measure of the accuracy is a more complex curve.

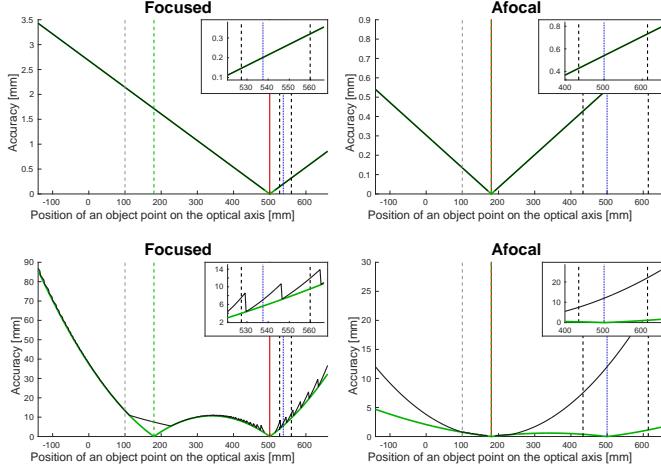


Fig. 12: Transversal (Top) and longitudinal (Bottom) accuracy for an evaluation point on the optical axis for a Back Focal Length of 100mm. The real curves (in black) have undefined values for positions where the baseline is either zero or undefined (close to the virtual camera center plane in red). The theoretical longitudinal accuracy (in green) and the real accuracy (in black) differ for positions where the baseline is clamped. Overall, the discontinuous behavior of the real curve is due to the discrete change of the baseline value. The vertical lines have the same definitions as in Fig. 11.

VI. VALIDATION

We validate our model through experimentation by obtaining the properties of the ECA for a real light-field camera. As a result, we can estimate some unknown physical specifications of the system. In contrast to our first-order model, we observe effects of non-linearities in the real data which we point out in the discussion of our results.

Our selected light-field camera is the first generation Lytro [30], which can be categorized under the afocal type systems similar to Fig. 3(f). We use this light-field camera to create a data set of correspondences between world rays and sensor pixels that are later used to construct a generalized imaging model as described in [31]. Given the interpretation of a light-field as an array of subview images of the scene, we obtain a generalized imaging model per subview that is directly analogous to our proposed ECA.

A. Experimental setup

Our experimental setup is displayed in Fig. 13. We modify the camera by removing the main lens from its encasing. This way, we can control its distance to the sensor and micro-lens array. The Lytro camera provides optical refocus and zoom. For simplification purposes, we fix the main lens settings, keeping its optical properties constant.

The separated lens and light-field sensor (micro-lens array-sensor couple) are set up independently. Given our mechanical

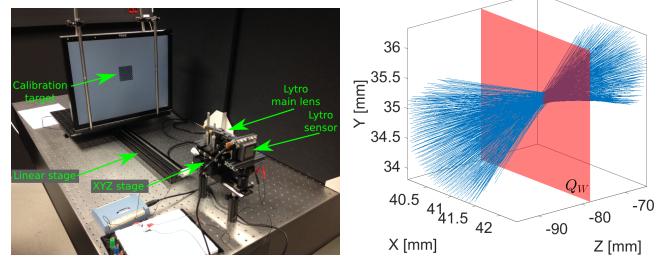


Fig. 13: (Left) Photograph of the experimental setup. (Right) Ray bundle for a single subview with a reduced number of rays for visualization. The intersecting plane identifies the location of the center of perspective.

conditions, we cannot guarantee an optimal alignment as compared to the original camera. As a consequence, our data shows some irregularities that we will discuss in the following subsection. We measure world space rays by recording multiple positions of a calibration target [32] displayed on a computer screen for a series of known distances.

B. Estimation of properties

The pre-processing of the acquired data involves decoding the light-fields [16], detecting the corners of the calibration target [32], and upsampling the resulting ray data set to obtain a corresponding ray for every pixel in each subview. In order to construct property plots such as Fig. 8, we repeat the experiment for multiple displacements of the main lens ΔZ , which produces property measurements for different back focal length values.

An example of the ray data is shown in Fig. 13 (right), for a single subview and main lens position. From the ray distribution, we can conclude that the imaging is effectively a perspective transformation.

We equate the locus point of each ray bundle with the center of perspective for each subview; these points should agree with our ECA positions in order to validate our model. The obtained centers for all subviews are plotted in Fig. 14(a), the plane containing the centers is the aperture plane Q_W of our model. We ignore the outer-most centers because of low contrast. The array of centers indicates regularity with the exception of the extreme subviews. This validates the ECA model for the central paraxial region but also indicates the presence of aberrations for the outer subviews.

Our ECA model predicts that the array of centers is located on a plane perpendicular to the optical axis as presented in the sketch in Fig. 14(c). However, our modification of the Lytro camera creates a misalignment between the axis of the lens and the micro-lens array, which effectively produces a Scheimpflug effect, tilting the imaging planes of the main lens. We do not account for these effects in our model and instead use our experimental data to obtain a new optical axis. We compute a unique optical axis for all lens positions ΔZ . This new optical axis is defined as the normal of the least squares fitted plane to all estimated centers of perspective. The result is shown in Fig. 14(d). The optical axis obtained with

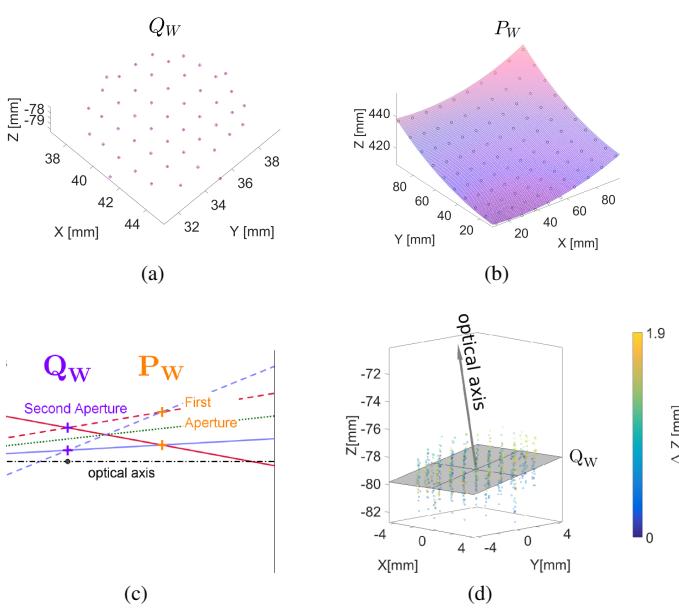


Fig. 14: (a) Centers of perspective for a single main lens position. (b) Best focus plane, with the markers as the intersections of rays and a surface as spherical fit. (c) Sketch for an afocal setting. (d) Plot of all centers for all lens positions and computed optical axis.

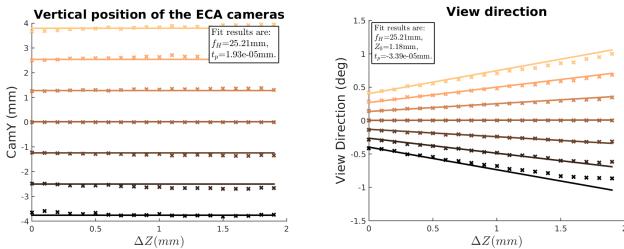


Fig. 15: (Left) Height of perspective centers *markers* and corresponding fit results *solid*. (Right) View direction of subviews *markers* and corresponding fit results *solid*. The central camera is excluded from the fit since it was used to re-center the other ones. ΔZ indicates the main lens displacement with respect to the micro-lens array.

this procedure is used to obtain the experimental properties of the subviews.

Plane positions: We compute the perspective centers, again, for multiple positions of the main lens represented by the distance ΔZ of the main lens with respect to its initial position. Furthermore, to cancel variations between different experiments for different main lens positions, that are due to mechanical influences, we re-center all cameras to the central one. This implies that we cannot measure the central camera, only the separation inside the array. The results are displayed in Fig. 15 (left). In order to match the format of plots employed in the previous sections, we restrict ourselves to the central column array of subviews.

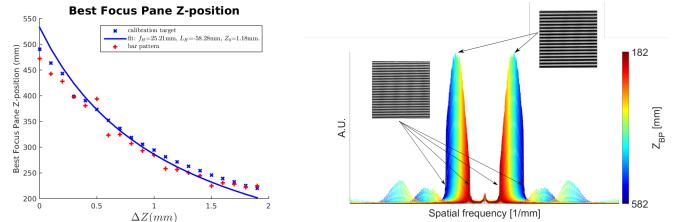


Fig. 16: (Left) Best focus plane location for the calibration target red markers corresponding fit solid line and for comparison, the bar pattern results. (Right) Superposition of the spectra at multiple target positions (with DC term removed). There are two subviews of the bar pattern indicating the positions for highest and lowest contrast. Colors indicate different target positions.

View Direction: Similar to its definition for our ECA model, the view direction is measured as the angle between the central ray of each subview and the optical axis of the ECA. For our experiment, we use the central ray of the bundle as shown in Fig. 13 (right) and the normal from the fitted plane to all centers as in Fig. 14(d). The view direction is the angle between these two vectors. For the same column of cameras as before, the experimental view direction is displayed in Fig. 15 (right).

Best Focus Plane distance: This plane corresponds to the plane P_W in our ECA model. It is simultaneously the plane of zero disparity between subviews. Therefore, each corresponding ray from the same pixel in all subviews must intersect in this plane. A computed example, for a smaller sample of rays, is displayed in Fig. 14(b). We observe that, in fact, we do not obtain a plane but a curved surface. This clearly indicates the presence of optical aberrations that are global to all subviews, in particular, a field curvature of the main lens. We compute the axial location of this surface as the average of the Z coordinates for all points in it. The resulting plot is displayed in Fig. 16 (left).

We further support our experimental findings by measuring the best focus plane location with an alternative method. We use the same experimental setup, now displaying a binary bar target for multiple positions of the computer screen Z_{BP} . Treating each subview independently, we use a metric for contrast to establish the best focus plane location and subsequently averaging for all subviews. Our metric for contrast is based on the Fourier spectrum of the bar target. We take an average of several line profiles of each image and compute the 2D spectrum. In it, we measure the height of the secondary spectral peaks as a function of the bar target position. The highest contrast corresponds to the maximum peak. The spectra for multiple bar pattern profiles at different distances is displayed in Fig. 16 (right). There are two subviews with high and low contrast to exemplify the peak heights to which they correspond. The frequency location of the peaks varies with the position of the target Z_{BP} due to the magnification effect of changing the object position.

We compare the results of these two strategies for the axial

position of the best focus plane in Fig. 16 (left). The agreement between the data points is a good indicator of the accuracy of our experiments.

Validation: We use all the accumulated experimental data to validate our ECA model. Since it is physically challenging to perfectly know the position of the main lens with respect to the sensor, we do not have information to forward-simulate our ECA model generating identical plots to superimpose with the experimental results.

Instead, we perform a global fit of the experimentally attained properties to their corresponding analytical equations. This fit delivers a list of coefficients which can be interpreted as the physical unknown specifications of our setup. The global fit distributes the fit errors among all chosen properties.

The least squares minimization problem for the system parameters \mathbf{p} is of the following form:

$$\min_{\mathbf{p}} \|g(\mathbf{p})\|_2^2 = \min_{\mathbf{p}} \left(\sum_i \alpha_i g_i(\mathbf{p})^2 \right), \quad (11)$$

with $g_i(\mathbf{p}) = Data_i - Model_i(\mathbf{p})$ the error function for the property i , and α_i being a weighting coefficient. The properties $Model_i(\mathbf{p})$ used in this global fit are the ECA camera position in X, Y and Z, the view direction and the Z position of the best focus plane. The results for the fit are shown together with the experimental data in the preceding Figs. 15 and 16.

The output coefficients \mathbf{p} of the global fit are the parameters of the main lens of the system: the focal length f_H (25.21mm), the thickness L_H (-58.3mm), its absolute distance to the sensor $f_H + Z_0$ (26.4mm from the back principal plane for $\Delta Z = 0$) and the alignment shift between the sensor and the micro-lens array t_p . However, this last parameter is lost in the fit due to the re-centering of the data for the view direction and the ECA camera position in X and Y. We use 3 independent pseudo parameters, one for each property to represent it instead.

To culminate our validation, we perform a characterization of the main lens, employing a Shack-Hartmann sensor [33]. This experiment delivers a focal length f_H for the main lens of $23 \pm 2\text{mm}$, which provides a reasonable error with respect to our estimation from the fit. Measuring the absolute smallest distance of the main lens to the sensor $f_H + Z_0$ (i.e. at $\Delta Z = 0$) is a challenging experimental endeavor. From our experiment, we can *qualitatively* confirm that the distance obtained from the fit corresponds to the experimental setup since we purposely place the main lens at the physically closest distance to the sensor. The last pseudo parameter t_p has a different value for each property fit and this value is close to zero because of the re-centering of the data.

VII. DISCUSSION

A. Camera array equivalence

If a real camera array were to be constructed with the characteristics of the ECA (position of the centers and shape and size of the apertures) corresponding to a light-field camera, the light-field measured by this real camera array would be the same to the first-order as the one measured by the light-field camera. One difference between the real and the virtual

array is that the main lens front plane is possibly at a different position than the virtual lens plane where the real array must be placed. In case where the array position is in front of the main lens as in Fig. 6, a physical array would not be able to see an object lying between the main lens front plane and itself, whereas a light-field camera can also measure this part of the object space.

The other notable difference with a real array is that the abstraction made with the two-aperture model implies a loss of the optical properties of the apertures. As such, the effect of refraction is ignored and consequently, the focal length of the cameras of the ECA is unknown. Moreover, the real pixel pitch and sensor position are also unknown. These parameters characterize the image side of the cameras and cannot be predicted by the ECA model. Actually, all of these parameters are linked together and are parameterized by the focal length value which is free to choose. The constraints are fixed by the position and pitch of the pixels in object space given by the ECA. In the end, multiple camera arrays can be made having the same sampling as the ECA.

Finally, the main condition to obtain an equivalent camera array is to create a virtual camera from the grouping of several two aperture elements sharing the same aperture (Section IV-C). The center of this common aperture is considered to be the center of projection of the virtual camera. For imaging systems that do not maintain the condition of having a common aperture, the ECA of the system does not exist. As an example, imaging systems that use components that create non-perspective views [9] or that are too disordered [34] break the condition.

B. Notes on the focused/afocal comparison

The two configurations provide two distinct solutions for the spatio-angular resolution trade-off linked to the arrangement of the virtual cameras. From the baseline and angle of view properties, the ECA of the focused configuration is made of many cameras with a small angle of view, each looking at a different location of the focus plane. On the contrary, the ECA in the afocal configuration is very similar to a physical array of cameras. Each camera has a large angle of view and they all observe the same part of the focus plane. The focused configuration has a lower number of cameras seeing a common region of space and so its angular resolution is less than for the afocal configuration. This distinction was described as an improvement to the afocal case to retrieve lost spatial resolution [6].

The previous simulation and analysis did not take into account the influence of the vignetting effect, where some finite apertures in the system are blocking the theoretical path of light rays, resulting in cropped two-aperture elements. Vignetting reduces the spatial and angular resolution of the system so the depth of field and the angle of view of outer cameras in the ECA are most affected.

C. Limitations of a first-order model

The equivalent camera array model is a first-order model that can accurately predict the properties of a light-field camera

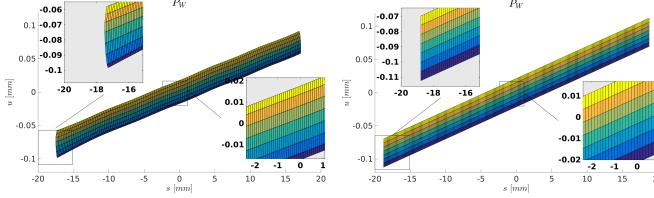


Fig. 17: (Left) Experimental phase space at the plane P_W . Inaccuracies of the ray estimation, stemming from corner detection and ray fitting, provoke the undulations observed in the light-field cells. (Right) Corresponding approximate simulated phase space.

if the Gauss conditions are respected. The apertures of the optical elements must remain small as well as the angles of the rays with respect to the optical axis. Using wider apertures goes with an increase of the effect of aberrations and degrades the quality of the measured light-field. This implies that light rays that in the ideal setting pass through the center of perspective of a given virtual camera, now do not converge to a single point. Non-converging rays correspond to a deformation of the phase space as displayed in the experimental phase space plotted in Fig. 17 (left). In contrast to the simulated phase space, the cells are no longer aligned in the plane P_W , as the aberrations redirect the ray bundle, distorting the phase space. This fact is more prominent on the edges of the subviews, corresponding to the left and right sides of Fig. 17 (left). The simulated phase space is only provided by the parameters obtained from the fit, as again, we lack some physical specifications to accurately model the real light-field camera. A future higher order model may be able to adequately describe the deformations of the phase space and correlate them with optical aberrations.

VIII. CONCLUSION

We propose a model based on constructing an Equivalent virtual Camera Array (ECA) to describe the characteristics of a light-field camera. Our model abstracts the physical components into a pair of apertures that can be grouped to define an individual virtual camera. We can quantify most imaging properties for each virtual camera with the exception of focal length, for which we can only determine a family of solutions. There is an analytic equation for each derived property which is parameterized by the back focal length of the main lens. Camera properties can be computed for the distances comprised by the scene volume, which makes them a useful tool for optical design.

We validated our model through experimentation with a real light-field camera. In the absence of a genuine ground truth, we fitted our model to the experimental data, extracting system parameters that were checked by alternative methods. This validates our model for the light-field camera selected. A similar procedure, as the one demonstrated in the paper, can be used without any difficulty for other camera architectures.

Our model is a first-order optical model constructed from simple ray geometry. However, aberrations are very common in all imaging devices as we demonstrated for the Lytro

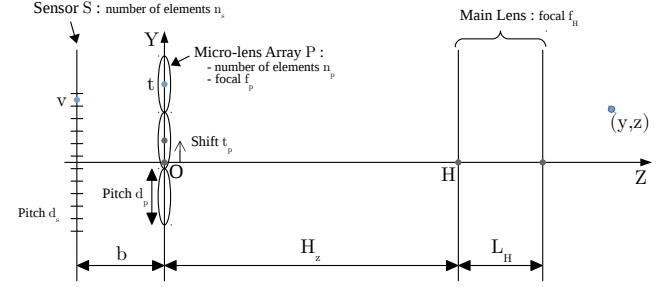


Fig. 18: Sketch of the light-field system indicating the used notation.

camera. Future work will focus on a theory of aberrations for light-field cameras.

APPENDIX EQUATIONS OF THE PROPERTIES

In this section, we list the analytical expressions used to plot the properties of the light-field cameras. The notations for the system parameters are defined in Fig. 18. The derivation can be found in the supplemental material.

A. Focused case

To simplify and shorten the equations, we introduce the following adimensional parameters:

$$\alpha = \frac{L_H}{f_H}, \gamma = \frac{H_z}{f_H}, \beta = \frac{b}{f_p}, \delta = \frac{f_p}{f_H}.$$

We now list the equations that describe the imaged planes in object space.

a) Positions:

$$\mathbf{P}_w = \begin{pmatrix} -\frac{t_p - d_p (\frac{n_p}{2} - t + \frac{1}{2})}{\gamma - 1} \\ f_H \left(\alpha + \frac{\gamma^2}{\gamma - 1} \right) \end{pmatrix}.$$

$$\mathbf{Q}_w = \begin{pmatrix} \frac{\beta (t_p - d_p (\frac{n_p}{2} - t + \frac{1}{2})) + d_s (\frac{n_s}{2} - v + \frac{1}{2})}{\beta \delta - (\beta - 1)(\gamma - 1)} \\ f_H \left(\alpha + \gamma - \frac{\beta - 1}{\beta \delta - (\beta - 1)(\gamma - 1)} + 1 \right) \end{pmatrix}.$$

b) Pitches: The pitch of the apertures of the micro-lens and the pixel image in object space are respectively:

$$D_p = -\frac{d_p}{\gamma - 1} \text{ and } D_s = -\frac{d_s}{\beta \delta - (\beta - 1)(\gamma - 1)}.$$

We now list the equations obtained for the ECA properties.

c) View direction:

$$VD = -\arctan \left(\frac{t_p - d_p (\frac{n_p}{2} - t + \frac{1}{2})}{f_H} \right)$$

d) Angle of view:

$$AOV = \arctan(A + B) - \arctan(A - B),$$

$$\text{where } A = \frac{t_p - d_p \left(\frac{n_p}{2} - t + \frac{1}{2} \right)}{f_H} \text{ and } B = \frac{d_p (\gamma - 1)}{2 \beta \delta f_H}.$$

e) Depth of field:

$$\text{DOF} = -\frac{2\beta\delta d_p f_H d_s}{d_s^2(\gamma-1)^2 - d_p^2(\beta\delta - (\beta-1)(\gamma-1))^2}.$$

f) Baseline: For a point $(z, 0)$ on the optical axis, the baseline is:

$$\text{Bln} = -\frac{d_p \left((\gamma-1) \left(\alpha + \gamma - \frac{z}{f_H} + 1 \right) + 1 \right)}{\beta\delta (\gamma-1) \left(\alpha + \gamma - \frac{z}{f_H} + 1 \right)}.$$

g) Accuracy: For a point $(z, 0)$ on the optical axis, the transversal and longitudinal accuracies are respectively:

$$\text{Trans. Acc.} = -\frac{d_s \left(\left(\alpha - \frac{z}{f_H} \right) (\gamma-1) + \gamma^2 \right)}{\beta\delta},$$

and

$$\text{Long. Acc.} = \frac{2 f_H d_s ((\gamma-1) C + 1) C}{d_p^2 - d_s^2 (\gamma-1)^2 C^2},$$

where $C = \left(\alpha + \gamma - \frac{z}{f_H} + 1 \right)$.

B. Afocal case

h) Positions:

$$\begin{aligned} \mathbf{P}_w &= \begin{pmatrix} -\frac{t_p - d_p \left(\frac{n_p}{2} - t + \frac{1}{2} \right)}{\gamma-1} \\ f_H \left(\alpha + \frac{\gamma^2}{\gamma-1} \right) \end{pmatrix}. \\ \mathbf{Q}_w &= \begin{pmatrix} \frac{t_p + d_s \left(\frac{n_s}{2} - v + \frac{1}{2} \right) - d_p \left(\frac{n_p}{2} - t + \frac{1}{2} \right)}{\delta} \\ f_H (\alpha + \gamma + 1) \end{pmatrix}. \end{aligned}$$

i) Pitches:

$$D_p = -\frac{d_p}{\gamma-1} \text{ and } D_s = -\frac{d_s}{\delta}$$

j) View direction:

$$\text{VD} = -\arctan\left(\frac{\delta t_p + (\gamma-1) D}{\delta f_H}\right),$$

where $D = t_p + d_s \left(\frac{n_s}{2} - v + \frac{1}{2} \right) - d_p \left(\frac{n_p}{2} - t + \frac{1}{2} \right)$.

k) Angle of view:

$$\text{AOV} = \arctan\left(E - \frac{n_p d_p}{2 f_H}\right) - \arctan\left(E + \frac{n_p d_p}{2 f_H}\right),$$

where $E = \frac{\delta t_p + (\gamma-1) D}{\delta f_H}$.

l) Depth of field:

$$\text{DOF} = \frac{2\delta d_p f_H d_s}{d_s^2(\gamma-1)^2 - \delta^2 d_p^2}.$$

m) Baseline: For a point $(z, 0)$ on the optical axis, the baseline is:

$$\text{Bln} = -\frac{n_p d_p \left(\alpha + \gamma - \frac{z}{f_H} + 1 \right)}{\left(\alpha - \frac{z}{f_H} \right) (\gamma-1) + \gamma^2}.$$

n) Accuracy: For a point $(z, 0)$ on the optical axis, the transversal and longitudinal accuracies respectively are:

$$\text{Trans. Acc.} = d_p \left(\alpha + \gamma - \frac{z}{f_H} + 1 \right),$$

and

$$\text{Long. Acc.} = \frac{2 n_p f_H \left(\left(\alpha - \frac{z}{f_H} \right) (\gamma-1) + \gamma^2 \right) C}{\left(\left(\alpha - \frac{z}{f_H} \right) (\gamma-1) + \gamma^2 \right)^2 - n_p^2}.$$

ACKNOWLEDGMENT

This work was supported by the German Research Foundation (DFG) through Emmy-Noether fellowship IH 114/1-1 and the ANR ISAR project.

REFERENCES

- [1] G. Lippman, “Épreuves réversibles photographies intégrales,” *CR Acad. Sci.*, vol. 146, pp. 446–451, 1908.
- [2] E. H. Adelson and J. Y. A. Wang, “Single lens stereo with a plenoptic camera,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 2, pp. 99–106, 1992.
- [3] J. C. Yang, M. Everett, C. Buehler, and L. McMillan, “A real-time distributed light field camera,” in *Proceedings of the 13th Eurographics workshop on Rendering*. Eurographics Association, 2002, pp. 77–86.
- [4] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, “High performance imaging using large camera arrays,” *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 765–776, 2005.
- [5] 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*, vol. 2, no. 11, 2005.
- [6] A. Lumsdaine and T. Georgiev, “The focused plenoptic camera,” in *Computational Photography (ICCP), 2009 IEEE International Conference on*. IEEE, 2009, pp. 1–8.
- [7] R. Ng, “Digital light field photography.” Ph.D. dissertation, Stanford University, Department of Computer Science, July 2006.
- [8] K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar, “Picam: An ultra-thin high performance monolithic camera array,” *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 166, 2013.
- [9] D. Lanman, D. Crispell, M. Wachs, and G. Taubin, “Spherical catadioptric arrays: Construction, multi-view geometry, and calibration,” in *3D Data Processing, Visualization, and Transmission, Third International Symposium on*. IEEE, 2006, pp. 81–88.
- [10] Y. Taguchi, A. Agrawal, A. Veeraraghavan, S. Ramalingam, and R. Raskar, “Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering,” *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2010)*, vol. 29, no. 6, pp. 172:1–172:8, Dec 2010.
- [11] J. Unger, A. Wenger, T. Hawkins, A. Gardner, and P. Debevec, “Capturing and rendering with incident light fields,” DTIC Document, Tech. Rep., 2003.
- [12] T. Georgiev and C. Intwala, “Light field camera design for integral view photography,” Adobe System, Inc, Tech. Rep., 2003.
- [13] K. Ueda, D. Lee, T. Koike, K. Takahashi, and T. Naemura, “Multi-focal compound eye: liquid lens array for computational photography,” in *ACM SIGGRAPH 2008 new tech demos*. ACM, 2008, p. 28.
- [14] A. Manakov, J. Restrepo, O. Klehm, R. Hegedüs, H.-P. Seidel, E. Eisemann, and I. Ihrke, “A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging,” *ACM Trans. on Graphics (SIGGRAPH’13)*, vol. 32, no. 4, pp. 47:1 – 47:14, 2013.
- [15] C. Thomason, B. Thurow, and T. Fahringer, “Calibration of a microlens array for a plenoptic camera,” in *52nd Aerospace Sciences Meeting: American Institute of Aeronautics and Astronautics*, 2014, p. 0396.
- [16] D. G. Dantereau, O. Pizarro, and S. B. Williams, “Decoding, calibration and rectification for lenselet-based plenoptic cameras,” in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 1027–1034.

- [17] O. Johannsen, C. Heinze, B. Goldluecke, and C. Perwass, "On the calibration of focused plenoptic cameras," in *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*. Springer, 2013, pp. 302–317.
- [18] C. Heinze, S. Spyropoulos, S. Hussmann, and C. Perwass, "Automated robust metric calibration of multi-focus plenoptic cameras," in *Instrumentation and Measurement Technology Conference (I2MTC), 2015 IEEE International*. IEEE, 2015, pp. 2038–2043.
- [19] T. E. Bishop, S. Zanetti, and P. Favaro, "Light field superresolution," in *Computational Photography (ICCP), 2009 IEEE International Conference on*. IEEE, 2009, pp. 1–9.
- [20] T. E. Bishop and P. Favaro, "The light field camera: Extended depth of field, aliasing, and superresolution," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 5, pp. 972–986, 2012.
- [21] C.-K. Liang and R. Ramamoorthi, "A light transport framework for lenslet light field cameras," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 2, p. 16, 2015.
- [22] I. Ihrke, J. Restrepo, and L. Mignard-Debise, "Principles of light field imaging," *IEEE Signal Processing Magazine*, no. 5, pp. 59–69, 2016.
- [23] C. Perwass and L. Wietzke, "Single lens 3d-camera with extended depth-of-field," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 829 108–829 108.
- [24] T. Feng and H.-Y. Shum, "An optical analysis of light field rendering," in *Proceedings of Fifth Asian Conference on Computer Vision*, 2000, pp. 394–399.
- [25] C. Hahne, A. Aggoun, S. Haxha, V. Velisavljevic, and J. C. J. Fernández, "Light field geometry of a standard plenoptic camera," *Optics Express*, vol. 22, no. 22, pp. 26 659–26 673, 2014.
- [26] R. Ng and P. Hanrahan, "Digital correction of lens aberrations in light field photography," vol. 6342, 2007, pp. 63 421E–63 421E–14. [Online]. Available: <http://dx.doi.org/10.1117/12.692290>
- [27] L.-Y. Wei, C.-K. Liang, G. Myhre, C. Pitts, and K. Akeley, "Improving light field camera sample design with irregularity and aberration," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, p. 152, 2015.
- [28] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996, pp. 31–42.
- [29] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen, "Programmable aperture photography: multiplexed light field acquisition," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, p. 55, 2008.
- [30] "Lytro, inc." <https://www.lytro.com/>, accessed: 2016-10-01.
- [31] M. D. Grossberg and S. K. Nayar, "The raxel imaging model and ray-based calibration," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 119–137, 2005. [Online]. Available: <http://dx.doi.org/10.1023/B:VISI.0000043754.56350.10>
- [32] B. Acheson, F. Heide, and W. Heidrich, "Caltag: High precision fiducial markers for camera calibration," in *VMV*, vol. 10. Citeseer, 2010, pp. 41–48.
- [33] B. C. Platt and R. Shack, "History and Principles of Shack-Hartmann Wavefront Sensing," *Journal of Refractive Surgery*, vol. 17, no. Sept/Oct, pp. 573–577, 2001.
- [34] A. Wender, J. Isenberghausen, B. Goldluecke, M. Fuchs, and M. B. Hullin, "Light field imaging through household optics," in *Vision, Modeling & Visualization*. The Eurographics Association, 2015, pp. 159–166.

Ivo Ihrke is currently a researcher at Carl Zeiss Inc. Before, he was a permanent researcher at Inria Bordeaux Sud-Ouest (2013–2016), an Emmy-Noether fellow (DFG) and headed a research group within the Cluster of Excellence Multimodal Computing and Interaction at Saarland University that was also associated with the MPI Informatik and the Max-Planck Center for Visual Computing and Communications (2010–2012). He spent his postdoc as a Humboldt fellow at the University of British Columbia, Vancouver, Canada (2008–2009). He received a MS degree in Scientific Computing from the Royal Institute of Technology (KTH), Stockholm, Sweden (2002) and a Ph.D. (summa cum laude) in Computer Science from Saarland University (2007). He co-organized the CVPR workshop on Computational Cameras and Displays in 2012, a Dagstuhl seminar on Computational Imaging in 2015 and the Zeiss Symposium Workshop on Computational Imaging 2016.

Loïs Mignard-Debise is currently a Ph.D. student in the Manao team at INRIA Bordeaux Sud-Ouest, France. He received a MS degree in Optics from the Institut d'Optique Graduate School, Palaiseau, France (2014).

John Restrepo received his Physics Engineering degree in 2008 at the Universidad Nacional de Colombia, Medellin campus in Colombia. He subsequently received his Masters degree in Physics, later in 2010 at the same university. He continued his research work with an internship of one and half years at the Cluster of Excellence "Multimodal Computing and Interaction" at Saarland University in Germany. Thereafter in 2013 and up to this date, he pursues his Doctorate in Computer Science at the Université de Bordeaux in France, working at the Inria Bordeaux Sud-Ouest research institute.