Light-field image acquisition from a conventional camera: design of a four minilens ring device

Stephane Bazeille, Yvan Maillot, Frédéric Cordier, Cécile Riou, Christophe Cudel

Light field image acquisition from a conventional camera: design of a four mini-lens ring device

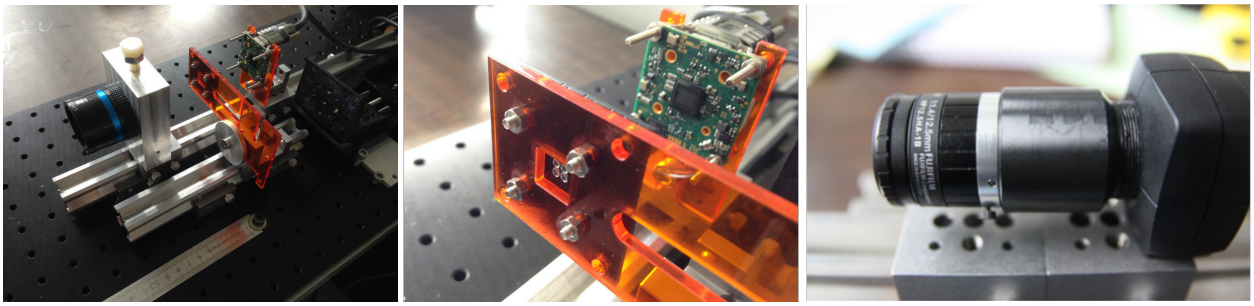
Stéphane Bazeille[a,*], **Yvan Maillot**[a], **Frédéric Cordier**[a], **Cécile Riou**[a], **Christophe Cudel**[a]
[a]Institute IRIMAS - Université Haute-Alsace, 61 rue Albert Camus, 68093, Mulhouse Cedex, France.

Abstract. Since few years, a new type of camera is emerging on the market: a digital camera capable of capturing both the intensity of the light emanating from a scene and the direction of the light rays. This new camera technology called light-field camera uses an array of lenses placed in front of a single image sensor, or simply, an array of cameras attached together. In this paper, a new optical device is proposed: a 4 mini-lens ring which is inserted between the lens and the image sensor of a digital camera. This device prototype is able to convert a regular digital camera into a light-field camera as it makes it possible to record 4 sub-aperture images of the scene. It is a compact and cost-effective solution to perform both post-capture refocusing and depth estimation. The mini-lens ring makes also the plenoptic camera versatile; it is possible to adjust the parameters of the ring so as to reduce or increase the size of the projected image. Together with the proof of concept of this new device, we propose a method to estimate the positions of each optical component depending on the observed scene (object size and distance) and the optics parameters. Real-world results are presented to validate our device prototype.

Keywords: Optics, light field, camera. *Stéphane Bazeille stephane.bazeille@uha.fr

1 Introduction

With the development of optical sensors and small powerful computers, 3D imaging is gaining more and more interest over the last decades. 3D imaging sensors are usually classified into two different modalities: active and passive. The principle of active sensors is to emit a light flow and measure its reflections from the scene; examples of active sensors are time-of-flight cameras or structured-light cameras. Unlike active sensors, passive sensors do not require any additional light sources; instead, the measurements are made using the ambient light. Examples of passive sensors are stereo cameras or light-field cameras. The light-field camera has been introduced by Lippmann.[1] The idea consists in capturing more information about the light emanating from a scene: the intensity and wavelength of the light and the direction of the light rays. Standard cameras are not able to capture the directional information. A light-field camera allows generating a series of different viewpoints of the same scene simultaneously called sub-aperture images.[2] It



(a) (b) (c)

Fig 1 The camera prototype. In (a) is shown the 3 components of the system: the camera lens (also identified as the main lens) on the right, the mini-lens ring (black square attached to the horizontal base in orange colour) in the middle, the image sensor (attached to the vertical base in orange colour) on the left. The image sensor is fixed and the lenses are able to slide along their rails to adjust the projected image. In (b) are shown a zoom of the mini-lens ring and the image sensor. The integration of the 4 the mini-lenses into a camera ring extension is shown in (c); the mini-lens is mounted in a tube which is placed between the main lens and the image sensor)

enables performing a wide range of processing such as depth estimation, post-capture digital refocusing, post-capture focal length or aperture adjustment, perspectives shift, image segmentation, video stabilization and material classification.[2,3] Until now, active systems (structured-light and laser technologies) have been the most widely-used technology for 3D visual inspection because of their accuracy and ease of use. However, with the emergence of light-field cameras, passive systems are gaining more interest since they offer new alternatives for accurate 3D measurements. For these reasons, light field cameras are increasingly used in a large variety of applications such as: digital photography, microscopy, robotics or machine vision. In this paper, we present a device, namely the mini-lens ring extension (see Fig. 1), that can be added to any standard digital camera to convert it into a light-field camera. This device bridges the gap between stereo vision systems and the first light field device (see Fig. 2). It is composed of a main lens and 4 mini-lens with a focused plenoptic configuration (2.0). This device is very simple to use, cost-effective and compact. In addition, its optical properties are modifiable so that it can be adapted to a wide range of scenes with different dimensions. This makes our device particularly suitable for automatic defect detection.

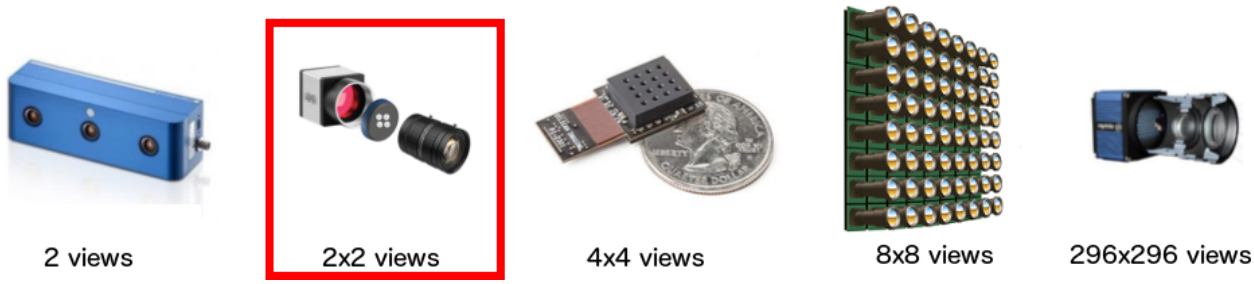


Fig 2 Different multi-views cameras. (a) stereo camera ENSENSO, (b) the mini-lens ring inserted between the image sensor and the main lens of a standard camera (c) Pelican imaging camera array, (d) Stanford multi camera array, (e) Raytrix light field camera.

The remaining of the paper is organized as follows: in Section 2 is given the state of the art in the domain of light-field devices; in Section 3, a description of our device and its mathematical model are given. In Section 4, the functioning of the device is demonstrated with a simulation and with real-world data. In Section 5, refocused images are generated with our device and compared with the ones obtained using the Raytrix camera. The last section concludes the paper and gives several suggestions for future work.

2 Related work

Light-field technology is an area of interest since many years, leading to the publication of several survey papers.[3–5] Several types of light-field cameras have been proposed in the literature. The first type is based on a temporal multiplexing and is composed of a single camera equipped with coded aperture mask placed in front of the main lens to capture different viewpoints sequentially. The second type makes use of spacial multiplexing to capture different points of view. This can be done using a camera array[6] or a lens array placed in front of a single image sensor.[7] In case of the plenoptic cameras, a micro-lenses array is inserted between the main lens and the image sensor of a camera. Two different configurations have been proposed for plenoptic cameras. The first one also

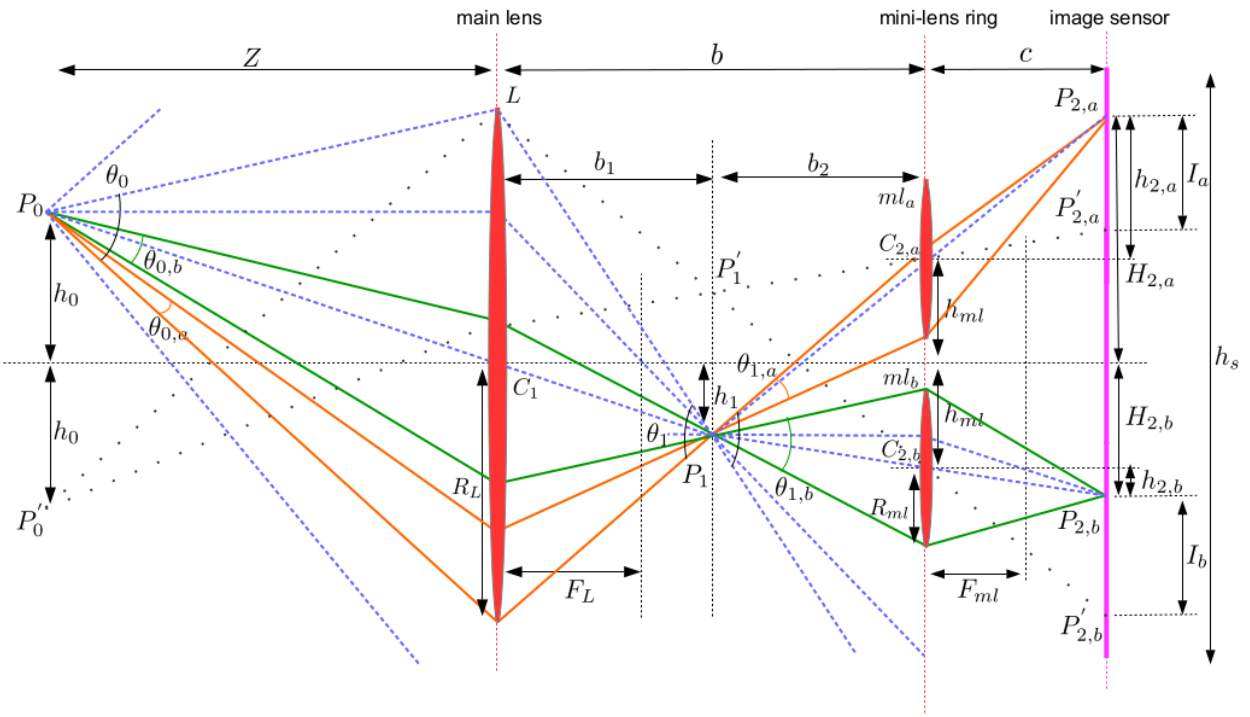


Fig 3 A 2D representation of our light-field camera. The point P_0 is located at the height h_0 with respect to the optical axis of the main lens and at the distance Z. P_0 is imaged on P_1 through the main lens L. Then, P_1 is imaged on $P_{2,a}$ through the mini-lens ml_a (orange rays) and on $P_{2,b}$ through the mini-lens ml_b (green rays). Our method estimates b and c such that the 2 sharp images of the object fit within the surface area of the camera sensor.

called configuration 1.0 is described by Adelson et al.[8] The micro-lens array is placed in the focal plane of the main lens. As a consequence, the main lens concentrates the light rays on the micro-lenses; depending on their orientation with respect to the optical axis, the rays are then distributed over the areas of the image sensor. The second configuration is described by[9] following the work of.[10] This model is called focused plenoptic camera or 2.0 because the raw image is composed of a series of focused sub-aperture images, each of these image being generated by a micro-lens. Other prototypes have also been proposed, such as the kaleidoscopic cameras, where several points of view are obtained using a tube of mirrors.[11] The raw data collected from light-field cameras can be used to produce different types of images: sub-aperture images that are a collection of synthetic points of view,[12] full focused images, stacks of refocused images, epipolar images and disparity images. Also, several algorithms have been proposed to perform depth estimation[13–16] using the

data from the light-field cameras.

In this paper, we present a light-field camera that belongs to the focused plenoptic configuration. The main drawback of these devices is the low spatial resolution that limits the use and applicability of the camera and the narrow baseline that reduces the depth estimation range and accuracy. The main reason is that every plenoptic camera suffers from the fundamental tradeoff between spatial and angular resolutions. Commercial products[17] or[18] based on this technology are (or were) available on the market but their main drawback is their cost. In this paper, we propose a new device: a simple 4 mini-lens ring that allows recording the light-field from a regular camera (see Fig. 1). On the market, this system can be classified between stereo systems with 2 views and the pelican imaging light field camera (see Fig. 2). Instead of having a large number of micro-lenses that project small sub-images of the scene such as,[19] our device is composed only of 4 mini-lenses with a large pitch. One important advantage of choosing a small number of lenses is that the device has a large baseline, which, in turn, provides a better depth resolution and the ability to record in one picture different modalities of the scene by adding polarimetric filters.[20] Unlike conventional light-field systems, this camera has also a larger spatial resolution (sub-aperture images of $\approx$1060x1060 pixels). This is an advantage because these raw sub-aperture images are readily usable by standard image processing (matching, tracking, etc.). On the other hand, our system has a smaller angular resolution which limits the refocusing capabilities and induces artefacts. Although the angular resolution is smaller compared to[18] or to the plenoptic ring proposed in,[19] we will show in this paper that performing post-capture refocusing calibration in depth with only 4 sub-aperture images is possible. This was proved in[21] where a simple calibration algorithm based on the properties of the variable homography was used to produce refocused images at a specific depth without any feature extraction or feature matching. The main advantage of this method is that the computation cost

is very small and thus the processing can be done on embedded computers (no GPU is required). Such refocused images calibrated in depth can be used efficiently in industrial visual applications such as highlighting a working plane, detecting measure defects in images, as it is illustrated in.[22] Another advantage is that the distance between the ring and the main lens and the distance between the ring and the image sensor are adjustable, which makes the device adaptable to different sizes and locations of objects in the scene.

3 Numerical modeling of the camera

3.1 System overview

The camera is composed of 2 components which are: the main lens (L in Fig. 3) with a focal length F_L and a radius R_L, the image sensor I whose dimensions are $h_s \times w_s$ with nb pixel and a ring of 4 mini-lens (see Fig. 1.b). These 4 mini-lenses are identical; their focal lengths and radius are noted F_{ml} and R_{ml} respectively. The vertical and horizontal distances between the centers of the mini-lens is $2h_{ml}$. The horizontal distance between the main lens and the mini-lens ring is b, the distance between the ring and the image sensor is c (see Fig. 3). For the sake of simplicity, only 2 of the 4 mini-lens are represented in this figure. A ray is originating from the point P_0 (highest point of the object located at a height h_0 and a distance Z from the main lens). P_0 is imaged in P_1 by the main lens (P_1 has a height h_1 and is located at a distance b_1 from the main lens). Then P_1 is imaged twice on the image sensor in $P_{2,a}$ and $P_{2,b}$ by the mini-lens. The portion of the rays emanating from P_0 and passing through the upper and lower mini-lenses are depicted in orange and green respectively.

3.2 Optical model assumptions

We assume that all the lenses are thin lenses i.e. the distance along the optical axis between the 2 lenses is negligible compared to the radius of curvature of these lenses. Thus, all computation is made based on the paraxial ray approximation. The paraxial ray approximation implies the following properties:

- all incoming rays parallel to the optical axis are directed through the lens focal point,
- all rays emanating from the focal point become parallel to the optical axis after passing through the lens,
- the rays that pass through the center of the lens do not change their direction,

The thin lens approximation gives the following relationship between the distances Z and b_1, b_1 being the distance between the lens and the main lens image plane,

$$\frac{1}{F_L} = \frac{1}{Z} + \frac{1}{b_1} \qquad \Longleftrightarrow \qquad b_1 = \frac{Z.F_L}{Z - F_L} \tag{1}$$

Also, the relationship between the height h_0 of the point P_0 and the height h_1 of its image P_1 is:

$$\frac{h_0}{Z} = \frac{h_1}{b_1} \tag{2}$$

Similar equations are derived for the 2 mini-lenses ml_a and ml_b:

$$\frac{1}{F_{ml}} = \frac{1}{b_2} + \frac{1}{c} \qquad \Longleftrightarrow \qquad c = \frac{b_2.F_{ml}}{b_2 - F_{ml}} \tag{3}$$

Equivalently, the equation for the relationship between the position of P_1 and its image $P_{2,b}$ through the second mini-lens ml_b is:

$$\frac{h_{ml} - h_1}{b_2} = \frac{h_{2,b}}{c} \tag{4}$$

where h_{ml} is the distance between the center of the mini-lens and the optical axis of the main lens. $h_{2,b}$ is the distance between $P_{2,b}$ and the optical axis of the mini-lens.

4 Lens position estimation in simulation

To simplify the experimental procedure as proposed in,[23] we developed a method to automatically estimate the lens positions depending on the scene definition and a ray-tracing simulator in 2D (presented on Fig. 4) to validate our results. The lens positions are represented by the parameters b and c as shown on Fig. 3, and the scene is defined by the object size and the working distance Z. It has to be noted that the diameter of the lenses is not considered in the equations to compute b and c. However, these parameters are taken into account in our ray-tracing simulation so that the results can be fully simulated in 2D before experimenting with the real camera prototype.

4.1 Lens positions estimation only considering the magnification factor

Some of the camera parameters are fixed by the geometric properties of the components chosen by the user. Thus, the estimation of the camera parameters using the simulator requires the user to provide the following information (see Fig. 3):

- the focal length F_L and the radius R_L of the main lens.
- the focal length F_{ml} and the radius R_{ml} of the mini-lenses.
- the width w_s and the height h_s of the image sensor.

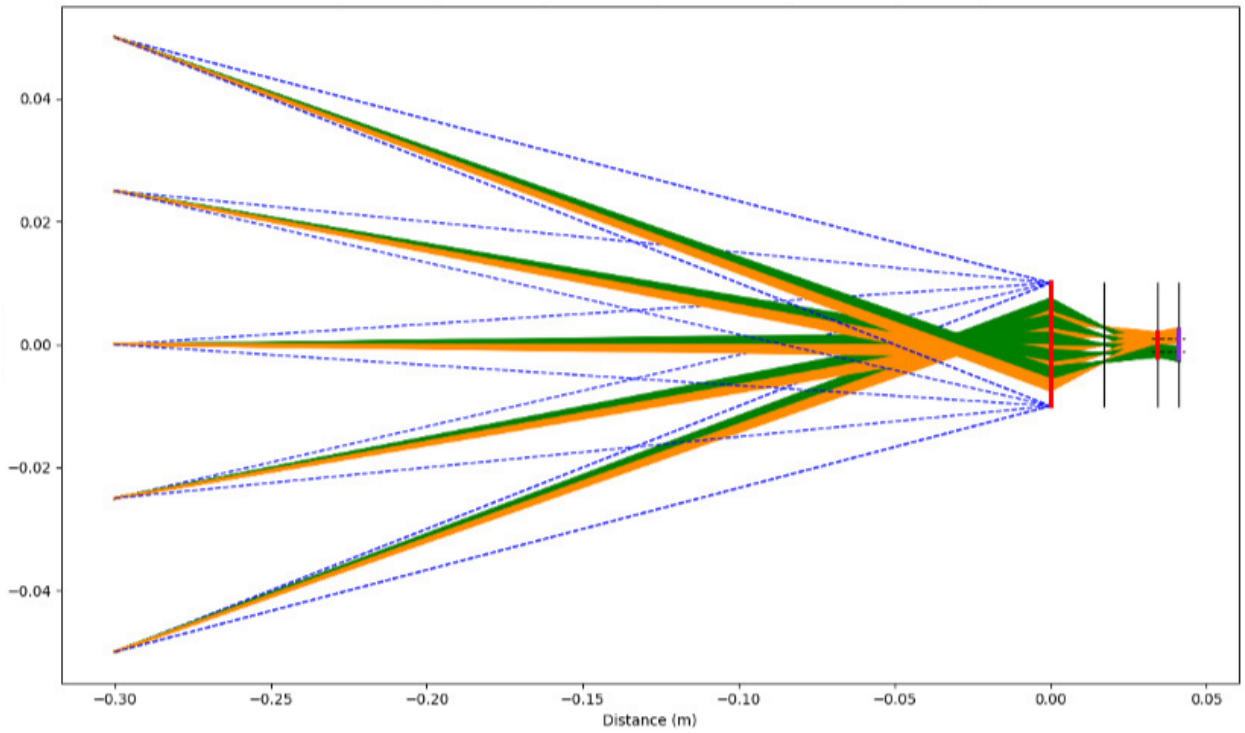


Fig 4 Simulated view of the whole system obtained with our 2D ray-tracing simulator. The parameters in meters are the following: $F_L = 0.016$m, $R_L = 0.021$m, $F_{ml} = 0.005$m, $R_{ml} = 0.001$m, $h_{ml} = 0.00105$m, $h_s = 0.0046$m, $h_0 = 0.05$m, $Z = 0.3$m, and $Ia = h_s/2$. All units are in meters. Orange rays going through the upper mini-lens. Green rays going through the lower mini-lens.

- the object size $2h_0$ and an estimatation of its distance Z to the main lens.

Then, the position b of the mini-lenses and the position c of the image sensor are calculated such that the images captured by the sensor are sharp and the magnification factor of the object is maximum. This magnification factor denoted by G is the ratio between the apparent size of an object on the image sensor and its real size. Its value is defined by:

$$G = G' \times G'' \tag{5}$$

where G' and G'' are the magnification factor of the main lens and the mini-lenses.

$$G' = \frac{h_1}{h_0} = \frac{b_1}{Z} \qquad \text{and} \qquad G'' = \frac{h_{2,b}}{h_1} = \frac{c}{b_2} \tag{6}$$

From Eq. 1 and Eq. 3 we obtain:

$$G' = \frac{F_L}{Z - F_L} \qquad \text{and} \qquad G'' = \frac{F_{ml}}{b_2 - F_{ml}} = \frac{F_{ml}}{b - b_1 - F_{ml}} = \frac{F_{ml}}{b - \frac{Z.F_L}{Z - F_L} - F_{ml}} \tag{7}$$

Hence, the magnification factor is:

$$G = \frac{F_L}{Z - F_L} \times \frac{F_{ml}}{b - \frac{Z.F_L}{Z - F_L} - F_{ml}} = \frac{F_L.F_{ml}}{(Z - F_L)(b - F_{ml}) - Z.F_L} \tag{8}$$

The position b of the mini-lenses is now expressed as a function of F_L, F_{ml}, h_0, Z and the desired magnification factor G:

$$b = \frac{F_L(\frac{F_{ml}}{G} + Z)}{(Z - F_L)} + F_{ml} \tag{9}$$

The position c of the image sensor is computed by using Eq. 3 and by replacing b_2 by $(b - b_1)$:

$$c = \frac{F_{ml}(b - \frac{Z.F_L}{Z - F_L})}{F_{ml} + b - \frac{Z.F_L}{Z - F_L}} \tag{10}$$

The magnification factor G is maximum when the 2 images placed one above the other exactly fit within the height of the image sensor: $I_a = I_b = h_s/2$ (see Fig. 3). In that case, the disparity for the points of the object is zero. The maximum magnification factor G_{max} is given by:

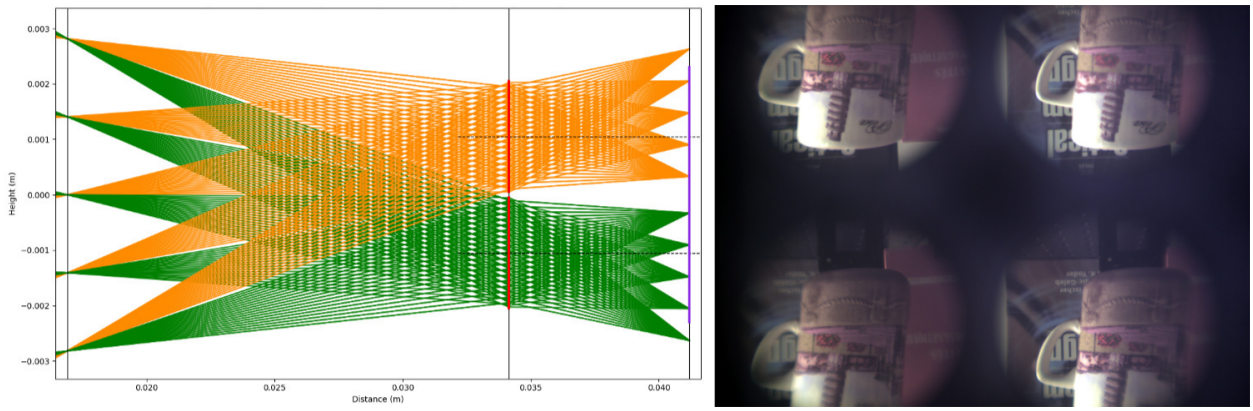
$$G_{max} = \frac{\frac{h_s}{2}}{2h_0} = \frac{h_s}{4h_0} \tag{11}$$

h_0 being half of the height of the real object and h_s the height of the image sensor. The position b of the mini-lenses is then obtained from Eq. 9 by replacing G by G_{max}. With the following

experimental parameters:

- an image sensor with width $w_s = 6.14$mm and height $h_s = 4.60$mm. The resolution is 4912×3684 pixels (IDS UI-3590LE) and the size of a pixel is $1.25\,\mu\text{m}$.
- 4 mini-lenses with focal length $F_{ml} = 5$mm, radius $R_{ml} = 1$mm and whose distance with respect to the optical axis of the main lens is $h_{ml} = 1.05$mm.
- a camera lens LINOS MeVIS-C with a focal length $F_L = 16$mm, a radius $R_L = 21$mm, a flange focal distance $FFD = 17.526$mm and the c-mount screw length $ST = 3.8$mm.
- an object with size $2h_0 = 0.1$mm. The distance between the object and the main lens is $Z = 300$mm.

The values of b and c corresponding to the maximum magnification factor are $b = 34.1$mm and $c = 6.59$mm. The result given by our ray-tracing simulator using these values is shown in Fig. 4. This ray-tracing simulation is computed with 5 source points located at different heights $(-0.05, -0.025, 0, 0.025, 0.05\text{m})$ on the object whose size is $2h_0 = 0.1$m. These 5 points are emitting light in the direction of the main lens; only the rays that pass through the main lens are drawn. In blue, we show the rays that hit the main lens but do not reach the mini-lenses. The lines in green are the rays going through the main lens and the lower mini-lens. The rays going through the main lens and the upper mini-lens are depicted in orange. A zoom of the image sensor of Fig. 4 in shown in Fig. 5.a and one can see that the 5 source points have an image through both mini-lenses. One can observe that the size of the images is equal to the value predicted by Eq. 9 and 11; we have I_a and I_b equal to $\frac{h_s}{2}$. The result of the experiment is presented in Fig. 5.b. It perfectly matches the simulated results, but unfortunately, the size of the object image exceeds the size of



(a) (b)

Fig 5 Results of the first parameter estimation (parameters detailed in Fig. 4 for an object of 0.1m at 0.3m. (a) Zoom on the simulated view, (b) Real image acquired with our experimental prototype.

the sensor because the central part of the sensor does not receive any rays. The details about the image acquisition prototype will be presented in the next section.

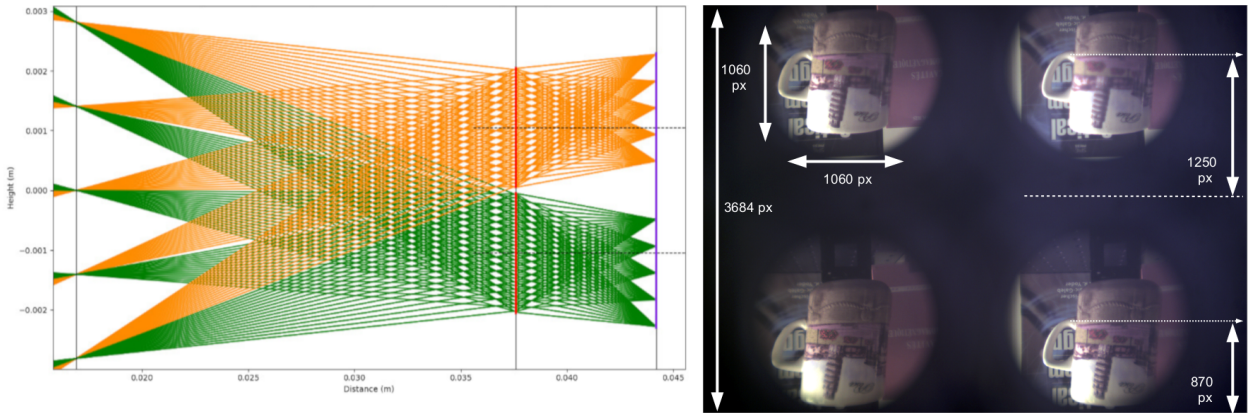
4.2 Estimating the lens positions using the magnification factor and the image position

In the previous estimation, we did not taken into account the position of the image. In order to put the object image inside the image sensor, we have to consider the height of the highest object pixel represented by the point $P_{2,a}$ in Fig. 3. Rewriting the equation presented in Sec. 3.2 for the upper mini-lens in the reference frame defined by the optical axis of the upper mini lens we obtain:

$$\frac{h_1}{b_2} = \frac{h_{2,a}}{c} \tag{12}$$

Then by combining this equation with the Eq. 1 and 2 we have for the upper and lower mini-lens:

$$h_{2,a} = \frac{c(h_{ml} + b_1.\frac{2h_0}{Z})}{b_2} \qquad h_{2,b} = \frac{c(-h_{ml} + b_1.\frac{2h_0}{Z})}{b_2} \tag{13}$$



(a) (b)

Fig 6 Results of the optimized parameter estimation for an object of 0.1m at 0.3m. In this case $I_a = I_b = 0.78 \times h_s$. (a) Zoom on the simulated view, (b) Real image. On this image, we plotted the sub-aperture image size for this particular example.

In the reference frame defined by the optical axis the main lens we have:

$$H_{2,a} = h_{ml} + \frac{c(h_{ml} + b_1.\frac{2h_0}{Z})}{b_2} \qquad H_{2,b} = -h_{ml} + \frac{c(-h_{ml} + b_1.\frac{2h_0}{Z})}{b_2} \tag{14}$$

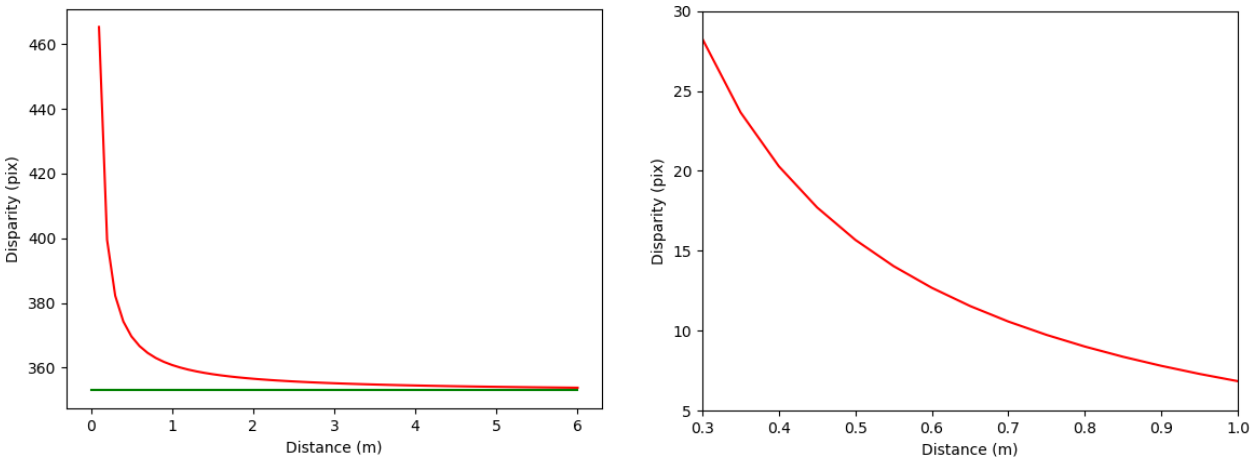
To compute the lens positions and to ensure that the magnification factor is maximum while keeping the image of the object inside the image sensor we need $H_{2,a} = h_s/2$ in Eq. 14. From Eq. 14 by replacing b_2 by $(b - b_1)$ and replacing b_1 by its value depending on Z and F_L we have:

$$h_{ml} + \frac{c(h_{ml} + b_1.\frac{2h_0}{Z})}{(b - b_1)} = \frac{h_s}{2} \qquad \Longleftrightarrow \qquad b = \frac{c(h_{ml} + \frac{Z.F_L}{Z-F_L}.\frac{2h_0}{Z})}{(\frac{h_s}{2} - h_{ml})} + \frac{Z.F_L}{Z - F_L} \tag{15}$$

We obtain the second result: $b = 37.4$mm, and $c = 6.61$mm which corresponds to the simulation presented in Fig. 6.a and the experiment with our prototype in Fig. 6.b.

4.3 Validation of our first use case: 10cm object at a distance of 0.3m

As shown in Fig. 6 the estimation gives correct values and the experimentation with our prototype confirmed the result obtained in simulation. Note that our system does not compute the amount



(a) (b)

Fig 7 Range of the disparity for the fixed optical parameters $b = 37.4$mm and $c = 6.61$mm, (a) When the distance Z is varying from 0.1m to 6.0m the disparity is around 400 pixels, (b) On the second plot we plotted only the disparity variations (that can be used to extract the depth) for Z varying from 0.3m to 1m.

of light passing through the lenses, and therefore cannot be used to estimate the vignetting effect. Once the size and position of the images have been calculated, we then compute the other parameters of the light-field prototype, which are the disparity and the overlapping ratio. The disparity is used to estimate the accuracy of the 3D information, and the overlapping ratio gives the portion of the image on which the 3D information is available. The disparity d is defined as the difference in pixels between 2 corresponding points in a couple of images acquired with our system (see $H_{2,a}$ and $H_{2,b}$ on Fig. 3). d and o can be computed as follows:

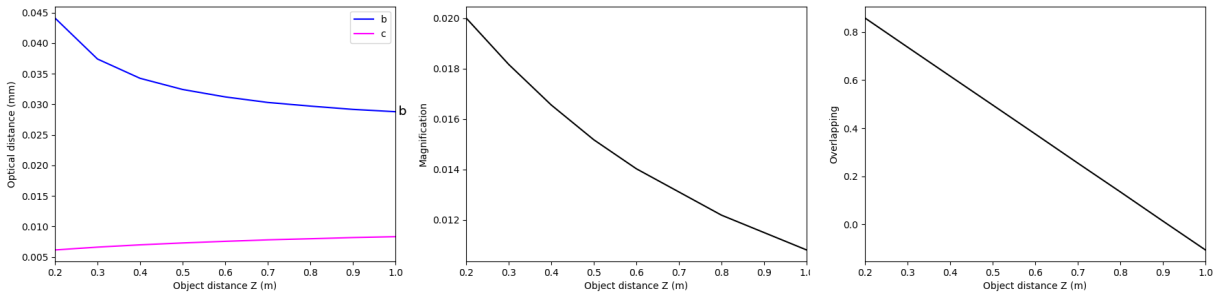
$$d = \frac{H_{2,a} - (\frac{h_s}{2} + H_{2,b})}{t_{pix}} = \frac{1}{t_{pix}}(\frac{4h_{ml} - h_s}{2} + \frac{2h_{ml}.c}{(b - \frac{Z.F_L}{Z - F_L})}) \qquad \text{and} \qquad o = 1 - \frac{d}{I_a} \tag{16}$$

where I_a is the size of the image produced by the upper mini-lens on the image sensor. These equations give $d = 382$ pixels and $o = 0.73$ (i.e. 73%) for the use-case we studied. The vertical size of the image I_a and I_b was equal to $0.78 \times h_s/2$. To complete these results, we estimated from Eq. 16 the range of the disparity for an interval of distances Z while keeping the distances b

and c fixed. This result is shown in Fig. 7. For distance values that range from 0.3m to 1.0m, the range of the disparity is $[382 - 361]$ i.e. slightly larger than 20 pixels for 70cm which seems good to get enough accuracy in depth. The disparity is important because there is a shift between the sub-aperture images explained by the overlapping ratio of 0.7. To modify this overlapping ratio, a new mini-lens ring with a smaller h_{ml} is necessary. In the future, we plan to investigate the effect of modifying h_{ml} in the system. For these experiments, we decided to put the 4 mini-lens as close as possible (0.5mm between the border of 2 mini-lenses) to maximize the portion of the sensor that is receiving light rays.

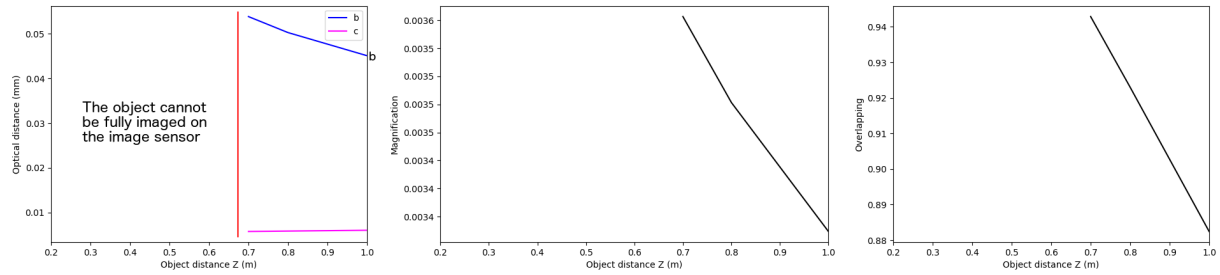
4.4 Example with different objects at different distances

In the previous section, we have given the equations to compute the position b of the mini-lenses and the position c of the image sensor for a given scene (a particular object at a given distance).To illustrate the versatility of our prototype, we show in Fig. 8 another object of size 10cm and placed at different distances with respect to the main lens. For each distance, we estimated the distances b and c and the corresponding magnification factor and overlapping ratio. One can observe that the magnification factor and overlapping ratio decrease when the distance between the main lens and the object increases. The reason is that the size of the camera sensor prevents the image to go up. In this case, to keep the entire image of the object inside the sensor area, the magnification has to be lowered. In Fig. 9 is shown the same result for a larger object (i.e. size 60cm). One can observe that the object image does not fit within the sensor area if its distance to the main lens is smaller than 0.7m.



(a) (b) (c)

Fig 8 An object of size 0.1m placed at different distances with respect to the main lens. (a) Values of b and c (b) Magnification factor (c) Overlapping ratio.



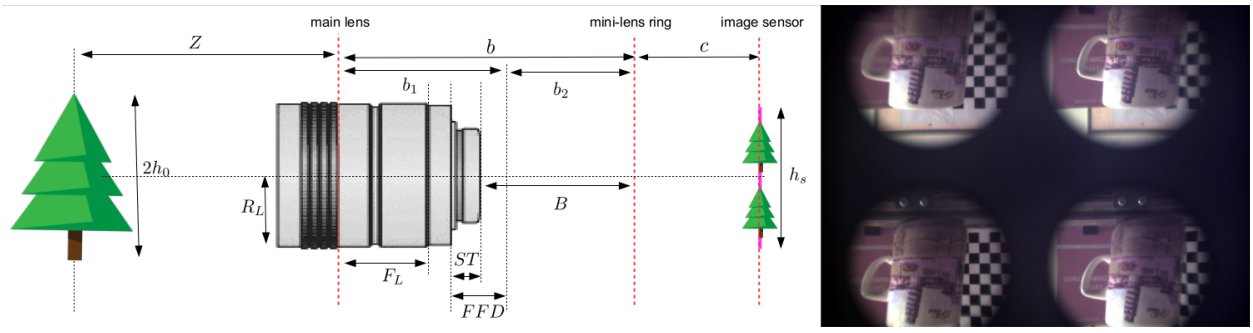
(a) (b) (c)

Fig 9 An object of size 0.6m placed at different distance with respect to the main lens. (a) Values of b and c (b) Magnification factor (c) Overlapping ratio.

5 Experiments with a camera prototype

5.1 Experimental setup

Our first experimental setup is presented in Fig. 1. It is composed of 2 rails on which the main lens and the mini-lens ring are able to slide forward and backward. The image sensor and the main lens are mounted on the first rail and the mini-lens ring on the parallel rail. The image sensor is attached to the first rail. The mini-lens ring is sliding on its rail; the position of this ring is precisely controlled with a micro-metric screw. Note that it is difficult to accurately measure the position of the center of the camera lens and therefore it is impossible to accurately set b. To overcome this problem, we measure the distance between the back of the camera lens and the image sensor. This



(a) (b)

Fig 10 (a) Illustration of the flange focal distance (FFD) and the screw thread length (ST), (b) Image obtained with our experimental prototype. The scene is composed of a coffee cup at 0.3m, a book at 0.5m and a chessboard at 1m.

distance is called B (see Fig. 10.a) and is defined as:

$$B = b_2 + FFD - ST = \frac{c.F_{ml}}{c - F_{ml}} + 0.017526 - 0.00038 = 0.0376\text{m} = 37.6\text{mm} \tag{17}$$

where FFD is the flange focal distance and ST the length of the screw thread. To demonstrate the efficiency of our simulation we have placed the object, the main lens and the mini-lens ring at the distances calculated in the previous section and the image has been recorded. In Fig. 5.b and 6.b, we show the 2 images obtained with our light-field prototype corresponding to the simulation results presented in Fig. 5.a and 6.b. The results from the experimental setup closely match the parameters estimated with our simulation program. However, the image produced by our camera prototype is reversed compared to our simulation (see Fig. 3). This is because the electronic circuit associated with the sensor generates an inverted image by default so that the image produced by a camera with a single lens is upright. In the remaining of the paper, the image is rotated.

5.2 Refocusing at a specific depth using a 4 sub-aperture images

In this section, we demonstrate post-capture refocusing using the method proposed in.[22] In Fig. 10.b is shown a raw image recorded with our light-field camera with a scene composed of 3 objects

(a) (b) (c) (d)

Fig 11 Refocusing experiments, (a) Refocusing on the Pisa Tower, (b) Refocusing on the chessboard, (c) Zoom on the Pisa tower in the refocused image, (d) Zoom on the chessboard in the refocused image.

located at different depths: a coffee cup in the foreground (at 0.3m from the main lens), a red book in the middle ground (at 0.5m) and a chessboard in the background (at a distance of about 1m). The distances b and c are the distances estimated previously for a distance $Z = 0.3$. Using this 4 sub-aperture images, we performed 2 different calibrated refocusing: the first on the foreground (see 11.a) and the second one on the background (see Fig. 11.b). Fig. 11.c and Fig. 11.d are a zoom of Fig. 11.b and Fig. 11.b respectively. These 2 examples show that this camera prototype allows performing numerical refocusing. However, it can be seen that the quality of refocused images is not perfect; this is due to our camera prototype design: the 4 mini-lenses inside the ring are not perfectly aligned and the components of the optical device (main lens, mini-lens ring, image sensor) are not exactly parallel and centered to the same axis. Finally, we have used mini-lenses with a value N of f-number of 2.5 (the f-number of an optical system is the ratio of the system's focal length to the diameter of the entrance pupil). This value is close but above the limit of diffraction defined by $d = 2.44.\lambda N = 3\,\mu\text{m}$. Even by activating the binning mode to record images, the value is still above the limit of diffraction with a pixel size of $1.25\,\mu\text{m}$.

5.3 Comparison with a Raytrix camera

This section provides a comparison between the refocusing with the 4 mini-lens ring and the refocusing using the Raytrix camera. The experiment with the 3 depth planes (coffee cup, book,

(a) (b) (c)

Fig 12 Results obtained with a Raytrix camera with a main lens focal length of 50mm. (a) Raw image, (b) Focusing on the background, (c) Focusing on the foreground.

chessboard) that has been done with the 4 mini-lens ring has been redone using the Raytrix camera model R5 with focal length of 50mm. The results are shown in Fig. 12. Fig. 12.a shows the raw image generated by the Raytrix camera. One can observe the artefacts caused by the micro-lens grid and which are particularly visible on the cup. They are mainly due to the limited angular resolution offered by our system. Using the Raytrix camera software, 2 refocused images have then been generated, the first image (Fig. 12.b) with a focus on the foreground (the cup) and the second one (Fig. 12.c) on the background (the chessboard). As one may observe, the quality of rendering is higher for the Raytrix camera than for the 4 mini-lens ring. In the future, we plan to improve the rendering quality by increasing the resolution. On the other hand, the quality of refocusing using the 4 mini-lens ring with the method by[22] is good. This is an advantage given the fact the 4 mini-lens ring is much cheaper and easy to use since it only requires 2 images at 2 known distances (see[22] for more details).

6 Conclusion and future work

A new device to convert any digital camera into a light-field camera has been described. Usually, light-field cameras are classified in 2 groups: the camera array and the plenoptic models. The cameras array only gives 4 points of views with large aperture angle. On the other hand, plenoptic

cameras which are composed of a micro-lens array placed between the main lens and the image sensor, produce a series of sub-images representing a small part of the scene due to the small aperture angle. Even if the light field is captured in both cases, these approaches leads to different methods of reconstructions. In case of the camera array, the raw data are exploited by using the "plane and parallax" approaches, while more complex algorithms must be used to process each sub-view of the plenoptic devices. In this paper, we proposed a different approach which consists of adding a plenoptic ring to transform a standard camera into a plenoptic one but keeping the advantage of the camera array (i.e. larger baseline and larger spatial resolution). Together with the proof of concept of this new device, a method is proposed to estimate the positions of each optical component depending on the working distance and the object size set by the user. Real-world results recorded with our 4 mini-lens ring prototype are presented. In the future, we plan to improve the design and increase the resolution of the 4 mini-lens ring by using a camera with a larger image sensor and larger pixel size. We will also make some experiments using mini-lenses with larger diameter. This should help to decrease the diffraction effects and improve the image quality of the 4 mini-lens ring.[24] We also plan to develop algorithms to compute depth images using the data from the 4 mini-lens ring.

6.1 Acknowledgments

This work has been funded by the french state, FUI project TRAC, grant number DOS0068206/00.

References

1 G. Lippmann, "Épreuves réversibles donnant la sensation du relief," J. Phys. Theor. Appl. **7**(1), 821–825 (1908).

2 E. Y. Lam, "Computational photography with plenoptic camera and light field capture: tutorial," J. Opt. Soc. Am. A **32**, 2021–2032 (2015).

3 G. Wu, B. Masia, A. Jarabo, et al., "Light field image processing: An overview," IEEE Journal of Selected Topics in Signal Processing **11**, 926–954 (2017).

4 M. Levoy, "Light fields and computational imaging," Computer **39**, 46–55 (2006).

5 I. Ihrke, J. Restrepo, and L. Mignard-Debise, "Principles of light field imaging: Briefly revisiting 25 years of research," IEEE Signal Processing Magazine **33**, 59–69 (2016).

6 B. Wilburn, N. Joshi, V. Vaish, et al., "High performance imaging using large camera arrays," ACM Trans. Graph. **24**, 765–776 (2005).

7 A. Levin, R. Fergus, F. Durand, et al., "Image and depth from a conventional camera with a coded aperture," ACM transactions on graphics (TOG) **26**(3), 70 (2007).

8 E. H. Adelson and J. Y. A. Wang, "Single lens stereo with a plenoptic camera," IEEE Transactions on Pattern Analysis and Machine Intelligence **14**, 99–106 (1992).

9 A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in IEEE International Conference on Computational Photography (ICCP), 1–8 (2009).

10 R. Ng, Digital light field photography. PhD dissertation, Standford University (2006).

11 I. Reshetouski, A. Manakov, H. P. Seidel, et al., "Three-dimensional kaleidoscopic imaging," in IEEE conference on computer vision and pattern recognition (CVPR), 353–360 (2011).

12 A. Isaksen, L. McMillan, and S. Gortler, "Dynamically reparameterized light fields," in Proceedings of the ACM SIGGRAPH Conference on Computer Graphics, 297–306 (2000).

13 H. Jeon, J. Park, G. Choe, et al., "Accurate depth map estimation from a lenslet light field

camera," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1547–1555 (2015).

14 M. W. Tao, S. Hadap, J. Malik, et al., "Depth from combining defocus and correspondence using light-field cameras," in IEEE International Conference on Computer Vision, 673–680 (2013).

15 S. Wanner and B. Goldluecke, "Globally consistent depth labeling of 4d light fields," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 41–48 (2012).

16 J. P. Lke, F. Rosa, J. G. Marichal-Hernndez, et al., "Depth from light fields analyzing 4d local structure," Journal of Display Technology **11**, 900–907 (2015).

17 "Lytro." [https://en.wikipedia.org/wiki/Lytro]. Accessed: 2018-07-01.

18 "Raytrix." [https://raytrix.de]. Accessed: 2018-07-01.

19 I. Montilla, M. Puga, J. P. Lke, et al., "Design and laboratory results of a plenoptic objective: From 2d to 3d with a standard camera," Journal of Display Technology **11**, 73–78 (2015).

20 L. Gendre, S. Bazeille, L. Bigue, et al., "Interest of polarimetric refocused images calibrated in depth for control by vision," in SPIE Photonics Europe, (2018).

21 C. Riou, B. Colicchio, J.-P. Lauffenburger, et al., "Calibration and disparity maps for a depth camera based on a four-lens device," Journal of Electronic Imaging **24**, 1–11 (2015).

22 C. Riou, B. Colicchio, J. Lauffenburger, et al., "Interests of refocused images calibrated in depth with a multi-view camera for control by vision," in International Conference on Quality Control by Artificial Vision (QCAV17), (2017).

23 C. Hahne, A. Aggoun, S. Haxha, et al., "Light field geometry of a standard plenoptic camera," Opt. Express **22**, 26659–26673 (2014).

24 T. Georgiev, "Plenoptic camera resolution," Imaging and Applied Optics 2015 (2015).

Stéphane Bazeille is an Assistant Professor at Université Haute-Alsace in France. He received his Master's Degree in mathematics and computer science from the University of La Rochelle in 2005 and his Ph.D. degree in perception for robots from the University of Western Britany in 2008. He is the author of more than 30 scientific papers in international journal and conferences. His current research interests include vision systems, computer vision, robot perception and artificial intelligence.

Yvan Maillot is a an Assistant Professor at Université Haute-Alsace in France. He received his Ph.D. degree in automatic and computer science from the University of Franche-Comté in 1997. He is the author of several papers in international journals and conferences. His main research interest is Computational Geometry including Shape Reconstruction.

Frédéric Cordier is an Associate Professor at Université Haute-Alsace. His research interests include 3D reconstruction, geometric modelling and shape processing. He has published more than 25 papers in the major journals in Computer Graphics. He is an associate editor of The Visual Computer journal (Springer).

Cécile Riou received her Ph.D. degree from Université Haute-Alsace in France in 2017. She worked during her thesis on the designs of light-field vision systems. She received her MS degree in optics, image and vision from University Jean Monnet, Saint Etienne and her engineering degree from Telecom Saint-Etienne, in 2013. Since February 2018, she is working as an engineer on optronic test benches with several application fields such as cameras, imaging, radiometry and optical design.

Christophe Cudel received his Ph.D. from the University of Reims Champagne Ardenne in 1995. He joined the Université Haute-Alsace and the laboratory MIPS in 1997, first as assistant professor and now as professor. His researches are focused on computer vision, eye-tracking, light-field and multi-view cameras. He teaches computer vision for industrial control, robotics and signal processing.