Plenoptic sensor: application to extend field-of-view

Benoît Vandame, Valter Drazic, Matthieu Hog, Neus Sabater

Plenoptic sensor: application to extend field-of-view

Benoît Vandame Valter Drazic Matthieu Hog Neus Sabater
Technicolor R & I
975 Avenue des Champs Blancs, 35576 Cesson-Sévigné, France
[benoit.vandame, valter.drazic, matthieu.hog, neus.sabater]@technicolor.com

Abstract

In this paper we study the light field sampling produced by ideal plenoptic sensors, an emerging technology providing new optical capabilities. In particular, we leverage its potential with a new optical design that couples a pyramid lens with an ideal plenoptic sensor. The main advantage is that it extends the field-of-view (FOV) of a main-lens without changing its focal length. To evince the utility of the proposed design we have performed different experiments. First, we demonstrate on simulated synthetic images, considering ideal and real lenses, that our optical design effectively doubles the FOV. Then, we show its feasibility with two different prototypes using plenoptic cameras on the market with very different plenoptic samplings, namely a Raytrix R5 and a Canon 5D MarkIV. Arguably, future cameras with ideal plenoptic sensors will be able to be coupled with pyramid lenses to extend its inherent FOV in a single snapshot.

1. Introduction

Plenoptic cameras are able to capture the Light-Field (LF), thanks to a micro-lens array (MLA) placed between the main-lens and the sensor. Depending on the MLA position, plenoptic cameras are divided in type-1 [19] (e.g. Lytro [1]) and type-2 or focused [17] (e.g. Raytrix [3]).

The first plenoptic cameras on the market have the particularity that the MLA is not well aligned with the sensor. Indeed, plenoptic camera manufacturers assemble individual optic components producing an unavoidable rotation offset between the MLA and the pixel matrix. For this reason, many research works in the field have focused on calibration and decoding methods [10, 15], as well as image processing algorithms taking into account such misalignments [7, 22, 12]. However, camera manufacturers rely on wafer-level fabrication to assemble micro-optical components like MLA's onto pixel matrix with great accuracy.
Thus, plenoptic cameras with the MLA ideally aligned with respect to the sensor shall be available in the near future Such *ideal plenoptic sensors* (simply called *plenoptic sensors* in the sequel) provide a new paradigm for LF processing since no camera calibration is required and sub-aperture images (SAI's) or Epipolar Plan Images (EPI's) are simply extracted without interpolation.

In fact, very simple plenoptic sensors already exist in the consumer market. This is the case of dual-pixels in high-end smartphones such as the Samsung S7 [8], and DSLR (Digital Single Lens Reflex) cameras such as the Canon 5D MarkIV. These devices provide a limited angular sampling of the LF comparable to the right and left views of a stereo camera. Maintaining a zero offset between these two views controls the autofocus of the camera.

Wide-FOV imaging is achieved by stitching multiple images that are recorded from the same center of projection [28]. Stitched images produce a better spatial resolution, but the parallax between the views produce artifacts on the resulting image. This issue is addressed by the LF panorama stitching [30]. Unfortunately, both strategies require a sequential capture, thus dedicated to static scenes. Changing the main-lens for a wider FOV lens is another solution, but these lenses are often bulky and require larger stack of lenses to produce sharp images. Alternatively, monocentric lenses [25, 26] have become increasingly popular for gigapixel imaging [9]. With these lenses, the light is collected on a spherical surface either with a curved sensor or a fiber coupling interface to a flat sensor. This approach has been pushed further using a plenoptic camera [11]. Finally, combining a prism array with common cameras has been proposed to double [6] or quadruple [27] the FOV of the main-lens. In [24], prisms and mirrors are combined for stereo capture out of a single lens camera.

Our contributions In this paper we analyze the advantages and constraints of plenoptic sensors. In particular, we focus on the so-called quad-pixel sensor where a micro-lens covers 2×2 pixels. In our study, we describe the relationship between the SAI's and the corresponding portions of the main-lens pupil through which light rays have travelled.

Besides, we propose an optical design for a single lens camera that doubles the FOV of the lens combining a

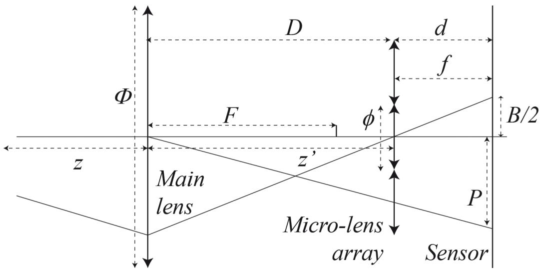


Figure 1: Schematic view of a type-1 plenoptic camera.

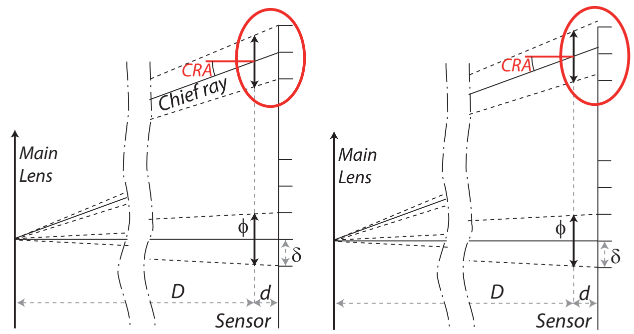


Figure 2: Type-1 plenoptic camera with $N = 2$. Left: symmetric sampling of the micro-images when $\phi = 2\delta/e$. Right: asymmetric sampling of the micro-images when $\phi = 2\delta$ (considering $e = 1$).

plenoptic sensor and a pyramidal lens made of four prisms. The main idea is that the prisms deviate the photons entering the main-lens creating four distinct views that can be demultiplexed into the SAI's thanks to the plenoptic sensor. Then, with a single snapshot, the stitching of the SAI's increases the FOV up to a factor of two in each direction without changing the main-lens focal length, which is an unprecedented capability. Furthermore, the use of plenoptic sensors simplifies the parameterization and the processing of the captured LF compared to existing designs in the literature. Our experiments include synthetic image simulations and real images captured with two different prototypes we have constructed.

2. From plenoptic cameras to plenoptic sensor

In this paper we focus on type-1 plenoptic sensors for its capacity to sample the main-lens exit pupil. Indeed, type-1 plenoptic cameras [19] are characterized by the fact that the distance d between the MLA and the sensor is equal to the micro-lenses focal length f (as illustrated in Fig. 1). In that case, considering a thin lens model, the micro-lenses focus at infinity. Now, the main-lens is considered at the micro-lenses optical infinity since ($D \gg d$). So, the micro-lenses are imaging the main-lens exit pupil which is equal to the lens aperture in a thin lens model. The image on the pixels underneath the micro-lens is called micro-image.

Captured micro-images with $N \times N$ pixels, $N \in \mathbb{N}$, correspond to sampling the main-lens exit pupil with a $N \times N$ grid. For instance, if $N = 2$, each pixel of the 2×2 micro-image integrates the light rays passing through one of the four quarter-discs of the aperture stop.

It turns out that the sharp sampling of the exit pupil with the same pixel grid for all micro-images requires the MLA to be perfectly aligned with the sensor, this is the same squared lattice and same orientation than the pixel matrix. By contrast, in type-2 or focused plenoptic cameras ($d \neq f$) micro-lenses focus at a plane that does not match with the main-lens exit pupil [13].

Regarding the optical design of type-1 plenoptic cameras, it is imposed that the F-numbers of the main-lens and micro-lenses are equal. In that case, the system is said to be aperture matched and it guarantees the micro-images to cover as many pixels as possible without overlapping. Indeed, using the Thales theorem, the distance P between two consecutive micro-images (in physical unit) is equal to $P = \phi e$, where $e = 1 + d/D$. Tuning the main-lens F-number, F/Φ, the diameter P of the micro-images is set to be equal to $P \approx \phi$.

2.1. Pupil sampling with the micro-images

Ideally the distance P between two micro-images should be equal to an integer number of pixels. This is, $P = N\delta$, with δ being the pixel size. This requires the micro-lenses to have a pitch $\phi = N\delta/e$ which is slightly smaller than N pixels and is function of the main-lens distance D. Fig. 2 illustrates the micro-images position with respect to the pixel array at the center of the sensor and at the border.

It is worth mentioning that sensor manufacturers use micro-lenses to guide photons to the middle of each pixel where the photo-diode is located. Recently, manufacturers have also designed a matrix of pixels mounted with a MLA having a pitch ϕ slightly smaller than the pixel size δ [23]. It corresponds to the case $N = 1$ with all the micro-images centered on the middle of the corresponding pixels. In other words, the chief rays, i.e. the rays passing through the exit pupil and the micro-lens centers, hit the photo-diodes despite the increasing Chief Ray Angle (CRA), which is the angle between a chief ray and the optical axis of a micro-lens. The so-called CRA correction decreases the pixel vignetting [20]. This fact emphasizes how ready are manufacturers to produce quad-pixels plenoptic sensor with CRA correction.

2.2. Study of SAI's on the Plenoptic Sensor

Collecting the SAI's is simple thanks to the integer size $N \times N$ of the micro-images. Let $L(x, y)$ be the image cap-

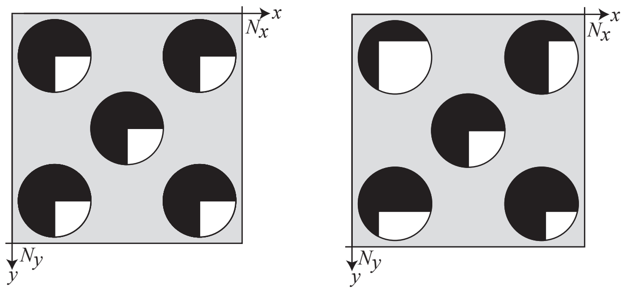


Figure 3: Left: ideal portion of the exit pupil being visible by the SAI $S_{0,0}$ when $P = 2\delta$. This portion is constant independently of the sensor coordinate (x, y). Right: portion of the exit pupil visible by $\widehat{S}_{0,0}$ for a camera with $\phi = 2\delta$. The visible portion is asymmetric and depends on the position of the micro-image on the sensor as well as e.

tured by the sensor with $(x, y) \in [0, N_x[\times[0, N_y[$. Then, the SAI's $S_{i,j},\ [i, j] \in [0, N[^2$ are obtained by simple demultiplexing:

$$S_{i,j}(k,l) = L\left(\left\{\left\lfloor\frac{x}{N}\right\rfloor + i\right\} \bmod N,\ \left\{\left\lfloor\frac{y}{N}\right\rfloor + j\right\} \bmod N\right) \quad (1)$$

where $(k,l) = (\lfloor x/N \rfloor, \lfloor y/N \rfloor) \in [0, N_x/N] \times [0, N_y/N]$. If the micro-lens pitch ϕ is not exactly equal to $N\delta/e$, then P is not an integer number of pixels, and computing the SAI's with Eq. 1 is incorrect. Indeed, by definition SAI's require to collect pixels at a fixed distance from the micro-image centers. So, if $P \neq N\delta$, SAI computation requires interpolation to extract pixels at non-integer coordinates from the sensor image [10]. Such interpolation averages micro-image pixels, mixing the angular information encoded in the micro-images. To prevent it, Eq. 1 is nevertheless used to extract approximate SAI's that we note $\widehat{S}_{i,j}$, even if P is not a multiple of δ. Such approximation amounts to interpolate with nearest neighbors instead of a more sophisticated method.

It is interesting to point out that the micro-images sample the main-lens exit pupil with a constant partition when $P = N\delta$ regardless of the micro-image position on the sensor. On the contrary, when $P \neq N\delta$, the approximate SAI's do not sample the main-lens exit pupil homogeneously. Fig. 3 illustrates the portion of the pupil visible by the SAI $S_{(0,0)}$ of a quad-pixel ($N = 2$) type-1 plenoptic sensor.

2.3. Considering real-lenses

Until now, we have modelled the main-lens with a thin lens model. However, considering the real-lens is mandatory to understand how the main-lens exit pupil is sampled. The real-lens pupil, also named the aperture stop, is the

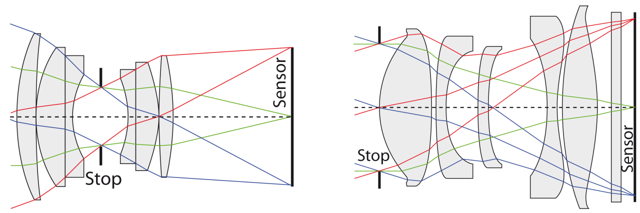


Figure 4: Typical main-lens objective for: single lens reflex camera (left), and smartphone (right).

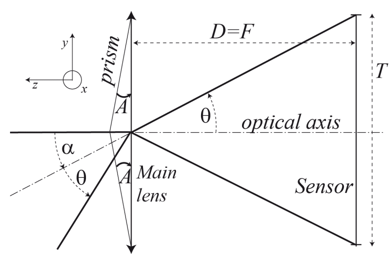


Figure 5: Sensor and main-lens mounted with two prisms.

physical stop delimiting the beam of photons entering the camera. The aperture stop is located where a diaphragm can reduce homogeneously the amount of light on the sensor. Most lenses designed for DSLR cameras have an aperture stop roughly located within the main-lens (as illustrated by a typical double-Gauss lens in Fig. 4-left). On the contrary, for smartphones, the aperture stop is mostly located at the entrance of the main-lens [18] (see Fig. 4-right), even though smartphones do not have diaphragms. In this paper, we consider main-lenses where the aperture stop is positioned at the first diopter on the opposite side of the sensor.

With real-lenses, the distance D is the distance between the so called exit pupil and the MLA. The position of the exit pupil is located at the imaging plane of the aperture stop by the rear lenses and depends on the individual lenses located between the aperture stop and the sensor.

3. Doubling the FOV

Extending the FOV with a pyramid lens With a common camera, θ the half FOV of an image is given by the focal length F of the main-lens and the physical size T of the sensor: $\theta = \arctan \frac{T}{2F}$.

To extend the FOV, a pyramid lens (*i.e.* four prisms assembled together) is placed at the main-lens aperture stop. The FOV per portion of the pupil is rotated in different orientations.

Each prism deviates the photons entering the main-lens by an angle α with respect to the x axis. α is chosen to be equal to the half FOV angle θ of the main-lens. With the two prisms of Fig. 5, the sensor records the superposition of two images, each one imaging a different part of the main-lens exit pupil. These two superposed images, combined together, double the FOV in the vertical direction of the camera. Similarly, using four prisms in front of the main-lens, each one covering a quarter of the main-lens aperture stop, doubles the FOV of the camera in vertical and horizontal orientations.

The angle α associated with the prism is function of its angle A and the refraction index n of its material. A good approximation gives $\alpha \approx (n-1)A$. Since α is set to be equal than θ, the prism angle A is easily computed knowing the refraction index of the prism material.

Extending the FOV with a plenoptic sensor
Discriminating the two or four images that have been summed at the sensor plane is not easy. It requires to know which prism the photon has crossed. This is done with a plenoptic sensor with an ideal pupil sampling. Considering a quad-pixel sensor, the four SAI's are stitched to produce a single image that has a double FOV, as many pixels as the sensor and a single exposure time. Note that the equivalent f-number of the stitched image is divided by two, since SAI's collect only a quarter of the incoming photons.

4. Image simulation

Simulated images are generated with PBRT [21] which has been extended to support real lenses. Our extension permits to define: a thin lens model, real-lenses with their interfaces (each one is defined with a mathematical model of its curvature and a glass material similar to [16]), an array of prisms and a MLA. In particular, a ray is cast within a cone which is defined by an apex located on the sensor and a disk-shaped base which is defined by the exit pupil diameter of the main-lens and is located at the exit pupil position from the sensor. PBRT defines 5D random coordinates: 2D for the sensor coordinates (cone apex), 2D for the exit pupil coordinates (within the cone apex) and 1D for the wavelengths. Using the Snell-Descartes law, a ray is refracted at each interface along its path.

4.1. Ideal Lens simulation

Primarily, synthetic images are generated with the ideal thin lens model (for both main- and micro-lenses). The plenoptic camera characteristics are defined in Table 1. The sensor size is $T = 3.6 \times 3.6mm^2$ as typically found in smartphones. By design ($P = \phi e$), the size of the micro-images is strictly equal to 2×2 pixels. A test chart of colored letters and numbers is located at $z = 2m$ from the camera. The main-lens produces a sharp image on the micro-lenses, D is computed using the thin lens equation. Four prisms forming a pyramid are positioned at the main-lens aperture stop. Each prism is defined by a material with a constant refraction index of $n = 1.74$ and an angle of $A = 16.7°$.

$F = 6.16mm$	$\Phi = 3.09mm$	$D = 6.18mm$
$f = 4.8\mu m$	$\phi = 2.3981\mu m$	$d = f$
$\delta = 1.2\mu m$	$N_x N_y = 3000^2$	$T = 3.6mm$

Table 1: Ideal plenoptic camera characteristics.

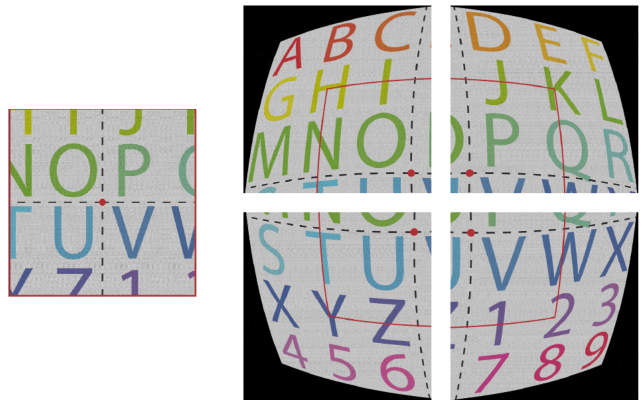


Figure 6: Left: One of the four identical SAI's when there is no prism. Right: Different simulated SAI's obtained with a pyramid lens. The red dot on the left and red square on the right of the test chart indicate respectively the optical axis of the main-lens.

Fig. 6 illustrates the SAI's extracted from the synthetic simulation. Without prisms the four SAI's are identical whereas with the four prisms, each SAI observes a deviated FOV. The observed distortions are due to the prism. Cancelling the constant geometrical distortion enables to stitch the four SAI's. The FOV of the main-lens was 32°, and with the prisms becomes 64° (or 77° considering the hypotenuse of the sensor with a $2/3$ ratio), competing with a typical wide angle lens for smartphones.

It is worth mentioning that the image sharpness of one SAI is quite equivalent to the image sharpness of the main-lens alone. The simulated main-lens is optimized for a field of 32°, beyond that field, the image quality degrades dramatically. This experiment also demonstrates that narrow-angle main-lenses with simple optical design can produce extended FOV with their native resolution.

4.2. Real-lens simulation

Using Zemax optics studio [5], we have designed and optimized a real main-lens made of three biconvex lenses. It has comparable characteristics as the ideal lens simulation shown above. The six interfaces (two interfaces per lens) are presented sequentially starting from the closest interface to the sensor (see Table 2). The shape of an interface is characterized by a central symmetric sag function $z(r)$

Thickness	Radius	Glass	Diameter	Conic
1.82133	3.08301	SK16	3.53718	0.17470
0.07071	-2.6425	air	3.38069	-18.7316
0.43551	-2.1581	F5	3.38110	-11.9393
1.51451	4.72854	air	3.24343	5.13491
3.79133	3.81723	SK16	4.38871	-4.97297
1.03669	6.23189	air	4.73167	-1.297

Table 2: Description of the six interfaces defining the real main lens.

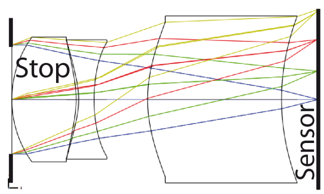


Figure 7: Main-lens with three biconvex lenses.

where the optic axis is presumed to lie in the z direction, and $z(r)$ is the sag (the z-component of the displacement of the surface from the vertex, at distance r from the axis). More precisely, the sag model is an aspheric surface characterized by a radius of curvature R and a conic term K:

$$z(r) = \frac{r^2}{R\left(1 + \sqrt{(1-K)\frac{r^2}{R^2}}\right)} . \tag{2}$$

An interface is fully defined by the thickness between itself to the next one measured on the main optical axis, a curvature radius (in mm), the physical material following the interface, a diameter (in mm) and a conic term.

The refractive index of a material is function of the wavelength. The function is often modeled by the Sellmeier equation made of 6 parameters [2]. In our simulation software, we use this model to estimate the refractive index of a given material for the three color channels RGB.

Fig. 7 shows the six interfaces and the sensor. The interfaces have been computed such that the main-lens aperture stop is located at the first interface where the four prisms should be positioned. The geometrical distortion of this main-lens is within $\pm 0.5\%$. The MLA and sensor are simulated with the same thin lens model than the ideal simulation. The four prisms are modelled with a perfect material with a constant refraction index (thus producing no chromatic aberrations).

The MLA is designed such that SAI's collect photons strictly from a corresponding quarter-pupil of the main-lens. However, with real-lenses, the rays are distorted due to the lens aberrations. As a result, some photons passing by a given quarter-pupil are not recorded by the expected SAI. We call this phenomenon sub-aperture cross-talk. To illustrate it, we simulate a white board located at $z = 2m$ and a mask at the main-lens aperture stop to mask out three of the four quarter-pupils. One expects only one SAI to record photons but we have observed that this is not the case (see Fig. 8). The cross-talk ratio r is the ratio between the maximum flux on another SAI versus the flux on the corresponding SAI. We have measured that $r = 13\%$, which is too high to completely isolate the photons passing from different prisms. To solve this problem, a cross shaped mask is positioned at the aperture stop of the main-lens. The thickness of the mask is set to have a ratio $r \approx 3\%$. The mask width is $1/15$ of the aperture stop diameter, which removes 16% of the photon flux entering the main-lens. Note that the cross-talk ratio depends on the main-lens design. According to our simulations, it seems correlated with the geometrical distortion of the main-lens. Fig. 9 compares two SAI's of a test chart with and without the cross-shaped mask.

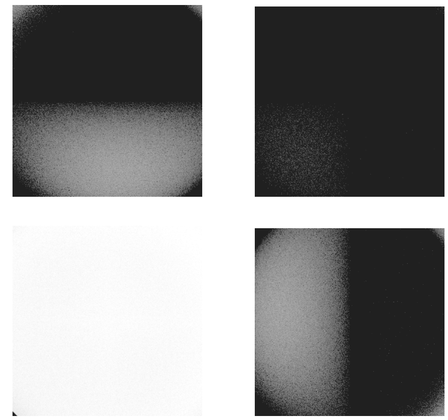


Figure 8: Four SAI's showing cross-talk. Ideally, only the bottom-left SAI would capture photons when three quarter-pupils are masked but other SAI's are contaminated due to lens aberrations.

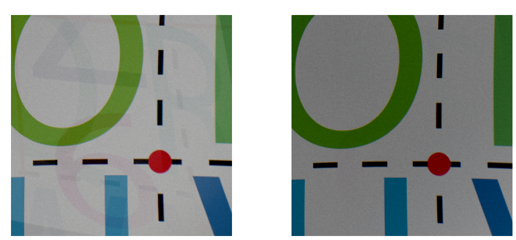


Figure 9: Crop of a SAI with (right) and without (left) the cross-shaped mask. The cross-talk decreases from $r = 13\%$ to $r = 3\%$ with the mask. The ghosting artifacts on the left are not visible on the right (both images have the same dynamic range).

5. Extended FOV with a type-2 camera

We ideally would experiment with a type-1 camera, but the Raytrix camera is the only available plenoptic camera which enables to change the main-lens. Indeed, a main-lens with an aperture stop located at its entrance is required for our experiment. In particular, the experiment we have set up allows to double horizontally the FOV of a main-lens mounted with the Raytrix R5 camera.

5.1. Type-2 pupil sampling

As mentioned before, micro-lenses in type-2 cameras do not focus at the main-lens exit pupil but at a plane S distant by d' from the MLA. Thus, the main-lens exit pupil is not sampled with a disjoint partition. Instead, corresponding pixel visibility areas overlap (see Fig.10). The overlap thickness between two juxtaposed portions is function of the ratio f/d and the pupil diameter Φ.

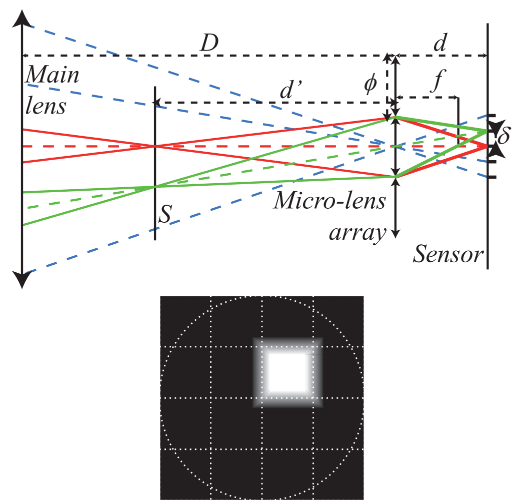
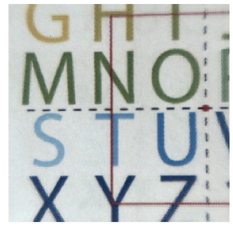


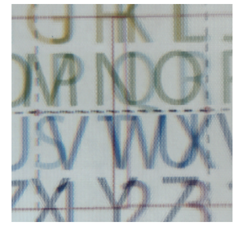
Figure 10: Type-2 plenoptic camera with $P = 4\delta = \phi e$ Top: Scheme in 1D. Red and green lines indicate the ray beams which exit at the border of the pixel of size δ, converging on two points at plane S and diverging on the main-lens. The colored dashed lines delimit the four projected pixels of one micro-lens into the main-lens pupil. Bottom: Pupil sampling in 2D with the portion of the pupil seen by one pixel in white. Fading regions represent visibility overlapping.

5.2. Converting a type-2 into a dual-pixel camera

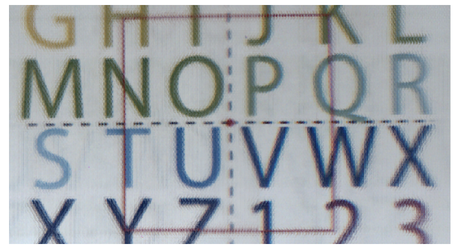
Dual-pixel refers to a plenoptic sensor having 2 rectangular pixels bellow each micro-lens. It enables to capture 2 SAI's. Our experiment converts the Raytrix camera into a dual-pixel camera. Our prototype has only one prism to deviate half of the photons entering the main-lens. The main-lens needs to be sufficiently thin, so the prism can be set as close as possible to it. We assume the prism to be placed at the main-lens aperture stop. Supplementary material illustrates the experiment with the Raytrix camera.



(a) Full pupil refocus without prism.



(b) Full pupil refocus with prism.



(c) SAI stitching.

Figure 11: Raytrix refocus and stitching.

The Raytrix R5 has many pixels per micro-lens ($P = 20.20\,\delta$). To convert it into a dual-pixel camera, stereo refocused images are computed splitting the left and right portions of the micro-images. The two refocused images are comparable to the SAI's extracted from a dual-pixel sensor. We have used the pipeline in [15] to compute the two refocused images from the Raytrix R5. We summarize the main steps bellow:

Micro-lens image calibration The MLA is an hexagonal lattice fully characterized by a radius, a rotation angle (with respect to the pixel matrix) and the offset between the first micro-image center and the origin of the sensor. These values are all computed using the Fourier transform of a white image, this is a shot of a flat white screen homogeneously illuminated. The calibration associates a pixel coordinate (x, y) with a micro-lens coordinate (i, j).

Micro-lens vignetting correction The white image is normalized by its average maximum and inverted to define a flux scaled correction per pixel that fixes the vignetting of the micro-images.

Stereo image refocusing Image refocusing with a type-2 camera is defined with Eq. 3:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = s\left(g\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} x_{i,j} \\ y_{i,j} \end{bmatrix}\right) + \begin{bmatrix} x_{i,j} \\ y_{i,j} \end{bmatrix}\right) \qquad (3)$$

Where (X, Y) is the projection on the refocused image of the sensor pixel (x, y) which belongs to micro-image

(i, j). Micro-image (i, j) is centered at pixel coordinate $(x_{i,j}, y_{i,j})$. g controls the refocalization distance, and s the relative size of the refocused image versus the input LF image. The LF pixel (x, y) is projected at the non-integer coordinate (X, Y). Lanczos interpolation is used to splat the pixel into the refocused image. A weight-map is maintained to count the projected pixels. The refocused image is normalized with the weight-map after all pixels are projected.

The left and right refocused images are computed with Eq. 3 and the following conditions respectively $(x - x_{i,j}) < -m/2$ and $(x - x_{i,j}) > m/2$. Where m is the thickness in pixels of the vertical masks located at the middle of the micro-images to cancel sub-aperture cross-talk.

The refocalization parameters are set to $s = 0.2$ (refocused images are 5 times smaller than the input image) and $g = 4$. The test chart is positioned such that the images appear as sharp as possible. Several prisms with different angles are tested such that the left and right images are almost juxtaposed with a tiny common vertical portion. Experimentally m is set to 2 to cancel the cross-talk.

Extended image stitching The prism produces small distortions and the two refocused images are stitched using a simple horizontal translation to match their two common portions. This process can further be totally automatized.

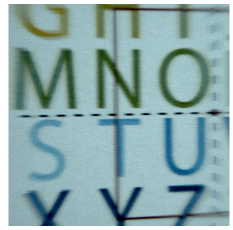
Common refocused images considering all the pixels below the micro-lens are shown at the top of Fig. 11 with and without the prism covering half of the main-lens pupil. The stitched SAI's are visible in Fig. 11c.

6. Application with a dual-pixel camera

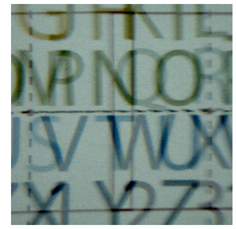
The primary goal of dual-pixel sensor is to perform live auto-focus for video shooting. Recently, with the 5D MarkIV, Canon gives access to the raw dual-pixel. A raw dual-pixel still image contains a first image being the sum of the dual-pixels (*i.e.* the conventional picture) and a second image being the left SAI. Both images are $30.4MPix$. The right SAI is computed by subtracting the second image to the first one.

The raw images are processed with the following steps: 1/ The raw files are read with [4], 2/ The SAI are bias-corrected thanks to over-scan areas which are then cropped to keep the responsive pixels, 3/ A white screen is observed with the same lens used with the Raytrix camera, the two SAI's are median filtered, normalized by their maxima, and inverted to define a flux-scale correction which is applied to all captured images. Due to the poor quality of the main-lens, only a central portion of $2000^2 pix$ is used.

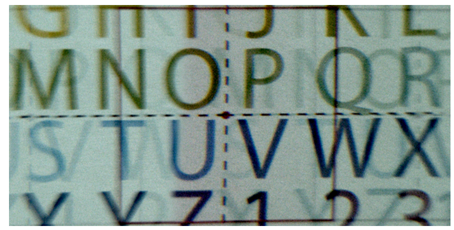
The images captured with and without the prism are shown at the top of Fig. 12 The two SAI's obtained with the prism are stitched to form an extended FOV image (see Fig. 12c).



(a) Full pupil refocus without prism.



(b) Full pupil refocus with prism.



(c) SAI stitching,

Figure 12: Canon 5D MarkIV refocus and stitching.

Note that on the stitched image, a ghost image is clearly visible. The sub-aperture cross-talk r is measured masking half of the main-lens while observing a white screen. The cross-talk is as high as $r \approx 50\%$. A central large mask, masking most of the aperture stop, is placed at the entrance of the main-lens to decrease the cross-talk up to $r \approx 18\%$, but it remains important. Indeed, we have discovered that the dual-pixel sensor designed by Canon does not produce a sharp sampling of the main-lens pupil. This is certainly done because a sharp sampling would ruin the light efficiency (photons not being collected), and would not improve the auto-focus accuracy. This experiment demonstrates the importance of jointly designing sensors and the corresponding computational algorithms.

7. Discussion and Conclusion

In the future, ideal quad-pixel sensors could be used in two distinct modes: 1/ the four SAI's are used with typical plenoptic algorithms (e.g. refocus[12], depth-estimation [29], partial lens aberration correction [14]) without previous micro-lens center estimation which is a cumbersome task, or 2/ as proposed in this paper, combined with a pyramid lens and stitching the different SAI's to double the FOV as well as the spatial resolution. The pyramid lens would be used as a conversion lens. These two options are especially suitable for smartphones which use a fixed focal lens, and presumably soon a quad-pixel sensor (dual-pixels already integrated).

To complete this study, several technical aspects are to be pointed out: 1/ The SAI's collect photons passing by a quarter disk in the case of a quad-pixel sensor. Thus the bokeh is not a rounded shape as for conventional cameras. Also, we effectively trade the amount of light dedicated to image a scene point against an increase in FOV. 2/ The main-lens has to be designed with an aperture stop located at its entrance and 3/ prisms produce strong chromatic aberrations that degrade the image quality of SAI's, even if achromatic prisms would decrease chromatic aberrations. Finally, using an optical simulation software, the two flat interfaces of the prism could be replaced with quarter spherical interfaces adapted to the main-lens to produce fewer aberrations.

All in all, in this paper we have demonstrated that plenoptic sensors offer a new paradigm for image processing thanks to the ideal pupil sampling they offer. We have presented a novel optical design coupling a pyramidal lens and an ideal plenoptic sensor. The main interest is the capacity to extend the FOV of the captured snapshot. Our design is validated with different experiments including simulated synthetic images and images captured with two different prototypes. In particular, we consider a type-2 plenoptic camera and a Canon DSLR with a dual-pixel sensor, being two available plenoptic cameras on the market. To our knowledge, this is the first time that the scientific community in the field presents experiments with a dual-pixel sensor.

References

[1] https://www.lytro.com.

[2] https://www.refractiveindex.info/.

[3] http://www.raytrix.de.

[4] Libraw library. `http://www.libraw.org/`.

[5] Optic studio. `http://www.zemax.com/`.

[6] C.-Y. Chen, T.-T. Yang, and W. S. Sun. Optics system design applying a micro-prism array of a single lens stereo image pair. *Opt. Express*, 16(20):15495–15505, Sep 2008.

[7] D. Cho, M. Lee, S. Kim, and Y.-W. Tai. Modeling the calibration pipeline of the lytro camera for high quality light-field image reconstruction. In *ICCV*, pages 3280–3287. IEEE, 2013.

[8] S. Choi, K. Lee, J. Yun, S. Choi, S. Lee, J. Park, E. S. Shim, J. Pyo, B. Kim, M. Jung, Y. Lee, K. Son, S. Jung, T. S. Wang, Y. Choi, D. k. Min, J. Im, C. R. Moon, D. Lee, and D. Chang. An all pixel pdaf cmos image sensor with $0.64\mu m$x$1.28\mu m$ photodiode separated by self-aligned in-pixel deep trench isolation for high af performance. In *Symp. on VLSI Technology*, pages T104–T105, 2017.

[9] O. S. Cossairt, D. Miau, and S. K. Nayar. Gigapixel computational imaging. In *ICCP*, pages 1–8, April 2011.

[10] D. G. Dansereau, O. Pizarro, and S. B. Williams. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *CVPR*, pages 1027–1034. IEEE, 2013.

[11] D. G. Dansereau, G. Schuster, J. Ford, and G. Wetzstein. A wide-field-of-view monocentric light field camera. In *Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jul. 2017.

[12] J. Fiss, B. Curless, and R. Szeliski. Refocusing plenoptic images using depth-adaptive splatting. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–9. IEEE, 2014.

[13] T. Georgiev and A. Lumsdaine. Focused plenoptic camera and rendering. *Journal of Electronic Imaging*, 19(2):021106–021106, 2010.

[14] P. Hanrahan and R. Ng. Digital correction of lens aberrations in light field photography. In *Int. Optical Design*, page WB2. Optical Society of America, 2006.

[15] M. Hog, N. Sabater, B. Vandame, and V. Drazic. An image rendering pipeline for focused plenoptic cameras. *IEEE Trans. on Computational Imaging*, 3(4):811–821, Dec 2017.

[16] C. Kolb, D. Mitchell, and P. Hanrahan. A realistic camera model for computer graphics. In *SIGGRAPH*, pages 317–324, New York, NY, USA, 1995. ACM.

[17] A. Lumsdaine and T. Georgiev. The focused plenoptic camera. In *ICCP*, pages 1–8. IEEE, 2009.

[18] R. Mercado. Small form factor telephoto camera, Apr. 30 2015. US Patent App. 14/069,027.

[19] R. Ng. *Digital Light Field Photography*. PhD thesis, Stanford, CA, USA, 2006. AAI3219345.

[20] A. E. G. Peter Bert Catrysse, Xinqiao Liu. Qe reduction due to pixel vignetting in cmos image sensors. volume 3965, pages 3965 – 3965 – 11, 2000.

[21] M. Pharr and G. Humphreys. *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2010.

[22] N. Sabater, M. Seifi, V. Drazic, G. Sandri, and P. Perez. Accurate disparity estimation for plenoptic images. In *ECCV Workshop LF4CV*, 2014.

[23] Y. Sekine. Solid state image sensor, method of manufacturing solid state image sensor, and image capturing system, Aug. 11 2016. US Patent App. 15/010,340.

[24] G. Somanath, M. V. Rohith, and C. Kambhamettu. *Single Camera Stereo System Using Prism and Mirrors*, pages 170–181. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[25] I. Stamenov, I. P. Agurok, and J. E. Ford. Optimization of two-glass monocentric lenses for compact panoramic imagers: general aberration analysis and specific designs. *Appl. Opt.*, 51(31):7648–7661, Nov 2012.

[26] I. Stamenov, A. Arianpour, S. J. Olivas, I. P. Agurok, A. R. Johnson, R. A. Stack, R. L. Morrison, and J. E. Ford. Panoramic monocentric imaging using fiber-coupled focal planes. *Opt. Express*, 22(26):31708–31721, Dec 2014.

[27] W.-S. Sun, C.-L. Tien, C.-Y. Chen, and D.-C. Chen. Single-lens camera based on a pyramid prism array to capture four images. *Optical Review*, 20(2):145–152, Mar 2013.

[28] R. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH*, pages 251–258, New York, NY, USA, 1997.

[29] M. W. Tao, P. P. Srinivasan, J. Malik, S. Rusinkiewicz, and R. Ramamoorthi. Depth from shading, defocus, and correspondence using light-field angular coherence. In *CVPR*, pages 1940–1948, 2015.

[30] P. P. M. V. Zhou Xue, Loic Baboulaz. Light field panorama by a plenoptic camera. In *SPIE*, volume 9020, 2014.