

8

Plenoptic Cameras

Fernando Pérez Nava¹, Alejandro Pérez Nava¹, Manuel Rodríguez Valido², Eduardo Magdaleno Castellò²

¹*Departamento de Estadística, Investigación Operativa y Computación, Universidad de La Laguna, Spain.*

²*Departamento de Ingeniería Industrial, Universidad de La Laguna, Spain.*

8.1 Introduction

The perceptual system provides to biological and artificial agents information about the environment in which they inhabit. This information is obtained by capturing one or more of the types of energy that surrounds them. By far the most informative is light, electromagnetic radiation in the visible spectrum. The sun fills the earth with photons that are reflected and absorbed by the environment and finally perceived by visual sensors. This initiates several complex information processing tasks that supply the agents with an understanding of their environment. There is an enormous interest to comprehend and replicate the biological visual perceptual system in artificial agents. The first step of this process is deciding how to capture light. To design visual sensors in artificial agents, we are not constrained by evolutionary or biological reasons and we may question if it is appropriate to replicate the human visual system. Although at first sight it seems reasonable that sensors resembling our eyes are more appropriate for the tasks that we are interested for computers to automate, other eye designs present in animals or not explored by the evolution may provide better performance for some tasks.

Eye design in animals began with simple photosensitive cells to capture light intensity. From these remote beginnings, eye evolution has been guided towards increasing the information that can be extracted from the environment [1]. When photosensitive cells were placed on a curved surface, directional selectivity was incorporated to the recorded light intensity. The curvature of this inward depression increased and finally formed a chamber, with an inlet similar to a pinhole camera. Pinhole cameras face a trade-off between low

light sensitivity (increasing with bigger holes) and resolution (increasing with smaller holes). The solution both in photography and animal eye evolution is the development of a lens. The resulting eye is the design found in vertebrates. The main advantages of single aperture eyes are high sensitivity and resolution, while its drawbacks are the small size of the field of view and its large volume and weight [2]. For small invertebrates, eyes are very expensive in weight and brain processing needs, so the compound eye was developed as a different solution toward achieving more information. Rather than curving the photoreceptive surface inward, evolution has led to a surface curved outwards in these creatures. This outward curvature result in a directional activation of specific photosensitive cells that is similar to the single chamber eye, except that this configuration does not allow for the presence of a single lens. Instead, each portion of the compound eye is usually equipped with its own lens forming an elemental eye. Each elemental eye captures light from a certain angle of incidence contributing to a spot in the final image. The final image of the environment consists of the sum of the light captured by all the elemental eyes, resulting in a mosaic like structure of the world. Inspired by their biological counterpart, artificial compound eye imaging systems have received a lot of attention in recent years. The main advantages of compound eye imaging systems over the classical camera imaging design are thinness, lightness and wide field of view. Several designs of compound eye imaging systems have been proposed over the last years [2].

In this chapter we will present the plenoptic camera and study its properties. As pointed out in [3], from the biological perspective, the optical design of the plenoptic camera can be thought of as taking a human eye (a camera) and replacing its retina with an compound eye (a combination of a microlens and a photosensor array). This allows the plenoptic camera to measure the 4D lightfield composed by the radiance and direction of all the light rays in a scene. Conventional 2D images are obtained by 2D projections of the 4D lightfield. The fundamental ideas behind the use of plenoptic cameras can be traced back to the beginning of the previous century. These ideas have been recently implemented in the field of Computational Photography and now there are several commercial plenoptic cameras available [4], [5].

In this chapter, we will expose the theory behind plenoptic cameras and their practical applications. In Section 8.2 we will introduce the lightfield concept and the devices that are used to capture it. Then, in Section 8.3 we will give a detailed description of the standard plenoptic camera, describing its design, properties and how it samples the 4D lightfield. In Section 8.4 we will show several applications of the plenoptic camera that extend the capabilities of current imaging sensors like perspective shift, refocusing the image after the shot, recovery of 3D information or extending the depth of field of the image. Plenoptic camera also has several drawbacks like a reduction in spatial resolution compared with a conventional camera. We will explain in Section 8.4 how to increase the resolution of the plenoptic camera using super-resolution techniques. In Section 8.5 we will introduce the generalized plenoptic camera as an optical method to increase the spatial resolution of the plenoptic camera and explore its properties. Plenoptic image processing requires high processing power so in Section 8.6 we will review some implementations in high performance computing architectures. Finally, in Section 8.7 we summarize the main outcomes of this chapter.

8.2 Lightfield representation of the plenoptic function

8.2.1 The plenoptic function

Every eye captures a subset of the space of the available visual information. The content in this subset and the capability to extract the necessary information determines how well the organism can perform a given task. To extract and process the visual information for a specific task, we need to know what is the information that we need and the camera design and image representation that optimally facilitates processing such visual information [6]. Therefore, our first step must be to describe the visual information available to an agent at any point in space and time. As pointed out in [7], complete visual information can be described in terms of a high dimensional plenoptic function (plenoptic comes from the Latin *plenus*, meaning full, and *opticus*, meaning related to seeing or vision). This function is a ray-based model for radiance. Rays are determined in terms of geometric spatial and directional parameters, time and colour spectrum. Since the plenoptic function concept is defined for geometric optics it is restricted to incoherent light and objects larger than the wavelength of light. It can be described as the 7D function:

$$P(x, y, z, \theta, \phi, \lambda, t), \quad (8.1)$$

where (x, y, z) are spatial 3D coordinates, (θ, ϕ) denote any direction given by spherical coordinates, λ is a wavelength and t stands for time. The plenoptic function is also fundamental for Image Based Rendering (IBR) in Computer Graphics that studies the representation of scenes from images. The dimensionality of the plenoptic function is higher than is required for IBR, and in [8] the dimensionality was reduced to a 4D subset, the lightfield, making the storage requirements associated to the representation of the plenoptic function more tractable. This dimensionality reduction is obtained through several simplifications. Since IBR uses several colour images of the scene, the wavelength dimension is described through three colour channels, i.e., red, green, and blue channels. Each channel represents the integration of the plenoptic function over a certain wavelength range. As the air is transparent and the radiances along a light ray through empty space remain constant, it is not necessary to record the radiances of a light ray on different positions along its path since they are all identical. Finally, it is considered that the scene is static, thus the time dimension is eliminated. The plenoptic function is therefore reduced to a set of 4D dimensional functions,

$$L(x, y, \theta, \phi). \quad (8.2)$$

Note that we have one 4D function for each colour plane making the plenoptic function more manageable.

8.2.2 Lightfield parametrizations

The lightfield can be described through several parametrizations. The most important are the spherical-cartesian parametrization, the sphere parametrization and the two planes parametrization. The spherical-Cartesian parametrization $L(x, y, \theta, \phi)$, also known as the one-plane parameterization, describes a ray's position using the coordinates (x, y) of its point of intersection with a plane and its direction using two angles (θ, ϕ) , for a total of four parameters. The sphere parametrization $L(\theta, \varphi, \phi, \psi)$ describes a ray through its intersection with a sphere. The 4D lightfield can also be described with the two plane parametrization as: $L(x, y, u, v)$, where the parameters (x, y) and (u, v) represent the intersection of the light ray with two parallel reference planes with axes denoted as x, y and u, v at one unit of distance [8]. The point on the x, y plane represents position and the point on the u, v plane represents direction, so the terms spatial and angular reference planes are used to describe each plane. The coordinates of the intersection point with the u, v plane may be defined as relative coordinates with respect to the intersection on the x, y plane or as absolute coordinates with respect to the origin of the u, v plane. Note that all previous parameterizations represent essentially the same information, the 4D lightfield, and conversions between them is possible except where rays run parallel to the reference planes in the two plane parametrization. In Figure 8.1 we show the different parametrizations in 3D space. In order to make the notation and graphical representations of the 4D lightfield simpler, we will use through the chapter a 2D version of the 4D lightfield and 2D versions of the 4D parametrizations. The generalization to the 4D case is straightforward. All parametrizations are shown on Figure 8.1

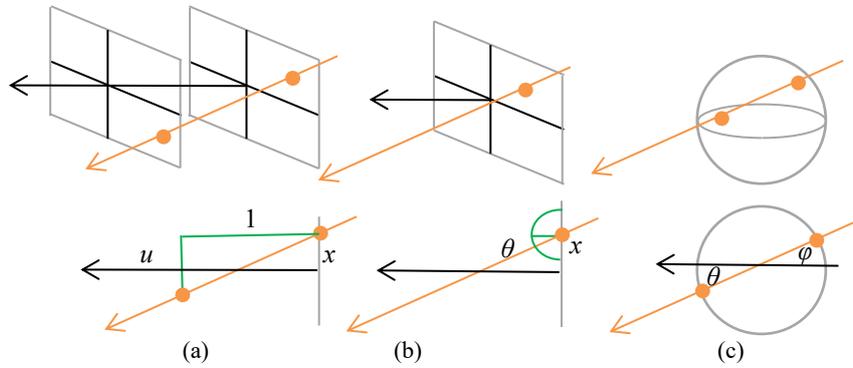


Figure 8.1 Lightfield parametrizations (a) two- plane parametrization, (b) spherical-cartesian parametrization (c) sphere parametrization

The lightfield can also be described in the frequency domain by means of its Fourier transform. Using the two-plane parametrization we have:

$$\hat{L}(\xi, \eta) = \int L(x, u) e^{-2\pi j(x\xi + u\eta)} dx du, \quad (8.3)$$

where (ξ, η) are spatial frequency variables and \hat{L} is the Fourier transform of the lightfield.

8.2.3 Lightfield reparametrization

When light traverses free space the radiance along the light ray is unchanged, but the lightfield representation depends on the position of the reference planes. Different reference planes will be fixed to the optical elements of the plenoptic camera to study how rays change inside the camera body. Thus, the reference planes can be moved to arbitrary distance d or be coincident with the plane of a lens with focal length f .

If we displace the reference planes by d units we can relate the original L_0 and new representation L_1 through a linear transform defined by a matrix \mathbf{T}_d . Similarly, we can describe the lightfield when light rays pass through a converging lens with focal length f aligning the spatial reference plane with the lens. In this case we also obtain a linear transform defined by a matrix \mathbf{R}_f considering the thin lens model in Gaussian optics [9].

The matrix description of the transforms is shown below in the spatial and frequency domain using the two plane parametrization with relative spatial coordinates (x, u) and relative frequency coordinates (ξ, η) [10].

$$L_1((x, u)^T) = L_0(\mathbf{T}_d(x, u)^T), \mathbf{T}_d = \begin{pmatrix} 1 & -d \\ 0 & 1 \end{pmatrix}, \hat{L}_1((\xi, \eta)^T) = \hat{L}_0(\mathbf{T}_d^{-T}(\xi, \eta)^T), \quad (8.4)$$

$$L_1((x, u)^T) = L_0(\mathbf{R}_f(x, u)^T), \mathbf{R}_f = \begin{pmatrix} 1 & 0 \\ 1/f & 1 \end{pmatrix}, \hat{L}_1((\xi, \eta)^T) = \hat{L}_0(\mathbf{R}_f^{-T}(\xi, \eta)^T), \quad (8.5)$$

Finally, it may be desirable to describe the lightfield by two reference planes separated by a distance B in absolute coordinates instead of the relative coordinates with unit distance separation. This is useful to parametrize the lightfield inside a camera where the spatial reference plane is aligned with the image sensor and the directional plane is aligned with the main lens. In this case, we can relate the original L_0 and new representation L_1 through a linear transform defined by a matrix \mathbf{S}_B .

$$L_1((x, u)^T) = \frac{1}{B} L_0(\mathbf{S}_B(x, u)^T), \mathbf{S}_B = \begin{pmatrix} 1 & 0 \\ -\frac{1}{B} & \frac{1}{B} \end{pmatrix}, \hat{L}_1((\xi, \eta)^T) = B \hat{L}_0(\mathbf{S}_B^{-T}(\xi, \eta)^T), \quad (8.6)$$

where the scaling factor is the Jacobian of the transform. Since the transformations described above are linear they can be combined to model more complex transforms simply by multiplying their associated matrices. Inverting the transform it is also possible to describe the original lightfield L_0 in terms of the new representation L_1

8.2.4 Image formation

An image sensor integrates all the incoming radiances into a sensed irradiance value $I(x)$:

$$I(x) = \int L(x, u) du \quad (8.7)$$

Therefore, a conventional image is obtained integrating the lightfield with relative spatial coordinates along the directional u -axis. In equation (8.7) there is also an optical vignetting term. In the paraxial approximation [9], this term can be ignored. Geometrically,

to obtain an image we project the 2D lightfield into the spatial reference plane. In the general case, for a n dimensional signal, integration along $n-k$ dimensions generates a k dimensional image. For a 4D lightfield, we integrate along the 2D directional variables to obtain a 2D image. According to the slice theorem [11] projecting a signal in the spatial domain is equivalent to extracting the spectrum of a signal in the Fourier domain. This process is known as a slicing operation and can be described as:

$$\hat{I}(\xi) = \hat{L}(\xi, 0) \quad (8.8)$$

The advantage of this formulation is that integration in the spatial domain is replaced by taking the slice $\eta = 0$ in the frequency domain that is a simpler operation.

8.2.5 Lightfield sampling

Lightfield recording devices necessarily capture a subset of the lightfield in terms of discrete samples. The discretization comes from two sources: the optical configuration of the device and the discrete pixels in the imaging sensor. The combined effect of these sources is to impose a specific geometry on the overall sampling of the lightfield. Understanding the geometry of the captured lightfield is essential to compare the different camera designs. In this section we will explore different solutions to sample the lightfield and explore its properties.

8.2.5.1 Pinhole and thin lens camera

The sampling geometry of an ideal pinhole camera is depicted in 2D in Figure 8.2. While this model treats pixels as having a finite spatial extent, the aperture is modelled as being infinitesimally small. The integration volume corresponding to each pixel of size Δp is a segment in the x axis, taking on a value of zero everywhere except for $u=0$. The sampling geometry of the thin lens model with an aperture of size D is shown in Figure 8.3. Now the integration volume for each pixel is a rectangle in the x, u plane, but as we can see from the sampling geometry, very low resolution is obtained in the directional reference plane u so several approaches have been proposed to improve the directional resolution of the lightfield.

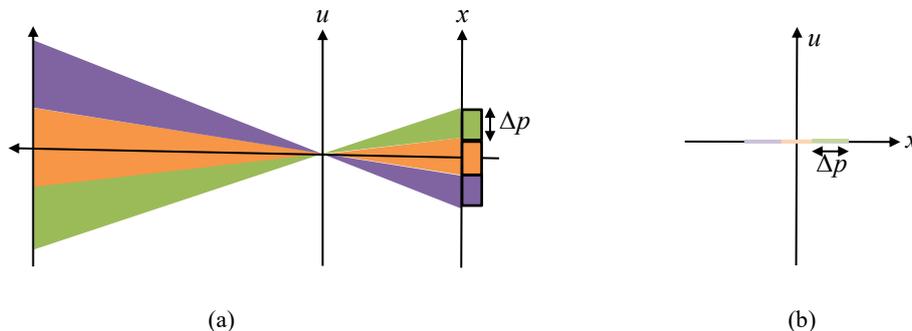


Figure 8.2 Pinhole camera. (a) Optical model, (b) Lightfield representation. Δp is the pixel size.

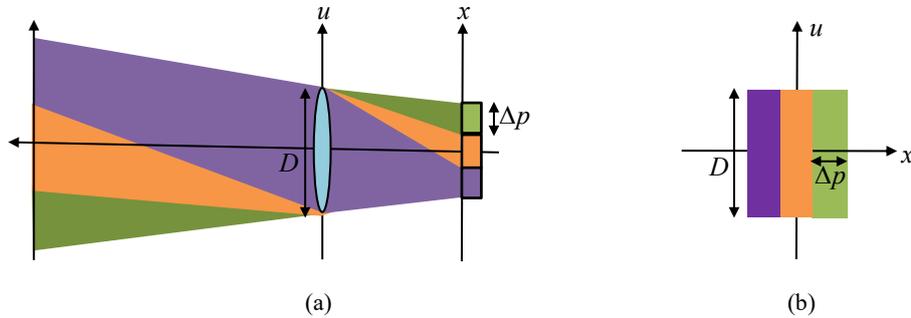


Figure 8.3 Thin lens camera. (a) Optical model, (b) Lightfield representation. D is the lens aperture and Δp is the pixel size.

8.2.5.2 Multiple devices

A straightforward approach to capture more information from the lightfield is the use of multiple devices. This is a solution also adopted in the biological domain that can be replicated with the use of an array of conventional 2D cameras distributed on a plane [12], [13]. If we adopt the pinhole approximation, sampling the light field from a set of images corresponds to inserting 2D slices into the 4D lightfield representation. Considering the image from each camera as a 2D slice of the 4D lightfield, an estimate of the lightfield is obtained from concatenating the captured slices. Camera arrays need geometric and color calibration and are expensive and complex to maintain, so there has been an effort to produce more manageable devices. In Figure 8.4 we show an example of a compact camera array composed of nine OV2640 cameras. The separation between them is 35 mm, the whole board is governed by a FPGA and data transmission is performed through and Ethernet 1000 MB under UDP/IP. More compact arrays for mobile phones have also been developed [14].



Figure 8.4 (a) A compact camera array (b) Raw images from the array. Note the need for geometric and colour calibration.

8.2.5.3 Temporal multiplexing

Camera arrays cannot provide sufficient lightfield resolution in the directional dimension for several applications. This is a natural result of the camera size that physically limits the cameras from being located close to each other. To handle this limitation, the temporal multiplexing approach uses a single camera to capture multiple images from the camera array by moving its location [8], [15]. The position of the camera can be controlled precisely using a camera gantry or estimated from the captured images by structure from motion (SFM) algorithms. The disadvantages of the time-sequential capture includes the difficulty of capturing dynamic environments and videos, its slowness because the position of the camera or the object has to be altered before each exposure and the necessity of either gantry or the SFM algorithms to work in controlled environments.

8.2.5.4 Frequency multiplexing

Temporal multiplexing solves some of the problems associated with camera arrays, but it has to be applied to static scenes so other means of multiplexing the 4D light field into a 2D image are required to overcome this limitation. In [16] the frequency multiplexing approach is introduced as an alternative method to capture the lightfield using a single device. The frequency multiplexing method encodes different slices of the plenoptic function in different frequency bands and is implemented by placing non-refractive light-attenuating masks in front of the image sensor. These masks provide frequency domain multiplexing of the 4D Fourier transform of the light field into the Fourier transform of the 2D sensor image. The main limitation of these methods is that the attenuating mask limits the light entering in the camera.

8.2.5.5 Spatial multiplexing

Spatial multiplexing produces an array of elemental images within the image captured with a single camera. This method began in the previous century with the works of Ives and Lippmann on integral imaging [17], [18]. Spatial multiplexing allows capturing dynamic scenes and videos, but transfers spatial sampling resolution in favour of directional sampling resolution. One implementation of a spatial multiplexing system uses an array of microlenses placed near the image sensor. This device is called a plenoptic camera and will be described in Section 8.3. Examples of such implementations can be found in [3], [19], [20], [21], [22].

8.2.5.6 Simulation

Light fields can also be obtained by rendering images from 3D models [23], [24]. By digitally moving the virtual cameras, this approach offers a low cost acquisition method and easy access to several properties of the objects like their geometry or textures. It is also possible to simulate imperfections of the optical device or the sensor's noise behaviour. Figure 8.5 shows an example of a simulated lightfield.



Figure 8.5 Simulated images from a lightfield.

8.2.6 Lightfield sampling analysis

In the previous section, we reviewed several devices designed to sample the lightfield. If enough samples are taken from them, we expect to reconstruct the complete continuous lightfield. However, the lightfield is a 4D structure so storage requirements are important and it is necessary to keep the samples to the minimum. The lightfield sampling problem is difficult because the sampling rate is determined by several factors, including the scene geometry, the texture on the scene surface or the reflection properties of the scene surface. In [25] it was proposed to perform the lightfield sampling analysis applying the Fourier transform to the lightfield, and then sampling it based on its spectrum for a camera array. Assuming constant depth, a Lambertian surface and no occlusions, the lightfield in the spatial domain is composed of constant slope lines and the Fourier spectrum of the lightfield is a slanted line. By considering all depths or equivalently all slanted lines, the maximum distance Δu between the cameras was found to be:

$$\Delta u = \frac{1}{Bf h_d}, \quad h_d = \frac{1}{z_{min}} - \frac{1}{z_{max}}, \quad (8.9)$$

where B is the highest spatial frequency of the lightfield, z_{min} and z_{max} are the depth bounds of the scene geometry and f is the focal length of the cameras. A more detailed analysis can be found in [26], where non-Lambertian and non-uniform sampling are considered. Sampling analysis of the plenoptic camera will be detailed in Section 8.4.1.

8.2.7 Lightfield visualization

In order to understand the lightfield properties, it is useful to have some way of visualizing its structure. Since it is not possible to display a 4D signal in a 2D image, the 4D light field will be represented joining multiple 2D slices from the 4D signal to obtain a 2D multiplexed image. The directional information for a lightfield with relative spatial coordinates (x, y, u, v) can be visualized extracting an u, v slice of the lightfield fixing the spatial variables x and y . In Figure 8.6 we see all the u, v slices spatially multiplexed. Another alternative shown in Figure 8.7 is to fix a direction u, v and let the spatial variables x and y vary. These images can be interpreted as images from a camera array. Finally, in Figure 8.8 the pair of variables x and u (or y and v) are fixed and the other pair vary. The resulting images are called epipolar images of the lightfield. These epipolar images can be used to estimate the depth of the objects using the slope of the lines that compose them [27]. In a plenoptic camera, a line with zero slope indicates that the object is placed on the focus plane while non-zero absolute values for slope indicates that the object is at a greater distance from the focus plane.

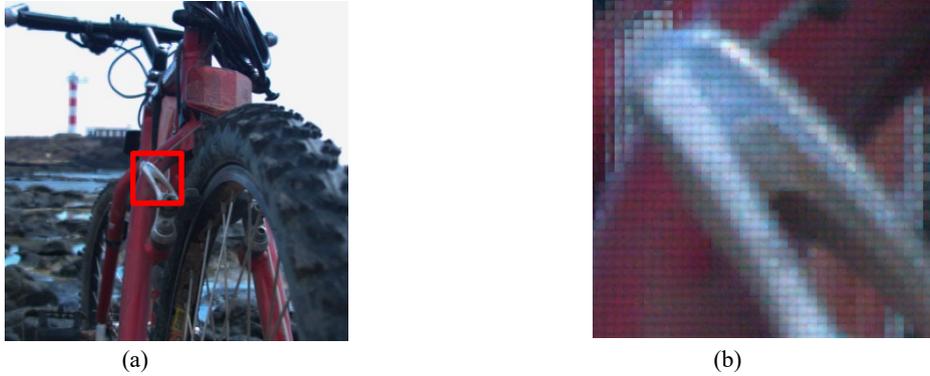


Figure 8.6 Fixing the spatial variables (x, y) and varying the directional variables (u, v) gives the directional distribution of the lightfield. On (a) we see the 4D lightfield as a 2D multiplexed image. The image in (b) is a detail from (a). Each small square in (b) is the (u, v) slice for a fixed (x, y) .



Figure 8.7 Fixing the directional variables (u, v) and varying the spatial variables (x, y) can be interpreted as taking images from a camera array. On (a) we see the 4D lightfield as another 2D multiplexed image. Image (b) is a detail from (a). Each image in (b) is the (x, y) slice for a fixed (u, v) .



Figure 8.8 Fixing one directional variable (u or v) and one spatial variable (x or y) generates another 2D multiplexed image. This is the epipolar image that gives information about the depth of the elements in the scene. On (b) we see a detail from (a). Note the different orientations corresponding to different depths in the scene.

8.3 The plenoptic camera

A plenoptic camera captures the 4D lightfield by placing a microlens array between the lens of the camera and the image sensor to measure the radiance and direction of all the light rays in a scene. The fundamental ideas behind the use of plenoptic cameras can be traced back to the beginning of the previous century with the works of Lippmann and Ives on integral photography [17], [18]. A problem of the Lippmann's method is the presence of pseudoscopic images (inverted parallax) when the object is formed at reconstruction. Different methods have been proposed for avoiding pseudoscopic images in real time applications [28], [29]. One of the first plenoptic cameras based on the principles of integral photography was proposed in the Computer Vision field by Adelson and Wang [19] to infer depth from a single image. Other similar systems were built by Okano et al. [30] and Naemura et al. [31] using graded-index microlens arrays. Georgiev et al. [32] have also used an integral camera composed of an array of lenses and prisms in front of the lens of an ordinary camera. Another integral imaging system is the Shack-Hartmann sensor [33] that is used to measure aberrations in optical systems. The plenoptic camera field experimented a great impulse when Ng et al. [20] presented a portable hand-held plenoptic camera. Its basic optical configuration, comprises a photographic main lens, a $N_x \times N_x$ microlens array and a photosensor array with $N_u \times N_u$ pixels behind each microlens. Figure 8.9 illustrates the layout of these components. Parameters D and d are the main lens and microlens apertures respectively. Parameter B is the distance from the microlens x -plane to the main lens u -plane and parameter b is the distance from the microlens plane to the sensor plane. Δp is the pixel size. The main lens may be translated along its optical axis to focus at a desired plane at distance A . As shown in Figure 8.9, rays of light from a single point on space converge on a single location in the plane of the microlens array. Then, the microlens at that location separates these light rays based on direction, creating a focused image of the main lens plane.

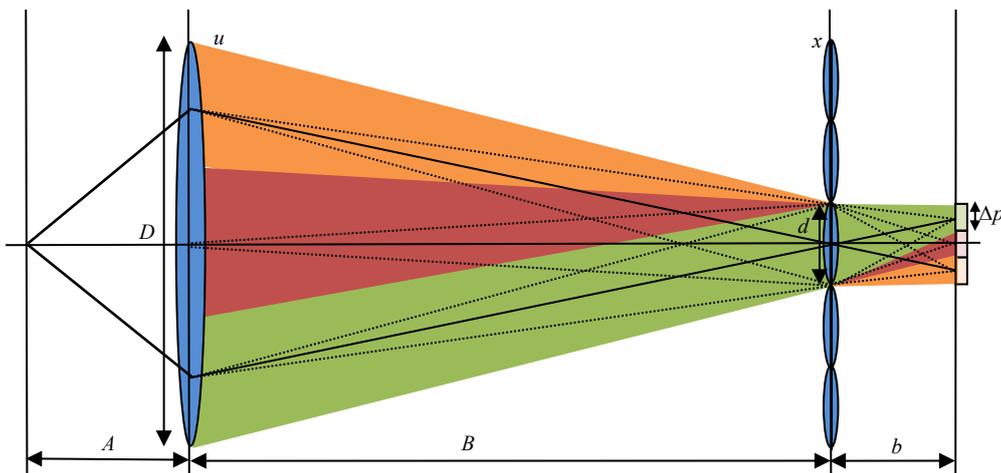


Figure 8.9 Optical configuration of the plenoptic camera (not shown at scale)

To maximize the directional resolution, the microlenses are focused on the plane of the main lens. This may seem to require a dynamic change of the separation between the photosensor plane and microlens array since the main lens moves during focusing and zooming. However, the microlenses are vanishingly small compared to the main lens, so independently of its zoom or focus settings the main lens is fixed at the optical infinity of the microlenses by placing the photosensor plane at the microlenses' focal depth [19]. To describe the light rays inside the plenoptic camera, an initial lightfield L_0 is defined aligning a directional reference plane with the image sensor and a spatial reference plane with the microlens plane. We can describe in absolute coordinates the lightfield inside the camera L_1 after light rays pass through the microlens s of length d and focal length f aligning the same spatial reference plane with the microlens plane and a directional reference plane with the main lens plane. The relation between L_0 and L_1 can be described using the transforms in Section 8.2.3 as:

$$L_1((x, u)^T) = (b/B) L_0(\mathbf{P}\mathbf{S}_b^{-1}\mathbf{T}_b\mathbf{R}_f\mathbf{S}_B(x - ds, u - ds)^T + (ds, ds)^T), \quad (8.10)$$

where \mathbf{P} is a permutation matrix that interchanges the components of a vector since rays in L_0 and L_1 point in a different direction.

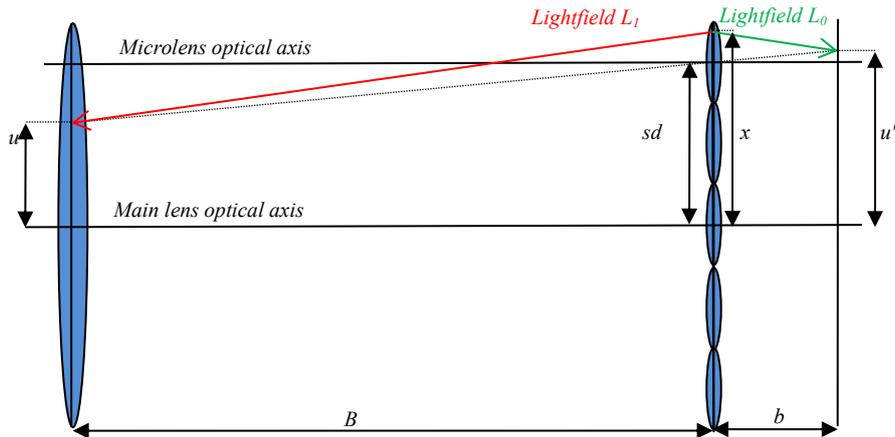


Figure 8.10 Ray transform in the plenoptic camera

Equation 8.10 can be explained as follows. We take a ray in lightfield L_1 defined from a point x on the microlens plane to a point u on the main lens plane and denote it by (x, u) and we want to obtain the corresponding ray (x, u') defined from the microlens plane to the sensor plane as seen in Figure 8.10. To obtain (x, u') we change the optical axis from the main lens to the microlens s of length d obtaining the ray $(x-ds, u-ds)$ with coordinates relative to the new optical axis. As shown in Figure 8.10 the distance between the microlens plane and the main lens plane is B , so we use \mathbf{S}_B to rewrite $(x-ds, u-ds)$ using a unit distance reference plane from the microlens plane in the main lens direction. Ray $\mathbf{S}_B(x-ds, u-ds)^T$ passes through the microlens s with focal length f that changes its direction, so we use \mathbf{R}_f to obtain the transformed ray $\mathbf{R}_f\mathbf{S}_B(x-ds, u-ds)^T$ defined from the microlens plane to the unit distance plane. To obtain the corresponding ray in L_0 , we have to move both

reference planes b units using \mathbf{T}_b to obtain the ray defined by $\mathbf{T}_b \mathbf{R}_f \mathbf{S}_B(x-ds, u-ds)^T$. Now, we separate the unit distance reference plane to distance b obtaining the ray $\mathbf{S}_b^{-1} \mathbf{T}_b \mathbf{R}_f \mathbf{S}_B(x-ds, u-ds)^T$ from the sensor plane to the microlens plane. Since we are interested that rays in L_0 point from the microlens plane to the sensor plane, we permute their coordinates using \mathbf{P} . The final step is to move again the optical axis to the main lens axis obtaining the ray $(x, u)^T = \mathbf{P} \mathbf{S}_b^{-1} \mathbf{T}_b \mathbf{R}_f \mathbf{S}_B(x-ds, u-ds)^T + (ds, ds)^T$ in L_0 . The relation between the radiance of rays (x, u) in L_1 and (x, u) in L_0 is defined in (8.10).

Computing the matrix products, equation (8.10) can be rewritten as:

$$L_1(x, u) = (b/B) L_0(x, ds(1 + b/B) - (b/B)u). \quad (8.11)$$

In order to use as many photosensor pixels as possible, it is desirable to choose the relative sizes of the main lens and microlens apertures so that the images under the microlenses are as large as possible without overlapping. Using (8.11) we see that the image of the main lens aperture in relative coordinates to microlens s is approximately $-(b/B)u$ since $(1 + b/B) \approx 1$. To use the maximum extent of the image sensor we must have the f -number condition $d/D = b/B$ since u varies within the aperture values $|u| \leq D/2$. Note that, the f -number condition is defined in terms of the image-side f -number, which is the diameter of the lens divided by the separation between the corresponding planes. Using (8.11) we can represent the sampling geometry of each pixel in L_1 as shown in Figure 8.11. Note the difference between the sampling pattern of Figures 8.3 and 8.10. In a plenoptic camera, the directional coordinate u is sampled while in a conventional camera the corresponding samples are integrated. Another result that can be seen from Figure 8.11 is that each row through the x -axis corresponds to an image taken from an aperture of size D/N_u . This implies that the depth of field of such images is extended N_u times relative to a conventional photograph.

The plenoptic model of this section represents a simplified representation of reality since it does not account for lens distortion or unwanted rotations between the microlens array and the photosensor plane. Therefore, a prior step to use this model in practical applications is to calibrate the plenoptic camera [34]. Images recorded from several plenoptic cameras can be obtained from [35], [36], [37], [38].

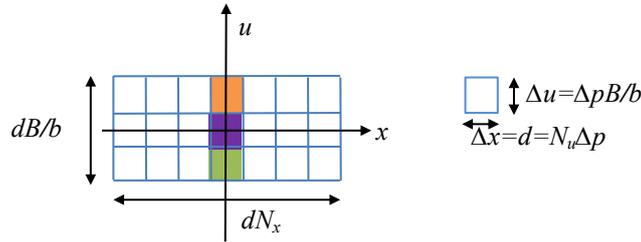


Figure 8.11 Sampling geometry of the plenoptic camera in L_1 . Each vertical column corresponds to a microlens. Each square corresponds to a pixel of the image sensor.

8.4 Applications of the plenoptic camera

8.4.1 Refocusing

In conventional photography the focus of the camera must be fixed before the shot is taken to ensure that selected elements in the scene are in focus. This is a result of all the light rays that reach an individual pixel being integrated as seen in Figure 8.3. With the plenoptic camera the lightfield is sampled directionally and the device provides the user with the ability to refocus images after the moment of exposure. If L_1 is a lightfield from the microlens plane to the main lens plane, the image is refocused by virtually moving the microlens plane to a distance $B\alpha$ from the main lens plane obtaining a lightfield L_r and applying the image formation operator in (8.7) as seen in Figure 8.12 (b). For a 2D lightfield we have:

$$L_r(\alpha x, u) = L_1(\mathbf{S}_B^{-1} \mathbf{T}_{(1-\alpha)B} \mathbf{S}_{B\alpha}(\alpha x, u)^T) = L_1((1 - 1/\alpha)u + x, u) \quad (8.12)$$

Then, applying the image formation operator, we obtain:

$$I_{\alpha B}(\alpha x) = \frac{1}{\alpha B} \int L_r(\alpha x, u) du = \frac{1}{\alpha B} \int L_1((1 - 1/\alpha)u + x, u) du \quad (8.13)$$

From equation (8.13), refocusing to a virtual plane is equivalent to integration along lines with slope $(1 - 1/\alpha)$. It can be interpreted as shifting for each u the subaperture image $L_1(x, u)$ with an amount of $(1 - 1/\alpha)u$ and adding all the shifted sub-aperture images. If we assume that a plenoptic camera is composed of $N_x \times N_x$ microlenses and each microlens generates a $N_u \times N_u$ image with $O(N_x) = O(N_u)$ then the full sensor resolution is $O((N_x N_u)^2) = O(N^4)$. In this case the shift and add approach needs $O(N^4)$ operations to generate a single photograph refocused on a determined distance.

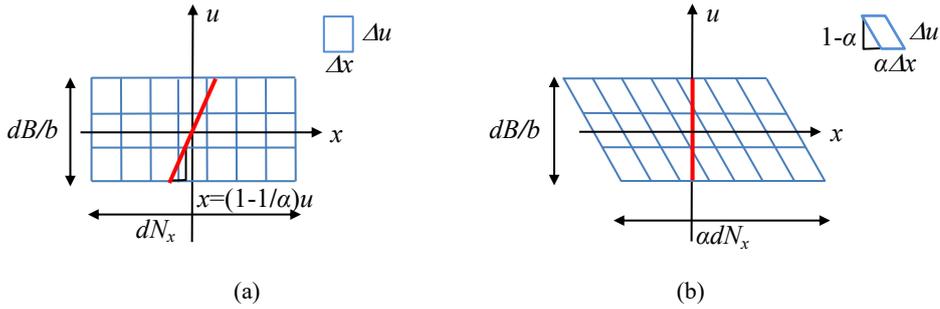


Figure 8.12 (a) Lightfield L_1 inside the camera. Refocusing is equivalent to integration along lines with $x - u$ slope $(1 - 1/\alpha)$. (b) Equivalent lightfield L_r with $x - u$ slope $(1 - \alpha)$ after virtually moving the focusing plane.

A significant improvement to this performance was obtained in [20] with the Fourier Slice Photography technique that reformulates the problem in the frequency domain. This

technique decreased the computational cost to $O(N^4 \log(N))$ to generate a set of photographs with size $O(N^2)$, giving $O(N^2 \log(N))$ operations to compute a $O(N^2)$ refocused image. This technique is based on the slicing operation in (8.8). Then, using (8.13) we obtain (8.14) indicating that a refocused photograph is a slanted slice in the Fourier Transform of the lightfield. This can be seen graphically in Figure 8.13

$$\begin{aligned} \hat{L}_1((1 - 1/\alpha)u + x, u)(\xi, \eta) &= \hat{L}_1(\xi, -(1 - 1/\alpha)\xi + \eta) \\ \hat{I}_{\alpha B}(\alpha x)(\xi) &= 1/(\alpha B) \hat{L}_1(\xi, -(1 - 1/\alpha)\xi) \end{aligned} \quad (8.14)$$

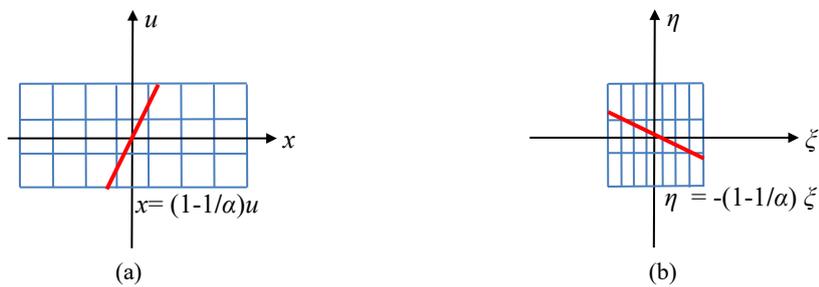


Figure 8.13 Fourier Slice theorem. (a) Spatial integration. (b) Fourier slicing

Then, in order to obtain a photograph we have to compute first the Fourier transform of the lightfield \hat{L}_1 . After that, we have to evaluate the slice $\hat{L}_1(\xi, -(1 - 1/\alpha)\xi)$ and finally we have to compute the inverse Fourier transform of the slice. A discrete version of the Fourier slice theorem can be found in [39]. An example of the refocusing technique is shown on Figure 8.14. To characterize the performance of the refocusing process it is necessary to obtain the range of depths for which we can compute an exact refocused photograph. If we assume a band-limited lightfield the refocusing range is [20]:

$$|1 - 1/\alpha| \leq \Delta x / \Delta u \quad (8.15)$$

In this slope interval the reconstruction of the photograph from the samples of the lightfield is exact.



Figure 8.14 Image refocusing at several distances using the discrete focal stack transform [39].

8.4.2 Perspective shift

In addition to the refocusing capability of the plenoptic camera, it is also possible to view the scene from different perspectives, allowing the viewer a better angular knowledge of the world. As explained in Section 8.3 the pixels under each microlens collects light rays from a certain direction. Thus, considering the same pixel under each microlens, it is possible to compose a picture of the scene viewed from that direction. Perspective images are shown in Figure 8.7.

8.4.3 Depth estimation

Plenoptic cameras can also provide 3D information about the scene indicating the depth of all pixels in the image [39], [40]. In fact, the plenoptic camera as a depth sensing device is equivalent to a set of convergent cameras [41]. Based on this 3D information it is also possible to obtain extended depth of field images in which all the elements in the image are in focus. In Figure 8.15 we show the basic geometry of depth estimation.

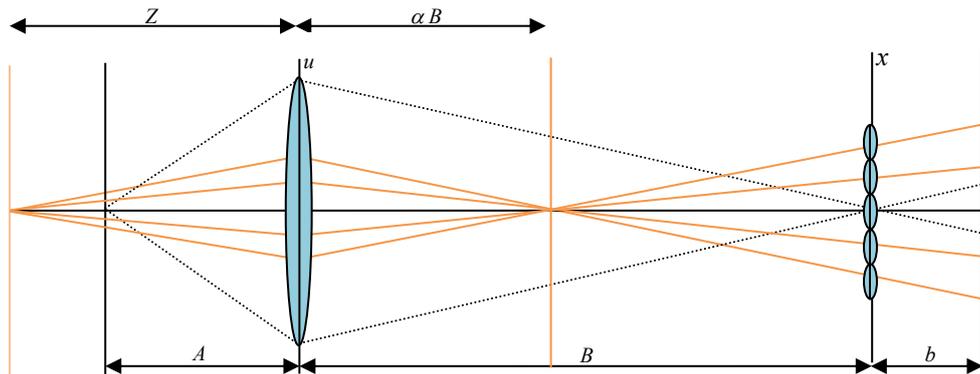


Figure 8.15 Geometry of depth estimation with a plenoptic camera.

It can be seen in the figure above that a Lambertian point at distance Z in the scene is transformed through the lens to a point at distance αB from the main lens. The set of rays emanating from this point is simply a line in 2D plenoptic space with equation $x = (1 - 1/\alpha)u$. Then, we can obtain the depth of the point estimating the slope of the line in the epipolar image that was presented in Section 8.2.7. This approach is widely employed due to its simplicity [19], [42]. For example, we may write the lightfield L for all points (x_i, u_i) in a neighbourhood $N(x, u)$ of a point (x, u) as: $L((x_i, u_i) + \mathbf{v}) \approx L((x_i, u_i)) + \nabla L(x_i, u_i)^T \mathbf{v}$. Then, noting that in the direction $\mathbf{v} = ((1 - 1/\alpha), 1)$ the lightfield is constant due to the Lambertian assumption, we obtain the set of equations $e(x_i, u_i) = L_x(x_i, u_i)(1 - 1/\alpha) + L_u(x_i, u_i) = 0$. Then, we can define a variance measure

$VM(x, u, \alpha)$ as the sum of the $e(x_i, u_i)^2$ in the neighbourhood $N(x, u)$ and we can estimate the unknown slope through minimizing $VM(x, u, \alpha)$ as [19]:

$$(1 - 1/\alpha) = - \sum_{(x_i, u_i) \in N(x, u)} L_x(x_i, u_i) L_u(x_i, u_i) / \sum_{(x_i, u_i) \in N(x, u)} L_x(x_i, u_i)^2 \quad (8.16)$$

Note that the estimation is unstable when the denominator is close to zero. This occurs in regions of the image with insufficient detail as is common in all passive depth recovery methods. Another alternative is to use a depth from focus approach. If we take a photograph of the Lambertian point at distance αB from the main lens with the sensor placed at a virtual distance of γB from the main lens we obtain a constant element with radius:

$$FM(x, \alpha) = D|(1 - 1/\alpha) - (1 - 1/\gamma)| \quad (8.17)$$

Then, we can obtain refocused images for several values of γ and select the sharpest one (the image with the smallest radius). There are several focus measures FM that can be employed which are basically contrast detectors within a small spatial window in the image. Unfortunately, the selection of the sharpest point is also unstable when the image has insufficient detail and some regularization method has to be employed to guide the solution. In [39] the Markov Random Field approach is used to obtain the optimal depth minimizing an energy function. Other regularization approaches include variational techniques [42]. Some examples of depth estimation are shown in Figure 8.16. In order to use the plenoptic camera as a depth recovery device, it is interesting to study its depth discrimination capabilities. This is defined as the smallest relative distance between two on-axis points, one at the focusing distance Z of the plenoptic camera and the other at Z' which would give a slope connecting the central pixel of the central microlens with a border pixel of an adjacent microlens. The relative distance can be approximated as [43]:

$$(Z' - Z)/Z \approx \pm dk(\theta/S)^2 Z \quad (8.18)$$

where d is the microlens size, k is the f -number of the camera, θ is the field of view and S is the sensor size. Therefore, to increase the depth discrimination capabilities, it is necessary to select small values of the microlens size, f -number, field of view and focusing distance and to select high values of the sensor size.

8.4.4 Extended depth of field images

Another feature of plenoptic cameras is their capability to obtain extended depth of field images also known as all-in-focus images [3], [39]. The approach in lightfield photography is to capture a lightfield, obtain the refocused images at all depths, and select for each pixel the appropriate refocused images according to the estimated depth. Extended depth of field images are obtained in classical photography reducing the size of the lens aperture. An advantage of the plenoptic extended depth of field photograph is the use of light coming

through a larger lens aperture allowing images with a higher signal to noise ratio. Example of extended depth of field images are shown on Figure 8.16



Figure 8.16 Depth estimation and extended depth of field images.

8.4.5 Super-resolution

One of the major drawbacks of plenoptic cameras is that capturing a certain amount of directional resolution requires a proportional reduction in spatial resolution. If the number of microlenses is $N_x \times N_x$ and the pixels behind each microlens are $N_u \times N_u$, the refocused images have a resolution of $N_x \times N_x$ and sensor resolution is reduced with a factor of $(N_u)^2$. Two approaches to increase the resolution can be found in plenoptic cameras. In this section we will present some software-based methods while in Section 8.5 hardware-based methods will be shown. Software based super-resolution methods discard the optimal sampling assumption of the lightfield of Section 8.4.1 and try to recover the lightfield using the aliased recorded plenoptic image. In [44] it is shown how to build super-resolution refocused photographs that has a common size of approximately $\frac{1}{4}$ of the full sensor resolution. It recovers the super-resolution photograph, assuming that the scene is Lambertian and composed of several planes of constant depth. Since depth in the lightfield is unknown, it is estimated by the defocus method in section 8.4.3. In [45] the super-resolution problem was formulated in a variational Bayesian framework to include blur effects reporting a $7\times$ magnification factor. In [46] the Fourier Slice technique of [39] that was developed for photographs with standard resolution is extended to obtain super-resolved images. Depth is also estimated by the defocus method in Section 8.4.3. In Figure 8.17 some results of the Fourier Slice super-resolution technique are shown.

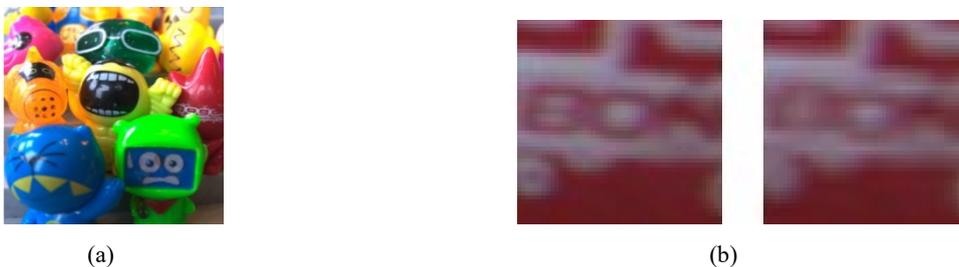


Figure 8.17 (a) Super-resolution photographs from a plenoptic camera. (b) Comparison of super-resolution (left) against bilinear interpolation (right).

That can be explicitly written as:

$$L_1(x, u) = (b\beta/B) L_0(x, ds(1 + b\beta/B) + (x - ds)(1 - \beta) - (b\beta/B)u) \quad (8.20)$$

Using (8.20) we can represent the sampling geometry of each pixel as shown in Figure 8.19 (a). Note the difference between the sampling pattern of the standard plenoptic camera in Figure 8.11 and the generalized plenoptic camera. As the β value changes from 1 the lightfield sampling pattern shears within each microlens. This generates an increase in spatial resolution and a decrease in the directional resolution. If the resolution of the sensor is $N_p \times N_p$ and there are $N_u \times N_u$ pixels behind each of the $N_x \times N_x$ microlenses, the refocused images have a resolution of $N_r \times N_r$, where:

$$N_r = \max((1 - \beta) N_p, N_x) \quad (8.21)$$

and the increment in resolution is $((1 - \beta) N_u)^2$ that is $(1 - \beta)^2$ times the number of pixels in each microlens [3],[47]. Note that the change of pixel size in the x axis is reflected in Figures 8.11 and 8.19 varying from d to $\Delta p / (1 - \beta)$ leading to the increment in resolution in (8.21)

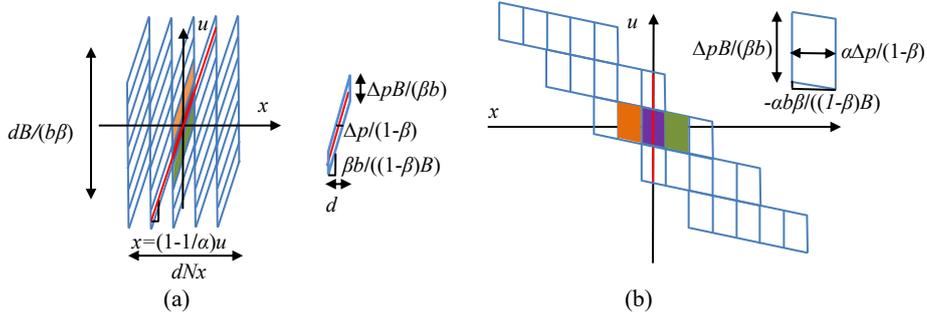


Figure 8.19 Sampling geometry of the generalized plenoptic camera. (a) Sampling in L_1 . Each vertical column corresponds to a microlens. Each sheared square corresponds to a pixel of the image sensor. Pixel slope in the x - u plane is $\beta b / ((1 - \beta) B)$. (b) Sampling in L_r , where the microlens plane is at distance $B\alpha$ from the main lens and α verifies (8.22).

In order to exploit the resolution capabilities of the generalized plenoptic camera, the focus of its main lens has to be done differently than in a conventional or plenoptic camera [3]. In conventional and standard plenoptic cameras best results are obtained by optically focusing on the subject of interest. On the other side, as shown in Figure 8.19 (a) the highest final image resolution occurs when the refocusing plane verifies:

$$1 - 1/\alpha = \beta b / ((1 - \beta) B), \quad (8.22)$$

or equivalently $\alpha = \beta b / (B(1 - \beta))$. The sampling pattern for a lightfield Lr where the microlens plane is at distance $B\alpha$ from the main lens plane and α verifies (8.22) is shown in Figure 8.19 (b). If we want to obtain the maximum resolution at depth B , we should optically focus the main lens by positioning it at depth B/α for an optimal focusing distance of:

$$B_{opt} = B - b\beta / (1 - \beta). \quad (8.23)$$

Refocused photographs from a generalized plenoptic camera are shown in Figure 8.20. Other generalizations of the standard design include the use of microlenses with different focal lengths [48], [49] to increase the depth of field of the generalized plenoptic camera or the use microlenses with different apertures to obtain high dynamic range imaging [50].



Figure 8.20 Refocusing from a generalized plenoptic camera

8.6 High Performance Computing with Plenoptic Cameras

Processing and rendering the plenoptic images requires significant computational power and memory bandwidth. At the same time, interactive rendering performance is highly desirable so that users can explore the variety of images that can be rendered from a single plenoptic image. For example, it is desirable to analyse the refocused photographs as a whole to estimate depth or to select a group of them with focus on determined elements. Therefore, it is convenient to use fast algorithms and appropriate hardware platform such as FPGA (Field Programmable Gate Array) or GPU (Graphics Processing Unit). A GPU is a co-processor that takes on graphical calculations and transformations so that the main CPU does not have to be burdened by them. The FPGAs are semiconductor devices that are based on a matrix of configurable logic blocks connected via programmable interconnects. This technology makes possible to insert the applications on a custom built integrated circuits. Also, FPGA devices offer extremely high-performance signal processing through parallelism and highly flexible interconnection possibilities. In [44] a method is presented for simultaneously recovering super-resolved depth and all-in-focus images from a standard plenoptic camera in near real time using graphics processing units (GPUs). In [51] a GPU-based approach for lightfield processing and rendering for a generalized plenoptic camera is detailed. The rendering performance allows to process 39Mpixel plenoptic data to 2Mpixel images with frame rates in excess of 500 frames per second. Another application example 244 × 168 Style File (Version 18 January 1999)

for FPGAs is presented in [52]. The authors describe a fast, specialized hardware implementation of a super-resolution algorithm for plenoptic cameras. Results show a time reduction factor of 50 in comparison with the solution on a conventional computer. Another algorithm designed for FPGA devices is shown in [53]. The authors describe a low-cost embedded hardware architecture for real-time refocusing based on a standard plenoptic camera. The video output is transmitted via High-Definition Multimedia Interface (HDMI) with a resolution of 720p at a frame rate of 60 fps conforming to the HD ready standard.

8.7 Conclusions

During the last years the field of Computational Photography continued developing new devices with unconventional structures that extends the capabilities of current commercial cameras. The design of these sensors is guided towards increasing the information that can be captured from the environment. The available visual information can be described through the plenoptic function and its simplified 4D lightfield representation. In this chapter, we have reviewed several devices to capture this visual information and shown how the plenoptic camera captures the lightfield by inserting a microlens array between the lens of the camera and the image sensor measuring the radiance and direction of all the light rays in a scene. We have explained how the plenoptic images can be used, for instance, to change the scene viewpoint, to refocus the image, to estimate the scene depth or to obtain extended depth of field images. The drawback of the plenoptic camera in terms of spatial resolution loss has also been detailed and some solutions to overcome these problems have also been presented. Finally, we have explored the computational part of the techniques presented in high performance computing platforms such as GPUs and FPGAs. The plenoptic camera has left the research laboratories and now several commercial versions can be found [4], [5], [13]. We expect that the extension of its use will increase the capabilities of current imaging sensors and drive imaging science to new directions.

References

- [1] E. B. Goldstein, *Encyclopedia of Perception*. Los Angeles: SAGE Publications, Inc, 2009.
- [2] J. W. Duparré and F. C. Wippermann, “Micro-optical artificial compound eyes,” *Bioinspir. Biomim.*, vol. 1, no. 1, pp. R1–16, Mar. 2006.
- [3] R. Ng, “Digital light field photography,” 2006.
- [4] “Lytro.” [Online]. Available: <https://www.lytro.com>. [Accessed: 04-Jul-2014].
- [5] “Raytrix.” [Online]. Available: <http://www.raytrix.de/>. [Accessed: 04-Jul-2014].
- [6] J. Neumann, C. Fermuller, and Y. Aloimonos, “Eye design in the plenoptic space of light rays,” in *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings*, 2003, pp. 1160–1167 vol.2.
- [7] E. H. Adelson and J. R. Bergen, “The Plenoptic Function and the Elements of Early Vision,” in *Computational Models of Visual Processing*, 1991, pp. 3–20.
- [8] M. Levoy and P. Hanrahan, “Light Field Rendering,” in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1996, pp. 31–42.

- [9] Born M. and Wolf E, *Principles of Optics*, 7th ed. Cambridge University Press, 1999.
- [10] C.-K. Liang, Y.-C. Shih, and H. H. Chen, "Light Field Analysis for Modeling Image Formation," *IEEE Trans. Image Process.*, vol. 20, no. 2, pp. 446–460, Feb. 2011.
- [11] R. N. Bracewell, "Numerical Transforms," *Science*, vol. 248, no. 4956, pp. 697–704, Nov. 1990.
- [12] 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," in *ACM SIGGRAPH 2005 Papers*, New York, NY, USA, 2005, pp. 765–776.
- [13] Y. Liu, Q. Dai, and W. Xu, "A Real Time Interactive Dynamic Light Field Transmission System," in *2006 IEEE International Conference on Multimedia and Expo*, 2006, pp. 2173–2176.
- [14] "Pelican Imaging." [Online]. Available: <http://www.pelicanimaging.com/>. [Accessed: 09-Jul-2014].
- [15] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The Lumigraph," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1996, pp. 43–54.
- [16] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing," in *ACM SIGGRAPH 2007 Papers*, New York, NY, USA, 2007.
- [17] F. Ives, "US 725567," US 725567.
- [18] G. Lippmann, "Epreuves reversibles donnant la sensation du relief," *J. Phys.*, vol. 7, no. 4, pp. 821–825, 1908.
- [19] E. H. Adelson and J. Y. A. Wang, "Single lens stereo with a plenoptic camera," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 99–106, Feb. 1992.
- [20] R. Ng, "Fourier Slice Photography," in *ACM SIGGRAPH 2005 Papers*, New York, NY, USA, 2005, pp. 735–744.
- [21] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *2009 IEEE International Conference on Computational Photography (ICCP)*, 2009, pp. 1–8.
- [22] J. M. Rodríguez-Ramos, B. Femenía Castellá, F. Pérez Nava, and S. Fumero, "Wavefront and distance measurement using the CAFADIS camera," in *Proc. SPIE 7015, Adaptive Optics Systems*, 2008, vol. 7015, p. 70155Q–70155Q–10.
- [23] F. P. Nava, J. P. Luke, J. G. Marichal-Hernandez, F. Rosa, and J. M. Rodriguez-Ramos, "A Simulator for the Cafadis Real Time 3DTV Camera," in *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008*, 2008, pp. 321–324.
- [24] S. Wanner, S. Meister, and B. Goldluecke, "Datasets and Benchmarks for Densely Sampled 4D Light Fields," in *Vision, Modelling and Visualization (VMV)*, 2013.
- [25] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic Sampling," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 2000, pp. 307–318.
- [26] C. Zhang and T. Chen, "Spectral analysis for sampling image-based rendering data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 11, pp. 1038–1050, Nov. 2003.
- [27] R. C. Bolles, H. H. Baker, David, and H. Marimont, "Epipolarplane image analysis: An approach to determining structure from motion," in *Intern..I. Computer Vision*, 1987, pp. 1–7.
- [28] J. Arai, H. Kawai, and F. Okano, "Microlens arrays for integral imaging system," *Appl. Opt.*, vol. 45, no. 36, pp. 9066–9078, Diciembre 2006.
- 244 × 168 Style File (Version 18 January 1999)

- [29] J. Yeom, K. Hong, Y. Jeong, C. Jang, and B. Lee, "Solution for pseudoscopic problem in integral imaging using phase-conjugated reconstruction of lens-array holographic optical elements," *Opt. Express*, vol. 22, no. 11, pp. 13659–13670, Jun. 2014.
- [30] F. Okano, J. Arai, H. Hoshino, and I. Yuyama, "Three-dimensional video system based on integral photography," *Opt. Eng.*, vol. 38, no. 6, pp. 1072–1077, 1999.
- [31] T. Naemura, T. Yoshida, and H. Harashima, "3-D computer graphics based on integral photography," *Opt. Express*, vol. 8, no. 4, pp. 255–262, Feb. 2001.
- [32] T. Georgeiv, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala, "Spatio-Angular Resolution Tradeoff in Integral Photography," in *In Eurographics Symposium on Rendering*, 2006, pp. 263–272.
- [33] B. C. Platt and R. Shack, "History and principles of Shack-Hartmann wavefront sensing," *J. Refract. Surg. Thorofare NJ 1995*, vol. 17, no. 5, pp. S573–577, Oct. 2001.
- [34] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras," in *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 1027–1034.
- [35] "Research - Raytrix GmbH." [Online]. Available: <http://www.raytrix.de/index.php/Research.html>. [Accessed: 28-Nov-2014].
- [36] "Plenoptic Imaging | ACFR Marine." [Online]. Available: <http://www-personal.acfr.usyd.edu.au/ddan1654/LFSamplePack1-r2.zip>. [Accessed: 28-Nov-2014].
- [37] "Heidelberg Collaboratory for Image Processing." [Online]. Available: http://hci.iwr.uni-heidelberg.de/HCI/Research/LightField/lf_benchmark.php. [Accessed: 28-Nov-2014].
- [38] "Todor Georgiev - Photoshop - Adobe." [Online]. Available: <http://www.tgeorgiev.net/Gallery/>. [Accessed: 28-Nov-2014].
- [39] F. P. Nava, J. G. Marichal-Hernández, and J. M. Rodríguez-Ramos, "The Discrete Focal Stack Transform," in *Proc. European Signal Processing Conference*, 2008.
- [40] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi, "Depth from Combining Defocus and Correspondence Using Light-Field Cameras," in *Computer Vision, IEEE International Conference on*, Los Alamitos, CA, USA, 2013, vol. 0, pp. 673–680.
- [41] Y. Y. Schechner and N. Kiryati, "Depth from Defocus vs. Stereo: How Different Really Are They?," *Int J Comput Vis.*, vol. 39, no. 2, pp. 141–162, Sep. 2000.
- [42] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 606–619, Mar. 2014.
- [43] V. Drazic, "Optimal depth resolution in plenoptic imaging," in *2010 IEEE International Conference on Multimedia and Expo (ICME)*, 2010, pp. 1588–1593.
- [44] F. Perez Nava and J. P. Luke, "Simultaneous estimation of super-resolved depth and all-in-focus images from a plenoptic camera," in *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2009*, 2009, pp. 1–4.
- [45] T. E. Bishop and P. Favaro, "The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 972–986, May 2012.
- [46] F. Pérez, A. Pérez, M. Rodríguez, and E. Magdaleno, "Super-Resolved Fourier-Slice Refocusing in Plenoptic Cameras," *J. Math. Imaging Vis.*, pp. 1–18, Oct. 2014.
- [47] A. Lumsdaine, T. G. Georgiev, and G. Chunev, "Spatial analysis of discrete plenoptic sampling," in *Proc. SPIE 8299, Digital Photography VIII*, 2012, vol. 8299, pp. 829909–829909–10.

- [48] C. Perwass and L. Wietzke, “Single lens 3D-camera with extended depth-of-field,” in *Proc. SPIE 8291, Human Vision and Electronic Imaging XVII*, 2012, vol. 8291, pp. 829108–829108–15.
- [49] T. Georgiev and A. Lumsdaine, “The multifocus plenoptic camera,” 2012, vol. 8299, pp. 829908–829908–11.
- [50] T. G. Georgiev, A. Lumsdaine, and S. Goma, “High Dynamic Range Image Capture with Plenoptic 2.0 Camera,” in *Frontiers in Optics 2009/Laser Science XXV/Fall 2009 OSA Optics & Photonics Technical Digest*, 2009, p. SWA7P.
- [51] A. Lumsdaine, G. Chunev, and T. Georgiev, “Plenoptic rendering with interactive performance using GPUs,” in *Proc. SPIE 8295, Image Processing: Algorithms and Systems X; and Parallel Processing for Imaging Applications II*, 2012, vol. 8295, pp. 829513–829513–15.
- [52] J. Pérez, E. Magdaleno, F. Pérez Nava, M. Rodríguez, D. Hernández, and J. Corrales, “Super-Resolution in Plenoptic Cameras Using FPGAs,” *Sensors*, vol. 14, no. 5, pp. 8669–8685, May 2014.
- [53] C. Hahne and A. Aggoun, “Embedded FIR filter design for real-time refocusing using a standard plenoptic video camera,” in *Proc. SPIE 9023, Digital Photography X*, 2014.