The Multi-Focus Plenoptic Camera

Todor Georgiev[a] and Andrew Lumsdaine[b]

[a]Adobe Systems, San Jose, CA, USA;
[b]Indiana University, Bloomington, IN, USA

Abstract Text for Online or Printed Programs: The focused plenoptic camera is based on the Lippmann sensor: an array of microlenses focused on the pixels of a conventional image sensor. This device samples the radiance, or plenoptic function, as an array of cameras with large depth of field, focused at a certain plane in front of the microlenses. For the purpose of digital refocusing (which is one of the important applications) the depth of field needs to be large, but there are fundamental optical limitations to this. The solution of the above problem is to use and array of interleaved microlenses of different focal lengths, focused at two or more different planes. In this way a focused image can be constructed at any depth of focus, and a really wide range of digital refocusing can be achieved. This paper presents our theory and results of implementing such camera. Real world images are demonstrating the extended capabilities, and limitations are discussed.

Keywords: plenoptic camera, rendering, multiple focus

1. INTRODUCTION

Photography is constantly innovating, expanding the range of its applicability in terms of both image capture and creative postprocessing. This is especially true in the age of computational photography, where the computer provides new possibilities. High dynamic range (HDR),[1] panoramas,[2] stereo 3D[3,4] and lightfield imaging[5,6] are some examples of innovations that extend photography beyond its traditional boundaries. Many more can be found in the recent literature (and we remark that each of these fields has its own vast literature—our citations here are merely representative, not definitive).

Integral/lightfield photography in particular provides the creative professional with powerful capabilities. Normally, when a photographer takes a picture, he or she must make a number of decisions about various camera parameters such as aperture, focus, and point of view. Then, once the picture is taken, those parameter choices are fixed for that photograph. If the photographer wishes to have a picture with a different set of parameter choices, a separate picture must be taken.

Instead of capturing one view or 2D image of the scene with a fixed set of parameters, integral photography captures the 4D radiance or plenoptic function associated with the scene. Pictures are rendered from the plenoptic function computationally. Most importantly for the creative professional, the decisions about camera settings are made when an image is rendered computationally from the captured radiance, and not at capture time, allowing the creative professional to render an infinite number of different images from the same captured data.

The plenoptic function, as originally defined in,[7] is a record of the geometric structure of the lightfield as well as its dependence on parameters such as wavelength, polarization, etc. Most work to this point has explored purely geometric aspects of the plenoptic function, enabling photographers to computationally manipulate the focus, depth of field, and parallax of a scene. Less attention has been paid to capturing and using other aspects of the plenoptic function (with some exceptions, e.g., HDR capture in[8]).

In this paper, we analyze in detail the specific case of multifocus radiance capture. This approach allows us to extend the plenoptic depth of field, i.e., the range in which we can create perfectly focused images with a single capture.

Our approach to plenoptic sampling is based on the focused plenoptic camera,[9] which, as the name implies, requires that the captured microimages be well-focused. Because of the wave nature of light, diffraction can be minimized only with large microlens apertures, which in turn corresponds to a shallow depth of field. In order to extend the depth of field of the focused plenoptic capture, our multifocus sensor has microlenses with different focal lengths, capturing in-focus images at different depths in front of the lenslet array. In the case of multifocus capture, our approach makes it possible to computationally focus on any depth, something that was not possible with the previous plenoptic camera based on a single focal length.

2. BACKGROUND

In this section, we briefly review terminology and notation used in the remainder of this paper.

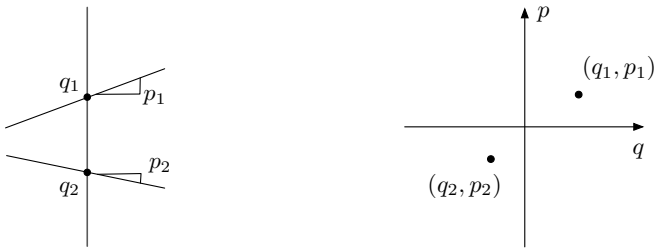
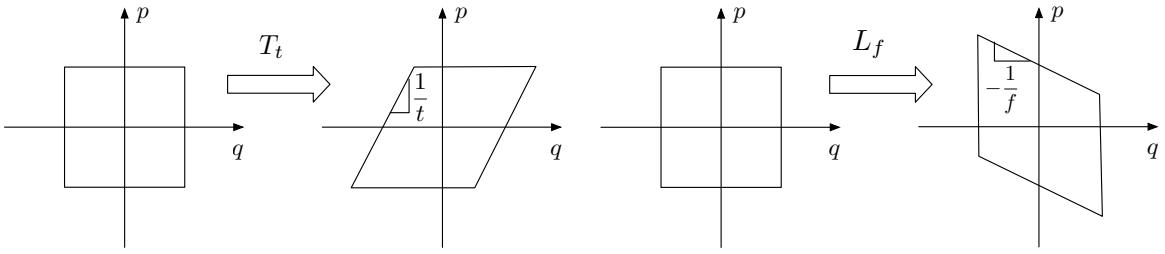


Figure 1: Rays are represented with coordinates q and p. The space of all rays comprises the q–p phase space.



(a) Shearing due to translation.

(b) Shearing due to refraction by a lens.

Figure 2: Translation and refraction by a lens act as shearing transformations in q–p phase space.

2.1 The Plenoptic Function

The plenoptic function[7] (also called the *lightfield*[10] or *radiance*[11]) is a density function describing the light rays in a scene. Since the radiance is a density function over the ray space, we describe radiance transformations via transformations applied to elements of the underlying ray space. Rays in three dimensional space are represented as four dimensional vectors: Two coordinates are required to describe position and two are required to describe direction. Following the physics-based convention of,[12] we denote the radiance at a given plane perpendicular to the optical axis as $r(q, p)$, where q describes the location of a ray in the plane and p describes its direction (see Figure 1). (These coordinates are also used in optics texts such as.[13,14]) For illustrative purposes, and without loss of generality, we adopt the convention of a two-dimensional q–p plane in this paper.

Translation and refraction by a lens are two fundamental transformations that can be applied to rays. Rays are transformed due to translation a distance t in the direction of the optical axis according to $(q', p') = (q + tp, p)$, corresponding to a linear transformation $x' = T_t x$, where

$$T_t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}. \tag{1}$$

Similarly, rays are transformed due to optical refraction of a lens with focal length f according to $(q', p') = (q, p - \frac{1}{f}q)$, the linear transformation for which is $x' = L_f x$ where

$$L_f = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}. \tag{2}$$

In phase space, translation and refraction by a lens are both *shearing transformations.* As shown in Figure 2, translation adds horizontal shear, while refraction by a lens adds vertical shear. The shearing property of translation in particular will be an important part of the analysis in the body of this paper.

Given a radiance $r(q, p)$ at the image plane of a camera, an image $I(q)$ is rendered for a given range of the available p values according to

$$I(q) = \int_p r(q, p) \mathrm{d}p. \tag{3}$$

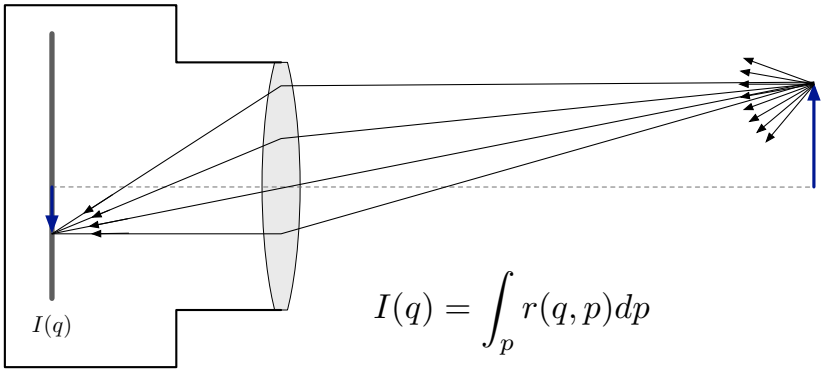
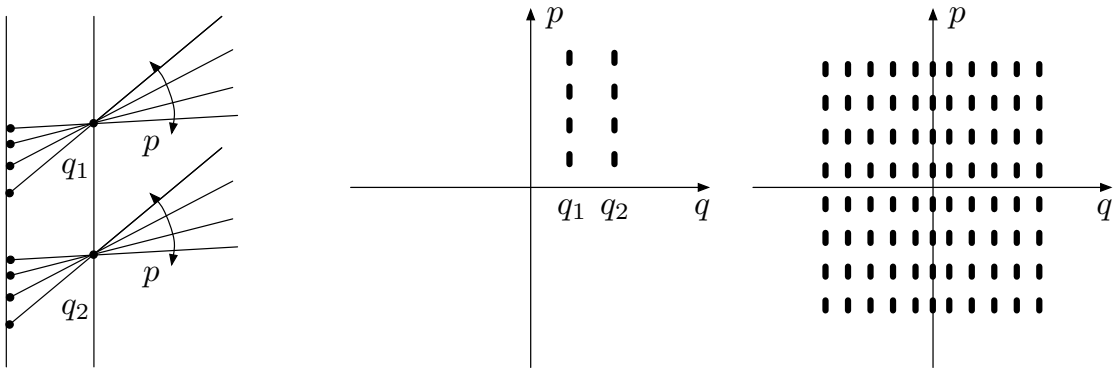


Figure 3: Each sensor pixel in a traditional camera captures its part of an image by physically integrating the intensities of all of the rays impinging on it.



(a) The rays that converge at a pinhole will separate from each other as they travel behind it and can therefore be captured individually by a sensor behind the pinhole.

(b) Individual pinholes sample one position in the q–p plane, while individual pixels sample different positions. A single image captured behind the pinhole thus samples a vertical stripe in the q–p plane

(c) An array of pinholes samples a grid.

Figure 4: An array of pinholes can be used to sample the plenoptic function by multiplexing angular information.

2.2 Plenoptic Cameras

As shown in Figure 3, a traditional camera captures an image by physically integrating the intensities of all of the rays impinging on each sensor pixel.

A plenoptic camera, on the other hand, captures each thin bundle of rays separately. One approach to separating, and then individually capturing, the rays in a scene is to put a pinhole where the sensor pixel would be while placing the sensor itself some distance b behind the pinhole. In this case, the rays that converge at the pinhole will diverge as they propagate behind the pinhole, as shown in Figure 4a. The separate pixels in the sensor now capture separate rays. That is, the intensity as a function of position at the sensor represents the radiance as a function of direction at the position q of the pinhole.

To see how the plenoptic function can be sampled, note that every ray can be uniquely described with a q and p coordinate, by a point in the q–p plane. Each pixel in the sensor behind a pinhole captures a distinct ray, i.e., it samples a distinct point in the q–p plane. Building up from the single pixel, a single image captured behind the pinhole samples a vertical stripe in the q–p plane, while an array of pinholes samples the plane. In practical cameras, pixels are discrete, so a small area in the q–p plane is sampled rather than a single point.

2.3 Lippmann Sensors

From a more abstract perspective, our approach is a general method of sampling the plenoptic function with respect to arbitrary parameters (or modes), (including, but not restricted to, the usual four dimensions of 2D position and 2D angle. The sensor we use to carry out this sampling consists of an array of lenslets that form microcameras focused at a given plane. We refer to this sensor as the "Lippmann sensor" since it was first proposed by Lippmann in his 1908 paper.[15] The original Lippmann sensor is shown in Figure 5. Our generalization is based on introducing different types of filters (or other modifiers) into the plenoptic function sampling process. In this regard, the filters serve a function similar to that of a Bayer array filter in a normal sensor, the difference being that the Lippmann sensor samples the full 4D radiance in optical phase space, as opposed to conventional sensors sampling the 2D irradiance.

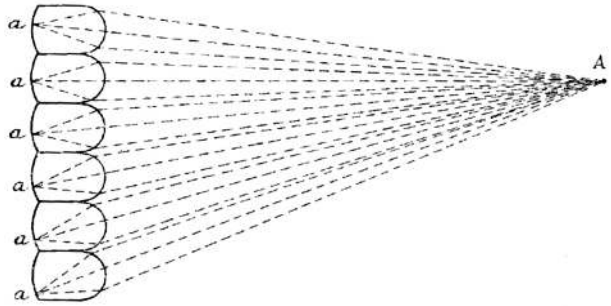


Figure 5: The Lippmann sensor capturing the image of a point A in the world, as an array of dots. The figure is taken from the original paper.[15]

Although the ideal pinhole makes an ideal "ray separator," microlenses are used in practice instead to gather sufficient light and to avoid diffraction effects. Figure 6 shows a diagram of such a "Lippmann sensor".[16] In the diagram, b is the distance from the sensor to the microlens plane, and a is the distance from the microlens plane to the main lens image plane. The microlens focal length is f; a, b, and f are assumed to satisfy the lens equation $1/a + 1/b = 1/f$. Sensor pixels have size δ, and, without loss of generality, we take d to be the microlens aperture and the spacing between microlenses.

The form of the Lippmann sensor that we use in this paper was proposed by Lumsdaine and Georgiev,[9] in which

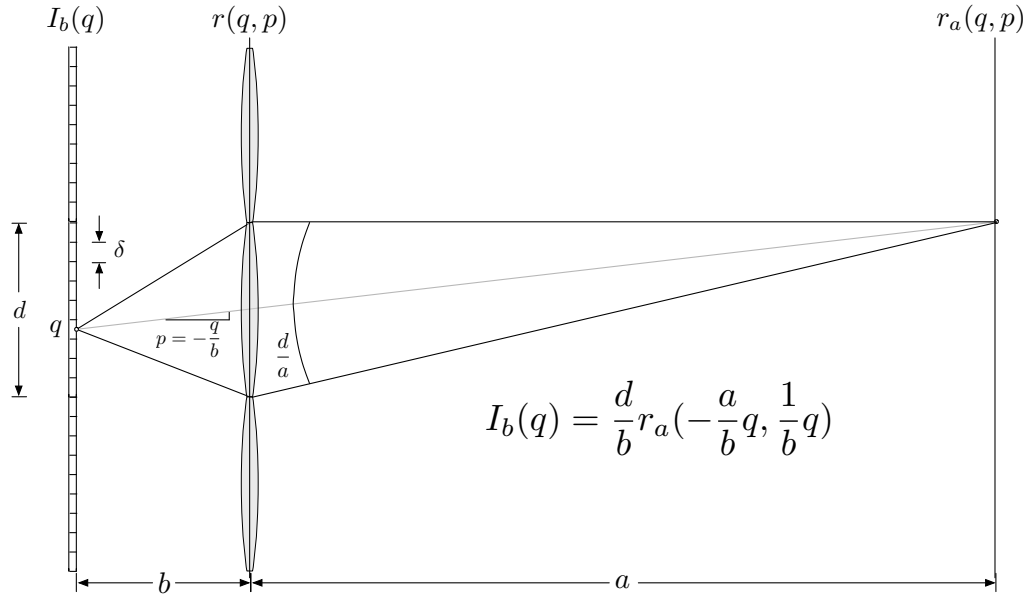


Figure 6: Geometry of the Lippmann sensor for a plenoptic camera. The CCD (or CMOS) sensor is placed at a distance b behind an array of microlenses. In one notable limit, $b \to f$ and $a \to \infty$.

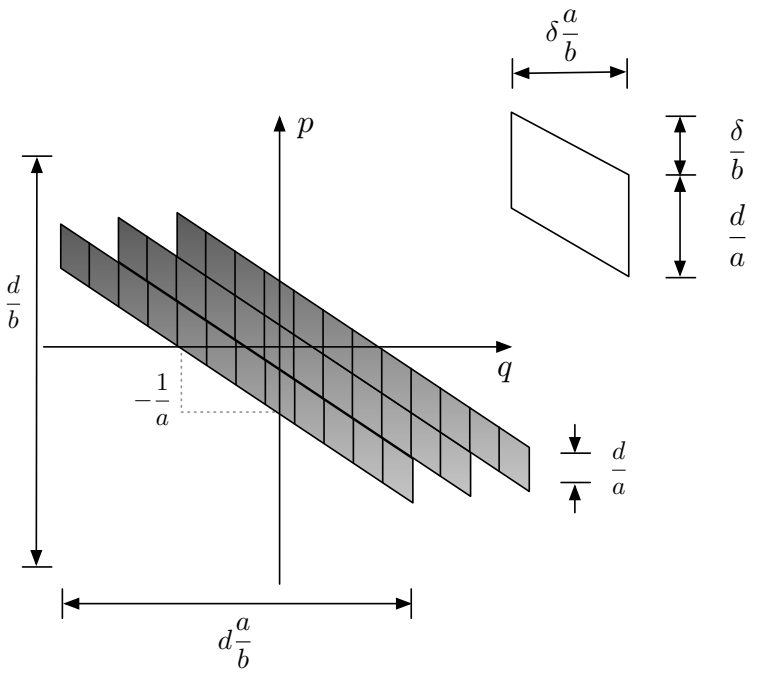


Figure 7: Radiance sampling by the focused Lippmann sensor at the imaging plane at distance a in front of the microlenses. Here, a, b, and f satisfy the lens equation. The geometry of a single pixel is shown in the upper right.

the distance b was chosen not to be equal to f in order to form a relay system with the camera's main lens. Since the microimages are focused in this case, we refer to the sensor as the focused Lippmann sensor and a camera using it as the "focused plenoptic camera". In this case, as derived in,[9] we have the following expression for how the image captured on the sensor (I_b) samples the radiance at the microlens front focal plane (r_a):

$$I_b(q) = \frac{d}{b} r_a(-\frac{a}{b}q, \frac{1}{b}q). \tag{4}$$

This sampling is shown graphically in Figure 7. As shown in the figure, and as reported in,[9] each focused Lippmann microimage captures a slanted stripe of the radiance at distance a in front of the microlens array.

3. MULTIFOCUS

As discussed in previous sections, the focused plenoptic camera uses an array of microlenses to re-image parts of the image plane of the main camera lens onto the CCD. This makes up the Lippmann sensor that samples the 4D radiance at a certain plane in front of the microlenses. For the purpose of computational focusing, the depth of field of those microcameras needs to be as large as possible: The range of computational focusing is defined by that depth of field. At the same time, there are fundamental optical limitations to the amount of depth of field that can be reached. Large depth of field is achieved in ray optics with small apertures, which in turn leads to high F-number and low light efficiency. This approach also has a fundamental limiting constraint: At high F-numbers diffraction plays an increasing role, blurring the images.

In order to provide a sensor that has both large apertures and a large depth of field, we propose a sensor design based on an array of interleaved microlenses having different focal lengths, so that they are focused at two or more different planes. If properly designed, the depths of field corresponding to differently focused microlenses will cover the entire space in front of the sensor (i.e., the complete interior of the camera body). Given an arbitrary point in the world, at least one of the sets of microlenses would be sharply focused on it. Now these microlenses could work at fully open apertures (low F-numbers), implementing a camera that has the highest resolution possible, and at the same time using all the light available.

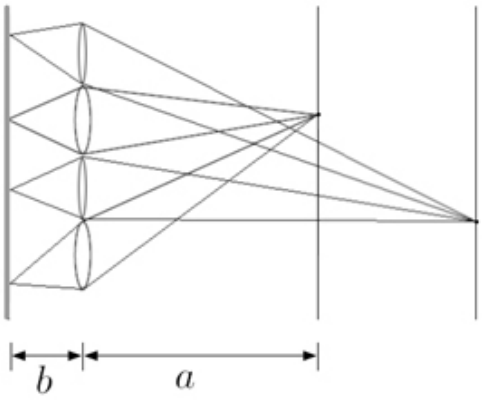


Figure 8: The multifocus Lippmann sensor capturing two types of images focused at two different depths.

4. RADIANCE CAPTURE WITH A MULTIFOCUS LIPPMANN SENSOR

In this section we analyze the radiance capture model in the case of an array of microlenses with a common CCD. This will motivate the need for a multifocus sensor. Capturing the full 4D radiance in object space is easily achieved with an array of identical microcameras. Each microcamera is focused on object space and records a different overlapping view of a piece of the scene. Note that at this overlap, two different cameras actually represent stereo views of the 3D object. Such a device for capturing the light intensity as a function in ray space (plenoptic function), was first proposed in.[15]

An optical phase-space diagram for the radiance as a function of ray position and angle in object space, recorded by the Lippmann sensor, was shown Figure 7. The size of the image created in a microcamera is d. This size is constrained by the main lens aperture, which needs to be chosen appropriately by matching the F-number of the main lens system to that of the microlenses.[17] The distances a and b define the minification of the microcamera. Considering this fact, the object-space size of the image for our microcamera is $d \times \frac{a}{b}$. The viewing angle of the microcamera is that size of the object divided by the distance to the object a, which can be shown to equal $\frac{d}{b}$. Each point is viewed by the microcamera from a range of angles depending on the microlens aperture. That range of angles is $\frac{d}{a}$.

Note that different points in object space are seen by the camera at different angles. This is represented by the tilt of the line of pixels depicting the microimage in Figure 7, and measured by the slope $-\frac{1}{a}$.

Figure 9 (right) shows how the Lippmann sensor samples the radiance at a plane different from the focal plane of the microlenses. Such a plane is a translation from the focal plane and so the sampled radiance is transformed by shearing. In other words, Figure 9 (right) is related to Figure 9 (left) by a shearing transformation. Note that since the pixels are now tilted in phase space, rendering from them (which is an integration in the vertical direction) will mix content from neighboring pixels, giving a blurred result.

To enable the Lippmann sensor to sample any location q with "vertical pixels", we use a multifocal version that interleaves microlenses with different focal lengths. The result is that one set of microlenses will be in focus for regions of the scene for which other sets are not in focus. The phase-space interpretation of this interleaving is shown in Figure 10. Focal lengths are chosen such that the distances from the microlens array to the focal planes are a and $a/3$.

4.1 Considering wave optics

There is an important constraint on spatial and angular resolution, the diffraction limit, that goes beyond ray optics and is fundamentally dictated by wave optics. Because of diffraction effects, arbitrarily thin pixels in spatial coordinates are unrealistic. Rather (as is well known), the smallest diffraction-limited spot resolved by a camera is $1.22\lambda\frac{a}{d}$, where $\frac{a}{d}$ represents the object-side F-number of the microlens in our notations. In the case of a 1D image (2D ray space), the diffraction limited spot size is $\lambda\frac{a}{d}$.

Considering that the vertical (angular) size of our ray-space pixels is $\frac{d}{a}$, and the horizontal size cannot be less than $\lambda\frac{a}{d}$, we come to a fundamental fact that the volume of a pixel in ray space (i.e., the product of spatial and angular extent of

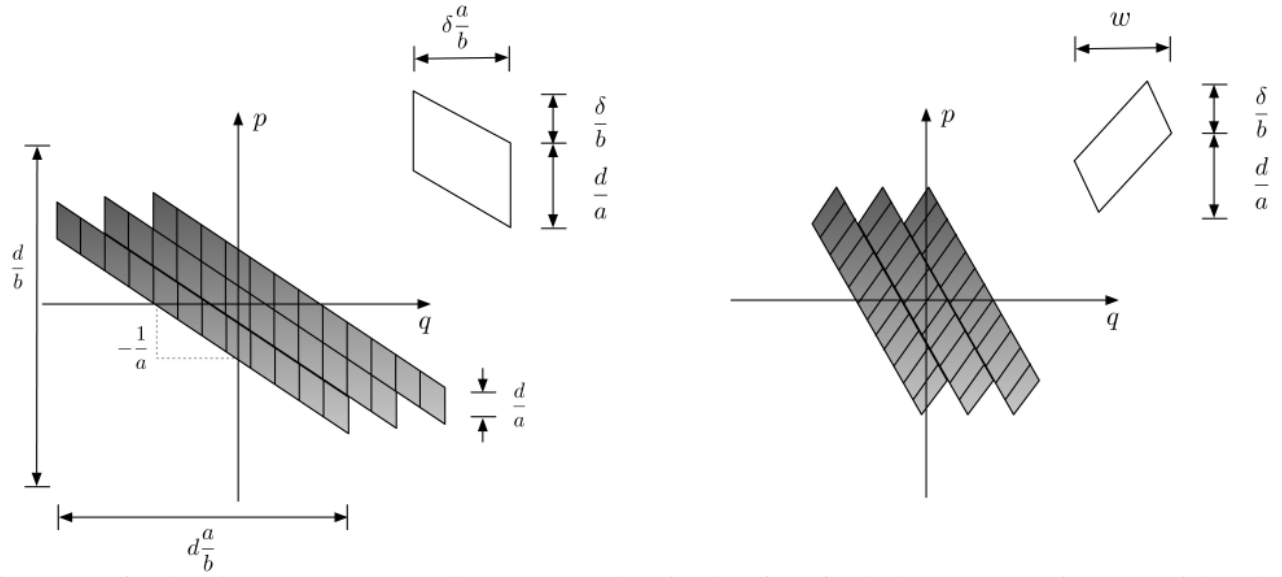


Figure 9: Left: The Lippmann sensor sampling pattern at the microlens front focal plane (same as Figure 7). Right: The same Lippmann sensor, now sampling the radiance at a position different from the microlens focal plane. The translation transformation shears the sampling pattern shown on the left to produce the sampling on the right.

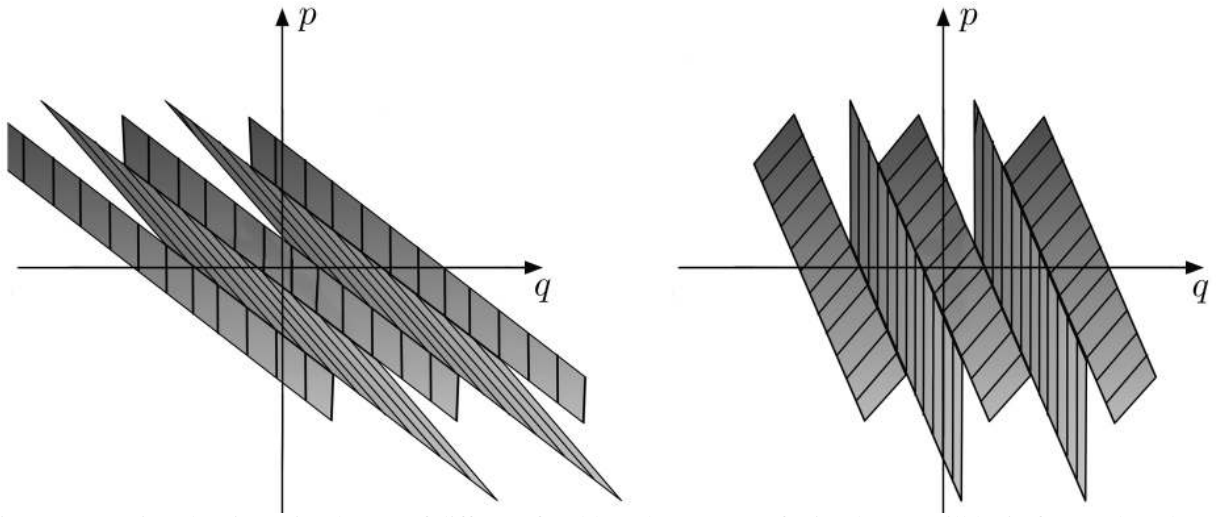


Figure 10: By interleaving microlenses of different focal lengths, one set of microlenses will be in focus when the other is not. Left: Sampling pattern at distance a in front of the microlenses. Right: Sampling pattern at distance $a/3$.

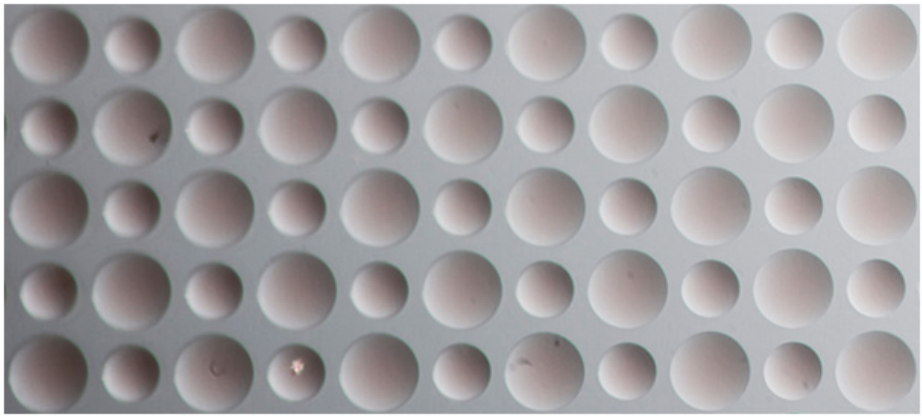


Figure 11: The array of microlenses used to capture the image in Figure 3. Notice the different focal lengths, which are coupled with the microlens diameter for our microlens array.

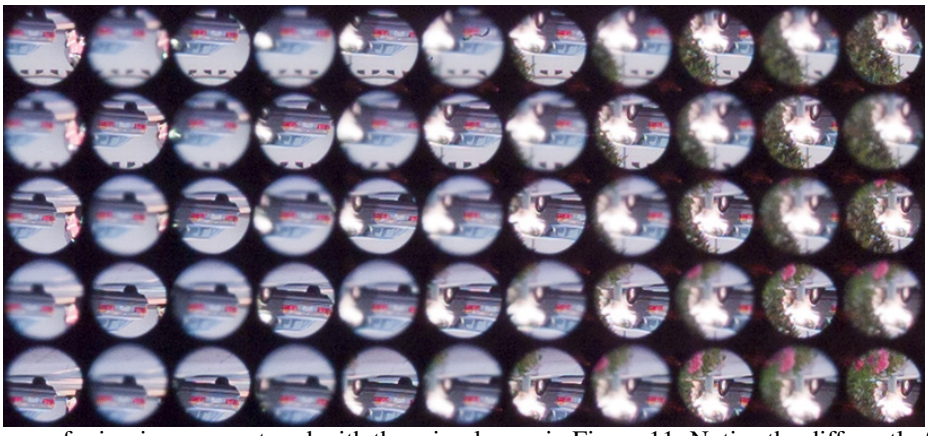


Figure 12: The array of microimages captured with the microlenses in Figure 11. Notice the differently focused microimages, interleaved.

a pixel) cannot be less than the wavelength λ. This fundamental fact has not been discussed in the context of plenoptic cameras.

The above diffraction constraint limits the depth of field of the microcameras in a Lippmann sensor. Only a certain range around a given plane of focus is imaged sharply. Points that remain outside that depth of field would be out of focus, and would appear blurry. What we achieve with the multifocus Lippmann sensor is that for any point in the world at least one microimage is exactly in focus. If that's the case, any object would be rendered in focus (using the appropriate stitching and blending techniques). This motivates our choice of different focal lengths for different microlenses.

5. EXPERIMENTAL RESULTS

In this section, we demonstrate interesting results that can be achieved in integral photography with a multifocus microlens array. Our point here is that all microimages need to be exactly in focus so that any object could be rendered in focus using the appropriate stitching and blending techniques. The problem with this is that microcameras have large but finite depth of field, so all-in-focus imaging is not really possible.

That's why we have chosen the approach of a microlens array with interleaved lenses of different focal lengths. A picture of our microlenses is shown in Figure 11. Notice the two different types of lenses having different focal lengths. An image captured with these microlenses is shown in Figure 12. It is important to notice how every other microimage in the array is in focus or out of focus.

We have applied plenoptic rendering to the microimages in Figure 12 to produce the left and right image in Figure 13. Rendering is best focused on the car (at optical infinity). Notice that the left image appears out of focus. That's because we

Figure 13: Images rendered in focus from the microimages in Figure 12. The image on the left is blurry.

Figure 14: Images rendered in focus from the microimages captured with the lenses in Figure 11. Notice that the flowers on the left are sharp. They are rendered from the same type of microlenses as in Figure 13 left. The flowers are close to the microlenses, so they are blurry. Also, the apertures are smaller, so a wider range of depths are in focus.

have rendered it from microimages that are themselves blurry. Their microlenses are focused at a plane closer than infinity.

Exactly the opposite phenomenon happens in Figure 14, which is rendered from the same input image. The left image comes from the type of microlenses used to create the left image in Figure 13. They are focused close in and the flowers are sharp. The right image is blurry because the microlenses are focused farther away.

Next, we show the two complete images, achieving best focus on both close and far objects. Such wide range of re-focusing is possible only because of the different focal lengths of our microlenses.

The full captured multifocus radiance is shown in Figure 17.

6. CONCLUSION

The seemingly ever-increasing resolution of image sensors opens up fascinating possibilities for the kinds of rich image data that might be captured by a camera. In particular, additional image dimensions, such as multiple views or multiple modes can be captured and then used computationally to create an infinite variety of rendered images. In this paper, we explored rich radiance capture for multiple depths of microlens focusing. With the multifocused Lippmann sensor we were able to extend the plenoptic depth of field to the whole scene, potentially covering every depth in focus. This enables the full power of refocusing and stereo that can be reached in ray and wave optics.

This is just one of the less trivial examples that demonstrate different aspects of what is possible with a device as versatile as the Lippmann sensor. Much more can be expected, including HDR, polarization, multispectral color, IR and UV and X-rays, and others. As sensor resolutions continue to grow even further, we look forward to being able to capture yet richer plenoptic data in new and unexpected so far ways, that give us the chance to render even richer images.

REFERENCES

1. E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting, Morgan Kaufmann, San Francisco, CA, USA, 2005.

Figure 15: Image rendered with the car in focus.

Figure 16: Image rendered with the flowers in focus.

2. R. Szeliski, H.-Y. Shum, Creating full view panoramic image mosaics and environment maps, in: SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, New York, NY, USA, 1997, pp. 251–258.
3. S. T. Barnard, M. A. Fischler, Computational stereo, ACM Comput. Surv. 14 (4) (1982) 553–572.
4. M. Z. Brown, D. Burschka, G. D. Hager, Advances in computational stereo, IEEE Trans. Pattern Anal. Mach. Intell. 25 (2003) 993–1008. doi:10.1109/TPAMI.2003.1217603.
5. M. Levoy, P. Hanrahan, Light field rendering, ACM Trans. Graph. (1996) 31–42.
6. S. J. Gortler, R. Grzeszczuk, R. Szeliski, M. F. Cohen, The lumigraph, ACM Trans. Graph. (1996) 43–54.
7. E. Adelson, J. Bergen, The plenoptic function and the elements of early vision, in: Computational models of visual processing, MIT Press, Cambridge, MA, 1991.

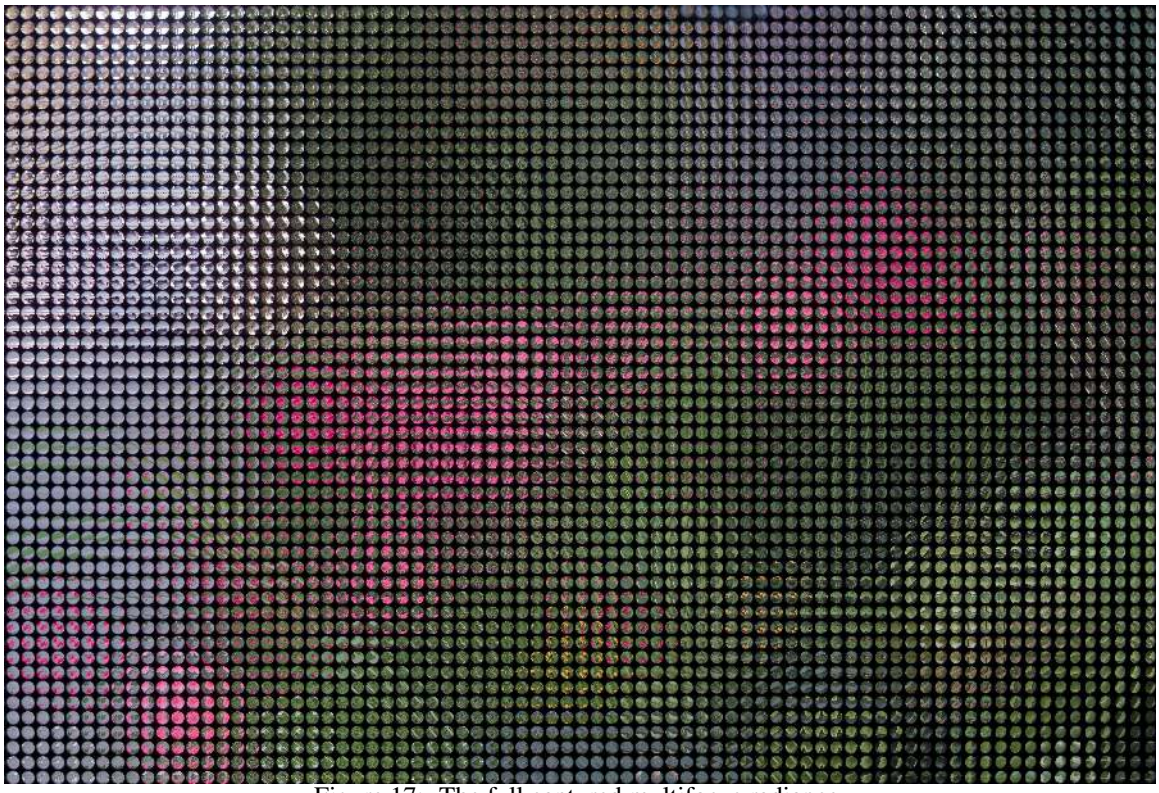


Figure 17: The full captured multifocus radiance.

8. T. Georgiev, A. Lumsdaine, S. Goma, High dynamic range image capture with plenoptic 2.0 camera, in: Signal Recovery and Synthesis, Optical Society of America, San Jose, CA, 2009.
9. A. Lumsdaine, T. Georgiev, The focused plenoptic camera, in: IEEE International Conference on Computational Photography (ICCP), 2009.
10. M. Levoy, P. Hanrahan, Light field rendering, Proceedings of the 23rd annual conference on Computer Graphics and Interactive Techniques.
    URL `http://portal.acm.org/citation.cfm?id=237170.237199`
11. F. E. Nicodemus (Ed.), Self-study manual on optical radiation measurements, National Bureau of Standards, 1978.
12. V. Guillemin, S. Sternberg, Symplectic techniques in physics, Cambridge University Press, New York, 1985.
13. A. Gerrard, J. M. Burch, Introduction to Matrix Methods in Optics, Dover Publications, Mineola, NY, 1994.
14. K. B. Wolf, Geometric Optics on Phase Space, Springer, New York, 2004.
15. G. Lippmann, Epreuves réversibles donnant la sensation du relief, J. Phys. 7 (4) (1908) 821–825.
16. G. Lippmann, Epreuves réversibles. Photographies intégrales., Académie des Sciences (1908) 446–451.
17. R. Ng, Digital light field photography, Ph.D. thesis, Stanford University, Stanford, CA, USA, adviser-Patrick Hanrahan (2006).