

Using Focused Plenoptic Cameras for Rich Image Capture

Todor Georgiev ■ Adobe Systems

Andrew Lumsdaine and Georgi Chunev ■ Indiana University

Edward Adelson and James Bergen introduced the *plenoptic function* as a means to completely represent the information necessary to characterize the radiance in a scene.¹ This formulation naturally includes a description of light rays' spatial distribution in 3D space. However, it also includes those rays' other properties, such as wavelength and polarization. Subsequent research

with the plenoptic function, however, has focused primarily on spatial distribution. One notable example is the light field, a 4D function that established a theoretical framework for describing and analyzing light rays' spatial distribution in a scene.² Cameras for capturing the plenoptic function have similarly emphasized capturing only the light field (the light-ray spatial distribution).

Here, we revisit the richer definition of the plenoptic function and present techniques to capture

information about radiance in a scene, in addition to the light field. In particular, we explore two complementary techniques for rich image capture with *focused plenoptic cameras*. (For more on plenoptic cameras, see the related sidebar.) The first technique multiplexes the captured plenoptic function at the microlens array, by interleaving microlenses with different properties so as to distinguish that plenoptic function's property. The second technique performs the same type of modulation at the main camera lens, yet still captures the multiplexed data

with a microlens array focused on the image. Figure 1 illustrates these techniques.

We illustrate our approach to rich image capture with the primary example of high-dynamic-range (HDR) imaging. We also extend the second technique to polarization and multispectral color capture.

HDR with the Focused Plenoptic Camera

The most popular HDR approach allocates pixels temporally. That is, it photographs the same scene multiple times at different exposures and merges the resulting images into one floating-point image (which can support extended dynamic range). In the simplest case, a final image consists of piecewise linear transfer curves with different slopes. For display purposes, this approach often uses tone mapping or other HDR compression techniques to produce a low-dynamic-range output image while preserving contrast and image details.³

For a moving scene, we must allocate pixels spatially because the camera can't take multiple exposures at different times. A plenoptic camera captures multiple views of the same scene. By appropriately filtering those views, the camera can capture multiple exposures of a scene at the same time. With traditional plenoptic cameras, one microlens captures all the different views of a particular point in the scene. So, you can separately filter different views by filtering at the main lens.⁴ Under the Lambertian assumption, radiance is constant for all views, so this approach is equivalent to filtering the same scene with different filters. With the focused plenoptic camera, different microlenses capture different views. So, you can

This approach uses a focused plenoptic camera to capture the plenoptic function's rich "non 3D" structure. It employs two techniques. The first simultaneously captures multiple exposures (or other aspects) based on a microlens array having an interleaved set of different filters. The second places multiple filters at the main lens aperture.

separately filter different views in two ways: at the microlenses or the main lens.

Filtering at the Microlenses

The focused plenoptic camera's main lens forms an image at its image plane. An array of microlenses behind that plane maps the image from the image plane to the sensor.⁵

Each microlens creates its own little image of part of the scene, as seen through the main lens aperture, which defines the image shape and size. In other words, each microlens works as a micro-camera, and we can determine its exposure by placing a neutral-density (ND) filter in front of it (see Figure 1a).

In the focused plenoptic camera, the main lens image is formed at distance a in front of the microlenses. We're careful to adjust and fix the appropriate distance b to the sensor, to satisfy the lens equation

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}, \quad (1)$$

with the main lens image at distance a ; f refers to the focal length of the microlens. Fine-tuning these distances can make the captured image smaller than the main lens image by a factor of N , where $N = a/b$.

So, the resolution depends on the distance between the microlenses and the images created in the camera (which, in turn, depends on the scene depth). Rendering an entire scene with a fixed resolution will thus produce artifacts if the scene has sufficiently different depths. We require that $N > 2$, so that every point in the scene is seen at least two times—that is, at least once for each filter. At the same time, because we lose a factor of N in our

Plenoptic Cameras

Plenoptic cameras have been used to capture the 4D light field by multiplexing it onto a conventional 2D sensor. The light field represents the light rays' positional and angular information. Capturing a richer plenoptic function involves fully using the camera's multiplexing capabilities to perform extended sampling so that more information is captured with the same limited sensor capabilities.

How you allocate sensor pixels for capturing additional data about the plenoptic function bears directly on the rendered images' achievable resolution. Ren Ng and his colleagues' handheld plenoptic camera used a microlens array in front of the camera sensor and produced one positional sample per microlens.¹ This scheme resulted in a resolution of the final image that was equal to the number of microlenses, 300×300 .

More recently, the focused plenoptic camera introduced an approach that sampled more sparsely in the directional coordinates.² This approach produces final rendered images comparable to those of regular, nonplenoptic, cameras. Our approach is based on this camera. Besides higher spatial resolution, it enables flexible trade-offs in how you sample the plenoptic function. Both these features are essential for efficiently capturing additional information about the plenoptic function.

References

1. R. Ng et al., *Light Field Photography with a Hand-Held Plenoptic Camera*, tech. report CSTR 2005-02, Computer Science Dept., Stanford Univ., 2005.
2. A. Lumsdaine and T. Georgiev, "The Focused Plenoptic Camera," *Proc. IEEE Int'l Conf. Computational Photography (ICCP 09)*, IEEE Press, 2009.

final rendered image's spatial resolution, we would like to keep N small.

Consider the phase-space diagram of image capture in Figure 2. Following a position-direction convention, we denote the 4D radiance at a given

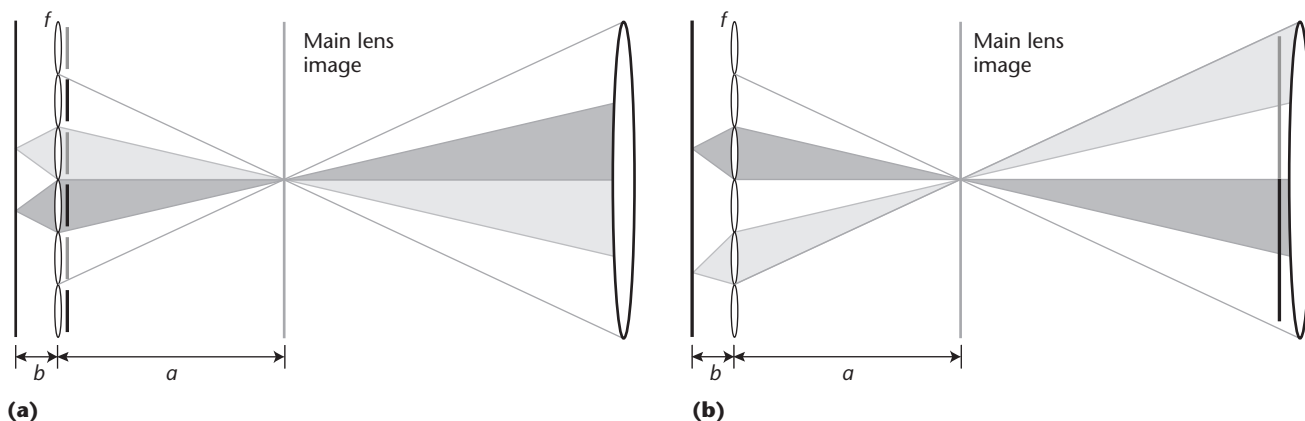


Figure 1. The focused plenoptic camera can be viewed as a relay system. We interleave filters on the (a) microlens apertures or (b) main camera lens aperture. f is the focal length of the microlens, b is the distance to the sensor, and a is the distance from the microlenses to the main lens image.

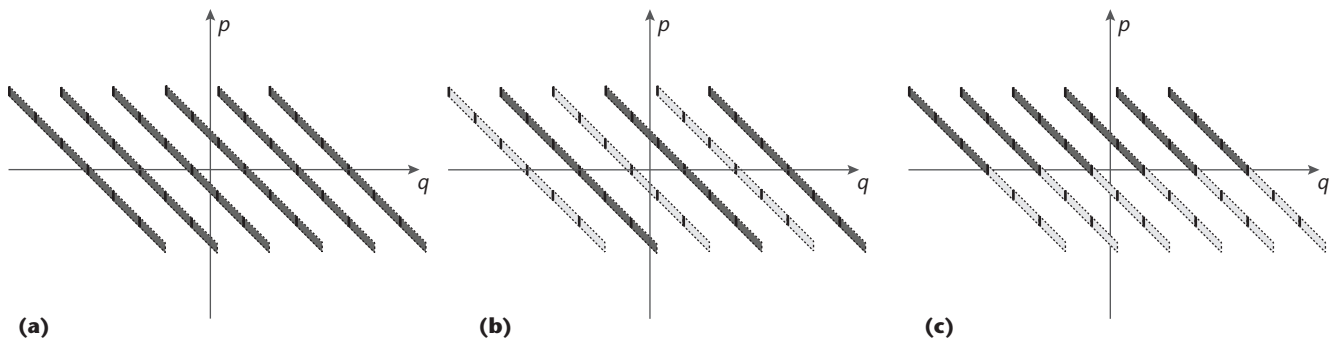


Figure 2. A phase-space diagram of radiance capture in the focused plenoptic camera. (a) The conventional plenoptic camera. (b) The camera with interleaved filters at the microlenses. (c) The camera with filters at the main lens. Similarly filtered samples are in the same color (black or gray). q describes a ray's location, and p describes the ray's direction.

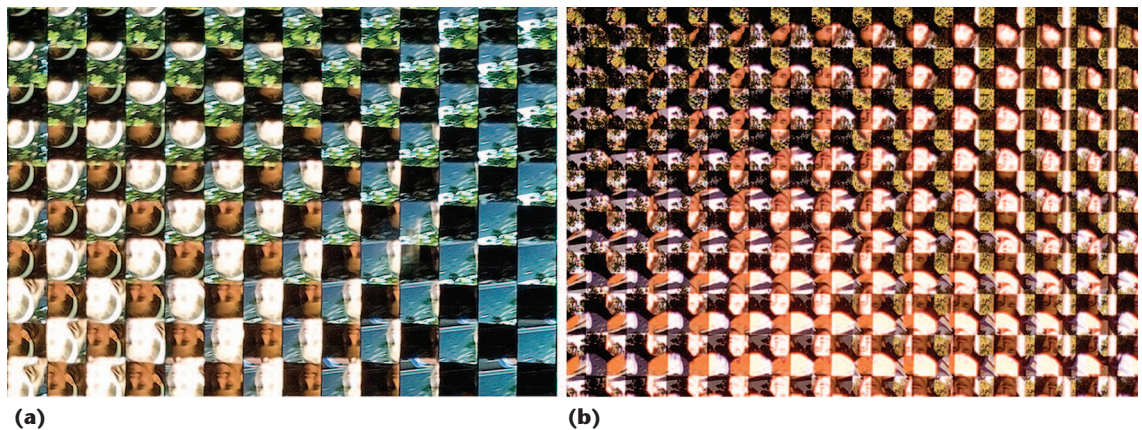


Figure 3. Multiple capture of an image by the microlenses with (a) interleaved filters at the microlenses and (b) four filters on the main camera lens, which split each microimage into four parts. The focused plenoptic camera captures multiple copies of parts of the image, one for each microlens. Modifications of those microimages through appropriate filters achieve the main effects described in this article.

plane perpendicular to the optical axis as $r(q, p)$, where q describes a ray's location in the plane and p describes the ray's direction. In 2D space, $p = \tan\theta$, where θ is the ray's angle relative to the optical axis. The image formed behind each microlens samples the plenoptic function at the main lens image plane.⁵ Sampling occurs in a tilted fashion defined by the system's optical transfer matrices. We adjust the parameters such that images in the microlenses overlap; that is, the same spatial coordinate (q) corresponds to two or more images. By placing ND filters on the microlenses, we interleave different sampling types, as the bright and dark slanted lines in Figure 2b illustrate. In this case, we show simple modulation with two filters, but the extension to more filters should be clear.

Figure 3a shows a crop from the array of microimages created by the microlenses. Because different microlens filters are interleaved, we observe respectively bright and dark microimages, each with the shape of the main lens aperture. This is the interleaved pattern in Figure 2b. Notice that we have made the main lens aperture square so

that microimages are square and tile together with little loss of space between them.

The microimages captured in this way with the thousands of microcameras show a decrease in the size of captured structures relative to those of the focal-plane image. As we discussed earlier, the reduction is by a factor of N , typically chosen to be between 2 and 5.

Filtering at the Main Lens

We can separately filter different views by placing filters at the main lens. This solution is easier to implement because the physical dimensions are on the order of millimeters and can be manipulated by hand. Considered at one group of microlenses, the light modulation pattern would be similar to that in the microlens filter design (see Figure 1b).

The phase-space diagram (Figure 2) shows the difference between the two filtering techniques. Filtering at the microlenses spreads the pattern of k elements over k microlenses. With filtering at the main lens, the whole pattern is visible in each

microimage. (Figure 2 shows simple modulation with $k = 2$ filters.)

Figure 3b shows a crop from the array of microimages created by the microlenses when using four filters on the main lens aperture. Each microimage has four bright and dark areas. Compare this with Figure 2b. The pattern change appears twice as fast along each axis, compared to a similar image captured with interleaved filters on the microlenses in Figure 3a. Again, we make the main lens aperture square, for the reasons we explained earlier.

Resolution Analysis with Main-Lens Filtering

You can also put filters at the main lens with traditional plenoptic cameras.⁴ Because the microlenses are focused at the main lens aperture, this technique is good for discriminating between different filters. For each pixel in the image, we know exactly which filter has influenced it. However, this technique doesn't produce optimal results.

In the focused plenoptic camera, the microlenses focus on the image the main camera lens created, not on the main camera lens. As Figure 4 shows, the main lens aperture will be blurred in each microimage, whereas the microimage itself is sharp. Any structures, such as filters, placed on the main lens aperture will be out of focus.

The microimages' sharp in-focus content makes our final resolution much higher than that of the traditional plenoptic camera. For instance, in the example in Figure 4, we rendered a sharp stereo image of 5 megapixels for each view, starting with a 39-megapixel input. A traditional plenoptic camera⁶ would produce a 0.22-megapixel image with the same sensor.

So, we have two trade-offs:

- Focusing the microlenses on the main image causes the filter boundaries to blur but keeps the image of the scene sharp. The ability to resolve the filtering for individual pixels is limited.
- Focusing the microlenses on the main lens keeps the filter boundaries sharp but causes the image of the scene to blur. The overall image resolution is limited.

To quantify these trade-offs, we estimate the main lens aperture blur as it appears in a micro-lens image. Consider the camera system in Figure 5. As in the traditional plenoptic camera, microlenses are small, and we can assume the main lens is at optical infinity. The main lens aperture P is imaged from optical infinity to P' , a distance f behind each microlens, whereas the image of the scene forms at a distance b behind the microlens,

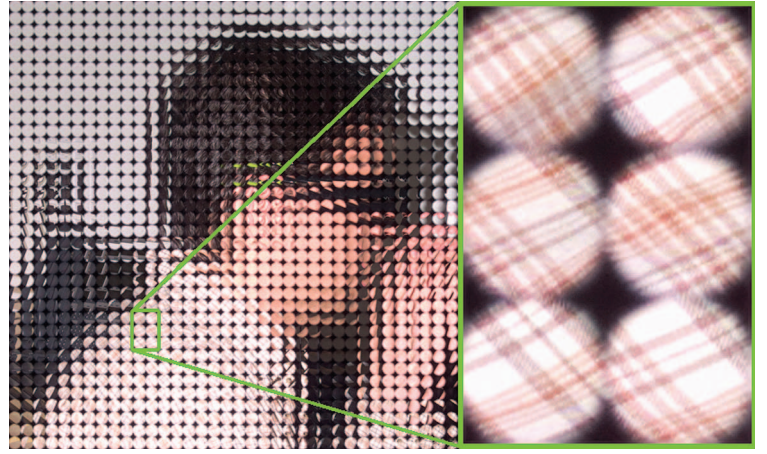


Figure 4. With the focused plenoptic camera, the main lens aperture is out of focus in each microimage. Note the fading edges of the microimages on the right.

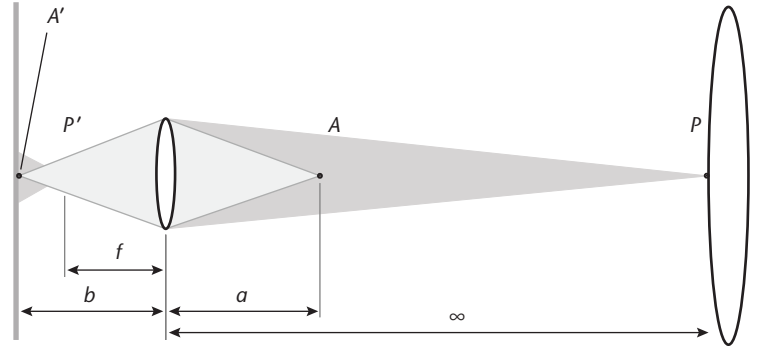


Figure 5. A point P on the main lens aperture (at optical infinity) is imaged to point P' behind the microlens and creates a circle of confusion on the sensor. At the same time, a point A from the main lens image is imaged to A' exactly on the sensor.

where the sensor is. We're interested in the blur caused by this misfocus of the main lens aperture.

Using the lens equation in Newton's form,

$$(a - f)(b - f) = f^2,$$

the distance between the image of the scene and the image of the main aperture is

$$b - f = \frac{f^2}{a - f}.$$

If the microlenses have microapertures of diameter d , the diameter of the circle of confusion for the main aperture's image will be

$$\Delta = \frac{f^2}{a - f} \frac{d}{f}.$$

From Equation 1, we have

$$\Delta = \frac{f^2}{af \left(\frac{1}{f} - \frac{1}{a} \right)} \frac{d}{f} = \frac{bd}{a}.$$

Another useful expression for the previous equation is in terms of the magnification factor $N = a/b$ —that is, $\Delta = d/N$.

This blur's upper limit is $\Delta < d/2$, which corresponds to the minimal applicable $N = 2$. A typical blur we get is at approximately $N = 5$, which would be approximately three pixels with our equipment (see the next section).

This blur is only at the boundaries between filtered regions in each microimage, so only a small percentage of pixels becomes unavailable for use in reconstructing the final image. Also, in terms of such blur, this method is typically five times better than the same camera using an array of pinholes,

For our 39-megapixel sensor, at least in theory, we could capture four separate modes at more than 9 megapixels each.

which has a blur of no less than the pinhole diameter, $\Delta = d$.

A similar calculation shows that if we focus the microlenses on the main lens aperture, the blur of the image of the scene will be approximately the same amount:

$$\Delta = \frac{d}{N+1}.$$

However, now the blur is across the entire microimage.

This analysis shows that focusing on the image has significant advantages over focusing on the main lens. As we mentioned before, when focusing on the image, we lose a few pixels at the filter edges, but the image itself is sharp. The loss in pixel count in our case is approximately 10 percent. However, we can compensate for this by computing the influence of the filters' defocus blur on individual pixel values at the filter image boundaries, and restoring their "unmixed" values. On the other hand, if we focus the microlenses on the main lens aperture, the filter edge boundaries are sharp, but the image itself is blurred. For example, a three-pixel blur would be equivalent to $3 \times 3 = 9$ times lower resolution in terms of pixel count. The difference in resolution is fundamental to these two methods, with significantly better resolution from focusing on the image.

So, the focused plenoptic camera efficiently uses sensor pixels. For example, without any optimiza-

tion for pixel use, our current camera produces 1.3-megapixel images ($1,300 \times 1,000$ pixels). Our images are slightly blurry (we've reduced image size for this article), but this is due to the particular optics we used and isn't a fundamental constraint of our imaging model. For comparison, Ren Ng and his colleagues reported 300×300 images,⁶ and Roarke Horstmeyer and his colleagues reported 177×177 images.⁴ With more careful camera design, we should be able to approach a factor of k reduction in the total sensor resolution when capturing k modes. For our 39-megapixel sensor, at least in theory, we could capture four separate modes at more than 9 megapixels each. In the case of Figure 4, we've been able to render four 5-megapixel views of the scene.

You could apply focused plenoptic rendering to the traditional plenoptic camera if you appropriately spaced the main lens image from the microlenses. However, because the traditional plenoptic camera uses defocused microlenses, this image would be blurred. You could achieve additional resolution by deblurring (for example, with deconvolution techniques), but with loss of quality. The bottom line is that the traditional and focused plenoptic cameras differ fundamentally. In the focused plenoptic camera, the sensor plane is conjugate to the object plane; in the traditional plenoptic camera, it isn't. This difference translates into the difference in the final rendered image's resolution and quality.

Regarding specular highlights (or other non-Lambertian scenes), main lens filtering and microlens filtering produce different results. From Figure 2b, we see that for large N , the configuration with filters at the microlenses will sample densely in the angular dimension. This means it will work well for specular and other non-Lambertian scenes, as well as for other plenoptic rendering tasks such as refocusing. The configuration with filters at the main lens always has certain views tied to certain filters, so it can't exploit large numbers of views in the same way (see Figure 2c).

The Experimental Setup

Our camera is medium format with an 80-mm lens and a 39-megapixel digital back from Phase One. We mounted the lens on the camera with a 13-mm extension tube, which provides the required spacing a between the microlens array and the main lens image. Our experiments evaluated two configurations at the microlens array: a microlens array with different apertures and an array of ND filters with different transmittance (see Figure 6).

The microlens array is custom-made by Leister

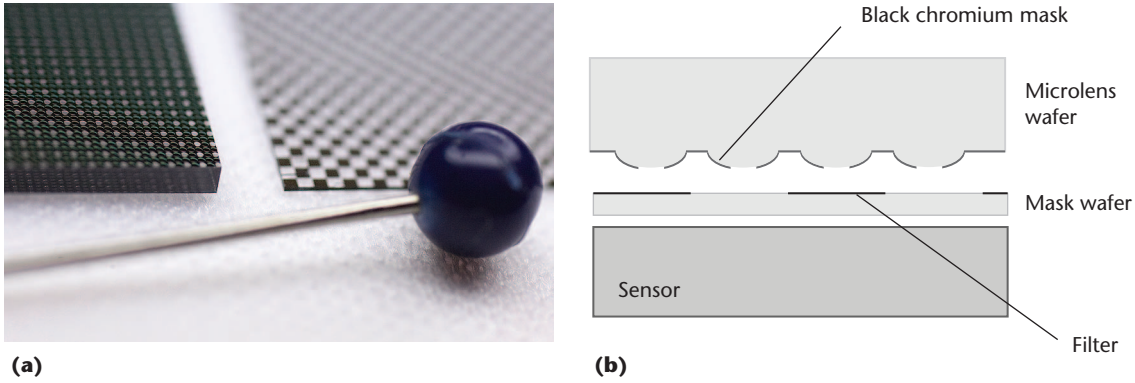


Figure 6. The experimental setup at the camera sensor. (a) A macrophotograph of our microlens array, with apertures and individual neutral-density microfilters on a separate mask wafer. A pin is placed in front for comparison of sizes. (b) A diagram of sensor, microfilter, and microlens placement.

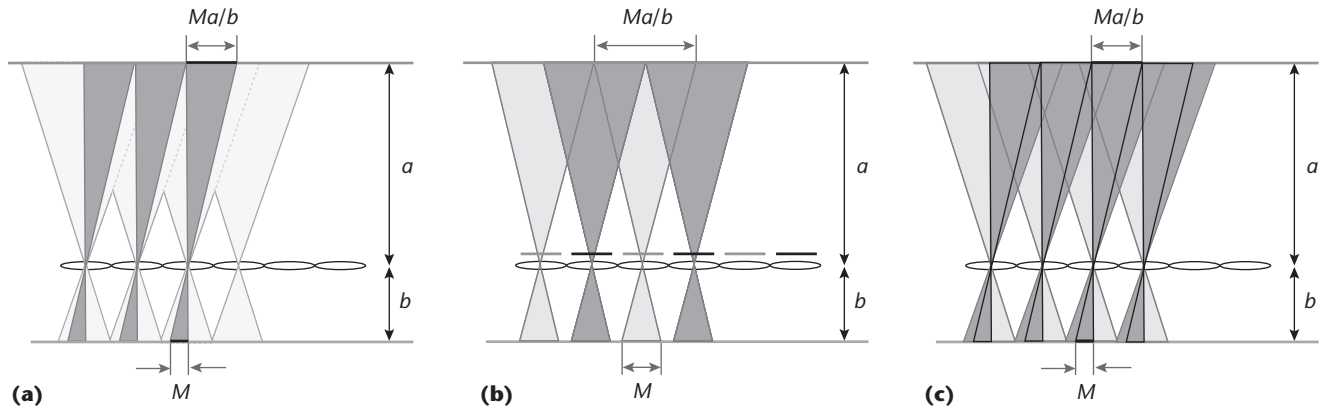


Figure 7. The geometry of image capture in the case of (a) basic focused-plenoptic rendering, (b) interleaved microlens filters, and (c) filters at the main camera lens (which is at infinity). With interleaved microlens filters, each filter affects an entire microimage. With filters at the main camera lens, each filter influences a given part of each microimage. M is the sample size.

Axetris. Each microlens has a focal length of 1.5 mm so that we can place it directly on the sensor's cover glass. We can provide additional spacing of up to 0.5 mm to enable fine-tuning of the microlens focus. The microlenses' pitch is 500 μm , with 1- μm precision. The sensor pixels are 6.8 μm . We estimated b to be approximately 1.6 mm with 0.1-mm precision from known sensor parameters and independently from the microlens images at different f-numbers.

The microlens apertures are formed by a black chromium mask deposited on the microlenses and are circular, each with a 100- μm diameter. This small size extends the depth of field and limits the microlenses' diffraction. However, it introduces a high f-number, with associated diffraction blur and longer required exposure times. These small apertures aren't generally required, especially with microlenses of good optical quality. We chose the microlenses' pitch to match the main camera lens's f-number, which can be as low as $f/3$. Because our microlenses are far (1.6 mm) from the sensor, as constrained by the cover glass, we need a relatively large pitch.

For our experiments, the ND filter array was an array of 0.5-mm square filters, deposited as chromium masks on a thin 0.21-mm silica wafer. The interleaved ND microfilters form a checkerboard pattern with transmission of 100 percent and 6.25 percent, enabling a dynamic-range expansion of four stops. We didn't apply antireflectant coating to the microlens array or the ND filter array so that we could evaluate the severity of any artifacts that might be produced without it.

To introduce filters at the main lens aperture, we opened the main camera lens and introduced the square aperture along with various filter types at the original main lens aperture location. We experimented with ND filters, polarizers, and a large set of color filters. In addition, we experimented with no filters, which you can consider as rich image capture of stereo views.

Experimental Results

To make rich image data captured by the focused plenoptic camera available for subsequent processing, we render separate images for each independently captured mode, using the focused-plenoptic



Figure 8. Patching together a final image, using the microlenses. (a) A bracketed image from a single exposure, using the filtered microlenses. (b) Another such image, using the filtered microlenses. (c) The final tone-mapped high-dynamic-range (HDR) image.

rendering algorithm we described in a previous paper.⁵ In this rendering process, we represent the captured radiance as a 4D array. The algorithm selects contiguous samples of size M from each microimage, spaced at equal distances, and patches them together to generate a view.

Figure 2 shows this process in phase space; Figure 7 shows it geometrically.

Interleaved Filters on the Microlenses

To render separate images when we use interleaved filters on the microlenses, we modify the focused-plenoptic rendering algorithm to accommodate allocation of pixels to differently filtered images. For interleaved filters, we have interleaved plenoptic images. We render the final image by applying the basic algorithm using only images of one mode with a modified pitch size. We multiply the pitch size by a constant factor—in this case, two. Consider the phase-space diagram in Figure 2b and the geometric diagram in Figure 7b. With sampling at twice the normal pitch, the rendering algorithm will create a final image from all the black samples or all the gray samples (depending on the choice of the initial sample).

In Figures 8a and 8b, we patch together a final image using the unfiltered and filtered aperture microlenses (see Figure 3a). Here, the ND filter reduces light by a factor of 16. Figure 8c shows the captured images merged into a final HDR image, using Photoshop. (For the merged image, we simply use a standard solution because our main point is image capture, not tone mapping and HDR processing.)

Filters on the Main Lens

Here, we interleave plenoptic images using the interleaving in the microlens images. So, we can render the final image simply by applying the basic focused-plenoptic rendering to specific portions of each microlens image. Consider the phase-space diagram in Figure 2c and the geometric diagram in Figure 7c. By choosing samples from the left side

of each microimage (positive p), we render a final image from the black samples. With samples from the right side of each microimage (negative p), we render a final image from the gray samples.

In Figures 9a and 9b, we patch together a final image using the main lens's unfiltered and filtered parts (see Figure 3b). Again, the ND filter reduces light by a factor of 16. Figure 9c shows the captured images merged into a final HDR image.

Analysis

Capturing dynamic rather than static scenes is a unique strength of our approach. Approaches using multiple samples of each pixel⁷ or variable pixel sizes⁸ can only partially match our results. Our approach is optical and is applied before digitizing the image, which brings the main advantage. Also, it can work in combination with any other approach providing an enhancement of the effect. (For more on other approaches, see the “Related Work in Rich Image Capture” sidebar.)

Our approach has some similarities with the variable-pixel-size approach. Both involve loss of resolution proportional to the number of different exposures. Whereas the variable-pixel-size approach resembles the Bayer array approach (see the “Related Work” sidebar) and requires a blurring (antialiasing) filter, our approach has perfect sharpness but produces an image size reduced by a factor equivalent to the blur. However, our approach has slightly better quality, similar to the manner in which a Foveon sensor is better than a Bayer array sensor, when sampling each color at exactly the same location. Under the Lambertian assumption, our “assorted pixels” are effectively on top of each other; that is, they differ in only angular, not spatial, coordinates. This works for the plane in focus. When parallax is visible, our approach might not necessarily exhibit these advantages.

Unlike other approaches, ours avoids blooming. If a pixel is brightly illuminated so that neighboring pixel values in the sensor are damaged by bloom-

Related Work in Rich Image Capture

Researchers have studied capturing richer information about a scene in the context of both traditional and plenoptic photography. Capturing more information means capturing more pixels, with some means of allocating pixels for the richer capture. Typically, approaches for rich image capture allocate pixels either spatially, which captures the information in a single exposure, or temporally, which captures the information across multiple exposures.

A familiar instance of spatial allocation is the Bayer filter, which captures color digital photographs by interleaving red, green, and blue pixels across a sensor. The *assorted-pixels* approach extends this idea to other rich image data, such as from high-dynamic-range (HDR) photography, polarization, and multiple spectra.¹ The *split-aperture* approach also allocates pixels spatially.² However, it performs a coarse allocation by splitting the incoming image, filtering the pixels separately, and capturing the filtered scenes separately.

The temporal-allocation technique of *generalized mosaicing* captures a sequence of images and applies different optical settings to each image.³ Researchers have applied generalized mosaicing to capture multispectral, HDR, and polarization image data.

Roarke Horstmeyer and his colleagues proposed using a plenoptic camera with filters on the aperture to capture HDR and other multimodal images.⁴ (For more on plenoptic cameras, see the other sidebar.) The multimodal camera uses pinholes rather than microlenses, and the main lens focuses on the pinholes. Because each modal image contains only one pixel from each pinhole, we can classify the camera as a traditional plenoptic camera. Amit Agrawal and his colleagues described a reinterpretable imager that uses a static mask at the camera sensor along with a dynamic mask, enabling the capture of the light field's temporally varying aspects.⁵

The approach we present in the main article extends some of our previous research.⁶ It employs spatial allocation; that is, we capture a single exposure that contains sets of pixels that have been exposed to different optical settings. Our approach is unique in that we developed it in the context of the focused plenoptic camera, which enables higher resolution and flexible trade-offs of captured image parameters.⁷

References

1. S.G. Narasimhan and S.K. Nayar, "Enhancing Resolution along Multiple Imaging Dimensions Using Assorted Pixels," *IEEE Trans. Pattern Analysis and Mach. Intelligence*, vol. 27, no. 4, 2005, pp. 518–530.
2. M. Aggarwal and N. Ahuja, "Split Aperture Imaging for High Dynamic Range," *Int'l J. Computer Vision*, Jan. 2004, pp. 7–17.
3. Y.Y. Schechner and S.K. Nayar, "Generalized Mosaicing," *Proc. 8th IEEE Int'l Conf. Computer Vision (ICCV 01)*, IEEE Press, 2001, pp. 17–25.
4. R. Horstmeyer et al., "Flexible Multimodal Camera Using a Light Field Architecture," *Proc. IEEE Int'l Conf. Computational Photography (ICCP 09)*, IEEE Press, 2009.
5. A. Agrawal, A. Veeraraghavan, and R. Raskar, "Reinterpretable Imager: Towards Variable Post-Capture Space, Angle and Time Resolution in Photography," *Computer Graphics Forum*, vol. 29, no. 2, 2010, pp. 763–772.
6. A. Lumsdaine and T. Georgiev, "Rich Image Capture with Plenoptic Cameras," *Proc. 2010 IEEE Int'l Conf. Computational Photography (ICCP 10)*, IEEE Press, 2010.
7. A. Lumsdaine and T. Georgiev, "The Focused Plenoptic Camera," *Proc. IEEE Int'l Conf. Computational Photography (ICCP 09)*, IEEE Press, 2009.

ing, we can pick those pixels' representation from another microlens image captured with a darker ND filter (albeit with a loss of dynamic range). For comparison, in approaches such as the variable-pixel-size approach, if a pixel with a darker filter is a direct neighbor to a bright one such that blooming occurs, the damage is irreparable.

Considering our microlens filter design (see Figure 6), the use of two separate wafers increases interreflectance and microlens glare. The filter (mask) wafer reflects light back toward the microlens wafer, which in turn reflects some of it back. Several interreflections can cause lens glare (see Figure 8).

This effect resulted from an intentional design decision. As we discussed earlier, we applied no antireflectant coating to the microlens array or the ND filter array. A comparison with the main lens aperture modulation approach shows the differ-

ence. Microlens glare's effect is especially strong in dark microimages that are close to bright ones. Because we used photographic-quality filters on the main lens, the images in Figure 9 don't have as strong of a glare problem and are generally sharper and of higher quality. We confirm that antireflectant coating is important for HDR imaging.

Another issue with our approach is that because the microimages sample at a tilt in phase space, some small parallax exists between images produced from different modes. Applications that combine the modes to form the final image (such as HDR) might need to account for the parallax. However, in many multimodal applications, properly aligning image features for the combining process is an established procedure. Parallax depends on the main aperture's diameter—it's small with point-and-shoot cameras and with



Figure 9. Patching together a final image, using the main lens. (a) A bracketed image from a single exposure, using the main lens's unfiltered part. (b) Another such image, using the main lens's filtered part. (c) The final tone-mapped HDR image.

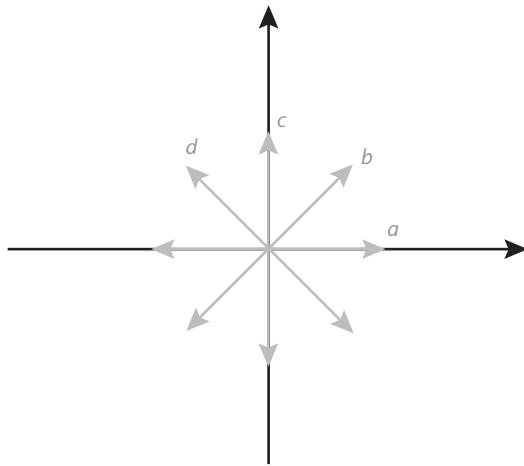


Figure 10. This diagram specifies the direction of polarization for the four images we capture.

any camera at large distances. In such cases our method works best.

Our experiment produced additional artifacts. The small white squares, mainly in the bottom left of the images, are due to defective microlenses. They aren't significant for evaluating our approach because better manufacturing can eliminate them.

Beyond HDR

Using filters to capture additional dimensions or data types in plenoptic cameras isn't restricted to HDR image capture; it's a general approach. Placing filters at the main lens aperture lets you easily (and economically) change the filter type. Here, we show other types of rich image capture that you can readily accomplish with our camera setup, simply by using different filter types. As with the HDR image capture, you can sample other modalities by placing appropriate filters at the main lens aperture or at the microlenses. Here, we report on our implementation with filters at the main lens aperture.

Polarization Capture

You can describe the complete polarization state of light from a natural scene with three images:

- one for the unpolarized light and
- one each for the two orthogonal components of the linearly polarized light.

(We assume there's no rotationally polarized light.) Accordingly, to capture sufficient information from a scene to reconstruct these basis images, we must sample at least three polarization components. Using four polarizing filters, we capture four images, a , b , c , and d , describing polarizations differing by $\pi/4$ (see Figure 10).

Let u be the unpolarized component, p the polarized component, and ϕ the polarization angle (we assume only one direction for the polarized component). Images a through d are

$$a = u/2 + p \cos^2 \phi$$

$$b = u/2 + p \cos^2(\pi/4 - \phi)$$

$$c = u/2 + p \cos^2(\pi/2 - \phi)$$

$$d = u/2 + p \cos^2(3\pi/4 - \phi).$$

A straightforward derivation then shows

$$p = \sqrt{(a - c)^2 + (b - d)^2}$$

$$u = a + c - p$$

$$\tan 2\phi = \frac{b - d}{a - c}.$$

With these relationships, we have all the information needed to compute arbitrary polarization images. In particular, we can generate an image captured with a filter with an arbitrary polarization angle α according to

$$I = u/2 + p \cos^2 \alpha. \quad (2)$$

The captured data contains a record of all four polarizations defined by the filters. This record is in the form of microimages, each containing four areas corresponding to the four polarization filters in the main lens aperture.

Figure 11 shows this polarization structure; parallel lines in each patch represent the direction of polarization transmitted through the individual filter. The top left of Figure 12 outlines the microimage captured by one microlens.

In the spirit of computational photography, we defer the decision about filter rotation until rendering time, freeing the photographer from making critical choices at capture time. Figure 13 demonstrates synthetic images generated according to Equation 2 at different angles of virtual rotation of the filter. We compute the images as combinations of the polarized and unpolarized components, according to Equation 2.

Figure 14 demonstrates polarization capture. We took a picture of two ducks on a lake, using a main lens filter with four polarization angles. That is, our capture of the plenoptic function included sampling four different polarization angles. Rendering a plenoptic function with polarization samples can produce interesting effects. For example, you can render final images corresponding to an arbitrary polarization angle. Or, you can render a final image with only the polarization component or the unpolarized component. Traditionally, photographers obtained similar effects by capturing consecutive images with different rotations of the polarization filter. However, as with HDR capture, this technique is problematic if any movement occurs between captures.

Multispectral Color Capture

Here, we use a main lens filter with four different colors (see Figure 15). The captured data contains a record of four colors defined by the filters, in the form of microimages, each containing four areas corresponding to the four color filters. Figure 16 shows the captured microimages.

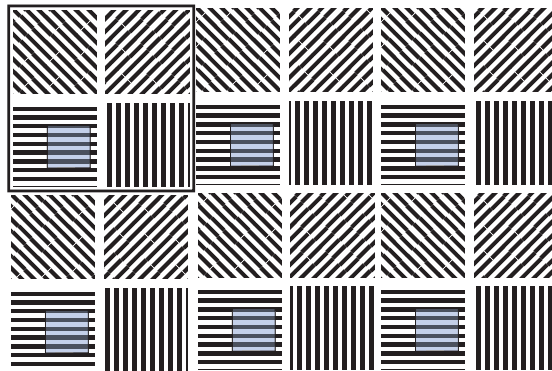


Figure 11. The polarization structure of our captured data. Each microlens image contains four areas of different polarization. Parallel lines represent the transmitted polarization's direction.

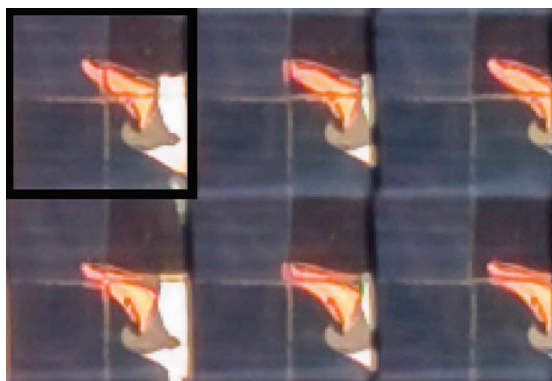


Figure 12. A crop with real polarization microimages. Looking at the swan's beak and part of its head, notice that the microimages are captured and inverted, as you would expect for the little camera formed with each microlens.

Figure 17 demonstrates multispectral color capture. We took a picture of some flowers, using a main lens filter with four colors. Because the camera already captures three colors, the effective image color space is 12D. The effect on colors in Figure 17 can't be achieved computationally from a single RGB image.

We see a great future for different methods of rich image capture. Based on sensors with

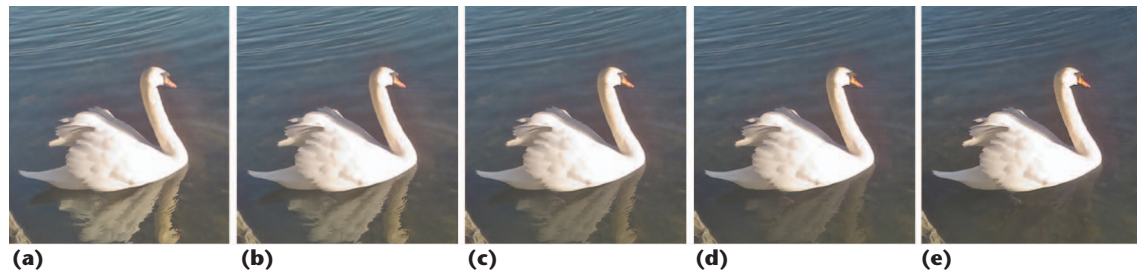
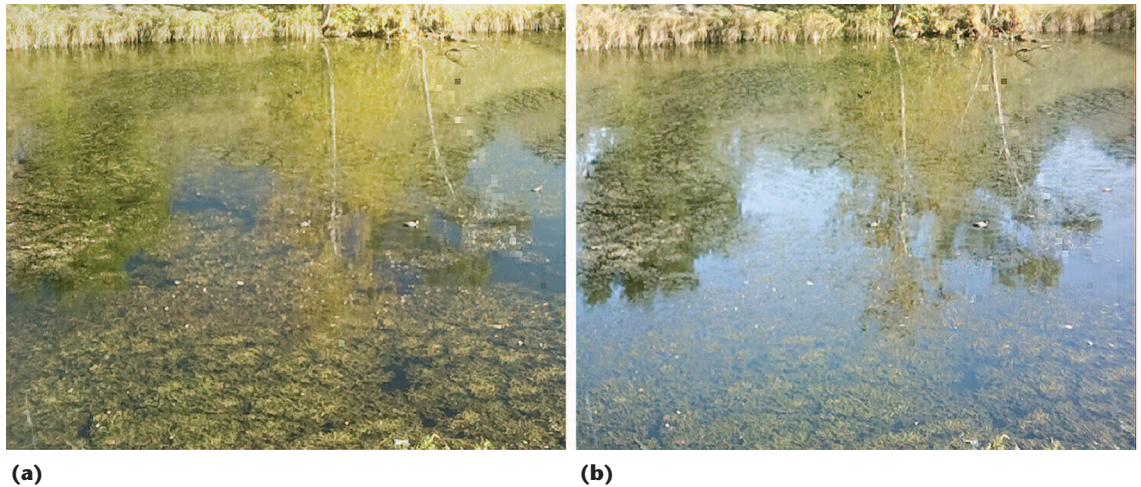


Figure 13. Our approach renders images with arbitrary polarization angles α . In these examples, α is (a) 0, (b) 30, (c) 45, (d) 60, and (e) 90 degrees.



(a)

(b)

Figure 14. Polarization capture. We captured the plenoptic image using a filter on the main lens with four different polarizations, letting us render the final images with (a) no polarization component or (b) an arbitrary polarization angle. In the second image, note the removal of the polarized components of light reflected from the water.

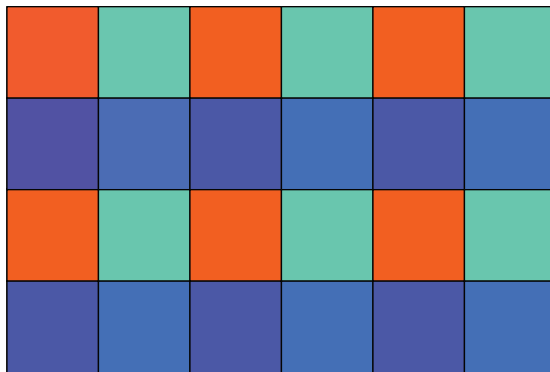


Figure 15. The color filters we applied to the four images we captured. Each microlens image contains four areas of different colors.

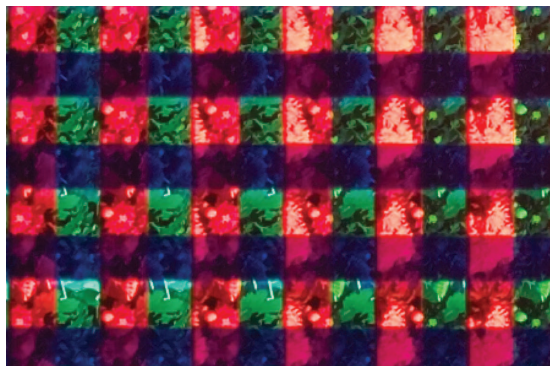



Figure 16. A crop from the raw captured image with four filters designed for multispectral capture. Each microimage contains four parts filtered differently at the main lens aperture.

improved characteristics and using plenoptic technology, this approach can increase image quality and diversity beyond what has been possible so far in real-time image capture.

From the perspective of advancing the technol-

ogy further, the method isn't limited to just applying different filters. Other modifiers are also possible, such as glass elements changing the optical path or additional microlenses or interleaved focal-length microlenses that change the focusing of individual microimages much like filters. This would increase the depth of field and final quality.

The basic rendering approach we've presented often produces artifacts between extracted samples when the samples don't exactly match at their boundaries (owing, for example, to parallax, incorrect scaling, or optical variations between microcameras). We're working on eliminating these artifacts. We also plan to extend our overall approach to additional modes. 

References

1. E. Adelson and J. Bergen, "The Plenoptic Function and the Elements of Early Vision," *Computational Models of Visual Processing*, M.S. Landy and J.A. Movshon, eds., MIT Press, 1991.
2. M. Levoy and P. Hanrahan, "Light Field Rendering," *Proc. Siggraph*, ACM Press, 1996, pp. 31–42.
3. P.E. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," *Proc. Siggraph*, ACM Press, 1997, pp. 369–378.
4. R. Horstmeyer et al., "Flexible Multimodal Camera Using a Light Field Architecture," *Proc. IEEE Int'l Conf. Computational Photography (ICCP 09)*, IEEE Press, 2009.
5. A. Lumsdaine and T. Georgiev, "The Focused Plenoptic Camera," *Proc. IEEE Int'l Conf. Computational Photography (ICCP 09)*, IEEE Press, 2009.
6. R. Ng et al., *Light Field Photography with a Hand-Held Plenoptic Camera*, tech. report CSTR 2005-02, Computer Science Dept., Stanford Univ, 2005.

7. J.H. Park et al., "An Ultra Wide Dynamic Range CMOS Image Sensor with a Linear Response," *Sensors, Cameras, and Systems for Scientific/Industrial Applications VII*, Proc. SPIE, vol. 6068, SPIE, 2006, pp. 94–101.
8. S.K. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," *Proc. 2000 IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR 00)*, vol. 1, IEEE CS Press, 2000, pp. 472–479.

Todor Georgiev is a senior research scientist at Adobe Systems, working closely with the Photoshop group. He concentrates on applying mathematical methods from physics to image processing, graphics, and vision. He's the author of Photoshop's Healing Brush tool, a method better known as Poisson image editing. Georgiev has a PhD in theoretical physics from XXXXXXXXXXXX. Contact him at tgeorgie@adobe.com.

Andrew Lumsdaine is a professor of computer science at Indiana University, where he also directs the Open Systems Laboratory. His research interests include computational science and engineering, parallel and distributed computing, mathematical software, numerical analysis, and radiance photography. Lumsdaine has a PhD in electrical engineering and computer science from the Massachusetts Institute of Technology. He's a member of IEEE, the IEEE Computer Society, the ACM, and the Society for Industrial and Applied Mathematics. Contact him at lums@cs.indiana.edu.



Figure 17. Multispectral color capture. The upper-left image is an unfiltered RGB image. We captured the plenoptic image using a filter on the main lens with four different colors, letting us render the final images from a 12D color space.

Georgi Chunev is a PhD student in computer science at Indiana University and a research assistant in the university's Open Systems Lab. His research interests include computer graphics and visualization, image processing, and plenoptic photography. Chunev has an MS in computer science from Indiana University. Contact him at gunchunev@cs.indiana.edu.