Abstract

Plenoptic cameras, constructed with internal microlens arrays, focus those microlenses at infinity in order to sample the 4D radiance directly at the microlenses. The consequent assumption is that each microlens image is completely defocused with respect to the image created by the main camera lens and the outside object. As a result, only a single pixel in the final image can be rendered from it, resulting in disappointingly low resolution. In this paper, we present a new approach to lightfield capture and image rendering that interprets the microlens array as an imaging system focused on the focal plane of the main camera lens. This approach captures a lightfield with significantly higher spatial resolution than the traditional approach, allowing us to render high resolution images that meet the expectations of modern photographers. Although the new approach samples the lightfield with reduced angular density, analysis and experimental results demonstrate that there is sufficient parallax to completely support lightfield manipulation algorithms such as refocusing and novel views.

1. Introduction

Recent developments in computational photography have made it possible for even amateur photographers to take and manipulate fully 3D photographs. The fundamental ideas behind 3D Photography can be traced back to Lippmann’s and Ives’ early 20th century works on integral photography [10, 15]. Though these ideas have been developed and refined throughout history [2, 3, 11, 12], it wasn’t until we were able to apply digital processing that these techniques have become fully practical as a tool for general or creative photography [1, 7, 9, 13, 14, 17, 18]. The plenoptic camera, as proposed by Adelson-Wang [1] and implemented by Ng [19], can be viewed as a “generic imaging device” that produces dense lightfield data. More specifically, the camera captures radiance with an array of microlenses, wherein each microlens samples a dense set of ray directions at a single spatial point. In this regard, the traditional plenoptic camera is a completely general device for capturing 4D radiance. However, the large proportion of sensor pixels that must be devoted to the directional dimensions in the 4D radiance results in low spatial resolution of the final rendered images. This limited resolution has been a drawback to radiance based photography.

In this paper we present detailed analysis of this new class of plenoptic cameras: The focused plenoptic camera or plenoptic camera 2.0. Like the traditional plenoptic camera, the focused plenoptic camera uses an array of microlenses internal to the camera to capture radiance. However, as the name implies, the focused plenoptic camera uses the microlens array as an array of microcameras, each of which captures a focused microimage. In this way, the new camera captures dense positional information, rather than capturing dense directional information, resulting in spatial resolution that is significantly higher than the number of microlenses used.

Different versions of this integral/plenoptic camera have been proposed by many authors throughout the years, starting with Lippmann himself [15], and including Ng [17], Fife [4], Lumsdaine [16]. In those modified plenoptic cameras, the radiance is sampled by an array of microcameras densely in the spatial dimensions, but sparsely in the angular dimensions. A discussion of the general issues involved in trading off spatial vs. angular sampling was reported by Georgiev et al. in [5].

2. Background

2.1. Basic Equations

The lightfield [14] or radiance [20] is a density function \( r(q, p) \) where \( q \) and \( p \) are vectors that respectively represent the position and direction of a ray intersecting a plane transverse to the optical axis. The pair \( (q, p) \) is also denoted by a single vector \( \mathbf{x} \). These coordinates are common in optics texts such as [6, 8, 21]. For rays in three dimensional space, \( q \) and \( p \) are two-dimensional vectors and \( \mathbf{x} \) is a four-dimensional vector. Without loss of generality, we adopt the convention of a two-dimensional \( q \cdot p \) plane in this paper, in which case \( p \) and \( q \) are one-dimensional and \( \mathbf{x} \) is two-dimensional.
Rays are transformed due to translation a distance \( t \) in the direction of the optical axis according to \((q', p') = (q + t p, p)\). This corresponds to a linear transformation \( x' = T_t x \), where

\[
T_t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.
\]

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

\[
L_f = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}.
\]

Optical systems created from cascades of translation and lens elements are described by the product of the matrices representing individual elements. For example, the operation on a ray of an optical system comprised of a lens followed by translation is described as \( x' = TLx \).

Since the radiance is a density function over the ray space, we can describe how radiance is transformed via transformations applied to elements of the ray space. To do this, we require a key property of the radiance—namely, that it is conserved (assuming non-absorbing transmission media). In that case,

\[
r'(x') = r(x).
\]

Since for any optical transformation \( A \) we describe transformation of rays according to \( x' = Ax \), or \( x = A^{-1} x' \), we have the following for the operation of optical transformations on radiance:

\[
r'(x) = r(A^{-1} x)
\]

Finally, an image is captured when light rays impinge on a sensor. The intensity at a given point of the image is the integral over all directions for the radiance at that point, i.e.,

\[
I(q) = \int_p r(q, p) dp.
\]

### 2.2. The Plenoptic Camera

As illustrated in Figure 1, a traditional plenoptic camera [17, 19] consists of a main lens and a microlens array placed at distance \( f \) in front of a sensor. Each microlens has aperture \( d \) and focal length \( f \). Without loss of generality, we take the microlenses to be equally spaced with spacing \( d \). The main lens of the camera is assumed to be focused at optical infinity (equivalently, the main lens).

The entire microlens therefore samples \( d \) times the number of microlenses.

![Figure 1: Traditional plenoptic camera. The main lens is focused at the microlens plane and the microlenses are focused at optical infinity (equivalently, the main lens).](image)

Figure 1: Traditional plenoptic camera. The main lens is focused at the microlens plane and the microlenses are focused at optical infinity (equivalently, the main lens).

The image captured by the sensor behind the microlenses is therefore

\[
I_f(q) = \int_p r_f(q, p) dp = \int_p r(q - fp, \frac{1}{f} q) dp
\]

To complete this integral, consider Figure 2. The range of \( p \) over which \( r \) is integrated spans \( d/f \). If we assume \( r \) is constant across the microlens for a given \( p \), then

\[
I_f(q) = \frac{d}{f} r(0, \frac{1}{f} q)
\]

From this equation we can see that each spatial point in a given microlens image corresponds to a different direction for the point at \( q = 0 \) at the microlens plane.

![Figure 2: The “1f” system represented by a single microlens. The image from the main lens is assumed to be focused on the plane of the microlenses. Different directions in the radiance function at \( q = 0 \) are mapped to different positions on the sensor.](image)

Figure 2: The “1f” system represented by a single microlens. The image from the main lens is assumed to be focused on the plane of the microlenses. Different directions in the radiance function at \( q = 0 \) are mapped to different positions on the sensor.
If \( f \) \((7, 6, 5, 4, 3, 2, 1, q, p)\) corrections radiance sampled a single direction in the directional coordinate and samples a span of \( d \) in the positional coordinate.

Figure 3: Sampling of the radiance \( r(p, q) \) by the microlens array represented in the two-dimensional \((q, p)\) plane. Each pixel samples a single direction in the directional coordinate and samples a span of \( d \) in the positional coordinate.

\[
I_f(q) = \int r(q, p) dp
\]

Figure 4: A final image is rendered from the captured lightfield in two steps. First, the sensor image \( I_f(q) \) is converted into the radiance \( r(q, p) \). Next, we integrate (average) over \( p \) at each value of \( q \) corresponding to microlens samples.

Sensor image \( I_f(q) \). Each microlens image captured at the sensor corresponds to a single position \( q \) and a range of directions \( p \). Thus, \( r(q, p) \) is obtained from \( I_f(q) \) simply by “stacking” the microlens images in the \((q, p)\) plane. This process is shown in Figure 4.

2.3. Rendering with the Plenoptic Camera

Images are rendered from the sampled radiance \( r(q, p) \) according to Equation (5). However, since each microlens image spans the same range of \( q \), integrating over \( p \) at each \( q \) simply means integrating all of the pixels under each microlens. Notice also that the resolution in this case is limited to the number of microlens images. In Figure 4, the rendered image \( I(q) \) is obtained by integrating (averaging) along each value of \( q \) corresponding to microlens samples taken from \( I_f(q) \).

3. High Spatial Resolution Radiance Capture

The high spatial resolution radiance approach is based on an alternative interpretation of the microlens array functionality. That is, rather than using the microlens array to sample the radiance by position at the plane of the microlens array, we will use it to sample the radiance by direction at a plane some distance \( a \) in front of the array. A different rendering algorithm that is appropriate for this interpretation of the image captured on the sensor will be required to render the final image. As we will see, a change to the camera is required to support this perfectly. Because this change is small, we will also see that reasonable results can even be obtained by applying the high resolution rendering approach to radiance images captured with a traditional plenoptic camera.

3.1. Focused Plenoptic Camera

Figures 5a and 5b show the two versions of the focused plenoptic camera. Note that rather than placing the sensor at distance \( f \) behind the microlenses, we place it at a distance \( b > f \) or \( b < f \). The microlenses therefore image the plane at distance \( a \) in front of the microlenses, assuming the lens equation holds, \( 1/a + 1/b = 1/f \). In the first case, we assume that the main lens of the camera creates a real image at this plane. In the second case, \( a < 0 \) and the image created by the main lens is virtual, on the other side of the microlens array. This image is picked up by the microlenses and reimaged onto the sensor.

3.2. The “ba” system

To show how the focused plenoptic camera captures radiance, we derive an expression for the image captured at the sensor in terms of the radiance at the focal plane of the main lens (that is, the radiance at a plane distance \( a \) in front of the microlenses). Consider one microlens and let \( r_a(q, p) \) be the radiance at the focal plane of the main lens and \( r_b(q, p) \) be the radiance at the sensor behind that microlens (see Figure 6).

Then, \( r_b \) is related to \( r_a \) according to \( r_b(x) = r_a(A^{-1}_{ba}x) \), where

\[
A_{ba} = T_b LT_a = \begin{bmatrix} -\frac{b}{f} & 0 \\ -\frac{a}{f} & -\frac{2}{b} \end{bmatrix}.
\]

I.e., \( r_b(x) = r_a(-\frac{a}{b}q, -\frac{b}{f}p - \frac{1}{f}q) \).

To complete this integral, consider Figure 6. The range of \( p \) over which \( r \) is integrated spans \( d/b \). However, now there is only a single value of \( q \) over which the integration takes place in \( r_a \).

\[
I_b(x) = \frac{d}{b} \int_{p} r_a(-\frac{a}{b}q, \frac{1}{f}p) dp
\]
(a) Real image in front of the microlens array.  
(b) Virtual image behind the microlens array.

Figure 5: Two configurations of the focused plenoptic camera. The microlenses image the focal plane of the main lens, which may be in front of the microlenses or behind the microlenses. In the former case, the microlenses are focused on a real image, in the latter case they are focused on a virtual image.

\[ r_a(q, p) = \frac{-q}{b} \]

Figure 6: The “ba” microlens system. The image from the main lens is assumed to be focused on the plane distance \( a \) in front of the microlenses, i.e., the imaging plane of the microlenses. Each microlens focuses a portion of the image plane onto the sensor.

\[ I_b(q) \rightarrow r_a(q, p) \]

Figure 8: Sampling of the radiance \( r_a(q, p) \) by the microlens array represented in the two-dimensional \((q, p)\) plane.

\[ \int_{p} r_a(q, p) dp \]

Figure 9: A final image is rendered from the captured lightfield in two steps. First, the sensor image \( I_b(q) \) is converted into the radiance \( r_a(q, p) \). Next, we integrate (average) at fixed \( q \) over \( p \). The resolution at which we can integrate is determined by the resolution of the microlens images, not by the number of microlenses.

Figure 7 represents the way this radiance is sampled by a single microlens (infinitely small pixels and microlens aperture \( d \) are assumed). Each pixel samples a single position in \( q \) and samples a span of \( d/a \) in \( p \). The entire microlens samples a span of \( da/b \) in \( q \).

Spatial resolution of focused plenoptic camera. The spatial resolution of the captured radiance \( r_a(q, p) \) is \( b/a \) times the resolution of the sensor image \( I_b(q) \). Consider the radiance at a plane a distance \( a \) in front of the microlenses.

Angular resolution of focused plenoptic camera. At any position \( q \), the radiance captured by the focused plenoptic camera has angular resolution \( a/b \), meaning it has only \( a/b \)
unique angular samples. In fact, there can only be $a/b$ non-overlapping directional ranges in $r_a(q,p)$. Consider Figure 7. Since the angular width of each pixel sample is $d/a$ and the entire microlens samples a range of $d/b$, more than $a/b$ samples would overlap.

**Conservation of spatio-angular information.** For any $a$ and $b$ satisfying the lens equation, the total amount of spatio-angular information remains constant. This is simply due to the fact that the spatial resolution increases exactly in proportion to the decrease in angular resolution.

Fundamentally, these properties are consequences of the reduction in size by $a/b$ during image formation in each of the microcameras (Figure 6). As a result, the spatial resolution of $r_a(q,p)$ is a function of the resolution of the microlens images and their overlap, not of the number of microlenses. This is a critical observation that distinguishes focused from traditional plenoptic cameras.

With the focused plenoptic camera configuration, it is a straightforward process to reconstruct the radiance $r_a(q,p)$ from the sensor image $I_b(q)$. Each microlens image captured at the sensor corresponds to a skewed line in the $(q,p)$ plane. Thus, $r(q,p)$ is obtained from $I_f(q)$ simply by “stacking” these skewed microlens images in the $(q,p)$ plane. This process is shown in Figure 9.

### 3.3. High Spatial Resolution Rendering

With the focused plenoptic camera, image rendering is defined by applying Equation (5) to $r_a(q,p)$. Thus, the resolution of the rendered image is the spatial resolution of $r_a(q,p)$, which, as described above, is the total available resolution of the sensor divided by $a/b$. The radiance at the image plane is related to the radiances at the microlenses $r(q,p)$ by a translation $a$, i.e.,

$$r_a(q,p) = r(q + ap,p),$$

which is a shearing transformation in ray space.

Note that the radiance $r(q,p)$ is directly sampled when it is captured by the microlens array. Under the transformation (13), these samples are sheared in the $q$ direction. As seen in Figure 8, the samples will overlap, i.e., adjacent microlens images will have samples at overlapping $q$ coordinates but at differing $p$ coordinates.

Since the different directional samples for a given positional sample are spread across microlenses, our rendering process must integrate across microlens images, rather than within microlens images. That is, assuming we are “imaging the image” that is in focus, we integrate the points in the microlenses that correspond to the same position in the image by overlapping them at a fixed pitch. Note that we do not need to determine a registration among microlens images. The focused image can be derived directly from the geometry of the camera which, together with the focusing depth, defines the pitch and the corresponding amount to overlap.

An algorithm for rendering a final image from the radiance is the following:

**Focused Plenoptic Rendering Algorithm.** Given: Discrete $N_x \times N_y \times n_x \times n_y$ radiances $r[i,j,k,l]$ Output: Rendered $M \times N_x \times M \times N_y$ image $I[s,t]$

For $(s,t)$ in $(M \times N_x, M \times N_y)$

$$I[s,t] = \sum_{i,j,k,l} r[i,j,k,l],$$

where $i = (s/Nx) \times M$, $j = (t/Ny) \times M$,

$k = (s \% Nx) \times M$, and $l = (t \% Ny) \times M$.

In practice, to mitigate the effects of poor image quality at the edges of microlens images [1, 5], we may use weighted averages that favor the center of the microlens image.

### 4. Experimental Results

In this section we demonstrate some of the capabilities of the focused plenoptic camera and the radiance that it captures. We compare the images produced by the focused plenoptic rendering approach to those produced by the traditional lightfield rendering approach. Since the image captured by this camera is a radiance, we also demonstrate that the usual rendering techniques (such as refocusing) can be accomplished at high resolution. Finally, we show that even the images captured by a traditional plenoptic camera are often sufficiently well-focused so that our rendering algorithm can be applied with reasonably good results. In other words, we can often achieve a factor of 3 to 5 times increase in resolution of images rendered from traditional lightfields.

#### 4.1. Experimental Setup

We conducted experiments using images captured with a focused plenoptic camera based on a medium format camera. The camera includes a Megavision E4 digital back with 16 megapixel sensor KAF-16803 from Kodak. The sensor cover glass was removed to properly accommodate our microlens array. The microlens array was manufactured by Leister as part of a fused silica wafer. Each microlens has diameter $d = 250\mu m$ and focal length $f = 750\mu m$. The array has $144^2$ microlenses. We have experimented with two different types of spacing from the microlenses to the silicon, $b < f$ and one longer $b > f$ with variations within 100$\mu m$ of $f$. The main lens of the camera was 140mm and the photographs reported here were taken with F number 2.8. Focusing of the main lens behind the microlens array was accomplished by lengthening the focal length of the main lens by inserting a piece of flat glass (optical window) behind the main lens. The software for rendering the images was written in Python, using Numpy and the Python Imaging Library.

#### 4.2. Radiance Characteristics

In Figure 10 we show some characteristics of the radiance captured by the focused plenoptic camera. In the figure, we first show the radiance as it is captured by the sensor.
Considered as a normal radiance, this would be “position major” order—each separate image is a different position and contains the directional information for that point.

Figure 10b shows the same radiance data, but in “direction major” order. That is, each individual image patch (and they are quite small in this radiance image) corresponds to a different direction and contains all the positional information for that direction.

Figure 10c shows the focused radiance data in “direction major” order. This is the same radiance data as for the other two images, but considered at the focal plane of the main lens. Again, each individual image patch corresponds to a different direction and contains all the positional information for that direction. Note, however, there are significantly fewer directional samples—but at the same time each directional sample contains significantly (and proportionally) more positional samples.

4.3. High Resolution Rendering

Figure 11 shows a comparison of an image rendered with the typical radiance rendering algorithm (Figure 11b) with an image rendered with the plenoptic 2.0 algorithm (Figure 11c). Note that since the sensor in the focused plenoptic camera still captures the radiance in front of the microlens array, application of the traditional rendering algorithm to this radiance is valid. In this case, the focused plenoptic camera has a $8 \times$ increase in resolution in both the $x$ and $y$ dimensions. Specifically, the normally rendered image has a resolution of $130 \times 122$ (the number of microlenses in the microlens array plus some trimming), whereas the high resolution image has resolution $1040 \times 976$.

4.4. Rendering Effects

Figure 12 shows refocused images rendered from our captured radiance. Even though the angular information is sparse, the redundancy due to our assumption of the Lambertian nature of the scene allows refocusing to still be accomplished.

4.5. High Resolution Rendering Applied to Traditional Radiance

Figure 13 shows the results of our high resolution rendering algorithm applied to a radiance image captured with a traditional plenoptic camera (image from [18], used with permission). In this case (which is representative of our general observations with similar images), even the images captured by a traditional plenoptic camera are sufficiently well-focused so that our rendering algorithm can be applied with reasonably good results. With this example we were able to achieve an increase in resolution of up to a factor of about 25 ($5 \times$ in each dimension). Figure 13 shows a $5 \times$ by $5 \times$ improvement.

4.6. Limitations

As shown in Section 3.2 there is a direct relationship between spatial resolution and depth of focus in the focused plenoptic camera. That is, a given resolution always correspond to a particular focal plane in the scene. When we render an image with one particular resolution, it may contain parts of the scene that would be fully resolved and correctly rendered at a different resolution. Portions of the scene (at smaller depth) that would be fully resolved with higher resolution would overlap and look blurred when rendered at a lower resolution. At the same time, portions of the scene (at higher depth) that would be fully resolved at lower resolutions will not overlap. They will be spread apart. If the microlenses responsible for this portion of the scene were focused at infinity, we would simply see a pixellation effect. However, since the focused plenoptic camera has focused microimages, we see disjoint repetitions of portions of each microimage—a “macropixellation” effect, if you will.

When macropixellation occurs, the macropixels are the same size as a magnified pixel would be in a normally rendered lightfield image, magnified (see Figures 11b and 11c). Thus, in this worst-case scenario, our approach is essentially equivalent to traditional lightfield rendering.

For better results even in this case, there are several approaches that one could take. The details and results of those approaches will be considered in future papers. Briefly, however, we note that issues related to sparsely sampled lightfields were developed in [5], where intermediate directional samples were synthesized via view morphing. Given the direct relationship between the focused plenoptic camera as an array of microcameras, and the systems in [5], view morphing would immediately be one approach for dealing with macropixellation in the focused plenoptic camera. Since resolution is related to depth of focus, macropixellation can also be eliminated by combining full-resolution rendering with depth estimation techniques whereby different depths are rendered at their appropriate resolutions.

The use of fewer directional samples in plenoptic 2.0 rendering means that there is less averaging being done to create final output pixels. As a result, noise or other imperfections in the captured radiance may be more readily evident. This can be seen, for example, in Figure 13, which is based on a noisy captured radiance. Exposing flaws is something that typically happens with higher resolution in any kind of setting.

5. Conclusion

In this paper we have presented the design and analysis of lightfield camera structure that provides new insight on the interactions between the main lens system and the microlens array system. By focusing the microlenses on the image produced by the main lens, our camera is able to fully
capture the positional information of the lightfield. Based on change of focusing, it can therefore also provide variable spatial-angular resolution tradeoff. In conjunction with the camera, we have developed new algorithms for rendering images from a captured radiance which produce images at a dramatically higher resolution than traditional techniques.

The focused plenoptic camera is making a tradeoff between spatial and angular resolution. Macropixellation was shown to occur because of the sparse angular sampling. Development of techniques for reducing macropixellation is a topic of ongoing work.

With the capability to capture and use more of the spatial resolution information available in a lightfield, we can now render images at a resolution expected in modern photography (e.g., 10 megapixel and beyond) without waiting for significant new advances in sensor or camera technologies. Lightfield photography is suddenly much more practical.

Acknowledgments

This work was supported in part by a grant from the Lilly Endowment and by a gift from Adobe Systems, Inc. The authors wish to thank Ren Ng for sharing his original lightfield data, from which we generated the image in Figure 13.

References

Figure 12: Refocused images rendered from our captured radiance.

Figure 13: High resolution rendering applied to traditional radiance.


