Characterization of Plenoptic Imaging Systems and Efficient Volumetric Estimation from Plenoptic Data

Paul Anglin, Stanley J. Reeves, Senior Member, IEEE, and Brian S. Thurow

Abstract—Plenoptic imaging is a rapidly growing field driven by the ever-declining cost of imaging systems and the promise of image focus, perspective, and depth of field manipulation during post-processing. While plenoptic systems are often limited to 2D image reconstruction and manipulation, plenoptic data and reconstruction algorithms can be extended to volumetric fields. An estimate of the imaged volume can be created by generating a stack of 2D images, but such an estimate can easily be dominated by image blur from neighboring focal planes. Tomographic algorithms have been shown to be effective in creating volumetric estimates from plenoptic data but are often prohibitively slow. The research presented here shows that the reconstruction is solvable through deconvolution. Unfortunately, the observation model is not shift-invariant. However, with appropriate transformations, the problem can be made shift-invariant so that deconvolution is a viable solution. Utilizing the computationally efficient fast Fourier transform (FFT) allows the reconstruction to be completed quickly while producing estimates exhibiting significantly reduced blur compared to a simple focal stack. This work describes a deconvolution algorithm designed to reconstruct a 3D volume from a 2D plenoptic image. The imaging system and refocusing algorithm are characterized with respect to shift-variance in order to identify potential sources of artifacts and propose potential mitigating steps. To demonstrate the efficacy of the algorithm, experimental data is presented with comparisons of the focal stack to the reconstructed volume.

Index Terms—deconvolution, plenoptic, light-field.

I. INTRODUCTION

Traditional imaging modalities such as a standard camera can be used to acquire spatial samples of a volume, these samples are limited to projections on a 2D plane. To obtain the angular information necessary to produce a 3D reconstruction, movable or multi-camera systems can be used, but these are often limited by size and cost restraints [3, 14, 26–28]. However, light-field imaging offers a potential solution. A light-field—or plenoptic—camera samples each of the four dimensions of the light field with each exposure. By sampling both the spatial and angular domains, the user is given the tools to manipulate focus, perspective, and depth of field (DOF) during post-processing. With traditional imaging, these attributes are largely fixed once the image is acquired. The concepts related to sampling the 4D light field are not new and have been presented by many others [27,29–31]. However, the works presented by [30, 32–34] and the introduction of the Lytro Camera have brought both these concepts and the technology to mainstream attention.

From plenoptic data, a crude estimate of the imaged volume can be generated simply by creating computationally refocused images at slightly different depths throughout the volume. This focal stack represents a sampled version of the imaged volume. However, such a reconstruction represents a poor estimate of the volume as each refocused plane will contain energy from out-of-focus structures in adjacent focal planes. This out-of-focus energy is commonly recognized as image blur in conventional 2D images. In some scenarios, the image blur can dominate the response and obscure smaller structures in the image. In order to provide an accurate estimate of the volume, a method of reducing or eliminating the out-of-plane-energy is desired.

Tomographic methods utilizing plenoptic data have been shown to produce largely blur-free estimates but are computationally expensive and may take hours to produce a single reconstruction [35–37]. This work demonstrates that the volume reconstruction problem is solvable via frequency-domain deconvolution methods. This is advantageous as such methods can utilize FFT-based processing techniques for computation of the frequency spectra. The result is a computationally efficient algorithm that can produce volume reconstructions in seconds or minutes when given the focal stack and PSF.

To demonstrate the efficacy of deconvolution when applied to plenoptic data, a shift-invariant model of the imaging system is first developed. Potential sources of shift-variance are explored in detail, and potential impacts to the reconstructions and possible methods of minimizing this impact are presented. An experimental volume consisting of two separate flames is imaged utilizing a plenoptic camera, and existing spatial-domain refocusing techniques are then utilized in creating the focal stacks from the simulated data. Deconvolution is then used to reconstruct the imaged volume, artifacts in the reconstructions are identified, and potential sources and methods of minimization are addressed.
II. BACKGROUND

The basic concepts related to light-field imaging, plenoptic cameras, and computational refocusing are presented here. However, the reader is encouraged to refer to works devoted to the subjects [27,29–31,33,34,38–42] for a fuller treatment of the topics.

A. The Plenoptic Function and Light-field Imaging

The geometric distribution of light in the world around us can be described by a 5D function often referred to as the plenoptic function given by \( P(x, y, z, \theta, \phi) \). This function has three spatial coordinates \((x, y, z)\) and two angular coordinates \((\theta, \phi)\) and is often used in computer vision and computer graphics. However, the measurement of light in a space is generally restricted to regions outside the convex hull, in which case the plenoptic function becomes redundant in one dimension. That is, the intensity of each ray does not change as it propagates in the \( z \) direction in the absence of occluders. The resulting 4D function is referred to as the light field and is given by \( P(x, y, \theta, \phi) \) [27, 29].

Traditional imaging modalities capture a 2D projection of the 4D light field by integrating over \((\theta, \phi)\). When the image is acquired by the camera sensor, only the \((x, y)\) spatial information is recorded and the angular information is lost. However, a simple modification to a traditional camera allows the system to sample all four dimensions of the light field. Placing a micro-lens array between the camera main lens and the camera sensor subdivides the sensor such that the pixels underneath each micro-lens correspond to angular samples and each micro-lens now corresponds to a spatial sample. This also establishes the fundamental tradeoff in plenoptic imaging. Subdividing the camera sensor in this way sacrifices spatial resolution for angular information. The system spatial resolution is now established by the micro-lens array while the angular resolution is set by the pixels underneath each micro-lens [43, 44]. Such a tradeoff may have previously been unacceptable, but the exponential growth of sensor resolution has resulted in cases where the system is resolution-limited by the optics rather than the sensor. For these cases, and those where exceedingly high spatial resolution is not required, plenoptic imaging offers an attractive method of utilizing the otherwise unused sensor resolution [44].

B. Computational Refocusing

Light-field cameras provide many post-processing opportunities that were not previously available to the photographer or researcher. The ability to computationally shift perspective and change the DOF are useful, but of particular interest to this work is the ability to create an image at any focal depth utilizing the data captured during a single acquisition. Computational refocusing is possible because a plenoptic camera captures not only the intensity at each spatial location but also the angle of arrival for the incident rays. With this information, the light field can be traced through space to simulate the equivalent sensor information at any arbitrary depth. The equivalent image at a chosen focal depth can be calculated by simply integrating the rays passing through each spatial location. In this way, images at any focal length are generated in the same way that a traditional camera captures a single image, that is, by integrating over the angular information and collecting the remaining spatial information.

C. Imaging Equations

A more rigorous description of the process can be achieved by dividing the camera into two planes [29]. The first plane represents the camera main lens or aperture and is assigned coordinates in \((u, v)\) while the second plane represents the micro-lens array and is assigned coordinates in \((x, y)\). This is convenient as any ray within the camera can now be described by its intersection with these two planes. The radiance at a given point on the \((x, y)\) plane is then given by

\[
E_F(x, y) = \frac{1}{F^2} \int \int L_F(x, y, u, v) \, du \, dv
\]

where \(F\) is the distance between the \((u, v)\) plane and the \((x, y)\) plane. \(E_F\) is the irradiance at \((x, y)\), and \(L_F\) is the light field given in terms of the two planes separated by \(F\). Determining the light field at the desired focal plane \(F'\) is accomplished by calculating the location where each ray intersects the new focal plane and is given by

\[
L_{F'}(x', y', u, v) = L_F \left( u + \frac{x' - u}{\alpha}, v + \frac{y' - v}{\alpha}, u, v \right)
\]

\[
= L_F \left( u \left( 1 - \frac{1}{\alpha} \right) + \frac{x'}{\alpha}, v \left( 1 - \frac{1}{\alpha} \right) + \frac{y'}{\alpha}, u, v \right),
\]

where the ratio of the new focal plane to the old focal plane is defined as \(\alpha = F'/F\). These equations provide the necessary relationships for integral-based refocusing.

D. Ray-space Diagrams and the Radiance Array

The use of a two-plane parameterization also introduces the concept of a Cartesian ray-space diagram and the radiance array. These concepts have proven to be invaluable as a visualization tool in determining the source of spatial variability in the refocusing algorithm. Figure 1 shows the relationship between the imaged point and the ray-space diagram in two dimensions. Light rays radiating from a discrete point in space trace a line through the \((x, u)\) plane, where the angle of the line is dependent on the distance of the point from the focal plane. A continuous distribution of light is sampled spatially by the microlens array and angularly by the camera sensor. That is, the pixels underneath each lenslet correspond to angular samples across the camera aperture. Each line in image space corresponds to a single sample of the plenoptic function, which is represented as a point in the ray-space diagram. When digitally refocusing to the plane containing this point, the pixel associated with the point is synthesized by summing the angular samples along this line in the \((x, u)\) plane. Moving the point in the \(x\) direction on the current focal plane simply shifts the line along the \(x\)-axis in the ray-space diagram, while moving off the current focal plane results in a shearing or rotation of the line/samples about the \(x\)-axis sample.
The physical application of the ray-space diagram can be seen when considering the sensor data acquired by a plenoptic camera. The camera sensor collects a 2D array of angular samples, but the spatial information is also encoded in the locations of these samples on the sensor. In order to simplify processing, it becomes beneficial to reorder the sensor data into a 4D radiance array that mimics the structure of the ray-space diagram. To clarify and demonstrate concepts, 2D simulated data is presented throughout this work. This data is generated by choosing a point or points in object space and tracing the path of the light rays originating at these points back to the camera lens, to the microlens array, and finally to the sensor. The process is described in greater detail in [45], and the simulation parameters used are presented in Table I. Utilizing this process, Fig. 2 shows the relationship between the sensor data for a simulated point 9 mm in front of the focal plane and the corresponding radiance array. Note the shearing/rotation of the ray-space samples that occurs for points not on the optical focal plane. With this representation, refocusing to an arbitrary focal plane becomes a matter of choosing the correct angular samples from the radiance array, and summing the values. Recognizing that the data acquired for an arbitrary point in space is a combination of a shifting and shearing within the radiance array, the samples for that location can be similarly chosen. The importance of understanding this relationship will become clear in subsequent sections.

III. Volume Reconstruction Via FFT-Based Deconvolution

When imaged by a plenoptic camera, an estimate of the imaged volume can be created by generating a focal stack utilizing computational refocusing. When objects are present within the volume, each focal plane within the stack will contain energy from out-of-plane features. FFT-based deconvolution offers a fast and efficient means of removing this volumetric blur, thereby improving the estimate of the volume. While volumetric deconvolution of light-field data has been applied to microscopy [46], the characteristics of these systems are governed by the wave nature of light where each point emits a spherical wave of light. At the macro level, the system is modeled based on ray approximations to light propagation which yields unique characteristics. Furthermore, microscopes are shift-invariant in the $x$ and $y$ directions [46], which is a significant advantage when applying 3D deconvolution. In contrast, existing refocusing algorithms for light-field cameras are inherently shift-variant, which must be considered when applying 3D deconvolution.

A. Imaging Model

In order to apply deconvolution to volumetric reconstruction, it is beneficial to model the system as a convolution in the forward direction. Denoting an arbitrary point in three-space as $f(x, y, z)$, approximating the shift-variant point spread function (PSF) as a shift-invariant function $h(x, y, z)$, noise as $\eta(x, y, z)$, and the resulting image as $g(x, y, z)$, the imaging operation can be approximated as

$$
\begin{align*}
\theta(x, y, z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y', z') h(x - x', y - y', z - z') \, dx' \, dy' \, dz' + \eta(x, y, z) \\
&= f(x, y, z) * h(x, y, z) + \eta(x, y, z).
\end{align*}
$$

The second line follows by denoting the convolution operator as $\ast$. Equation (4) describes the relationship between the imaged point and the focal stack generated by refocusing the plenoptic data.

The PSF is the system impulse response. Ideally, the PSF chosen will completely characterize the imaging system. This would include the blurring that occurs as objects are moved further from the focal plane as well as impacts due to optical aberrations or other system anomalies. Unfortunately, neither the plenoptic imaging system nor the refocusing algorithm are shift-invariant. However, under certain circumstances and with appropriate strategies, the system can be closely approximated as shift-invariant. Sources of shift-variance must be identified and addressed where possible to provide the best possible agreement between the shift-invariant model and the shift-variant system being analyzed.
1) Shift-Invariant Image Synthesis: Identifying sources of shift-variance begins with analyzing the refocusing equation
\[
E_{\alpha,F}(x', y') = \frac{1}{\alpha^2 F^2} \int L_F(u ) \frac{1}{\alpha} + \frac{x'}{\alpha}, \\
v \left( \frac{1}{\alpha} + y'/\alpha, u, v \right) dudv. \tag{5}
\]
A compact notation is adopted by letting \( x = [ x \ y ]' \) and \( u = [ u \ v ]' \), which gives
\[
E_{\alpha,F}(x') = \frac{1}{\alpha^2 F^2} \int L_F(u ) \frac{1}{\alpha} + \frac{x'}{\alpha}, u dudv. \tag{6}
\]

The analysis begins by selecting a point \( x_0' \) in image space on some arbitrary focal plane a distance \( z_0' \) from the microlens array. This can be modeled by an impulse \( A(z_0') \delta(x' - x_0') \) with an arbitrary amplitude given by \( A(z_0') \). The object-space impulse is observed through the camera aperture which limits the range of angular samples and is modeled by applying a windowing function \( W(u) \) to the impulse that models the effects of the aperture. The light field for a point source at \( x_0' \) at depth \( z_0' \) from the microlens is defined by:
\[
L_{F - z_0'}(x', u) = A(z_0') \delta(x' - x_0') W(u) \tag{7}
\]

The light field \( L_F \) at the microlens plane can then be expressed in terms of the light field \( L_{F - z_0'} \) at the plane \( F - z_0' \) by solving for \( x' \) in terms of \( x \), where \( \alpha = (F - z_0')/F \):
\[
x = \frac{1}{\alpha} [ u(\alpha - 1) + x'] \tag{8}
\]
\[
x' = \alpha x - u(\alpha - 1) \tag{9}
\]

Let \( \alpha_0 = (F - z_0')/F \). The light field at the microlens array is then
\[
L_F(x, u) = L_{F - z_0'}(\alpha_0 x - u(\alpha_0 - 1), u) \tag{10}
\]
\[
= A(z_0') \delta(\alpha_0 x - u(\alpha_0 - 1) - x_0') W(u) \tag{11}
\]
\[
= A(z_0') \delta(x - u(1 - \frac{1}{\alpha_0}) - \frac{x_0'}{\alpha_0}) W(u) \tag{12}
\]
\[
= \delta(x - \frac{x_0'}{\alpha_0} - u(1 - \frac{1}{\alpha_0}) W(u) \tag{13}
\]

where the last line is obtained by setting \( A(z_0') = \alpha_0^2 \).

Note that the light field at \( F - z_0' \) is shifted by \( x_0' \), while the light field at the micromems array is shifted by a function dependent upon the location of the impulse along \( z \). Therefore, the imaging equation based on this light field parameterization is shift-variant. However, this can be remedied by scaling the argument when evaluating functions of the light field at the micromems.

Define a mapping from an image space \( x'' \)-plane, where \( x'' \) is used to denote an \( x \) position in the scaled image space, to the light-field plane with the scaling removed (\( \hat{x} \)):
\[
\hat{x} = u(\alpha - 1) + x'' \tag{14}
\]
\[
x'' = \hat{x} - u(\alpha - 1) \tag{15}
\]

where \( \alpha = (F - z_0')/F \). The modified light field \( L''_{F - z''}(x'', u) \) at an arbitrary focal plane \( z'' \) can now be calculated in terms of the light-field pattern created by the impulse:
\[
L''_{F - z''}(x'', u) = L_F(\hat{x}, u) \big|_{\hat{x}=u(\alpha-1)+x''} \tag{16}
\]
\[
= \delta(u(\alpha - 1) + x'' - u(1 - \frac{1}{\alpha_0}) - \frac{x_0'}{\alpha_0}) W(u) \tag{17}
\]
\[
= \delta(u[\alpha - 1 - 1 + \frac{1}{\alpha_0}] + x'' - \frac{x_0'}{\alpha_0}) W(u) \tag{18}
\]

Observe that
\[
\alpha - 1 - 1 + \frac{1}{\alpha_0} = \frac{F - z''}{F} - \frac{F}{F - z_0'} + \frac{F}{F - z_0'} \tag{19}
\]
\[
= -\frac{1}{F} \left( \frac{z'' - \frac{z_0'}{\alpha_0}}{F} \right) \tag{20}
\]
\[
= -\frac{1}{F} \left( \frac{z'' - \frac{z_0'}{\alpha_0}}\delta \right) \tag{21}
\]

Further defining the following transformed coordinates:
\[
x''_0 = \frac{x_0'}{\alpha_0} \tag{22}
\]
\[
z''_0 = \frac{z_0'}{\alpha_0} \tag{23}
\]

we substitute into the previous expression to obtain
\[
L''_{F - z''}(x'', u) = \delta(u(\alpha - 1) + x'' - \frac{x_0'}{\alpha_0}) W(u) \tag{24}
\]

Note that evaluating this modified function at \( z'' = z''_0 \) yields
\[
L''_{F - z''_0}(x'', u) = \delta(x'' - \frac{x_0'}{\alpha_0}) W(u) \tag{25}
\]

which is the light-field function arising from an impulse in scaled image space \( \delta(x'' - \frac{x_0'}{\alpha_0}) \). The integral over the aperture is then calculated by integrating over \( u \) to find modified refocused images at each \( z''_0 \)-plane:
\[
E''_{\alpha,F}(x'', y'') = \int L''_{F - z''_0}(x'', u) dudv \tag{26}
\]
\[
= \int \delta(u(\alpha - 1) + x'' - \frac{x_0'}{\alpha_0}) W(u) dudv \tag{27}
\]
\[
= \frac{F^2}{(z'' - z''_0)^2} W \left( \frac{x'' - \frac{x_0'}{\alpha_0}}{z'' - z''_0} \right) \tag{28}
\]

Thus, a shift of the input impulse by \( (x'', z'') \) results in a shift of the refocused image by the same amount, which demonstrates shift-invariance. The resulting algorithm for calculating a shift-invariant focal stack is
\[
E_{\alpha,F}(x', y') = \int L_F(\alpha - 1 + x', v) \frac{1}{\alpha} + y', v) dudv. \tag{29}
\]
samples integrated to obtain the point, likewise denoted by an
for example regions of the PSF. To illustrate, a point along one
when considering the samples of the radiance array responsible
optical centerline. The source of these bands becomes apparent
light and dark bands that appear as the PSF decays from the
Fig. 4. PSF (left) showing banding with two sample points denoted by an * and a ○ in the PSF with the associated samples in the radiance array (right).

These artifacts result in variations in the reconstructed points. Other
samples from the radiance array demonstrates the source of
Fig. 5. PSF (left) obtained by oversampling the radiance array (right). Increasing the samples used in the reconstruction algorithm produces a smoother response.

Fig. 3. Unscaled PSFs (top) showing spatial variability in \((x)\) (left) and in \((z)\) (right) and scaled PSFs (bottom) showing shift-invariance in \((x, z)\).

depicting how the shape of the impulse response will change
throughout the image in both \(x\) and \(z\)-directions. Equations
\((5)\) and \((9)\) show that the slope of the PSF boundary is set by
\((x - u)\) which will vary for every point in the imaged space.

The boundary of the PSF in scaled coordinates is obtained
similarly from \((15)\). The slope of the PSF boundary in scaled
coordinates is now set by \(u\) which is constant for every point
throughout the volume. This can be seen in the second row of
Fig. 3 which depicts the PSFs in scaled coordinates.

2) Additional Sources of Shift Variance: Unfortunately, the
sampling and reconstruction process introduces further sources
of shift-variance. One source occurs when objects or features
are imaged across multiple micro-lenses. This can occur when
objects are located at the boundary of multiple micro-lenses or
when objects off the focal plane are imaged. Quantization
artifacts result in variations in the reconstructed points. Other
sources of shift-variance result from the reconstruction algo-

The most prominent feature of the PSF is the alternating
light and dark bands that appear as the PSF decays from the
optical centerline. The source of these bands becomes apparent
when considering the samples of the radiance array responsible
for example regions of the PSF. To illustrate, a point along one
of the bright regions in the focal stack is denoted by an * . The
samples integrated to obtain the point, likewise denoted by an
*, are overlaid on the radiance array. Of note is the single
sample that falls within the data (i.e. non-zero region) of the
radiance array. When the samples are summed, this sample is
the most significant contributor to the intensity of the point in
the PSF. A second point is chosen that falls within one of the
dark bands and is denoted by a ○ along with the corresponding
samples from the radiance array. The samples for this point fall
on either side of the data in the radiance array, and therefore
produce a much lower intensity in the corresponding point in
the PSF.

The significance of this observation is that this banding is
not the result of inadequate data collection, or aliasing, but
of inadequately sampling the data that has been recorded. To
address the issue, at least in part, simply increasing the number
of samples used in the reconstruction can significantly reduce
the banding present in the focal stack. Figure 5 depicts the case
where four times as many data samples are used to reconstruct
the PSF from the same radiance array. Once again the points in
the PSF and the corresponding samples in the radiance array
are identified.

While oversampling improves the reconstruction signifi-
cantly, some asymmetric banding and dark regions persist
in the PSF. Both artifacts are identified in Fig. 6. Again
comparing the points in the PSF with the corresponding
samples from the radiance array demonstrates the source of
these artifacts.

First, the persistent banding is caused when the samples for
a point in the PSF have the opposite slope compared to the
radiance data. This can be seen when comparing the radiance
array samples shown in Fig. 5 and Fig. 6. In the former,
the discontinuities in the data are masked by the overlapping
angular information between adjacent micro-lenses. Because
the slope of the samples has the same sign, more samples
consistently fall on regions containing data, which results in
a more uniform intensity. In the latter, because the slope of
the samples has the opposite sign as the data and the data
has a stair step structure, as the slope changes the number of
samples falling on data regions varies more dramatically. This
causes the alternating light and dark regions in the refocused
image. This also explains why this source of banding appears
only on one side of the optical focal plane.
Second, the dark regions on either side of the focused point are related to the persistent banding described above. The point denoted by an * in Fig. 6 falls in one of the dark regions, and its corresponding samples are shown on the radiance array. It can be seen that the samples cross a region in the data where no overlap exists between angular samples. As a result, fewer samples in the reconstruction collect data which results in a lower intensity.

These final artifacts are of particular interest because they are spatially variant. The persistent banding will vary depending on the position of the point and is asymmetric about the point center depending on which side of the optical focal plane the point appears. The dark regions appear when angular data is collected by a single micro-lens, which again is a spatially variant phenomena. Addressing these issues requires direct manipulation of the data or the imaging system and is the subject of future research. For the purposes of this work, these are sources of spatial variability within the reconstruction.

With the imaging system modeled and sources of variability identified and addresses where possible, implementation of a deconvolution algorithm can be pursued.

B. Deconvolution

Deconvolution of the imaging model given in (4) requires that the presence of noise be taken into account. The PSF is a finite bandwidth approximation of an infinite bandwidth function. The result is a frequency response with zeros, which would unacceptably amplify the noise term if direct inversion is attempted. To limit the impacts of noise amplification and PSF mismatch by viewing the image as a random process, we employ the multidimensional Wiener filter given by

$$\hat{F}(\omega_x, \omega_y, \omega_z) = \frac{H^*(\omega_x, \omega_y, \omega_z)}{\left|H(\omega_x, \omega_y, \omega_z)\right|^2 + K} G(\omega_x, \omega_y, \omega_z),$$  

(30)

where $S_n$ and $S_f$ are the power spectrum of the noise and the original image volume respectively. When the power spectra are not known, as is the case here, the ratio of the power spectra can be replaced by a regularization parameter $K$, usually with little performance degradation. Substituting this into (30) gives

$$\hat{F}(\omega_x, \omega_y, \omega_z) = \frac{H^*(\omega_x, \omega_y, \omega_z)}{\left|H(\omega_x, \omega_y, \omega_z)\right|^2 + K} G(\omega_x, \omega_y, \omega_z)$$  

(31)

C. Effects of Regularization and Artifact Identification

To demonstrate the efficacy of deconvolution when applied to plenoptic imaging and to identify artifacts in the reconstruction, two example cases are presented. The first is a simplified 2D case which is used to identify artifacts in the deconvolved reconstruction. The second is a 3D case used to demonstrate results when the algorithm is applied to experimentally acquired volumetric data.

D. 2D Reconstruction

For the 2D example case, the system is modeled with a native focal plane at 100 mm and the remaining system parameters as identified in Table I. The simulations presented here utilize a lenslet pitch that is not an integer multiple of the pixel pitch. As a result, some overlap exists between pixels associated with adjacent micro-lenses. This accurately simulates the physical scenario and these ambiguous pixels must be discarded before creating the radiance array.

Another important aspect of generating the radiance array is properly registering the micro-lens array with respect to the sensor. This step accounts for shifts and/or rotations of the micro-lens array relative to the sensor as well the case where the micro-lens pitch is not an integer multiple of the pixel pitch. While the details of such a process are not considered here, the simulations provided here are calibrated via a calibration image in the same way experimental data would be calibrated. This process identifies the center of each micro-lens with respect to the sensor. The sensor data is then interpolated onto a uniform grid which simulates the perfectly aligned/integer multiple scenario. This ensures that each row of the radiance array corresponds to the same angle while each column represents the same spatial location throughout the array. Without this step, each entry within the 2D array (for 2D images) or the 4D array (for 3D volumes) must be uniquely identified within the $(x, u)$ space or the $(x, y, u, v)$ space respectively. Utilizing a plaid—that is, separably sampled—array significantly improves computational efficiency of the refocusing algorithm. This plaid data is then used by the refocusing algorithm to generate the focal stack. Inclusion of these steps is intended to more accurately model the physical
system, its parameters, and potential sources of error from the algorithm.

A PSF must be selected in order to apply deconvolution to a simulated point field. The PSF chosen for the first simulation is based on a point located at the optical focal plane along the optical axis. This ensures that the point is imaged by a single micro-lens. This is a unique point in the system in that the banding and dark regions present in PSFs located elsewhere in the imaged space are not present. A particle located at this point will exhibit a smooth decay in all directions. This is due to the structure of the data in the radiance array, which does not exhibit the step-like discontinuities associated with points off the focal plane. These characteristics can be seen in Fig. 7. Furthermore, the simulated point is ideal in that each ray used in the simulation originates from an infinitesimally small point. While such a point is not likely to be used for experimental data, it serves to show that the artifacts resulting from the deconvolution are inherent in the system, and not specifically the result of the artifacts in the PSF.

With a PSF chosen, the next step is to simulate an example field of point sources. A test case is selected with three points placed within the imaged space at (1.3mm,91.2mm), (1.9mm,97.5mm), and (−2.5mm,112.5mm). The corresponding focal stack is then generated, and the image is reconstructed utilizing deconvolution. The resulting focal stack is shown in Fig. 8, which clearly depicts the blur resulting from out-of-plane points. Figure 9 compares the results of the deconvolution when various levels of regularization are applied and highlights the importance of selecting the appropriate value for $K$. Too little regularization results in a poor reconstruction dominated by artifacts. Only two of the particles are easily identifiable as can be seen in the leftmost image of Fig. 9. Applying too much regularization results increased blurring of the focal stack as shown by the rightmost image in Fig. 9. Selecting the appropriate level of regularization results in three clearly identifiable points, with significantly reduced blur as shown in the center image of Fig. 9. In practice, this parameter can be set experimentally and held constant as long as the imaging conditions remain the same from one acquisition to another.

Unfortunately, selection of the optimal regularization parameter does not ensure complete elimination of artifacts. These can be seen in the center image in Fig. 9. The deconvolved points display an X-like pattern resulting from mismatch between the PSF and the object response. The uppermost point displays the persistent banding in both the object response and the deconvolved image. These artifacts cannot be eliminated when implementing a single PSF as the banding phenomena is shift-variant. While techniques exist for addressing spatially variant PSFs [47], they increase the computational burden of the algorithm. Nevertheless, these results show a significant improvement over the estimate produced by simply generating a focal stack from the plenoptic data. As a result, we are able to achieve much improved depth resolution after deconvolution processing.

E. Volumetric Flame Reconstruction

A final experimental case is presented to demonstrate how the system characteristics and artifacts demonstrated in the previous sections impact a volumetric reconstruction with a real-world system. The plenoptic camera parameters are presented in Table II. This configuration places the optical focal plane at 50 cm in front of the camera CCD, or 35.1 mm in front of the camera aperture. For the experiment, two flames are utilized as shown in Fig. 10. The first is from a Bunsen burner located to the right of centerline and 6 cm in front of the camera focal plane. The second is from a handheld torch located to the left of centerline and 10 cm behind the camera focal plane. The plenoptic camera is then used to capture an image of both flames burning. The sensor data is shown in Fig.
TABLE II
PLENOPTIC CAMERA PARAMETERS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lenslet Focal Length</td>
<td>$f_l = 0.500$ mm</td>
</tr>
<tr>
<td>Lenslet Pitch</td>
<td>$p_l = 0.125$ mm</td>
</tr>
<tr>
<td>Pixel Pitch</td>
<td>$p_p = 0.0074$ mm</td>
</tr>
<tr>
<td>Number of Pixels</td>
<td>$n_p = 3280 \times 4904$</td>
</tr>
<tr>
<td>Sensor Size</td>
<td>$24$ mm $\times 36$ mm</td>
</tr>
<tr>
<td>Number of Lenslets</td>
<td>$n_l = 193 \times 289$</td>
</tr>
<tr>
<td>Main Lens Focal Length</td>
<td>$f_m = 100$ mm</td>
</tr>
</tbody>
</table>

Next, a PSF must be selected. Many possibilities exist for creating a PSF including imaging a small light source such as a fiber optic light and generating the resulting focal stack. This poses challenges with regard to adequate light collection and preventing imaging the fiber optic cable as well. Simulated sensor data can be used, but these often neglect characteristics specific to the camera and lens being used as well as micro-lens/sensor misalignment artifacts. For this reconstruction, a combination of the experimentally acquired data and simulated data is used.

When viewed closely, the pixels associated with each microlens can be identified across the camera sensor. A subset of the pixels from the central portion of Fig. 11 is shown in Fig. 12. The pixel values are normalized in the subset to clearly show the illuminated pixels. The sensor response to a point source can be simulated by first realizing that a point centered on the optical focal plane, centered on one of the micro-lenses shown in Fig. 12 will illuminate all of the underlying pixels uniformly. The sensor response is then simulated by choosing a centrally located micro-lens, setting the associated pixels to the maximum value, and setting the remaining sensor pixels to zero. The result is shown in Fig. 13. This method offers a suitable compromise for the purposes of this experiment.

The deconvolution begins by creating a focal stack of both the flame image and the PSF. Figure 14 shows an $x-y$ slice of the volume, a single focal plane, at 44 cm where the Bunsen burner is in focus. This focal plane clearly shows the image blur resulting from the torch flame 16 cm further from the camera. In a blur-free image, we would expect to see only the in-focus Bunsen burner flame on the right side of the image. An $(x, z)$ slice through the flames shown in Fig. 15 provides another view demonstrating how the energy from each flame is spread into adjacent focal planes. Again, in the absence of image blur, each flame would be localized to the planes where the flame is present, rather than smeared into adjacent planes as shown in the figure. Note that the scaling of the $z$-axis is the result of choosing a linear spacing between image-space samples, which results in a nonlinear sampling in object space. Sample spacing in object space is determined using the thin lens approximation

\[
\frac{1}{f_m} = \frac{1}{s_i} + \frac{1}{s'}
\]

where $s_i$ is the image distance and $s'$ is the object distance and $f_m$ is the main lens focal length.
Performing the deconvolution, and considering the same \((x, y)\) and \((x, z)\) slices as before shows a significant reduction in out-of-plane energy. Of particular note, the blur from the torch flame is almost completely eliminated in the \((x, y)\) slice at 44 cm shown in Fig. 16, while the structure of the Bunsen burner flame is now well defined. Examining the same \((x, z)\) slice as before shows how the image blur is drastically reduced in the \(z\)-direction in Fig. 17, and again, the structure of the flames is well defined in-plane. The out-of-plane blur is not completely removed in the reconstruction, which is primarily the result of mismatch between the PSF and the object response.

The deconvolution for the 3D case presented required less than 10 seconds on a typical laptop without the benefit of parallel processing. Unfortunately, the integral-based refocusing algorithm used to generate the focal stacks for both the example volume and the PSF is processor intensive and can require an order of magnitude more time to complete. This algorithm is easily parallelizable, and such algorithms have been demonstrated in our lab to generate similar focal stacks twice as fast as a non-parallel algorithm. Nonetheless, a faster
method is desirable, and Fourier-based refocusing algorithms can significantly reduce the total time required to generate an estimate of the imaged volume. The development of such algorithms is the subject of future research.

IV. Conclusions

Deconvolution has been shown to be a highly efficient means of volume reconstruction from plenoptic data. Given the PSF and focal stack, the volume can be reconstructed in seconds. When compared to existing tomographic methods such as MART which may require several hours to reconstruct modestly sized volumes [35–37], this is a significant improvement.

However, limitations remain. The algorithm presented here relies on integral-based refocusing techniques. These integral based algorithms are computationally intensive and can require hours to produce the necessary PSF and focal stacks. Fortunately, FFT-based refocusing algorithms can also reduce this computation time to seconds as well and is the subject of future work.

Finally, the quality of the reconstructions obtained via deconvolution is limited by the mismatch between the PSF and the object response. While the plenoptic imaging system impulse response can be modeled as shift-invariant, sampling and quantization effects result in a highly variable object response. As a result, a single PSF cannot provide a perfect reconstruction of the volume. Future work will explore the possibility of filtering the acquired data and/or the reconstructed response, as well as the development of iterative algorithms that can be used either exclusively or in conjunction with deconvolution to provide improved reconstructions. However, despite the residual blur, such a reconstruction is a significant improvement over the initial focal stack. We have demonstrated the value of the technique for a single example—reconstructing combustion events that could be used to improve burner design. This example illustrates the potential of plenoptic imaging for volume reconstruction in many other applications as well.

Acknowledgments

This work has been partially supported through funding provided by the Air Force Office of Scientific Research, specifically grant FA9550-100100576 (program manager: Dr. Doug Smith). We would also like to thank Tim Fahringer for his assistance in developing the integral-based algorithms utilized in this research and continued discussions regarding plenoptic imaging. Many thanks to Kyle Johnson and Jeffrey Bolan for providing experimental data.

References


