Multi-Camera Plenoptic Particle Image Velocimetry

Roderick R. La Foy¹ and Pavlos Vlachos²

¹ Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA, United States
rlafoy@vt.edu

² Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA, United States
pvlachos@vt.edu

ABSTRACT

Recently, plenoptic or lightfield cameras have been proposed for performing single camera three-dimensional fluid velocity field measurement due to their unique nature of recording both the spatial variation and the propagation direction of the incoming light. The combination of both the spatial and angular information of the light yields the lightfield function, which in turn can be used to create three-dimensional reconstructions. Unfortunately single camera plenoptic PIV measurements suffer from the same low angular resolution problems as do the single camera defocusing PIV and holographic PIV techniques. To increase the quality of the volumetric reconstructions a multiplicative, multiple plenoptic camera reconstruction algorithm is derived and it is shown in lightray simulations that using multiple plenoptic cameras can produce three-dimensional intensity fields with much higher fidelity to the true field than the reconstructions produced by a single camera. Moreover, the intensity fields produced using multiple plenoptic cameras have comparable or higher fidelity than those produced using tomographic PIV techniques with the same number and orientation of cameras.

Introduction

In many fluid systems measuring the full three-dimensional velocity field is essential to understanding the flow dynamics. However, due to increased complexity and computational costs typically associated with three-dimensional measurement techniques, these experimental methods still remain much less common than two-dimensional techniques. Many different experimental techniques have been proposed for measuring three-dimensional velocity fields using PIV including defocusing PIV [1], holographic PIV [2, 3], and tomographic PIV [4, 5] among several others. Recently, so-called plenoptic or lightfield cameras have been proposed for performing single camera three-dimensional fluid velocity field measurement [6-9]. Typical cameras only record the spatial variation of light, however plenoptic cameras have the unique property of additionally measuring the propagation direction of the incoming light. The combination of both the spatial and angular information of the light yields an approximation to the lightfield function, which can be used to create three-dimensional reconstructions. Unfortunately single camera plenoptic PIV measurements suffer from the same problems as do single camera defocusing PIV and holographic PIV techniques due to the low angular resolution.

In this work, the optical properties of the plenoptic camera are used to develop a camera calibration function relating the world coordinates to the lightfield imaged on the camera sensor. Previous work studying plenoptic cameras has primarily focused on producing computationally refocused planes which are then combined to yield a focal stack, but the equations derived here may be used to directly perform volumetric refocusing. This function is then used to develop a single camera multiplicative reconstruction algorithm that produces higher quality reconstructions than previous additive refocusing algorithms. This algorithm is then expanded to allow the use of multiple plenoptic cameras in performing volumetric reconstructions, while previous work has only investigated the use of single plenoptic cameras for PIV measurements [6-9]. It is shown in simulations that using multiple cameras produces three-dimensional intensity fields with much higher fidelity to the true field than are produced by a single camera. Moreover, the intensity fields produced using multiple plenoptic cameras have comparable or higher fidelity than those produced using tomographic PIV techniques with the same camera configuration.

While a camera sensor ostensibly records only two-dimensional data, the lightfield function is in fact seven-dimensional. This function was first described by Arun Gershun as the quantity of light \( L \), of a particular wavelength \( \lambda \), passing through a point in space \((x,y,z)\), traveling in a specific direction \((\theta,\phi)\), at a precise moment in time [10]. However the lightfield function \( L(x,y,z,t,\lambda,\theta,\phi) \) may be described by a four-dimensional approximation to
enable measurement with standard camera sensor. For typical PIV measurements the color of the light $\lambda$ may be neglected. Further assuming that attenuation of the light is negligible, then the radiance along a particular light ray may be assumed to be constant, thus eliminating the need for three spatial dimensions. Finally, since the speed of light is much greater than processes found in typical PIV, the temporal dimension may be assumed to be constant for a particular image frame. Therefore the lightfield may be described by the simplified function $L(x, y, \theta, \phi)$ without loss of generality. This function may be parameterized by several different methods, however in this analysis the function will be represented by the intersection of the light rays with two parallel planes. This may be interpreted as being the two points of intersection $(u, v)$ and $(s, t)$ within the planes. A diagram illustrating this parameterization is shown in Figure 1a.

A variety of experimental techniques have been developed to record lightfield data. Stereo camera setups used to extract depth information from images were the first approximations to lightfield measurements [11, 12]. However stereo systems do not allow for computational refocusing or synthesis of new imaging viewpoints. The first practical lightfield measurements were made using cameras moving on gantries [13], but since the angular information of these systems is encoded in time, they are impractical for fluid measurements. To enable time-resolved measurements of lightfields, arrays of cameras were designed in which each camera measures a different set of lightfield angles [14]. This technique was applied to experimental fluid measurements to develop the Synthetic Aperture PIV technique which allows the fluid particle lightfield to be computationally refocused to produce a volumetric reconstruction [15]. To develop the plenoptic camera system to measure lightfields, an array of micro-lenses was placed directly in front of the sensor in a standard camera allowing for both spatial and angular information to be recorded by a single camera [16].

A plenoptic camera measures both the position and angle of incoming light rays to produce a four-dimensional radiance function [16-19]. This function can be used to computationally refocus the camera to a range of different focal lengths to produce a focal stack. In PIV applications, this focal stack consists of a three-dimensional intensity field of tracer particles upon which standard cross-correlation may be performed to yield a fluid velocity field. There are currently two primary designs of lightfield cameras that are referred to as plenoptic 1.0 cameras and plenoptic 2.0 cameras. Plenoptic 1.0 cameras densely sample the angular information of the lightfield, while sampling at a relatively low spatial resolution. In contrast, plenoptic 2.0 cameras sample the angular information at a low resolution, while sampling the spatial information at a high resolution [20]. While high spatial resolution images have benefits for PIV measurements, the angular resolution is vital for collecting volumetric PIV data. For this reason, the cameras simulated in this work are all plenoptic 1.0 cameras. In Figure 1b the design of a typical plenoptic 1.0 camera is shown consisting of a main lens and an array of microlenses in front of the camera sensor. This arrangement generates an array of images in which each image corresponds to a different set of angles of the lightfield. Typically the set of lightfield angles for each lenslet is relatively small compared to the total angular resolution of the camera. The spatial information of the lightfield is encoded by the position of the pixels on which each microlens image is projected.

The remainder of this work first describes the optical system used to the model the lightfield camera. These models are then used to derive the volumetric refocusing algorithm for a single and multiple plenoptic cameras. Next the results of
Methods

Lightray Simulation

In the simulations, we computationally model between one and four cameras are computationally modeled. The simulated cameras consist of a single large primary lens that focuses the incoming lightrays onto an array of small lenses, referred to as the lenslet array. Figure 1b shows the design of the simulated cameras. The individual lenslets are focused at infinity and thus produce out-of-focus images upon the simulated camera sensor. The tracer particles are simulated by computationally creating a series of lightrays emanating from points surrounding each tracer particle. The intensity profile of the lightrays is Gaussian with respect to the radial coordinate from the tracer particle; this is based upon the assumption that the illuminated tracer particles will produce approximately Gaussian shaped intensity distributions upon the image sensor. The lightray source points have a uniform spatial distribution within a radius corresponding to 1% and greater of the peak intensity; no lightrays are simulated that emanate from outside of this radius. The standard deviation of the Gaussian particles is set equal to 0.7 times the voxel diameter to produce particle images that are consistent in diameter with traditional PIV/PTV measurements. The direction of the lightrays is randomly selected from a uniform distribution that exactly covers the primary lens. This increases the computational efficiency of the simulations since lightrays that do not enter the simulated camera are not created.

The lightrays are propagated through free-space and through the camera lenses using standard optical matrix operations. In this system, the lightray is represented by the vector \( \mathbf{L} = (x, y, \theta, \phi) \) where \( x \) and \( y \) are the position of the lightray from the \( z \)-axis and \( \theta \) and \( \phi \) are the angles of the lightray to the \( z \)-axis in the \( x \) and \( y \) directions respectively. All optical operations then correspond to a simple matrix operation \( \mathbf{L}' = \mathbf{T} \cdot \mathbf{L} \) on the lightray vector \( \mathbf{L} \) where the matrix \( \mathbf{M} \) is determined by the type of optical operation. Using this notation, the simplest operation is given by a lightray propagating through a free-space of length \( d \) which is described by the transformation

\[
\mathbf{L}' = \mathbf{T} \cdot \mathbf{L} = \begin{bmatrix} 1 & 0 & d & 0 \\ 0 & 1 & 0 & d \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \theta \\ \phi \end{bmatrix}
\]

(1)

which can be seen to change the position of the lightray, but not the angle of the lightray as would be expected from a non-refracting medium. This matrix operation is applied three times for each lightray: first, when the lightray passes from the source particle to the primary lens; second, when the lightray passes from the primary lens to the lenslet array; and finally, when the lightray passes from the lenslet array to the camera sensor. The second matrix operation corresponds to a lightray refracting through a thin lens with a focal length \( f \) and is given by

\[
\mathbf{L}' = \mathbf{R} \cdot \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/f & 0 & 1 & 0 \\ 0 & -1/f & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \theta \\ \phi \end{bmatrix}
\]

(2)

which changes the angle of the lightray, but not the position of the lightray. This operation is used when the lightrays pass through the main lens that lies on the optical axis. However since the lenslets do not lie on the optical axis, a modified form of this operation must be used to calculate the transformation of a ray passing through the lenslets. Thus for a thin lens with a focal length \( f \) that is centered at the location \( (s_x, s_y) \) the ray transformation is given by
\[
L' = R \cdot L + S = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/f & 0 & 1 & 0 \\ 0 & -1/f & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \theta \\ \phi \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ s_x/f \\ s_y/f \end{bmatrix}
\]

which can be seen to be equivalent to a refracting lens followed by a prism such that the position and direction of the lightrays changes due to the transformation [17]. The center locations of the lenslets \((s_x, s_y)\) are set to lie on a rectilinear grid with the spacing between the lenslets equal to the lenslet’s pitch. In practice, the precise value of \((s_x, s_y)\) must be calculated individually for each lightray. The total lightray propagation transformation is given by

\[
L' = T_{LS} \cdot (R_L \cdot T_{ML} \cdot R_M \cdot T_{EM} \cdot L + \tilde{S}_L)
\]

where \(T_{EM}\) is the propagation matrix exterior to the camera from the source point to the main lens, \(R_M\) is the refraction matrix for the main lens, \(T_{ML}\) is the propagation matrix from the main lens to the lenslet array, \(R_L\) is the refraction matrix of the lenslet array, \(\tilde{S}_L\) is a shifting vector due to the lenslets being off-axis, and finally \(T_{LS}\) is the propagation matrix from the lenslet array to the sensor. This operation is applied to every simulated lightray. Additionally, the shifting vector \(T_{LS}\) depends upon the particular lenslet that the lightray intersects and thus needs to be calculated after propagating the lightray to the lenslets.

During the computational propagation of the lightrays, it was assumed that there was no attenuation of the lightray intensity due to absorption by either the free-space or the lenses. However, the rays that do not intersect the optical elements are assumed to be lost as though the interior of the camera was coated with 100% absorbing paint. Stops are not specifically simulated within the camera, but the lenses act as stops by programming the simulation to remove all the lightrays that do not intersect the lenses. The intensity field produced on the simulated camera sensor is calculated by integrating all the lightrays that intersect the sensor.

The simulated sensor consists of a rectilinear grid of square pixels that return a weighted sum of the lightrays’ irradiance. The lightrays intersect the simulated sensor at relatively small angles with respect to the optical axis, so a paraxial approximation is used to assume that the irradiance is independent of the incident angle. Since the lightrays typically intersect the pixels at non-integral positions, the intensity produced by each lightray must be interpolated. This is accomplished by assuming that the lightrays have square profiles with dimensions equal to the dimensions of a single pixel. Then the intensity of the lightray is split between the nearest four pixels to the point of intersection based upon the area of overlap. The total intensity of each pixel in the simulated camera sensor is equal to the sum of the intensities of all the interpolated lightrays that intersect the particular pixel. The final sensor image is saved both as a double precision array as well as 16 bit unsigned integer array. To accurately model discretization effects upon the volumetric reconstruction, the 16 bit images were used in all subsequent processing. Figure 2 shows the qualitative agreement between the simulated and real plenoptic sensor images of a PIV field.

**Volumetric Image Reconstruction**

To perform the volumetric reconstructions from the plenoptic camera data, a four-dimensional lightfield function is extracted from the camera sensor and integrated over a specified domain. The camera may then be computationally refocused on different regions by varying the integration domain of the lightfield function. The matrix optics transformations described in the previous section are used to analytically derive the reconstruction algorithm.

Qualitatively, the reconstruction algorithm is equivalent to refocusing the image formed on the camera lenslet array to a new position either in front of the lenslet array or behind the lenslet array. To computationally refocus the lightfield onto the new focal plane requires knowing both the spatial and angular information of the lightrays. This knowledge can be extracted from the camera sensor image produced by the lenslets. The spatial information of the lightfield is encoded by the position of the lenslets. This intuitively makes sense, since the main lens is focused onto the lenslet array; the sensor in a standard camera is found where the lenslet array is in a plenoptic camera, so the lenslets are essentially acting as large pixel sensors. The angular information of the lightfield is encoded within the images produced by each lenslet. For example, the light that falls on the pixels on the right side of a lenslet image may be
predominantly from the left side of the main lens (the actual orientation will depend upon where the lightrays originated with respect to the focal plain of the main lens).

Analytically, refocusing the lightfield may be represented by re-parameterizing the spatial and angular information from the sensor into \((u,v,s,t)\) coordinates. In this parameterization the \((u,v)\) coordinates correspond to the location that a lightray intersects the main lens, while the \((s,t)\) coordinates give the location that the same lightray intersects the lenslet plain. All coordinates are measured with respect to the optical axis of the camera. Using this parameterization, the irradiance on the \((s,t)\) plane, or equivalently the image produced by a standard camera, is given by integrating the lightfield across the \((u,v)\) coordinates

\[
E(s,t) = \frac{1}{D^2} \iint L(u,v,s,t) \cdot A(u,v) \cdot \cos^4(\alpha) \, du \, dv
\]

where \(D\) is the distance from camera aperture to the \((s,t)\) plane, \(A(u,v)\) is an aperture function returning one within the aperture and zero elsewhere, and finally \(\alpha\) is the angle the lightray makes with the \((s,t)\) plane. Since only the relative magnitude within the refocused volume is important for measuring velocities, the \(1/D^2\) factor may be eliminated. Additionally, by assuming that the main lens acts as the camera aperture stop, the aperture function \(A(u,v)\) may be assumed to be uniformly equal to one. While in general a camera many have multiple lenses as well as a separate aperture stop, the camera’s optical system can be assumed to be equivalent to a single thin lens with it’s outer edge acting as the aperture stop without loss of generality. Finally, by assuming that the lightrays arrive on the \((s,t)\) plane with a small angle to the optical axis, the paraxial approximation can be used to assume that \(\cos^4(\alpha) = 1\). With these assumptions, the measured image irradiance equation on the \((s,t)\) plane becomes

\[
E(s,t) = \iint L(u,v,s,t) \, du \, dv.
\]

However, to volumetrically reconstruct the image requires integrating the lightfield over a range of different domains, so a relationship between the \((u,v,s,t)\) coordinates and the lab coordinates \((x,y,z)\) must be calculated. This is accomplished by using the optical matrix transformations to propagate a lightray emanating from the coordinate \((x,y,z)\) first to the main lens to find the relationship with the \((u,v)\) coordinates and second, to the lenslet array to find the relationship with the \((s,t)\) coordinates. A system of equations is then constructed from these relationships to transform the lab coordinates into the lightfield coordinates. Propagation of a lightray to the main lens is given by the transformation

\[\text{(a) Real sensor image \hspace{1cm} (b) Simulated sensor image}\]
\[
\begin{bmatrix}
    u \\
    v \\
    \theta_{(u,v)} \\
    \phi_{(u,v)}
\end{bmatrix}
= \begin{bmatrix}
    x + \theta z \\
    y + \phi z \\
    \theta \\
    \phi
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & z & 0 \\
    0 & 1 & 0 & z \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    \theta \\
    \phi
\end{bmatrix},
\]

which gives the relationship between the lab coordinates and the \((u,v)\) coordinates. Additionally, the propagation of the lightray to the lenslet array is given by the operation

\[
\begin{bmatrix}
    s \\
    t \\
    \theta_{(s,t)} \\
    \phi_{(s,t)}
\end{bmatrix}
= \begin{bmatrix}
    \left(1 - \frac{d_{ML}}{f_M}\right) x + \left(\frac{d}{f_M} + \left(1 - \frac{d_{ML}}{f_M}\right) z\right) \theta \\
    \left(1 - \frac{d_{ML}}{f_M}\right) y + \left(\frac{d}{f_M} + \left(1 - \frac{d_{ML}}{f_M}\right) z\right) \phi \\
    -\frac{x}{f_M} + \left(1 - \frac{d_{ML}}{f_M}\right) \theta \\
    -\frac{y}{f_M} + \left(1 - \frac{d_{ML}}{f_M}\right) \phi
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & d_{ML} & 0 \\
    0 & 1 & 0 & d_{ML} \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    1 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    \theta \\
    \phi
\end{bmatrix},
\]

which gives the relationship between the lab coordinates and the \((s,t)\) coordinates. Then extracting the first two equations from Equations (7) and (8) gives the system of equations relating the lab coordinates and the lightray coordinates as

\[
\begin{align*}
    u &= x + \theta z \\
    v &= y + \phi z \\
    s &= \left(1 - \frac{d_{ML}}{f_M}\right) x + \left(d_{ML} + \left(1 - \frac{d_{ML}}{f_M}\right) z\right) \theta , \\
    t &= \left(1 - \frac{d_{ML}}{f_M}\right) y + \left(d_{ML} + \left(1 - \frac{d_{ML}}{f_M}\right) z\right) \phi
\end{align*}
\]

which being a simple linear system, may then be solved for the original lab coordinates lightray vector to yield

\[
\begin{align*}
    x &= u \cdot \left(1 - \frac{z}{f_M}\right) + \frac{z \cdot (u - s)}{d_{ML}} \\
    y &= v \cdot \left(1 - \frac{z}{f_M}\right) + \frac{z \cdot (v - t)}{d_{ML}} \\
    \theta &= \frac{s \cdot f_M + u \cdot (d_{ML} - f_M)}{d_{ML} \cdot f_M} \\
    \phi &= \frac{t \cdot f_M + v \cdot (d_{ML} - f_M)}{d_{ML} \cdot f_M}
\end{align*}
\]

To perform the volumetric reconstructions, the relationship between the \((u,v,s,t)\) coordinates of the lightfield and the pixel coordinates on the sensor must also be known. To calculate this transformation, the lab coordinates in Equation (10) are substituted into the full camera lightray Equation (4) to yield the sensor pixel coordinates.
Then if the camera sensor image is denoted by the function $I_s(x,y)$, the coordinates in Equation (11) may be used to extract the lightfield from the sensor according to the function $L(u,v,s,t) = I_s(x_s,y_s)$. Since $u$ and $v$ correspond to the coordinates that the lightrays intersect the main lens, they must satisfy $p_M / 2 \geq \sqrt{u^2 + v^2}$ where $p_M$ is the pitch of the main lens. Additionally, since the lenslets are focused at infinity, the images they form on the camera sensor are of the main lens. Thus each pixel under an individual lenslet corresponds to one $(u,v)$ coordinate. The $s$ and $t$ coordinates correspond to the locations of the centers of the individual lenslets. The lightfield function $L(u,v,s,t)$ is then calculated by varying the $(u,v,s,t)$ coordinates over their respective domains to produce the sensor coordinates $(x_s,y_s)$. Since these coordinates likely do not correspond to integral pixel values in the sensor, cubic interpolation is used to calculate $L(u,v,s,t)$.

A calibration process is required to refocus an actual plenoptic camera. This process involves several steps. First the location of the individual lenslet images on the sensor must be determined to calculate the $(s,t)$ coordinates. This process is relatively straightforward since the images projected by the lenslets appear as periodic circles on a uniformly dark background and this format is commonly used in standard camera calibration procedures. The second step in performing the calibration process is to calculate estimates for the values of main lens focal length $f_L$ and the main lens to lenslet array distance $d_{ML}$. Both of these values should be approximately known from the camera design, however precise estimates may be determined by refocusing the plenoptic camera on a calibration grid. The parameters are then calculated by finding the distance from the plenoptic camera at which the calibration grid becomes focused. Additionally higher order calibration terms may be accounted for using the positional information in the calibration grid.

While the camera produces a lightfield with a focused image on the $(s,t)$ plane, by shifting this plane to a new position on the $z$-axis denoted by $(s',t')$, an image focused on a different depth may be produced. The image at this new plane may be calculated from Equation (6). To computationally refocus the image on an arbitrary lab coordinate $(x,y,z)$ requires finding a relationship to the computational focal plane $(s',t')$. Using the first two equations in the system of Equations (10) the transformed $(s',t')$ coordinates may be calculated as

$$
\begin{align*}
s' &= u + d_{ML} \left( \frac{u-x}{z} - \frac{u}{f_M} \right) \\
t' &= v + d_{ML} \left( \frac{v-y}{z} - \frac{v}{f_M} \right)
\end{align*}
$$

(12)

The $(u,v)$ plane that is coincident with the main lens may also be transformed to a new location along the optical axis. However, this increases the complexity of the volumetric reconstruction algorithm while not conferring a particular advantage. So the transformed coordinates are set equal to the original coordinates $(u',v') = (u,v)$.

Once the transformed coordinates $(u',v',s',t')$ are calculated, Equation (6) may be used to computationally refocus the image to the lab coordinate $(x,y,z)$ by the function

$$
V(x,y,z) = \iiint L\left( u,v,u+d_{ML} \left( \frac{u-x}{z} - \frac{u}{f_M} \right),v+d_{ML} \left( \frac{v-y}{z} - \frac{v}{f_M} \right) \right) du dv.
$$

(13)
Evaluating the integral in Equation (13) requires interpolating the lightfield from the discretely sampled function $L(u,v,s,t)$. However, it was found that the results produced by using a simple discrete sum were nearly identical to the results produced by using numerical quadrature, so the lightfield function only needs to be interpolated at a limited number of points for each evaluation.

Using a form of Equation (13) is suitable for performing computational refocusing that closely resembles the focusing process in a standard camera, however, the quality of the volumetric reconstructions can be improved by taking into account the unique nature of fluid measurement images. In typical PIV images, there are a large number high intensity points surrounded by a low intensity background. For this reason, limited angle-number tomographic reconstruction techniques such as the MLOS and MART techniques can be used to reconstruct volumetric images from PIV data. This reasoning applies equally well to the images collected by lightfield cameras, so an MLOS based lightfield reconstruction algorithm was developed.

By extracting the same pixel under every lenslet image and combining these pixels together, a sub-aperture image is formed. These images can be extracted from the lightfield $L(u,v,s,t)$ by evaluating the function over the domain of $(s,t)$ while holding $(u,v)$ constant [16]. Qualitatively, the plenoptic camera refocusing algorithm is equivalent to overlapping and adding the sub-aperture images together. To process the lightfield data with the MLOS algorithm requires overlapping and multiplying the images together. However, directly performing this operation is problematic since typically the lenslet images will contain some vignetting near the edge of the images. To illustrate this effect, a collection of sub-aperture images is shown in Figure 3a. In this figure it can be seen that the vignetting causes some of the sub-aperture images to have nearly uniformly zero intensities. Thus a simple multiplication of these images together would yield a uniformly zero volumetric reconstruction.

To overcome the effect of the vignetting, a weighting function is introduced. To calculate the weighting function, an image of a uniformly bright background is taken with the plenoptic camera and the lightfield $L(u,v,s,t)$ is extracted. The weighting function is set equal to the sum of the intensities of each sub-aperture image $w(u,v) = \sum_{s,t} L(u,v,s,t)$.

Then the weighting function is linearly scaled to lie in the range $0 \leq w(u,v) \leq 1$. This results in the weighting function being approximately equal to one for the sub-aperture images taken from the center of the lenslet images and approximately zero for the sub-aperture images taken from the edge of the lenslet images. Figure 3b shows the weighting function calculated from the sub-aperture image in Figure 3a. The weighting function only needs to be calculated once for a particular plenoptic camera configuration since it is independent of the scene being imaged by the camera.
The MLOS plenoptic camera reconstruction is then performed by raising each sub-aperture image to the power of the weighting function. This scaling results in the low intensity sub-aperture images being nearly uniformly equal to ones and as a result, contributing relatively little to the total reconstruction. The specific formula used to calculate the MLOS volumetric reconstruction is

$$V(x,y,z) = \left( \prod_{u_{viss}} \prod_{v_{viss}} L(u,v,u+d_{ML}, \frac{u-x}{z} - \frac{u}{f_M}, v+d_{ML}, \frac{v-y}{z} - \frac{v}{f_M}) \right)^{\gamma_r}$$

where $\gamma_r$ is used to rescale the reconstruction to ensure that the intensity histogram of the reconstruction approximately equals the histogram of the individual sub-aperture images. The value of the exponent $\gamma_r$ will vary with the lightfield camera configuration and focal distance, but typically $\gamma_r$ will be approximately equal to the reciprocal of the number of pixels under each lenslet.

**Multiple Camera Reconstructions**

While a single plenoptic camera can produce a volumetric reconstruction suitable for performing PIV measurements, the fidelity of the reconstructions, and thus the accuracy of the velocity field measurements, may be dramatically improved by using multiple cameras as will be shown later. Since the computational refocusing of a single lightfield camera directly yields a three-dimensional intensity field, a basic process for combining the data from multiple cameras is a simple extension of the MLOS algorithm to three-dimensional data. It is possible that other algorithms may exist that yield better reconstructions, but the analysis here focuses on the MLOS algorithm.

In a similar manner to tomographic reconstruction algorithms, the first step in the multiple camera reconstruction is to perform an intensity normalization of the individual camera reconstructions so that the intensity histograms of the individual reconstructions are approximately equal. This ensures that each camera contributes equally to the composite reconstruction and no information is lost. Then by denoting each plenoptic camera's reconstruction as $V_i(x,y,z)$, the composite reconstruction is given by

$$V_i(x,y,z) = \prod_{i=1}^M V_i(x,y,z)^{\gamma_c}$$

where $M$ is the number of lightfield cameras and $\gamma_c$ is a rescaling exponent. Since in Equation (14) the single camera reconstruction is raised to the exponent $\gamma_r$, the two exponents are dependent upon one-another. Thus only one exponent needs to be controlled during the reconstruction process. For this reason, the multiple camera reconstruction exponent is set equal to the reciprocal of the camera number $\gamma_c = 1/M$ as is typically done in MLOS reconstructions [4].

**Reconstruction Fidelity Metrics**

A variety of different quality measurements are used to evaluate the accuracy of the lightfield reconstructions and to compare them to standard tomographic reconstructions. The fidelity of the reconstructions was measured using the following metrics: the zero mean reconstruction quality factor, the error in the reconstructed particle positions, the RMS error of the measured velocity field with respect to the true velocity field, and the percentage of outlier vectors as defined by the Universal Outlier Detector (UOD) [21]. The zero mean reconstruction quality factor is a modified form of the reconstruction quality factor defined by Elsinga, et al [5]. The quality factor as defined by Elsinga is

$$Q = \frac{\sum_{x,y,z} V(x,y,z) \cdot T(x,y,z)}{\sqrt{\sum_{x,y,z} V(x,y,z)^2 \cdot \sum_{x,y,z} T(x,y,z)^2}}$$

where the summation is taken over all coordinates in the reconstructed field. This metric is essentially a normalized cross-correlation between the reconstructed intensity field $V(x,y,z)$ and the true intensity field $T(x,y,z)$.
function will produce a value of $Q$ approximately equal to one for nearly perfect reconstructions while poorer quality reconstructions will yield values approaching zero. However, it was found that this quality factor produced artificially high values for high seeding densities due to the fact that as the seeding density increased, the functions $V(x,y,z)$ and $T(x,y,z)$ approached nearly uniform values of one. Thus regardless of the difference between the true field and the reconstructed field, the correlation produced a value nearly equal to one.

To overcome this effect a zero mean reconstruction quality is used in analyzing the lightfield reconstruction results. This quality factor is defined as

$$Q^* = \frac{\sum_{x,y,z} V(x,y,z) \cdot T(x,y,z)}{\sqrt{\sum_{x,y,z} V(x,y,z)^2 \cdot \sum_{x,y,z} T(x,y,z)^2}}$$

where $V(x,y,z)$ and $T(x,y,z)$ are given by zero mean reconstruction field and zero mean true intensity field respectively. The effect this has on the reconstruction quality is shown in Figure 4 where the standard reconstruction quality $Q$ and the zero mean reconstruction quality $Q^*$ are plotted as functions of particle density in particles per image pixels (ppp) for standard tomographic reconstructions. Below a particle density of approximately $ppp = 0.1$ the two reconstruction quality factors differ by less than 10%. For higher particle densities, the zero mean quality factor $Q^*$ continues to decrease while the standard quality factor increases beyond a particle density of approximately $ppp = 0.2$. Intuitively higher particle densities will make volumetric reconstructions more difficult due to the rapidly increasing number of particles overlapping in the camera images, thus the zero mean reconstruction quality factor produces more reasonable values at high densities. In all following analysis the zero mean reconstruction quality factor $Q^*$ will be used.

The error in the reconstructed particle positions is measured by fitting a Gaussian intensity profile to all reconstructed particles and then comparing these fits to the known true particle positions. Typically ghost particles that have no corresponding true particles will be created during the reconstruction process, so the nearest true particle to each reconstructed particle is identified after performing the Gaussian fit. Then the particle position error is measured for the nearest reconstructed particle to the true particle. This process is repeated for the set of all true particles to produce a distribution of particle position errors. Typically the position errors followed a normal, zero mean distribution as can be seen in Figure 5 which shows a scatter plot of the reconstructed particle position errors for the $x$ and $z$ axes.

**Figure 4** A comparison between the standard tomographic reconstruction quality factor $Q$ and the zero mean reconstruction quality factor $Q^*$ showing the artificially high values produced by $Q$ at very high particle seeding densities.
The uncertainty in the reconstructed particle positions is used as the primary position-based metric to assess the quality of the lightfield reconstructions. The uncertainty is reported as the distance in voxels from the center of the reconstructed particles in which there is a 95% chance that the true particle is located assuming a normal distribution of errors. The $z$-axis uncertainty is specifically reported since this tends to be larger than either the $x$ or $y$ axis uncertainty and sets an upper bound in the particle position error.

The simulated particle positions are advected using an analytical solution for a vortex ring described in [22]. The simulated vortex ring translates through the particle volume with a constant velocity without experiencing dissipation or changing shape. Three-dimensional PIV measurements are made on the reconstructed particle volume to return a measured velocity field that can be compared to the analytical solution of the vortex ring to produce an RMS velocity field error defined by

$$\epsilon_{\text{RMS}} = \frac{1}{N} \sum_{i=1}^{N} \left( (U_i - U_{Ti})^2 + (V_i - V_{Ti})^2 + (W_i - W_{Ti})^2 \right)$$

where $(U_i, V_i, W_i)$ are the velocity vectors measured using the PIV correlations. Two passes are completed using the Robust Phase Correlation algorithm [23]; during the first pass, the UOD is used to locate likely outlier vectors. These vectors are then subsequently replaced before the second PIV pass is processed. The final reconstruction quality metric is given by the ratio of vectors flagged by the UOD in the second pass using the UOD residual threshold $r_0 > 2$.

### Simulation Parameters

Due to the complex design of plenoptic cameras and the nature of the reconstruction algorithms, the parameter space that can be studied to analyze the use of plenoptic cameras for fluid measurements is very large. For this reason, the

<table>
<thead>
<tr>
<th>Simulation Parameters</th>
<th>Main Lens</th>
<th>Lenslets</th>
<th>Sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal Length</td>
<td>$f_M = 10 \text{ mm}$</td>
<td>$f_i = 28 \mu m$</td>
<td>—</td>
</tr>
<tr>
<td>Pitch</td>
<td>$p_M = 5 \text{ mm}$</td>
<td>$p_L = 14 \mu m$</td>
<td>$p_S = 1.4 \mu m$</td>
</tr>
<tr>
<td>Aperture</td>
<td>$f_M / # = 2$</td>
<td>$f_i / # = 2$</td>
<td>—</td>
</tr>
<tr>
<td>Number</td>
<td>1</td>
<td>$n_{x_s} \times n_{y_s} = 328 \times 328$</td>
<td>$n_{x_s} \times n_{y_s} = 3280 \times 3280$</td>
</tr>
</tbody>
</table>

Table 1: A table listing the plenoptic camera parameters used in the lightfield simulations. These parameters were held constant for all completed simulations.
simulations are designed to approximate the Lytro (www.lytro.com) plenoptic camera design. The results of the simulations will then be used to design a complementary series of experiments using the Lytro cameras. Table 1 lists the parameters used in the lightfield simulations. These parameters are slightly different than the actual design parameters used in the Lytro cameras to simplify the simulations, however, the differences are minor and the results should still apply to the actual cameras.

While the parameters described here may appear to restrict the results of simulations by assuming that diffraction is negligible, the results that will be presented are scalable to arbitrary camera dimensions. If diffraction is accounted for in the design of the lightfield camera, then the reconstruction fidelity will in fact be higher for cameras with dimensions larger than the simulated camera.

Results

Before investigating the reconstruction fidelity of the lightfield cameras over a large parameter range, the reconstruction algorithm was specifically studied. To validate the multiplicative reconstruction algorithm given by Equation (14), both experimental and simulated lightfields are computationally refocused. The experimental data was collected using a single Lytro camera viewing a field of fluorescent particles illuminated with a 1 cm thick 532 nm laser sheet. The data was computationally refocused using both the additive refocusing algorithm and the multiplicative refocusing algorithm as is shown by the volume slices in Figure 6. The background noise in the multiplicatively reconstructed field appears much lower than the noise in the additively reconstructed field. The simulations conducted using both the additive and multiplicative refocusing algorithms uniformly showed the multiplicative algorithm performing better.

A range of different parameters are studied in the lightfield camera simulations including varying (1) the distance of the cameras from the particle volume, (2) the thickness of the particle volume, (3) the number of simulated cameras, (4) the angle of the simulated cameras to the particle volume, and (5) the density of the particles inside the volume. Additionally the fidelity of the reconstructions produced by the lightfield cameras are compared to the fidelity of reconstructions produced using standard cameras. These reconstructions are completed using the MART tomographic reconstruction algorithm [4, 5]. A series of calibration images are generated to provide calibration data for the tomographic reconstructions. These results show that single lightfield cameras perform relatively poorly for measuring particle velocity fields, but high-quality measurements may be achieved by combining the data from two or more cameras. Additionally it is shown that there is an optimal focal distance for using lightfield cameras due to trade-offs in the optical resolution of the cameras. At this optimal distance, the lightfield cameras perform better than traditional cameras for PIV measurements.

Since the lightfield cameras depth-of-field decreases with the focal distance of the camera, but the angular resolution of

![Figure 6](image-url)  
**Figure 6** Computationally refocused images taken using the Lytro lightfield camera during a PIV experiment. (a) The PIV field refocused using the standard additive refocusing algorithm described in Equation (13). There is relatively high-magnitude background noise in the image due to out-of-focus particles. (b) The same PIV field refocused using the multiplicative refocusing algorithm described in Equation (14). The background noise level in this image is much lower than the noise level produced by the additive reconstruction.
the camera increases, it was speculated that there might be an optimal distance to perform volumetric measurements with the lightfield cameras. To test this hypothesis, a series of simulations were completed for a range of different camera focal distances and particle volume thicknesses. Figure 7 shows the results of these simulations. In these tests the zero mean normalized cross correlation quality factor is measured for reconstructions produced by a single lightfield camera located on the particle volume axis. The camera focal distance $s_{O_m}$ is normalized by the aperture of the main lens $p_M$ so that the results may be scaled to arbitrarily sized primary lenses. Additionally the particle volume was scaled to fill the entire field-of-view of the camera at each different focal distance. Due to the resolution trade-off, the quality factor $Q^*$ has a local maximum near $s_{O_m}/p_M \approx 10$. The quality factor also decreases as the particle volume increases due to the particles near the edge of the volume becoming more out-of-focus. For all subsequent lightfield camera simulations a volume thickness-to-width ratio of 1:4 is used as this is a common ratio in tomographic PIV experiments; additionally, a dimensionless focal distance of $s_{O_m}/p_M = 8$ is also used since the quality factor experiences a peak at this relative distance.

Once an optimal distance was determined for completing volumetric measurements with the lightfield camera, the number of cameras and their configuration was investigated. Between one and four camera is simulate in these tests. The cameras are located equal distances from one-another with angles measured from the $z$-axis of the particle volume. Using this system, two cameras are located along a line on the $x$-axis, three cameras are located at the vertices of an equilateral triangle on the $xy$-plane, and four cameras are located at the vertices of a square on the $xy$-plane. The angles of the lightfield cameras were varied from 0 to 60 degrees. Standard cameras are also simulated in addition to the lightfield cameras for these tests. The standard camera images are used to perform tomographic reconstructions using the MART algorithm.

The results of these simulations are shown in Figure 8. The single plenoptic cameras perform relatively poorly using all four quality metrics. However, reconstruction fidelity dramatically increased by adding a single additional plenoptic camera. The $z$ particle position uncertainty in particular is between two and three times as high for a single camera as for the cases using multiple cameras. Additionally, from this data it is apparent that using more than two lightfield cameras to perform volumetric reconstructions only marginally increases the fidelity of the reconstructions. This data also shows that the fidelity of the reconstructions created using multiple plenoptic cameras has a maximum value around 25 to 35 degrees, agreeing well with the results found in tomographic PIV simulations [5].

The tomographic MART reconstruction qualities created using the simulated standard camera is also shown in Figure 8. For these tests, two different apertures on the standard cameras were simulated: an aperture of $f/2$, which is the same aperture used on the plenoptic cameras, and an aperture of $f/20$, which is the aperture necessary to have an

---

**Figure 7** A graph showing the zero-mean normalized cross correlation quality factor as a function of the dimensionless camera focal distance for a range of different volume width to thickness ratios. The camera focal distance $s_{O_m}$ is normalized by the aperture of the main lens $p_M$. The quality factor has a peak at $s_{O_m}/p_M \approx 10$ due to the camera’s angular resolution decreasing with distance while the depth-of-field increases.
equivalent depth-of-field as the plenoptic cameras. The reconstruction quality of the images suffers from a low depth-of-field for the large aperture case. The depth-of-field is large for the small aperture case, but diffraction effects start causing uniform blurring of the particles. The results shown in the figure are from tomographic reconstructions created using two simulated cameras. The quality of the MART reconstructions performed poorer than the plenoptic camera reconstructions using two or more cameras for all cases with the exception of the MART $f/2$ UOD outlier ratio. In this case it is likely that a large number of ghost particles are reconstructed due to many out-of-focus particles caused by the low depth-of-field. These ghost particles may then act as a low-pass filter on the measured velocity field, thus decreasing the number of outlier vectors. This effect needs further investigation.

The final series of simulations investigated the effects of particle density on the lightfield camera reconstruction quality. In these tests the particle density of the simulated volume was varied over a range typically used in volumetric PIV experiments. The particle density $ppp$ is generally measured as the ratio of the number of particles imaged to the total number of pixels covered by the particle volume in the camera image. However, the number of used pixels is poorly defined for a plenoptic camera since the images produced by the lenslet array typically does not cover 100% of the sensor. Additionally for Plenoptic 1.0 cameras, the final image produced by the camera is based upon the number of lenslets rather than the number of pixels on the sensor. So to provide a fair comparison to particle densities reported in other volumetric PIV papers, the particle density $ppp$ is measured as the ratio of the number of particles to the number of reconstructed pixels.

![Graphs](image)

**Figure 8** Graphs showing several different reconstruction fidelity metrics for a range of lightfield camera numbers and angles. Tomographic MART reconstructions of the same particle field are shown for comparison. (a) A graph showing the quality factor as a function of the camera configuration. (b) A graph showing the $z$ particle position uncertainty as a function of the camera configuration. Two or more cameras dramatically increase the $z$ resolution. (c) The RMS velocity error as a function of the camera configuration. Two or more cameras perform equally well in reconstructing the velocity field. (d) The UOD outlier ratio as a function of the camera configuration.
Figure 9  Graphs showing several different reconstruction fidelity metrics for a range of particle densities. Tomographic MART reconstructions using both $f/2$ and $f/20$ apertures of the same particle field are shown for comparison. (a) A graph showing the quality factor as a function of the particle density. (b) A graph showing the $z$ particle position uncertainty as a function of the particle density. (c) The RMS velocity error as a function of the particle density. (d) The UOD outlier ratio as a function of the particle density.

The particle density tests are completed using two plenoptic cameras and two standard cameras positioned 25 degrees off the particle volume axis. The reconstruction qualities are measured for particle densities over the domain $1 \cdot 10^{-3} \leq ppp \leq 3 \cdot 10^{-1}$. The standard cameras used to create the MART reconstructions have simulated apertures of $f/2$ and $f/20$.

The particle density simulations showed that the reconstruction fidelities generally decreased as the particle density increased. However, the PIV velocity metrics showed an increase in quality with the particle density as is shown in Figure 9. The decrease in the measured velocity field error with increasing particle density is likely due to the relatively simple nature of the prescribed velocity field. While the vortex ring does produce three-dimensional motion relative to the simulated cameras, the motion has sufficiently large scales when compared to the total particle volume that it is likely that the ghost particles produced in the reconstructions only contribute relatively little noise to the PIV correlations. At the same time, increasing the particle density results in the true particles increasing the signal in the correlations. It is thus possible that increasing the particle density will only decrease the measured velocity field noise in cases where the flow is relatively simple, but determining these cases may be difficult. This effect needs to be investigated in further studies.

The standard camera tomographic MART reconstructions performed more poorly than the plenoptic camera reconstructions for nearly all cases. The exception to this is again found for the case of the large aperture UOD outlier ratio that may still be explained by the reconstructed ghost particles acting as a low-pass filter on the measured velocity field.
Conclusions

This work presents advancements to the field of volumetric PIV developing novel tools based on lightfield (plenoptic) imaging. First, by using ray optics, a volumetric reconstruction algorithm for refocusing single plenoptic cameras is developed. Then it is shown that due to the unique properties of PIV data, that a multiplicative refocusing algorithm may be used to improve the quality of the refocused images. This multiplicative algorithm is then modified so that data from several plenoptic cameras may be combined together. This reconstruction algorithm is next used to calculate volumetric reconstructions from artificial sensor data created using lightray simulations of an illuminated particle field. These reconstructions are analyzed in terms of several metrics measuring both the reconstruction fidelity to the true particle intensity field and the accuracy with which the particle motion is measured.

The results of the simulations show that multiple camera plenoptic reconstructions can yield higher quality data than traditional volumetric measurement techniques under some circumstances. Single plenoptic camera reconstructions are shown to yield relatively low quality data, but it is shown that high quality volumetric data may be taken by simultaneously using two plenoptic cameras. The simulations show that there is only a small benefit from using more than two cameras, however. Additionally, it is shown that under the optimal conditions for plenoptic cameras, standard tomographic cameras perform relatively poorly.

Future work on the plenoptic camera PIV system will focus on several primary areas. First, the parameter range of the simulations will be expanded. The simulations carried out for this current work were mainly focused on one particular camera design used in a limited number of configurations. However, it is unlikely that this particular camera design or the simulated configurations are optimal for all PIV measurements. Thus the design of the plenoptic cameras will be specifically investigated in future research in an attempt to determine the optimal parameters for volumetric particle measurements.

The information learned from these simulations will then be used to design complementary experiments to validate the simulation results. Experiments will initially be conducted using the commercial Lytro cameras, but since these cameras have very low frame rates and cannot be externally triggered, they are poorly suited for experimental work. Thus the results of the simulations will also be used to design and build high-speed plenoptic cameras that can directly be used for fluid dynamics research.

Finally, more advanced reconstruction algorithms will be developed and validated using the simulation tools that were developed for this work. While traditionally the images from plenoptic cameras are additively refocused, it was shown that combining the data from the lenslet images using a multiplicative algorithm increases the reconstruction fidelity. It is likely that the reconstruction fidelity could be further increased by using a MART or other similar limited-angle tomography algorithm. Moreover, since each camera can independently generate a volumetric reconstruction, it is also possible that these reconstructions could be combined together in a more complex manner to produce better results.

Acknowledgments

The partial support of the NSF “IGERT Multi-scale Transport in Environmental and Physiological Systems (MultiSTEPS)” and NSF grant “IDBR: Instrument development for three-dimensional fluid flow measurements of freely-flying animals” are gratefully acknowledged.

REFERENCES


