Title
Single-Shot Diffuser-Encoded Light Field Imaging

Authors
Antipa, Nicholas
Necula, Sylvia
Ng, Ren
et al.

Publication Date
2016

Peer reviewed

Single-Shot Diffuser-Encoded Light Field Imaging

Nicholas Antipa
UC Berkeley
naantipa@gmail.com

Sylvia Necula
Georgia Institute of Technology
sylvia.necula@gmail.com

Ren Ng
UC Berkeley
ren@berkeley.edu

Laura Waller
UC Berkeley
waller@berkeley.edu

Abstract

We capture 4D light field data in a single 2D sensor image by encoding spatio-angular information into a speckle field (caustic pattern) through a phase diffuser. Using wave-optics theory and a coherent phase retrieval method, we calibrate the system by measuring the diffuser surface height from through-focus images. Wave-optics theory further informs the design of system geometry such that a purely additive ray-optics model is valid. Light field reconstruction is done using nonlinear matrix inversion methods, including ℓ_1 minimization. We demonstrate a prototype system and present empirical results of 4D light field reconstruction and computational refocusing from a single diffuser-encoded 2D image.

1. Introduction

Recording and processing of 4D light fields offers new capabilities over traditional 2D imaging. These include the ability to compute a different focus and depth of field after the fact, change the viewpoint slightly, compute depth and computationally correct for optical aberrations in the camera's lens. Various approaches have been studied to project the 4D light field onto a 2D sensor such that the light field can be inferred from a single-shot. The two most common approaches are based on microlens arrays and attenuation masks placed at a small distance from the image sensor. Microlens arrays use regular grid of lenslets to refractively encode the light field onto the sensor. Attenuation masks encode the light field into shadow patterns on the sensor, providing enhanced resolution, but at the cost of absorbing a portion of the light [40] [32] [23].

In this paper we generalize both of these approaches to use a transparent phase plate (e.g. a diffuser). Compared with microlens arrays, we allow arbitrary height maps. Compared with attenuation masks, we allow a similar coding effect, but with higher light throughput. Such diffusers provide an inexpensive and flexible means for single-shot light field recording using an off-the-shelf diffuser. A challenge in utilizing such diffusers is that they are generally diffractive, producing speckles that exhibit significant wave-optical effects due to interference. We show here a theoretical analysis of when it is appropriate to use wave-optics versus ray optics models for interpreting light fields encoded by a phase plate. Leveraging this, our camera is designed such that the image synthesis and reconstruction are correctly described by a ray optics model, independent of the object or illumination coherence. We further present a wave-optics calibration routine, based on the Transport of Intensity Equation (TIE) [38], that recovers the phase surface height map from images captured through focus. We use ray tracing to build a linear model for the system, based on the phase measurements, which we then invert to recover the 4D light field. Lastly, we demonstrate the efficacy of these methods by showing experimental results.

2. Related Work

Single-shot capture of in-camera light fields using microlenses is an active area of research, with a long history dating back to the early 1900s [28, 21, 35, 2, 33, 34, 30, 27, 43]. The technique is variously referred to as integral, plenoptic and light field imaging. The main tradeoff is reduced image resolution in order to sample angular information. Microlens approaches have also been applied in microscopy [25], where a wave-optics model can be used to exploit diffraction effects for higher-resolution reconstruction at some distances [9]. This approach recognizes that microlens arrays become subject to wave-optics effects as their sizes shrink. In this paper, we extend this line of thinking to a new framework in which the phase-encoded surface need not be a pre-designed periodic array of lenslets, but can be any phase (or amplitude) mask, even with irregular surface relief and diffractive properties.

978-1-4673-8623-4/16/$31.00 © 2016 IEEE

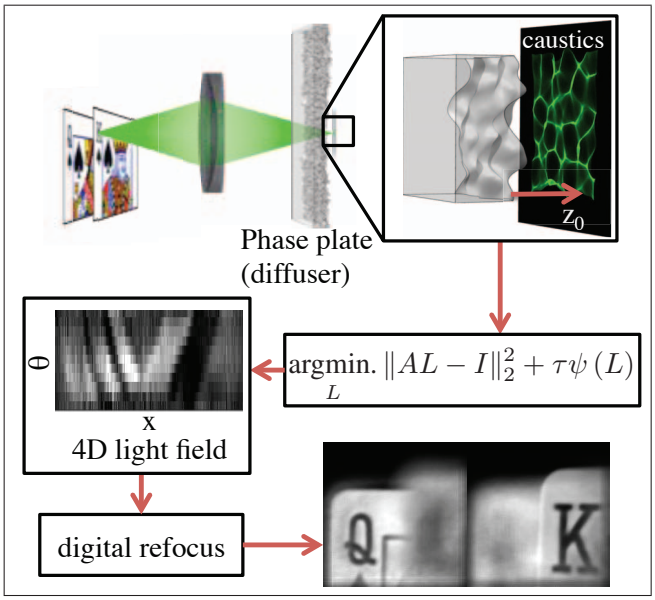

Figure 1. Pipeline for recording and reconstructing light fields with phase plates (a diffuser). The object light passes through an imaging lens and the phase plate, then propagates to the sensor, where caustics encode spatial and angular information. A linear inverse problem is solved to reconstruct the light field, which contains 3D information, enabling digital refocus, among other benefits.

Attenuation masks are another approach for encoding light fields in a single-shot [40]. Recent work has shown that these systems can be modeled using a matrix method that is amenable to compressive sensing, possibly overcoming the resolution trade-offs inherent in 4D imaging with 2D sensors [32, 23]. These techniques require significantly more computational resources than microlens systems in order to infer the light field from the sensor measurements, and the masks attenuate a significant amount of light. In this paper, we use phase encoding rather than attenuation encoding, thereby avoiding light loss. However, our system is similar to the mask-based methods in that it implements a multiplexed-type measurement that is able to exploit a priori information (e.g. sparsity) for possible resolution benefits.

Microlens arrays and attenuation masks have also been used without a main imaging lens to create very flat cameras for 2D imaging [37, 5], since the lens function can be achieved in computation. Very small 2D cameras have also been created with diffractive optics and computation [36] using a purely wave-optics model. Although it is not shown in this paper, our approach of using diffusers could also be extended to flat imaging device designs, but with volumetric reconstructions resulting from the light field data.

Other approaches to light field capture include angle sensitive pixels [42, 20], aperture scanning with time multiplexing [26], macroscopic lens arrays [15], and camera arrays [46, 41]. Various attempts at obtaining higher image resolution have been proposed, for example depth-aware splatting [13] and hybrid imaging by also using a high-resolution 2D camera [7].

Designed phase plates have been used in 2D imaging to extend depth of field [11, 10] and improve signal for phase retrieval [3, 31]. Random planar refractive masks have been used to reconstruct 2D images and estimate depth [12]. Recent efforts have attempted to image objects through unknown random diffusers [24, 19, 4], which suggests possible future applications for our work.

3. Theory

Consider the 4D light field, $L(x, y, \theta, \phi)$, inside a camera body after passing through a primary imaging lens having numerical aperture less than 1. Each ray is described by two lateral coordinates, (x, y), and two angular coordinates (θ, ϕ), corresponding to x and y, respectively. The light at the imaging plane then passes through a non-absorbing phase plate (e.g. a diffuser), and propagates a distance z_0 to the 2D sensor plane, where it is recorded as an intensity image. The system architecture is illustrated in Figure 1.

As an example application of this approach, we use an inexpensive off-the-shelf Light Shaping Diffuser [1]. These diffusers are thin pieces of polymer with refractive index $n \approx 1.5$ that are planar on the input side and have an output surface that can be modeled as a smooth random Gaussian surface, described by a height field, $D(x, y)$:

$$D(x, y) = s\left[R(x, y) * K(\sigma)\right], \tag{1}$$

where s is a unitless scaling factor, $K(\sigma)$ is a zero-mean Gaussian smoothing kernel having full-width half-maximum (FWHM) value of σ and $R(x, y)$ a set of random height values chosen from the normal distribution at each discrete sample location (x, y). It is assumed that σ is greater than the wavelength of light, λ, thereby avoiding sub-wavelength scattering effects. We show in Section 3.3 that diffusers create high-contrast intensity patterns (caustics) at certain distances. These patterns are unique to particular regions on the diffuser surface, thus they encode multiplexed spatial and angular information in an invertible way.

3.1. Wave Optics Model

The phase plate can be thought of as a thin transparency which imparts a spatially-varying phase delay, $\phi_D(x, y)$, onto any wave passing through it. Consider a coherent incident wave having amplitude $A(x, y)$ and phase $\phi_i(x, y)$. The wavefront exiting the phase plate (at $z = 0$) will be the product of the incident wave's complex-field and the complex transmittance of the phase plate [17],

$$E(x, y, z = 0) = A(x, y)e^{i[\phi_i(x,y)+\phi_D(x,y)]}. \tag{2}$$

Assuming the phase plate has homogeneous index of refraction, n, the phase delay is directly proportional to the height map of the phase plate

$$\phi_D(x,y) = \frac{2\pi\Delta n}{\lambda} D(x,y), \tag{3}$$

where Δn is the refractive index difference between the phase plate and surrounding medium. For plane wave illumination, $A(x,y)$ and $\phi_i(x,y)$ are constant, so Equation (2) simplifies to

$$E(x,y,z=0) = \exp\left[i\phi_D(x,y)\right]. \tag{4}$$

The resulting complex-field at the sensor, $E(x_0,y_0,z_0)$, is predicted using Fresnel diffraction theory [17]. Finally, the intensity at the sensor is proportional to absolute value squared of the complex-field, $I(\xi,\eta;z_0) \propto |E(\xi,\eta;z_0)|^2$.

In this wave-optical model, the gradient of the incident beam's phase describes the local angle of propagation, according to the Poynting vector description of energy flow. For partially coherent (or incoherent) light, however, the optical field cannot be described by a single complex field; rather, it is the incoherent superposition of many [8]. The Wigner function provides a wave-optical analog to the light field [47]. Still, it is significantly more complicated than a ray optics model, which is preferred where accurate.

3.2. Ray Optics Model

Ray tracing approaches are generally thought to be only valid for incoherent imaging, whereas diffractive effects require wave optics. However, the diffractive nature of the phase masks used here does not always imply that a full wave-optical model is necessary. We will show here that for sufficiently small diffuser-to-sensor distances, ray optics is a suitable approximation, irrespective of whether the object is coherent or incoherent.

To model ray transport through a dielectric interface described by (1), a full 3D ray tracing approach is suitable. We assume that the diffuser is flat enough that we can neglect self-shadowing, total internal reflection and multiple refractions. However, for weak diffusers (on the order of 1°) and apertures below $F/2.8$, the maximum angle of incidence at the phase surface will be on the order of 11°, which is small enough to adopt the paraxial (small angle) approximation to Snell's law. While the paraxial model is not necessary for our methods to work, it provides valuable insight into the behavior of phase diffusers.

For simplicity, we describe our model with a 2D paraxial light field, $L(x,\theta)$, traveling along the optical axis of the system ($+z$ direction)[1]. In the paraxial regime, refraction at each interface becomes a linear form of Snell's law,

[1]Extending to 4D non-paraxial optics is straightforward with the ray tracing approaches used here, based on the surface normals in $\mathbb{R}^3$

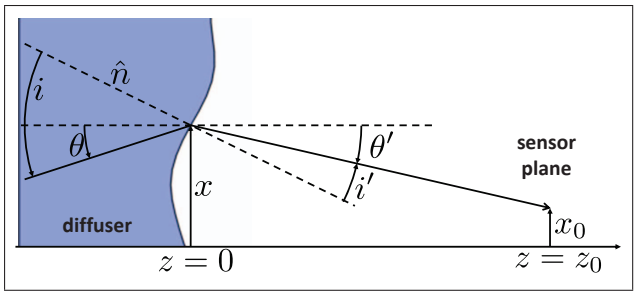

Figure 2. Ray geometry for a single ray hitting a diffuser surface and refracting before reaching the sensor plane.

$ni = n'i'$, where i is the incident angle and i' is the output angle, both measured relative to the interface normal; n and n' are the input and output refractive index, respectively. The diffuser surface gradient, $D_x(x) = -1/\hat{n}$, with $\hat{n}$ being the surface normal. As shown in Figure 2, we can write i in terms of θ and the surface gradient: $i = \theta + D_x(x)$. Similarly, the exit angle is given by $i' = \theta' + D_x(x)$. Plugging this into Snell's Law, the output ray angle is

$$\theta' = \frac{n}{n'}\theta + \left(\frac{n}{n'} - 1\right) D_x(x). \tag{5}$$

Since refraction changes the ray angle but not its position, the output position does not change ($x' = x$). Substituting (5) into the paraxial ray propagation equation, $x_0 = x + \theta' z_0$, results in the ray position, x_0, being

$$x_0 = x + z_0\left[\frac{n}{n'}\theta + \left(\frac{n}{n'} - 1\right) D_x(x)\right], \tag{6}$$

where z_0 is the diffuser-sensor separation distance.

The final irradiance on the sensor is the sum of the radiance along all rays that fall within each pixel from any angle. This is equivalent to projecting the resulting light field (at the sensor) along the angle dimension before sampling.

3.3. Caustics

To illustrate how Equation (6) behaves, we explore the refraction and propagation of a plane wave through a diffuser described by Equation (1) with index n, in air ($n' = 1$). In (x,θ) space, a plane wave is represented by equal radiance at all positions and a single angle, as shown in Figure 3a. After applying Equation (5), each ray in L is displaced in angle by $(n-1)D_x(x)$. Notice that immediately after the diffuser, the irradiance is unchanged (projecting through angle will not change irradiance). However, propagation of the warped light field shears this curve (by shifting each point in x by an amount $z_0 n\theta'$ in accordance with Equation (6)). Because the shear is proportional to θ, the angular ripples created by the diffuser result in structure in the final projected irradiance (Figure 3b). Another way

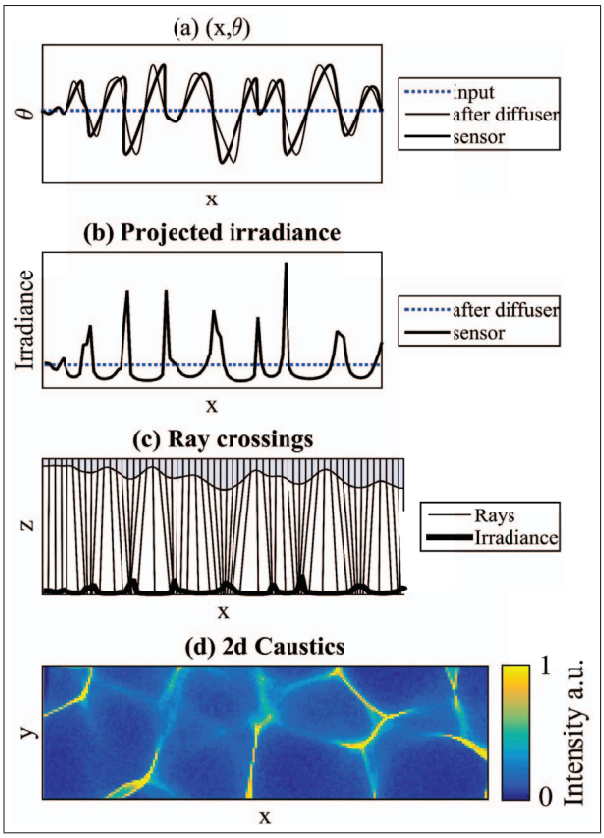

Figure 3. Simulation of diffuser caustics from plane wave illumination. (a) Space-angle plots for the input plane wave, post-diffuser, and sensor plane. (b) The resulting irradiance at the sensor, generated by integrating over θ. (c) Axial cross-section of rays passing through the diffuser to form caustic patterns at the sensor plane. (d) 2D caustics predicted by 4D ray tracing.

to visualize this is via the peaks that arise from the bunching of rays under regions where the local diffuser curvature causes it to act like a positive lens (see Figure 3c).

The intensity peaks induced by the combination of refraction and propagation, known as caustics, create an intensity pattern at the sensor that is directly related to the local structure of the mask surface. Intuitively, these peaks will be located under the strongly convex regions of the diffuser. Because Equation (5) is linear in θ, changing the incident illumination angle leads to a linear shift of the caustic pattern. Thus, the intensity pattern formed by light striking any part of the diffuser is uniquely determined by the incident angle and the local diffuser structure. This is also true for amplitude-coded masks and is the fundamental building block for the invertible linear model in Section 3.5.

For a 4D treatment of the light field, we can treat (x, θ) and (y, ϕ) independently and apply Equation (6) to each direction separately. This leads to 2D caustic patterns demonstrated in Figure 3d, which was simulated using our ray tracing model (described in Section 3.5).

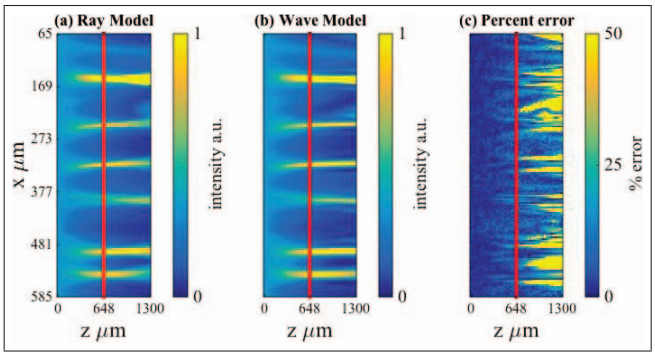

Figure 4. Simulated axial (x-z) slices of a plane wave after passing through a diffuser, under both our wave optics and ray optics models. The red line corresponds to a Fresnel number of $F = 1$ (at $z = 648\mu m$ for our system), which demarcates approximately the propagation distance at which the ray and wave models diverge. For smaller propagation distances, the models agree.

3.4. Coherence

Diffusers generally produce diffractive speckle patterns that depend greatly on the coherence properties of the illumination [16]. Therefore, to design our system to be valid across a broad range of objects and illumination, we must consider the effects of diffraction and coherence.

Diffraction depends on the wavelength and distance propagated, as compared to the size scale of the object. These effects are captured by the Fresnel number, F, which provides a guideline for describing the amount of diffraction (larger F means less diffraction effects) [17]:

$$F = \frac{a^2}{z\lambda}, \tag{7}$$

where z is the propagation distance and a is the size of the object under consideration (for diffusers, we use $a = \sigma$). Generally, diffraction becomes important when $F < 1$, for example outside of the small defocus regime [18]. Our ray optics model should thus be accurate as long as the diffuser-to-sensor propagation distance corresponds to a Fresnel number larger than $F = 1$.

Figure 4 compares the intensity evolution of a diffuser illuminated by a coherent plane wave, using both wave and ray optics models. Our wave-optics model applies Fresnel propagation to the electric field in Equation (4) with $\lambda = 532nm$, then computes irradiance at propagation distance z as $I_c(x, y, z) = |E(x, y; z)|^2$. Our ray optics model uses the methods discussed in Section 3.2 to compute the output positions of each ray, then calculates irradiance, I_r, by binning the rays onto the same grid used in the wave model. Figure 4 shows an x-z slice of the irradiance pattern generated by each method. Since wave optics captures interference and diffraction effects, discrepancies are considered errors due to the ray optics approximation.

Clearly, high-contrast caustic patterns arise before $F = 1$, though the rms error between the wave and ray models is small, indicating that the ray optics model is valid. For larger propagation distances ($F < 1$), the error is dominated by the interference fringes surrounding each caustic peak. This suggests that a good choice of propagation distance[2] is slightly less than $F = 1$, providing strong caustic patterns (thus good signal to noise), while also making the system independent of illumination coherence properties[3]. This greatly simplifies computation by allowing us to ignore lighting conditions and coherence, while still preserving the phase-coding behavior of the diffuser.

3.5. Linear Model

Because we designed the system to operate in a regime where interference is negligible, we effectively treat the object as temporally and spatially incoherent. That is, all light striking a single detector pixel adds linearly in intensity. This enables the optical system to be represented via a linear mapping from the 4D light field (radiance) before the diffuser, $L(x, y, \theta, \phi)$, to the 2D sensor irradiance, $I(x_0, y_0)$.

To derive the linear forward mapping as a matrix, $\mathbf{A}$, we consider sampled versions of both $L(x, y, \theta, \phi)$ and $I(x_0, y_0)$. We discretize $L(x, y, \theta, \phi)$ into 4D boxes, each with spatial extent Δx by Δy and angular extent $\Delta\theta$ by $\Delta\phi$. Figure 5 shows a 2D example with N spatial samples and P angular samples at each position.

To construct $\mathbf{A}$, consider that the first column is mapped by multiplying with a column vector of zeros and a 1 in the first element. Physically, this corresponds to a bundle of light rays striking the diffuser across an area Δx centered at x_n from angles spanning $\Delta\theta$ centered at θ_p, then propagating to the sensor. Therefore, column $j = Pn + p$ of $\mathbf{A}$ is the sensor image due to uniform illumination at x between $x_n - \frac{\Delta x}{2}$ and $x_n + \frac{\Delta x}{2}$, and θ between $\theta_p - \frac{\Delta\theta}{2}$ and $\theta_p + \frac{\Delta\theta}{2}$. Each entry of $\mathbf{A}$, $a_{i,j}$, is the fraction of light that strikes pixel i from the light field point indexed by j. This is equivalent to the fractional area of each sensor pixel in (x, θ) space that falls within box j, as shown in Figure 5(a).

Only the 2D phase plate shape and refractive index are needed for computing the entire $\mathbf{A}$ matrix. Extending to 4D implies a convenient method for generating $\mathbf{A}$: for the j^{th} column, we generate many rays randomly distributed across light field bundle j, then compute their output positions using (6). Finally, we bin rays into sensor pixels, then column-stack the resulting image as column j of $\mathbf{A}$. Extending this method to multiple colors is done by repeating the above procedure at different wavelengths, accounting for the diffuser material dispersion curve.

[2]This refers to the propagation distance from the diffuser to the sensor and is independent of amount of depth present in the original scene.

[3]Note that this metric does not account for phase height, which may lead to larger discrepancies for strong phase objects.

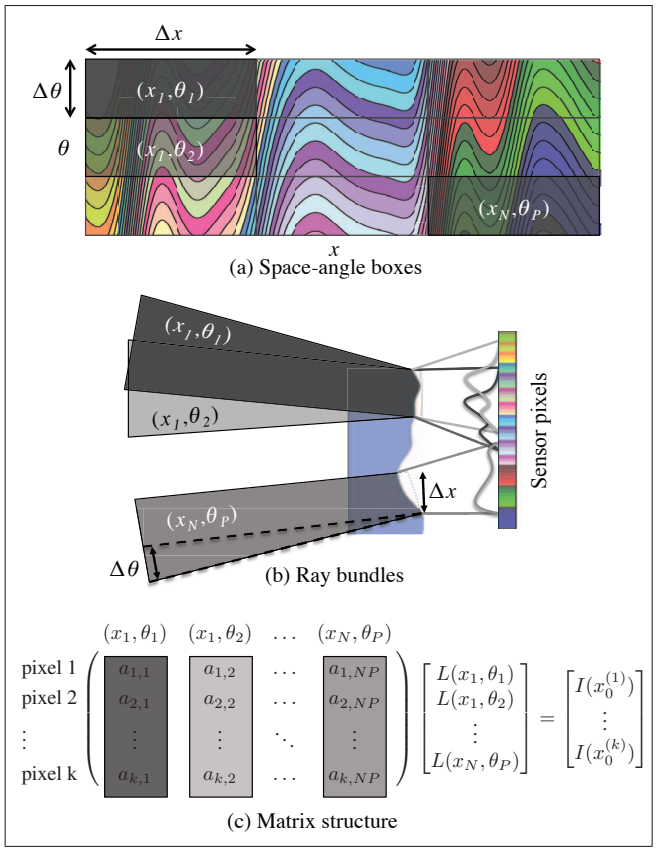

Figure 5. (a) Finite-sized boxes (in grey) of the light field correspond to ray bundles hitting the diffuser. The structure of each sensor pixel in (x, θ) takes on the shape of the sheared diffuser gradient. Here, each band of color corresponds to all the (x, θ) pairs that strike a single sensor pixel. A bundle will span multiple pixels in (x, θ) space. (b) Each ray bundle creates a unique caustic pattern on the sensor, which shifts according to the input angle. The set of sensor pixels illuminated matches those within each bundle's box in (a). (c) The corresponding matrix structure for a light field consisting of N spatial samples with P angular samples at each x. $I \in \mathbb{R}^k$ and L is a 1D vector $L \in \mathbb{R}^{NP}$.

3.6. Inverse Problem

The ultimate goal of the linear forward model is to recover the 4D light field from a single sensor measurement by solving the inverse problem. Here, we explore the properties of $\mathbf{A}$ that enable stable inversion and how these numerical requirements map to the physical system. Equation (6) states that output ray position is linearly related to the input angle and the diffuser gradient. Bundles sharing common (x, y) coordinates will exhibit the same caustic structure for all incident angles; the only difference will be a lateral shift at the sensor due to their relative input angle differences. This shifting behavior is visualized in Figure 5b.

In the matrix, this shift behavior imparts a block circulant structure where columns belonging to a common (x, y)

are shifted copies of each other. The critical implication of this is that it is possible for two different light field regions to strike the exact same set of output pixels, while remaining distinguishable. Hence, the diffuser-encoded light field recording process can be thought of as a multiplexing approach, since each pixel measures a linear combination of points in light field space. This is equivalent to two columns of $\mathbf{A}$ having nonzero elements in identical rows. If these rows have proportional *values* at each nonzero pixel, the rows are linearly dependent and the problem is ill-posed.

In order to prevent this issue, the phase mask must have sufficient structure within each bundle so that it creates distinct values in each row. This is the key to why a diffuser works: its random surface has extremely low probability of repeating a pattern. However, this also implies that we are not free to make the light field sampling arbitrarily dense in the lateral direction. In other words, we must choose Δx and Δy to be sufficiently large compared to σ so that distinct caustics are present within each bundle at the sensor. In practice, setting Δx and Δy to be at least σ leads to reasonable inversion behavior. Additionally, the angular sampling is limited by the sensor pixel pitch, since the change in angle between two consecutive bundles must cause a shift at the sensor that is large enough to be sampled correctly.

Once a well-conditioned $\mathbf{A}$ matrix has been constructed, we recover the 4D light field, L_{rec}, from a 2D sensor image by solving the following least squares inverse problem:

$$L_{rec} = \arg\min_{L} \|\mathbf{A}L - I\|_2^2 + \tau\psi(L), \tag{8}$$

where $\psi(L)$ is a regularization function and τ is a scalar regularization parameter. As a baseline, we solve the ℓ_2 regularized problem using $\psi(L) = \|L\|_2^2$. In the experimental section, we also explore the use of two nonlinear regularizers for exploiting sparsity: 3D Total Variation (3DTV) regularization from Tian et al. [39] and ℓ_1 regularization in the 2D wavelet domain, similar to Veeraraghavan et al. [40]:

$$\alpha_{rec} = \arg\min_{\alpha} \|\mathbf{A}W^{-1}\alpha - I\|_2^2 + \tau\|\alpha\|_1, \tag{9}$$

where α is a vector of the wavelet coefficients for each subaperture image, and W^{-1} is the inverse 2D wavelet transform operator. It follows that $L_{rec} = W^{-1}\alpha_{rec}$.

4. Implementation

To implement our approach in practice, two key pieces are needed. One is the imaging system itself, which is described in Figure 1. The other is a diffuser height map; we must either use a known surface shape, or measure it. We propose here a phase-from-focus method for measuring the diffuser surface shape *in situ*. The height field for ray tracing can be computed from the phase map using Equation 3. From this one-time calibration, we can then computationally reconstruct light fields according to Section 3.5.

4.1. Experimental Setup

In our experiments, we use a 1 degree Light Shaping Diffuser (Luminit, LLC). Because the diffuser-to-sensor propagation distance needed is short (approximately $650\mu m$) and our physical sensor's packaging does not allow the diffuser to be placed so close, we add a $4f$ relay system with $1.33\times$ magnification to image the diffuser onto the sensor. This introduces unwanted aberrations in the diffuser wavefront; however, our TIE measurement incorporates these into the result, thereby mitigating their effects on our final images. To record images, we use an imaging lens with f=125mm stopped down to f/16. The total field of view is approximately 25 mm laterally. Our PCO Edge 5.5 sCMOS monochrome camera has 5 Megapixels with pixel pitch 6.5 μm. The image of the diffuser is placed $648\mu m$, corresponding approximately to $F = 1$, in front of the sensor.

4.2. Phase imaging for diffuser calibration

Since phase is linearly related to surface profile and can be measured with sub-wavelength accuracy, phase retrieval methods are a practical means for calibrating the diffuser surface *in situ*. In particular, our method uses only a few images taken at different focus positions [22], which is easy to implement in our existing system by translating the camera between images. The TIE describes how intensity evolves axially with respect to phase [38]

$$\frac{\partial I(x,y)}{\partial z} = -\frac{1}{k}\nabla_\perp \bullet [I(x,y)\nabla_\perp \phi(x,y)], \tag{10}$$

where $\nabla_\perp$ is the gradient operator in the lateral (x, y) dimensions only and $k = 2\pi/\lambda$ is the wave vector magnitude. Using this equation, a few images taken with small defocus can be used to solve for phase. The algorithm we use is a GP-TIE solver [22] which is offered open-source on Laura Waller's Computational Imaging Lab website.[4]

Experimentally, we calibrate the system using a coherent collimated plane wave from a 532 nm laser diode. The camera is mounted on a micrometer axial translation stage, which we use to take a focus stack of 100 images with z step size of 25.4 μm. In fact, only 5 through-focus images are necessary for a good phase result which correctly predicts the other intensity images; however, we use the full stack to ensure robustness. A few raw images are shown in Figure 6, along with the recovered phase map. Notice that the diffuser becomes invisible at focus, where it is a pure phase delay that does not change intensity. The phase profile recovers an average angle of $1.4\,\mathrm{deg}$ after magnification, and $\sigma = 18\mu m$, with rms surface height of $1.15\mu m$. We observe good agreement between the measured caustics within the Fresnel range and the caustics predicted using our matrix.

[4] http://www.laurawaller.com/opensource/

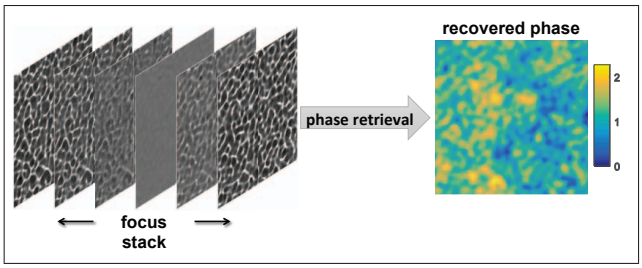

Figure 6. A stack of irradiance images collected at different focus positions in our experimental setup are used to recover the phase map of the diffuser surface, which directly relates to height.

4.3. Simulations

In order to account for the fact that light field radiance values may change over the span of a single bundle in the inversion matrix, we use a high spatial resolution forward matrix to simulate the sensor data and a lower resolution one for inversion. Our synthetic input light field, $L(x, y, \theta, \phi)$, is rendered using POV-Ray with 5×5 angular samples and 512×512 spatial samples using scene files shared from [40]. We then use the ray tracing approach described in Section 3.5 in conjunction with the experimentally measured diffuser data from Section 4.2 to create a forward matrix, $\mathbf{A}_f$, that projects $L(x, y, \theta, \phi)$ onto a 1024×1024 pixel sensor. We simulate the sensor data as the matrix-vector product $\mathbf{A}_f L$, then add 5% Gaussian noise. To invert the problem, we trace a second lower resolution matrix, $\mathbf{A}$, that projects back to 128×128 spatial samples and 5×5 angular samples in light field space. We then use $\mathbf{A}$ to solve for L.

4.4. Inverse Problem

The inverse problem is solved by the gradient descent solver LSMR [14] for the ℓ_2 problem and Two Step Iterative Shrinkage/Thresholding (TwIST) for 3DTV and ℓ_1 [6] to solve equations (8) and (9). We choose the angular resolution such that each successive caustic pattern shifts by at least 1 pixel compared to the neighboring bundles. The spatial sampling is determined heuristically, but we find it must be at least σ to obtain good results.

5. Results

5.1. Simulation Results

The rendered light field is represented in RGB form, so we solve each color independently using the 2D wavelets with ℓ_1 regularizer. Figure 7 shows the simulated sensor data as well as the caustic patterns and the original and reconstructed (x, θ) plots along one line in the image. Figure 7 (e) shows the irradiance detected at the diffuser, computed by summing the recovered light field over θ and ϕ. We use the shift-and-add technique [34] to digitally refocus the reconstructed light field to several planes and compare with the original refocused light field. We achieve good agreement between the original light field and our reconstruction, with PSNR of 25 dB in the synthetic focus images, and 18.5 dB in the (x, θ) images.

5.2. Experimental Images

With the calibration complete, individual diffuser-blurred images are recorded through the imaging path of the system. We restrict the primary lens aperture to $f/16$ solely for the purpose of controlling aberrations in the $4f$ system. We solve the inverse problem using TwIST with the 3DTV regularizer or with an ℓ_1 regularizer on in the 2D wavelet coefficients of each sub-aperture image. We are able to reconstruct a light field with 11×11 angular samples in each direction and 170×170 lateral samples from a 2048×2048 sensor measurement, and demonstrate a large refocus distance. Figure 8 shows experimental results for a pair of playing cards placed $+40$ mm and -40 mm from the native focal plane, using 3DTV regularization. In the (x, θ) plot, strong angular variation is visible, including occlusion effects. Figure 9 shows another experimental set of images from a ruler tilted at an angle to the optical axis, using wavelet regularization.

We find that the ℓ_2 regularizer performs poorly on real-world data, destroying angular structure before it brings noise under control. 3DTV is extremely good at suppressing noise, but imparts a distinctive piecewise-constant look to the sub-aperture images that is only suitable for piecewise-constant scenes. For natural objects, we find the most robust approach to be ℓ_1 regularized 2D wavelets. Figure 10 shows an example of all three regularizers applied to the playing card image from Figure 8.

6. Limitations

While the thin phase plate approach provides high light throughput, we find that significant noise is present in the sub-aperture images, which we attribute to ray error due to aberrations in the $4f$ relay lenses. Although the TIE phase measurement helps overcome this by measuring the aberrated diffuser phase for monochromatic on-axis illumination, it does not compensate for off-axis or chromatic aberrations induced by the relay optics. This severely limits the F-numbers and wavelengths we are able to use, but is not a fundamental limitation of our approach. To overcome this, we plan to place the diffuser directly in front of the sensor in the future. This will dramatically improve consistency of our phase measurement for different illumination angles, and will enable us to model polychromatic illumination using the diffuser's material dispersion curve. This will also have the added benefit of making the system more compact.

Because we solve for the radiance within an entire ray

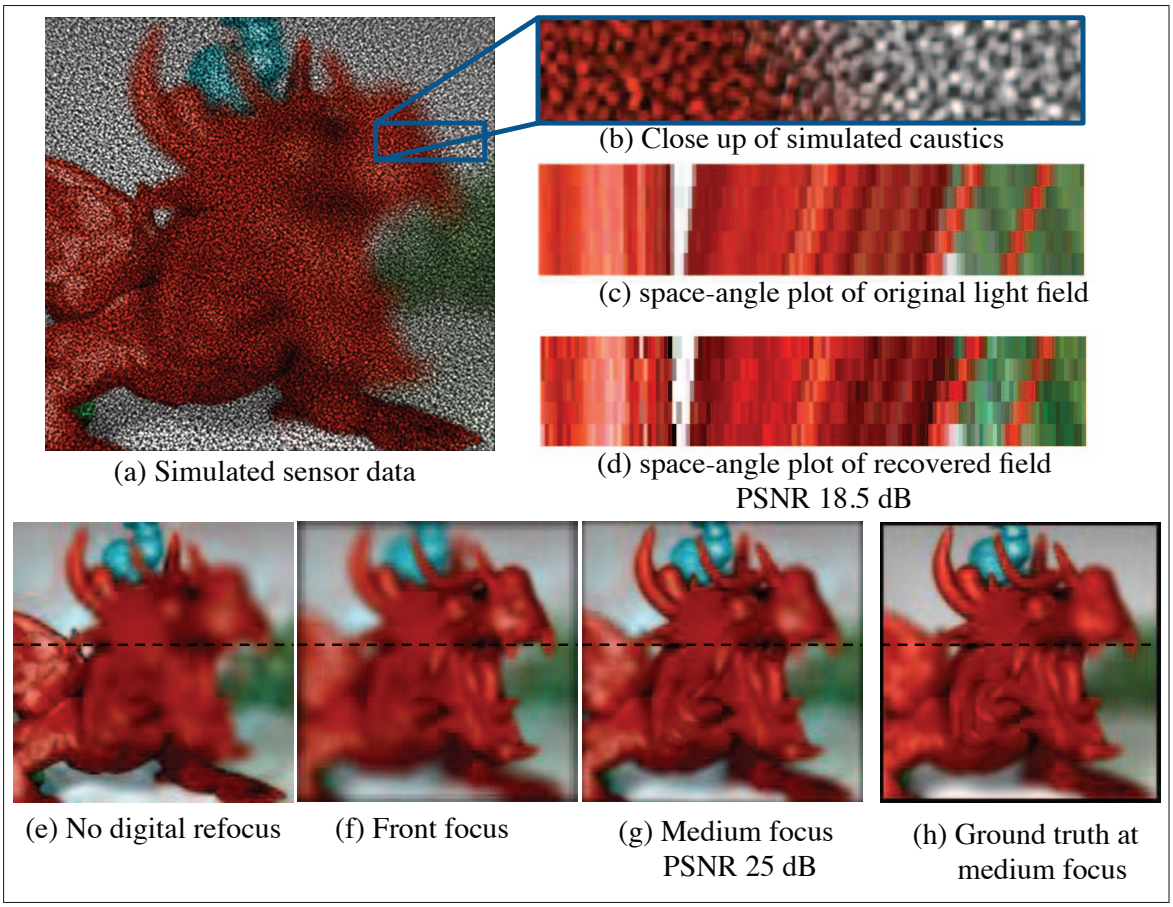

Figure 7. (a) Simulated sensor data with a zoom-in to show caustics shown in (b). We achieve good qualitative agreement between our simulated caustics and those shown in figure 8(b). (c) An (x, θ) plot from the original light field along the black line in (e)-(h), with 5% Gaussian noise added. (d) Image at same (x, θ) from our recovered light field. We are able to recover full parallax and occlusion effects. (e)-(g) Reconstructed synthetic-focus images generated from recovered light field. (e) No digital refocus, (f) refocused at the front plane, and (g) refocused on the blue bunny in the mid-focus. (h) Ground image of original light field refocused to same plane as (g).

bundle of **A**, we are unable to resolve variations in the light field that happen across spatial scales smaller than each bundle. We observe, empirically, that this leads to artifacts in the recovered sub-aperture images at very strong edges.

Lastly, our approach requires significant computational time as compared to microlens systems. Creating an entire matrix requires roughly 20 billion rays and takes 20-60 minutes to create the calibration matrix. However, once it is computed, solving the inverse problem takes 1-2 minutes.

7. Future Work

Because the angular sampling is determined by the diffuser-sensor distance, we believe it may be possible to adjust the focus distance to compensate for changes in the main lens F-number. The matrices for various focus distances could be precomputed, enabling F-number flexibility in a way that lenslets cannot accomplish.

In our prototype system, the spatial and angular sampling has been determined heuristically. The impact of the discretization in the matrix representation is still an open problem the warrants future work.

We also believe it would be possible to build a camera that operates lens-free, provided the diffuser characteristics are appropriately chosen. Finally, we plan to explore compressive sensing as applied to 4D light field data [40, 29].

8. Acknowledgments

The authors thank Lei Tian and Jingshan Zhong for help with algorithms, and Gordon Wetzstein [45] [44] for help with generating high resolution light fields.

References

[1] http://www.luminitco.com/products/light-shaping-diffusers.

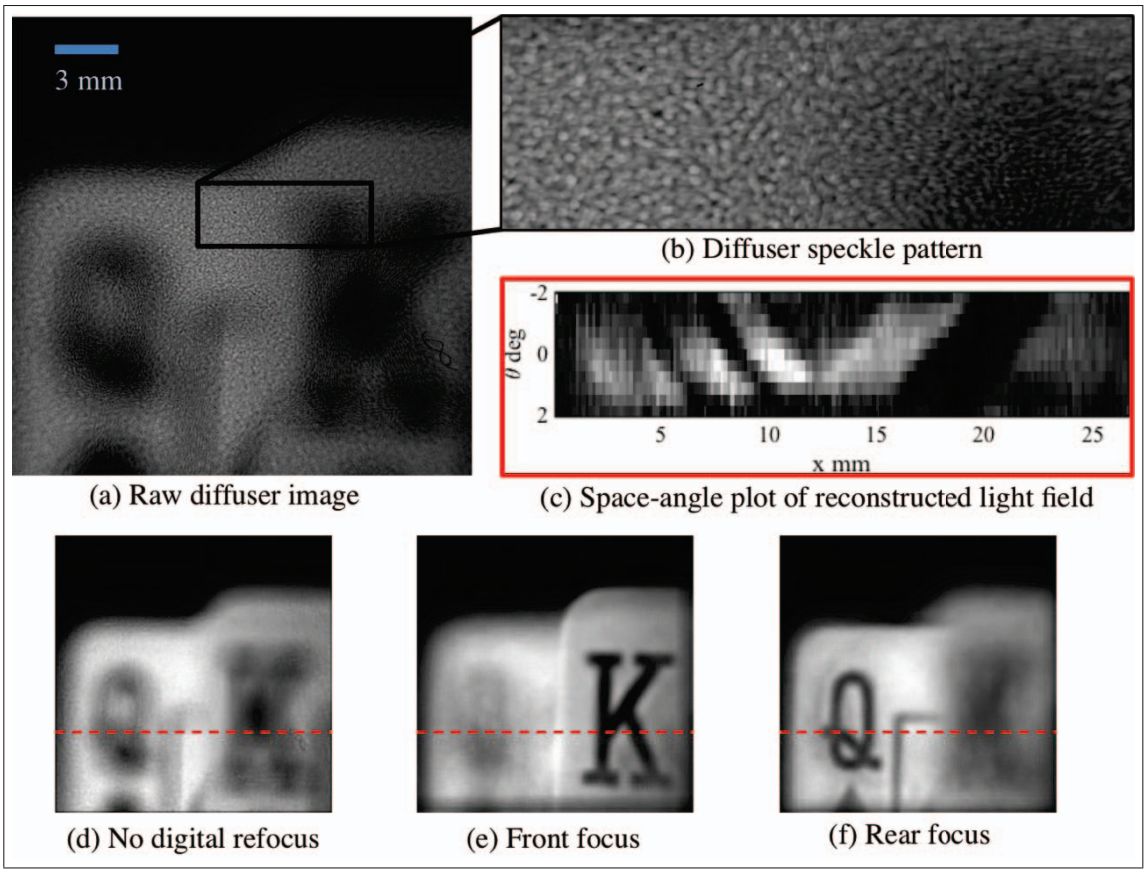

Figure 8. Experimental light field reconstruction from two playing cards using wavelet denoising. (a) Raw data, (b) close-up of diffuser caustics. (c) An x-θ plot along the red line in (d)-(f). Notice that the parallax due to the depth differences manifests as strong angular variations, and we also observe occlusion effects in the center. (d) Shows the reconstructed light field projected to $z = 0$ (no refocusing). (e) and (f) are the digitally refocused images at $+40$ mm and -40 mm, respectively.

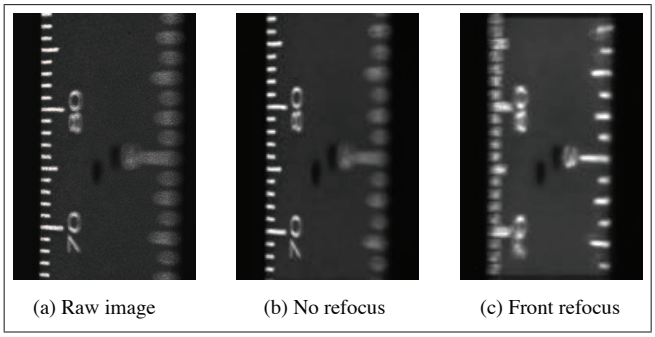

Figure 9. Experimental light field reconstruction of a ruler that is tilted relative to optical axis by approximately 30 degrees. (a) Raw data, (b) no refocus, and (c) focused to -20 mm.

[2] E. H. Adelson and J. Y. A. Wang. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 14(2):99–106, 1992.

[3] P. F. Almoro, L. Waller, M. Agour, C. Falldorf, G. Pedrini, W. Osten, and S. G. Hanson. Enhanced deterministic phase retrieval using a partially developed speckle field. *Optics Letters*, 37(11):2088–2090, Jun 2012.

[4] A. Anand, V. K. Chhaniwal, P. Almoro, G. Pedrini, and W. Osten. Shape and deformation measurements of 3d objects using volume speckle field and phase retrieval. *Optics Letters*, 34(10):1522–1524, 2009.

[5] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. Baraniuk. Flatcam: Thin, bare-sensor cameras using coded aperture and computation. *arXiv preprint arXiv:1509.00116*, 2015.

[6] J. Bioucas-Dias and M. Figueiredo. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *Image Processing, IEEE Transactions on*, 16(12):2992–3004, Dec 2007.

[7] V. Boominathan, K. Mitra, and A. Veeraraghavan. Improving resolution and depth-of-field of light field cameras using a hybrid imaging system. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–10. IEEE, 2014.

[8] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, 1999.

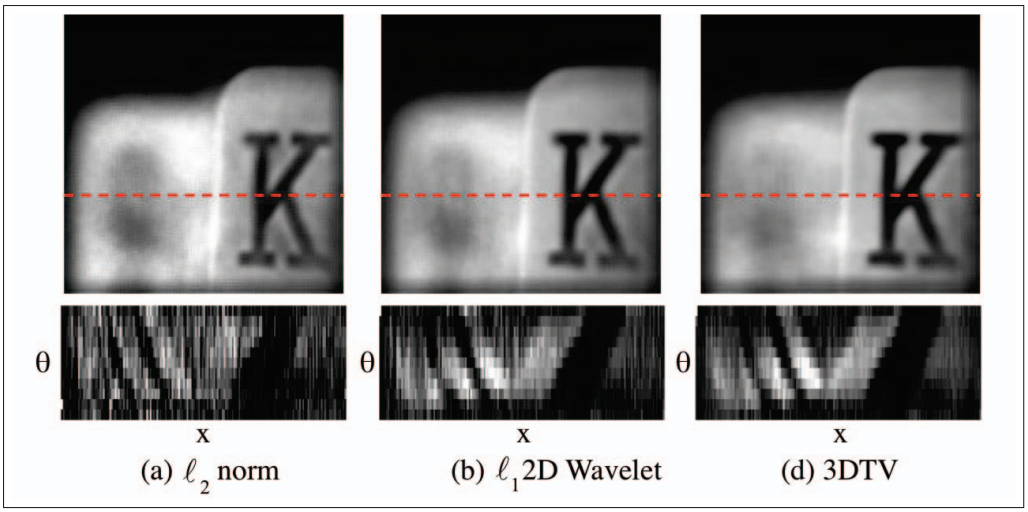

Figure 10. The effects of different regularizers on experimental reconstructions. (Top row) The reconstructed light fields refocused at the front plane. (Bottom row) (x, θ) plots along the red line. (a) ℓ_2 regularization suffers from noise artifacts, and increasing τ destroys angular information before adequately reducing noise. (b) ℓ_1 regularized 2D Wavelets is able to reduce noise significantly without destroying angular information. (c) 3DTV qualitatively performs the best in this case, due to the piecewise constant nature of this object.

[9] M. Broxton, L. Grosenick, S. Yang, N. Cohen, A. Andalman, K. Deisseroth, and M. Levoy. Wave optics theory and 3-D deconvolution for the light field microscope. *Optics Express*, 21(21):25418–25439, 2013.

[10] O. Cossairt, C. Zhou, and S. Nayar. Diffusion Coding Photography for Extended Depth of Field. *ACM Trans. on Graphics (also Proc. of ACM SIGGRAPH)*, Aug 2010.

[11] E. R. Dowski and W. T. Cathey. Extended depth of field through wave-front coding. *Applied Optics*, 34(11):1859–1866, Apr 1995.

[12] R. Fergus, A. Torralba, and W. T. Freeman. Random Lens Imaging. Technical report, Massachusetts Institute of Technology, 09 2006.

[13] J. Fiss, B. Curless, and R. Szeliski. Refocusing plenoptic images using depth-adaptive splatting. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–9. IEEE, 2014.

[14] D. C.-L. Fong and M. Saunders. LSMR: An iterative algorithm for sparse least-squares problems. *SIAM Journal on Scientific Computing*, 33(5):2950–2971, 2011.

[15] T. Georgiev, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala. Spatio-angular resolution tradeoffs in integral photography. *Rendering Techniques*, 2006:263–272, 2006.

[16] J. W. Goodman. Statistical properties of laser speckle patterns. In *Laser speckle and related phenomena*, pages 9–75. Springer, 1975.

[17] J. W. Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[18] T. Gureyev, A. Pogany, D. Paganin, and S. Wilkins. Linear algorithms for phase retrieval in the fresnel region. *Optics Communications*, 231:53 – 70, 2004.

[19] W. Harm, C. Roider, A. Jesacher, S. Bernet, and M. Ritsch-Marte. Lensless imaging through thin diffusive media. *Optics express*, 22(18):22146–22156, 2014.

[20] M. Hirsch, S. Sivaramakrishnan, S. Jayasuriya, A. Wang, A. Molnar, R. Raskar, and G. Wetzstein. A switchable light field camera architecture with angle sensitive pixels and dictionary-based sparse coding. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–10. IEEE, 2014.

[21] F. E. Ives. Parallax stereogram and process of making same., Apr. 14 1903. US Patent 725,567.

[22] Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller. Transport of intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes. *Optics Express*, 22(9):10661–10674, May 2014.

[23] M. H. Kamal, B. Heshmat, R. Raskar, P. Vandergheynst, and G. Wetzstein. Tensor low-rank and sparse light field photography. *Computer Vision and Image Understanding*, 145:172 – 181, 2016. Light Field for Computer Vision.

[24] O. Katz, P. Heidmann, M. Fink, and S. Gigan. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature Photonics*, 8(10):784–790, 2014.

[25] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz. Light Field Microscopy. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 25(3), 2006.

[26] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen. Programmable aperture photography: multiplexed light field acquisition. *ACM Transactions on Graphics (TOG)*, 27(3):55, 2008.

[27] C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. *ACM Transactions on Graphics (TOG)*, 34(2):16, 2015.

[28] G. Lippmann. La photographie intégrale. *Comptes-Rendus, Académie des Sciences*, 146:446–551, 1908.

[29] H.-Y. Liu, E. Jonas, L. Tian, J. Zhong, B. Recht, and L. Waller. 3D imaging in volumetric scattering media using phase-space measurements. *Optics Express*, 23(11):14461–14471, 2015.

[30] A. Lumsdaine and T. Georgiev. The focused plenoptic camera. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–8. IEEE, 2009.

[31] A. Maiden, G. Morrison, B. Kaulich, A. Gianoncelli, and J. Rodenburg. Soft x-ray spectromicroscopy using ptychography with randomly phased illumination. *Nature communications*, 4:1669, 2013.

[32] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive Light Field Photography using Overcomplete Dictionaries and Optimized Projections. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 32(4):1–11, 2013.

[33] T. Naemura, T. Yoshida, and H. Harashima. 3-d computer graphics based on integral photography. *Optics express*, 8(4):255–262, 2001.

[34] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Stanford University Computer Science Tech Report*, pages 3418–3421, April 2005.

[35] F. Okano, J. Arai, H. Hoshino, and I. Yuyama. Three-dimensional video system based on integral photography. *Optical Engineering*, 38(6):1072–1077, 1999.

[36] D. Stork and P. Gill. Optical, mathematical, and computational foundations of lensless ultra-miniature diffractive imagers and sensors. *International Journal on Advances in Systems and Measurements*, 7(3):4, 2014.

[37] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka. Thin observation module by bound optics (tombo): concept and experimental verification. *Applied optics*, 40(11):1806–1813, 2001.

[38] M. Teague. Deterministic phase retrieval: a Green's function solution. *Journal of the Optical Society of America*, 73(11):1434–1441, Nov 1983.

[39] L. Tian, J. C. Petruccelli, Q. Miao, H. Kudrolli, V. Nagarkar, and G. Barbastathis. Compressive x-ray phase tomography based on the transport of intensity equation. *Opt. Lett.*, 38(17):3418–3421, Sep 2013.

[40] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.*, 26(3), July 2007.

[41] K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar. Picam: An ultra-thin high performance monolithic camera array. *ACM Transactions on Graphics (TOG)*, 32(6):166, 2013.

[42] A. Wang, P. R. Gill, and A. Molnar. An angle-sensitive CMOS imager for single-sensor 3D photography. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*, pages 412–414. IEEE, 2011.

[43] L.-Y. Wei, C.-K. Liang, G. Myhre, C. Pitts, and K. Akeley. Improving light field camera sample design with irregularity and aberration. *ACM Transactions on Graphics (TOG)*, 34(4):152, 2015.

[44] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4), 2011.

[45] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):1–11, 2012.

[46] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics (TOG)*, 24(3):765–776, 2005.

[47] Z. Zhang and M. Levoy. Wigner distributions and how they relate to the light field. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–10, April 2009.