Diffusion Coded Photography for Extended Depth of Field

Oliver Cossairt[*]
Columbia University

Changyin Zhou[†]
Columbia University

Shree Nayar[‡]
Columbia University

(a) An image captured with a F/1.8 lens   (b) Detail from (a)   (c) Detail from (d)   (d) A diffusion coded image (deblurred)
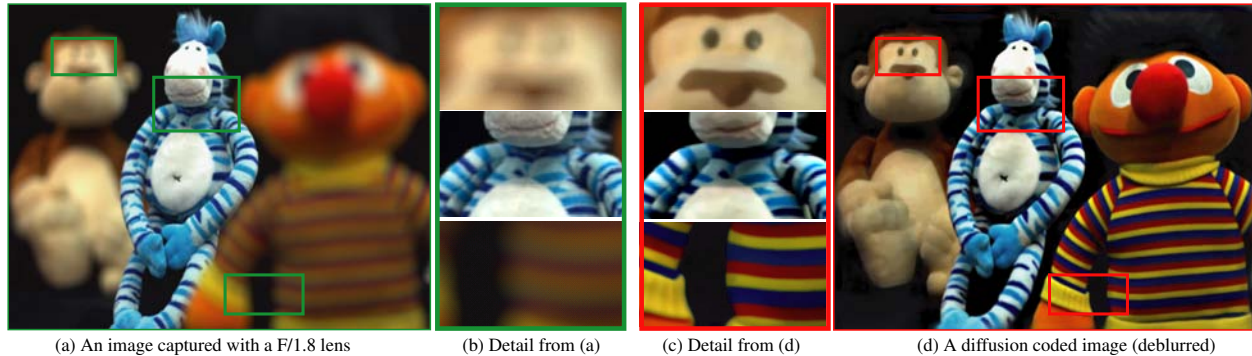
Figure 1: *Extending depth of field with diffusion coding for a scene consisting of three stuffed animals placed at different depths. (a) An image captured with a 50mm F/1.8 Canon Lens. The foreground and background objects exhibit severe defocus blur. (d) The diffusion coded image after deblurring. The image was captured with the diffuser from Section 6 placed in the lens aperture. (b-c) Magnified regions from (a) and (d) that show that diffusion coding preserves details in foreground and background objects.*

Abstract

In recent years, several cameras have been introduced which extend depth of field (DOF) by producing a depth-invariant point spread function (PSF). These cameras extend DOF by deblurring a captured image with a single spatially-invariant PSF. For these cameras, the quality of recovered images depends both on the magnitude of the PSF spectrum (MTF) of the camera, and the similarity between PSFs at different depths. While researchers have compared the MTFs of different extended DOF cameras, relatively little attention has been paid to evaluating their depth invariances. In this paper, we compare the depth invariance of several cameras, and introduce a new camera that improves in this regard over existing designs, while still maintaining a good MTF.

Our technique utilizes a novel optical element placed in the pupil plane of an imaging system. Whereas previous approaches use optical elements characterized by their amplitude or phase profile, our approach utilizes one whose behavior is characterized by its scattering properties. Such an element is commonly referred to as an optical diffuser, and thus we refer to our new approach as *diffusion coding*. We show that diffusion coding can be analyzed in a simple and intuitive way by modeling the effect of a diffuser as a kernel in light field space. We provide detailed analysis of diffusion coded cameras and show results from an implementation using a custom designed diffuser.

CR Categories: I.4.3 [Image Processing and Computer Vision]: Enhancement—Sharpening and Deblurring;

Keywords: computational photography, extended depth of field

1 Introduction

For conventional cameras, there is a fundamental trade-off between depth of field (DOF) and noise. Cameras have a single focal plane, and objects that deviate from this plane are blurred due to defocus (see Figure 1(a)). The amount of defocus blur depends on the aperture size and the distance from the focal plane. To decrease defocus blur and increase DOF, the aperture size must be decreased, reducing the signal strength of the recorded image as well.

In many cases, it is desirable to have a DOF that is as large as possible so that all details in the scene are preserved. This is the case, for instance, in machine vision applications such as object detection and recognition, where it is desirable that all objects of interest be in focus. However, stopping down the lens aperture is not always an option, especially in low light conditions, because it increases noise and corrupts the signal.

To obtain an extended DOF (EDOF) image without stopping down the aperture, there are two strategies that can be employed: (1) Measure the depths of objects in the scene, use this to calculate depth-dependent PSFs, and apply spatially varying deconvolution to the captured image. (2) Engineer a PSF that is depth-invariant so that a single PSF can be used for deblurring. It is well known that high precision depth estimation is error prone, and difficult (if not impossible) without the aid of additional hardware, such as that used in structured light or laser scanning systems. For this reason, we focus our attention in this paper on extending DOF by engineering a depth-invariant PSF.

Typically, when deblurring a noisy image, a larger magnitude in the PSF spectrum (MTF) will result in less deblurring reconstruction error. However, this is only the case if the correct PSF is used to deblur the image. If a different PSF is used for deblurring, it is

[*]e-mail: ollie@cs.columbia.edu
[†]e-mail: changyin@cs.columbia.edu
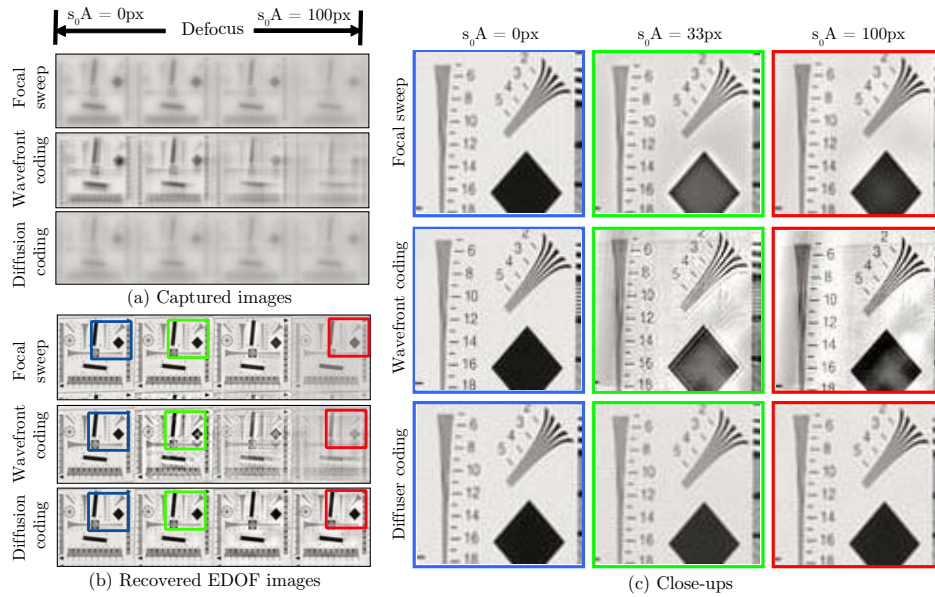[‡]e-mail: nayar@cs.columbia.edu

Figure 2: *Simulated image performance for three EDOF cameras. An IEEE resolution chart is placed at different depths. The aperture size A and defocus slope in light field space s_0 are chosen so that the maximum defocus blur diameter is 100 pixels. The center PSF is used for deblurring, producing the images shown in (b). Close-ups in (c) show that the sharpest image is produced by wavefront coding at the center depth ($s_0 A = 0$). However, wavefront coding produces significant deblurring artifacts for defocus values as small as $s_0 A = 33$ pixels, while diffusion coding produces near identical results for the entire depth range.*

also important to consider the similarity between the original PSF and the one used for deblurring. This consideration is of utmost importance in the context of EDOF cameras because, in practice. it is only possible to produce a PSF that is *approximately* depth-invariant, and the amount of variation determines the severity of the artifacts that are introduced in the deblurring process.

Two well-studied techniques that produce a depth-invariant PSF are wavefront coding [Dowski and Cathey 1995], which uses a cubic phase plate, and focal sweep [Nagahara et al. 2008] [Häusler 1972], where object, sensor position, or lens focus setting is mechanically varied during exposure. Recently, Baek compared the degree of depth-invariance of these two techniques, and observed that focal sweep gives a near-optimal tradeoff between MTF and depth-invariance at all frequencies [Baek 2010], while wavefront coding is only guaranteed to be optimal at a single frequency. We introduce a new diffusion coding camera that produces near identical performance to focal sweep, but without the need for moving

parts (see Figures 1,15-17).

In this work, we focus our attention on the use of optical diffusers whose behavior is predominately defined by scattering properties alone. Like phase-plates, these surfaces have the advantage of being almost completely non-absorptive, and thus do not sacrifice signal intensity. We coin the term *diffusion coding* to mean a camera with a diffuser placed in the pupil plane. Diffusers are commonly implemented as surfaces with random thickness variations. Such a surface can either be analyzed as a phase-plate using Fourier optics, or using geometric optics. This is discussed further in Sections 4 and 6, and in both cases the statistics of the surface variations are used to calculate the expected distribution of light scattered by the diffuser. While the former is more common, it also results in much more complicated analysis. In this paper, we focus on the latter, which allows us to characterize diffusers as kernels that operate on a 4D light field propagating from a camera lens to sensor. As a result, we are able to obtain an analytical solution for the PSF of our diffusion coded camera, which is given in Section 4.
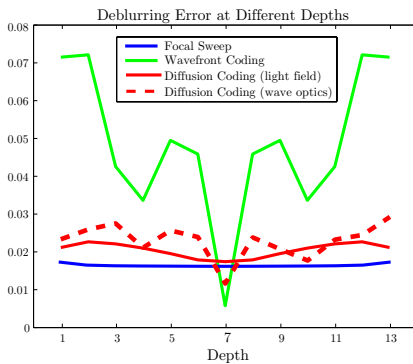
Levin et al. show that wavefront coding produces better results than focal sweep if variation in the PSF is not taken into account [Levin et al. 2009]. As can be seen from Figure 2, wavefront coding recovers more detail than other methods for objects at the focal plane when the correct PSF is used for deblurring. However, the method also introduces noticeable artifacts for objects at different depths because the PSF varies significantly with depth. To measure the degree of depth-invariance of a camera, we compute the deblurring reconstruction error for objects at different depths. The result is shown in Figure 3, where a flatter curve signifies more similarity between PSFs at different depths. We note that the focal sweep camera produces a PSF that is more depth-invariant than wavefront coding, and furthermore that our diffusion coded camera produces near identical results to that of focal sweep. The comparison of EDOF Cameras is discussed further in Section 5.

We focus our attention on the use of diffusers with predefined scattering properties, and do not address the task of diffuser design. Much work has been done in recent years to develop cus-
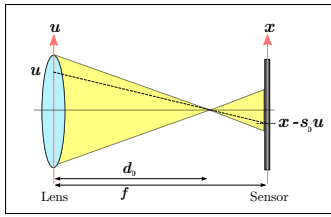

Figure 3: *The deblurring error (based on simulations in Section 5) as a function of depth for three EDOF cameras. A flatter curve denotes less PSF variation. The diffusion coding curves are very similar to that of focal sweep.*

Figure 4: *The geometry of an image point focused at a distance d_0 from the camera lens aperture. A sensor is located a distance f from the aperture. A ray at piercing the aperture at location u intersects the sensor at location $x - s_0 u$. where $s_0 = \frac{d_0 - f}{f}$.*

Figure 5: *For the diffuser defined by the kernel in Equation 7, the diffusion angle does not vary across the aperture. Each ray is blurred so that it covers an area on the sensor determined by the the diffuser parameter w.*

tom diffusers with tailored scattering profiles. These diffusers are frequently used in lighting and display applications to produce uniform illumination or arbitrary beam shaping. The popularity of these diffusers has also led to much innovation in replication techniques, so that today several companies sell off-the-shelf diffusers reproduced onto plastic sheets up to 36" wide [Luminit][RPC][Wavefront]. In Section 6, we introduce our implementation of a diffusion coded camera using a custom diffuser manufactured by RPC Photonics [RPC]. We conclude with examples of EDOF images taken with our implementation in Section 7.

2   Previous Work

Optical diffusers and other random surfaces have been used to assist in a variety of imaging tasks, including super-resolution [Ashok and Neifeld 2003][Ashok and Neifeld 2007], lenseless imaging [Freeman et al. 2006], and extended DOF [García-Guerrero et al. 2007]. In this work, we focus on the task of using diffusers to extend DOF.

Several radially symmetric phase masks have been introduced to extend DOF [Chi and George 2001] [Ojeda-Castaneda et al. 2005][García-Guerrero et al. 2007]. The work most similar to ours is by Garcia-Guerrero et al., who also use a radially symmetric diffuser. To design their diffuser, the authors take a completely different approach than the one used in this paper. They derive a random surface that on average produces a PSF whose value at the center is constant over a large depth range, while we derive a diffuser whose entire PSF is approximately depth-invariant. The Garcia-Guerrero diffuser consists of annular sections of quadratic surfaces, where the width of the annulus decreases quadratically with distance from the optical axis. This design requires the feature size to decrease from the center to the edge of the diffuser. The minimum feature size is limited by the fabrication technology that is used to make the diffuser. In this paper we consider the use of laser machining technology that has a minimum spot size on the order of 10μm. The result is that the performance of one instance of the Garcia-Guerrero diffuser varies significantly from the expected performance while the diffuser we introduce in Section 6 performs very close to the expected performance (see Figure 12). This difference is discussed further in Section 6.

Wavefront coding was introduced by Dowski and Cathey [Dowski and Cathey 1995], who place a cubic phase plate (CPP) in the pupil plane of a camera system. Dowski et al. show analytically that a camera with a cubic phase plate produces a PSF that is approximately invariant to defocus. Although the CPP does produce a PSF that is approximately depth-invariant, the PSF is not as invariant as the focal sweep camera or our diffusion coded camera (see Figures 2 and 3).

Focal sweep cameras produce a depth-invariant PSF by sweeping either the object [Häusler 1972] or sensor [Nagahara et al. 2008] along the optical axis during exposure. The PSFs for these tech-
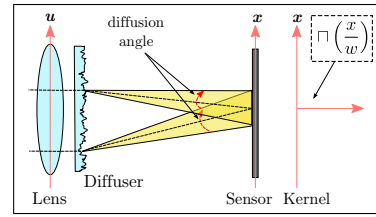
niques preserves high frequencies because each object is instantaneously in focus at one point during exposure. Focal sweep techniques require the use of moving parts and introduce limitations on the minimum exposure time.

Levin et al. compare the performance of focal sweep and wavefront coding cameras without considering the effect of depth-invariance [Levin et al. 2009]. Hasinoff et al. analyzed the SNR characteristics of both focal sweep and wavefront coding cameras when multiple exposures with different focus settings are used [Hasinoff et al. 2009], and Baek compared the MTF and depth-invariance of focal sweep and wavefront coding cameras [Baek 2010].

Other works exist in the vision community which recover an EDOF image after first estimating scene depth [Levin et al. 2007] [Levin et al. 2009] [Zhou and Nayar 2009]. The quality of these techniques, however, is closely coupled to the precision of depth estimation, since each region in the image is deblurred using an estimated defocus PSF.

In this paper, we analyze the light field [Levoy and Hanrahan 1996] to understand the properties of imaging systems. Several researchers have analyzed the image formation of camera systems as projections of light fields [Ng 2005][Veeraraghavan et al. 2007][Levin et al. 2009]. In addition, several authors have looked at light fields in the frequency domain, including image formation and interactions between transmissive and reflective objects [Ng 2005][Durand et al. 2005][Veeraraghavan et al. 2007].

3   Light Field Analysis

A light field $L(\bar{u}, \bar{x})$ can be used to represent the 4D set of rays propagating from an ideal lens with effective focal length (EFL) f to a sensor. The vector $\bar{u} = (u, v)$ denotes the coordinates on the u-v plane, which is coincident with the exit pupil of the lens. The vector $\bar{x} = (x, y)$ denotes the coordinates on the x-y plane that is coincident with the sensor. Note that this is a slightly different convention than used by Levin et al., where the x-y plane is defined in object space [Levin et al. 2009]. The irradiance $E(\bar{x})$ observed on the sensor is simply the light field integrated over all ray angles:

$$E(\bar{x}) = \int_{\Omega_{\bar{u}}} L(\bar{u}, \bar{x}) d\bar{u}, \tag{1}$$

where, $\Omega_{\bar{u}}$ is the domain of $\bar{u}$. For a scene with smooth depth variation, locally, the captured image $E(\bar{x})$ can be modeled as a convolution between a depth-dependent PSF kernel $P(\bar{x})$ and an all-in-focus image $I(\bar{x})$. The goal of this paper is to shape the camera PSF so that the *entire* image $I(\bar{x})$ can be recovered from the captured image $E(\bar{x})$ by deblurring with a *single* PSF $P(\bar{x})$. We
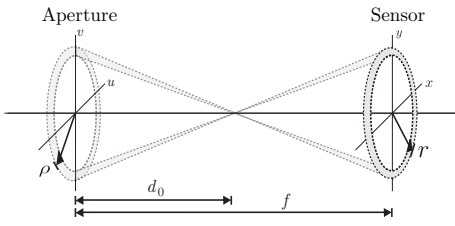
Figure 6: *The geometry of a radially symmetric light field using reduced coordinates. The light field consists of a point source focused a distance d_0 from the lens aperture. Because the point source is on-axis and isotropic, the light field can be represented as a 2D function $L(\rho, r)$. A 2D slice of the light field $L(\rho, r)$ represents the set of rays traveling from a circle with radius ρ in the aperture plane to a circle with radius r on the sensor. This set of rays forms a conic surface.*

analyze the depth-dependence of the camera PSF by considering the image produced by a unit energy point source. Consider a point source whose image comes to focus at a distance d_0 from the aperture of the lens (see Figure 4). Assuming a rectangular aperture of width A, the light field produced by this point is

$$L_\delta(\bar{u}, \bar{x}) = \frac{1}{A^2} \sqcap \left(\frac{\bar{u}}{A} \right) \delta(\bar{x} - s_0 \bar{u}), \qquad (2)$$

where $s_0 = \frac{d_0 - f}{d}$ is the defocus slope in light field space, and $\sqcap$ is the box function

$$\sqcap \left(\frac{\bar{x}}{w} \right) = \begin{cases} 1 & \text{if } |\bar{x}_i| < \frac{w}{2}, \ \forall i \\ 0 & \text{otherwise} \end{cases} . \qquad (3)$$

The image of this point is the camera PSF at the depth d_0, which is the familiar box shaped PSF with defocus blur width $s_0 A$:

$$P(x) = \frac{1}{s_0^2 A^2} \sqcap \left(\frac{\bar{x}}{s_0 A} \right). \qquad (4)$$

We now analyze the effect of a general kernel D applied to a light field L, which represents the effect of a diffuser placed in the aperture of a camera lens. The kernel produces a new filtered light field $\hat{L}$, from which we can derive the modified PSF $\hat{P}$:

$$\hat{L}(\bar{u}, \bar{x}) = \int_{\Omega_{\bar{u}'}} \int_{\Omega_{\bar{x}'}} D(\bar{u}, \bar{u}', \bar{x}, \bar{x}') L(\bar{u}', \bar{x}') d\bar{u}' d\bar{x}', \qquad (5)$$

$$\hat{P}(\bar{x}) = \int_{\Omega_{\bar{u}}} \hat{L}(\bar{u}, \bar{x}) d\bar{u}, \qquad (6)$$

where $\Omega_{\bar{x}}$ is the domain of $\bar{x}$. This approach allows us to express a large class of operations applied to a light field. For instance, consider a kernel of the form

$$D(\bar{u}, \bar{u}', \bar{x}, \bar{x}') = \frac{1}{w^2} \delta(\bar{u} - \bar{u}') \sqcap \left(\frac{\bar{x} - \bar{x}'}{w} \right). \qquad (7)$$
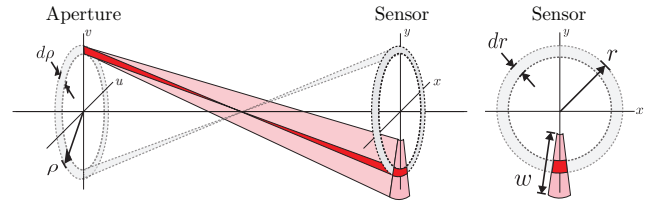
Figure 7: *The geometry of a radially symmetric diffuser. The diffuser scatters light only in the radial direction, and has no effect in the tangential direction. A thin annulus of light is emitted from the aperture of width $d\rho$ and radius ρ. In the absence of the diffuser, the emitted light projects to an annulus on the sensor of width dr and radius r. When the diffuser is present, the width of the annulus on the sensor becomes w, the diffuser scatter width.*

Note that here D takes the form of a separable convolution kernel with finite support in the $\bar{x}$ domain. The geometric meaning of this kernel is illustrated in Figure 5. Each ray in the light field is blurred so that, instead of piercing the sensor at a single location, it contributes to a square of width w. In order to understand the effect of the diffuser, we compare an image E captured without the diffuser to an image $\hat{E}$ captured with it. For this diffuser kernel, substituting Equation 7 into Equations 5 and 6 gives (see Supplementary Material for a detailed derivation):

$$\hat{P}(\bar{x}) = \frac{1}{w^2} \sqcap \left(\frac{\bar{x}}{w} \right) \otimes P(\bar{x}), \qquad (8)$$

where $\otimes$ denotes convolution. The modified PSF is simply the camera PSF blurred with a box function. Therefore, the effect of the diffuser is to blur the image that would be captured were it not present. Introducing the diffuser given by the kernel in Equation 7 is clearly not useful for extending depth of field since it it does not increase depth independence or preserve high frequencies in the camera PSF. We note that, in general, the kernel for any diffuser that is placed in the aperture takes the form

$$D(\bar{u}, \bar{u}', \bar{x}, \bar{x}') = \delta(\bar{u} - \bar{u}') f(\bar{u}, \bar{x} - \bar{x}'), \qquad (9)$$

where f is called the scatter function. That is, the diffuser has no effect in the $\bar{u}$ domain, but has the effect of a convolution in the $\bar{x}$ domain. For the diffuser given by Equation 7, the scatter function is the 2D box function $f(\bar{u}, \bar{x}) = \frac{1}{w^2} \sqcap \left(\frac{\bar{x}}{w} \right)$.

4 Radially Symmetric Light Fields

We now change from rectangular coordinates (u, v, x, y) to polar coordinates (ρ, ϕ, r, θ) using the relations $u = \rho cos\phi$, $v = \rho sin\phi$, $x = rcos\theta$, and $y = rsin\theta$. We consider a polar system where $\rho, r \in (-\infty, \infty)$ and $\theta, \phi \in (0, \pi)$ and a circular aperture with diameter A. The light field representing a unit-energy point source located at distance d_0 in this new system can be written as

$$L_\delta(\rho, r) = \frac{4}{\pi A^2} \sqcap \left(\frac{\rho}{A} \right) \frac{\delta(r - s_0 \rho)}{\pi |r|}, \qquad (10)$$

which is independent of both θ and ϕ because the source is isotropic. Note that verifying unit-energy can be carried out trivially by integrating $L_\delta(\rho, r)$ in polar coordinates (see Supplementary Material). Comparing the parameterizations for the light field

of a point source in Equations 2 and 10, we can see that a slice of $L_\delta(\bar{x}, \bar{u})$ represents a single ray, while a slice $L(\rho, r)$ represents a 2D set of rays. In the radially symmetric parameterization, a slice of the light field represents a conic surface connecting a circle with radius ρ in the aperture plane to a circle of radius r on the sensor (see Figure 6).

We now consider the effect of a radially symmetric diffuser on the camera PSF. Somewhat surprisingly, a diffuser that is parameterized in these reduced 2D coordinates produces a drastically different effect than the diffuser given by Equation 7. When a radially symmetric diffuser is introduced, neither the diffuser nor the lens deflects rays tangentially, and therefore we can represent the diffuser kernel and modified light field using the reduced coordinates (ρ, r). Equations 5 and 6 then become

$$\hat{L}(\rho, r) = \pi^2 \int_{\Omega_\rho} \int_{\Omega_r} D(\rho, \rho', r, r') L(\rho', r) |\rho'| d\rho' |r'| dr', \quad (11)$$

$$E(r) = \pi \int_{\Omega_\rho} \hat{L}(\rho, r) |\rho| d\rho, \quad (12)$$

and the general form of the diffuser kernel becomes

$$D(\rho, \rho', r, r') = \frac{\delta(\rho - \rho')}{\pi |\rho'|} \frac{f(r - r', \rho)}{\pi |r|}. \quad (13)$$

We use the same box-shaped scattering function as we did for the diffuser kernel in Equation 7:

$$f(r, \rho) = \frac{1}{w} \sqcap \left(\frac{r}{w}\right). \quad (14)$$

However, the physical interpretation of this diffuser is drastically different than for the previous diffuser. For the previous one, each ray in the light field is scattered so that it spreads across a square on the sensor. The effect of the scattering function in Equation 14 is illustrated in Figure 7. In the absence of the diffuser, light from an annulus of width $d\rho$ and radius ρ in the aperture plane projects to an annulus of width dr and radius r on the sensor. The effect of the scatter function in Equation 14 is to spread the light incident on the sensor so that it produces an annulus of width w instead. We can also consider the scattering from the perspective of a single ray, as illustrated by the pink and red volumes in Figure 7. In polar coordinates, a ray is a small annular section that travels from the aperture plane to the sensor plane, illustrated by the red volume in Figure 7. The pink volume illustrates the effect of the diffuser, which is to scatter a ray along a radial line of width w. We note that a box-shaped scatter function is used here for notational convenience, but we found that a Gaussian scattering function is superior for extended DOF imaging (see Figure 11(d)).

The light field of a point source filtered by this diffuser kernel and PSF can be shown to be (see Supplementary Material for a complete derivation)

$$\hat{L}(\rho, r) = \frac{4}{\pi A^2} \sqcap \left(\frac{\rho}{A}\right) \frac{\sqcap\left(\frac{r - s_0 \rho}{w}\right)}{\pi w |r|}, \quad (15)$$

$$\hat{P}(r) = \frac{4}{\pi s_0^2 A^2} \frac{1}{w |r|} \left[\sqcap\left(\frac{r}{w}\right) \otimes \left(\sqcap\left(\frac{r}{s_0 A}\right) \cdot |r| \right) \right]. \quad (16)$$

The analytic solution for the PSF is a piecewise function due to the contribution from the term in brackets, which is a convolution between the two rect functions (one weighted by $|r|$). Note that as the scattering width w is reduced to zero, the first rect (combined with $\frac{1}{w}$) approaches a delta function and the result is the familiar pillbox shaped defocus PSF. Also note that if a different scattering function is used, the first rect is simply replaced with the new function. However, the convolution term is far less significant than the $\frac{1}{|r|}$ term, whose effect dominates, resulting in a PSF which is strongly depth-independent while still maintaining a strong peak and preserving high frequencies.

The solution for the PSF may be interpreted in the following way. Please refer to Figure 7. Suppose we have a pillbox defocus PSF, and we want to know how a small annular region of width δr and radius r will be affected by the diffuser. Light incident on this region emanates from an annulus in the aperture, and its energy will be proportional to ρ or equivalently r/s_0. This explains the presence of the $|r|$ multiplier within the term in brackets. The term in brackets states that the energy in the PSF annulus is spread uniformly along radial lines of width w, as shown on the right hand side of Figure 7. The $\frac{1}{|r|}$ term in Equation 16 can be attributed to the fact that the energy density becomes larger for light that is scattered closer to the center of the PSF.

Figure 9 shows several PSF/MTF pairs for a camera with and without the diffuser given by Equation 16. The defocus blur diameter $s_0 A$ varies from 0 to 100 pixels. The scatter function of Equation 14 is a Gaussian instead of a box function, and the diffuser parameter w (the variance of the gaussian) is chosen so that $w = 100$ pixels. Note that when the diffuser is present, there is little variation with depth for either the PSF or MTF. Introducing the diffuser also eliminates the zero crossings in the MTF. For smaller defocus values, the diffuser suppresses high frequencies in the MTF. However, because the diffuser MTF does not vary significantly with depth, high frequencies can be recovered via deconvolution. Figure 8 shows a simulated light field filtered by the radially symmetric diffuser given by Equation 14. On the far right of the figure, we show a high contrast, extended depth of field image that is recovered after deconvolution is applied.
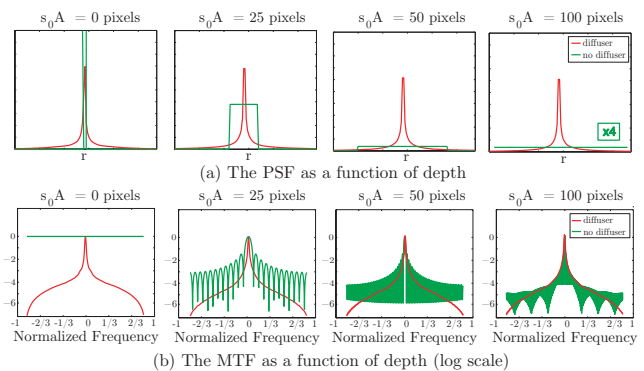

(a) The PSF as a function of depth

(b) The MTF as a function of depth (log scale)

Figure 9: *PSF plots (top) and MTF (bottom) plots for a camera with (red) and without (green) the diffuser kernel defined in Equation 14. The defocus blur diameter $s_0 A$ is varied across columns from 0 to 100 pixels, and the diffuser parameter $w = 100$ pixels. Both the PSF and MTF exhibit negligible variation when the diffuser is present.*
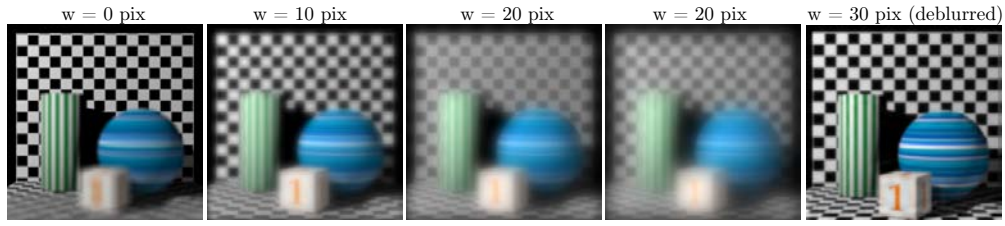
Figure 8: *Simulated photographs taken of of a light field filtered by the diffuser kernel in Equation 14. The parameter w of the diffuser kernel is varied across the columns. The rightmost figure shows a deblurred diffusion coded image with a $10\times$ increase in DOF.*

5   Comparison between EDOF Cameras

All EDOF cameras sacrifice MTF response at high frequencies in order to achieve depth-invariance. High frequencies in captured images are recovered via deconvolution, but this process also amplifies sensor noise which degrades the recovered image. In addition, any variation in the PSF/MTF as a function of depth will result in deblurring artifacts due to a mismatch between the actual PSF and the PSF used for deblurring. The quality of an edof camera can be represented by the deblurring reconstruction error R, which takes into account the camera MTF, the degree of depth-invariance of the PSF/MTF, and sensor noise. To calculate the deblurring error R we compute the $L2$ norm on the difference between the ground truth (focused) image and the captured image deblurred by a PSF P_d. The captured image is the ground truth image I blurred by a PSF P_b plus noise η.

$$R = \left|\left|(I \otimes P_b + \eta) \otimes P_d^{-1} - I\right|\right|^2. \qquad (17)$$

This measure takes into account the camera MTF, since it includes the term $\eta \otimes P_d^{-1}$, which represents the amplification of sensor noise due to small MTF values. In addition, the measure takes into account the degree of depth-invariance of the camera PSF/MTF because it includes the term $I - (I \otimes P_b) \otimes P_d^{-1}$, which is the difference between a ground truth image and the same image blurred by one PSF and then deblurred by another.

To evaluate the performance of an EDOF camera, we calculate the deblurring error over a range of depths. If an EDOF camera performs well, it will have a small deblurring error over all depths. For each camera, we calculated the camera PSF at a variety of discrete depths and used this as the blurring PSF P_b. For the deblurring PSF P_d, we used the camera PSF at the center of the depth range. In all simulations, η was set to be Gaussian white noise with variance $\sigma = .005$. Since the deblurring error can vary with I, we compute the value over a variety of natural images and take the average. In Figure 3, we show the deblurring error for three EDOF methods. Wavefront coding achieves the minimum deblurring error for all cameras when the defocus blur diameter $s_0 A = 0$ pixels. This is because the wavefront coding MTF is greater and therefore preserves more information when deblurred with the correct PSF. However, both diffusion coding and focus sweep produce a flatter curve that results in less deblurring error at all other depth locations.

To demonstrate the performance of our EDOF method, we simulated a scene consisting of an IEEE resolution chart. Simulated defocused images are shown in Figure 2(a), where the maximum defocus blur diameter is $s_0 A = 100$ pixels. We apply Wiener deconvolution with the PSF at the center depth to obtain the EDOF images shown in (b). Close-ups of the deblurring results are shown in (c). As expected, the sharpest image is produced by wavefront coding for the center depth. However, wavefront coding produces significant deblurring artifacts for defocus values as small as $s_0 A = 33$

pixels, while diffusion coding produces near identical results for the entire depth range.

To generate the PSFs for Figures 2 and 3, we used the analytical solution for the diffusion coding PSF from Equation 16. For the focal sweep camera, we numerically integrated a sequence of defocus discs which, for the center PSF, represents a range of defocus blur diameters from 0 to 120 pixels. We performed a numerical search to find the focal sweep range that produces a local minimum in average deblurring error for this simulation. We used the raytracing engine in Zemax to numerically compute the wavefront coding PSFs without the effect of diffraction. To generate the Zemax raytrace, a cubic refractive surface was used such that the light field integration curve takes the form $(x = au^2, y = av^2)$. The optimal value for a was chosen to be $a = S/(2A)$ [Levin et al. 2009], where S is the maximum value of the defocus parameter s_0. Furthermore, we performed a numerical search to verify that this a produces a local minimum in average deblurring error for this simulation.

6   Implementing the Diffuser

We consider diffusers of the "kinoform" type [Caulfield 1971], where the scattering effect is caused entirely by roughness variations across a surface. Such a diffuser can be considered a random phase screen, and according to statistical optics, for a camera with effective focal length f, and center wavelength $\bar\lambda$, the effect of placing this screen in the aperture of the camera results in the following
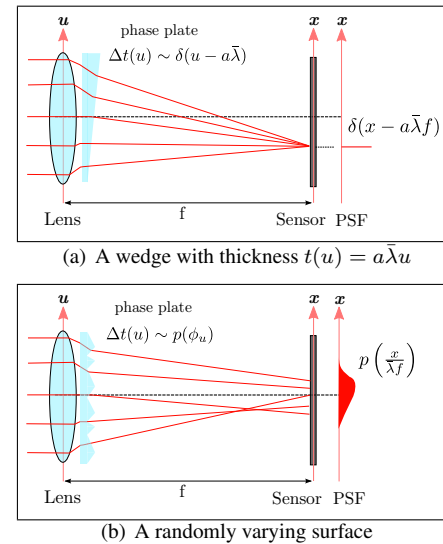

(a) A wedge with thickness $t(u) = a\bar\lambda u$

(b) A randomly varying surface

Figure 10: *A wedge can be thought of as a having a slope drawn from a probability density function which is a delta function. A diffuser can be thought of as a phase plate with a randomly varying thickness with a slope that is drawn from a more general probability density function.*
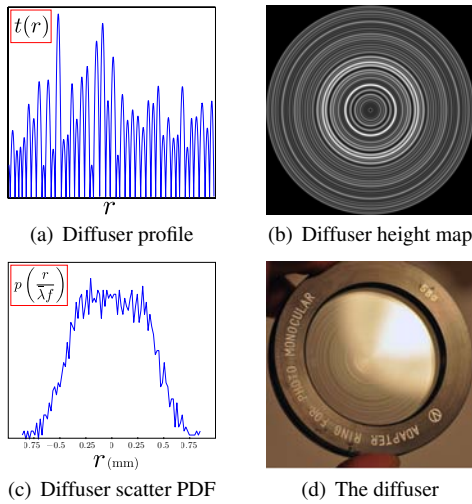
(a) Diffuser profile

(b) Diffuser height map

(c) Diffuser scatter PDF

(d) The diffuser

Figure 11: *An implementation of the diffuser defined by the kernel in Equation 14. (a), (b), and (c) show the radial profile, height-map, and radial scatter function of the diffuser surface, respectively. (d) shows the fabricated diffuser.*

PSF [Goodman 1985]:

$$\hat{P}(x,y) \propto p_{\phi_u, \phi_v}\left(\frac{x}{\bar{\lambda} f}, \frac{y}{\bar{\lambda} f}\right), \qquad (18)$$

where ϕ_u and ϕ_v are the u and v derivatives of the phase shift induced by the surface, and p_{ϕ_x, ϕ_y} is the joint probability of these derivatives. The result of Equation 18 is that we can implement a diffuser simply by creating an optical element with thickness $t(u,v)$, where the gradient of this surface $\bigtriangledown t(u,v)$ is sampled from a probability distribution which is also our desired PSF. Intuitively, we can understand this equation as follows: p_{ϕ_u, ϕ_v} denotes the fraction of the surface $t(u,v)$ with slope (ϕ_u, ϕ_v). For small angles, all incoming rays incident on this fraction of the surface will be deflected at the same angle, since the slope is constant over this region. Thus the quantity p_{ϕ_u, ϕ_v} also reflects the portion of light that will be deflected by the slope (ϕ_x, ϕ_y).

In fact, kinoform diffusers can be thought of as generalized phase
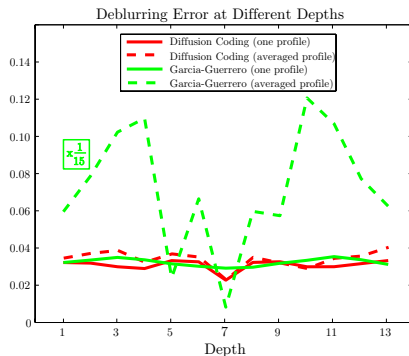

Figure 12: *The deblurring error as a function of depth for both diffusion coding and the Garcia-Guerrero diffuser. The dotted lines show the deblurring error for a single instance of the diffuser surface. The solid lines show the deblurring error averaged over 100 realizations of the diffuser surfaces. A single instance of the diffusion coding surface performs significantly better than the Garcia-Guerrero diffuser.*

plates, as shown in Figure 10. In Figure 10(a), a wedge with thickness $t(u) = a\bar{\lambda} u$ is placed in the aperture of a lens system. The effect of the wedge is to shift the PSF away from the optical axis. The wedge can be thought of as a having a slope drawn from a probability function $p(\phi_u)$ which is a delta function. The result of placing a wedge in the pupil plane of a camera is to shift the PSF, which can be thought of as convolving $p(\phi_u)$ with the PSF. A kinoform diffuser has a randomly varying surface with a more general probability distribution of slopes (Figure 10(b)).

To implement the diffuser defined in Equation 14, we follow the procedure in [Sales 2003], which simply implements a diffuser surface as a sequence of quadratic elements whose diameter and sag is drawn from a random distribution. The scatter function is designed to be roughly Gaussian with 0.5mm variance (corresponding to $w = 1$mm in Equation 16) as shown in Figure 11(c). To create a radially symmetric diffuser, we create a 1D random profile and then apply a polar transformation to create the final 2D surface (see Figures 11(a) and 11(b)). The maximum height of the surface is 3μm. The diffuser was fabricated using a laser machining technology which has a minimum spot size of about 10μm. To ensure that each quadratic element was fabricated with high accuracy, the minimum diameter of a single element was chosen to be 200μm, resulting in a diffuser with 42 different annular sections. The diffuser used in all our experiments (see Figures 1,15-17) is shown in Figure 11(d), and was fabricated by RPC Photonics [RPC].

To compare the performance of our diffuser surface relative to the analytic PSF from Equation 16 derived using light field analysis, we calculated PSFs for the diffuser surface using wave optics, and used them to create a deblurring error curve. The resulting curve is shown as the dotted red line in Figure 12, and it is very close to the light field curve shown in solid red.

We also used wave optics to compare the deblurring error for our diffuser and the diffuser proposed by Garcia-Guerrero et al. [García-Guerrero et al. 2007]. For a fair comparison, we also restricted the feature size of the Garcia-Guerrero diffuser to be 200μm. Since this design requires features to reduce in size from the center to the edge of the diffuser, only 21 annular sections could be made to fit within a 22mm aperture. The results are shown in Figure 12. The solid red and green lines show the deblurring errors for the diffusion coding and Garcia-Guerrero diffuser, respectively, for PSFs that are averaged over 100 surface realizations. The two curves are very similar, however, a single realization of the diffusion coding surface performs much closer to the average, as seen from the dotted red and green lines. In short, given the imposed fabrication limitations, diffusion coding significantly outperforms the Garcia-Guerrero diffuser.

7 Experimental Results

Figure 13 shows the PSFs produced when using the diffuser shown in Figure 11(d). The PSFs closely resemble the shape predicted by Equation 16 as is evident from the depth-invariance shown in the figure. The PSFs are normalized to unit intensity by color channel. The defocus range is chosen so that the normal lens PSF blur diameter ranges between 0 and 1 mm.

Figure 15 shows two images taken with a normal lens (Figure 15(a) taken with f/4.5 and Figure 15(b) taken with f/29) and two images (Figure 15(c) before deblurring, and Figure 15(d) after deblurring) taken with the diffuser from Section 6. All images are taken with a 50ms exposure time and the brightness in the f/29 image is normalized. The example shows that diffusion coding does indeed give far superior results in comparison to stopping down a lens. The deblurred image in Figure 15(d) extends depth of field by roughly a factor of six.
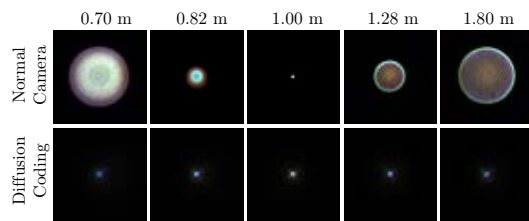
Figure 13: *Measured PSFs for a 50mm f/1.8 lens without (top) and with diffusion coding (bottom). Almost no variation is visible in the diffusion coding PSF.*

Figures 1 and 14-17 compare images taken with a normal lens to diffusion coded images taken with the diffuser from Section 6. The depth range of each scene is chosen so that the normal lens PSF blur diameter ranges between 0 and 1 mm. Within each figure, all images have the same exposure time and aperture setting. In each figure, three images are taken with the normal lens focusing on the background, middle, and foreground. These three images are then compared to the diffusion coded image(s). In all examples, the deblurred diffusion coded images exhibit a significant increase in DOF. Please see the Supplementary Material for additional diffusion coding examples.

All images were captured with a Canon 450D sensor. To capture diffusion coded images, the 22mm diameter diffuser from Figure 11(d) was inserted into the aperture of a 50mm f/1.8 Canon lens. Deblurring of all diffusion coded images was performed using the BM3D deblurring algorithm [Dabov et al. 2008]. The BM3D deblurring algorithm enforces a piecewise smoothness prior that suppresses the noise amplified by the deblurring process. Note that, as discussed in Section 5, all EDOF cameras amplify noise in the deblurring process, and the amount of amplification can be measured by the deblurring error. The result of using the BM3D algorithm is that while our deblurred images do not look noisy in comparison to images captured without the diffuser, some of the fine details in the deblurred images are not preserved.

8   Conclusions and Future Work

We have introduced a new method for extending the DOF of a camera system. We first formulated diffusion as a kernel applied to a light field. We then used this notation to guide the design of a depth-invariant diffuser. We showed that a radially symmetric diffuser produces a PSF which achieves a similar performance to a focal sweep camera, but without the need for mechanical motion. We fabricated a diffuser and verified that it functions as predicted by our theoretical analysis. We showed through examples the effectiveness of our diffusion coding technique in extending DOF. Our examples demonstrated a significant extension of DOF.

While the diffuser design in this paper achieves a similar performance to focal sweep, it remains an open question whether or not the design is optimal. An interesting direction for future work is to determine how to simultaneously maximize depth-invariance in the PSF while also maximizing MTF. Diffusion coding may prove to be a useful tool for implementing cameras that satisfy this optimality criterion.

9   Acknowledgements

(a) Normal camera at three focus settings

(b) Diffusion coded camera

Figure 14: *Extending DOF with diffusion coding. All images were taken with a 16ms exposure time. (a) The top, middle, and bottom images were captured using a a 50mm f/1.8 Canon lens focused on the background, middle, and foreground, respectively. The depth of field is too narrow for all objects to be in focus simultaneously. (b) The diffuser from Section 6 is inserted into the lens aperture and deblurring is applied to recover the EDOF image in (b). Diffusion coding results in a roughly 10× increase in DOF.*

(a) Normal camera (f/4.5)

(c) Diffusion coded camera (captured)

(b) Normal camera (f/29)

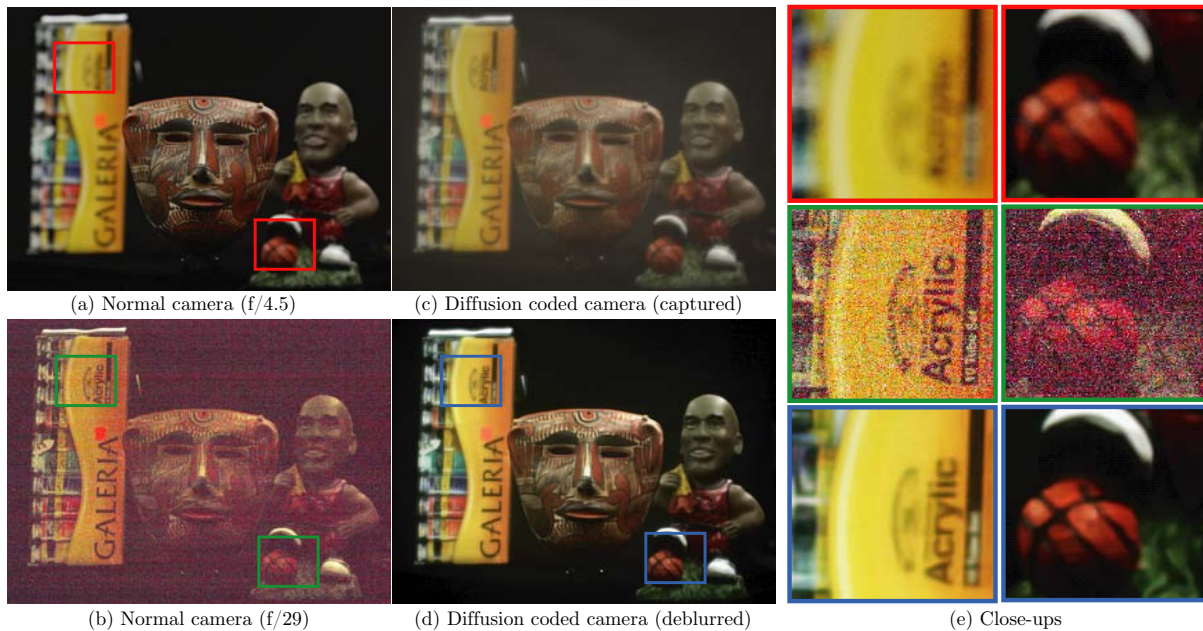(d) Diffusion coded camera (deblurred)

(e) Close-ups

Figure 15: *Noise comparison between a diffusion coded camera and a normal camera. All images were taken with a 20ms exposure time. (a) Image taken with a f/4.5 camera. The DOF is too narrow for all objects to be in focus. (b) Image taken with the lens stopped down to f/29. All the objects are in focus but the noise is significantly increased. (c) Image taken with the same settings as in (a), but with the diffuser from Section 6 inserted into the lens aperture. All objects are in focus, but the image exhibits a slight haze. (d) Image obtained by deblurring the one in (c). The image preserves similar detail as in (b), but with significantly less noise. (e) Close-ups of the images in (a),(b), and (d).*

References

ASHOK, A., AND NEIFELD, M. 2003. Information-based analysis of simple incoherent imaging systems. *Opt. Express 11*, 18, 2153–2162.

ASHOK, A., AND NEIFELD, M. A. 2007. Pseudorandom phase masks for superresolution imaging from subpixel shifting. *Appl. Opt. 46*, 12, 2256–2268.

BAEK, J. 2010. Transfer Efficiency and Depth Invariance in Computational Cameras. In *to appear in Proc. ICCP*.

CAULFIELD, H. J. 1971. Kinoform Diffusers. In *SPIE Conference Series*, vol. 25, 111.

CHI, W., AND GEORGE, N. 2001. Electronic imaging using a logarithmic asphere. *Opt. Lett. 26*, 12, 875–877.

DABOV, K., FOI, A., KATKOVNIK, V., AND EGIAZARIAN, K. 2008. Image restoration by sparse 3D transform-domain collaborative filtering. In *SPIE Conference Series*, vol. 6812, 681207.

DOWSKI, E. R., AND CATHEY, W. T. 1995. Extended depth of field through wave-front coding. *Appl. Opt. 34*, 11, 1859–1866.

DURAND, F., HOLZSCHUCH, N., SOLER, C., CHAN, E., AND SILLION, F. X. 2005. A frequency analysis of light transport. In *SIGGRAPH '05*, 1115–1126.

FREEMAN, W., FERGUS, R., TORRALBA, A., AND FREEMAN, W. 2006. Random Lens Imaging. *MIT CSAIL*.

GARCÍA-GUERRERO, E. E., MÉNDEZ, E. R., ESCAMILLA, H. M., LESKOVA, T. A., AND MARADUDIN, A. A. 2007. Design and fabrication of random phase diffusers for extending the depth of focus. *Opt. Express 15*, 3, 910–923.

GOODMAN, J. W. 1985. *Statistical optics*. Wiley series in pure and applied optics.

HASINOFF, S., KUTULAKOS, K., DURAND, F., AND FREEMAN, W. 2009. Time-constrained photography. In *Proc. ICCV*, 1–8.

HÄUSLER, G. 1972. A method to increase the depth of focus by two step image processing. *Optics Communications 6* (Sep), 38–42.

LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and depth from a conventional camera with a coded aperture. In *SIGGRAPH '07*, 70.

LEVIN, A., HASINOFF, S., GREEN, P., DURAND, F., AND FREEMAN, W. T. 2009. 4d frequency analysis of computational cameras for depth of field extension. In *SIGGRAPH '09*, 1–14.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *SIGGRAPH '96*, 31–42.

NAGAHARA, H., KUTHIRUMMAL, S., ZHOU, C., AND NAYAR, S. 2008. Flexible Depth of Field Photography. In *ECCV '08*, 73.

NG, R. 2005. Fourier slice photography. In *SIGGRAPH '05*, 735–744.

OJEDA-CASTANEDA, J., LANDGRAVE, J. E. A., AND ESCAMILLA, H. M. 2005. Annular phase-only mask for high focal depth. *Opt. Lett. 30*, 13, 1647–1649.

SALES, T. R. M. 2003. Structured microlens arrays for beam shaping. *Optical Engineering 42*, 11, 3084–3085.

VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In *SIGGRAPH '07*, 69.

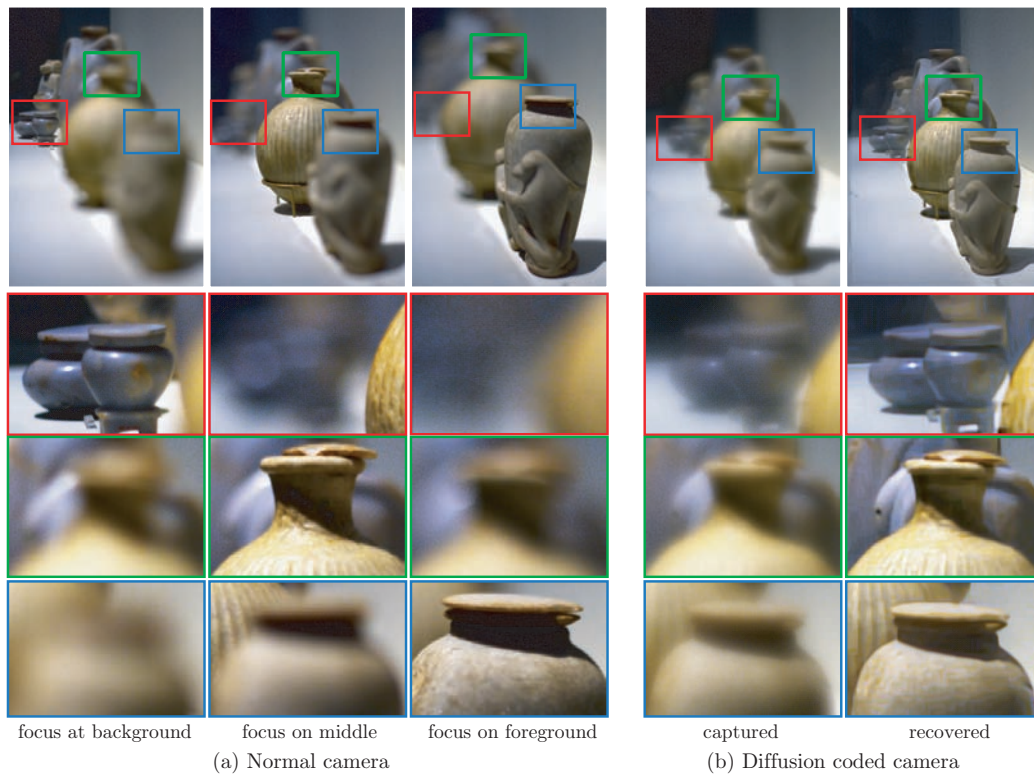ZHOU, C., AND NAYAR, S. 2009. What are Good Apertures for Defocus Deblurring? In *ICCP '09*.

focus at background   focus on middle   focus on foreground

(a) Normal camera

captured   recovered

(b) Diffusion coded camera

Figure 16: *Images of a scene consisting of several vases at different depths shot with a 50mm f/1.8 Canon lens. All images were taken with a 12ms exposure time. (a) Images focused on the background, middle, and foreground from left to right. (b) Images captured using the diffuser from Section 6. The right column shows the result after deblurring. Close-ups at the bottom show that the recovered image significantly increases DOF.*
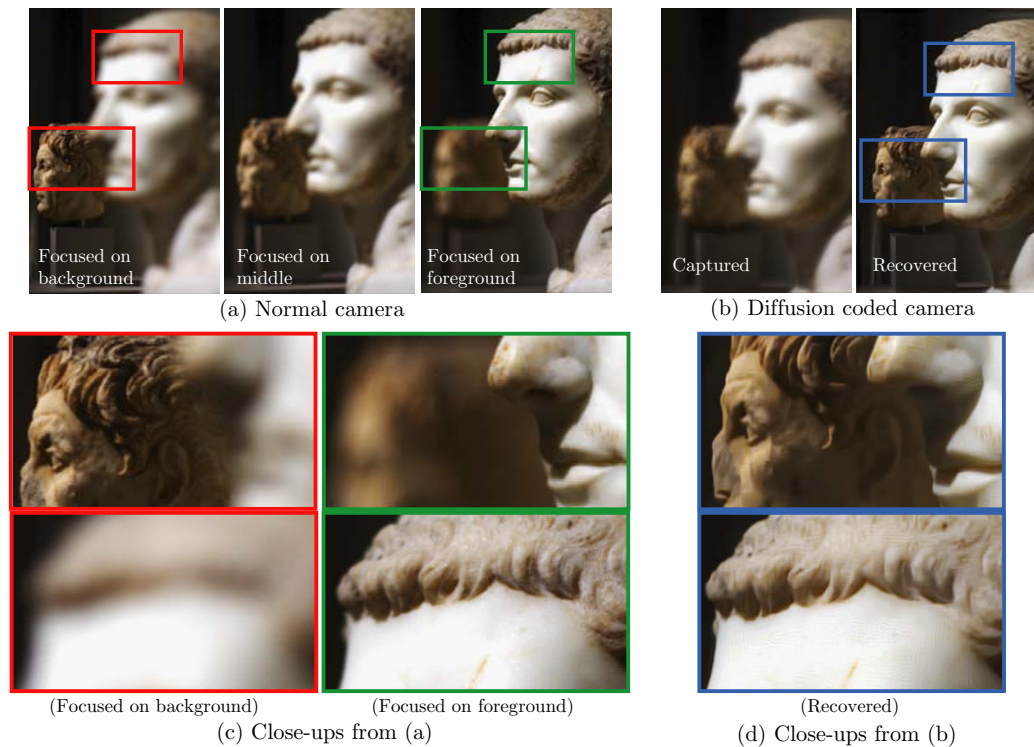

(Focused on background)   (Focused on foreground)

(c) Close-ups from (a)

(Recovered)

(d) Close-ups from (b)

Figure 17: *Images of a scene consisting of two statues at different depths shot with a 50mm f/1.8 Canon lens. All images were taken with a 10ms exposure time. (a) Images are focused on the background, middle, and foreground from left to right. (b) Images captured using the diffuser from Section 6. The right image shows the result after deblurring. Close-ups at the bottom show that the recovered image significantly increases DOF.*