Computer Science and Artificial Intelligence Laboratory

Technical Report

MIT-CSAIL-TR-2009-019 May 8, 2009

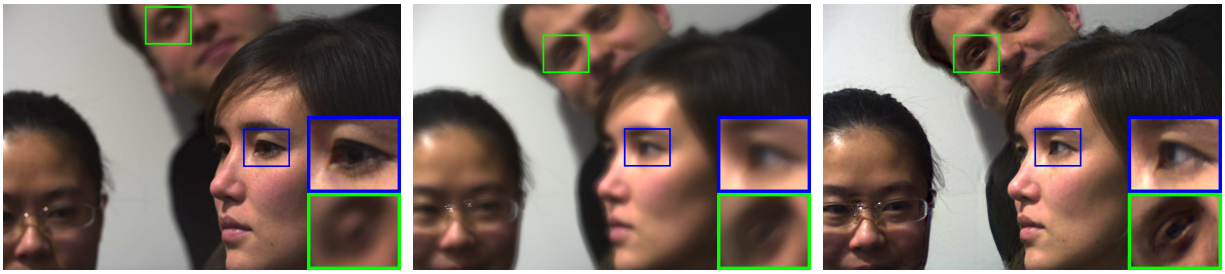
4D Frequency Analysis of Computational Cameras for Depth of Field Extension

Anat Levin, Samuel W. Hasinoff, Paul Green, Fredo Durand, and William T. Freeman

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

4D Frequency Analysis of Computational Cameras for Depth of Field Extension

Anat Levin[1,2] Samuel W. Hasinoff[1] Paul Green[1] Frédo Durand[1] William T. Freeman[1]
[1]MIT CSAIL [2]Weizmann Institute



Standard lens image | Our lattice-focal lens: input | Lattice-focal lens: all-focused output

Figure 1: *Left: Image from a standard lens showing limited depth of field, with only the rightmost subject in focus. Center: Input from our lattice-focal lens. The defocus kernel of this lens is designed to preserve high frequencies over a wide depth range. Right: An all-focused image processed from the lattice-focal lens input. Since the defocus kernel preserves high frequencies, we achieve a good restoration over the full depth range.*

Abstract

Depth of field (DOF), the range of scene depths that appear sharp in a photograph, poses a fundamental tradeoff in photography—wide apertures are important to reduce imaging noise, but they also increase defocus blur. Recent advances in computational imaging modify the acquisition process to extend the DOF through deconvolution. Because deconvolution quality is a tight function of the frequency power spectrum of the defocus kernel, designs with high spectra are desirable. In this paper we study how to design effective extended-DOF systems, and show an upper bound on the maximal power spectrum that can be achieved. We analyze defocus kernels in the 4D light field space and show that in the frequency domain, only a low-dimensional 3D manifold contributes to focus. Thus, to maximize the defocus spectrum, imaging systems should concentrate their limited energy on this manifold. We review several computational imaging systems and show either that they spend energy outside the focal manifold or do not achieve a high spectrum over the DOF. Guided by this analysis we introduce the lattice-focal lens, which concentrates energy at the low-dimensional focal manifold and achieves a higher power spectrum than previous designs. We have built a prototype lattice-focal lens and present extended depth of field results.

Keywords: Computational camera, depth of field, light field, Fourier analysis.

1 Introduction

Depth of field, the depth range over which objects in a photograph appear acceptably sharp, presents an important tradeoff. Lenses gather more light than a pinhole, which is critical to reduce noise, but this comes at the expense of defocus outside the focal plane. While some defocus can be removed computationally using deconvolution, the results depend heavily on the information preserved by the blur, as characterized by the frequency power spectrum of the defocus kernel. Recent advances in computational imaging [Dowski and Cathey 1995; Levin et al. 2007; Veeraraghavan et al. 2007; Hausler 1972; Nagahara et al. 2008] modify the image acquisition process to enable extended depth of field through such a deconvolution approach.

Computational imaging systems can dramatically extend depth of field, but little is known about the maximal frequency magnitude response that can be achieved. In this paper, we use a standard computational photography tool, the light field, e.g., [Levoy and Hanrahan 1996; Ng 2005; Levin et al. 2008a], to address these issues. Using arguments of conservation of energy and taking into account the finite size of the aperture, we present bounds on the power spectrum of all defocus kernels.

Furthermore, a dimensionality gap has been observed between the 4D light field and the space of 2D images over the 1D set of depths [Gu et al. 1997; Ng 2005]. In the frequency domain, only a 3D manifold contributes to standard photographs, which corresponds to *focal* optical conditions. Given the above bounds, we show that it is desirable to avoid spending power in the other *afocal* regions of the light field spectrum. We review existing camera designs and find that some spend significant power in these afocal regions, while others do not achieve a high spectrum over the depth range.

Our analysis leads to the development of the lattice-focal lens—a novel design which allows for improved image reconstruction. It is designed to concentrate energy at the focal manifold of the light field spectrum, and achieves defocus kernels with high spectra. The design is a simple arrangement of lens patches with different focal powers, but the patches' size and powers are carefully derived. The defocus kernels of a lattice-focal lens are high over a wide depth range, but they are not depth invariant. This both requires and enables coarse depth estimation. We have constructed a prototype and demonstrate encouraging extended depth of field results.

1.1 Depth of field evaluation

To facilitate equal comparison across designs all systems are allocated a fixed time budget and maximal aperture width, and hence can collect an equal amount of photons. All systems are expected to cover an equal depth range $d \in [d_{min}, d_{max}]$.

Similar to previous work, we focus on Lambertian scenes and assume locally constant depth. The observed image B of an object at depth d is then described as a convolution $B = \phi_d \otimes I + N$, where I is the ideally sharp image, N is the imaging noise, and ϕ_d is the defocus kernel, commonly referred to as the point spread function (PSF). The defocus PSF ϕ_d is often analyzed in terms of its Fourier transform $\hat{\phi}_d$, known as the optical transfer function (OTF). In the frequency domain, convolution is a multiplication $\hat{B}(\omega) = \hat{\phi}_d(\omega)\hat{I}(\omega) + \hat{N}(\omega)$ where hats denote Fourier transforms. In a nutshell, deblurring divides every spatial frequency by the ker-

nel spectrum, so the information preserved at a spatial frequency ω depends strongly on the kernel spectrum. If $|\hat{\phi}_d(\omega)|$ is low, noise is amplified and image reconstruction is degraded. To capture scenes with a given depth range $d \in [d_{min}, d_{max}]$, we want PSFs ϕ_d whose modulation transfer function (MTF) $|\hat{\phi}_d|$ is as high as possible for every spatial frequency ω, over the full depth range. Noise is absent from the equations in the rest of this paper, because whatever noise is introduced by the sensor gets amplified as a monotonic function of $|\hat{\phi}_d(\omega)|$.

In this paper, we focus on the stability of the deblurring process to noise and evaluate imaging systems according to the spectra they achieve over a specified depth range. We note, however, that many approaches such as coded apertures and our new lattice-focal lens involve a depth-dependent PSF ϕ_d and require a challenging depth identification stage. On the positive side, such systems output a coarse depth map of the scene in addition to the all-focused image. In contrast, designs like wavefront coding and focus sweep have an important advantage: their blur kernel is invariant to depth.

While the tools derived here apply to many computational cameras, our focus is on designs capturing only a single input image. In appendix B we present one possible extension to multiple measurement strategies like the focal stack and the plenoptic camera.

1.2 Related work

Depth of field is traditionally increased by reducing the aperture, but this unfortunately lowers the light collected and increases noise. Alternatively, a focal stack [Horn 1968; Hasinoff and Kutulakos 2008] captures a sequence of images with narrow depth of field but varying focus, which can be merged for extended depth of field [Ogden et al. 1985; Agarwala et al. 2004]. Our new lattice-focal lens can be thought of as capturing all the images from a special focal stack, shifted and summed together in a single photo.

New designs have achieved improved frequency response together with a depth invariant PSFs, allowing for deconvolution without depth estimation. Wavefront coding achieves this with a cubic optical element [Dowski and Cathey 1995]. Others use a log asphere [George and Chi 2003] and focus sweep approaches modify the focus configuration continuously during the exposure [Hausler 1972; Nagahara et al. 2008].

In contrast, coded aperture approaches [Veeraraghavan et al. 2007; Levin et al. 2007] make the defocus blur more discriminative to depth variations. Having identified the defocus diameter, blur can be partially removed via deconvolution. One disadvantage of this design is that some light rays are blocked. A more serious problem is that the lens is still focused only at one particular depth and objects located away from the focus depth are still heavily blurred.

Other designs [Ben-Eliezer et al. 2005] divide the aperture into subsquares consisting of standard lenses, similar to our lattice-focal lens. But while these methods involve redundant focal lengths, our analysis lets us optimize the combination of different focal powers for improved depth of field.

We build on previous analysis of cameras and defocus in light field space [Ng 2005; Adams and Levoy 2007; Levin et al. 2008a]. A related representation in the Fourier optics literature is the Ambiguity function [Rihaczek 1969; Papoulis 1974; Brenner et al. 1983; FitzGerrell et al. 1997], allowing a simultaneous analysis of defocus over a continuous depth range.

2 Background on defocus in light field space

Our main analysis is based on geometric optics and the light field, but appendix C provides complementary derivations using wave optics. We first review how the light field can be used to analyze cameras [Ng 2005; Levin et al. 2008a]. It is a 4D function $\ell(x,y,u,v)$

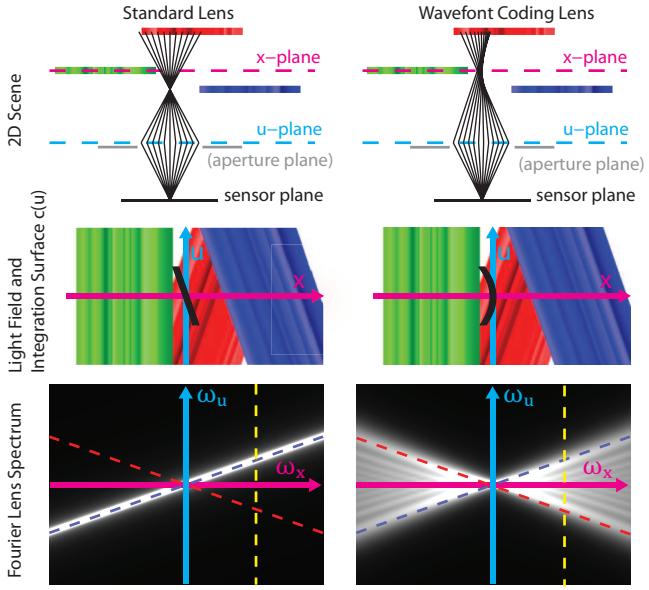


Figure 2: *Integration surfaces in flatland. Top: Ray mapping diagrams. Middle: The corresponding light field and integration surface $c(u)$. Bottom: The lens spectrum $\hat{k}$. The blue/red slices represent OTF-slices of the blue/red objects respectively. The vertical yellow slices represent ω_{x_0} slices discussed in Sec. 3. Left: Standard lens focused at the blue object. Right: Wavefront coding.*

u, v	aperture plane coordinates
x, y	spatial coordinates (at focus plane)
$\omega_{x,y}$	spatial frequencies
Ω	max spatial frequency
$\phi(x,y)$	point spread function (PSF)
$\hat{\phi}(\omega_x, \omega_y)$	optical transfer function (OTF)
$k(x,y,u,v)$	4D lens kernel
$\hat{k}(\omega_x, \omega_y, \omega_u, \omega_v)$	4D lens spectrum
A	aperture width
εA	hole/subsquare width
$\alpha(\omega_{x,y})$, $\beta(\omega_{x,y})$	bounded multiplicative factors (Eqs. (43,11))

Table 1: *Notation.*

describing radiance for all rays in a scene, where a ray is parameterized by its intersections with two parallel planes, the uv-plane and the xy-plane [Levoy and Hanrahan 1996]. Figure 2 shows a 2D flatland scene and its corresponding 2D light field. We assume the camera aperture is positioned on the uv-plane, and xy is a plane in the scene (e.g., the focal plane of a standard lens). x, y are spatial coordinates and the u, v coordinates denote the viewpoint direction.

An important property is that the light rays emerging from a given physical point correspond to a 2D plane in 4D of the form

$$x = su + (1-s)p_x,\ y = sv + (1-s)p_y\ , \tag{1}$$

whose slope s encodes the object's depth:

$$s = (d - d_o)/d\ , \tag{2}$$

where d is the object depth and d_o the distance between the uv, xy planes. The offsets p_x and p_y characterize the location of the scene point within the plane at depth d.

Each sensor element gathers light over its 2D area and the 2D aperture. This is a 4D integral over a set of rays, and under first order

optics (paraxial optics), it can be modeled as a convolution [Ng 2005; Levin et al. 2008a]. A shift-invariant kernel $k(x,y,u,v)$ determines which rays are summed for each element, as governed by the lens. Before applying imaging noise, the value recorded at a sensor element is then:

$$\tilde{B}(x_0,y_0) = \iiiint k(x_0-x, y_0-y, -u, -v)\ell(x,y,u,v)\,dxdydudv\ . \tag{3}$$

For most designs, the 4D kernel is effectively non-zero only at a 2D integration surface because the pixel area is small compared to the aperture. That is, the 4D kernel is of the form

$$k(x,y,u,v) = \delta(x-c_x(u,v), y-c_y(u,v))R(u/A)R(v/A)\ , \tag{4}$$

where R is a rect function, δ denotes a Dirac delta, and $c(u,v) \rightarrow (x,y)$ is a 2D$\rightarrow$2D surface describing the ray mapping at the lens's aperture, which we assume to be square and of size $A \times A$. The surface c is shown in black in the middle row of Figure 2.

For example, a standard lens focuses rays emerging from a point at the focus depth and the integration surface c is linear $c(u,v) = (su, sv)$. The integration slope s corresponds to the slope of the focusing distance (Fig. 2, left). When integrating a light field with the same slope (blue object in Fig. 2), all rays contributing to a sensor element come from the same 3D point. In contrast, when the object is misfocused (e.g., red/green objects), values from multiple scene points get averaged, causing defocus. Wavefront coding [Dowski and Cathey 1995] involves a cubic lens. Since refraction is a function of the surface normal, the kernel is a parabolic surface [Levin et al. 2008b; Zhang and Levoy 2009] (Fig. 2, right) defined by

$$c(u,v) = (au^2, av^2)\ . \tag{5}$$

Finally, the kernel of the focus sweep is not a 2D surface but the integral of standard lens kernels with different slopes/depths.

Consider a Lambertian scene with locally constant depth. If the local scene depth, or slope, is known, the noise-free defocused image $\tilde{B}$ can be expressed as a convolution of an ideal sharp image I with a PSF ϕ_s: $\tilde{B} = \phi_s \otimes I$. As demonstrated in [Levin et al. 2008c], for a given slope s this PSF is fully determined by projecting the 4D lens kernel k along the slope s:

$$\phi_s(x,y) = \iint k(x,y,u+sx,v+sy)dudv\ . \tag{6}$$

That is, we simply integrate over all rays $(x,y,u+sx,v+sy)$ corresponding to a given point in the xy-plane (see Eq. 1).

For example, we have seen that the 4D kernel k for a standard lens is planar. If the slope s of an object and the orientation of this planar k coincide, the object is in focus and the projected PSF ϕ_s is an impulse. For a different slope the projected PSF is a box filter, and the width of this box depends on the difference between the slopes of the object and that of k. For wavefront coding, the parabolic 4D kernel has an equal projection in all directions, explaining why the resulting PSF is invariant to object depth [Levin et al. 2008b; Zhang and Levoy 2009].

Now that we have expressed defocus as a convolution, we can analyze it in the frequency domain. Let $\hat{k}(\omega_x,\omega_y,\omega_u,\omega_v)$ denote the 4D lens spectrum, the Fourier transform of the 4D lens kernel $k(x,y,u,v)$. Figure 2 visualizes lenses spectra $\hat{k}$ in flatland for a standard and wavefront coding lenses. As the PSF ϕ_s is obtained from k by projection (Eq. (6)), by the Fourier slice theorem, the OTF (optical transfer function) $\hat{\phi}_s$ is a slice of the 4D lens spectrum $\hat{k}$ in the orthogonal direction [Ng 2005; Levin et al. 2008c]:

$$\hat{\phi}_s(\omega_x,\omega_y) = \hat{k}(\omega_x,\omega_y,-s\omega_x,-s\omega_y)\ . \tag{7}$$

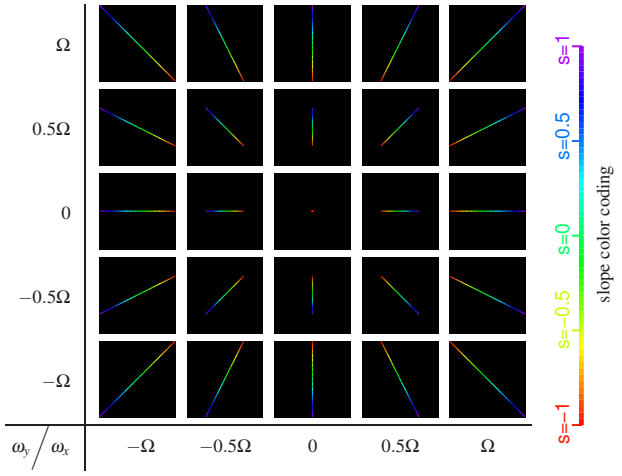


Figure 3: *Layout of the 4D lens spectrum, highlighting the focal manifold. Each subplot represents a ω_{x_0,y_0}-slice, $\hat{k}_{\omega_{x_0,y_0}}(\omega_u,\omega_v)$. The outer axes vary the spatial frequency ω_{x_0,y_0}, i.e., the slicing position. The inner axes of each subplot, i.e., of each slice, vary $\omega_{u,v}$. The entries of $\hat{k}$ along each focal segment are color coded, so that the 2D set of points sharing the same color corresponds to an OTF with a given depth/slope (e.g., the red points define an OTF for the slope $s=-1$). This illustrates the dimensionality gap: the set of entries contributing to an OTF at any physical depth occupies only a 1D segment in each 2D ω_{x_0,y_0}-slice. In the flatland case (Fig. 2), each ω_{x_0,y_0}-slice corresponds to a vertical column.*

Below we refer to slices of this form as *OTF-slices*, because they directly provide the OTF, describing the frequency response due to defocus at a given depth. OTF-slices in flatland are illustrated in the last row of Figure 2 (dashed red/blue). These are slanted slices whose slope is orthogonal to the object slope in the primal light field domain. Low spectrum values in $\hat{k}$ leads to low magnitudes in the OTF for the corresponding depth. In particular, for a standard lens, only the OTF-slice corresponding to the focusing distance (dashed blue, Fig. 2 left) has high values.

Notations and assumptions: All systems in this paper are allocated a fixed exposure time, w.l.o.g. 1. The aperture size is $A \times A$. Δ denotes a pixel width back-projected onto the focal xy-plane. In the frequency domain we deal with the range $[-\Omega,\Omega]$, where $\Omega = 1/(2\Delta)$. $\omega_{x,y}, \omega_{u,v}$ are shortcuts for the 2D vectors (ω_x,ω_y), (ω_u,ω_v). Table 1 summarizes notations.

We seek to capture a fixed depth range $[d_{min}, d_{max}]$. To simplify the light field parameterization, we select the location of the xy plane according to the harmonic mean $d_o = \frac{2d_{min}d_{max}}{d_{min}+d_{max}}$, corresponding to the point at which one would focus a standard lens to equalize defocus diameter at both ends of the depth range, e.g., [Hasinoff and Kutulakos 2008]. This maps the depth range to the symmetric slope range $[-S/2, S/2]$, where $S = \frac{2(d_{max}-d_{min})}{d_{max}+d_{min}}$ (Eq. (2)). Under this parameterization the defocus diameter (on the xy-plane) of slope s can be expressed simply as $A|s|$.

We also assume that scene radiance is fairly constant over the narrow solid angle subtended by the camera aperture. This assumption is violated by highly specular objects or at occlusion boundaries.

3 Frequency analysis of depth of field

We now analyze the requirements, strategies, and limits of depth of field extension. We show that a key factor for depth of field optimization is the presence of a *dimensionality gap* in the 4D light field: only a manifold of the 4D spectrum, which we call *focal*,

contributes to focusing at physical depths. Furthermore, we show that the energy in a 4D lens spectrum is bounded. This suggests that to optimize depth of field, most energy should be concentrated on the focal manifold. We discuss existing lens designs and show that many of them spend energy outside the focal manifold. In Sec. 4 we propose a novel design which significantly reduces this problem.

3.1 The dimensionality gap

As described above, scene depth corresponds to slope s in the light field. It has, however, been observed that the 4D light field has a *dimensionality gap*, in that most slopes do not correspond to a physical depth [Gu et al. 1997; Ng 2005]. Indeed, the set of all 2D planes $x = s_u u + p_x,\ y = s_v v + p_y$ described by their slope s_u, s_v and offset p_x, p_y is 4D. In contrast, the set corresponding to real depth, i.e., where $s = s_u = s_v$, is only 3D, as described by Eq. (1). This makes sense because scene points are 3D. The dimensionality gap is a property of the 4D light field, and does not exist for the 2D light field in flatland. The other slopes where $s_u \neq s_v$ are *afocal* and represent rays from astigmatic refractive or reflective surfaces, which are surfaces with anisotropic curvature [Adams and Levoy 2007], e.g., the reflection from a cylindrical mirror. Since we consider scenes which are sufficiently Lambertian over the aperture, afocal light field orientations hold no interesting information.

The dimensionality gap is particularly clear in the Fourier domain [Ng 2005]. Consider the 4D lens spectrum $\hat{k}$, and examine the 2D slices $\hat{k}_{\omega_{x_0,y_0}}(\omega_u, \omega_v)$, in which the the spatial frequencies $\omega_{x_0}, \omega_{y_0}$ are held constant (Fig. 3). We call these ω_{x_0,y_0}*-slices*. In flatland, ω_{x_0,y_0}-slices are vertical slices (yellow in Fig. 2). Following Eq. (7), we note that the set of entries in each $\hat{k}_{\omega_{x_0,y_0}}$ participating in the OTF for any depth is restricted to a 1D line:

$$\hat{k}_{\omega_{x_0,y_0}}(-s\omega_{x_0}, -s\omega_{y_0}) \ , \tag{8}$$

for which $\omega_u = -s\omega_{x_0},\ \omega_v = -s\omega_{y_0}$. For a fixed slope range $s \in [-S/2, S/2]$ the set of entries participating in any OTF $\hat{\phi}_s$ is a 1D segment. These segments, which we refer to as *focal segments*, are highlighted in Figure 3. The rest of the spectrum is *afocal*. This property is especially important, because it implies that *most entries of $\hat{k}$ do not contribute to an OTF at any depth*.

As an example, Figure 4(b-e) shows the 2D families of 2D ω_{x_0,y_0}-slices for a variety of cameras. A standard lens has a high response for an isolated point in each slice, corresponding to the focusing distance. In contrast, wavefront coding (Fig. 4(e)) has a broader response that spans more of the focal segment, but also over the afocal region. While the spectrum of the focus sweep (Fig. 4(d)) is on the focal segment, its magnitude is lower magnitude than that of a standard lens.

3.2 Upper bound on the defocus MTF

In this section we derive a bound on the defocus MTF. As introduced earlier, we pose depth of field extension as maximizing the MTFs $|\hat{\phi}_s(\omega_{x,y})|$ over all slopes $s \in [-S/2, S/2]$ and over all spatial frequencies $\omega_{x,y}$. Since the OTFs are slices from the 4D lens spectrum $\hat{k}$ (Eq. (7)), this is equivalent to maximizing the spectrum on the focal segments of $\hat{k}$.

We first derive the available energy budget, using a direct extension of the 1D case [FitzGerrell et al. 1997; Levin et al. 2008c].

Claim 1 *For an aperture of size $A \times A$ and exposure length* 1*, the total energy in each ω_{x_0,y_0}-slice is bounded by A^2:*

$$\iint |\hat{k}_{\omega_{x_0,y_0}}(\omega_u, \omega_v)|^2 d\omega_u d\omega_v \leq A^2 \ . \tag{9}$$

The proof, provided in appendix A, follows from the finite amount of light passing through a bounded aperture over a fixed exposure. As a consequence of Parseval's theorem, this energy budget then applies to every ω_{x_0,y_0}-slice $\hat{k}_{\omega_{x_0,y_0}}$. While Claim 1 involves geometric optics, similar bounds can be obtained with Fourier optics using slices of the ambiguity function [Rihaczek 1969; FitzGerrell et al. 1997]. In appendix C we derive an analogous bound under Fourier optics, with a small difference—the budget is no longer equal across spatial frequencies, but decreases with the diffraction-limited MTF.

As in the 1D space-time case [Levin et al. 2008c], optimal worst-case performance can be realized by spreading the energy budget uniformly over the range of slopes. The key difference in this paper is the dimensionality gap. As shown in Figure 3, the OTFs $\hat{\phi}_s$ cover only a 1D line segment, and most entries in an ω_{x_0,y_0}-slice $\hat{k}_{\omega_{x_0,y_0}}$ do not contribute to any OTF. Therefore, the energy budget should be spread evenly over the 1D focal segment only.

Given a power budget for each ω_{x_0,y_0}-slice, the upper bound for the defocus MTF concentrates this budget on the 1D focal segment only. Distributing energy over the focal manifold requires caution, however, because the segment effectively has non-zero thickness due to its finite support in the primal domain. If a 1D focal segment had zero thickness, its spectrum values could be made infinite while still obeying the norm constraints of Claim 1. As we show below, since the primal support of k is finite (k admits no light outside the aperture), the spectrum must be finite as well, so the 1D focal segment must have non-zero thickness. Slices from this ideal spectrum are visualized in Figure 4(a).

Claim 2 *The worst-case defocus MTF for the range $[-S/2, S/2]$ is bounded. For every spatial frequency $\omega_{x,y}$:*

$$\min_{s\in[-S/2,S/2]} |\hat{\phi}_s(\omega_x, \omega_y)|^2 \leq \frac{\beta(\omega_{x,y})A^3}{S|\omega_{x,y}|} \ , \tag{10}$$

where the factor

$$\beta(\omega_{x,y}) = \frac{|\omega_{x,y}|}{\max(|\omega_x|, |\omega_y|)} \left(1 - \frac{\min(|\omega_x|, |\omega_y|)}{3 \cdot \max(|\omega_x|, |\omega_y|)}\right) \tag{11}$$

is in the range $[\frac{5\sqrt{5}}{12}, 1] \approx [0.93, 1]$.

Proof: For each ω_{x_0,y_0}-slice $\hat{k}_{\omega_{x_0,y_0}}$ the 1D focal segment is of length $S|\omega_{x_0,y_0}|$. We first show that the focal segment norm is bounded by A^3, and then the worst-case optimal strategy is to spread the budget evenly over the segment.

To simplify notations, we consider the case $\omega_{y_0} = 0$ since the general proof is similar after a basis change. For this case, the 1D focal segment is a horizontal line of the form $\hat{k}_{\omega_{x_0,y_0}}(\omega_u, 0)$, shown in the central row of Figure 3. For a fixed value of ω_{x_0}, this line is the Fourier transform of:

$$\iiint k(x,y,u,v) e^{-2i\pi(\omega_{x_0}x + 0y + 0v)} dxdydv \ . \tag{12}$$

By showing that the total power of Eq. (12) is bounded by A^3, Parseval's theorem gives us the same bound for the focal segment.

Since the exposure time is assumed to be 1, we collect unit energy through every u, v point lying within the clear aperture[1]:

$$\iint k(x,y,u,v)dxdy = \begin{cases} 1 & |u| \leq A/2,\ |v| \leq A/2 \\ 0 & \text{otherwise} \end{cases} \ . \tag{13}$$

[1] If an amplitude mask is placed at the aperture (e.g., a coded aperture) the energy will be reduced and the upper bound still holds.

Camera type	Squared MTF
a. Upper bound	$\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2 \;\leq\; \frac{A^3}{S\lvert\omega_{x,y}\rvert}$
b. Standard lens	$\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2 \;=\; A^4\mathrm{sinc}^2(A(s-s_0)\omega_x)\mathrm{sinc}^2(A(s-s_0)\omega_y)$
c. Coded aperture	$\mathrm{E}[\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2] \;\approx\; \frac{\varepsilon^2A^4}{2}\mathrm{sinc}^2(\varepsilon A(s-s_0)\omega_x)\mathrm{sinc}^2(\varepsilon A(s-s_0)\omega_y)$
d. Focus sweep	$\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2 \;\approx\; \frac{A^2\alpha(\omega_{x,y})^2}{S^2\lvert\omega_{x,y}\rvert^2}$
e. Wavefront coding	$\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2 \;\approx\; \frac{A^2}{S^2\lvert\omega_x\rvert\lvert\omega_y\rvert}$
f. Lattice-focal	$\mathrm{E}[\lvert\hat{\phi}_s(\omega_{x,y})\rvert^2] \;\approx\; \frac{A^{8/3}\beta(\omega_{x,y})}{S^{4/3}\Omega^{1/3}\lvert\omega_{x,y}\rvert}$

Table 2: *Squared MTFs of computational imaging designs. See Table 1 for notation. The optimal spectrum bound falls off* linearly *as a function of spatial frequency, yet existing designs such as the focus sweep and wavefront coding fall off* quadratically *and do not utilize the full budget. The new lattice-focal lens derived in this paper achieves a higher spectrum, closer to the upper bound.*

A phase change to the integral in Eq. (13) does not increase its magnitude, therefore, for every spatial frequency ω_{x_0,y_0},

$$\left|\iint k(x,y,u,v)e^{-2i\pi(\omega_{x_0}x+\omega_{y_0}y)}dxdy\right| \leq 1\ . \tag{14}$$

Using Eq. (14) and the fact that the aperture is width A along on the v-axis, we obtain:

$$\left|\iiint k(x,y,u,v)e^{-2i\pi\omega_{x_0}x+0y+0v}dxdydv\right|^2 \leq A^2\ . \tag{15}$$

On the u-axis, the aperture has width A as well. By integrating Eq. (15) over u we see the power is bounded by A^3:

$$\int\left|\iiint k(x,y,u,v)e^{-2i\pi(\omega_{x_0}x+\omega_{y_0}y)}dxdydv\right|^2 du \leq A^3\ . \tag{16}$$

Since the left-hand side of Eq. (15) is the power spectrum of $\hat{k}_{\omega_{x_0,y_0}}(\omega_u,0)$, by applying Parseval's theorem we see that the total power over the focal segment is bounded by A^3 as well:

$$\int|\hat{k}_{\omega_{x_0,y_0}}(\omega_u,0)|^2 d\omega_u \leq A^3\ . \tag{17}$$

Since the focal segment norm is bounded by A^3, and since we aim to maximize the worst-case magnitude, the best we can do is to spread the budget uniformly over the length $S|\omega_{x_0,y_0}|$ focal segment, which bounds the worst MTF power by $A^3/S|\omega_{x_0}|$. In the general case, Eq. (16) is bounded by $\beta(\omega_{x,y})A^3$ rather than A^3, and Eq. (10) follows. □

3.3 Analysis of existing designs

We analyze the spectra of existing imaging designs with particular attention paid to the spectrum on the focal manifold since it is the portion of the spectrum that contributes to focus at physical depths. Figure 4 visualizes ω_{x_0,y_0}-slices through a 4D lens spectrum $|\hat{k}|$ for recent imaging systems. Figure 5 shows the corresponding MTFs (OTF-slices) at a few depths. A low spectrum value at a point on the focal segment leads to low spectrum content at the OTF of the corresponding depth. Examining Figures 4 and 5, we see that some designs spend a significant portion of the budget on afocal regions. The MTFs for the previous designs shown in Figure 5 are lower than the upper bound. We have analytically computed spectra for these designs. The derivation is provided in appendix A and summarized in Table 2. We observe that no existing spectrum reaches the upper bound. Below we review the results in Table 2b-e and provide some intuitive arguments. In the next section we introduce a new design whose spectrum is higher than all known designs, but still does not fully meet the bound.

Standard lens: For a standard lens focused at depth s_0 we see in Figure 4(b) high frequency content near the isolated points $\hat{k}_{\omega_{x_0,y_0}}(-s_0\omega_{x_0},-s_0\omega_{y_0})$, which correspond to the in-focus depth $\hat{\phi}_{s_0}$. The spectrum falls off rapidly away from these points, with a sinc whose width is inversely proportional to the aperture. When the deviation between the focus slope and the object slope $|s_0-s|$ is large, this sinc severely attenuates high frequencies.

Coded aperture: The coded aperture [Levin et al. 2007; Veeraraghavan et al. 2007] incorporates a pattern blocking light rays. The integration surface is linear, like that of a standard lens, but has holes at the blocked areas. Compared to the sinc of a standard aperture, the coded aperture camera has a broader spectrum (Fig. 4(c)), but is still far from the bound. To see why, assume w.l.o.g. that the lens is focused at $s_0=0$. The primal integration surface lies on the $x=0,y=0$ plane and $\hat{k}$ is constant over all $\omega_{x,y}$. Indeed, all ω_{x_0,y_0}-slices in Figure 4(c) are equal. Since the union of focal segment orientations from all ω_{x_0,y_0}-slices covers the plane, to guarantee worst-case performance, the coded aperture spectrum should be spread over the entire 2D plane of each ω_{x_0,y_0}-slice. This implies significant energy away from focal segments.

Focus sweep: For a focus sweep camera [Hausler 1972; Nagahara et al. 2008], the focus distance is varied continuously during exposure and the 4D lens spectrum is the average of standard lenses spectra over a range of slopes s_0 (Figs. 4(d) and 5(d)). In contrast to the isolated points covered by a static lens, this spreads energy over the entire focal segment, since the focus varies during exposure. This design does not spend budget away from the focal segment of interest. However, as discussed in appendix A, since the lens kernel describing a focus sweep camera is not a Dirac delta, phase cancellation occurs between different focus settings and the magnitude is lower than the upper bound (Fig. 4(a)).

Wavefront coding: The integration surface of a wavefront coding lens [Dowski and Cathey 1995] is a separable 2D parabola [Levin et al. 2008b; Zhang and Levoy 2009]. The spectrum is a separable extension of that of the 1D parabola [Levin et al. 2008c]. However, while the 1D parabola achieves an optimal worst-case spectrum, this is no longer the case for a 2D parabola in 4D, and the wavefront coding spectrum (Table 2e, Figs. 4(e) and 5(e)) is lower than the bound. The ω_{x_0,y_0}-slices in Figure 4(e) reveal why. Due to the separability, energy is spread uniformly within the minimal rectangle bounding the focal segment. For another perspective, consider the wavefront coding integration surface in the primal domain, which is a separable parabola $c(u,v)=(au^2,av^2)$. A local planar approximation to that surface around an aperture point u_0,v_0 is of the form $c(u,v)=(s_uu,s_vv)$, for $s_u=\frac{\partial c_x}{\partial u}=2au_0$, $s_v=\frac{\partial c_y}{\partial v}=2av_0$. For $u_0\neq v_0$ the lens is locally astigmatic, and as discussed in Sec. 3.1, this is an *afocal* surface. Thus, the only focal part of the wavefront coding lens is the narrow strip along its diagonal, where $u_0=v_0$.

Still, the wavefront coding spectrum is superior to that of coded apertures at low-to-mid frequencies. It spreads budget only within the minimal rectangle bounding the focal segment, but not up to the maximal cutoff spatial frequency. The wavefront coding spectrum and that of a focus sweep are equal if $|\omega_x|=|\omega_y|$. However, the wavefront coding spectrum is significantly improved for $|\omega_x|\to 0$

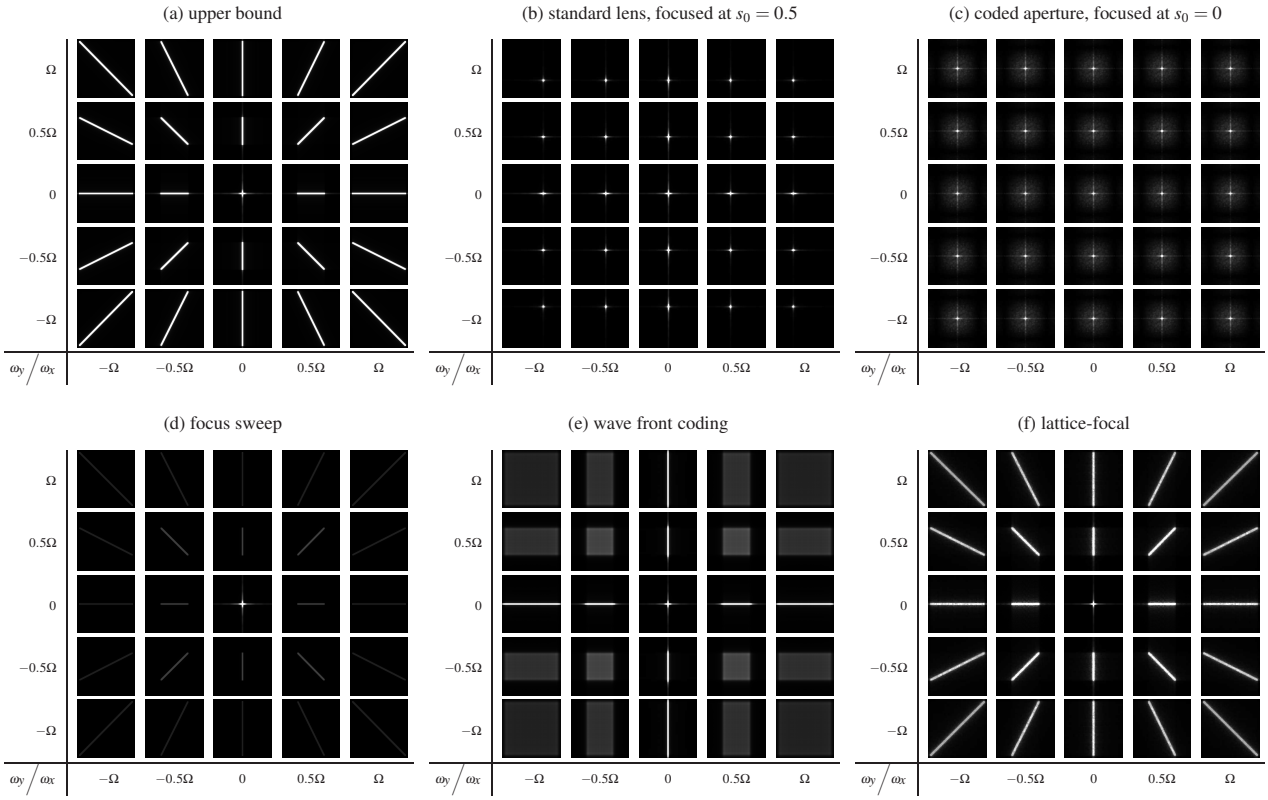


Figure 4: *4D lens spectrum for different optical designs. Each subplot is an ω_{x_0,y_0}-slice as described in Figure 3. In the flatland case of Figure 2, these ω_{x_0,y_0}-slices correspond to vertical columns. An ideal design (a) should account for the dimensionality gap and spend energy only on the focal segments. Yet, this bound is not reached by any existing design. A standard lens (b) devotes energy only to a point in each subplot. A coded aperture (c) is more broadband, but its spectrum is constant over all ω_{x_0,y_0}-slices, so it cannot cover only the focal segment in each ω_{x_0,y_0}-slice. The focus sweep camera (d) covers only the focal segments, but has reduced energy due to phase cancellations and does not achieve the bound. A wavefront coding lens (e) is separable in the ω_u, ω_v directions and spends significant energy on afocal areas. Our new lattice-focal lens (f) is an improvement over existing designs, and spreads energy budget over the focal segments. Note that all subplots show the numerical simulation of particular design instances, with parameters for each design tuned to the depth range (see Sec. 5.1), approximating the analytic spectra in Table 2. The intensity scale is constant for all subplots.*

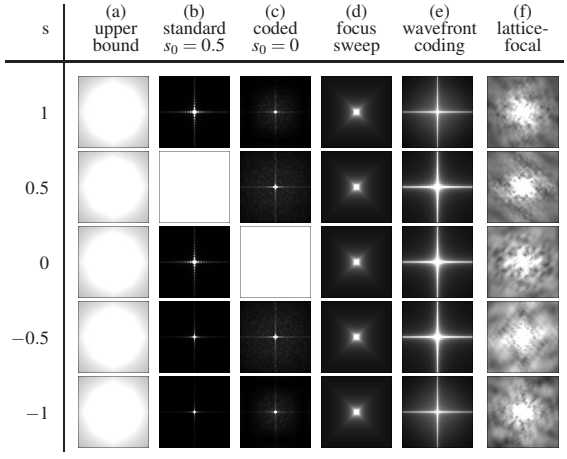


Figure 5: *Spectra of OTF-slices for different optical designs over a set of depths. The subplots represent the MTF of a given imaging system for slope s, $|\hat{\phi}_s(\omega_x, \omega_y)|$, where the subplot axes are $\omega_{x,y}$. These OTF-slices are the 2D analog of the slanted red and blue slices in Figure 2. Our new lattice-focal lens design best approximates the ideal spectrum upper bound. Note that all subplots show the numerical simulation of particular design instances, with parameters for each design tuned to the depth range (see Sec. 5.1), approximating the analytic spectra in Table 2.*

or $|\omega_y| \to 0$, because the rectangle becomes compact, as shown in the central row and column of Figure 4(e).

In appendix B we also analyze the plenoptic camera and the focal stack imaging models. Note that despite all the sinc patterns mentioned so far, the derivation in this section and the simulations in Figures 4 and 5 model pure geometric optics. Diffraction and wave optics effects are also discussed in appendix C. In most cases Fourier optics models lead to small adjustments to the spectra in Table 2, and the spectra are scaled by the diffraction-limited OTF.

Having reviewed several previous computational imaging approaches to extending depth of field, we conclude that none of them spends the energy budget in an optimal way. In a standard lens the entire aperture area is focal, but light is focused only from a single depth. A wavefront coding lens attempts to cover a full depth range, but at the expense that most aperture area is afocal. In the next section we propose a new lens design, the lattice-focal lens, with the best attributes of both—all aperture area is focal, yet it focuses light from multiple depths. This lets our new design get closer to the upper bound compared to existing imaging systems.

4 The lattice-focal lens

Motivated by the previous discussion, we propose a new design, which we call the lattice-focal lens. The spectrum it achieves is higher than previous designs but still lower than the upper bound. In this design, the aperture is divided into $1/\varepsilon^2$ subsquares of

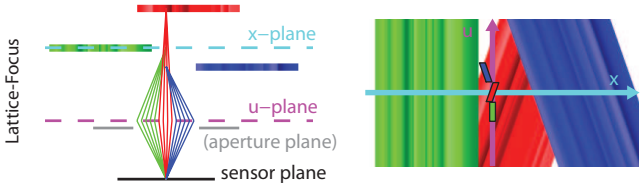
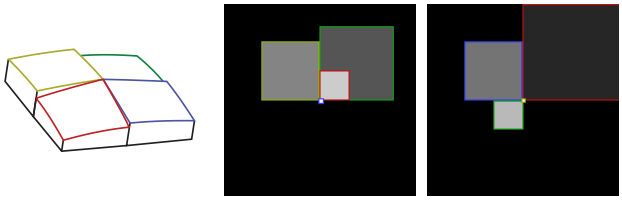


Figure 6: *Left: Ray mapping for a lattice-focal lens in flatland. The aperture is divided into three color-coded sections, each focused on a different depth. Right: In the 2D light field the integration surface is a set of slanted segments, shown with corresponding colors.*



(a) Lattice-focal lens (b) PSFs

Figure 7: *(a) Toy lattice-focal lens design with only 4 subsquares. (b) The PSFs ϕ_s in the primal domain, at two different depths. Each subsquare (color-coded) corresponds to a box in the PSF. The width of each box is a function of the deviation between the subsquare focal depth and the object depth.*

size $\varepsilon A \times \varepsilon A$ each (for $0 < \varepsilon < 1$). Each subsquare is a focal element cropped from a standard lens focused at some slope $s_j \in [-S/2, S/2]$. That is, the integration surface is defined as:

$$c(u,v) = (s_j u, s_j v) \quad \text{for } (u,v) \in W_j \ , \tag{18}$$

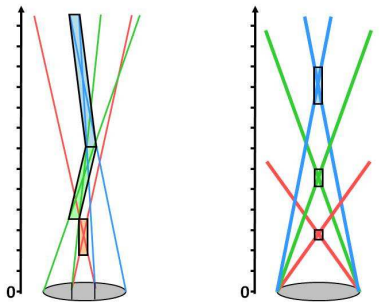
where W_j denotes the area of the j-th subsquare. Figure 6 visualizes the integration surface of a lattice-focal lens, composed of linear surfaces with different slopes (compare with Figure 2, left). Figure 7 illustrates a toy four-element lattice-focal lens and its PSF for two different depths. In the primal domain, the PSF is a superposition of scaled and shifted boxes corresponding to the various aperture subsquares. For this example, one of the subsquares is focused at the correct depth for each scene depth, so the PSF consists of an impulse plus three defocused boxes. The box width is a function of the deviation between the lens focal depth and the object depth.

The OTF $\hat{\phi}_s(\omega_x, \omega_y)$ of a lattice-focal lens is a sum of sincs corresponding to the different subsquares:

$$\sum_j \varepsilon^2 A^2 e^{2\pi i(\gamma_{j,x}\omega_x + \gamma_{j,y}\omega_y)} \operatorname{sinc}\left(\varepsilon A \omega_x (s_j - s)\right) \operatorname{sinc}\left(\varepsilon A \omega_y (s_j - s)\right) \ . \tag{19}$$

For a subsquare centered at aperture point (u_j, v_j), $(\gamma_{j,x}, \gamma_{j,y}) = (u_j(s_j - s), v_j(s_j - s))$ denotes the phase shift of the j-th subsquare, corresponding to its translated center. The 4D spectrum of a single aperture subsquare is a sinc around one point in the focal segment: $\hat{k}_{\omega_{x_0,y_0}}(-s_j\omega_{x_0}, -s_j\omega_{x_0})$. However since each subsquare is focused at a different slope s_j the summed spectra cover the focal segment (Figure 4(f)). In contrast to the spectrum for wavefront coding, the lattice-focal spectrum does not spend much budget away from the focal manifold. This follows from the fact that the subsquare slopes in Eq. (18) are set to be equal in u and v, therefore the entire aperture area is *focal*.

The lattice-focal design resembles the focus sweep in that both distribute focus over the DOF—focus sweep over time, and the lattice-focal design over aperture area. The crucial difference is that since each lattice-focal subsquare is *smaller* than the full aperture, its effective DOF is larger than the DOF for the full aperture (Figure 8). As shown in Fig. 4(d,f) and Fig. 5(d,f), the lattice-focal lens achieves significantly higher spectra than focus sweep. Mathematically, by discretizing the exposure time into N bins, each bin of the focus sweep (focused at slope s_j) contributes $\frac{A^2}{N}\operatorname{sinc}(A(s-s_j)\omega_x)\operatorname{sinc}(A(s-s_j)\omega_y)$ to the OTF. By contrast, by dividing the aperture into N bins, each bin of the lattice-focal lens contributes $\frac{A^2}{N}\operatorname{sinc}(AN^{-1/2}(s-s_j)\omega_x)\operatorname{sinc}(AN^{-1/2}(s-s_j)\omega_y)$. In both cases each bin collects $1/N$ of the total energy (and the sincs' height is A^2/N), but the lattice-focal sinc is wider. While coincidental phase alignments may narrow the sincs, these alignments occur in isolation and do not persist across all depths and all spatial frequencies. Therefore, the lattice-focal lens has a higher spectrum when integrating over s_j.



(a) Lattice-focal lens (b) Discrete focus sweep

Figure 8: *Focus sweep vs. the lattice-focal lens. (a) Lattice-focal lens whose aperture is divided into 3 differently-focused bins. (b) Discrete focus sweep, dividing the integration time into 3 bins, each focusing on a different depth (note that an actual focus sweep camera varies focus continuously). Depth ranges with defocus diameter below a threshold are colored. While in both cases each bin lets in $1/3$ of the energy, the sub-apertures for the lattice-focal lens are narrower than the full aperture used by the focus sweep, hence the effective DOF for each of the lattice-focal bins is larger.*

The ω_{x_0,y_0}-slices in Figure 4(f), and the OTF-slices in Figure 5(f) suggest that the lattice-focal lens achieves a higher spectrum compared to previous designs. In the rest of this section we develop an analytic, average-case approximation for the lattice-focal spectrum, which enables order-of-magnitude comparison to other designs. We then discuss the effect of window size ε and show it is a critical parameter of the construction, and implies a major difference between our design and previous multi-focus designs [George and Chi 2003; Ben-Eliezer et al. 2005].

Spectrum of the lattice-focal lens: The spectrum of a particular lattice focal lens can be computed numerically (Eq. (19)), and Figures 4 and 5 plot such a numerical evaluation. However, to allow an asymptotic order-of-magnitude comparison between lens designs we compute the expected spectrum over random choices of the slopes s_j and subsquare centers (u_j, v_j) in Eq. (18) (note that to simplify the proof, the subsquares in a generic random lattice-focal are allowed to overlap and to leave gaps in the aperture area). Given sufficiently many subsquares, the law of large numbers applies and a sample lattice-focal lens resembles the expected spectrum. While this analysis confers insight, the expected spectrum should not be confused with the spectrum of a particular lattice-focal lens. The spectrum of any particular lattice-focal instance is not equal to the expected one.

Claim 3 *Consider a lattice-focal lens whose subsquare slopes in Eq. (18) are sampled uniformly from the range $[-S/2, S/2]$, and subsquares centers sampled uniformly over the aperture area*

$[-A/2,A/2]\times[-A/2,A/2]$. *For* $|\omega_x|,|\omega_y| > (\varepsilon SA)^{-1}$*, the expected power spectrum asymptotically approaches*

$$\mathrm{E}[|\hat{\phi}_s(\omega_x,\omega_y)|^2] \approx \frac{\varepsilon A^3}{S|\omega_{x,y}|}\beta(\omega_{x,y}) \,, \tag{20}$$

where β *is defined in Eq. (11).*

Proof: Let s denote a particular scene depth of interest and let $\hat{\phi}_s^j$ denote the OTF of the j-th subsquare focused at slope s_j, so that the lattice-focal OTF is $\hat{\phi}_s = \sum_j \hat{\phi}_s^j$. For a subsquare size of $\varepsilon A \times \varepsilon A$, the aperture area is covered by $m = 1/\varepsilon^2$ subsquares. Since the m random variables $\hat{\phi}_s^j$ are drawn independently from the same distribution, it follows that

$$\mathrm{E}[|\hat{\phi}_s|^2] = m\mathrm{E}[|\hat{\phi}_s^j|^2] + m(m-1)|\mathrm{E}[\hat{\phi}_s^j]|^2 \,. \tag{21}$$

The second term in Eq. (21) is positive, and one can show it is small relative to the first term. For simplicity we make the conservative approximation $\mathrm{E}[|\hat{\phi}_s|^2] \approx m\mathrm{E}[|\hat{\phi}_s^j|^2]$, and show how to compute $\mathrm{E}[|\hat{\phi}_s^j|^2]$ below. Note that the exact lattice-focal spectrum (Eq. (19), and the right-hand side of Eq. (21)) involves interference from the phase of each subsquare. An advantage of our approximation $m\mathrm{E}[|\hat{\phi}_s^j|^2]$ is that it bypasses the need to model phase precisely.

Recall that the PSF from each subsquare is a box filter and the OTF is a sinc. If the j-th subsquare is focused at s_j,

$$|\hat{\phi}_s^j(\omega_{x,y})|^2 = \varepsilon^4 A^4 \mathrm{sinc}^2(\varepsilon A\omega_x(s-s_j))\mathrm{sinc}^2(\varepsilon A\omega_y(s-s_j)) \,. \tag{22}$$

Since the subsquare slopes are drawn uniformly from $[-S/2,S/2]$, the expected spectrum is obtained by averaging Eq. (22) over s_j.

$$\mathrm{E}[|\hat{\phi}_s^j|^2] = \frac{\varepsilon^4 A^4}{S}\int_{-S/2}^{S/2} \mathrm{sinc}^2\left(\varepsilon A\omega_x(s_j-s)\right)\mathrm{sinc}^2\left(\varepsilon A\omega_y(s_j-s)\right) ds_j \,. \tag{23}$$

To compute this integral we make use of the following identity: for a 2D vector $r=(r_1,r_2)$,

$$\int_{-\infty}^{\infty} \mathrm{sinc}^2(r_1 t)\mathrm{sinc}^2(r_2 t)dt = \frac{\beta(|r|)}{|r|} \,. \tag{24}$$

If $-S/2 < s < S/2$ and S is large, we can assume that the integration boundaries of Eq. (23) are sufficiently large[2], and asymptotically approximate Eq. (23) with the unbounded integration of Eq. (24):

$$\begin{aligned}\mathrm{E}[|\hat{\phi}_s^j|^2] &= \frac{\varepsilon^4 A^4}{S}\int_{-S/2}^{S/2} \mathrm{sinc}^2\left(\varepsilon A\omega_x(s_j-s)\right)\mathrm{sinc}^2\left(\varepsilon A\omega_y(s_j-s)\right) ds_j \\ &= \frac{\varepsilon^4 A^4}{S}\int_{-S/2+s}^{S/2+s} \mathrm{sinc}^2\left(\varepsilon A\omega_x s_j\right)\mathrm{sinc}^2\left(\varepsilon A\omega_y s_j\right) ds_j \\ &\approx \frac{\varepsilon^3 A^3\beta(\omega_{x,y})}{S|\omega_{x,y}|} \,.\end{aligned} \tag{25}$$

Eq. (20) now follows from Eq. (25), after multiplying by the number of subsquares, $m = \frac{1}{\varepsilon^2}$. □

[2]Note that the approximation in Eq. (25) is reasonable for $|\omega_x|,|\omega_y| > (S\varepsilon A)^{-1}$. The approximation is crude at the low frequencies but becomes accurate at higher frequencies, for which the MTF approaches the desired fall off. Furthermore, note that at the exact integration boundaries ($s = \pm S/2$) one gets only half of the contrast. Thus, in practice, one should set S a bit higher than the actual depth range to be covered.

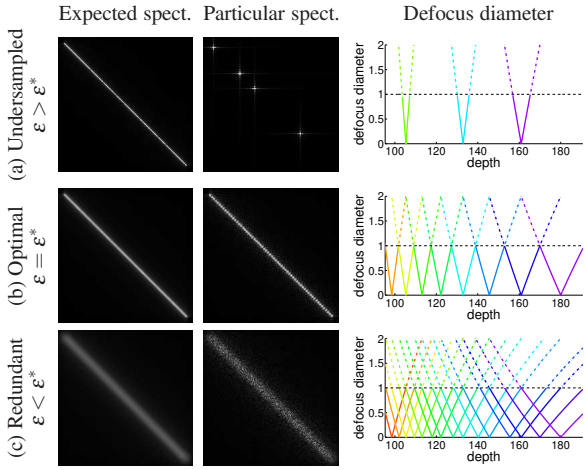


Figure 9: *The lattice-focal lens with varying window sizes. Left:* ω_{x_0,y_0}*-slice at* $\omega_x = 0.9\Omega, \omega_y = -0.9\Omega$*, through the* expected *spectrum. Middle:* ω_{x_0,y_0}*-slice from a* particular *lattice-focal lens instance. Right: The defocus diameter over the depth of field. The expected spectrum improves when the windows number is reduced, but every particular lattice-focal lens becomes undersampled and does not cover the full depth range.*

Optimal subsquare size: According to Claim 3, the expected power spectrum of a lattice-focal lens increases with window size ε (Fig. 9). For larger subsquares the sinc blur around the central focal segment is narrower, so more energy is concentrated on the focal segment. However, it is clear that we cannot make ε arbitrarily large. When the number of subsquares is small, the expected power spectrum is high, but there are not enough samples to cover the full focal segment (Figure 9(a)). On the other hand, when the number of subsquares is too large, every subsquare has wide support around the main focal segment, leading to lower energy on the focal segment (Fig. 9(c)).

Posed another way, each subsquare is focused at a different point in the depth range, and provides reasonable coverage over the subrange of depths for which it achieves a defocus diameter of less than 1 pixel (Fig. 9, rightmost column). The subsquares' arrangement is undersampled if the minimum defocus diameter for some depth range is above 1 pixel, and redundant when the subsquares' effective depth coverage overlap. In the optimal arrangement each depth is covered by exactly one subsquare.

We derive the minimal number of windows providing full coverage of the depth of field, resulting in an optimal ε^*.

Claim 4 *The maximal subsquare size which allows full spectrum coverage is*

$$\varepsilon^* = (AS\Omega)^{-1/3} \,. \tag{26}$$

Proof: If the spacing between spatial samples is Δ the maximal frequency we need to be concerned with is $\Omega S/2 = S/(4\Delta)$. For window size ε we obtain $1/\varepsilon^2$ subsquares. If the slopes of the subsquares are equally spaced over the range $[-S/2,S/2]$, the spacing between samples in the frequency domain is $\tau = \Omega S\varepsilon^2$. Using subsquares of width εA, we convolve the samples with $\mathrm{sinc}(\varepsilon A\omega_x)\mathrm{sinc}(\varepsilon A\omega_y)$. For full coverage, we thus require $\varepsilon A \le 1/\tau$, implying:

$$S\Omega\varepsilon^2 \le \frac{1}{\varepsilon A} \quad\Rightarrow\quad \varepsilon \le (AS\Omega)^{-1/3} \,. \tag{27}$$

□

If we plug the optimal ε^* from Eq. (26) into Eq. (20) we conclude that the expected power spectrum of a lattice-focal lens with optimal window size is:

$$\mathrm{E}[|\hat{\phi}_s(\omega_x,\omega_y)|^2] \approx \frac{A^{8/3}}{S^{4/3}\Omega^{1/3}|\omega_{x,y}|}\beta(\omega_{x,y})\ . \qquad (28)$$

Discussion of lens spectra: The lattice-focal lens with an optimal window size achieves the highest power spectrum (i.e., closest to the upper bound) among all computational imaging designs listed in Table 2. While the squared MTFs for wavefront coding and focus sweep fall off *quadratically* as a function of $\omega_{x,y}$, for the lattice-focal lens the squared MTF only falls off *linearly*. Furthermore, while the squared MTFs for wavefront coding and focus sweep scale with A^2, for the lattice-focal lens the squared MTF scales with $A^{8/3}$. Still, there exists a gap of $(AS\Omega)^{1/3}$ between the power spectrum of the lattice-focal lens and the upper bound. It should be noted that the advantage of the lattice-focal lens is asymptotic and is most effective for large depth of field ranges. When the depth range of interest is small the difference is less noticeable, as demonstrated below.

Compact support in other designs: From the above discussion, the aperture area should be divided more or less equally into elements focused at different depths. However, beyond equal area we also want the aperture regions focused at each depth to be grouped together. Eq. (20) indicates that the expected power spectrum is higher if we use few wide windows, rather than many small ones. This can shed some light on earlier multi-focus designs. For example, [George and Chi 2003] use annular focus rings, and [Ben-Eliezer et al. 2005] use multiplexed subsquares, but multiple non-adjacent subsquares are assigned the same focal length. In both cases, the support of the aperture area focused at each depth is not at all compact, leading to sub-optimal MTFs.

5 Experiments

We first perform a synthetic comparison between extended depth of field approaches. We then describe a prototype construction of the lattice-focal lens and demonstrate real extended-DOF images.

5.1 Simulation

We start with a synthetic simulation using spatially-invariant first order (paraxial) optics. The OTFs in this simulation are computed numerically with precision, and do not rely on the approximate formulas in Table 2 .

Our simulation uses $A = 1000\Delta$ and considers two depth of field ranges given by $S = 2$ and $S = 0.1$. Assuming a planar scene, we synthetically convolved an image with the PSF of each design adding i.i.d. Gaussian noise with standard deviation $\eta = 0.004$. Non-blind deconvolution was performed using Wiener filtering and the results are visualized in Figures 10 and 11. We set the free parameters of each design to best match the depth range—for example, we adjust the parabola width a (in Eq. (5)), and select the optimal subsquare size of the lattice-focal lens. The standard and coded lenses were focused at the middle of the depth range, at $s_0 = 0$.

In Figure 10 we simulate the effect of varying the depth of the object. Using cameras tuned for depth range $S = 2$, we positioned the planar object at $s = 0$ (Fig. 10, top row) and $s = -0.9$ (Fig. 10, bottom row). As expected, higher spectra improve the visual quality of the deconvolution. Standard and coded lenses obtain excellent reconstructions when the object is positioned at the focus slope $s = 0$, but away from the focus depth the image deconvolution cannot recover much information. Focus sweep, wavefront coding and the lattice-focal lens achieve uniform reconstruction quality across depth. The best reconstruction is obtained by our lattice-focal PSF, followed by wavefront coding, then focus sweep. Note that since we use a square aperture, several imaging systems have more horizontal and vertical frequency content. This leads to horizontal and vertical structure in the reconstructions of Figure 10, particularly noticeable in the standard lens and the wavefront coding results.

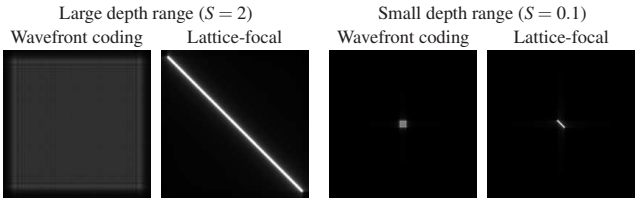


Figure 12: *ω_{x_0,y_0}-slice (at $\omega_{x_0} = 0.9\Omega, \omega_{y_0} = -0.9\Omega$) for two depth ranges defined by slope bounds $S = 2$ (left) and $S = 0.1$ (right). For the smaller range, the difference between the focal segment and the full bounding square is lower, and the spectra for wavefront coding and the lattice-focal lens are more similar.*

In Figure 11 we simulate the effect of varying the depth range. The planar object was positioned at $s = -0.5$, and the camera parameters were adjusted to cover a narrow depth range $S = 0.1$ (Fig. 11, top row) and a wider range $S = 2$ (Fig. 11, second row). When the focus sweep, wavefront coding and lattice-focal lens are adjusted to a narrower depth range their performance significantly improves, since they now distribute the same budget over a narrower range.

The difference between the designs becomes more critical when the depth range is large. Figure 12 visualizes a ω_{x_0,y_0}-slice for both S values. For $S = 0.1$, the length of the focal segment is so short that there is little difference between the segment and its bounding square. Thus, with a smaller depth range the wavefront coding lens incurs less of penalty for spending its budget on afocal regions.

Mapping slope ranges to physical distances: Assume that the camera has sensor resolution $\Delta_0 = 0.007$mm, and that we use an $f = 85$mm focal length lens focused at depth $d_o = 70$cm. This depth also specifies the location of the xy light field plane. The DOF is defined by the range $[d_{min}, d_{max}]$ corresponding to slopes $\pm S/2$. From Eq. (2), the depth range can be expressed as $d_o/(1 \pm S/2)$, yielding a DOF of $[35,\infty]$cm for $S = 2$ and $[66.2, 74.3]$cm for $S = 0.1$. The pixel size in the light field is $\Delta = \Delta_0/M$, where $M = f/(d_o - f) = 0.13$ is the magnification. We set the effective aperture size A to $1000\Delta = 1000\Delta_0/M = 50.6$mm, which corresponds to $f/1.68$. The subsquares number and focal lengths are selected such that for each point in the depth range, there is exactly one subsquare achieving defocus diameter of less than one pixel. The subsquare number is given by Eq. (26), in this simulation $m = 100$ aperture subsquares with $S = 2$, and $m = 16$ subsquares with $S = 0.1$. To set the focal lengths of each subsquare we select m equally spaced slopes s_j in the range $[-S/2, S/2]$. A slope s_j is mapped to a physical depth d_j according to Eq. (2). To make the j-th subsquare focus at depth d_j we select its focal length f_j according to the Gaussian lens formula: $1/f_j = 1/d_j + 1/d_s$ (where d_s denotes the sensor-to-lens distance).

5.2 Implementation

Hardware construction: To demonstrate our design we have built a prototype lattice-focal lens. Our construction provides a proof of concept showing that a lattice-focal lens can be implemented in practice and lead to reasonably good results, however it is not an optimized or fully-characterized system.

As shown in Figure 13, our lattice-focal lens mounts to a main lens using the standard threaded interface for a lens filter. The subsquares of the lattice-focal lens were cut from BK7 spherical plano-convex lens elements using a computer-controlled saw. The squares are of size 5.5×5.5mm and thickness 3mm. By attaching our

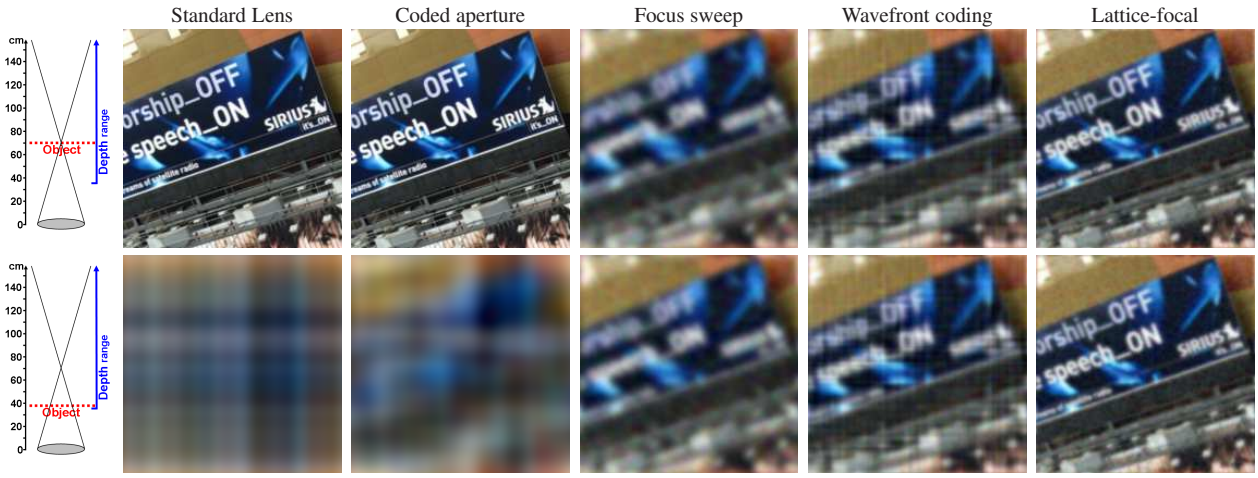


Figure 10: *Synthetic comparison of image reconstruction at different object depths Top row: object depth $s = 0$, Bottom row: object depth $s = -0.9$ Standard and coded lenses produce high quality reconstruction for an object at the focus depth, but a very poor one away from the focus depth. Focus sweep, wavefront coding and the lattice focal lens perform equally across depth. The highest quality reconstruction produced by our lattice-focal lens.*

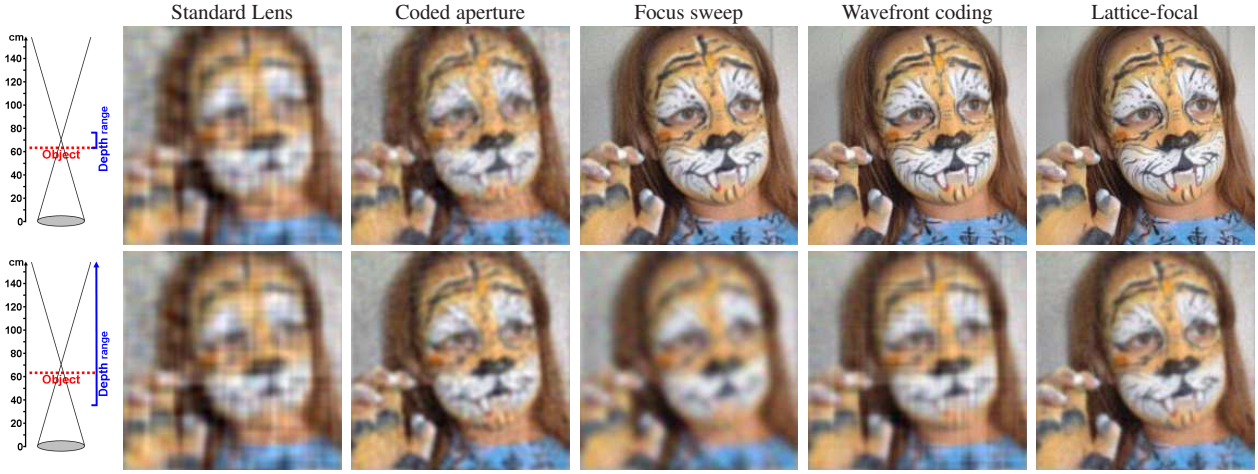


Figure 11: *Synthetic comparison of image reconstruction when camera parameters are adjusted for different depth ranges. Top row: narrow depth range bounded by $S = 0.1$, Bottom row: wider range bounded by $S = 2$. Most designs improve when they attempt to cover a narrower range. The difference between the designs is more drastic at large depth ranges.*

lattice-focal lens to a high-quality main lens (Canon 85mm f1.2L), we reduce aberrations. Since most of the focusing is achieved by the main lens, our new elements require low focal powers, and correspond to very low-curvature surfaces with limited aberrations (in our prototype, the subsquare focal lengths varied from 1m to 10m).

In theory the lattice-focal element should be placed in the plane of the main lens aperture or at one of its images, e.g., the entrance or exit pupils. To avoid disassembling the main lens to access these planes, we note that a sufficiently narrow stop in front of the main lens redefines a new aperture plane. This lets us attach our lattice-focal lens at the front, where the stop required to define a new aperture still let us use 60% of the lens diameter.

The minimal subsquare size is limited by diffraction. Since a normal lens starts being diffraction-limited around an $f/12$ aperture [Goodman 1968], we can fit about 100 subsquares within an $f/1.2$ aperture. To simplify the construction, however, our prototype included only 12 subsquares. The DOF this allowed us to cover was small and, as discussed in Sec. 5.1, in this range the lattice-focal lens advantage over wavefront coding is limited. Still, our prototype demonstrates the effectiveness of our approach.

Given a fixed budget of m subsquares of a given width, we can invert the arguments in Sec. 4 and determine the DOF it can cover in the optimal way. As discussed at the end of Sec. 5.1 and illustrated in Figure 9(b), for every point in the optimal DOF, there is exactly one subsquare achieving defocus diameter of less than 1 pixel. This constraint also determines the focal length for each of these subsquares. For our prototype we focused the main lens at 180cm and chose subsquare focal lengths covering a depth range of $[60, 180]$cm. Given the limited availability of commercial plano-convex elements, our subsquares' coverage was not perfectly uniform, and we used focal lengths of 10000,5000,4000,3000,2500,2000,1750,1500,1300,1200,1000mm, plus one flat subsquare (infinity focal length). However, for a custom-manufactured lens this would not be a limitation.

Calibration: To calibrate the lattice-focal lens, we used a planar white noise scene and captured a stack of 30 images for different depths of the scene. Given a blurred and sharp pair of images B_d, I_d at depth d, we solved for the kernel ϕ_d minimizing $|\phi_d \otimes I_d - B_d|$. We show the recovered PSF at 3 depths in Figure 13. As discussed in Sec. 4, the PSF is a superposition of boxes of varying sizes, but

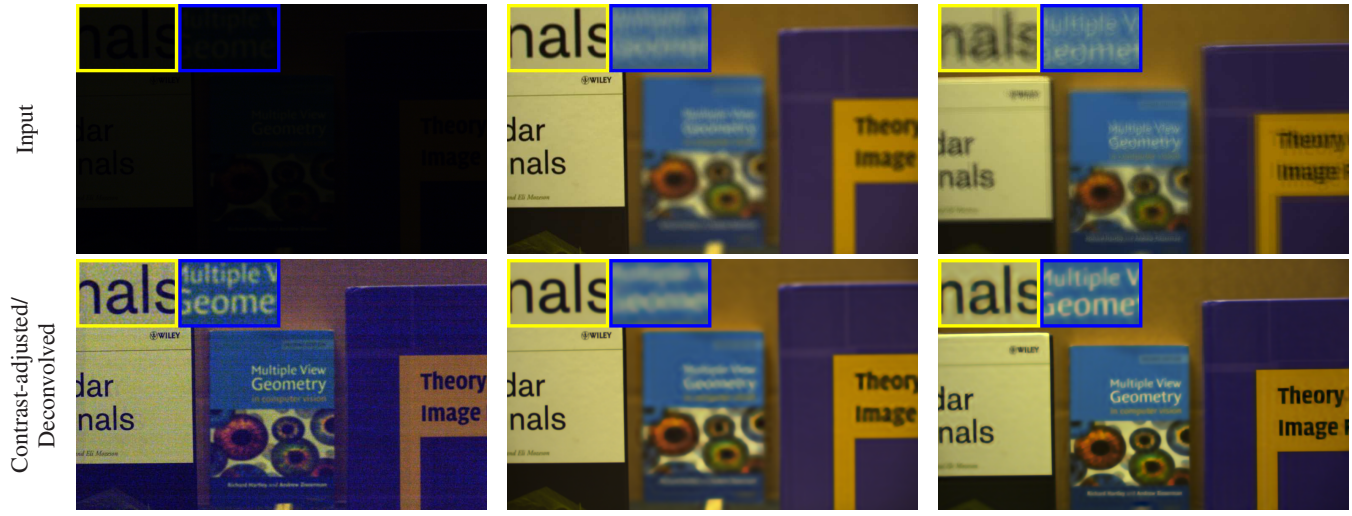


Figure 14: *Comparison between a lattice-focal lens and a standard lens, both for a narrow aperture ($f/16$) and for the same aperture size as our lattice-focal lens prototype ($f/4$). All photos were captured with equal exposure time, so the $f/16$ image is very noisy. The standard $f/4$ image is focused at the white book, but elsewhere produces a defocused image. The lattice-focal output is sharper over the entire scene.*

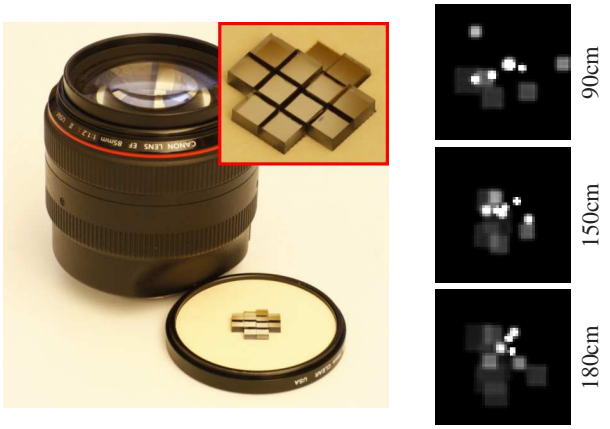


Figure 13: *Our prototype lattice-focal lens and PSFs calibrated at three depths. The prototype attaches to the main lens like a standard lens filter. The PSFs are a sum of box filters from the different subsquares, where the exact box width is a function of the deviation between the subsquare focal depth and the object depth.*

the exact arrangement of boxes varies with depth. For comparison, we did the same calibration using a standard lens as well.

Depth estimation: Given the calibrated per-depth PSFs, we deblur an image using sparse deconvolution [Levin et al. 2007]. This algorithm computes the latent image I_d as

$$I_d = \arg\min_I |\phi_d \otimes I - B|^2 + \lambda \sum_i \left[\rho(g_{x,i}(I)) + \rho(g_{y,i}(I))\right] , \quad (29)$$

where $g_{x,i}, g_{y,i}$ denote horizontal and vertical derivatives of the i-th pixel, ρ is a robust function, and λ is a weighting coefficient.

Since the PSF varies over depth, rough depth estimation is required for deblurring. If an image region is deconvolved with a PSF corresponding to the incorrect depth, the result will include ringing artifacts. To estimate depth, we start by deconvolving the entire image with the stack of all depth-varying PSFs, and obtain a stack of candidate deconvolved images $\{I_d\}$. Since deconvolution with the wrong PSF leads to convolution error, we can locally score the explanation provided by PSF ϕ_d around pixel i as:

$$E_i(d) = |B_i - \tilde{B}_{d,i}|^2 + \lambda \left[\rho(g_{x,i}(I_d)) + \rho(g_{y,i}(I_d))\right] , \quad (30)$$

where $\tilde{B}_d = \phi_d \otimes I_d$[3]. We regularize the local depth scores using a Markov random field (MRF), then generate an all-focus image using the Photomontage algorithm of Agarwala et al. [2004].

Results: In Figure 14 we compare the reconstruction using our lattice-focal lens with a standard lens focused at the middle of the depth range (i.e., the white book). Using a narrow aperture ($f/16$), the standard lens produces a very noisy image, since we held exposure time constant over all conditions. Using the same aperture size as our prototype ($f/4$), the standard lens resolves a sharp image of the white book, but the rest of the scene is defocused. For the purpose of comparison, we specified the depth layers manually and deconvolved both the standard and lattice-focal images with PSFs corresponding to the true depth. Because the spectrum of the lattice-focal lens is higher than a standard lens across the depth range, greater detail can be resolved after deconvolution.

Figure 15 shows all-focus images and depth maps captured using our lattice-focal lens. More results are available online[4]. Since the MRF of Agarwala et al. [2004] seeks invisible seams, the layer transitions usually happen at low-texture regions and not at the actual contours. Despite the MRF's preference for piecewise-constant depth structures we handle continuous depth variations, as shown in the rightmost column of Figure 15.

The results in Figure 15 were obtained fully automatically. However, depth estimation can fail, especially next to occlusion boundaries, which present a general problem for all computational extended-DOF systems [Dowski and Cathey 1995; Nagahara et al. 2008; Levin et al. 2007; Veeraraghavan et al. 2007]. While a principled solution to this problem is beyond the scope of this paper, most artifacts can be eliminated with simple manual layer refinement.

[3] Note that despite the discussion in [Levin et al. 2009], we employ a $\text{MAP}_{x,k}$ approach that scores a depth d based on the best I_d explanation alone. The reason this approach works here is that a delta explanation is absent from the search space, and there is a roughly equal volume of solutions around all PSFs ϕ_d.

[4] www.wisdom.weizmann.ac.il/~levina/papers/lattice

Figure 15: *Partially defocused images from a standard lens, compared with an all-focused image and depth map produced by the lattice-focal lens.*

Figure 16: *Synthetic refocusing using the coarse depth map estimated with the lattice-focal lens.*

Relying on depth estimation to decode an image from a lattice-focal lens is a disadvantage compared to depth-invariant solutions, but it also allows coarse depth recovery. In Figure 16 we used the rough depth map to synthetically refocus a scene post exposure.

6 Discussion

This paper analyzes extended depth of field systems in light field space. We show that while effective extended DOF systems seek high spectrum content, the maximal possible spectrum is bounded. The dimensionality gap between the 4D light field and the 3D focal manifold is a key design factor, and to maximize spectrum content lenses should concentrate their energy in the focal manifold of the light field spectrum. We analyze existing computational imaging designs and show that some do not follow this principle, while others do not achieve a high spectrum over the depth range. Guided by this analysis we propose the lattice-focal lens, accounting for the dimensionality gap. This allows us to achieve defocus PSFs with higher spectra compared to previous designs.

However, the lattice-focal lens does not fully achieve the upper bound. One open question is whether better designs exist, whether the upper bound could be tighter, or both. Our intuition is that the upper bound could be tighter. The proof of Claim 2 is based on the assumption that an $A \times A$ primal support is devoted to every frequency point. However, the fact that the integration surface has to "cover" a full family of slopes implies that the aperture area has to be divided between all slopes. Thus the primal support of each slope is much smaller than A, which implies a wider frequency support around the focal segment, reducing the height of the spectrum on the focal segment itself.

We have focused on spectra magnitude, which dominates the deconvolution quality. However, the accuracy of depth estimation is important as well. Wavefront coding and focus sweep cameras have an important advantage that they bypass the need to estimate depth. On the other hand, the lattice-focal lens has the benefit of recovering a rough depth map in addition to an all-focused image. One future research question is whether the higher spectrum of the lattice-focal lens can also be achieved with a depth-invariant design.

Acknowledgments: We thank the Israel Science Foundation, the Royal Dutch/Shell Group, NGA NEGI-1582-04-0004, MURI Grant N00014-06-1-0734, NSF CAREER award 0447561. F. Durand acknowledges a Microsoft Research New Faculty Fellowship and a Sloan Fellowship. S. Hasinoff acknowledges the NSERC PDF program.

Appendix A: Spectra derivations

Below we complete the budget and spectra derivation of Sec. 3.

Claim 5 *For an aperture of size $A \times A$ and exposure length* 1*, the total energy in each ω_{x_0,y_0}-slice is bounded by A^2:*

$$\iint |\hat{k}_{\omega_{x_0,y_0}}(\omega_u, \omega_v)|^2 d\omega_u d\omega_v \leq A^2 \,. \tag{31}$$

Proof: The proof reviews the budget proof in [Levin et al. 2008c]. Note that $\hat{k}_{\omega_{x_0,y_0}}(\omega_u,\omega_v)$ is the 2D Fourier transform of:

$$\iint k(x,y,u,v)e^{-2i\pi(\omega_{x_0}x+\omega_{y_0}y)}dxdy\ . \tag{32}$$

For every clear aperture point $|u|\le A/2,\ |v|\le A/2$:

$$\left|\iint k(x,y,u,v)e^{-2i\pi(\omega_{x_0}x+\omega_{y_0}y)}dxdy\right|^2 \le \left|\iint k(x,y,u,v)dxdy\right|^2 \le 1\ . \tag{33}$$

Where the first inequality follows from the fact that a phase change does not increase magnitude, and the second inequality follows from the unit energy through every clear aperture point (see also Eqs. (13) and (14)).

Since the aperture size is A^2, the total norm is bounded by A^2:

$$\iint\left|\iint k(x,y,u,v)e^{-2i\pi(\omega_{x_0}x+\omega_{y_0}y)}dxdy\right|^2 dudv \le A^2\ . \tag{34}$$

By Parseval's theorem, the square integral is the same in the dual and the primal domains, thus:

$$\iint|\hat{k}_{\omega_{x_0,y_0}}(\omega_u,\omega_v)|^2 d\omega_u d\omega_v \le A^2\ . \tag{35}$$

□

Standard lens:

Claim 6 *The power spectrum of a standard lens focused at depth s_0 with aperture $A\times A$ is*

$$|\hat{\phi}_s(\omega_x,\omega_y)|^2 = A^4\mathrm{sinc}^2(A(s-s_0)\omega_x)\mathrm{sinc}^2(A(s-s_0)\omega_y)\ . \tag{36}$$

Proof: A lens focused at slope s_0 is modeled by a linear integration surface $c(u,v)=(s_0u,s_0v)$. If the surface were infinite, the Fourier transform would be an impulse at $\omega_u=-s_0\omega_x,\omega_v=-s_0\omega_y$. Given the finite aperture we need to convolve that with a sinc, and thus

$$\hat{k}(\omega_x,\omega_y,\omega_u,\omega_v) = A^2\mathrm{sinc}(A(\omega_u-s_0\omega_x))\mathrm{sinc}(A(\omega_v-s_0\omega_y))\ . \tag{37}$$

Eq. (36) follows by selecting an OTF-slice. □

The ω_{x_0,y_0}-slices in Figure 4(b) reveals a sinc around the point $\omega_u=-s_0\omega_x,\omega_v=-s_0\omega_y$. Note that reducing the aperture size A increases the sinc width and minimizes defocus blur. However, given a fixed exposure length it also reduces the amount of light collected, which reduces the MTF. Indeed, the sinc height in Eq. (36) decreases for smaller A values.

Coded aperture: A coded aperture is constructed with a standard lens, w.l.o.g. focused at slope $s_0=0$. We construct a coded aperture by dividing the aperture into squares of size $\varepsilon A\times\varepsilon A$ and randomly blocking each subsquare with probability $1/2$. The expected power spectrum can then be computed analytically.

Claim 7 *For a lens focused at $s_0=0$, the expected power spectrum of a random coded aperture with holes size $\varepsilon A\times\varepsilon A$ is*

$$\mathrm{E}[|\hat{\phi}_s(\omega_{x,y})|^2] \approx \frac{\varepsilon^2A^4}{2}\mathrm{sinc}^2(\varepsilon As\omega_x)\mathrm{sinc}^2(\varepsilon As\omega_y)\ . \tag{38}$$

Proof: We express $\hat{k}=\sum_j\hat{k}^j$ where $\hat{k}^j$ is the 4D spectrum of an individual subsquare. For an unblocked hole centered at u_0,v_0 we can express $\hat{k}^j$ analytically as the transform of a box times a phase shift (resulting from the translation of the subsquare center):

$$\hat{k}^j(\omega_x,\omega_y,\omega_u,\omega_v) = \varepsilon^2A^2e^{-2i\pi(\omega_uu_0+\omega_vv_0)}\mathrm{sinc}(\varepsilon A\omega_u)\mathrm{sinc}(\varepsilon A\omega_v)\ . \tag{39}$$

As in the proof of Claim 3, we note that $\mathrm{E}[\hat{k}^j]$ affects very low frequencies only, so we use Eq. (21) to approximate

$$\begin{aligned}\mathrm{E}[|\hat{k}|^2] &\approx \frac{1}{\varepsilon^2}\mathrm{E}[|\hat{k}^j|^2] &(40)\\ &= \frac{\varepsilon^2A^4}{2}\mathrm{sinc}^2(\varepsilon A\omega_u)\mathrm{sinc}^2(\varepsilon A\omega_v)\ , &(41)\end{aligned}$$

where the number of subsquares is $1/\varepsilon^2$ and the factor $1/2$ represents the probability of a blocked subsquare. By selecting an OTF-slice, Eq. (38) follows. □

Eq. (41) suggests that, ignoring diffraction, the sensor spatial resolution implies a tradeoff in selecting the optimal hole size. If we use small holes, the power spectrum of the aperture is wider, and wider spectrum implies that more budget is spread away from the main focal segment (indeed Eq. (38) shows that the expected spectrum is multiplied by ε and decreases when ε is small). On the other hand, the expected power spectrum of $\hat{k}$ falls off like $\mathrm{sinc}^2(\varepsilon A\omega_u)\mathrm{sinc}^2(\varepsilon A\omega_v)$. That is, since the lens is focused only at a single depth, to have spectral content at slopes corresponding to other depths, the spectrum of the aperture must be sufficiently wide, implying that a small hole size ε is needed.

Focus sweep:

Claim 8 *For $|\omega_x|,|\omega_y|>(SA)^{-1}$, the power spectrum of the focus sweep camera asymptotically approaches*

$$|\hat{\phi}_s|^2 \approx \frac{A^2\alpha(\omega_{x,y})^2}{S^2|\omega_{x,y}|^2}\ , \tag{42}$$

where $\alpha(|\omega_{x,y}|)$ is a bounded multiplicative factor in the range $[1,\sqrt{2}]$:

$$\alpha(|\omega_{x,y}|) = \frac{|\omega_{x,y}|}{\max(|\omega_x|,|\omega_y|)}\ . \tag{43}$$

Proof: The spectrum of a standard lens focused at slope s_0 is

$$A^2\mathrm{sinc}(A\omega_x(s_0-s))\mathrm{sinc}(A\omega_y(s_0-s))\ . \tag{44}$$

The spectrum of a focus sweep is obtained by averaging Eq. (44) over s_0. To compute this integral we make use of the following identity: for a 2D vector $r=[r_1,r_2]$,

$$\int_{-\infty}^{\infty}\mathrm{sinc}(r_1t)\mathrm{sinc}(r_2t)dt = \frac{\alpha(|r|)}{|r|}\ . \tag{45}$$

If $-S/2<s<S/2$ and S large enough[5], we use Eq. (45) and get:

$$\begin{aligned}\hat{\phi}_s(\omega_{x,y}) &= \frac{A^2}{S}\int_{-S/2}^{S/2}\mathrm{sinc}(A\omega_x(s_0-s))\mathrm{sinc}(A\omega_y(s_0-s))ds_0\\ &= \frac{A^2}{S}\int_{-S/2+s}^{S/2+s}\mathrm{sinc}(A\omega_xs_0)\mathrm{sinc}(A\omega_ys_0)ds_0\\ &\to \frac{A\alpha(\omega_{x,y})}{S|\omega_{x,y}|}\ . &(46)\end{aligned}$$

Taking the power of Eq. (46) provides Eq. (42). □

[5]The approximation is reasonable for $|\omega_x|,|\omega_y|>(SA)^{-1}$.

Figure 4(d) displays ω_{x_0,y_0}-slices from the power spectrum of a focus sweep camera. On one hand, this spectrum is concentrated around the main focal segment, with the same narrow width achieved by the upper bound in Fig. 4(a). However, the magnitude of the focus sweep is significantly lower. In fact, the total energy at every ω_{x_0,y_0}-slice is much lower than the budget of Claim 5, that is:

$$\iint |\hat{k}_{\omega_{x_0,y_0}}(\omega_u,\omega_v)|^2 d\omega_u d\omega_v \ll A^2 \ . \tag{47}$$

To understand why, recall that the upper bound in Claim 5 is obtained by noting that when x,y are integrated, the magnitude of the projection integral is bounded by 1 (Eq. (33)). And indeed, when the 4D lens kernel is a delta function of u,v, that integral is equal to 1. By contrast, the effective 4D kernel for a focus sweep camera is the average of standard-lens 4D kernels over all depths, and therefore is not a delta function. When such a non-delta kernel is multiplied by a wave of the form $e^{-2i\pi(\omega_x x+\omega_y y)}$, interference and phase cancelations significantly reduce the magnitude of the integral, and Eq. (33) is far below 1.

Wavefront coding:

Claim 9 *For a slope $s \in [-S/2,S/2]$, the power spectrum of a wavefront coding lens asymptotically approaches*

$$|\hat{\phi}_s(\omega_x,\omega_y)|^2 \approx \frac{A^2}{S^2|\omega_x||\omega_y|} \ . \tag{48}$$

Proof: A wavefront coding lens is a cubic refractive element (as reported in [Dowski and Cathey 1995]). From Snell's law, the integration surface is determined by the lens normal. Therefore the integration surface is a separable parabola $c(u,v) = (au^2, av^2)$. The parabola width a can be set such that the parabola slope covers the slope range of interest $[-S/2,S/2]$, implying $a = S/(2A)$. The power spectrum of a 1D parabola as computed in [Levin et al. 2008c] is

$$|\hat{k}(\omega_x,\omega_u)|^2 \approx \frac{A}{S|\omega_x|}\delta_{|\omega_u|<S/2|\omega_x|} \ . \tag{49}$$

The 2D parabola case is a separable extension:

$$|\hat{k}(\omega_x,\omega_y,\omega_u,\omega_v)|^2 \approx \frac{A^2}{S^2|\omega_x||\omega_y|}\delta_{|\omega_u|<S/2|\omega_x|}\delta_{|\omega_v|<S/2|\omega_y|} \ . \tag{50}$$

If $s \in [-S/2,S/2]$, we can slice Eq. (50) to get Eq. (48). □

Appendix B: Cameras with multiple measurements

Many imaging models in the literature, like a focal stack, or a plenoptic camera involve multiple measurements per pixel. Analyzing these cameras requires additional assumptions about the noise model. One possible analysis is described below.

We consider additive Gaussian noise at the sensor and Wiener reconstruction, with a Gaussian zero-mean prior on the signal. That is, we assume N different measurements of each spatial image frequency, and denote the measurement j as

$$\hat{B}_j(\omega_{x,y}) = \hat{\phi}_j(\omega_{x,y})\hat{I}(\omega_{x,y}) + n_j \ . \tag{51}$$

We denote the noise variance by η^2, and the signal $\hat{I}(\omega_{x,y})$ variance with $\sigma^2_{\omega_{x,y}}$. The N-measurement Wiener estimate $\hat{I}_{\text{est}}(\omega_{x,y})$ is (see, for example, [Levin et al. 2009]):

$$\begin{aligned}\hat{I}_{\text{est}}(\omega_{x,y}) &= \arg\min \sum_j \frac{1}{\eta^2}|\hat{\phi}_j(\omega_{x,y})\hat{I}(\omega_{x,y}) - \hat{B}_j(\omega_{x,y})|^2 + \frac{|\hat{I}(\omega_{x,y})|^2}{\sigma^2} \\ &= \frac{\frac{1}{\eta^2}\sum_j \hat{\phi}_j(\omega_{x,y})^* \hat{B}_j(\omega_{x,y})}{\frac{1}{\eta^2}\sum_j |\hat{\phi}_j(\omega_{x,y})|^2 + \frac{1}{\sigma^2}} \ .\end{aligned} \tag{52}$$

One can also compute the expected reconstruction error (expectation over all $\hat{I}(\omega_{x,y}), n_j$ samples from the prior):

$$E[|\hat{I}_{\text{est}}(\omega_{x,y}) - \hat{I}(\omega_{x,y})|^2] = \left(\frac{1}{\eta^2}\sum_j |\hat{\phi}_j(\omega_{x,y})|^2 + \frac{1}{\sigma^2}\right)^{-1} \ . \tag{53}$$

Eq. (53) states that the reconstruction error is a function of the sum of power spectra of the individual OTFs. In the following analysis we evaluate two multiple measurement configurations, the focal stack and the plenoptic camera, by examining the summed power spectrum $\sum_j |\hat{\phi}_j|^2$.

Note that this suggests how to merge multiple independent measurements of the *same* spatial frequency, and should not be confused with summing multiple elements of a single measurement. All other cameras considered in the main body of this paper have a single measurement per spatial frequency.

Focal stack: With additive noise and Wiener filtering, the reconstruction error for a focal stack depends on the summed power spectra of all images in the stack. Suppose we divide exposure time into N bins and spread the focus setting of all time bins evenly over the depth range. The OTF of a time bin j focused at slope s_j is

$$\hat{\phi}_s(\omega_{x,y}) = \frac{A^2}{N}\text{sinc}\left(A\omega_x(s_j-s)\right)\text{sinc}\left(A\omega_y(s_j-s)\right) \ . \tag{54}$$

By substituting $\varepsilon = 1$ into Eq. (25), and multiplying by $1/N^2$ to account for the shorter exposure, the integration over all slopes s_j provides

$$E[|\hat{\phi}_s(\omega_{x,y})|^2] = \frac{1}{N^2}\frac{A^3}{S|\omega_{x,y}|} \ . \tag{55}$$

Since we have N such bins, the total sum is

$$\sum_j |\hat{\phi}_{j,s}(\omega_{x,y})|^2 = \frac{1}{N}\frac{A^3}{S|\omega_{x,y}|} \ . \tag{56}$$

Note that the number of bins N does not cancel out in the integration. By adjusting Claim 4, the minimal number of bins required to cover the depth range without aliasing is $N = AS\Omega$, resulting in summed power spectrum of

$$\sum_j |\hat{\phi}_{j,s}(\omega_{x,y})|^2 = \frac{A^2}{S^2\Omega|\omega_{x,y}|} \ . \tag{57}$$

As we have seen in Eq. (53), the reconstruction error is a function of the summed power spectrum. From Eq. (57) the summed power spectrum of a focal stack is lower than the focus sweep spectrum of $A^2/(S^2|\omega_{x,y}|^2)$, since $|\omega_{x,y}| < \Omega$. Thus, under additive noise, the reconstruction error of a focus sweep camera which measures different focus depths simultaneously is better than the focal stack approach which measures each focus depth separately.

Plenoptic camera: The MTF of a plenoptic camera cannot be compared directly to the other cameras because the imaging model

is different and involves tradeoffs between spatial and angular resolution. As a result, assumptions about the sensor layout of the plenoptic camera must also be made. One possible analysis is described below, which concludes that reconstruction error for a plenoptic camera is higher than for a lattice-focal lens. Intuitively, the plenoptic camera captures the entire 4D light field and hence spends energy away from the focal manifold.

Assume other cameras use N sensor elements. To keep the same spatial resolution in the final image we assume the plenoptic camera has N lenslets and $1/\varepsilon^2$ elements below each lenslet, resulting in N/ε^2 sensor elements in total. We assume the lens is focused at the middle of the depth range, at slope $s_0 = 0$.

The OTF from the j-th directional view is:

$$\hat{\phi}_s^j(\omega_{x,y}) = e^{2\pi i(\gamma_{j,x}\omega_x + \gamma_{j,y}\omega_y)}\varepsilon^2 A^2 \mathrm{sinc}(\varepsilon A s\omega_x)\mathrm{sinc}(\varepsilon A s\omega_y)\ , \quad (58)$$

where $(\gamma_{j,x}, \gamma_{j,y})$ denotes the phase shift of the j-th view.

Recall that the reconstruction error in Eq. (53) is a function of the sum of power spectra of the individual OTFs. For the case of a plenoptic camera,

$$\sum_j |\hat{\phi}_j(\omega_{x,y})|^2 = \varepsilon^2 A^4 \mathrm{sinc}^2(\varepsilon A s\omega_x)\mathrm{sinc}^2(\varepsilon A s\omega_y) \quad (59)$$

since there are $1/\varepsilon^2$ terms in the sum.

This suggests that there is a tradeoff in the number of directional views: more directional views increase the noise contribution and the reconstruction error. By contrast, if we do not divide the directional views finely enough we have defocus in the depth range.

To avoid defocus over the depth range, the worst-case slope S and the worst-case frequency Ω should be included in the first sinc lobe, hence

$$\varepsilon \leq \frac{1}{AS\Omega}\ . \quad (60)$$

This implies that

$$\sum_j |\hat{\phi}_j(\omega_{x,y})|^2 \leq \frac{A^2}{S^2\Omega^2} \quad (61)$$

at *all* spatial frequencies $\omega_{x,y}$, and the plenoptic camera will have increased noise contribution compared to most single-measurement cameras we have analyzed.

Appendix C: Fourier optics derivation

Our goal in this appendix is to derive an analogy to the light field results, while modeling wave optics effects and not only geometric optics. We use a traditional optics representation- the Wigner distribution [Wigner 1932], and build on the tight relation between light fields and Wigner distribution derived by Zhang and Levoy [2009].

We will show that an analogy to the budget bounds exist, but while the geometric optics bound allocates an equal budget at all spatial frequencies, Fourier optics implies that the bound drops as the diffraction limited OTF. Yet, if the system dimensions are kept well above the diffraction limit, the approximation provided by the geometric model is reasonable enough.

In our derivation we consider a lens with an $A \times A$ aperture and focal length f, we denote the distance from the lens to the camera sensor by d_i and the focus depth by d_o. By the Gaussian lens formula,

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}. \quad (62)$$

This lens is then expressed as

$$R(Au)R(Av)e^{-i\pi\frac{u^2+v^2}{\lambda f}} \quad (63)$$

where R is a rect function and λ is the wavelength. A general refractive element is expressed relative to this lens as

$$R(Au)R(Av)e^{-i\pi\frac{u^2+v^2}{\lambda f}}\varphi(u,v) \quad (64)$$

where $\varphi(u,v)$ is the aberration, or the deviation from a normal lens behavior. We also denote by $\varphi_A(u,v) = R(Au)R(Av)\varphi(u,v)$ the aberration multiplied by the aperture.

In our derivation we study the classical ambiguity function, the Fourier dual of the Wigner distribution. The ambiguity function of φ_A is defined as

$$\mathscr{A}(u,v,s_u,s_v) = \iint \xi_{u,v}(t_1,t_2)e^{i2\pi(s_u t_1 + s_v t_2)}dt_1 dt_2. \quad (65)$$

where

$$\xi_{u,v}(t_1,t_2) = \varphi_A(t_1+u/2, t_2+v/2)\varphi_A^*(t_1-u/2, t_2-v/2). \quad (66)$$

We now define a 4D analog of the 4D lens spectrum as

$$\tilde{\hat{k}}(\omega_x,\omega_y,\omega_u,\omega_v) = \mathscr{A}(\rho\omega_x, \rho\omega_y, \omega_u, \omega_v) \quad (67)$$

where $\rho = \lambda d_o$.

It is well known in optics that OTFs can be obtained as slices from the ambiguity function [Brenner et al. 1983]. Below we provide a revised derivation of this result adjusted to our notation. That is, we show that the Fourier optics OTF is a slice through $\tilde{\hat{k}}$ exactly in the same way that the geometric optics OTF is a slice through $\hat{k}$. We later show that we can place upper bounds on the norm of $\tilde{\hat{k}}$ in the same way we have specified bounds the norm of $\hat{k}$.

Claim 10 *Consider an object at depth d_a and its light field slope $s = (d_a - d_o)/d_a$. The Fourier optics OTF $\hat{\phi}_s(\omega_{x,y})$ is a slice through $\tilde{\hat{k}}$:*

$$\hat{\phi}_s(\omega_{x,y}) = \tilde{\hat{k}}(\omega_x,\omega_y,-s\omega_x,-s\omega_y). \quad (68)$$

Proof: If we are imaging an on axis point at distance d_a, the wavefront reaching the lens is

$$e^{i2\pi\frac{d_a}{\lambda}}e^{i\pi\frac{u^2+v^2}{\lambda d_a}} \quad (69)$$

The wavefront leaving the lens is a product of the lens with the wave reaching the lens.

A wavefront reaching a point (x',y') on the image plane (the image plane is located at distance d_i from the lens) is thus the product of the wave leaving the lens with a spherical wavefront centered at

(x',y',d_i), which reduces to

$$\begin{aligned}
&h(x',y')\\
&= \iint \varphi_A(u,v) e^{-i\pi\frac{u^2+v^2}{\lambda f}} e^{i2\pi\frac{d_a}{\lambda}} e^{i\pi\frac{u^2+v^2}{\lambda d_a}} e^{i2\pi\frac{d_i}{\lambda}} e^{i\pi\frac{(u-x')^2+(v-y')^2}{\lambda d_i}} dudv\\
&= e^{i2\pi(d_i+d_a)/\lambda} e^{i\pi\frac{x'^2+y'^2}{\lambda d_i}}\\
&\quad \iint \varphi_A(u,v) e^{i\pi(1/d_a+1/d_i-1/f)\frac{u^2+v^2}{\lambda}} e^{-i2\pi\frac{x'u+y'v}{\lambda d_i}} dudv\\
&=^1 e^{i2\pi(d_i+d_a)/\lambda} e^{i\pi\frac{x'^2+y'^2}{\lambda d_i}}\\
&\quad \iint \varphi_A(u,v) e^{-i\pi s\frac{u^2+v^2}{\rho}} e^{-i2\pi\frac{x'u+y'v}{\lambda d_i}} dudv
\end{aligned} \tag{70}$$

where [1] follows from the Gaussian lens formula $1/d_o+1/d_i=1/f$, implying that the spherical wave coefficient above can be expressed as $1/\lambda(1/d_a+1/d_i-1/f)=1/\lambda(1/d_a-1/d_o)=-s/\rho$.

Now let

$$\psi_s(u,v) = \varphi_A(u,v) e^{-i\pi s\frac{u^2+v^2}{\rho}}, \tag{71}$$

and let Ψ_s denote its Fourier transform. Then

$$h(x',y') = \Psi_s\left(\frac{x'}{\lambda d_i}, \frac{y'}{\lambda d_i}\right) e^{i\pi\cdot const}. \tag{72}$$

To map x' to the spatial coordinates in the rest of this document, we note that we used coordinates on the xy-plane, and not on the image. That is, we need to divide by the magnification factor $M=d_i/d_o$, and transfer to the coordinate system $x=x'/M, y=y'/M$. That is:

$$h_s(x,y) = h_s(x'/M, y'/M) = \Psi_s\left(\frac{x}{\rho}, \frac{y}{\rho}\right) e^{i\cdot const} \tag{73}$$

The incoherent PSF is now equal to

$$\phi_s(x) = |h_s(x)|^2 = |\Psi_s(x/\rho)|^2 \tag{74}$$

The OTF $\hat{\phi}_s$ can thus be obtained from the auto correlation of ψ, scaling the coordinates by ρ, which reduces to the definition of $\tilde{\tilde{k}}$ in Eq. (67), since:

$$\begin{aligned}
&\hat{\phi}_s(\omega_x,\omega_y)\\
&= \iint \psi(t_1+\rho\omega_x/2, t_2+\rho\omega_y/2)\psi^*(t_1-\rho\omega_x/2, t_2-\rho\omega_y/2) dt_1 dt_2\\
&= \iint \xi_{\rho\omega_{x,y}}(t_1,t_2) e^{i2\pi s(\omega_x t_1+\omega_y t_2)} dt_1 dt_2\\
&= \iint \xi_{\rho\omega_{x,y}}(t_1,t_2) e^{i2\pi(\omega_u t_1+\omega_v t_2)} dt_1 dt_2\\
&= \tilde{\tilde{k}}(\omega_x,\omega_y,\omega_u,\omega_v).
\end{aligned} \tag{75}$$

□

We now let $\mathscr{H}(\omega_x,\omega_y)$ denote the diffraction limited OTF of the system

$$\mathscr{H}(\omega_x,\omega_y) = \Lambda_{A/\rho}(\omega_x)\Lambda_{A/\rho}(\omega_y) \tag{76}$$

where $\Lambda(\omega/c)$ is a width c tent function, and we use the shortcut notation $\Lambda_c(\omega)=\Lambda(\omega/c)$.

The previous claim states that the Fourier optics OTF is a slice through the ambiguity function. The radars literature has also derived the energy of slices via the ambiguity function [Rihaczek 1969], which can be translated to energy bounds on $\tilde{\tilde{k}}_{\omega_{x_0,y_0}}$ [FitzGerrell et al. 1997]. We provide a revised derivation of this result adjusted to our notation. In an analogy to the geometric optics derivation, the norm of a ω_{x_0,y_0}-slice through $\tilde{\tilde{k}}$ is bounded. However, the more accurate Fourier optics bound is no longer equal across spatial frequencies, it is falling off like the diffraction limited OTF $\mathscr{H}(\omega_{x_0},\omega_{y_0})$. This result also suggests that if the system dimensions are kept well above the diffraction limit, the approximation provided by the geometric model is reasonable enough.

Claim 11 *For every spatial frequency ω_{x_0,y_0} the norm of a ω_{x_0,y_0}-slice $\tilde{\tilde{k}}_{\omega_{x_0,y_0}}$ is bounded by the aperture area times the diffraction limited OTF:*

$$\iint |\tilde{\tilde{k}}_{\omega_{x_0,y_0}}|^2 \le A^2 \mathscr{H}(\omega_{x_0},\omega_{y_0}). \tag{77}$$

Proof: We have seen in the proof of Claim 10 that

$$\tilde{\tilde{k}}_{\omega_{x_0,y_0}}(\omega_u,\omega_v) = \iint \xi_{\omega_{x_0,y_0}}(t_1,t_2) e^{i2\pi(\omega_u t_1+\omega_v t_2)} dt_1 dt_2. \tag{78}$$

That is, $\tilde{\tilde{k}}_{\omega_{x_0,y_0}}$ is the Fourier transform of $\xi_{\omega_{x_0,y_0}}$. Therefore we can use Parseval's theorem to bound the norm of $\tilde{\tilde{k}}_{\omega_{x_0,y_0}}$:

$$\begin{aligned}
&\iint |\tilde{\tilde{k}}_{\omega_{x_0,y_0}}(\omega_u,\omega_v)|^2 d\omega_u d\omega_v\\
&= \iint |\xi_{\omega_{x_0,y_0}}(t_1,t_2)|^2 dt_1 dt_2\\
&\le \iint R(A(t_1+\rho\omega_{x_0}/2)) R(A(t_1-\rho\omega_{x_0}/2))\\
&\qquad \cdot R(A(t_2+\rho\omega_{y_0}/2)) R(A(t_2-\rho\omega_{y_0}/2)) dt_1 dt_2\\
&= A^2 \Lambda_{\frac{A}{\rho}}(\omega_{x_0}) \Lambda_{\frac{A}{\rho}}(\omega_{y_0})\\
&= A^2 \mathscr{H}(\omega_{x_0},\omega_{y_0}).
\end{aligned} \tag{79}$$

□

We have seen that using the Fourier optics model, the budget bound is decreasing as the diffraction limited OTF. To understand how specific computational imaging designs are affected, we briefly review the spectrums derivation in the previous section and describe the required adjustments for the Fourier optics model. In a nutshell, for most cases Fourier optics implies that the results in Table 2 should be multiplied by the diffraction limit OTF and $A \to A\Lambda\left(\frac{\rho\omega_x}{A}\right)$. For designs with holes (coded aperture and lattice-focal lens), the hole size reduces the diffraction limit and $\varepsilon A \to \varepsilon A\Lambda\left(\frac{\rho\omega_x}{\varepsilon A}\right)$. Therefore, if the aperture and hole sizes are kept well above the diffraction limit, the approximation provided by the geometric optics model is good enough.

The simplest case is that of a standard lens. The diffraction limited OTF of a squared aperture is derived in many Fourier optics textbooks, e.g., [Goodman 1968]. We briefly review this in the following claim.

Claim 12 *The Fourier optics squared MTF of a standard lens focused at s_0 is*

$$\begin{aligned}
|\hat{\phi}_s(\omega_x,\omega_y)|^2 \quad = \quad & A^4 \Lambda^2_{\frac{A}{\rho}}(\omega_x) \mathrm{sinc}^2\left(A\Lambda_{\frac{A}{\rho}}(\omega_x)(s-s_0)\omega_x\right)\\
& \cdot \Lambda^2_{\frac{A}{\rho}}(\omega_y) \mathrm{sinc}^2\left(A\Lambda_{\frac{A}{\rho}}(\omega_y)(s-s_0)\omega_y\right).
\end{aligned} \tag{80}$$

Proof: The derivation above considered a reference lens with focal length f focused at depth d_o, and with sensor to lens distance d_i

such that $1/f = 1/d_o + 1/d_i$. We now consider another lens with focal length f_a. We maintain the lens to sensor distance d_i. This lens is then focused at depth d_a such that $1/f_a = 1/d_a + 1/d_i$. If s_0 is the light field slope at that depth then $s_0 = d_o(1/d_o - 1/d_a) = d_o(1/f - 1/f_a)$. The aberration φ for this lens is

$$\varphi(u,v) = e^{-i\pi\frac{u^2+v^2}{\lambda}(1/f_a - 1/f)} = e^{-i\pi s_0 \frac{u^2+v^2}{\rho}} \tag{81}$$

and $\xi_{u,v}(t_1,t_2) = R(\Lambda_A(u)\frac{t_1}{A})R(\Lambda_A(v)\frac{t_2}{A})e^{-i2\pi s_0(t_1u+t_2v)}$. Taking the Fourier transform of ξ and adjusting the coordinate system provides $\tilde{\tilde{k}}$:

$$\begin{aligned}\tilde{\tilde{k}}(\omega_x,\omega_y,\omega_u,\omega_v) &= A^2\Lambda_{\frac{A}{\rho}}(\omega_x)\mathrm{sinc}\left(A\Lambda_{\frac{A}{\rho}}(\omega_x)(\omega_u - s_0\omega_x)\right) \\ &\quad\cdot\Lambda_{\frac{A}{\rho}}(\omega_y)\mathrm{sinc}\left(A\Lambda_{\frac{A}{\rho}}(\omega_y)(\omega_v - s_0\omega_y)\right)\end{aligned} \tag{82}$$

Eq. (80) now follows by taking a ω_{x_0,y_0}-slice through $\tilde{\tilde{k}}$. □

Claim 13 *For a lens focused at $s = 0$, the expected Fourier optics squared MTF of a random coded aperture with hole size $\varepsilon A \times \varepsilon A$ is*

$$\begin{aligned}|\hat{\phi}_s(\omega_x,\omega_y)|^2 &= \varepsilon^2A^4\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_x)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_x)(s-s_0)\omega_x\right) \\ &\quad\cdot\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_y)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_y)(s-s_0)\omega_y\right).\end{aligned} \tag{83}$$

Proof: Referring to the notation in the proof of Claim 9, we need to compute $E|\tilde{K}^j|^2$. Based on Claim 12, the expected MTF of an $\varepsilon A \times \varepsilon A$ aperture hole is

$$\begin{aligned}\mathrm{E}[|\tilde{\tilde{k}}^j|^2] &= \frac{1}{2}\varepsilon^4A^4\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_x)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_x)(\omega_u - s_0\omega_x)\right) \\ &\quad\cdot\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_y)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_y)(\omega_v - s_0\omega_y)\right).\end{aligned} \tag{84}$$

(the fact that the holes can be off-centered only affects the phase of $\tilde{K}^j$ but that phase cancels as we measure only the power spectrum). We now multiply Eq. (84) by the windows number $1/\varepsilon^2$ and take an OTF-slice, to get Eq. (83). □

Claim 14 *The Fourier optics squared MTF of a focus sweep camera is*

$$|\hat{\phi}_{s_0}(\omega_x,\omega_y)|^2 \approx A^2\Lambda^2_{\frac{A}{\rho}}(\omega_x)\Lambda^2_{\frac{A}{\rho}}(\omega_y)\frac{\alpha^2(\omega'_{x,y})}{S^2|\omega'_{x,y}|^2} \tag{85}$$

where

$$\omega'_{x,y} = (\Lambda_{\frac{A}{\rho}}(\omega_x)\omega_x, \Lambda_{\frac{A}{\rho}}(\omega_y)\omega_y). \tag{86}$$

Proof: Based on Claim 12, the OTF at slope s is

$$\begin{aligned}\hat{\phi}_s(\omega_x,\omega_y) &= A^2\Lambda_{\frac{A}{\rho}}(\omega_x)\mathrm{sinc}\left(A\Lambda_{\frac{A}{\rho}}(\omega_x)(s-s_0)\omega_x\right) \\ &\quad\cdot\Lambda_{\frac{A}{\rho}}(\omega_y)\mathrm{sinc}\left(A\Lambda_{\frac{A}{\rho}}(\omega_y)(s-s_0)\omega_y\right).\end{aligned} \tag{87}$$

The OTF of a focus sweep camera is obtained by averaging Eq. (87) over s. Using Eq. (45), this conve2rges to

$$\hat{\phi}_{s_0}(\omega_x,\omega_y) \to A\Lambda_{\frac{A}{\rho}}(\omega_x)\Lambda_{\frac{A}{\rho}}(\omega_y)\frac{\alpha(\omega'_{x,y})}{S|\omega'_{x,y}|}. \tag{88}$$

Eq. (85) follows by taking the square of Eq. (88). □

Claim 15 *The expected Fourier optics squared MTF of a lattice-focal lens with holes size $\varepsilon A \times \varepsilon A$ is*

$$|\hat{\phi}_{s_0}(\omega_x,\omega_y)|^2 \approx \varepsilon A^3\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_x)\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_y)\frac{\beta(\omega'_{x,y})}{S|\omega'_{x,y}|}. \tag{89}$$

where

$$\omega'_{x,y} = (\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_x)\omega_x, \Lambda_{\frac{\varepsilon A}{\rho}}(\omega_y)\omega_y). \tag{90}$$

Proof: The squared MTF for a sublens focused at slope s is

$$\begin{aligned}|\hat{\phi}_{s_0}(\omega_x,\omega_y)|^2 &= \varepsilon^4A^4\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_x)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_x)(s-s_0)\omega_x\right) \\ &\quad\cdot\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_y)\mathrm{sinc}^2\left(\varepsilon A\Lambda_{\frac{\varepsilon A}{\rho}}(\omega_y)(s-s_0)\omega_y\right).\end{aligned} \tag{91}$$

As in the proof of Claim 3, the expected squared MTF of a lattice-focal lens is obtained by averaging Eq. (91) over s. Using Eq. (24), this converges to

$$|\hat{\phi}_{s_0}(\omega_x,\omega_y)|^2 \approx \varepsilon^3A^3\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_x)\Lambda^2_{\frac{\varepsilon A}{\rho}}(\omega_y)\frac{\beta(\omega'_{x,y})}{S|\omega'_{x,y}|}. \tag{92}$$

Multiplying Eq. (92) by the number of subsquares $1/\varepsilon^2$ produces Eq. (89). □

The Fourier optics OTF of a wavefront coding lens can be computed as well, and we refer the reader to [Zhang and Levoy 2009; Dowski and Cathey 1995] for derivation.

References

ADAMS, A., AND LEVOY, M. 2007. General linear cameras with finite aperture. In *EGSR*.

AGARWALA, A., DONTCHEVA, M., AGRAWALA, M., DRUCKER, S., COLBURN, A., CURLESS, B., SALESIN, D., AND COHEN, M. 2004. Interactive digital photomontage. In *SIGGRAPH*.

BEN-ELIEZER, E., ZALEVSKY, Z., MAROM, E., AND KONFORTI, N. 2005. Experimental realization of an imaging system with an extended depth of field. *Applied Optics*, 2792–2798.

BRENNER, K., LOHMANN, A., AND CASTENEDA, J. O. 1983. The ambiguity function as a polar display of the OTF. *Opt. Commun. 44*, 323–326.

DOWSKI, E., AND CATHEY, W. 1995. Extended depth of field through wavefront coding. *Applied Optics 34*, 1859–1866.

FITZGERRELL, A. R., DOWSKI, E., AND CATHEY, W. 1997. Defocus transfer function for circularly symmetric pupils. *Applied Optics 36*, 5796–5804.

GEORGE, N., AND CHI, W. 2003. Computational imaging with the logarithmic asphere: theory. *J. Opt. Soc. Am. A 20*, 2260–2273.

GOODMAN, J. W. 1968. *Introduction to Fourier Optics*. McGraw-Hill Book Company.

GU, X., GORTLER, S. J., AND COHEN, M. F. 1997. Polyhedral geometry and the two-plane parameterization. In *EGSR*.

HASINOFF, S., AND KUTULAKOS, K. 2008. Light-efficient photography. In *ECCV*.

HAUSLER, G. 1972. A method to increase the depth of focus by two step image processing. *Optics Communications*, 3842.

HORN, B. K. P. 1968. Focusing. Tech. Rep. AIM-160, Massachusetts Institute of Technology.

LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. 2007. Image and depth from a conventional camera with a coded aperture. *SIGGRAPH*.

LEVIN, A., FREEMAN, W., AND DURAND, F. 2008. Understanding camera trade-offs through a Bayesian analysis of light field projections. In *ECCV*.

LEVIN, A., FREEMAN, W., AND DURAND, F. 2008. Understanding camera trade-offs through a Bayesian analysis of light field projections. *MIT CSAIL TR 2008-049*.

LEVIN, A., SAND, P., CHO, T. S., DURAND, F., AND FREEMAN, W. T. 2008. Motion invariant photography. *SIGGRAPH*.

LEVIN, A., WEISS, Y., DURAND, F., AND FREEMAN, W. 2009. Understanding and evaluating blind deconvolution algorithms. In *CVPR*.

LEVOY, M., AND HANRAHAN, P. M. 1996. Light field rendering. In *SIGGRAPH*.

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

NG, R. 2005. Fourier slice photography. *SIGGRAPH*.

OGDEN, J., ADELSON, E., BERGEN, J. R., AND BURT, P. 1985. Pyramid-based computer graphics. *RCA Engineer 30*, 5, 4–15.

PAPOULIS, A. 1974. Ambiguity function in fourier optics. *Journal of the Optical Society of America A 64*, 779–788.

RIHACZEK, A. W. 1969. *Principles of high-resolution radar*. McGraw-Hill.

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. *SIGGRAPH*.

WIGNER, E. 1932. On the quantum correction for thermodynamic equilibrium. *Phys. Rev. 40*, 5 (Jun), 749–759.

ZHANG, Z., AND LEVOY, M. 2009. Wigner distributions and how they relate to the light field. In *ICCP*.