

Digital Refocusing with Incoherent Holography

Oliver Cossairt
Northwestern University
Evanston, IL

ollie@eecs.northwestern.edu

Nathan Matsuda
Northwestern University
Evanston, IL

nathanmatsuda@u.northwestern.edu

Mohit Gupta
Columbia University
New York, NY

mohitg@cs.columbia.edu

Abstract

Light field cameras allow us to digitally refocus a photograph after the time of capture. However, recording a light field requires either a significant loss in spatial resolution [10, 20, 9] or a large number of images to be captured [11]. In this paper, we propose incoherent holography for digital refocusing without loss of spatial resolution from only 3 captured images. The main idea is to capture 2D coherent holograms of the scene instead of the 4D light fields. The key properties of coherent light propagation are that the coherent spread function (hologram of a single point source) encodes scene depths and has a broadband spatial frequency response. These properties enable digital refocusing with 2D coherent holograms, which can be captured on sensors without loss of spatial resolution. Incoherent holography does not require illuminating the scene with high power coherent laser, making it possible to acquire holograms even for passively illuminated scenes. We provide an in-depth performance comparison between light field and incoherent holographic cameras in terms of the signal-to-noise-ratio (SNR). We show that given the same sensing resources, an incoherent holography camera outperforms light field cameras in most real world settings. We demonstrate a prototype incoherent holography camera capable of performing digital refocusing from only 3 acquired images. We show results on a variety of scenes that verify the accuracy of our theoretical analysis.

1. Introduction

Digital refocusing is the ability to refocus photographs after capture. In computational photography literature, digital refocusing is mainly associated with plenoptic cameras, i.e., cameras that capture 4D light fields [10, 20, 9, 11]. Once the light field is captured on the sensor, refocusing is achieved by propagating the light field to a virtual plane (refocus plane) by means of a 4D convolution. This is illustrated in Fig. 1(a). Finally, the refocused 2D images are computed from linear combinations of the 4D light field.

Why does refocusing require capturing a 4D light field, instead of only 2D images? Intuitively, refocusing requires

‘adding’ and ‘subtracting’ blur to the point spread functions (PSF) of different scene points, depending on their depth. In a conventional 2D image, while it is possible to add blur to a scene point’s PSF, it is not possible to subtract or remove blur without introducing artifacts in the image. This is because 2D imaging is an inherently *lossy* process; the angular information in the 4D light field is lost in a 2D image. Thus, typically, if digital refocusing is desired, 4D light fields are captured. Unfortunately, light field cameras sacrifice spatial resolution in order to capture the angular information. The loss in resolution can be significant, up to 1-2 orders of magnitude. While there have been attempts to improve resolution by using compressive sensing techniques [14], these often rely on strong scene priors [1, 16], which may limit their applicability.

In this paper, we propose using *incoherent holography*, an image capture technique that allows digital refocusing without loss of spatial resolution. The main idea is to capture the coherent 2D hologram of the scene on the sensor, instead of the 4D light field. The hologram is a lossless representation of the optical wave-front radiated by the scene. Since the hologram is 2D, it can be captured on a sensor without sacrificing spatial resolution. This is because holography is an inherently 2D process - the coherent electric field and propagation operator are 2D functions. The coherent electric fields can be propagated between two parallel planes without any loss of information simply by a 2D convolution. This is illustrated in Fig. 1(b).

The key properties of coherent light propagation are that the 2D coherent spread function (CSF)¹ is depth dependent and has a broadband frequency response. This enables refocusing by allowing either adding and subtracting blur for a scene point, depending on its depth. This is illustrated in Fig. 2. These properties taken together enable us to solve the problem of digital refocusing entirely in the 2D domain, completely bypassing the problem of 4D light field acquisition and resolution tradeoffs therein.

Despite the obvious advantages, why isn’t holography

¹Coherent spread function is the hologram due to a point source on the sensor plane. This is analogous to incoherent PSF which is the incoherent intensity field due to a point source on the sensor plane.

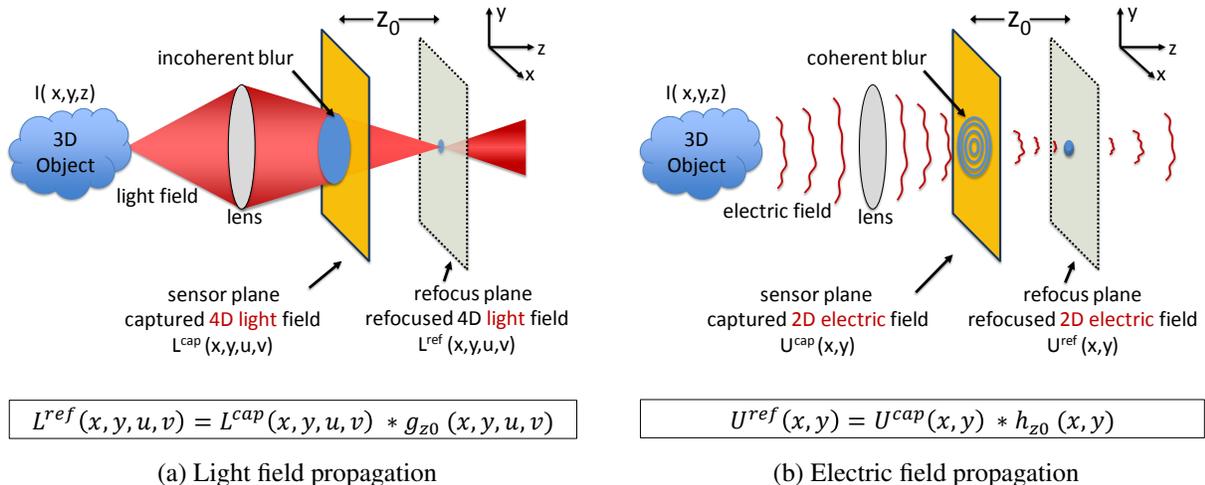


Figure 1. **Light field propagation vs. electric field propagation.** (a) Incoherent light propagation through free space is represented by 4D light fields (in some literature, propagation is expressed as a shearing of the 4D light field, which can be represented as a 4D convolution). Light field cameras capture 4D light fields on 2D sensors, resulting in loss of spatial resolution. (b) Coherent light wave propagation is represented by 2D electric fields. The 2D field can be propagated between parallel planes by a 2D convolution. The coherent blur encodes depth information and has broadband frequency response. This enables digital refocusing without loss of spatial resolution.

used more frequently in photographic settings? The reason is that coherent holography requires interference with a spatially and temporally coherent reference beam. A high power coherent reference beam can only be generated with a laser, making coherent holography an active illumination technique. Fortunately, it is possible to capture holograms without the use of a coherent reference beam, thus opening up the possibility of *incoherent holography*, i.e., holographic image acquisition for passively illuminated scenes. The primary idea behind incoherent holography, first proposed by Lohmann [13], is to divide each point source in the scene into two beams using a beam splitter (see Fig. 8). The first beam is left unchanged (reference) while in the second we introduce a change in focus. In effect, we create a reference beam individually for each point source. Since each point source is spatially coherent relative to itself, there is no longer a need for a spatially coherent reference beam².

Incoherent holography has been suggested as an alternative to light field cameras in order to perform digital refocusing [13, 6, 19, 5, 7, 8]. However, the resolution and SNR tradeoffs of incoherent holography have not been discussed before. In this paper, for the first time, we analyze the digital refocusing capabilities of incoherent holography systems in terms of the achieved resolution and SNR. We perform detailed comparisons with existing techniques that achieve digital refocusing (light field imaging and focal stack imaging). Incoherent holography systems have been built using both phase-based liquid crystal spatial light modulators [6, 19, 5] and Michelson-type interferometers [7, 8]. We employ the latter technique to build our prototype. The main contributions of this paper are:

- We propose incoherent holography as a way to achieve

²The technique eliminates the need for spatial incoherence, but still requires a moderate degree of temporal coherence since it relies on interference. The amount of temporal coherence is typically much less than for coherent holography (e.g. 10nm vs. 0.01nm bandwidth).

full resolution digital refocusing capability.

- We analyze the signal-to-noise ratio (SNR) performance of incoherent holography systems, deriving expressions for the SNR as a function of the maximum depth range of the scene.
- We provide the first in-depth performance comparison between incoherent holography cameras and conventional light field and focal stack cameras.
- We demonstrate a prototype incoherent holography camera capable of performing digital refocusing from only 3 acquired images. We show refocusing results captured using our prototype camera.

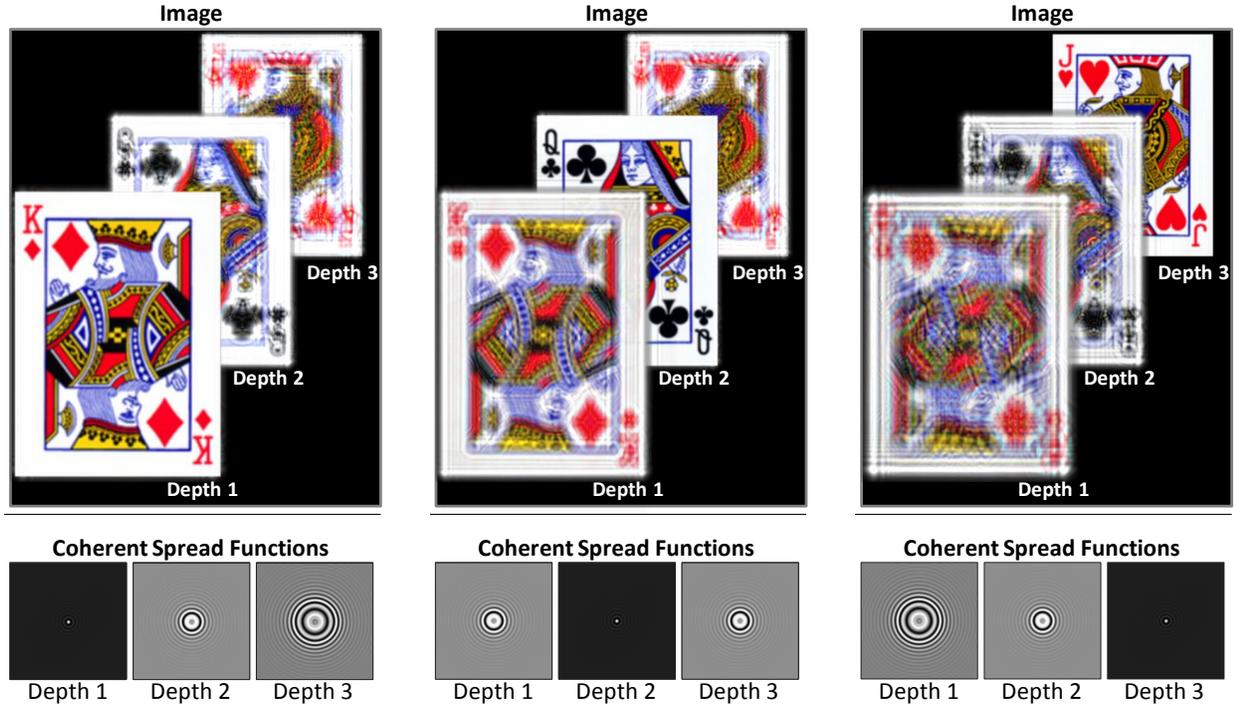
Limitations: The blur kernels in images captured using incoherent holography have a broadband frequency response. Thus, the ‘bokeh’ appears like concentric rings, and not the familiar ‘disc-shaped’ incoherent PSFs. On the flip side, because high frequencies are preserved, it is possible to handle very large image blur sizes. In our experiments, we show results with blurs as large as half the sensor size.

2. Image Formation Model

We now explain the mathematical model for incoherent holography. The idea is based on the optical propagation of coherent fields as seen in Fig. 1(b). We represent a 3D scene as a spatial distribution of point sources $i(\mathbf{r}) = i(x, y, z)$, where depth is measured relative to a reference plane aligned to the $x - y$ axis. Note that while this does not explicitly model the effects of occlusions, it is sufficient to express the depth dependent properties of the optical field. Under the paraxial approximation, the diffracted field intensity $u_0(\mathbf{r})$ is given by the 3D convolution [4]

$$u(\mathbf{r}) = \int_V i(\mathbf{r}_0)h(\mathbf{r} - \mathbf{r}_0)d\mathbf{r}_0, \quad (1)$$

$$h(x, y, z) = \frac{1}{i\lambda z} e^{ikz} e^{\frac{i\pi}{\lambda z}(x^2 + y^2)}. \quad (2)$$



(a) Captured hologram (focused at Depth 1) (b) Refocused at Depth 2 (c) Refocused at Depth 3

Figure 2. **Digital refocusing with incoherent holography.** The scene consists of three playing cards placed at different depths. (a) The 2D hologram is captured with the camera focused at Depth 1. The coherent spread functions (CSF) for the cards vary in size according to their depths. The CSFs have the shape of a sinusoidal zone plate. Consequently, high spatial frequencies are preserved in the hologram even though the maximum blur diameter is nearly half the image width. In contrast, the incoherent PSFs in conventional imaging act as low-pass filters. (b, c) Refocusing on Depth 2 and Depth 3 is achieved by 2D deconvolution of the image in (a) with the second and third CSFs in (a), respectively. After deblurring, the size of the CSFs in the refocused images increases or decreases depending on the scene depths. In conventional imaging, such refocusing cannot be achieved with a 2D deconvolution.

where $h(\mathbf{r}) = h(x, y, z)$ is the 3D Coherent Spread Function (CSF), V is the volume occupied by the scene, $k = 2\pi/\lambda$ is the wavenumber, and λ is the optical wavelength. The field in the reference plane is then simply $u_0(x, y) = u(x, y, z = 0)$. The CSF is a phase-only Fresnel pattern whose scale depends on the depth. The physical interpretation of the CSF is a spherical wave with wavelength λ with radius of curvature equal to z . Convolution of a wavefront with the CSF expresses the fact that each point on the wavefront produces a spherical wave propagating from that point, a phenomena known as the Huygens-Fresnel Principle [4]. The CSF has the following two useful properties:

Property 1: Propagation of the optical field from one plane to the next can be expressed in terms of a 2D convolution

$$u_{z_0}(x, y) = u_0(x, y) * h_{z_0}(x, y). \quad (3)$$

where $*$ is the 2D convolution operator, and we have used the short hand $h_{z_0}(x, y) = h(x, y, z = z_0)$ to represent the 2D CSF. Eqn. 3 is commonly referred to as the *Fresnel diffraction integral*.

Property 2: The Fourier transform of the 2D CSF $H_{z_0}(k_x, k_y)$ is also a phase-only pattern, and its modulus (i.e. the MTF) is constant

$$H_{z_0}(k_x, k_y) = \mathcal{F}\{h_{z_0}(x, y)\} = e^{ikz} e^{i\pi\lambda z(k_x^2 + k_y^2)}, \quad (4)$$

$$|H_{z_0}(k_x, k_y)| = 1. \quad (5)$$

The consequence of Property 1) is that the propagating the field to a plane with arbitrary depth $z = z_0$ can be expressed compactly as a 2D convolution between the field at the reference plane and the 2D CSF $h_{z_0}(x, y)$. If a sensor at the reference plane captures the coherent field, then Eqn. 3 expresses how to ‘refocus’ the field to a sensor located at a distance $z = z_0$ from the sensor plane. The consequence of Property 2) is that the refocusing operation can be performed with *zero noise amplification* since the MTF is unity and all frequencies are preserved exactly.

2.1. Coherent Holography

Unfortunately, visible light frequencies of hundreds of THz prevent direct capture of the coherent field. We can only measure the intensity of the optical field. The core idea behind holography is to use a reference beam $u_{ref}(x, y)$ that interferes with the coherent field $u_0(x, y)$ on the sensor [3]. The measured intensity of the combined field on the sensor is then

$$i_0(x, y) = |u_{ref}(x, y) + u_0(x, y)|^2. \quad (6)$$

$$= \underbrace{|u_{ref}|^2 + |u_0|^2}_{\text{intensity terms}} + \underbrace{u_{ref}^* u_0 + u_{ref} u_0^*}_{\text{interference terms}}, \quad (7)$$

where $*$ is the complex conjugate operator. The captured image $i(x, y)$ consists of the sum of four terms. The first two terms combine the intensities of the reference and object fields. The last two terms contain interference that gives relative phase between the reference $u_{ref}(x, y)$ and object field $u_0(x, y)$. A multitude of techniques have been devised to recover the desired complex object field from the captured image, in this paper we use the method of phase shifting [21]. The idea is to capture a sequence of K images $u_0^{(k)}$ with phase shifted reference beams $u_{ref}^{(k)}(x, y) = e^{i\phi_k}$, where the phase for the k^{th} reference beam is given by

$$\phi_k = \frac{2\pi}{\lambda} \frac{k}{K}. \quad (8)$$

Phase shifting the reference beam is typically accomplished with high precision using either a mirror actuated by a piezo-electric stack [7, 8] or a phase modulating LCD [6, 19, 5]. The complex field can then be recovered exactly from the K captured images using the relation

$$\sum_{k=1}^K i^{(k)}(x, y) e^{-i\phi_k} = u_0(x, y). \quad (9)$$

As few as $K = 3$ images is sufficient for exact recovery.

2.2. Incoherent Holography

Coherent holography relies on creating a reference beam as expressed by Eqn. 8. Unfortunately this can only be done with a high power spatially and temporally coherent source, (i.e. a laser). Incoherent holography overcomes this limitation by splitting each scene point into two beams, effectively creating a reference for each scene point individually.

2.2.1 System Definition

For the purposes of this paper, we consider the incoherent holography camera geometry shown in Fig. 3. We begin by following beam path 1 (Fig. 3, top). Consider a point source at a distance of z from the camera aperture of width w . The main lens with focal length f_1 images the point source onto the sensor. The camera aperture is located one focal length in front of the lens, and the sensor is located one focal length behind the lens. This is an infinite-conjugate geometry, so scene points at infinity (or just beyond the hyperfocal distance) will come to focus on the sensor. Closer scene points will come to focus a distance δ_1 in front of the sensor, producing a blur circle on the sensor with diameter b_1 .

For the second beam path (Fig. 3, bottom), a lens with focal length f_2 is placed in the aperture. The effect is to increase the focal shift to $\delta_1 + \delta_2$, producing a larger blur width b_2 . The two beams overlap on the sensor, producing

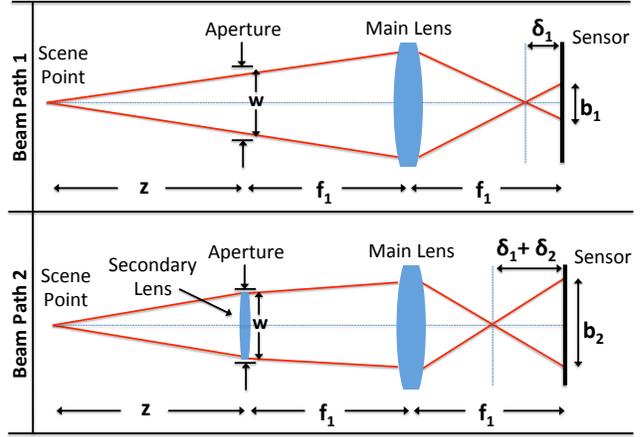


Figure 3. **Incoherent Holography Geometry.** The figure shows the geometric parameters of an incoherent holographic camera. The Michelson interferometer setup from Fig. 8 is unfolded and the two paths are shown independently.

an interference pattern. Incoherent holography is in no way restricted to such a $2f$ setup; it is used here merely as a convenience, resulting in the convenient relations

$$\delta_1 = -f_1^2/z \quad (10)$$

$$\delta_2 = f_2^2/f_2 \quad (11)$$

$$b_1 = w \cdot \delta_1/f_1 \quad (12)$$

$$b_2 = w \cdot (\delta_1 + \delta_2)/f_1, \quad (13)$$

where a negative blur size indicates the beam comes to focus behind the sensor. Scene points located at $z = \infty$ come in to focus on the sensor (i.e. $\delta_1 = 0$). If the scene depths are in the range $z \in \{z_{min}, \infty\}$, the focal length of the secondary lens is chosen so that $f_2 = z_{min}/2$, resulting in a range of blur sizes $b_1 \in \{-b_{max}, 0\}$ and $b_2 \in \{b_{max}, 2b_{max}\}$ for the first and second beam, respectively, where $b_{max} = w \cdot f_1/z_{min}$. The CSFs for beams 1 and 2 have a finite support that is determined by the optical cutoff frequency $f_0 = w/(2\lambda f_1)$. Spatial frequencies greater than f_0 will be suppressed by the optical system. This will naturally limit the support of the CSFs since spatial frequencies increase outwardly from the center for Fresnel patterns. It follows that a reasonable choice for the pixel size is $\Delta = 1/(2f_0)$. The two waves incident on the sensor may then be expressed as

$$h_1(x, y, z) = \frac{1}{b_1} e^{\frac{i\pi}{\Delta b_1}(x^2+y^2)} \quad (14)$$

$$h_2(x, y, z) = \frac{1}{b_2} e^{\frac{i\pi}{\Delta b_2}(x^2+y^2)} \quad (15)$$

where the depth dependence is implicit because b_1, b_2 , depend on the depth z .

2.2.2 Incoherent Holography PSF

The aggregate system CSF is equal to the sum of the CSFs. However, for incoherent imaging, scene points no longer add coherently. The impulse response becomes the Point

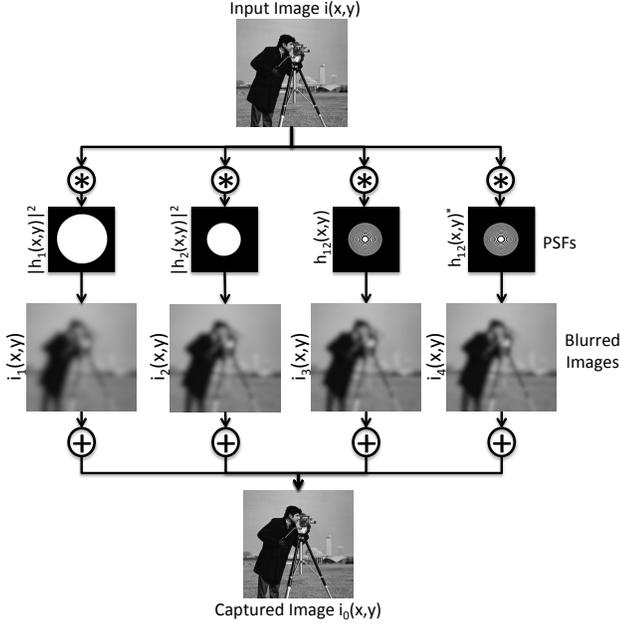


Figure 4. **Incoherent Holography Capture Decomposition.** The incoherent holography PSF consists of a sum of four different PSFs. Phase shifting is used to remove the blurred images produced from the three unwanted PSFs.

Spread Function (PSF), which is the squared modulus of the CSF. The image intensity $i_0(\mathbf{r}) = i_0(x, y, z)$ for a sensor located at a depth z will be

$$i_0^{(k)}(\mathbf{r}) = \int_V i(\mathbf{r}_0) p^{(k)}(\mathbf{r} - \mathbf{r}_0) d\mathbf{r}_0, \quad (16)$$

$$p^{(k)}(x, y, z) = \left\| h_1(x, y, z) + e^{i\phi_k} h_2(x, y, z) \right\|^2 \quad (17)$$

$$= \underbrace{\left\| h_1 \right\|^2}_{\text{Beam 1 PSF}} + \underbrace{\left\| h_2 \right\|^2}_{\text{Beam 2 PSF}} + \underbrace{e^{i\phi_k} h_{12} + e^{-i\phi_k} h_{12}^*}_{\text{Interference PSFs}}$$

where $p(\mathbf{r})$ is the system PSF and we have added a phase offset to the second beam (e.g. by using a piezo-actuated mirror as shown in Fig. 8). The aggregate PSF is a sum of four individual PSFs. The intensity image $i_0^{(k)}(x, y, 0)$ measured on the sensor at a depth $z = 0$ consists of a sum of four blurred images, as visualized in Figure 4. The first two terms in Eqn. 18 are the incoherent PSFs produced by beams 1 and 2. The third and fourth terms are PSFs produced from the interference of beams 1 and 2.

2.2.3 Synthesizing the Coherent Field

As in the coherent holography case, we capture K phase shifted images, and process to compute our final result.

$$\tilde{u}(\mathbf{r}) = \sum_{k=1}^K i_0^{(k)}(\mathbf{r}) e^{-i\phi_k} \quad (18)$$

For incoherent holography, the effect of phase shifting is to remove the images blurred by PSFs 1, 2, and 4. The result is an image that is blurred only by a synthesized PSF $h_{12}(x, y)$, where the synthesized PSF has a very remarkable property: *The synthesized 3D PSF is a scaled version of the 3D CSF from Eqn. 2.*

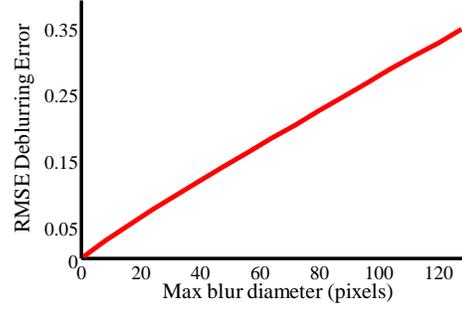


Figure 6. **RMS deblurring error for incoherent holography.** The deblurring error grows linearly as a function of the maximum blur size in the image.

$$\tilde{u}(\mathbf{r}) = \int_V i(\mathbf{r}_0) h_{12}(\mathbf{r} - \mathbf{r}_0) d\mathbf{r}_0 \quad (19)$$

$$h_{12}(x, y, z) = \frac{1}{b_1 b_2} e^{\frac{i\pi}{\Delta} \left(\frac{1}{b_1} - \frac{1}{b_2} \right) (x^2 + y^2)} \quad (20)$$

As in coherent holography, the recovered image is given by $\tilde{u}_0(x, y) = \tilde{u}(x, y, z = 0)$. As a result, digital refocusing can be applied using Eqn. 3 to synthetically propagate the field to a distance $z = z_0$ from the sensor (see Fig. 1).

3. SNR Analysis

3.1. Full-resolution Light Fields

Holographic cameras face the challenge of mapping a 2D complex quantity onto a 2D sensor, a problem commonly referred to as *phase retrieval* in the optics literature. However, light field cameras face the considerably more difficult challenge of mapping a 4D (real) quantity onto a 2D sensor. Plenoptic sensor designs [10, 20, 9] sacrifice spatial resolution for angular resolution. Capturing a full resolution light field must be done sequentially using a mask placed either in the lens aperture or near the sensor. For instance, to capture $N \times N$ angular samples, a $N \times N$ mask may be placed in the lens aperture [11]. N^2 images must be then acquired, changing the mask between exposures, and reducing the total amount of light captured.

According to the theory introduced by Ng [17], a light field camera with $N \times N$ angular resolution can be used to refocus images within a Depth of Field (DoF) that is $N \times$ greater than the DoF of the camera.

Sticking with the same 2-f geometry as in Section 2, in order to be able to digitally refocus scene depths in the range $z \in \{z_{min}, \infty\}$, we must ensure $\Delta/b_{max} \leq N$, where Δ is again the pixel size. For large light levels, the SNR for the light field camera measured relative to a conventional camera is then given by

$$SNR_{lf} = \left(\frac{\Delta}{b_{max}} \right)^2. \quad (21)$$

3.2. Incoherent Holography

Section 2 introduced the concept of synthesizing a 3D CSF from an incoherent imaging system. The analysis cov-

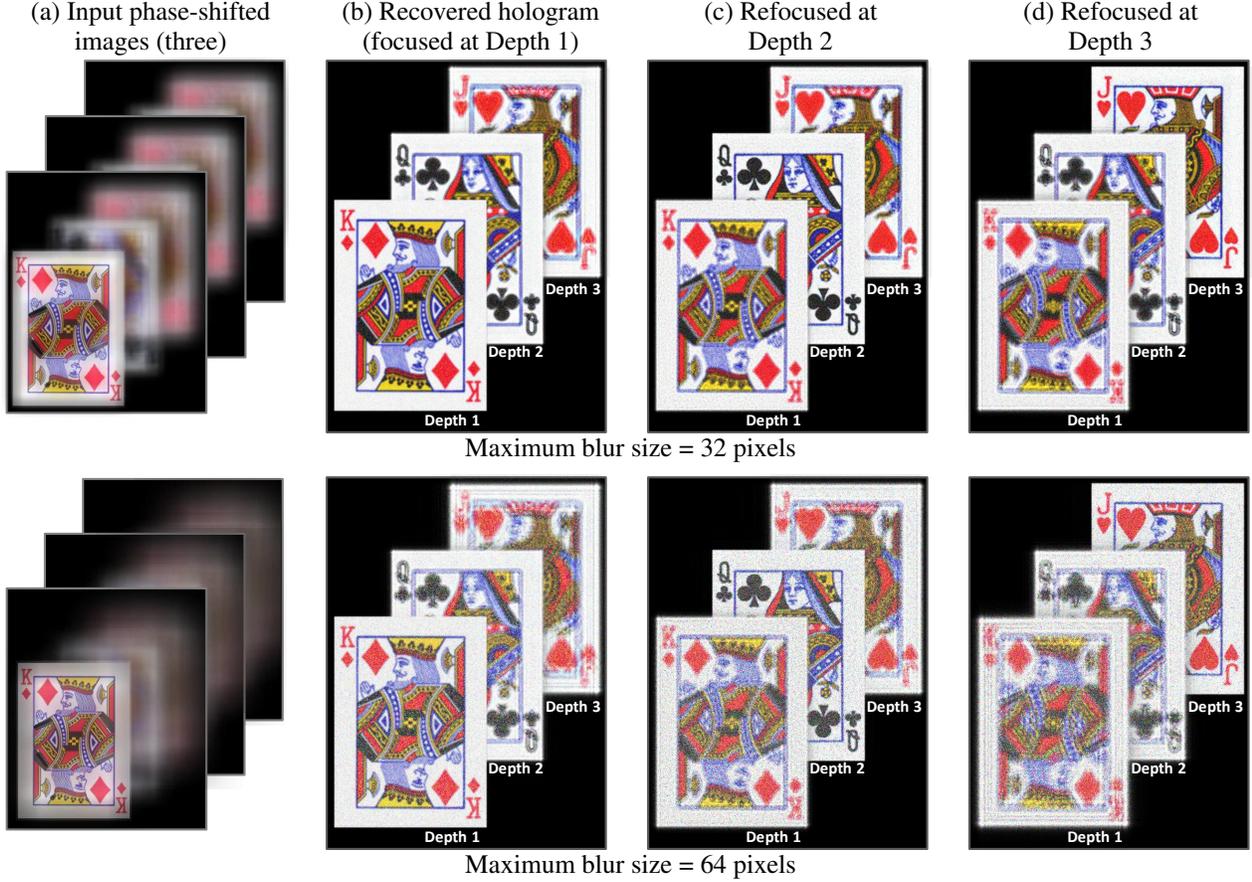


Figure 5. **Simulation results.** The scene consists of three playing cards placed at different depths. The camera is focused on the front card (Depth 1). The scene depths are chosen so that the maximum blur size is 32 pixels (top row) and 64 pixels (bottom row). (a) The three input phase-shifted images. (b) The recovered holograms. The front card is in focus, and the other two cards are blurred. (c, d) Images refocused on the middle card (Depth 2) and the back card (Depth 3), respectively. Image noise is higher for the larger blur (bottom row).

ered the basic principle of capturing the 2D field on the sensor and digitally refocusing to a new synthetic sensor location. Digital refocusing is achieved via a noise-free 2D deconvolution process. However, the light captured by the sensor was divided between the four blurry images. At no point did we quantify the fraction of captured signal corresponding to the recovered field $u_0(x, y)$. However, we can do so by calculating the OTF and MTF from Eqn. 20

$$H(k_x, k_y) = \mathcal{F}\{h_{12}(x, y)\} \quad (22)$$

$$= \frac{i\Delta}{b_2 - b_1} e^{i\pi\Delta(\frac{b_1 b_2}{b_2 - b_1})(k_x^2 + k_y^2)}, \quad (23)$$

$$|H(k_x, k_y)| = \frac{\Delta}{b_{max}}. \quad (24)$$

The MTF is inversely proportional to the maximum blur size. Since the MTF is flat over the passband, this means that SNR falls off exactly in the same way, *producing significantly greater SNR than a full resolution light field camera*. Fig. 5 shows simulation results for an incoherent holography system with two blur sizes, $b_{max}/\Delta = 32$ (top), and $b_{max}/\Delta = 64$ (bottom), and a noise variance of $\sigma = .005$. The deblurring error is seen to increase with the maximum blur size. The increase in noise is linear, as shown in Fig. 6,

where it is evident that the ratio of RMSE deblurring error is approximately equal to the ratio of blur sizes.

3.2.1 The Effect of Temporal Coherence

The analysis thus far assumed that the two beams produced by each scene point in an incoherent holography system will interfere. However, there are necessary conditions conditions that must be ensured before interference is observed: spatial and temporal coherence. For Incoherent holography, spatial coherence is preserved between the the two beams because they originate from the same scene point. However, spatial coherence is not maintained for neighboring scene points, and as a result, no interference is observed between them. Narrowband light sources, such as a laser or LED, may be used to illuminate the scene to ensure there is sufficient temporal coherence. On the other hand, if natural light sources, such as sunlight or incandescent bulbs, are used to illuminate the scene, color filters must be placed in front of the lens to reduce the range of wavelengths incident on the sensor to within a sufficiently narrow band $\Delta\lambda$, reducing the amount of captured light, and hence also the SNR. More precisely, the maximum difference in optical path length (OPD) between the two beams must not exceed

	SNR	#/Images
Light Field (photon noise limited)	1/2	N^2
Light Field (read noise limited)	$1/N^2$	N^2
Incoherent Holography (broadband lighting)	$1/N^2$	3
Incoherent Holography (coherent flash)	$1/N$	3
Focal Stack	$1/N$	$1/N$

Figure 7. **Full resolution refocusing comparison.** A comparison between full-resolution light field cameras, incoherent holography, and focal stack imaging. The SNR given is for the refocused images after digital processing.

the coherence length of the source $l_{src} = \lambda^2/\Delta\lambda$. The required coherence length l_{IC} for our incoherent holographic camera is determined by the maximum OPD

$$l_{IC} = OPD_{max} = \frac{\lambda b_{max}}{\Delta} \quad (25)$$

The coherence lengths for some common light sources are 30cm for a Helium-Neon Laser, and 1-10mm for a laser diodes. Blackbody emitters have coherence lengths on the order of a single wavelength. For instance, the sun has a coherence length of $l_{src} = .6\mu m$ and incandescent lights with a color temperature of 3,000K have a coherence length of $1.2\mu m$ [2]. The fraction of light transmitted after color filtering to achieve the desired coherence length can roughly be approximated by $T = l_{src}/l_{IC} \approx \Delta/b_{max}$ ³. As a result, the SNR of an incoherent holography systems after applying the color filter is

$$T \cdot |H(k_x, k_y)| \approx \left(\frac{\Delta}{b_{max}} \right)^2, \quad (26)$$

which is the same as for full resolution light field capture.

3.3. When to use Incoherent Holography

In extremely low light conditions, Hadamard Multiplexed light field cameras achieve a higher SNR as compared to incoherent holography camera, albeit at a cost of more input images. However, as light level increase and photon noise dominates (which is common in photography), incoherent holography achieves the same SNR as Hadamard light field cameras with significantly fewer images. Furthermore if a narrowband flash is used, the SNR for incoherent holography is even higher, better than a full resolution light field camera, as shown in Table 7. We include a comparison with focal stack imaging [22, 15]. Refocusing with focal stacks requires capturing N images for each of the N different blur sizes between 1 and $N = b_{max}/\Delta$. Note that incoherent holography with a narrowband flash can achieve the same SNR as focal stack imaging, however with significantly fewer images.

³This is a worst case approximation since it assumes the source spectrum is flat, and the center wavelength of the filter can typically be chosen for peak efficiency.

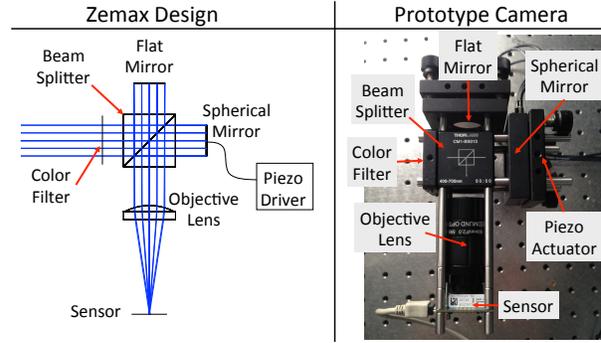


Figure 8. **Prototype Incoherent Holography Camera.** Our prototype camera utilizes a Michelson interferometer setup. The left figure shows the Zemax design and the right figure shows our prototype camera. The camera parameters are given in the text.

4. Prototype Camera

We designed and built a prototype incoherent holography camera based on a Michelson-type interferometer [7, 8], as shown in Fig. 8. A cube beam splitter divides incoming light from the scene into two beams. Two mirrors then reflect the beams back towards an objective lens that images the wavefront onto the sensor. One of the mirrors has a small amount of curvature to shift the focus of the beam relative to the other path. A piezo-actuated mirror creates sub-micron displacements that are used for phase shifting. We used a Kinematic mirror mount (Thorlabs KC1-T-PZ) controlled by three T-Cube piezo controllers (Thorlabs TPZ001). The mirrors were both $w = 25.4mm$ diameter and the objective lens was an 50mm focal length lens (Edmunds Optics 59-873), resulting in an aperture setting of $f_1/w \approx F/2$. For the sensor, we used a 1/2" monochrome 10 Mpix sensor with $1.67\mu m$ pixels (Edmunds Optics 86-749), producing a Nyquist frequency just slightly larger than the diffraction cutoff frequency. For the color filter, we used a 10nm bandwidth interference filter with center frequency $632.8nm$ (Thorlabs FL632.8-10). The focal length of the curved mirror (f_2) is chosen according to the desired depth range. As mentioned in Section 2.2.1, the depth range is given by $\{2f_2, \infty\}$. In our experiments, we chose $f_2 = 1000mm$, resulting in a focus shift of $\delta_2 = 2.5mm$, in order to cover a large depth range. A large depth range comes at a cost of low SNR. If higher SNR and lower depth range are desired, a mirror with lower f_2 should be chosen.

Two examples scenes captured with our camera are shown in Fig. 9. **Each scene was captured at full 10 Mpix resolution.** The exposure time was long $\approx 1/3sec$ due to the amount of light blocked by the interference filter. The small pixels have a very small dynamic range, and we averaged 100 frames and used denoising software to increase the SNR to a reasonable level.

The top row shows a scene with playing cards placed at different depths. The right image shows the coherent field recovered after phase shifting, which is focused on the King in the background. The center image shows the results after refocusing on the Jack in the foreground. The Jack is seen to come into crisp focus while the King becomes sig-

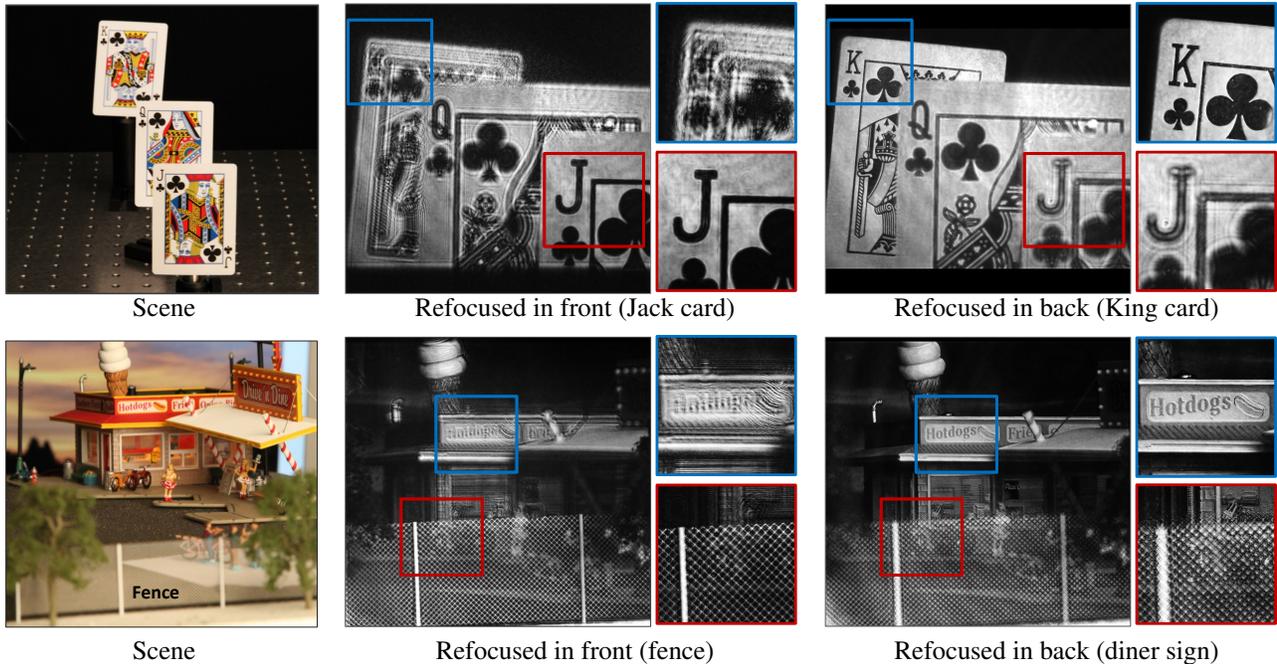


Figure 9. **Experimental results for digital refocusing.** (Top row) Playing cards scene. Depth range is approximately 75cm and the maximum blur size is approximately 1200 pixels. (Bottom row) Diner scene. The depth range is approximately 50cm and the maximum blur size is approximately 600 pixels. When image is refocused on the back, the fence is blurred, but maintains high frequencies, and thus, cannot be seen through.

nificantly blurred. The bottom row shows a miniature diner scene with a fence in the foreground. The right image shows the coherent field recovered after phase shifting, which is focused on the diner in the background, while the fence is severely blurred. Note that high frequencies in the fence are still visible even with an extremely large blur size. The center image shows the the fence come into focus while the diner becomes blurred.

5. Discussion

We have shown that incoherent holography provides a viable alternative to full-resolution light field capture for digital refocusing applications. The results from our prototype camera demonstrate refocusing over very large blur sizes, and we have shown high quality refocusing results at a full 10 Mpix resolution. However, as we have derived in the paper, the SNR of incoherent holography becomes small with large blur sizes. The images we captured required long exposures, frame averaging, and denoising to achieve a good SNR. Still, the performance we demonstrate is similar to what can be achieved with a full-resolution light field camera in photon limited settings. The significant difference is that we only require 3 images for full-resolution refocusing while light field cameras require significantly more. Even better performance may be achieved using a narrowband coherent flash, such as an LED or Laser diode, which can be easily incorporated into camera designs.

There are several future research directions. Analysis of how the focal length f_2 effects resolution and SNR was outside the scope of this paper but we plan to include this in future publications. One important challenge is to de-

velop new designs that overcome the FOV limitation imposed by the use of a beam splitter cube. Another interesting research question is how to develop image processing techniques that produce the same 'bokeh' as an incoherent imaging system. Since resolution has been a primary concern, it may be fruitful to consider compressive acquisition schemes [18] in addition to comparing with compressive light field [14] and focal stacks [12]. It will be interesting to more closely compare the information contained in light fields and holograms, for we expect that holograms will more faithfully capture view-dependent effects such as those caused by occlusion boundaries and highly specular BRDFs. Finally, we believe there are still significant advancements to be made in various applications for incoherent holography such as aberration correction [7], free-viewpoint rendering [23], and 3D scene estimation from a minimal number of measurements [18].

References

- [1] O. Cossairt, M. Gupta, and S. K. Nayar. When does computational imaging improve performance? *IEEE transactions on image processing*, 22(1-2):447–458, 2013.
- [2] A. Donges. The coherence length of black-body radiation. *European Journal of Physics*, 19(3):245, 1998.
- [3] D. Gabor. A new microscopic principle. *Nature*, 4098(161):777–778, 1948.
- [4] J. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 2nd edition, 1996.
- [5] B. K. J. Rosen and G. Brooker. Finch: Fresnel incoherent correlation hologram. In *Holography, Research and Technologies*, pages 135–154. Intech, 2011.

- [6] G. B. J. Rosen. Digital spatially incoherent fresnel holography. *Optics Letters*, (32):912–914, 2007.
- [7] M. K. Kim. Adaptive optics by incoherent digital holography. *Opt. Lett.*, 37(13):2694–2696, Jul 2012.
- [8] M. K. Kim. Full color natural light holographic camera. *Opt. Express*, 21(8):9636–9642, Apr 2013.
- [9] D. Lanman, R. Raskar, A. Agrawal, and G. Taubin. Shield fields: modeling and capturing 3d occluders. In *SIGGRAPH*, 2008.
- [10] M. Levoy. Light fields and computational imaging. *IEEE Computer*, 39(8):46–55, 2006.
- [11] C. Liang, T. Lin, B. Wong, C. Liu, and H. Chen. Programmable aperture photography: multiplexed light field acquisition. In *SIGGRAPH*, 2008.
- [12] X. Lin, J. Suo, G. Wetzstein, Q. Dai, and R. Raskar. Coded focal stack photography. In *ICCP*, pages 1–9, 2013.
- [13] A. W. Lohmann. Wavefront reconstruction for incoherent object. *J. Opt. Soc. Am.*, 1555(55):1555–1556, 1965.
- [14] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Trans. Graph.*, 32(4):46:1–46:12, July 2013.
- [15] D. Miao, O. Cossairt, and S. Nayar. Focal sweep videography with deformable optics. In *ICCP*, pages 1–8, 2013.
- [16] K. Mitra, O. Cossairt, and A. Veeraraghavan. A framework for the analysis of computational imaging systems with practical applications. *arXiv preprint arXiv:1308.1981*, 2013.
- [17] R. Ng. Fourier slice photography. In *SIGGRAPH*, pages 735–744, New York, NY, USA, 2005.
- [18] Y. Rivenson, A. Stern, and B. Javidi. Overview of compressive sensing techniques applied in holography. *Appl. Opt.*, 52(1):A423–A432, Jan 2013.
- [19] J. Rosen and G. Brooker. Fluorescence incoherent color holography. *Opt. Express*, 15(5):2244–2250, Mar 2007.
- [20] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In *SIGGRAPH*, 2007.
- [21] I. Yamaguchi and T. Zhang. Phase-shifting digital holography. *Opt. Lett.*, 22(16):1268–1270, Aug 1997.
- [22] C. Zhou, D. Miao, and S. Nayar. Focal Sweep Camera for Space-Time Refocusing. Technical report, Nov 2012.
- [23] R. Ziegler, S. Bucheli, L. Ahrenberg, M. Magnor, and M. Gross. A bidirectional light field - hologram transform. *Computer Graphics Forum*, 26(3):435–446, 2007.