Wave optics theory and 3-D deconvolution for the light field microscope

Michael Broxton,* Logan Grosenick, Samuel Yang, Noy Cohen, Aaron Andalman, Karl Deisseroth, and Marc Levoy

Departments of Computer Science and Bioengineering, Stanford University, Stanford, CA 94305, USA

*broxton@stanford.edu

Abstract: Light field microscopy is a new technique for high-speed volumetric imaging of weakly scattering or fluorescent specimens. It employs an array of microlenses to trade off spatial resolution against angular resolution, thereby allowing a 4-D light field to be captured using a single photographic exposure without the need for scanning. The recorded light field can then be used to computationally reconstruct a full volume. In this paper, we present an optical model for light field microscopy based on wave optics, instead of previously reported ray optics models. We also present a 3-D deconvolution method for light field microscopy that is able to reconstruct volumes at higher spatial resolution, and with better optical sectioning, than previously reported. To accomplish this, we take advantage of the dense spatio-angular sampling provided by a microlens array at axial positions away from the native object plane. This dense sampling permits us to decode aliasing present in the light field to reconstruct high-frequency information. We formulate our method as an inverse problem for reconstructing the 3-D volume, which we solve using a GPU-accelerated iterative algorithm. Theoretical limits on the depth-dependent lateral resolution of the reconstructed volumes are derived. We show that these limits are in good agreement with experimental results on a standard USAF 1951 resolution target. Finally, we present 3-D reconstructions of pollen grains that demonstrate the improvements in fidelity made possible by our method.

OCIS codes: (180.6900) Three dimensional microscopy; (180.2520) Fluorescence microscopy; (100.1830) Deconvolution; (100.6950) Tomographic image processing; (100.3190) Inverse problems.

References and links

1. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," in *Proceedings of ACM SIGGRAPH*. (2006) 924–934.
2. M. Levoy, Z. Zhang, and I. McDowell, "Recording and controlling the 4D light field in a microscope using microlens arrays," Journal of Microscopy **235**, 144–162 (2009).
3. S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," International Journal of Imaging Systems and Technology **14**, 47–57 (2004).
4. R. Ng, "Fourier slice photography," in *Proceedings of ACM SIGGRAPH* (2005). 735–744.
5. M. Bertero and C. de Mol, "III Super-resolution by data inversion," in *Progress in Optics* (Elsevier, 1996) pp. 129–178.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

6. T. Pham, L. van Vliet, and K. Schutte, "Influence of signal-to-noise ratio and point spread function on limits of superresolution," Proc. SPIE **5672**, 169–180 (2005).
7. S. Baker and T. Kanade, "Limits on super-resolution and how to break them," IEEE Trans. Pattern Anal. Mach. Intell. **24**. 1167–1183 (2002)
8. K. Grochenig and T. Strohmer, "Numerical and theoretical aspects of nonuniform sampling of band-limited images," in "Nonuniform Sampling," , F. Marvasti, ed.. Information Technology: Transmission, Processing, and Storage, 283–324 (Springer US, 2010).
9. T. Bishop and P. Favaro, "The light field camera: extended depth of field, aliasing and super-resolution," IEEE Trans. Pattern Anal. Mach. Intell. **34**. 972–986 (2012).
10. W. Chan, E. Lam, M. Ng, and G. Mak, "Super-resolution reconstruction in a computational compound-eye imaging system," Multidimensional Systems and Signal Processing **18**. 83–101. (2007).
11. M. Gu, *Advanced Optical Imaging Theory* (Springer, 1999).
12. D. A. Agard, "Optical sectioning microscopy: cellular architecture in three dimensions," Annual review of biophysics and bioengineering **13**, 191–219. (1984).
13. M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University, 1999).
14. M. R. Arnison and C. J. R. Sheppard, "A 3D vectorial optical transfer function suitable for arbitrary pupil functions," Optics communications **211**, 53–63 (2002).
15. A. Egner and S. W. Hell, "Equivalence of the Huygens–Fresnel and Debye approach for the calculation of high aperture point-spread functions in the presence of refractive index mismatch," Journal of Microscopy **193**, 244–249 (1999).
16. J. Breckinridge, D. Voelz, and J. B. Breckinridge, *Computational Fourier Optics: a MATLAB Tutorial* (SPIE Press, 2011).
17. J. M. Bardsley and J. G. Nagy, "Covariance-preconditioned iterative methods for nonnegatively constrained astronomical imaging," SIAM journal on matrix analysis and applications **27**, 1184–1197 (2006).
18. M. Bertero, P. Boccacci, G. Desidera, and G. Vicidomini, "Image deblurring with Poisson data: from cells to galaxies," Inverse Problems **25**, 123006 (2009).
19. S. Shroff and K. Berkner, "Image formation analysis and high resolution image reconstruction for plenoptic imaging systems," Applied optics, **52**, D22D31, (2013).
20. J. Rosen, N. Siegel, and G. Brooker, "Theoretical and experimental demonstration of resolution beyond the Rayleigh limit by FINCH fluorescence microscopic imaging," Opt. Express **19**, 1506–1508 (2011).
21. I. J. Cox and C. J. R. Sheppard, "Information capacity and resolution in an optical system," J. Opt. Soc. Am. A **3**, 1152 (1986).
22. R. Heintzmann, "Estimating missing information by maximum likelihood deconvolution," Micron **38**, 136–144 (2007)
23. P. Favaro, "A split-sensor light field camera for extended depth of field and superresolution," in "SPIE Conference Series," **8436**. (2012).
24. C. H. Lu, S. Muenzel, and J. Fleischer, "High-resolution light-field microscopy," in "Computational Optical Sensing and Imaging, Microscopy and Tomography I (CTh3B)," (2013).
25. S. Abrahamsson, J. Chen, B. Hajj, S. Stallinga, A. Y. Katsov, J. Wisniewski, G. Mizuguchi, P. Soule, F. Mueller, C. D. Darzacq, X. Darzacq, C. Wu, C. I. Bargmann, D. A. Agard, M. G. L. Gustafsson, and M. Dahan, "Fast multicolor 3D imaging using aberration-corrected multifocus microscopy," Nat. Meth. 1–6. (2012).
26. M. Pluta, *Advanced Light Microscopy, Vol. 1.* (Elsevier, 1988).
27. J. Goodman, *Introduction to Fourier Optics*, 2nd ed. (MaGraw-Hill, 1996).

1. Introduction

Light field microscopy [1, 2] is a new approach for rapid, scanless volumetric imaging using optical microscopy. By introducing a microlens array into the light path of a conventional microscope, it is possible to image both the lateral and angular distribution of light passing through the specimen volume. This raw spatio-angular data, referred to as a light field, can be post-processed in software to produce a full 3-D reconstruction of the object. Because the light field is captured in a snapshot at a single instant in time, light field microscopy makes it possible to perform scanless volumetric imaging at the frame rate of the camera.

However, scanless 3-D imaging with a light field microscope (LFM) comes at a cost: lateral spatial resolution must be sacrificed to record angular information. This angular information is what permits volumetric reconstruction, so this is equivalent to a trade off between lateral and axial resolution. In previous work [1], we concluded that this loss of lateral resolution is exactly proportional to the number of discrete angular samples collected. This in turn is proportional

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

to the number of pixels behind each lenslet in the array. Since it is desirable to record many angular samples at each spatial position (typically greater than 10 in each angular direction), this represents a considerable resolution loss relative to the diffraction limited performance of a conventional wide-field fluorescence microscope.

In this paper we present a novel approach for light field reconstruction with improved lateral spatial resolution compared to ordinary light field microscopy. We further show that, when constraining the reconstruction to a single z-plane in the case of a planar test target, we achieve up to an 8-fold improvement in resolution over previously reported limits (see Fig. 1). This is achieved by modeling the spatially varying point spread function of the LFM using wave optics and using this model to perform 3-D deconvolution. Such deconvolution mitigates the problem of decreased lateral resolution, thereby addressing one of the main drawbacks of light field microscopy.

This resolution enhancement is possible because the sampling rate provided by the lenslets is well below the optical band limit of the microscope. As a result, the microlens array samples the

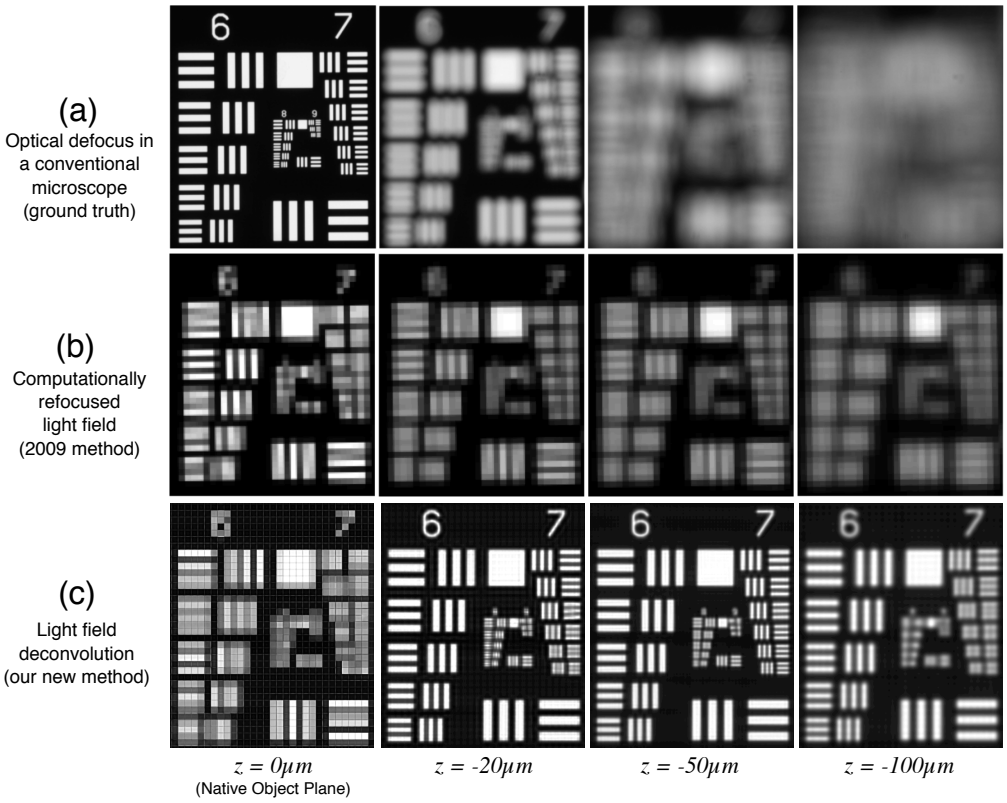

Fig. 1. USAF 1951 resolution test target translated to depths below the native object plane ($z = 0$ μm) and imaged using a light field microscope with a 20x 0.5NA water-dipping objective. **(a)** Photographs taken with a conventional microscope as the target is translated to the z-heights denoted below each image. **(b)** Computational re-focusing using our 2009 method [2] while the microscope was defocused to the same heights as (a). While some computational refocusing is possible, there has been a significant loss of lateral resolution. **(c)** The reconstruction algorithm presented in this paper brings the target back into focus, achieving up to an 8-fold improvement in lateral resolution compared to (b) except at the native object plane (left image).

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

optical signal below the Shannon-Nyquist limit, causing high frequency features to be aliased in the light field. In conventional imaging, such aliasing is undesirable because it irreversibly corrupts the recorded image. However, in light field imaging, aliasing is actually beneficial. In particular, the light field's angular samples, when projected into the volume, create a dense, non-uniform sampling pattern that allows us to "decode" aliased information during 3-D deconvolution.

The reconstruction technique we have developed is closely related to "computational super-resolution" methods in the field of computer vision [3]. This signal processing technique combines multiple under-sampled and aliased images of a scene to recover an image with sub-pixel (or in our case, sub-lenslet) resolution. Note that the computational super-resolution we describe here is not to be confused with optical "super-resolution" or "super-localization" methods in microscopy such as Structure Illumination Microscopy (SIM), Photo Activated Localization Microscopy (PALM), and Stochastic Optical Reconstruction Microscopy (STORM), which seek to surpass the diffraction limit. Computational super-resolution has recently been explored in light field photography, and shown to be effective at recovering resolution in the manner described above. However, these algorithms make assumptions typical at macroscopic photographic scales: namely opaque scenes and diffusely reflecting objects. They also model light using ray optics. These assumptions do not hold when imaging microscopic samples.

This paper builds on this past work, making several contributions. First, we present a wave optics model that accounts for the diffraction effects that occur when recording the light field with a microscope. We then explain how this optical model can be used in a deconvolution algorithm for 3-D volume reconstruction. This approach is suitable for fluorescence microscopy, where the volume to be reconstructed is largely transparent (i.e. with minimal scattering or absorption). Finally, we characterize the lateral resolution limit in light field microscopy, and discuss how optical design choices affect these resolution limits. To aid in exploring these trade-offs, we propose a theoretical band limit that should prove useful when choosing which objective, microlens array, and camera sensor to use for a given 3-D imaging scenario.

2. Background

2.1. 3-D imaging with the light field microscope

A conventional microscope can be converted into a light field microscope by placing a microlens array at the native image plane as shown in Fig. 2(a). Light field imaging can be performed using any microscope objective so long as the f-number of the microlens array is matched to the numerical aperture of the objective. This design choice ensures that the imaging circles behind the lenslets exactly tile the image plane without overlapping or leaving spaces in between them [1]. The camera sensor is placed behind the lenslets at a distance of one lenslet focal length—typically 1 to 10-mm. This is not practical for most commercial sensors, which are recessed inside the body of the camera, but it is possible to use a relay lens system placed in between the camera and the microlens array to re-image the lenslets' back focal plane onto the sensor. For example, we use a relay lens formed by a pair of photographic lenses mounted face-to-face via a filter ring adaptor.

Figures 2(b) and 2(c) depict simplified ray optics diagrams that are useful for building intuition about light propagation through the LFM. Figure 2(b) shows a ray bundle from a point emitter on the native object plane (i.e. the plane the is normally in focus in a widefield microscope). Rays emitted at different angles within the numerical aperture of the objective are recorded by separate pixels on the sensor. Summing the pixels shown in red yields the same measurement that would have resulted if the camera sensor itself had been placed at the native image plane (i.e. the image plane of a widefield microscope that is conjugate to the native

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

object plane) and binned to have a pixel that is the size of a lenslet. That is, summing the pixels behind each lenslet yields an image focused at the native object plane, albeit with greatly reduced lateral resolution equal to the diameter of one lenslet. Figure 2(c) shows how light is collected from a point emitter below the native object plane. Here the light is focused by the microlens array into a pattern of spots on the sensor. This pattern spreads out as the point moves further from the native object plane. Given a model of this spreading pattern, it is possible to sum together the appropriate light field pixels to produce a computationally refocused image (refer to [1] and [4] for more details).

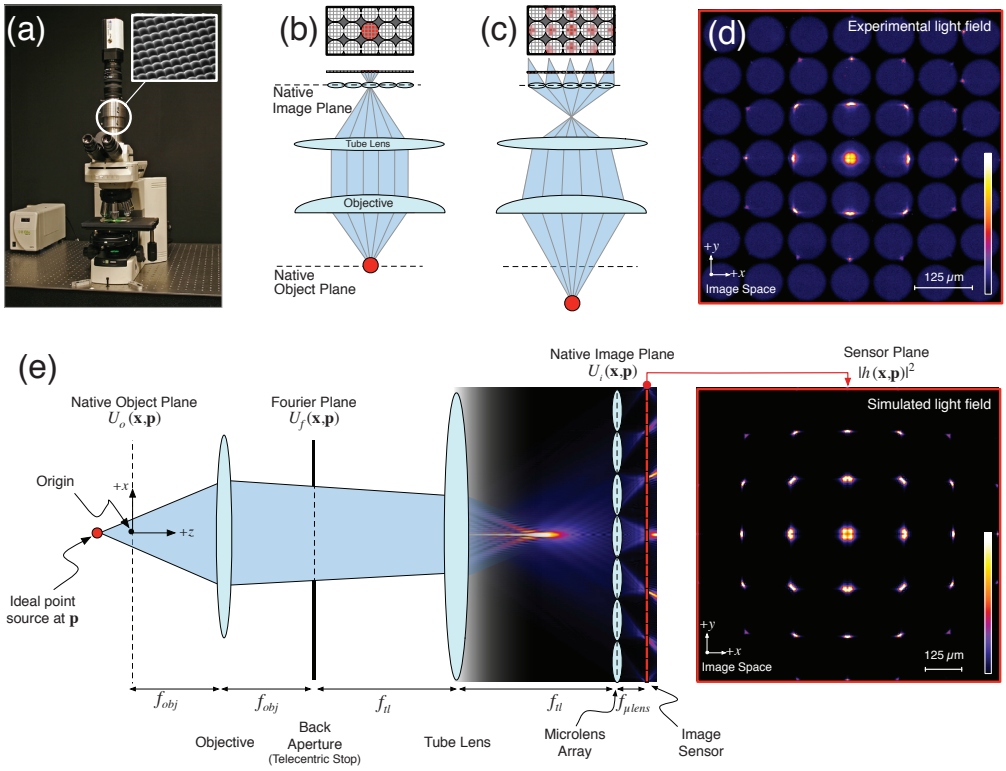

Fig. 2. Optical model of the light field microscope. **(a)** A fluorescence microscope can be converted into a light field microscope by placing a microlens array at the native image plane. **(b)** A ray-optics schematic indicating the pattern of illumination generated by one point source. The gray grid delineates pixel locations, and the white circles depict the back-aperture of the objective imaged onto the sensor by each lenslet. Red level indicates the intensity of illumination arriving at the sensor. For a point source on the native object plane (red dot), all rays pass through a single lenslet. **(c)** A point source below the native object plane generates a more complicated intensity pattern involving many lenslets. **(d)** A real light field recorded on our microscope using a 60x 1.4NA oil objective and a 125 μm pitch f/20 microlens array of a 0.5 μm fluorescent bead placed 4 μm below the native object plane. Diffraction effects are present in the images formed behind each lenslet. **(e)** A schematic of our wave optics model based on the LFM optical path (not drawn to scale). In this model, the microlens array at the native image plane is modeled as a tiled phase mask operating on this wavefront, which is then propagated to the camera sensor. The *xy* cross section on the far right shows the simulated light field generated at the sensor plane. The simulation is in good agreement with the experimentally measured light field in (d).

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

Of course, light is not collected in the simple manner that the schematics in Figs. 2(b) and 2(c) would suggest. Diffraction effects are evident upon inspection of a real light field recorded by a LFM (Fig. 2(d)), so a full wave optics treatment is necessary. Figure 2(e) shows a more realistic optical diagram, along with a simulated intensity distribution of light for a point source off the native object plane. This diagram was produced using our wave optics model. In Section 3 we will turn our attention to the details of this model and its implementation, but first we build more intuition about the 3-D reconstruction problem.

2.2. *Light field imaging as limited-angle tomography*

In light field microscopy, as with other imaging modalities, removing out of focus light from a reconstructed volume is accomplished by characterizing the microscope's point spread function (PSF) and using it to perform 3-D deconvolution. In the appendix of [1] it is shown that, for light field microscopy of a relatively transparent sample, 3-D deconvolution is equivalent to solving a limited-angle tomography problem. Figure 3(a) illustrates this connection. An oblique parallel projection through the sample is collected along chief rays (blue lines) corresponding to the pixels on the left-most side of each lenslet. If we take these blue pixels and use them to form an image, then we will have a low-resolution view of the specimen volume at a certain angle. In essence, this is the image that would be captured by placing a pinhole at a certain location in the back aperture of the microscope. Hence, we refer to this image as a "pinhole view."

A light field with $N \times N$ pixels behind each lenslet will contain pinhole views at N^2 different angles covering the numerical aperture of the microscope objective. This suggests that light field microscopy is essentially a simultaneous tomographic imaging technique in which all N^2 projections are collected at once as pinhole views. Thus, successful deconvolution amounts to

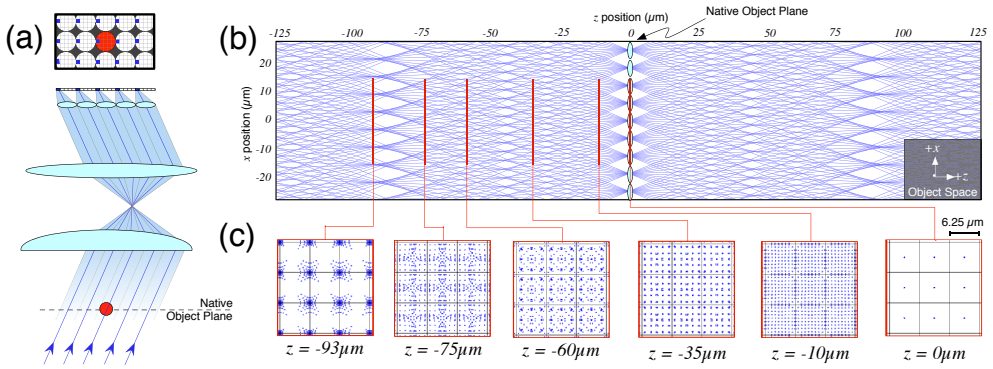

Fig. 3. Sampling of the volume recorded in a microlens-based light field. **(a)** A bundle of lenslet chief rays captured by the same pixel position relative to each lenslet (blue pixels) form a parallel projection through the volume, providing one of many angular views necessary to perform 3-D deconvolution. **(b)** When lenslet chief rays passing through every pixel in the light field are simultaneously projected back into the object volume, these rays cross at a diversity of x-positions (readers are encouraged to zoom into this figure in a PDF file to see how this pattern evolves with depth). This dense sampling pattern permits 3-D deconvolution with resolution finer than the lenslet spacing. The only place where this diversity does not occur is close to the native object plane; here resolution enhancement is not possible. **(c)** The distribution of the lenslet chief rays in xy cross-sections of the object volume changes at different distances from the native object plane. The outline of the lenslets are shown in light gray for scale. At $z = 0$ μm (rightmost image), the lack of spatial diversity in sample locations is evident.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

fusing these low resolution views to create a high resolution volumetric reconstruction.

How much resolution can be recovered in such a reconstruction? This depends in part on the band limit of the optical signal recorded by each pinhole view; i.e. the amount of low-pass filtering that the optical signal undergoes before it is sampled and digitized. In particular, each "pixel" in a pinhole view corresponds to one lenslet in the microlens array. Lenslets have a square aperture that acts as a (rect-shaped) band limiting filter. This induces optical blurring, and it is this blurring as well as the blurring due to diffraction that limits the highest spatial frequency that can be resolved in a pinhole view. When performing full 3-D deconvolution, high frequency details at a certain depth in the volume can only be reconstructed if they can be resolved in the pinhole views. Therefore, the band limits of pinhole views directly determine how much resolution can be recovered at any given z-plane.

In addition to the band limit, the spatial sampling rate within a pinhole view plays a key role in determining how high frequency content is recorded in the light field. In each pinhole view, the effective sampling rate is equal to the lenslet pitch divided by the objective magnification. For example, a 125 μm pitch array and a 20x microscope objective will result in a sampling period of 6.25 μm. When we present our results in Section 4, we will show that this samples the light field significantly below its band limit. As a result, high frequency details are aliased in individual pinhole views. While it may seem that such aliasing would be highly undesirable, it turns out that this is the key to enhancing resolution during 3-D deconvolution. To understand why this is the case, we now turn for insight to the field of computational super-resolution.

2.3. Aliasing and computational super-resolution

Computer vision researchers know that multiple aliased, low-resolution images of a scene can be computationally combined to recover a higher resolution image [3]. Early incarnations of this computational super-resolution involved capturing several images while a camera sensor was translated through random, sub-pixel movements. During each acquisition, high frequency information from the scene is aliased and recorded as low frequency image features in a way that is uniquely determined by the camera position. If the camera's trajectory can be accurately estimated, then the different, aliased copies of the scene can be combined to form a high resolution image. Although not required, deconvolution is often carried out as part of the reconstruction to de-blur the super-resolved image and further enhance its resolution [5].

However, as alluded to in the previous section, there are fundamental limits on the amount of recoverable resolution. In photography, the band limit is determined by the diffraction limit and the size and fill factor of the camera pixels [6]. Since modern image sensors with high fill factors are very close to band-limited, the achievable super-resolution factor is roughly 2x with deconvolution, although higher super-resolution factors can be achieved by leveraging priors that capture statistical properties of the image being reconstructed [7]. In microscopy, camera pixels are typically chosen to be smaller than the diffraction limit of the microscope objective, so little if any aliasing occurs and computational super-resolution is of limited use.

Of course, in order for high frequency information to be recoverable, enough distinct low resolution images (each with a different pattern of aliasing) must be recorded. This amounts to a sampling requirement: the sub-pixel shifts between the images must result in a sampling pattern that is dense enough to support reconstruction of the super-resolution image. An in-depth discussion of this sampling requirement is beyond the scope of this paper, but we refer the interested reader to [8].

Super-resolution methods have recently been explored in light field photography, with several papers demonstrating that a significant resolution enhancement can be achieved by combining aliased pinhole views [9, 10]. In microlens-based light fields, the geometry of the optics results in a fixed, sub-lenslet sampling pattern that takes the place of the sub-pixel camera movements

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

in traditional computational super-resolution. This pattern evolves with depth as depicted in Fig. 3(b). At z-planes where samples are dense relative to the spacing between the lenslets, it is possible to combine pinhole views to recover resolution up to the band limit. However, there are depths where the samples are redundant, most notably at the native object plane (although partial redundancy can also be seen in the figure at z = 72 μm and z = 109 μm). At these depths the sampling requirement may not be met, and super-resolution cannot always be fully realized. This is why the z = 0 μm plane in Fig. 1(c) remains a low-resolution, aliased image despite having been processed by our deconvolution algorithm.

3. Light field deconvolution

With this background in mind, we turn our attention to reconstructing a 3-D volume from a light field. In particular, we will solve the following inverse problem: given a light field recorded by a noisy imaging sensor, estimate the radiant intensity at each point in the volume that generated the observation. Concretely, we seek to invert the discrete linear forward imaging model,

$$\mathbf{f} = H\,\mathbf{g}, \tag{1}$$

where the vector $\mathbf{f}$ represents the light field, the vector $\mathbf{g}$ is the discrete volume being reconstructed, and H is a measurement matrix modeling the forward imaging process. The coefficients of H are largely determined by the point spread function of the light field microscope. Our first task will be to develop a model for this point spread function.

3.1. The light field PSF

The diffraction pattern generated by an ideal point source when passed through an optical system is the system's impulse response function, more commonly referred to as the point spread function. In an optical microscope this is the well-known Airy pattern (or its generalization as a double-cone for 3-D imaging) [11].

In the light field microscope a point source generates a more complex diffraction pattern. An example of one such "light field point spread function" is shown in Fig. 2(e). These patterns carry considerable information about the 3-D position of a point emitter in the volume, and this is the basis of our 3-D reconstruction algorithm. Unlike the Airy pattern, which is invariant with respect to the position of the point source, the light field PSF is translationally-variant. Specifically, the diffraction pattern behind the microlens array changes depending on the 3-D position of the point source. This gives rise to two challenges: (1) we must compute a unique PSF for each point in the volume, and (2) we cannot model optical blurring from the PSF as a convolution, as is commonly done in the case of conventional image formation models [12]. Instead, the wavefront recorded at the image sensor plane in Fig. 2(e) is described using a more general linear superposition integral

$$f(\mathbf{x}) = \int |h(\mathbf{x},\mathbf{p})|^2\, g(\mathbf{p})\, d\mathbf{p}, \tag{2}$$

where $\mathbf{p} \in \mathbb{R}^3$ is the position in a volume containing isotropic emitters whose combined intensities are distributed according to $g(\mathbf{p})$. When imaged, this volume gives rise to continuous 2-D intensity pattern $f(\mathbf{x})$ at the image sensor plane. The optical impulse response $h(\mathbf{x},\mathbf{p})$ is a function of both the position $\mathbf{p}$ in the volume being imaged as well as the position $\mathbf{x} \in \mathbb{R}^2$ on the sensor plane.

Notice that Eq. (2) is the continuous version of Eq. (1). We use the squared modulus of $h(\mathbf{x},\mathbf{p})$ in Eq. (2) because fluorescence microscopy is an incoherent, and therefore linear, imaging process. Although light from a single point emitter produces a coherent interference pattern

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

(and therefore the function $h(\mathbf{x},\mathbf{p})$ is a complex field containing both amplitude and phase information), coherence effects between any two point sources average out to a mean intensity level due to the rapid, random fluctuations in the emission time of different fluorophores. As a result, there are no interference effects when light from two sources interact; their contributions on the image sensor are simply the sum of their intensities.

Additionally, in the interest of practicality and computational efficiency we make two assumptions about the nature of the light and volume being imaged. First, we assume that the light is monochromatic with a wavelength λ. This approximation is reasonably accurate for fluorescent imaging with a narrow wavelength emission filter. Second, our model adopts the first Born approximation [13], i.e. it assumes that there is no scattering in the volume. Once emitted, the light from a point is assumed to radiate as a perfect spherical wavefront until it arrives at the microscope objective. This approximation holds up well when imaging a volume free of occluding or heavily scattering objects. However, performance does degrade when imaging deep into weakly scattering samples or into samples with varying indices of refraction. Modeling these effects is the subject of future work.

Figure 2(e) shows the optical path of the light field microscope in detail. With the objective and tube lens configured as a 4-f system, the back aperture of the objective serves as a telecentric stop, making the microscope both object side and image side telecentric. The focal length of the tube lens f_{tl} varies by microscope manufacturer ($f_{tl} = 200mm$ for our Nikon microscope), and the focal length of the objective can be computed from the magnification of the objective lens: $f_{obj} = f_{tl}/\mathrm{M}$.

An analytical model for the wavefront at the native image plane $U_i(\mathbf{x},\mathbf{p})$ generated by a point source at $\mathbf{p}$ can be computed using scalar Debye theory [11]. For an objective with a circular aperture, a point source at $\mathbf{p} = (p_1, p_2, p_3)$ produces a wavefront at the native image plane described by the integral,

$$U_i(\mathbf{x},\mathbf{p}) = \frac{\mathrm{M}}{f_{obj}^2 \lambda^2} \exp\left(-\frac{iu}{4\sin^2(\alpha/2)}\right) \int_0^\alpha P(\theta) \exp\left(-\frac{iu\sin^2(\theta/2)}{2\sin^2(\alpha/2)}\right) J_0\left(\frac{\sin(\theta)}{\sin(\alpha)} v\right) \sin(\theta)\, d\theta \quad (3)$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, and ρ is the normalized radius from the center of the pupil. The variables v and u represent normalized radial and axial optical coordinates,

$$\begin{aligned} v &\approx k\sqrt{(x_1-p_1)^2+(x_2-p_2)^2}\,\sin(\alpha) \\ u &\approx 4k\,p_3\,\sin^2(\alpha/2). \end{aligned}$$

The half-angle of the numerical aperture $\alpha = \sin^{-1}(\mathrm{NA}/n)$ and the wave number $k = 2\pi n/\lambda$ are computed using the emission wavelength λ and the index of refraction n of the sample. The function $P(\theta)$ is the apodization function of the microscope. For Abbe-sine corrected objectives, $P(\theta) = \sqrt{\cos(\theta)}$. Note that Eq. (3) only holds for low to moderate NA objectives. A vectorial diffraction theory [14, 15] could instead be substituted into our model to enable light field reconstruction with high NA objectives. A derivation of these equations and discussion of potential extensions, such as the ability to model aberrations, can be found in [11].

Having computed the scalar wavefront at the native image plane, we next account for the focusing of light through the microlens array. Our microlens arrays contain square-truncated

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

lenslets (meaning that their aperture is square) with a 100% fill factor. Consider a single lenslet centered on the optical axis with focal length $f_{\mu\text{lens}}$ and pitch d. This lenslet can be modeled as an amplitude mask representing the lenslet aperture and a phase mask representing the refraction of light through the lenslet itself:

$$\phi(\mathbf{x}) = \text{rect}(\mathbf{x}/d)\,\exp\left(\frac{-ik}{2f_{\mu\text{lens}}}\|\mathbf{x}\|_2^2\right). \tag{4}$$

The same amplitude and phase mask is applied in a tiled fashion to the rest of the incoming wavefront. Application of the full, tiled microlens array can be described as a convolution of a 2-D comb function with $\phi(\mathbf{x})$:

$$\Phi(\mathbf{x}) = \phi(\mathbf{x}) * \text{comb}(\mathbf{x}/d).$$

Next, the wavefront propagates a distance of $f_{\mu\text{lens}}$ from the native image plane to the sensor plane. The lenslets used in the LFM have a Fresnel number between 1 and 10, so Fresnel propagation is an accurate and computationally attractive approach for modeling light transport from the microlens array to the sensor [16, p.55]. The final light field PSF can thus be computed using the Fourier transform operator $\mathscr{F}\{\cdot\}$ as

$$h(\mathbf{x},\mathbf{p}) = \mathscr{F}^{-1}\left\{\mathscr{F}\left\{\Phi(\mathbf{x})\,U_i(\mathbf{x},\mathbf{p})\right\}\exp\left[-\frac{i}{4\pi}\lambda\, f_{\mu\text{lens}}\,(\omega_x^2+\omega_y^2)\right]\right\}, \tag{5}$$

where the exponential term is the transfer function for a Fresnel diffraction integral, and ω_x and ω_y are spatial frequencies along the x and y directions in the sensor plane.

3.2. Discretized optical model

We now describe how to discretize Eq. (2) to produce Eq. (1), which can then be solved on a digital computer. The sensor plane wavefront $f(\mathbf{x})$ is sampled by a camera containing N_p pixels. To simplify the notation, we re-order these pixels into a vector $\mathbf{f} \in \mathbb{Z}_+^{N_p}$. Similarly, during 3-D deconvolution the volume $g(\mathbf{p})$ is sub-divided into $N_v = N_x \times N_y \times N_z$ voxels and re-ordered into a vector $\mathbf{g} \in \mathbb{R}_+^{N_v}$ (see Fig. 4).

Although the dimensionality of $\mathbf{f}$ is fixed by the number of pixels in the image sensor, the dimensionality of $\mathbf{g}$ (i.e. the sampling rate of the reconstructed volume) is adjustable in our algorithm. Clearly, the sampling rate should be high enough to capture whatever information can be reconstructed from the light field. However, oversampling will lead to a rapid increase in computational cost without any real benefit. We will be explicit about our choice of volume sampling rates when we present results in Section 4, but here we will establish the following useful definition. A volume sampled at the "lenslet sampling period" has voxels with a spacing equal to the lenslet pitch d divided by the objective magnification M. For example, when imaging with a 125 µm pitch microlens array and a 20x microscope objective, the lenslet sampling period would be 6.25 μm. In this paper, we will sample the volume more finely at a rate that is a "super-sample factor" $s \in \mathbb{Z}$ times the lenslet sampling rate (where the lenslet sampling rate is the reciprocal of the lenslet sampling period). The sampling rate of the volume is therefore $s\text{M}/d$. Continuing our example, a reconstruction with super-sample factor $s = 16$ would result in a volume with voxels spaced by 0.39 μm. We refer to this as a volume that is sampled at 16x the lenslet sampling rate.

To complete our discrete model, we form a measurement matrix H whose coefficients h_{ij} indicate the proportion of the light arriving at pixel j from voxel i. Voxels in the reconstructed volume and pixels on the sensor have finite volume and area, respectively. Therefore the coefficients of H are computed via a definite integral over the continuous light field PSF,

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

$$h_{ij} = \int_{\alpha_j} \int_{\beta_i} w_i(\mathbf{p}) \, |h(\mathbf{x}, \mathbf{p})|^2 \, d\mathbf{p} \, d\mathbf{x} \tag{6}$$

where α_j is the area for pixel j, and β_i is the volume for voxel i. We assume that pixels are square and have have a 100% fill factor, which is a good approximation for modern scientific image sensors. Voxel i is integrated over a cubic volume centered at a point $\mathbf{p}_i$.

Equation (6) is a resampling operation, thus we must be careful not to introduce aliasing in the volume when discretizing Eq. (2). The resampling filter $w_i(\mathbf{p})$ that weights the PSF contribution at the center of a voxel more than the at its edges is introduced for this purpose. In our implementation we use a Hann (i.e. raised cosine) window with width equal to two times the volume sampling period, although a different resampling filter could be used if desired. We found the exact choice of resampling filter to be incidental to the algorithm; once the voxel sampling rate surpasses the band limit discussed in Sec. 2.2, the choice of resampling filter has little impact on the quality of the reconstruction.

Figure 4(a) shows that the columns of H contain the discrete versions of the light field point spread functions. That is, column i contains the *forward projection* generated when a single non-zero voxel i is projected according to Eq. (1). In Fig. 4(b) we see that the rows of H (or alternatively the columns of H^T) also have an interesting interpretation. We call these *pixel*

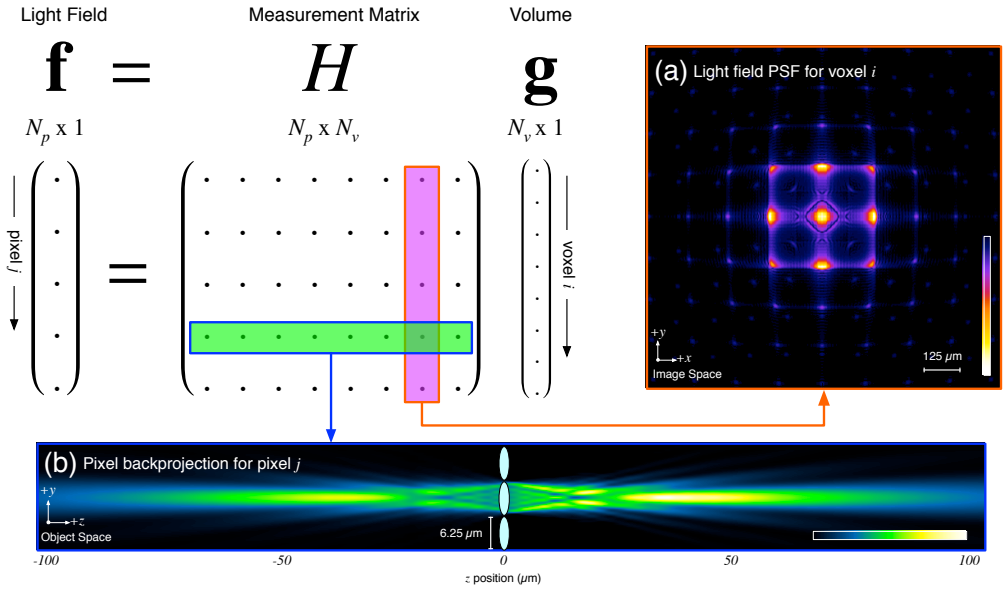

Fig. 4. The discrete light field imaging model, prior to accounting for sensor noise. The dimensionality of the light field **f** is fixed by the image sensor, but the dimensionality (or equivalently, the sampling rate) of the volume **g** is a user-selectable parameter in the reconstruction algorithm. **(a)** Column i of the measurement matrix H (purple) contains the discretized light field point spread function for voxel i, which corresponds to the forward projection of that point in the volume. **(b)** Row j of the measurement matrix (green) contains a pixel back projection: a visualization that shows how much each voxel in the volume contributes to a single pixel j in the light field. The cross sections in this figure were computed using our wave optics model for a 20x 0.5NA water dipping objective and a 125μm pitch f/20 microlens array.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

back projections by analogy to back projection operators in tomographic reconstruction algorithms, where the transpose of the measurement matrix is conceptualized as a projection of the measurement back into the volume. A pixel back projection from column j shows the position and proportion of light in the volume that contributes to the total intensity recorded at pixel j in the light field. In essence, the pixel back projection allows us to visualize the relative weight of coefficients in a single row of H when it operates on a volume $\mathbf{g}$.

3.3. Sensor noise model

Equation (1) is a strictly deterministic description of the light field imaging process and assumes no noise. We next consider how our imaging model can be extended to model the noise that is necessarily added by real sensor systems. With modern scientific cameras the dominant source of noise is photon shot noise, which means that the ith pixel follows Poisson statistics with a rate parameter equal to $\mathbf{f_i} = (H\mathbf{g})_i$ – i.e. the light intensity incident on the ith pixel. Read noise can be largely ignored for modern CCD and sCMOS cameras, although it can be added to the model below if desired. If we also consider photon shot noise arising from a background fluorescence light field $\mathbf{b}$ measured prior to imaging, then the stochastic, discrete imaging model is given by

$$\hat{\mathbf{f}} \sim \text{Pois}(H\,\mathbf{g} + \mathbf{b}), \tag{7}$$

where the measured light field $\widehat{\mathbf{f}}$ is a random vector with Poisson-distributed pixel values measured in units of photoelectrons e^-.

Due to the Poisson nature of the measurements, a light field with high dynamic range will have high variance at bright pixels and low variance at dark pixels. It is common in fluorescence microscopy to encounter bright objects against a dark field, and we have found in practice that it is critical that our model correctly account for this heterogeneity of variance across pixels.

3.4. Solving the inverse problem

Having now replaced the deterministic imaging model of Eq. (1) with its stochastic counterpart in Eq. (7), we can now perform 3-D deconvolution by seeking to invert the noisy measurement process. This can be formulated as a maximum likelihood estimation (MLE) problem in which the Poisson likelihood of the measured light field $\widehat{\mathbf{f}}$ given a particular volume $\mathbf{g}$ and background $\mathbf{b}$ is

$$\Pr(\widehat{\mathbf{f}}|\mathbf{g},\mathbf{b}) = \prod_i \left(\frac{(H\mathbf{g}+\mathbf{b})_i^{\widehat{\mathbf{f}}_i} \exp(-(H\mathbf{g}+\mathbf{b})_i)}{\widehat{\mathbf{f}}_i!} \right), \tag{8}$$

where $i \in \mathbb{Z}^{N_p}$ is the sensor pixel index. As the Poisson likelihood is log-concave, maximizing the negative log-likelihood over $\mathbf{g}$ and $\mathbf{b}$ yields a convex problem with the following gradient descent update:

$$\mathbf{g}^{(k+1)} = \text{diag}(H^T\mathbf{1})^{-1}\,\text{diag}(H^T\,\text{diag}(H\,\mathbf{g}^{(k)} + \mathbf{b})^{-1}\,\mathbf{f})\,\mathbf{g}^{(k)}, \tag{9}$$

where the $\text{diag}(\cdot)$ operator returns a matrix with the argument on the diagonal and zeros elsewhere. This is the well-known Richardson-Lucy iteration scheme.

With the exception of the measurement matrix H, whose structure captures the unique geometry of the light field PSF, this model is essentially identical to those that have been proposed in image de-blurring and deconvolution algorithms in astronomy and microscopy [17, 18]. However, due to the large size of both the imaging sensor and the volume being reconstructed (a typical imaging sensor will be on the order of several megapixels, and deconvolved volumes will often be sampled at tens to hundreds of millions of voxels), it is not feasible to store a dense matrix H in memory, much less apply it in the iterative updates of Eq. (9). We therefore must

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

exploit the specific structure of H in order represent it as a linear operator that can be applied to a vector without explicitly constructing a matrix.

Conveniently, the columns of H, which contain the discrete versions of the point spread functions as described in section 3.2, have sparse support on the camera sensor. Thus, H is sparse and can be applied efficiently using only its non-zero entries. More importantly, the repeating pattern of the lenslet array gives rise to a periodicity in the light field PSFs that dramatically reduces the computational burden of computing the coefficients of the H matrix. Consider a light PSF $h(\mathbf{x}, \mathbf{p_0})$ for a point $\mathbf{p_0} = (x, y, z)$ in the volume. The light field PSF $h(\mathbf{x}, \mathbf{p_1})$ for any other point $\mathbf{p_1} = (x + ad/\mathrm{M}, y + bd/\mathrm{M}, z)$ for any pair of integers $a, b \in \mathbb{Z}$ is identical up to a translation of (ad, bd) on the image sensor. Therefore, for a fixed axial depth, the columns of H can be described by a limited number of repeating patterns. Consequently, the application of the columns of H corresponding to a particular z-depth can be efficiently implemented as a convolution operation on a GPU. This accelerates the reconstruction of deconvolved volumes from measured light fields, with reconstruction times on the order of seconds to minutes depending on the size of the volume being reconstructed. Note that the relatively heavy-weight 3-D deconvolution algorithm is carried out as a post-processing step; acquisition of raw light field images can still be acquired at the frame rate of the camera sensor.

We close this section by noting that the model we have presented here resembles one recently introduced in [19]. Both are incoherent wave propagation models for light field imaging that make assumptions appropriate for fluorescence imaging. However, our approach differs from [19] in several ways: (1) The scope of [19] is photography, whereas ours is microscopy. (2) The model in [19] only considers imaging and reconstruction of the native object plane, whereas we show reconstructions for a variety of z-planes. (3) The formulation in [19] assumes white Gaussian image noise. While this is a reasonable assumption for a photographic light field camera, in microscopy it is important to treat image noise using the Poisson statistics we discussed above. Finally, (4) the results in [19] are based on computer simulations, whereas ours are based on empirical results that will be presented in the next section.

4. Experimental results

We now present experimental light field data processed using the wave optics model and 3-D deconvolution algorithm described above. In this section, the term "resolution" will be used to mean lateral resolution unless specified otherwise. We define resolution as the maximum spatial frequency appearing at a particular plane in the 3-D reconstructed volume with sufficient contrast that it is resolvable. In essence, we seek to measure the depth-dependent band limit of the light field microscope.

4.1. Experimental characterization of lateral resolution

Our experiments used an upright fluorescence microscope (Nikon Eclipse 80i) modified for light field imaging using a lenslet array (RPC Photonics) and digital camera sensor (QImaging Retiga 4000R). Two optical configurations were evaluated. In Figs. 5(b) and 5(c) we show results for a 20x 0.5NA water-dipping objective and a 125 μm pitch, f/20 microlens array. These lenslets have a square aperture and a spherical profile with a focal length of 2432 μm (at 525 nm). In Figs. 5(c) and 5(d), we show results for a 40x 0.8NA water-dipping objective and a 125 μm pitch, f/25 microlens array with a focal length of 3040 μm (at 525 nm). In Fig. 5(a) we show the baseline case: a widefield microscope imaging the target without any microlens array.

We imaged a high resolution US Air Force (USAF) 1951 test target (Max Levy DA052). The target was placed under the objective and trans-illuminated from below with a diffused halogen light source. We note that although the resolution target we used is not fluorescent (because we could not find a fluorescent target with sufficient spatial resolution for this characterization), the

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

fact that the light source emits incoherent light is sufficient for our imaging model to hold. We vertically translated the target in z from +100 μm to -100 μm relative to the native object plane in 1 μm increments and collected a light field for each z-plane. In other words, we deliberately mis-focused the microscope relative to the target, but then captured a light field rather than a simple photograph. The question, then, is to what extent we can computationally reconstruct the target (despite this mis-focus) using our 3-D deconvolution algorithm, and what resolution do we obtain?

3-D deconvolution was carried out as follows. For each light field with the USAF target at depth z, the Richardson-Lucy scheme described in Section 3.4 was run for 30 iterations to reconstruct a volume sampled at 16x the lenslet sampling rate. The "volume" being reconstructed was restricted to one z-plane known to contain the USAF target. In essence, this implicitly leveraged our prior knowledge of the z-position of the test target and the fact that it is planar (i.e. that there is no light coming from other z-planes). This approach, which we refer to as a "single-plane reconstruction," allows us to see how much resolution can be recovered at a particular z-plane under ideal circumstances. We will note that, although we knew the axial location of the target being reconstructed in these tests, this knowledge is probably not necessary. For example, our technique could be used for post-capture autofocusing when imaging a planar specimen at an unknown depth. This is a topic for future work.

Our method for analyzing these data follows the one described in [20]. We first registered each slice containing a deconvolved USAF pattern to a high-resolution ground-truth image of the USAF target, and then extracted the regions of interest (ROI) automatically (from group 6.1, representing a spatial frequency of $64\,[lp/mm]$ to group 9.3 with a spatial frequency of $645\,[lp/mm]$). For each ROI, the local contrast is calculated as

$$C_{thr} = (I_{max} - I_{min}) \,/\, (I_{max} + I_{min}), \tag{10}$$

where I_{max} and I_{min} are the maximal and minimal signal levels along a line drawn perpendicular to the stripes in each ROI. The final contrast measure is the the average of contrast for the horizontally and vertically oriented portion of the ROI.

Figure 5(b) shows, for single-plane reconstructions at different depths, the high resolution portion of the test target. Resolution is evidently depth-dependent, with peak resolution achieved at $z = -15\,\mu$m when imaging with the 20x configuration. Contrast in high resolution ROIs decreases gradually with increasing z-position. As previously discussed, enhanced resolution is not possible at the native object plane, and we can see this reflected in the reconstruction at $z = 0\,\mu$m where large, lenslet-shaped "pixels" are spaced equal to the lenslet sampling period. At the $z = -15\mu m$ plane in Fig. 5(b), one can also observe an apparent anisotropy in the resolution of the image. Spatial frequencies along the horizontal and vertical axis of the image are reconstructed at a higher resolution than those at other orientations. This artifact is due to the square aperture of our lenslets: the wider effective aperture measured across the diagonal of the lenslet results in a slightly lower band limit for spatial frequencies along that orientation. We note that this anisotropy could easily be avoided by using a microlens array with circular lenslet apertures.

The heat map in Fig. 5(c) shows the contrast of all ROIs as a function of z-position normalized so that the maximum contrast observed at any depth is assigned the value of 1.0. In essence, this shows the lateral modulation transfer function (MTF) as a function of depth in the 3-D reconstruction. To the left and right of z = 0 μm, we see an asymmetric dip in high frequency contrast. We conjecture that this dip in the MTF and apparent asymmetry around the native object plane is due to diffraction effects that play a particularly important role near to the native object plane. We believe that these irregularities in the MTF are related to the irregular intensity patterns in the pixel backprojections (see Fig. 4(b)), since they both occur over approximately the same z-range.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

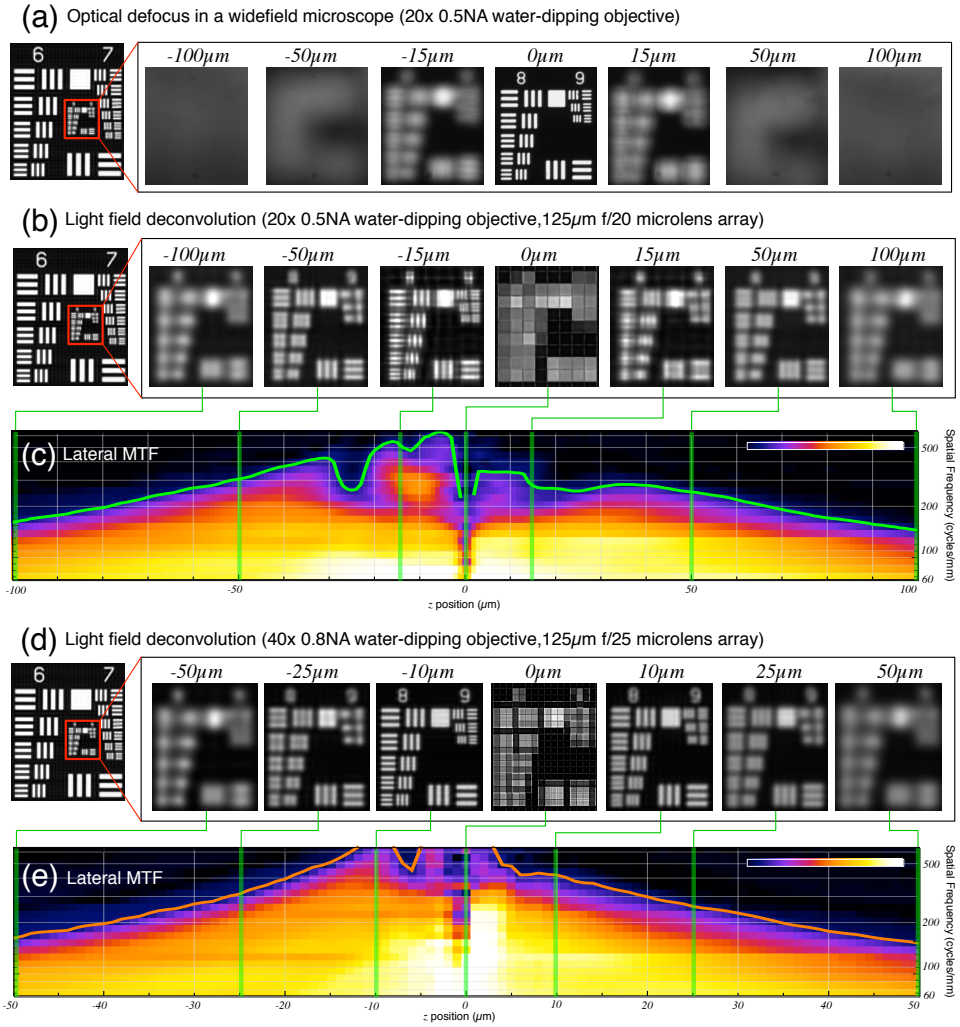

Fig. 5. Experimentally characterized resolution limits for two optical configurations of the light field microscope. **(a)** In a widefield microscope with no lenslet array, the target quickly goes out of focus when it is translated in z. **(b)** In a 3-D deconvolution from a light field, we lose resolution if the test target is placed at the native object plane (z = 0 μm), but we can reconstruct the target at much higher resolution than the spacing between lenslets when it is moved $z = -15$ μm (see also Fig. 1). Resolution falls off gradually beyond this depth (z = $\pm$50 μm and $\pm$100 μm). **(c)** Experimental MTF measured by analyzing the contrast of different line pair groupings in the USAF reconstruction. The colormap shows normalized contrast as measured using Eq. (10). The region of fluctuating resolution from $z = -30$ μm to 30 μm show that not all spatial frequencies are equally well reconstructed at all depths. **(d)** A slightly higher peak resolution ($z = \pm 10$ μm) can be achieved in the light field recorded with a 40x 0.8NA objective. However, the $z = \pm 25$ μm and $\pm$50 μm planes in (d) have the same apparent resolution as the $z = \pm 50$ μm and $\pm$100 μm planes in (b). **(e)** The experimental MTF for the 40x configuration shows that the region of fluctuating resolution (from -7.5 μm to 7.5 μm) is one quarter the size compared to (c). The solid green line in (c) and orange line in (e) are a 10% contrast cut-off representing the band limit of the reconstruction as a function of depth. Note that these plots are clipped to 645 cycles/mm, which is the highest resolution group on the USAF target.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

Similar trends can be seen in Figs. 5(d) and 5(e) for the 40x configuration, except that a slightly higher peak resolution is achieved at $z = -10\,\mu$m, and resolution falls of twice as quickly as in the 20x configuration (the z-range plotted in Fig. 5(b) is $\pm 50\,\mu$m, only half of that in Fig. 5(a)).

The solid green and orange lines superimposed on the MTF plots represent the band limit of the reconstruction. These lines are reproduced in Fig. 6, where they are plotted alongside the lenslet sampling rate (dotted black horizontal line) and the Nyquist rate at the diffraction limit of a conventional widefield fluorescence microscope (dashed blue horizontal line). Our key observations are (1) at its peak, the band limit we measure comes within a factor of 4x of the widefield diffraction limit; and (2) throughout most of the axial extent of these reconstructions, the resolution exceeds the lenslet sampling rate, and hence the resolution in our previous work [1], by 2x-8x.

These plots hint at trade-offs that might be made when choosing whether to image with the 20x configuration vs. the 40x configuration. For example, the 40x configuration achieves a higher resolution at its peak, but resolution falls off more rapidly with depth. The resolution fall-off in the 20x configuration is more gradual, but peak resolution is lower and there are more fluctuations in the resolution near the native object plane. We will now explore these trade-offs in more depth.

4.2. A theoretical band limit for the light field microscope

The experiments presented in Section 4.1 showed that there is a gradual, depth-dependent resolution fall-off in a light field reconstruction. In this section we explore how the choice of microscope objective and microlens array affect this fall-off. To aid in this discussion, we propose the following lateral resolution criterion for the light field microscope:

$$\nu_{\mathrm{lf}}(z) = \frac{d}{0.94\,\lambda\,\mathrm{M}\,|z|}. \tag{11}$$

In this equation, ν_{lf} is the depth-dependent band limit (in cycles/m), M and NA are the magnification and numerical aperture of the objective, λ is the emission wavelength of the sample,

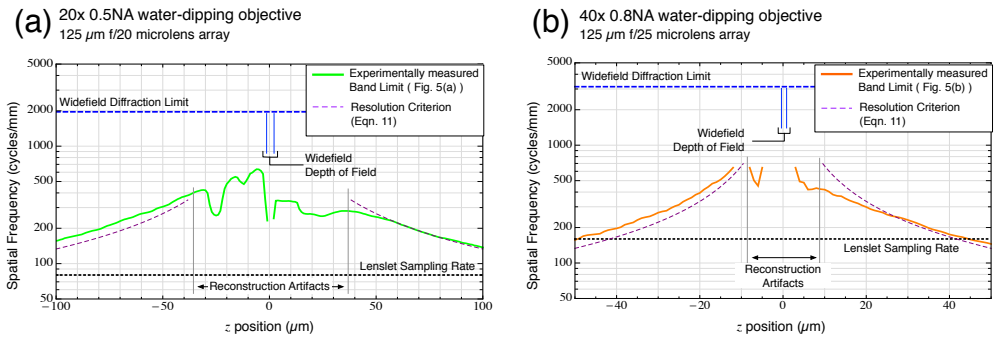

Fig. 6. Experimentally measured band-limits from Fig. 5 re-plotted along with the lenslet sampling rate (dotted black line) and the Nyquist sampling rate at the diffraction limit of a widefield fluorescence microscope (dashed blue line). For comparison, we have plotted the depth of field of a widefield microscope (thin blue lines). The theoretical band limit we propose in Eq. 11 (dotted purple line) is in good agreement with the experimental band limit. However, this criterion only predicts the resolution fall-off at moderate to large z-positions, and not near the native object plane where diffraction and sampling effects cause the band limit to fluctuate.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

d is the lenslet pitch, and z is the depth in the sample relative to the native object plane. The criterion applies only for depths where $|z| \geq d^2/(2\mathrm{M}^2\lambda)$. These equations, which are based on simple geometric calculations, are derived in Appendix 1.

The theoretical band limit of Eq. (11) is plotted as the purple dotted lines on Fig. 6, and is in good agreement with our experimentally determined band limit. However, its relative simplicity enables us to better understand the trade-offs we observed in our USAF experiments. As expected, the band limit predicted by Eq. (11) decreases in inverse proportion to z. More surprisingly, the predicted band limit does not depend on numerical aperture, as the diffraction limit of a widefield microscope does. Instead, the it is determined largely by the objective magnification and lenslet pitch. In Appendix 1, we explain why NA does not appear in the theoretical band limit. Here we will briefly mention that our microscope design assumes that NA is used to determine the optimal (diffraction limited) sampling rate behind each lenslet (i.e. the size of camera pixels relative to the pitch of the lenslet array). As such, NA plays an important role in determining the microscope's optical design, but once this design choice is made it has no direct impact on lateral resolution. We have also found that increasing NA improves axial resolution and signal to noise ratio in a 3-D light field reconstruction, just as it does it a widefield microscope. Thus, NA is still an important optical design parameter, just not one that effects the theoretical band limit presented here.

Equation (11) is plotted in Fig. 7(a) for several microscope objectives where the lenslet pitch has been fixed at 125 µm. Because Eq. (11) does not hold for small values of z, the resolution very near the native object plane cannot be predicted using the theoretical band limit. As we have seen in the USAF experiments, this region is often subject to reconstruction artifacts that arise due to diffraction and sampling effects. However, we can still use the criterion to understand the highest predicted resolution (which we will call "peak" resolution in this figure) as well as the resolution fall-off (which we will define as the rate at which relative resolutions between two optical recipes change as z is varied). With these definitions in mind, we make the following observations.

With pitch held constant, increasing the magnification results in higher peak resolution, but more rapid resolution fall-off as depth increases. Figure 7(b) shows a similar set of trends when lenslet pitch is varied but the magnification is fixed. In fact, a notable aspect of Eq. (11) is that the effect of doubling the objective magnification can be canceled out by halving the lenslet

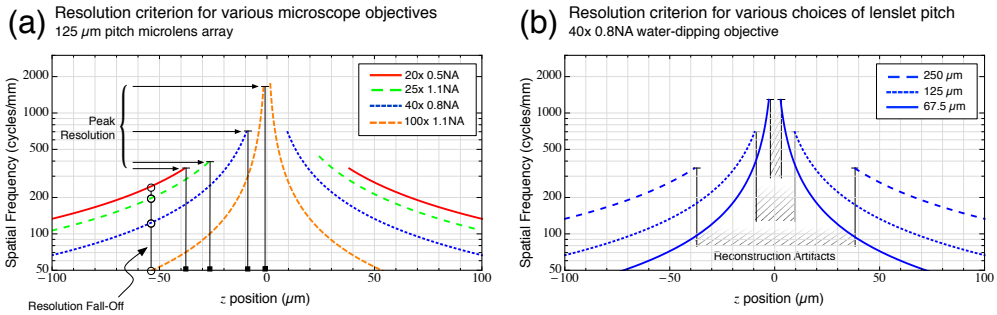

Fig. 7. Lateral resolution fall-off as a function of depth for various optical design choices. **(a)** For a fixed 125µm pitch lenslet array, a larger objective magnification results in better peak resolution, but a more rapid fall-off and hence a diminished axial range over which good resolution can be achieved. **(b)** For a fixed magnification factor, decreasing the lenslet pitch achieves the same trade-off as in (a). In fact, Eq. (11) shows that multiplying the lenslet pitch by some constant has the same effect on the resolution criterion as dividing the objective magnification by that same amount.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

pitch, and vice versa. For example, a 20x objective and 125 μm pitch microlens array would be expected to produce a reconstruction with the same resolution as with a 40x objective and a 250 μm pitch microlens array. However, the lateral field of view of the 3-D reconstruction does change when switching magnification factors or NA. Doubling the magnification factor will halve the field of view. Changing NA may change the field of view if camera pixels are magnified or demagnified (e.g. using relay optics) to achieve diffraction limited sampling. This suggests that the magnification factor and NA should first be selected to achieve a desired field of view, and then the lenslet pitch can be adjusted to trade off the rate of the resolution fall-off vs. the peak resolution achieved near the native object plane.

4.3. Reconstruction of a 3-D specimen

As an example of a more complicated 3-D specimen, we now present reconstructed light fields of pollen grains. Unlike the single-plane reconstructions in Section 4.1, these data were reconstructed as full 3-D volumes without any prior knowledge about the location of the specimen.

We collected these data on an inverted microscope (Nikon Eclipse TI) using a 60x 1.4NA oil objective with a 125 μm pitch f/20 microlens array. The light field was collected with the camera exposure time set to 40 ms. As a baseline for comparison, we first processed the light field using our previously published computational refocusing algorithm ("2009 method" [2]). These results are shown as maximum intensity projections in Fig. 8(a). This volume was reconstructed at the lenslet sampling rate exactly as in our previously published approach. The 3-D structure of the pollen grains can be made out in this computational focal stack; however contrast is low due to the presence of out-of-focus light.

Figure 8(b) we show the volume reconstructed using the 3-D deconvolution algorithm presented in Section 3. In this case, we sampled the volume at 8x the lenslet sampling rate. Processing at this high super-sampling rate took 137 minutes. High resolution features of the pollen grain that are not apparent in the focal stacks, such as spikes on the surface and chambers inside, can be readily discerned in the deconvolved volume. These differences are also clearly seen in Fig. 8(c) and Fig. 8(d), which show *xy* slices from the two respective reconstruction methods.

There are some reconstruction artifacts near the native object plane where the sampling constraint discussed in Section 2.3 is not met (see Fig. 3). Here, pinhole views in the light field contain largely redundant angular information that does not support recovery of resolution up to the band limit. However, despite being limited in resolution, the reconstruction at $z = 0$ μm is still improved relative to Fig. 8(a) thanks to the removal of out-of-focus light by the deconvolution algorithm. The highest resolution in the deconvolved volume is achieved at $z = -4$ μm and $z = 4$ μm. However, these planes are still close enough to the native object plane to be subject to some reconstruction artifacts. These artifacts are no longer present at $z = \pm 10$ μm, although the resolution here is already somewhat reduced from its peak at $z = \pm 4$ μm.

At 8x the lenslet sampling rate, the pollen grain volume pictured in Fig. 8(b) contains 648 x 648 x 61 voxels, or a total of roughly 25 million voxels. This is larger that the 4 million pixel measurements in the light field. As a result, the measurement matrix H is rank deficient and the inverse problem is underdetermined. Fortunately, the Richardson-Lucy algorithm favors sparse solutions to the inverse problem [22], and we believe that this provides a degree of regularization that allows it to produce reasonable volumes even in this underdetermined case. However, reconstructing this many voxels may not be necessary or efficient. While our experiments have shown that an 8x sampling rate may be needed to reconstruct the highest resolution z-planes in the volume, planes farther from the native object plane have a lower band limit and can therefore be sampled at a lower rate. In a more efficient implementation of our algorithm, one could choose a different sampling rate for each z-plane based on its band limit. This would dramatically lower the total number of voxels to be reconstructed. The measurement matrix would then

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

have fewer columns, resulting in a better conditioned inverse problem and a better performing algorithm.

Relatedly, the question arises: how much information can be recovered in a 3-D reconstruction from a light field captured my a camera sensor with a fixed number of pixels. Lacking the 3-D equivalent of a USAF test target, it is difficult to make a quantitative comparison between the resolution of this 3-D reconstruction and the resolution we achieved when performing the single plane reconstructions described in Section 4.1. Qualitatively, we observe about a 2-4x

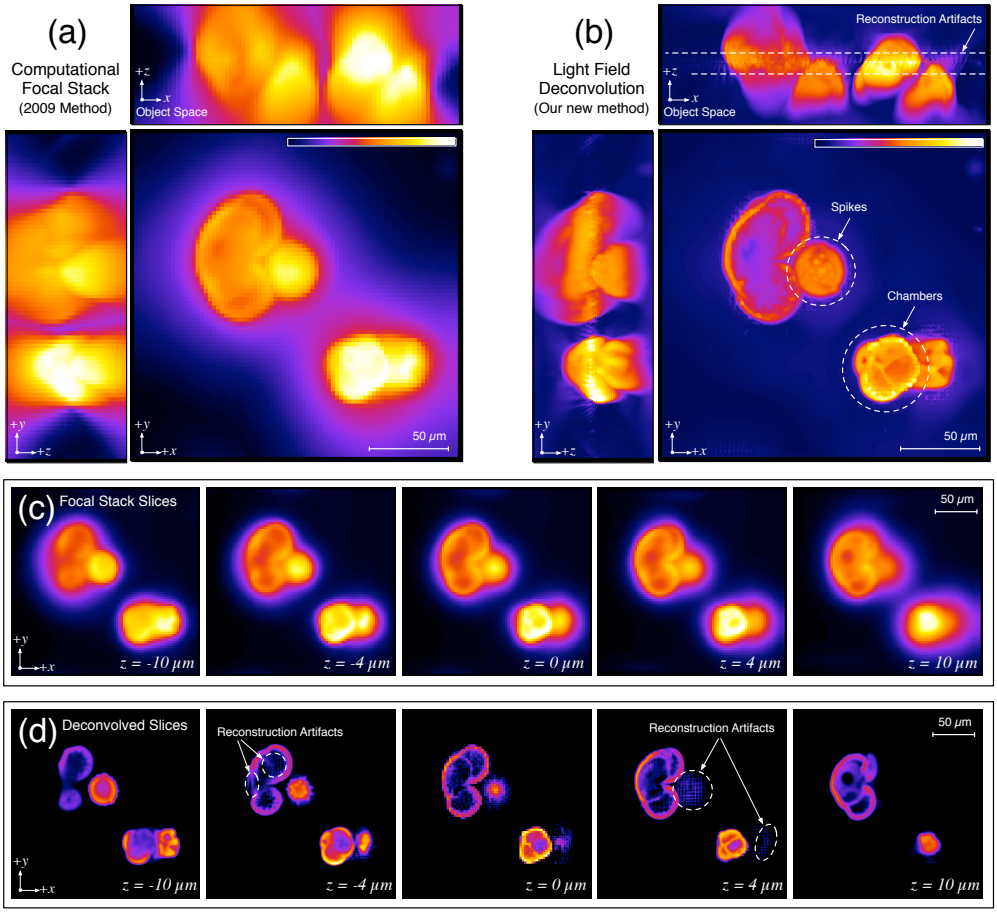

Fig. 8. Pollen grains imaged with a 60x 1.4NA oil dipping objective and a 125 μm f/20 microlens array. **(a)** Max-intensity projections of a volume reconstructed using the computational refocusing algorithm presented in our 2009 paper [2] shows the irregular shape of the pollen grains, but low contrast and little high frequency detail. **(b)** Maximum intensity projections of a volume reconstructed using the 3-D deconvolution algorithm introduced in this paper shows the structures of the pollen grain more clearly. A small region of reconstruction artifacts appears around the native object plane. **(c)** Individual *xy* slices from the computational refocusing algorithm. **(d)** Slices from the 3-D deconvolved volume. A gamma correction of 0.6 was applied to panels (a), (b), (c), and (d) to help visualize their full dynamic range. After gamma correction, each panel was separately normalized so that the 99% of the intensity range was represented by the colormap shown by the scale bar.

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

improvement over the lenslet resolution in Fig. 8(b). While useful, this is less than the 8x improvement in our single plane reconstructions. Why is this so? We conjecture that there is a firm relationship between the number of samples in a measured light field and the number of voxels one can expect to reconstruct from it without the application of priors. One can think of the former as a count of the degrees of freedom present in the captured light field [21]. Analysis of this conjecture is a topic of ongoing work.

5. Conclusion and future directions

In this paper we have presented a wave optics model for light propagation through the light field microscope, and we have shown how it can be used to produce deconvolved single plane reconstructions and 3-D volumes at a higher spatial resolution than the sampling period of the lenslet array. We have experimentally characterized the band limit of deconvolved light fields and found that lateral resolution decreases in inverse proportion to distance from the native object plane of the microscope. The theoretical band limit in Eq. (11) suggests that the magnification and pitch of lenslet array are optical design parameters that can be adjusted to change the rate of this resolution fall-off.

Even with the improved spatial resolution presented here, light field microscopy does not achieve the diffraction limited spatial resolution of other imaging modalities, such as confocal, 2-photon, or widefield deconvolution microscopy. However, these methods all require acquiring many observations over time, so they are not well-suited to recording dynamic phenomenon on the time scale of milliseconds. With light field imaging, an entire 3-D volume can be reconstructed from a single light field captured in a single exposure of the camera sensor. Therefore, frame rate alone determines how rapidly full 3-D volumes can be imaged. Hence, light field microscopy is well-suited to high-speed volumetric imaging where the object(s) of interest are inherently three-dimensional, have sub-second time dynamics, and are larger than the diffraction limit.

Today, light field imaging is possible thanks to the availability of low-noise, high pixel count, high frame-rate scientific image sensors. As pixel counts on these sensors continues to increase, it will become easier to simultaneously achieve dense angular sampling (in terms of the number of pixels per lenslet) and a wide lateral field of view. Such improvements will be a practical necessity for light field imaging at high resolution over large 3-D volumes. Improvements in the performance of general-purpose graphics processing hardware will also decrease the time it takes to run computationally expensive post-processing algorithms like the 3-D deconvolution algorithm presented in this paper.

There are several limitations and future avenues of research that are worth mentioning. The aliasing and low resolution that occurs near the native object plane is problematic, as it separates and isolates the regions above and below the native object plane that have relatively high lateral resolution. This limitation could be circumvented by splitting the optical path and imaging with two microlens arrays and two cameras focused on slightly different z-planes in the volume. One such design using a pair of prisms was recently proposed for light field cameras [23]. Alternatively, a light field could be captured along with a normal, high-resolution widefield image. This would improve resolution at the native object plane and possibly at other planes as well if the two were combined as proposed in [24]. Finally, a lenslet array could be placed at the native image plane of a multi-focal microscope [25] to create many overlapping regions of high resolution and extend the useful axial extent of the 3-D reconstruction.

More generally, lateral resolution could be improved (perhaps even beyond the limits discussed in this paper) by incorporating priors into the reconstruction algorithm. Work on 3-D deconvolution in light field photography suggest that this may prove to be particularly fruitful avenue for future research [9]. Finally, the imaging model presented here is only suitable when

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

imaging a sample emitting incoherent light. A generalization of the wave optics model that accounts for illumination with (partially) coherent light sources, refraction and scattering in the sample, or polarization effects would enable fundamentally new 3-D imaging modalities with the light field microscope.

Appendix 1: Derivation of the theoretical band limit

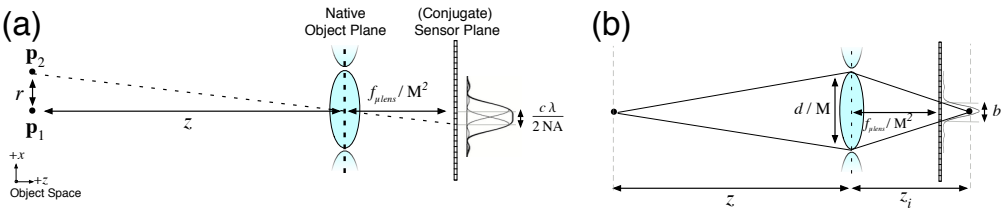

Fig. 9. Geometric construction of the light field resolution criterion introduced in Eq. (11). In these figures conjugate images of the microlens array and image sensor are depicted on the object side of the microscope. Taking the magnification factor into account, the effective object-side pitch of the conjugated lenslet is d/M and its effective focal length is $f_{\mu\mathrm{lens}}/\mathrm{M}^2$. **(a)** The resolution criterion holds when point sources $\mathbf{p}_1$ and $\mathbf{p}_2$ are at sufficient depth $|z|$ that they form diffraction limited spots on the sensor. Under this condition, the discriminability of the spots can be determined by simple geometric construction using similar triangles. Here, c is a constant that allows us to specify a particular resolution criterion (e.g. , $c = 1.22$ selects the Rayleigh 2-point criterion, and $c = 0.94$ selects the Sparrow 2-point criterion [26, p. 340]). In this paper, we use the Sparrow 2-point criterion since it is best suited for measuring where contrast drops to zero between two point projections measured in a digital image [26, p. 340]. **(b)** To see where this resolution criterion holds, we measure the diameter b of the blur disk predicted using geometric optics. Our approximation is valid when this diameter is less than the diameter of a diffraction-limited spot (i.e. when $b < c\,\lambda/\mathrm{NA}$).

Here we derive the theoretical band limit introduced as Eq. (11) in Section 4.2. Figure 9(a) shows the geometric intuition that gives rise to this criterion. Consider a point at $\mathbf{p}_1$ at a depth z on the object side of microscope. In the limit as $z \to \infty$, the image of this point behind the lenslet approaches a diffraction-limited spot whose size is determined by the numerical aperture of the system. Now consider a second point $\mathbf{p}_2$ at the same depth, but displaced by a distance r from the optical axis. If $|z|$ is large enough to produce two diffraction limited spots, the two points will be just barely distinguishable if they are separated by a distance determined by an appropriate 2-point resolution criterion. This occurs when

$$\frac{r}{|z|} = \frac{c\,\lambda}{2\,\mathrm{NA}} \cdot \frac{\mathrm{M}^2}{f_{\mu\mathrm{lens}}}. \tag{12}$$

To place this formula in a more generally useful form, we make the assumption that the f-number of the lenslet array is matched to the NA of the microscope objective [1]. For an objective satisfying the Abbe sine condition, the cut-off frequency for incoherent imaging is $\nu_{obj} = 2\,\mathrm{NA}/\lambda$ (see [27, p. 143]). This limit is set by the diameter of the microscope objective back aperture, i.e. the exit pupil of the system. Lenslets are focused on the back aperture, and thus each lenslet forms an image of this exit pupil. If their numerical apertures are matched, then $\nu_{obj} = \mathrm{M}\,\nu_{\mu\mathrm{lens}}$, where $\nu_{\mu\mathrm{lens}} = d/(\lambda\, f_{\mu\mathrm{lens}})$ is the maximum spatial frequency that can be

#193936 - $15.00 USD Received 16 Jul 2013; revised 30 Sep 2013; accepted 4 Oct 2013; published 17 Oct 2013

represented on the focal plane behind a lenslet for coherent imaging [27, p. 103]. Combining these three equations and solving for the lenslet focal length yields,

$$f_{\mu\text{lens}} = \frac{d\,\mathrm{M}}{2\,\mathrm{NA}}. \tag{13}$$

This is the focal length of a lenslet with pitch d that has been matched to the numerical aperture of a given microscope objective. Substituting this into Eq. (12) yields the distance r between $\mathbf{p}_1$ and $\mathbf{p}_2$ where they would be just discernible as two points on the image sensor: $r = \mathrm{c}\,\lambda\,\mathrm{M}\,|z|/d$.

Finally, we compute the theoretical band limit by taking the reciprocal of r. This is the spatial frequency where we can just barely discern features spaced a distance r at depth z.

$$\nu_{\mathrm{lf}}(z) = \frac{d}{c\,\lambda\,\mathrm{M}\,|z|}.$$

It is noteworthy that combining Eq. (12) and Eq. (13) has caused NA to cancel out. The lack of dependence on NA arises from our design choice to match the f-number of the microlens array with the numerical aperture of the objective. In doing so, we have implicitly determined number of pixels behind each lenslet needed to achieve diffraction-limited sampling. Although the optimal number of pixels per lenslet increases proportionally with NA, the range of angles collected by each pixel (i.e. the angular sampling rate) remains fixed. It is this angular sampling rate that determines our ability to discriminate between points on the sensor when z is relatively large. Since the angular sampling rate is independent of NA, so too is our theoretical band limit.

Finally, Fig. 9(b) shows the conditions under which the criterion is valid; i.e. where $|z|$ is sufficiently large that $\mathbf{p}_1$ and $\mathbf{p}_2$ form diffraction limited spots. This occurs approximately where the diameter of the blur disk b predicted via geometric optics is less than the diameter of a diffraction limited spot,

$$b < \frac{c\,\lambda}{\mathrm{NA}}. \tag{14}$$

Using similar triangles, we can compute $b = d\,(z_i - f_{\mu\text{lens}}/\mathrm{M}^2)/(\mathrm{M}\,z)$. Substituting this along with the lensmaker's formula $1/z_i + 1/z = \mathrm{M}^2/f_{\mu\text{lens}}$ and Eq. (13) into Eq. (14), we see that the criterion applies only for depths where $|z| \geq d^2/(2\,\mathrm{M}^2\,\lambda)$.

Acknowledgments

This work has been supported by NIH Grant #1R01MH09964701 and NSF Grant #DBI-0964204. M.B is supported by a National Defense Science & Engineering Graduate fellowship, L.G by NSF Integrative Graduate Education and Research Traineeship (IGERT) fellowship, S.Y. by NSF Graduate Research fellowship, A.A by the Helen Hay Whitney Foundation. We would also like to gratefully acknowledge the following people who over the course of this research have shared their time, ideas, and resources with us: Sara Abrahamson, Todd Anderson, Adam Backer, Matthew Cong, Mark Horowitz, Isaac Kauvar, Ian McDowell, Shalin Mehta, Dave Nicholson, Rudolf Oldenbourg, John Pauly, Judit Pungor, Tasso Sales, Stephen J. Smith, Rainer Heintzman, Ofer Yizhar, Andrew York, and Zhengyun Zhang.