Light Field Microscopy

Marc Levoy [1] Mark Horowitz [2] Ren Ng [1] Andrew Adams [1] Matthew Footer [3]

[1]Computer Science Department
Stanford University

[2]Electrical Engineering Department
Stanford University

[3]Department of Biochemistry
Stanford University

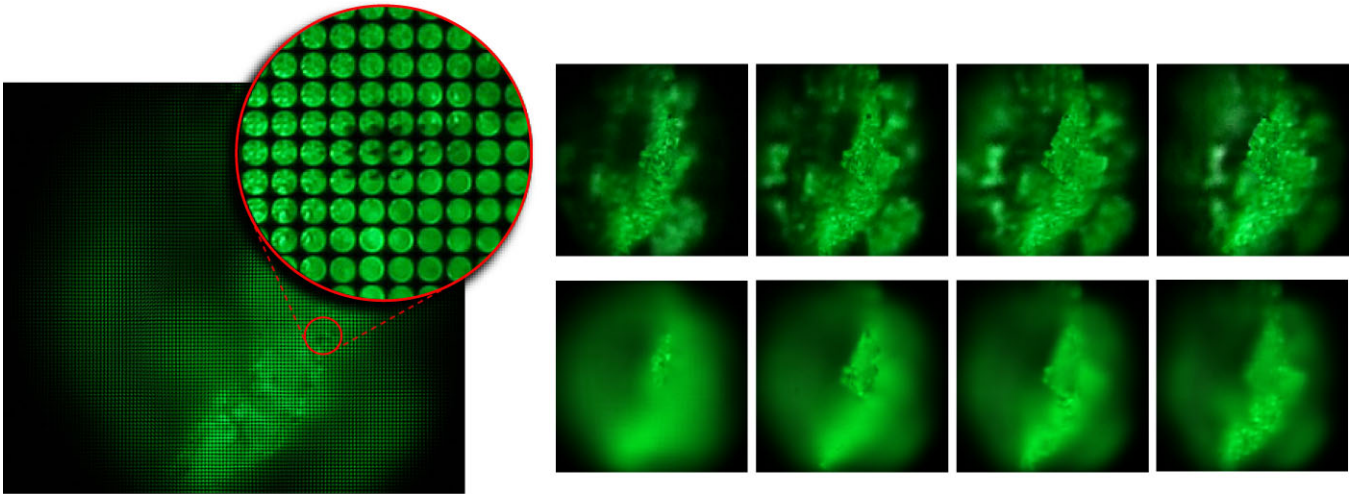


Figure 1: At left is a light field captured by photographing a speck of fluorescent crayon wax through a microscope objective and microlens array. The magnification is 16×, and the field of view is 1.3mm wide. The image consists of 170^2 subimages, one per microlens, each depicting a different part of the specimen. An individual subimage contains 20^2 pixels, each representing a different point on the objective lens and hence a unique direction of view of the specimen. By extracting one pixel from each subimage, we can produce perspective views of the specimen, a sequence of which is shown at top-right. Alternatively, by summing the pixels in each subimage, we can produce orthographic views with a shallow depth of field, like an ordinary microscope but of lower spatial resolution. Shearing the light field before summing, we can focus at different depths, as shown in the sequence at bottom-right. These images were computed in real-time on a PC.

Abstract

By inserting a microlens array into the optical train of a conventional microscope, one can capture light fields of biological specimens in a single photograph. Although diffraction places a limit on the product of spatial and angular resolution in these light fields, we can nevertheless produce useful perspective views and focal stacks from them. Since microscopes are inherently orthographic devices, perspective views represent a new way to look at microscopic specimens. The ability to create focal stacks from a single photograph allows moving or light-sensitive specimens to be recorded. Applying 3D deconvolution to these focal stacks, we can produce a set of cross sections, which can be visualized using volume rendering. In this paper, we demonstrate a prototype light field microscope, analyze its optical performance, and show perspective views, focal stacks, and reconstructed volumes for a variety of biological specimens. We also argue that synthetic focusing followed by 3D deconvolution is equivalent to applying limited-angle tomography directly to the 4D light field.

Keywords: Light fields, synthetic aperture, microscopy, deconvolution, tomography, volume rendering.

1. Introduction

Microscopes are the primary scientific instrument in many biological laboratories. Although their performance and ease of use have improved dramatically over their 400-year history, microscopes suffer from several limitations. First, diffraction limits their spatial resolution, especially at high magnifications. This limit can be ameliorated by enlarging the lens opening (called the numerical aperture) while keeping the lens strongly curved, but we reach a practical limit when the lens becomes a half-sphere. Second, in a microscope objects are seen in orthographic projection from a single direction (see figure 3). Moving the specimen laterally on the microscope stage does not produce parallax, making it hard to disambiguate superimposed features. Third, microscopes have a very shallow depth of field, particularly at high magnifications and numerical apertures. This "optical sectioning" is useful when viewing thick specimens, but examining the entire specimen requires moving the stage up and down, which is slow and may not be possible on live or light-sensitive specimens.

While the first limitation is intrinsic to the physics of light, the others arise from the current design of microscopes. These limits can be removed - albeit at a cost in spatial resolution - by capturing light fields instead of images. The scalar light field is defined as radiance as a function of position and direction in free space. The problem of recording the light field in a microscope was first addressed by Gabor using the interference of two coherent light beams [1948]. Subsequent development of the laser made his

technique practical, leading to holography and holographic microscopy [Ellis 1966]. Unfortunately, holomicrographs record not only the specimen but also the interior of the microscope, limiting its usefulness [Inoué 1997]. Researchers have also proposed capturing sequences of images in which the direction of illumination [Chamgoulov 2004] or direction of view [Kawata 1987] is restricted in each image; however, these methods are slow.

The use of lens arrays to capture light fields has its roots in Lippman's 1908 invention of integral photography. It is beyond the scope of this paper to survey that field; the reader is therefore referred to [Okoshi 1976]. Of particular relevance here are arrays based on microlenses (smaller than 1mm). Spurred by improvements in the technology for manufacturing these arrays, researchers have proposed using them for computer vision [Adelson 1992], 3D video [Javidi 2002] and photography [Ng et al. 2005]. In this paper, we show that by placing such an array at the intermediate image plane of a microscope, we can capture a light field of the specimen in a single photograph. As in Lippman's original proposal, we sacrifice spatial resolution to obtain angular resolution. We can then employ light field rendering [Levoy 1996] to generate perspective flyarounds, at least up to the angular limit of rays we have captured. Similarly, we can use synthetic focusing [Isaksen 2000] to produce a focal stack, a sequence of images each focused on a different plane.

Most devices for capturing light fields operate on macroscopic scenes (centimeters to meters). In a microscope our scenes are 3-4 orders of magnitude smaller. This change in scale has two important implications:

- In macroscopic scenes, light field capture and display can be analyzed using geometrical optics. Under the microscope wave optics must also be considered. As we shall see in section 3, *diffraction places an upper limit on the product of lateral and axial resolution in a light field microscope.*
- In macroscopic scenes, most objects scatter light, making them opaque. Under the microscope scattering no longer dominates, and most objects become partially transparent. This means that while the 3D structure of macroscopic scenes can only be analyzed using computer vision algorithms (such as shape-from-stereo), *the 3D structure of microscope light fields can be analyzed using algorithms for reconstruction from projections.*

Two algorithms in this class are tomography and 3D deconvolution. In the latter, the observation is made that each slice in a focal stack contains blurred contributions by features off the plane on which it was focused, and if we know the nature of this blurring, we can apply inverse filtering to remove it [Agard 1984]. Although sensitive to noise, 3D deconvolution can be performed using images from ordinary microscopes, making it an inexpensive alternative to laser confocal microscopy [Corle 1996]. In this paper we show that focal stacks produced synthetically from light fields can also be deconvolved, as described in section 4. The number of slices in the resulting volume is limited by the axial resolution of the light field, but volumes can be produced from a single photograph. Alternatively, since a microscope light field is a set of projections over a range of viewing angles, limited-angle tomography [Kak 1988] can be applied directly to the light field, producing a volume. As we show in Appendix A, 3D deconvolution and limited-angle tomography − techniques that developed in different fields − are really the same thing.

2. The light field microscope

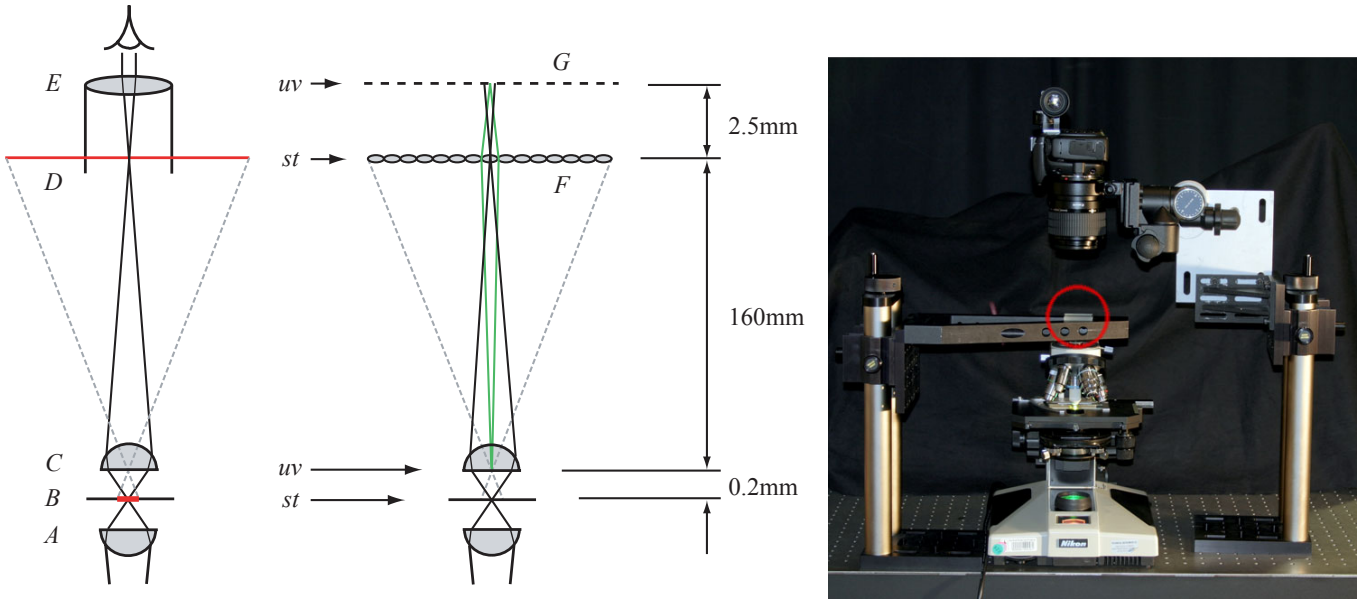
If one places a sensor behind an array of small lenses (lenslets), each lenslet records the scene observed from that position on the array. Interpreted as a two-plane light field, its angular, or *uv*, resolution depends on the number of lenslets, and its spatial or *st*, resolution depends on the number of pixels behind each lenslet. Placing a "field" lens on the object side of the lenslet array, and positioning this lens so that the scene is focused on the array, transposes the light field. Now its spatial resolution depends on the number of lenslets and its angular resolution on the number of pixels behind each lenslet.

The first arrangement has the advantage of being physically thin. However, humans are more tolerant of low angular resolution than low spatial resolution, and it is easier to build one field lens of high quality than thousands of lenslets of high quality, so if system thickness is not an issue, the latter arrangement is preferred. For a microscope, where resolution of the image is critical, a field lens is essential. This choice also allows us to reuse the most critical and highly optimized component in a conventional microscope, its objective (see figure 2(a)). The microscope objective is a highly-corrected optical subsystem capable of capturing rays that strike it at angles of up to 70 degrees from the optical axis. It would be challenging to replace this subsystem with an array of microlenses and achieve satisfactory optical performance. We therefore decided to leave the objective alone and place a microlens array at the intermediate image plane, as shown in figure 2(b).

Figure 2(c) shows our prototype light field microscope. The microlens array was manufactured to our specifications by Adaptive Optics, molded in epoxy, and mounted on an optical-quality glass window. The microlenses were plano-convex and square in plan view. The array size, microlens curvature, and microlens size are discussed in section 3. Microscopists normally employ a cooled-pixel camera to gather more light, especially in fluorescence studies, but we were not light-starved in our proof-of-concept experiments, and we needed higher resolution than is easily available in these cameras, so we employed a 12.8-megapixel digital SLR and used long exposures when necessary.

Figure 1 shows an example light field recorded by our prototype after color demosaicing, rotation to axis-align the microlens subimages, and scaling to make them an integral number of pixels on a side. Each subimage records $L(\bullet, \bullet, s, t)$, and within a subimage each pixel records $L(u, v, \bullet, \bullet)$. To compute perspective flyarounds and focal stacks, we used a real-time synthetic aperture focusing program. Example images are shown in figures 1 and 3. In the latter figure, note that the oblique view (c) is noisier than the focused view (d). This is because the former uses only one pixel from each microlens, while the latter uses all of them. This point is worth stressing; our synthetically focused views use all the light present in the microscope. They sacrifice spatial resolution relative to an ordinary microscope, but they do not sacrifice light-gathering ability.

The optical systems of microscopes differ from those of cameras, making our light field microscope different than a light field camera in several ways. Two of these differences are particularly worth discussing:



(a) ordinary microscope (b) light field microscope (c) our prototype

Figure 2: Optical layout of our light field microscope. (a) In a transmission-mode light microscope, an illumination source is focused by a condenser lens at *A* onto a specimen at *B*. An objective lens at *C* magnifies the specimen, creating a real image at intermediate image plane *D*. In older microscopes, this plane is located inside the microscope tube. An ocular (eyepiece) at *E* further magnifies the central portion of this image, creating a second image focused at infinity. (b) In our design the ocular is removed, a microlens array *F* is placed at the intermediate image plane, and a camera sensor is placed behind this at *G*, positioned so that each microlens records an in-focus image of the objective (green rays). In light field parlance, if the objective aperture and specimen constitute the *uv* and *st* planes, then the camera sensor and microlens array constitute a reimaging of these two planes. This drawing is not to scale; typical distances are shown beside it. (c) Our prototype consists of a Nikon Optiphot and custom microlens array (red circle). To avoid building a special camera, we re-image *G* using a Canon 5D 35mm SLR with a 1:1 macro lens.

Microscopes are orthographic. In a normal photographic lens, all portions of the lens participate in the imaging of each point in the scene (figure 3(a)). Microscope objectives, by contrast, incorporate a stop (a physical ring) one focal length behind the lens (figure 3(b)). This makes them *object-space telecentric* [Kingslake 1983], which means:[1]

(1) microscopes produce orthographic views, not perspective,

(2) translating the stage in x or y provides no parallax,

(3) features do not shift position when they come into focus,

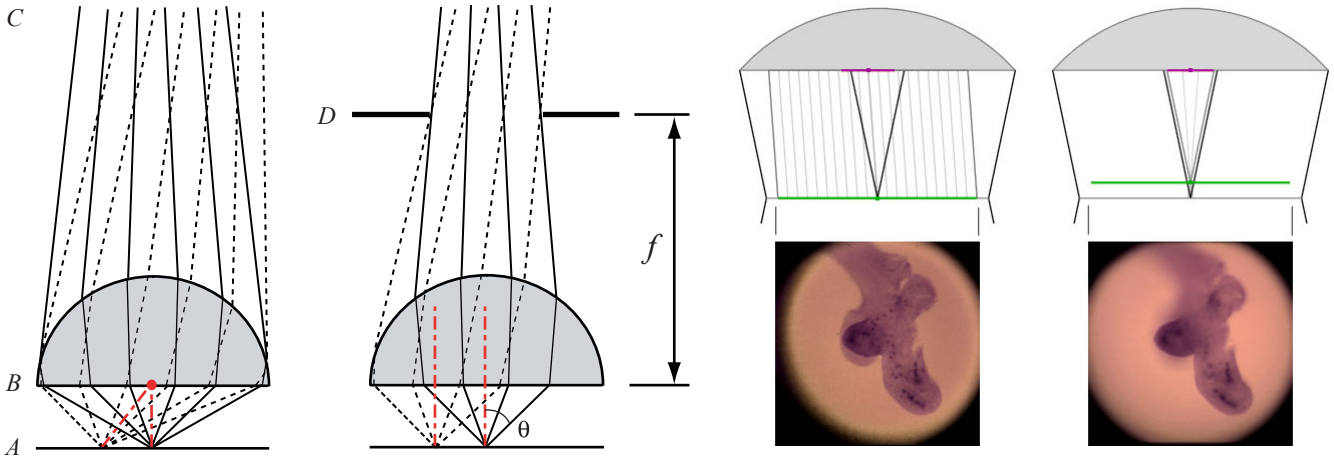
(4) the diameter of the lens is not given by its aperture, but by the sum of its aperture (base of a cone in figure 3(b)) and field of view ($2 \times$ distance between the two cones), and

(5) the defocus blur (double cone *ABC* in figure 2(a)), which becomes the point spread function (PSF) used in deconvolution, is shift-invariant with respect to x and y.

The first and fourth properties means that we must incorporate this telecentricity into our light field rendering software, if we are to generate correct views. Fortunately, this is straightforward; the *st* plane is set to the field of view, and the *uv* plane is set to the base of the cone in figure 3(b) but with W (the homogeneous coordinate) equal to zero. This places the *uv* plane at infinity, which causes the input data to be treated as orthographic. The second and third properties means that microscopes lack parallax-based depth cues. As a result, microscopists are excited when they see our perspective views and ability to shift the virtual viewpoint. The fifth property is essential to the functioning of 3D deconvolution microscopy, as will be discussed in section 4.

Illumination must be spatially and angularly uniform. When recording light fields of semi-transparent specimens under transmitted light, any non-uniformity in the spatial or angular distribution of illumination will appear in the light field, either as a falloff of intensity towards the periphery of the captured image or of each microlens subimage, respectively. Given a spatially non-uniform light source such as an incandescent bulb or arc lamp, it is possible by adjusting the height of the microscope substage condenser to force the illumination to be spatially uniform at the specimen (so-called Köhler illumination) or angularly uniform, but not both at once. In our prototype we used diffusers to approximate the desired condition, but this created some unwanted stray light; a fiber optic scrambler would have been better. When recording light fields of most fluorescent specimens, the angular distribution of fluorescence is independent of the angular distribution of the stimulating radiation, which therefore needs only to be spatially uniform. This condition is satisfied by most research-grade fluorescent light sources.

[1]Textbooks describe this aperture stop [Pluta 1988; Inoué 1997], but none that we can find mention "telecentric" or "orthographic", nor link the stop to the other properties listed above. We ascribe these omissions to differences in culture and focus.



(a) unrestricted optics (b) telecentric optics (c) oblique view (d) refocused view

Figure 3: The geometry of microscope light fields. (a) In an unrestricted optical system, cones of rays from points *A* on the object are refracted by the front surface of the lens *B*, reconverging to points on the intermediate image plane (off-page) at *C*. In an ideal lens, the centroids of these cones (red lines) converge at the first principal point (red dot); this is the effective center of perspective of the lens. (b) In a microscope objective, a stop *D* is placed at the rear focal plane of the lens (one focal length behind the second principal point). This so-called *telecentric stop* blocks certain peripheral rays, causing the cones to become symmetric and of constant width everywhere within the field of view. The centroids of these altered cones are vertical and parallel, creating an orthographic view of the object. (c) An oblique orthographic view of an embryo mouse lung, computed by extracting an off-center pixel from each microlens image. The accompanying diagram, displayed by our interactive viewer, is drawn to scale except for the objective lens (shaded gray shape at top). The objective was a 16×/0.4*NA* (dry), which allows views up to 15 degrees away from vertical; the current view (parallel gray rays) is about half of this. Alternatively, perspective views may be generated. (d) A synthetically focused view, computed by summing rays that converge at points on a plane (horizontal green line) 100 microns above the original focal plane (horizontal gray line).

3. Design trade-offs and optical performance

Unlike light field cameras, where the design can be worked out from ray diagrams [Ng et al. 2005], the behavior of a light field microscope is governed primarily by the wave nature of light. One of the principal results of Fourier optics is that light propagating from an object through a lens is bandlimited to NA/λ [Goodman 1996], where λ is the wavelength of light. *NA*, or numerical aperture, is the frequency response of a lens and is equal to $n \sin(\theta)$, where θ is the angle made by rays near the edge of the lens (marginal rays) with respect to the optical axis (see figure 3(b)), and n is the index of refraction of the medium adjacent to the lens.[2] Thus, the information in the image (or at any plane above the lens aperture) is a bandlimited and Nyquist sampled version of the original object with a sample spacing of $\lambda / (2NA)$.

The limited information in an image means that single samples (on these planes) have no directional information associated with them. Each emits a spherical wave, which has intensity and phase but no directional information. Directional information comes from aggregating a number of samples together − the larger the number of samples aggregated, the larger the number of distinguishable ray directions. This trade-off between spatial and directional resolution is one for one, so their product is fixed. In normal microscopy, we use spatial information only, so the number of resolvable spots in the image is the same as the number of samples as just defined. If we place a microlens array at the image plane, we can preserve some of the directional information, but only by making each microlens capture a number of samples, lowering the spatial resolution of the image.

Three main parameters govern the design of a light field microscope: the size of the microlens array and the curvature and size of an individual microlens. With the foregoing discussion in mind, the first two parameters should be chosen so as to make optimal use of the sensor resolution and the information in the light field. The third parameter may be freely chosen, and it allows us trade off lateral and axial resolution in the microscope.

Microlens array size. Since we are placing our microlens array at the intermediate image plane, its size should be chosen to cover the image created there by the objective. This equals the linear field of view on the specimen multiplied by the lateral magnification *M* of the objective (ratio of the two red lines in figure 2(a)). For typical objectives, *M* is in the range 10× to 100×, and the intermediate image is a circle 50-60mm in diameter. Of this only the innermost 19-23mm is well-corrected for aberrations. The ocular clips the intermediate image to this smaller diameter, which is then called the field number of the microscope. Since the spatial resolution of our images is limited by the (relatively large) distance between microlens centers, small aberrations in the objective don't affect us as severely as they do in a full-resolution system. We therefore felt justified in pushing the field number slightly; the array in our prototype measures 36×24mm, not coincidentally the size of a full-frame 35mm camera sensor.

[2] Typical values of n are 1.00 for air, 1.33 for water, and 1.52 for microscope immersion oil, and the numerical apertures of microscope objectives range from 0.2 to 1.4. For readers familiar with photography, numerical aperture can be converted to F-number (f/stop) N using the approximate formula $N = 1 / (2\ NA)$.

Microlens curvature. The resolving power of a multi-lens optical system is governed by the smallest *NA* among its lenses. To ensure that the microlenses do not limit the resolution of our light field microscope, and also to ensure that their images of the objective exactly fill the sensor plane without overlapping or leaving space between them [Ng et al. 2005], the numerical aperture of these microlenses should match the image-side numerical aperture of the objective, which is its (object-side) *NA* divided by its magnification M. Expressing this constraint as an F-number (focal length divided by diameter) gives

$$N = \frac{M}{2NA} \tag{1}$$

For example, for a 40×/0.95*NA* objective, the largest *NA* possible in air, we should employ f/20 microlenses. The microlenses in our prototype are 125μ on a side, so their focal length is 125μ × f/20 = 2.5mm, which is the distance from F to G in figure 2(b).

Microscopes typically allow the user to select among several objectives by rotating a mechanical turret. In future microscopes one can imagine a second turret filled with microlens arrays. Fortunately, microscope designers typically match the magnification of an objective with its *NA*, which has the effect of holding image-side numerical aperture constant. Thus, it is possible to find objectives with different magnifications that all use the same F-number microlenses. To illustrate this point, figure 4 shows how exit pupil size and microlens F-number vary as functions of objective magnification and numerical aperture.[3] Notice that the dashed horizontal line denoting our f/20 array intersects many popular objectives, including 20×/0.5, 40×/0.95, and 60×/1.4.

Another observation we can make from this figure is that, to accommodate low-magnification high-*NA* objectives, we must use stronger microlenses. Such combinations allow us to make larger images or to capture more angular parallax, assuming we use appropriate microlens arrays and sensors. However, building such systems will be challenging since they require larger exit pupils, hence physically larger objective lenses. Eventually, we exceed the diameter of the threaded barrel by which the objectives are screwed into the turret, denoted by the upper dashed horizontal line in the figure. To employ combinations above this line, new microscopes would need to be designed.

Microlens size and lateral resolution. In integral photography with a lenslet array and field lens, we control spatial and angular resolution by selecting the number of lenslets $N_s \times N_t$ and the number of pixels $N_u \times N_v$ behind each lenslet, respectively. In microscopy, total resolution $N_u \times N_v \times N_s \times N_t$ is limited by the number of resolvable sample spots in the specimen, as noted earlier. A commonly used measure of this resolution is the Sparrow limit [Inoué 1997], which is defined as the smallest spacing between two points on the specimen such that intensity along a line connecting their centers in the image barely shows a measurable dip. The Sparrow limit is nearly the same as the Nyquist criteria on sample spacing discussed earlier. Expressed as a distance on the intermediate image plane, this limit is

[3]The exit pupil gives the diameter of the image you see if you look through the back of an objective. Its size is 2 L (NA/M), where L is the distance to the image plane, sometimes called the tube length. In the figure we assume L = 160mm.

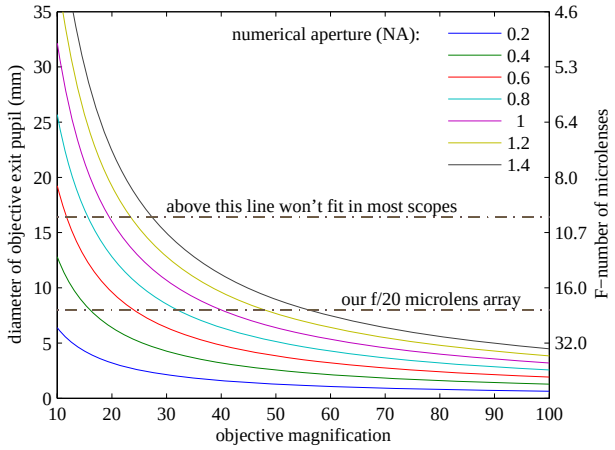


Figure 4: Microlens curvature for different objectives. For an objective having a given magnification and numerical aperture, the required F-number of the microlenses is given by the colored curves.

$$R_{obj} = \frac{0.47\ \lambda}{NA} M \tag{2}$$

where λ is the wavelength of light. Since the microlens array does not change the number of resolvable spots, our only remaining choice is how much angular resolution is desired, or how much spatial resolution can be sacrificed. This relation is

$$N_u \times N_v = \frac{W \times H}{R_{obj}} \tag{3}$$

where $W \times H$ are the microlens dimensions and $N_u \times N_v$ are the number of resolvable spots behind each microlens. For example, under green light (535nm), a 40×/0.95 objective has $R_{obj} = 10.59\mu$. Using our 36×24mm sensor, we obtain an upper limit of 3400 × 2266 resolvable spots.[4] Having chosen microlenses 125μ on a side, we obtain images of 288×192 pixels, our pixel size in object space is 125μ / 40× = 3.1μ, and we should expect $N_u = N_v$ = 125μ / 10.59μ = 11.8 resolvable spots per microlens. 3.1μ is our lateral resolution on the specimen, and 11.8 spots is our angular resolution. The effect of this angular resolution on axial resolution is considered next.

Axial resolution. After lateral resolution, microscopists are most concerned about axial resolution - the ability to distinguish features at different depths by refocusing the microscope. The commonly accepted measure of axial resolution is depth of field. In photography, formulas for depth of field can be derived using geometrical optics alone. In a microscope wave optics adds an additional term to these formulas to account for the fact that when light is perfectly focused, it still has spatial extent. The sum of these two terms gives the total depth of field [Inoué 1995]:

[4] The Canon 5D in our prototype has 8.2μ pixels and 4368 × 2912 pixels, which is adequate, but falls slightly short if the Bayer color mosaic is considered.

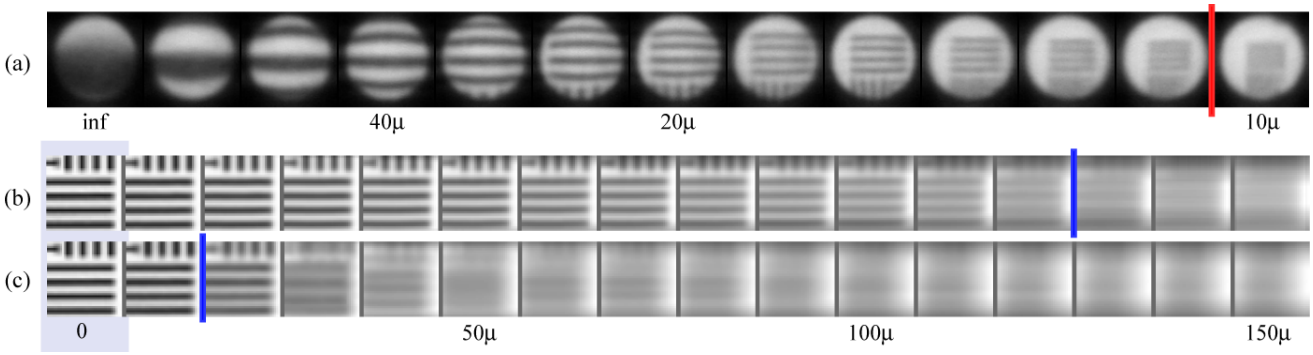


Figure 5: Optical performance of our light field microscope. The subject is a set of 11.6μ bars spaced 23.2μ apart from a Richardson Test Slide, imaged using a 16×/0.4 dry objective. (a) To measure light field resolution, we focused on the bars, then lowered the stage in increments, capturing a light field at each increment. This caused the bars to move closer together. Shown here is a single microlens image from each light field, labeled with inter-bar spacings in image space. The bars became indistinguishable when they were $10.5\mu \pm 1\mu$ apart (red line), in good agreement with R_{obj} in column 3 of table 1. (b) To demonstrate focusing range, we focused on the bars, then lowered the stage in 10μ increments, capturing a new light field at each increment. We then attempted to synthetically bring the bars back into focus. We succeeded until the bars had moved about 120μ. This represents the far side of our extended depth of field. The blue line represents the theoretical limit, based on D_{tot_3} from the table. (c) To demonstrate optical sectioning, we focused on the bars, captured one light field, and refocused it synthetically in 10μ increments. As expected, it went out of focus quickly. The blue line represents the theoretical limit D_{tot_2}. The relative distances between the base case (mauve rectangle) and the blue bars in (b) and (c) represent the extension in depth of field provided by our microscope.

$$D_{tot} = D_{wave} + D_{geom} = \frac{\lambda n}{NA^2} + \frac{n}{M\ NA}\, e \qquad (4)$$

where e is the spacing between samples in the image. In our microscope there are three cases to consider:

(1) *Microscope alone.* In an ordinary microscope having a camera at the intermediate image plane (i.e. no projection lens), e is the spacing between sensor pixels. If we assume these are small enough to not limit resolution, we obtain an expression in which the wave term dominates:

$$D_{tot_1} \approx \frac{\lambda n}{NA^2}\,. \qquad (5.1)$$

(2) *With microlenses.* In this situation the effective pixels are $N_u \times N_v$ times larger than the diffraction-limited spot size, since we are imaging this many spots behind each microlens. Let us assume $N_u = N_v$. This leads to an expression in which the geometry term now dominates:

$$D_{tot_2} \approx \frac{(2 + N_u)\ \lambda n}{2NA^2}\,. \qquad (5.2)$$

(3) *Single microlens pixel.* If we look at a single pixel behind a microlens, the wave term remains the same, since the NA of the optical system is unchanged, but the geometric term is increased by restricting the set of rays that form each pixel to $1/N_u^2$ of the available angular information. As is also proven in [Ng 2005], this increases depth of field by a factor of N_u, leading to

$$D_{tot_3} \approx \frac{(2 + N_u^2)\ \lambda n}{2NA^2}\,. \qquad (5.3)$$

To illustrate these formulae, the depths of field for three objectives are given in table 1. In the context of our application, D_{tot_3} is the depth of field observed when manipulating viewpoint (top row of images in figure 1), D_{tot_2} is the depth of field when focusing synthetically (bottom row of images in figure 1), and N_u is the number of slices with non-overlapping depths of field contained in any focal stack computed from a light field. [5]

Although the third column in this table corresponds to the arrangement used to capture figure 1, the pixel spacing in our Canon 5D fell short of the required resolution, as noted previously. This limited N_u in that figure to about 7 resolvable spots and the observed depths of field to about half of the values shown in the table. To confirm the validity of our model, we performed a separate experiment using a 5:1 macro lens. This experiment is summarized in figure 5. Although we were able to confirm R_{obj} (and N_u) to within measurement error, the synthetically focused images in figure 5(b) are softer than we would like, probably due to uncorrected spherical aberrations in the objective.

	50×/1.3NA oil	40×/0.95NA dry	16×/0.4NA dry
R_{obj}	9.67μ	10.59μ	10.06μ
N_u	12.93 spots	11.81 spots	12.43 spots
D_{tot_1}	0.48μ	0.59μ	3.34μ
D_{tot_2}	3.59μ	4.09μ	24.12μ
D_{tot_3}	40.67μ	41.91μ	261.57μ
figures	6		1, 3

Table 1: Depths of field computed using equations (5.1-5.3). We assume green light (λ = 535nm), f/20 microlenses, and n = 1.52 for the oil immersion objective. We also assume a spacing between camera pixels small enough to not be a limiting factor.

[5] To a microscopist, the number of useful slices is higher than N_u because focal stacks, especially if used in 3D deconvolution, are typically made of slices whose depth of field overlap by 50% or more to avoid aliasing [Schechner 1999].

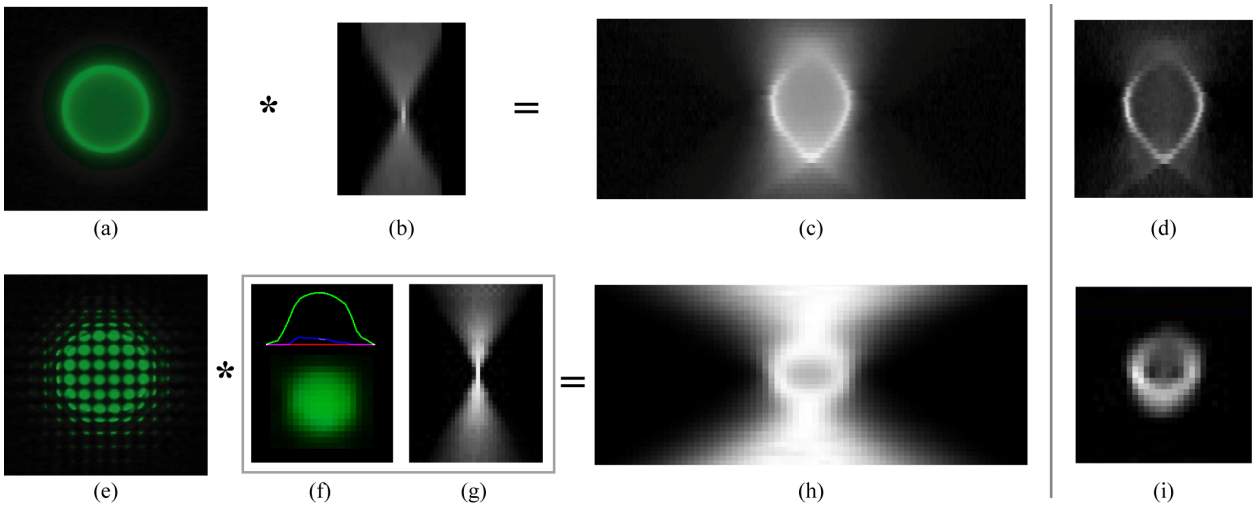


Figure 6: 3D reconstruction of a fluorescent bead using deconvolution microscopy. (a) Micrograph of a 15-micron polystyrene bead surface stained with BODIPY (Invitrogen FocalCheck microspheres #F-7235), mounted in immersion oil, and imaged by a 40×/1.3 oil Nikon Fluor objective (operating at 49× due to an intervening tube lens) with epi-illumination at 480nm and stimulated emission at 535nm. (b) Theoretical PSF for this objective, enhanced for display. (c) Vertical cross-section of a (grayscale) focal stack, obtained by moving the stage through 40 microns in 1-micron increments. The focal stack represents real (optical) convolution of the physical bead by the PSF. (d) Cross-section of deconvolved volume. Note the darker center and reduced blurring relative to (c). The turnip-shaped top and bottom arise from the limited set of rays captured by the objective. (e) Light field of the same bead, photographed with the same objective and our microlens array. Each microlens covers 3 microns on the specimen, so the bead covers only 5 microlenses. (f),(g) PSF of our synthetic focusing procedure. (g) was computed from an empirically measured "core" (f) using the method described in the text. (h) Cross-section of our synthetic focal stack. (i) Cross-section of deconvolved volume. Qualitatively the result compares favorably to (d), albeit at much lower spatial resolution. These reconstructions took several minutes each to compute on a PC.

4. 3D reconstruction

In pinhole imaging of a semi-transparent object, values of the object function (either its linear attenuation or emission) are integrated along rays passing through the pinhole. If the pinhole is replaced with an ideal lens, then integrals along rays are replaced with integrals over double cones centered at points in the object that are in focus [Swedlow 1997]. This double cone is the point-spread-function (PSF) of the lens, and it represents blurring due to defocus. If one includes the effects of diffraction, this PSF becomes hourglass shaped, having a finite waist at the plane of focus. Moving the lens along its optical axis, we form a sequence of images focused at different depths, i.e. a focal stack. This process is a convolution of the object by the PSF to form a 3D image.

Reconstruction by 3D deconvolution attempts to reassign that portion of the light recorded in each pixel that is due to blurring back to the voxels from which they came, i.e. it tries to estimate the object given the 3D image and PSF. Although the blurring convolution can be expressed as a matrix multiplication and directly inverted [Castleman 1979], the matrix is usually too ill-conditioned for this method to succeed. Instead, iterative algorithms are employed. To ensure convergence to a meaningful result, constraints are applied, for example that the object must be positive everywhere (see Appendix A). In this paper we employed an implementation of this algorithm in which the first guess for the object is formed from a Wiener filtering of the 3D image, and the PSF is provided empirically [Holmes 1995].

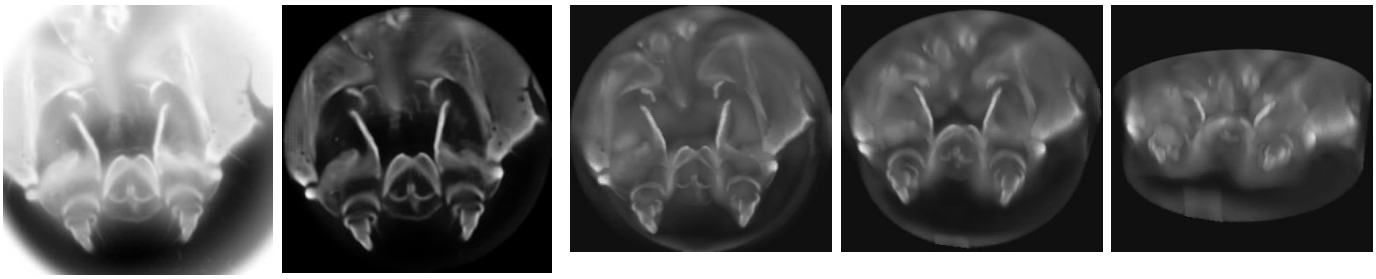
Empirical measurement of microscope PSFs is normally performed by recording a focal stack of a subresolution fluorescent bead. This focal stack shows the effects of both diffraction and defocus. In our case focusing is performed synthetically, so a different procedure must be used. We first record a light field of a 0.2μ bead using standard protocol. Since our microlenses are larger than the diffraction limit, the bead fills only one microlens subimage, as shown in figure 6(f). We then perform synthetic focusing on this light field, producing the PSF in figure 6(g). This PSF is free of noise and aberrations, since the focusing system is synthetic, but it correctly reflects the angular range of rays captured by the objective, including any falloff in intensity.

The rest of figure 6 compares 3D reconstruction of a 15-micron fluorescent bead − a standard test object − based on focal stacks produced by repositioning the microscope stage (top row) or by capturing and synthetically focusing a light field (bottom row). Figure 7 illustrates 3D reconstruction of a thicker specimen - the mouth of a silkworm. **[Watch the video!]**

5. Discussion

Since the original input to our focusing and deconvolution pipeline is a 4D light field rather than a 3D focal stack, one is led to ask: are we discarding information by generating a focal stack that could be used to make better reconstructions? The answer is no, but the reason is surprising.

If we approximate each resolvable spot on the intermediate image plane as a point and the samples in our light field as rays, then in the absence of scattering these samples represent line integrals of transmittance or emission through the specimen. Under these assumptions, it should be possible to reconstruct volumes



(a) focused image (b) cross section (c)-(e) volume renderings

Figure 7: 3D reconstruction of a silkworm mouth photographed under transmitted light. The objective was a 40×/1.3 Nikon Fluor oil objective (operating at 1.25*NA* due to the condenser). A photograph was captured through our microlens array, synthetically focused to create a 80-micron stack with 1-micron spacing, and deconvolved as described in the text. (a) Slice from focal stack. (b) Slice from volume after deconvolution. (c)-(e) Maximum-projection volume renderings from different viewpoints. All images are shown in inverted grayscale to improve clarity in the renderings.

from light fields using tomography [Kak 1988]. The standard input dataset for tomographic reconstruction of a volume is a 3D set of line integrals, organized into 2D subsets lying on parallel planar slices. In this scenario there are no lines passing through multiple slices, as would be present in a 4D light field. However, by gathering these out-of-plane lines into slices in other directions, one can extend the standard algorithm to accommodate them [Colsher 1980]. In the extended algorithm the extra rays act to reduce noise. The same is true in microscope deconvolution, but with a twist. By restricting illumination to a 1D slit, one can produce a focal stack whose PSF is a planar double-triangle rather than a volumetric double-cone. Ignoring any diffraction that might arise due to this slit illumination, 3D deconvolution still works and produces a valid volume. Thus, fully illuminating the microscope aperture is not strictly necessary, and the extra dimension of information in a light field acts mainly to reduce noise.

Looking deeper into the relationship between deconvolution and tomography, it is proven in Appendix A that, for the (unrealistic) case of an infinite aperture (i.e. no missing rays), synthetic focusing followed by 3D deconvolution using inverse filtering [Castleman 1977] is equivalent to tomographic reconstruction using the 3D Fourier projection-slice theorem. It is also argued that, for the more realistic case of a finite aperture, synthetic focusing followed by constrained iterative (non-blind) deconvolution [Swedlow 1997] is equivalent to tomographic reconstruction from limited-angle projections using the Simultaneous Algebraic Reconstruction Technique (SART) [Anderson 1984]. Although others have observed the similarity in artifacts produced by deconvolution and tomography [McNally 1994], we have not found these two approaches compared formally.

These relationships lead naturally to a second question: aside from reducing noise, do 4D light fields contain any information beyond that needed for 3D reconstruction? Under the assumption of weak scattering, the answer is again no, with the notable exception that some fluorescent materials emit polarized light, and the direction of polarization is a useful signal [Inoué 1997]. For more strongly scattering materials the answer is yes - the 4D light field encodes the directional dependence of scattering, i.e. the scattering phase function. However, if the specimen is too thick, then reconstruction by deconvolution is no longer strictly valid. In this case, 3D reconstruction may be still possible using other algorithms, such as diffuse optical tomography [Arridge 2001] or, in the case of complete opacity, computer vision algorithms.

As a final point of discussion, we have focused on the geometric trade-offs of our microscope; however microscopists are equally concerned about light-gathering power. This is especially true in fluorescence studies, where the level of emission may be low, or intense or repeated illumination may bleach the photopigment or kill the organism. With this criterion in mind, let us compare the performance of light field microscopy with conventional microscopy. Assuming the microscopist can tolerate an N_u^2 loss in lateral resolution and wishes to obtain a focal stack containing N_u slices, this can be done two ways:

(1) turn down the light by a factor of N_u^2, employ a camera whose pixels are N_u^2 larger, and shoot N_u images, moving the stage between each image, or

(2) turn down the light by a factor of N_u^2, employ the original camera and a microlens array having N_u^2 pixels behind each microlens, shoot one image, and use synthetic focusing to produce the N_u slices.

The pixels in both methods each gather the same amount of light and will therefore exhibit equivalent SNR (at least due to photon shot noise). However, the second method can be generated in $N_u\times$ less clock time and with $N_u\times$ less light. The secret, of course, is that the N_u slices produced using the light field microscope re-use the captured pixels N_u times, re-organizing them to focus at different depths. This re-use of data is one of the key arguments for light field microscopy.

6. Conclusions and future work

We have described a new kind of microscope, which uses an array of microlenses to capture the 3D structure of microscopic objects in a single snapshot (and therefore at a single instant in time). The sacrifice we make to obtain this capability is a reduction in lateral resolution. Specifically, for a reduction by a factor of $N_u \times N_v$, we can compute roughly $N_u \times N_v$ unique oblique views of the specimen, we can generate a focal stack containing N_u slices with non-overlapping depths of field, and using deconvolution algorithms, we can convert this focal stack to a volumetric model containing N_u slices.

One way to think about our microscope is that it separates image capture from the selection of viewpoint and focus. Thus, a specimen could be photographed today and (virtually) examined tomorrow, or photographed in one hospital and examined in another. Using our approach automated microscopes wouldn't need to focus perfectly before photographing each sample in a multi-well plate, leading to higher throughput. Here are a few variants on the light field microscope that we've thought of:

- *One-shot focusing.* To photograph light-sensitive specimens, the microscopist could insert a microlens array, capture a light field, and use it to find the most interesting features and best plane of focus. A motorized stage would be driven to the appropriate position, the microlens array removed, and a second image captured. This image would have high lateral resolution, and we have reduced light damage to the specimen.
- *Real-time follow-focus.* To capture video of moving specimens, one could employ an array with a large number of small microlenses, e.g. 800×800 microlenses with 3×3 spots behind each microlens. While not sufficient to perform 3D reconstruction, such an arrangement would have sufficient axial resolution to determine whether the specimen is moving up or down. This information could be used to drive a motorized z-stage. Similar systems based on 1D microlens array are employed in photographic cameras [Goldberg 1992].
- *Light field dissecting microscopes.* When examining large specimens (tens or hundreds of microns across), appreciation of perspective is as important as lateral resolution. This is why low-powered ("dissecting") microscopes provides stereo viewing. A light field microscope offers even more appreciation of 3D structure, if we can devise suitable ways to present the information, such as an autostereoscopic display.

Regarding other future work, we would like a better understanding of the angular distribution of light in our microlens subimages. Since current microscopes are not designed to image angular information, one does not expect them to do it well. One factor already mentioned is polarization. In polarized light microscopy, one can observe either the specimen directly (orthoscopy) or the angular distribution of its polarization properties (conoscopy [Pluta 1988]), but not both at once. We also believe that light fields can be used to correct for lens aberrations. Aside from improving the imagery from our existing system, this may pave the way toward microscopes with larger field numbers, or suggest a new methodology for designing objectives, in which certain aberrations are optimized while others are ignored.

Thinking beyond the selection of viewpoint and focus, by placing spatially graduated filters at the objective aperture plane (or condenser plane), one can modulate the exposure, spectral selectivity, polarization or phase delay of each microlens pixel, leading to a multi-variable mosaic [Schechner 2001; Nayar 2002] or to specialized devices, such as a single-snapshot imaging spectrophotometer. Finally, we conjecture that microlens arrays might be useful in the illumination path of microscopes, not only in the imaging path. If coupled with a video projector or other spatial light modulator, this would permit generation of arbitrary incident light fields, which could be used to implement confocal microscopy [Levoy 2004] or to provide structured illumination for the improvement resolution [Gustafsson 2005].

7. Acknowledgments

We thank Gordon Kino, Olav Solgaard, and Joseph Goodman for advice on optics, Julie Theriot, Stephen Smith, and Rudolf Oldenbourg for advice on microscopy and applications, and David Biggs for tips on 3D deconvolution algorithms. We also thank AutoQuant Imaging for use of their deconvolution software.

8. References

ADELSON, T., and WANG, J.Y.A. 1992. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 2, 99-106.

AGARD, D.A. 1984. Optical Sectioning Microscopy: Cellular Architecture in Three Dimensions, *Ann. Rev. Biophys. Bioeng* 13, 191-219.

ANDERSEN, A.H., KAK, A.C.. 1984. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm, *Ultrasonic Imaging* 6, 81-94.

ARRIDGE, S.R. 2001. Methods for the Inverse Problem in Optical Tomography, *Proc. Waves and Imaging Through Complex Media*, Kluwer, 307-329.

CASTLEMAN, K.R. 1979. Digital image processing, Prentice Hall. (or 2nd edition, 1996.)

CHAMGOULOV, R.O., LANE, P.M., MACAULAY, C.E. 2004. Optical computed-tomography microscope using digital spatial light modulation, *Proc. SPIE* 5324, 182-190.

COLSHER, J.G. 1980. Fully three-dimensional positron emission tomography, *Phys. Med. Biol.* 25, 1, 103-115.

CORLE, T.R.. KINO, G.S. 1996. *Confocal Scanning Optical Microscopy and Related Imaging Systems*, Academic Press.

ELLIS, G.W. 1966. Holomicrography: transformation of image during reconstruction a posteriori, *Science* 143, 1195-1196.

GOLDBERG, N. 1992. Camera technology: the dark side of the lens, Academic Press.

GOODMAN, J. 1996. *Introduction to Fourier optics*, 2nd edition, McGraw-Hill.

GUSTAFSSON, M.G.L. 2005. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution, *Proc. National Academy of Sciences* 102, 37.

HOLMES, T.J., BHATTACHARYYA, S., COOPER, J.A., HANZEL,D.V., KRISHNAMURTHI, LIN, W., ROYSAM, B., SZAROWSKI, D.H., TURNER, J.N. 1995. Light microscopic images reconstructed by maximum likelihood deconvolution, In *Handbook of Biological Confocal Microscopy*, ed. J.B. Pawley, Plenum Press, 389-402.

INOUE, S., OLDENBOURG, R. 1995. Microscopes, In *Handbook of Optics*, 2nd edition, McGraw-Hill.

INOUE, S. and SPRING, K.R. 1997. *Video Microscopy*, 2nd edition, Plenum Press.

GABOR, D. 1948. A New Microscopic Principle, *Nature* 161, 777-778.

ISAKSEN, A., MCMILLAN, L., GORTLER, S.J. 2000. Dynamically Reparameterized Light Fields, *Proc. SIGGRAPH 2000*.

JAVIDI, B., and OKANO, F., eds. 2002. *Three-Dimensional Television, Video and Display Technologies*, Springer-Verlag.

KAWATA, S., NAKAMURA, 0., MINAMI, S. 1987. Optical microscope tomography. I. Support constraint, *J. Opt. Soc. Am. A* 4, 1, 292-297.

KINGSLAKE, R. 1983. *Optical system design*, Academic Press.

LEVOY, M., HANRAHAN, P. 1996. Light Field Rendering, *Proc. SIGGRAPH 1996*.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., BOLAS, M. 2004. Synthetic aperture confocal imaging, *Proc. SIGGRAPH 2004*.

MCNALLY, J.G., PREZA, C., CONCHELLO, J.A., THOMAS, L.J. JR. 1994. Artifacts in computational optical-sectioning microscopy, *J Opt Soc Am A* 11, 3, 1056-67.

NAYAR, S.K., NARASIMHAN, S.G. 2002. Assorted Pixels : Multi-Sampled Imaging With Structural Models, *Proc. ECCV 2002*.

NG, R., LEVOY, M., BREDIF, M., DUVAL, G., HOROWITZ, M., HANRAHAN, P. 2005. Light Field Photography with a Hand-Held Plenoptic Camera, Stanford Tech Report CTSR 2005-02.

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

OKOSHI, T. 1976. *Three-Dimensional Imaging Techniques*, Academic Press.

PLUTA, M. 1988. *Advanced Light Microscopy* (in 3 volumes), Elsevier.

SCHECHNER, Y., KIRYATI, N. 1999. The optimal axial interval for estimating depth from defocus, *Proc. ICCV 1999*.

SCHECHNER, Y., NAYAR, S. 2001. Generalized Mosaicing, *Proc. ICCV 2001*.

SWEDLOW, J.R. , SEDAT, J.W., AGARD, D.A. 1997. Deconvolution in optical microscopy. In *Deconvolution of Images and Spectra* , ed. P.A. Jansson, Academic Press, 284-309.

Appendix A: 3D deconvolution and tomography

In this appendix we show the equivalence between 3D deconvolution methods used in microscopy and limited-angle tomography methods used in medical imaging.

Let V be the volume we are imaging, and I_ω be an orthographic view of this volume centered on the origin along direction ω. A light field of V consists of the set of I_ω where ω lies on the unit hemisphere Ω. Let I_ω be P_ω, the integral projection of V along ω, multiplied by A_ω, which can describe radiometric falloff over the aperture or restriction of viewing angle (if $A_\omega = 0$).

To show the equivalence between tomographic reconstruction and deconvolution, we need to define two operators: $E_\omega(I)$, which extrudes (backprojects) the plane I in direction ω to form a volume, and $R_\omega(F(I))$, the equivalent operation in Fourier space, which places the Fourier transform of I in 3-space centered on the origin so that its normal points in direction ω. This equivalence is a trivial application of the Fourier projection-slice theorem.

Deconvolution by inverse filtering [Castleman 1977] takes a focal stack FS and known point spread function PSF and solves the following equation in the Fourier domain:

$$V * PSF = FS, \quad F(V) = \frac{F(FS)}{F(PSF)}.$$

Assuming V is transmissive or emissive (no scattering), then a focal stack can be constructed from orthographic views by backprojecting them along their centers of projection and adding them:

$$FS = \int_\Omega E_\omega(I_\omega)\, d\omega = \int_\Omega E_\omega(A_\omega P_\omega)\, d\omega$$

As described in section E, the PSF is typically measured by capturing a focal stack of a subresolution bead, which we can treat as an impulse at the origin ($P_\omega = \delta$), leading to

$$PSF = \int_\Omega E_\omega(A_\omega \delta)\, d\omega$$

Assuming the PSF is spatially invariant, deconvolution by inverse filtering from a light field then consists of evaluating

$$V = F^{-1}\left(\frac{F\left(\int_\Omega E_\omega(A_\omega P_\omega)\, d\omega\right)}{F\left(\int_\Omega E_\omega(A_\omega \delta)\, d\omega\right)}\right)$$

Since the Fourier transform is linear and F(δ) = 1, this is equivalent to

$$V = F^{-1}\left(\frac{\int_\Omega R_\omega(A_\omega F(P_\omega))d\omega}{\int_\Omega R_\omega(A_\omega 1)\, d\omega}\right) \qquad \text{(A.1)}$$

In the full-angle case ($A_\omega = 1$), this deconvolution is the same as fully three-dimensional tomography using the projection-slice theorem [Colsher 1980]. The normalization factor, written here as the sum of all constant valued planes passing through the origin, evaluates to division by $1/r$, where r is radius from the origin.

Now consider the case when A_ω contains zeros. In deconvolution this can be interpreted as the PSF being bandlimited; its Fourier transform has zeros. In tomography this corresponds to missing data at certain angles. Applying equation (A.1) directly would give large noise amplification, so microscopy and medical imaging have developed spatial domain algorithms to replace it. These algorithms use constrained gradient descent, and are called constrained iterative deconvolution or Simultaneous Algebraic Reconstruction Technique (SART), respectively. These two algorithms are again equivalent, as we now show.

Constrained iterative deconvolution can be expressed as iteratively improving an estimated volume V by comparing the known focal stack to the estimated focal stack ($V * PSF$), correcting the estimate volume by the difference, then constraining the volume to be non-negative [Swedlow 1997]:

1. $\Delta^{(k+1)} \leftarrow FS^{(k)} - V^{(k)} * PSF$
2. $V^{(k+1)} \leftarrow \max\,(V^{(k)} + \Delta^{(k+1)}, 0)$

SART can be expressed as comparing known projections of the volume in each direction (I_ω) to computed projections of the estimate volume in the same directions with the same radiometry ($V * E_\omega(A_\omega \delta)$), correcting the estimate volume by the difference, then constraining the volume to be non-negative [Kak 1988]:

1. $\forall \omega, \Delta_\omega^{(k+1)} \leftarrow E_\omega(I_\omega) - V^{(k)} * E_\omega(A_\omega \delta)$
2. $V^{(k+1)} \leftarrow \max\,(V^{(k)} + \int_\Omega \Delta_\omega^{(k+1)}\, d\omega, 0)$

SART can be converted to iterative deconvolution by moving the integral inside each expression in step 1. It is only commuted with other linear operations, so this remains the same algorithm:

1. $\Delta^{(k+1)} \leftarrow \int_\Omega E_\omega(I_\omega)\, d\omega - V^{(k)} * \int_\Omega E_\omega(A_\omega \delta)\, d\omega$
2. $V^{(k+1)} \leftarrow \max\,(V^{(k)} + \Delta^{(k+1)}, 0)$

Thus the update rule is identical to constrained iterative deconvolution, since we showed previously that $\int_\Omega E_\omega(I_\omega)\, d\omega$ is the focal stack and $\int_\Omega E_\omega(A_\omega \delta)\, d\omega$ is the PSF.