24

Blind Deconvolution

Timothy J. Holmes, David Biggs, and Asad Abu-Tarif

INTRODUCTION

Purpose of Deconvolution

Deconvolution[1] is a process most often applied to three-dimensional (3D) data sets from widefield fluorescence (WFF) microscopy. Additionally, it can be applied to 3D data for transmitted light bright-field (TLB) and other widefield modalities. It improves resolving power and signal-to-noise ratio in data from confocal laser-scanning microscopes (CLSM) and confocal spinning-disk microscopes, and it provides a transformation to data from differential interference contrast (DIC) images to effectively extract a 3D image of refractive index distribution.

With WF microscopes, 3D data are collected by optical sectioning (Castleman, 1979; Agard, 1984, 1989). As illustrated in Figures 24.1 and 24.2, an automatic focus controller and digital camera record a sequence of images of a thick specimen. The focus controller drives either a stepper motor focus or a piezo crystal embedded within a customized objective lens mount. When a microscope having a narrow depth of field (DOF) is focused on a plane, the image contains both sharp, in-focus features originating from the plane of focus and blurred features originating from planes above and below. After storing an image, the microscope is focused to an adjacent plane, which is about one DOF away, and another image is recorded. This process is repeated until the entire specimen is scanned. The result is a 3D data set that contains the out-of-focus blur. Deconvolution eliminates the blur to restore clarity of the image. An example is shown in Figure 24.3. After deconvolution, the image may be rendered for 3D viewing and quantitation by methods discussed in Chapters 14 and 15.

When using confocal microscopes, including both the CLSM and the spinning-disk microscopes (such as the Yokogawa CSU22, PerkinElmer Life Sciences, Boston, MA; the CARV−2 from Atto Biosciences, Rockville, MD; or the DSU from Olympus), images of adjacent planes are optically sectioned in a similar way, except that each frame is collected using either a single, scanned light spot and a photodetector or a structured array of scanned pinholes that serve as both light spots and point detectors. Each confocal image is, by itself, already partially deconvolved because the optics reject most of the out-of-focus light. Most of the features recorded in a

3D confocal data set are in focus. However, the CLSM has its own limitations. While it rejects most of the out-of-focus light, it does not reject all of it. Some haze remains. In addition, the 3D data set has axial smearing. Point objects look like American footballs oriented along z. As fewer photons are detected compared to WFF, the raw data contains much quantum or Poisson noise. This noise not only obscures real structures in the image, as explained in Chapter 4, it can actually create impossible features. As Poisson noise affects the intensity recorded in each pixel separately, random positive excursions can give the appearance of small bright features that are only one pixel big. Such a feature is impossible because, assuming that the data was recorded using Nyquist sampling, no real feature should be smaller than 4 to 5 pixels across. Because deconvolution suppresses data corresponding to features smaller than the point-spread function (PSF), it eliminates such noise features from the data and thereby meets the Nyquist reconstruction criterion. Simply put, properly sampled, noisy 3D data sets are almost always improved by deconvolution. Deconvolution reduces or eliminates all four of these problems and thereby improves the confocal image. An example is shown in Figure 24.4.

To a first approximation, DIC contrast is produced by a set of polarizers and Wollaston prisms that form an image that is the differential in the direction of shear of the refractive index of the specimen at the focus plane. The image contains information about the 3D refractive index distribution, although this information is convolved and non-linearly distorted by the objective lens, polarizers, and differential interference optics. Deconvolution can be used to invert this process and to produce an image whose intensity is proportional to the 3D refractive index distribution. An example is shown in Figure 24.5.

Advantages and Limitations

Early deconvolution methods used a calculated or measured PSF (Hiraoka *et al.*, 1990; see also Chapter 25, *this volume*). The PSF was measured by collecting a 3D image of a subresolution fluorescent microsphere. The accuracy of this process is limited by the fact that the biological sample itself is part of the optics of the system. It has its own heterogeneous refractive index, and this changes the PSF. In principle, the only way to measure the correct PSF is to measure it within the biological sample itself, by placing microspheres in the sample at the depth being imaged. The depth is important because the blurring caused by spherical aberration is the most prominent factor that changes in the PSF, and spherical aberration changes with depth (Gibson, 1991; Visser, 1992; see also Chapter 20, *this volume*). Although people have succeeded at

---

[1] Deconvolution is a mathematical process that is the opposite to convolution. Convolution is described in the Appendix at the end of this chapter. Convolution is useful because it can mimic the blurring effects of diffraction and geometrical optics. Deconvolution is important because, to the extent that the blurring effect can be described mathematically, it can reverse this blurring effect.

Timothy J. Holmes • AutoQuant Imaging, Inc., Troy, New York 12180 and Rensselaer Polytechnic Institute, Troy, New York 12180
David Biggs and Asad Abu-Tarif • AutoQuant Imaging, Inc., Troy, New York 12180
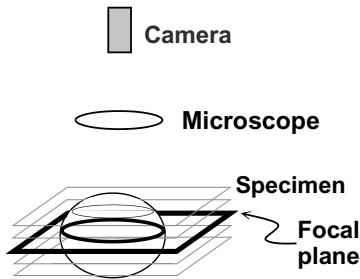
FIGURE 24.1. Optical sectioning schematic. This setup is used with the wide-field modalities, including WFF, TLB, DIC, and polarized light.

FIGURE 24.2. Hardware components needed for optical sectioning.

inserting microspheres into some samples, it is impractical to do so routinely.

The approach highlighted in this chapter does not use a calculated or measured PSF and is therefore called *blind deconvolution* (Ayers and Dainty, 1988; Bhattacharyya *et al*., 1996; Holmes, 1992; Krishnamurthi *et al*., 1995). The actual PSF is degraded by any heterogeneity in the refractive index of the specimen (Gibson and Lanni, 1991). These complications are compounded in CLSM fluorescence, where light levels are very low and the effect of

refractive index variations are even more severe (Visser *et al*., 1992; Chapter 20, *this volume*).

Blind deconvolution uses an iterative process to produce both a reconstructed estimate of the PSF as well as the deconvolved data. Realistic mathematical constraints on the PSF are used. Figures 24.6 and 24.7 demonstrate a blind deconvolution of a data set with severe spherical aberration, including the PSFs at the beginning and end of the reconstruction. Figures 24.6(A) and 24.7(A) show *xz* views that have a non-symmetrical flare along *z*


A

B
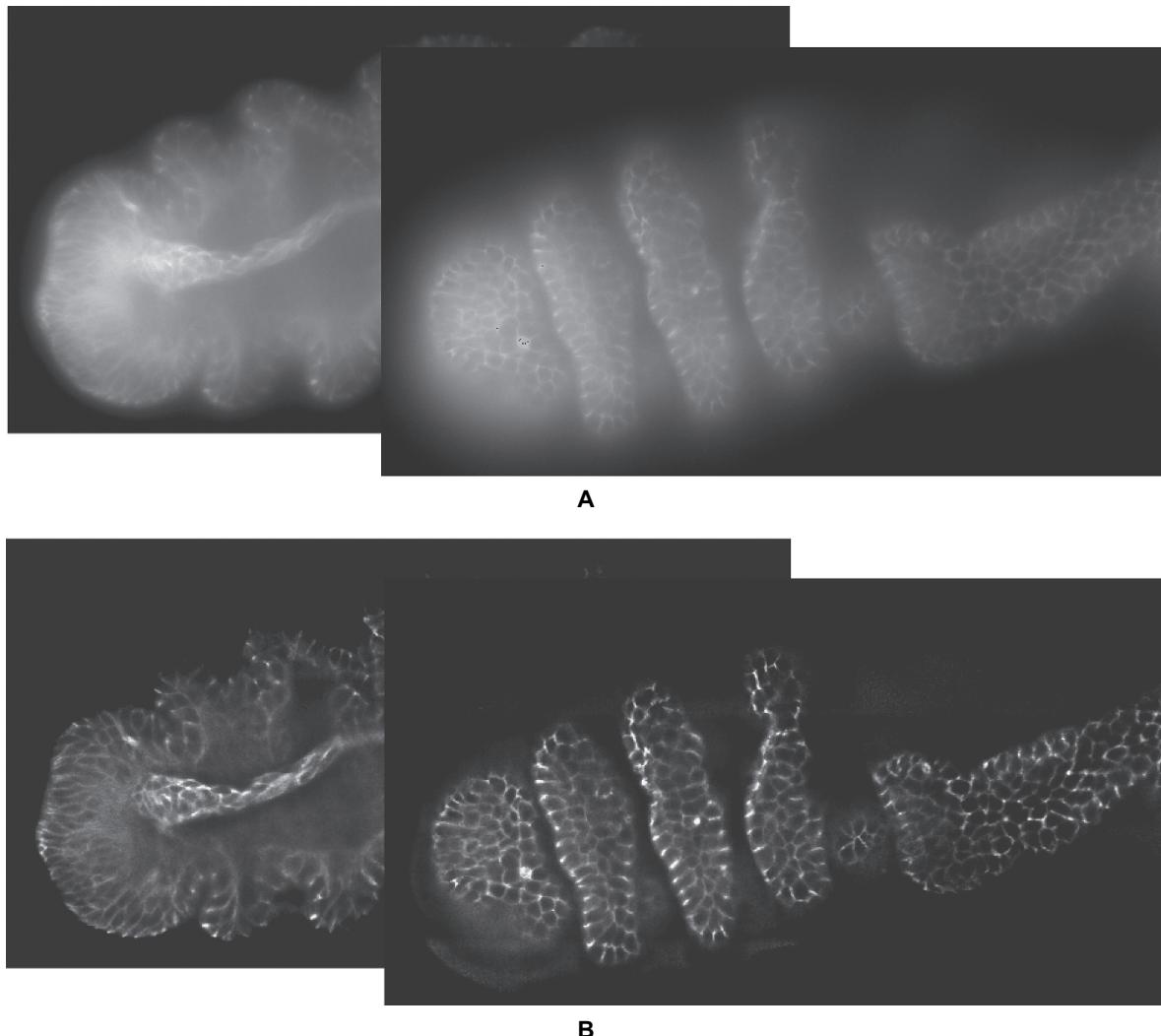
FIGURE 24.3. (A) Two slices from an unprocessed WFF stack. (B) Deconvolved slices from the blind deconvolution. (Image courtesy of M. Schechter, Leica Microsystems.)
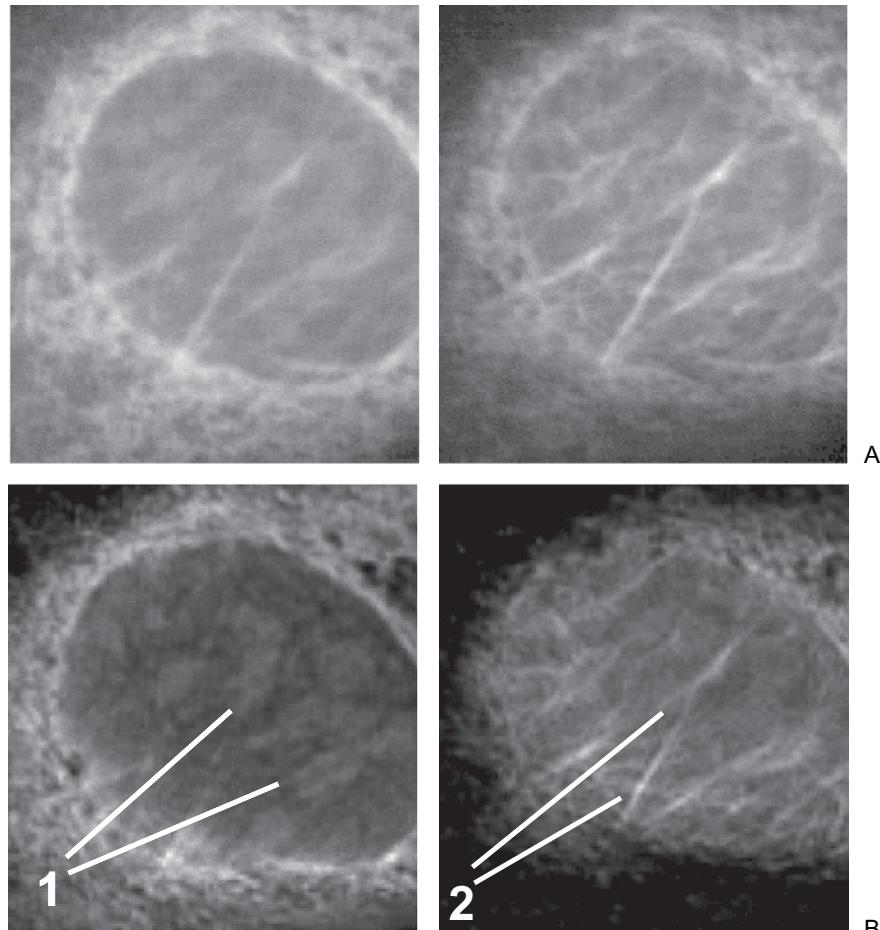
FIGURE 24.4. (A) Two sections from a confocal stack. (B) After deconvolution. Resolving power along z, signal-to-noise ratio and contrast are improved. Chromosomes are now apparent (1) and vimentin fibers are seen above and below the nucleus (2), as they should be, but are not seen within the nucleus [left frame of (B)]. (Image courtesy of M. Hoppe, Leica Microsystems.)
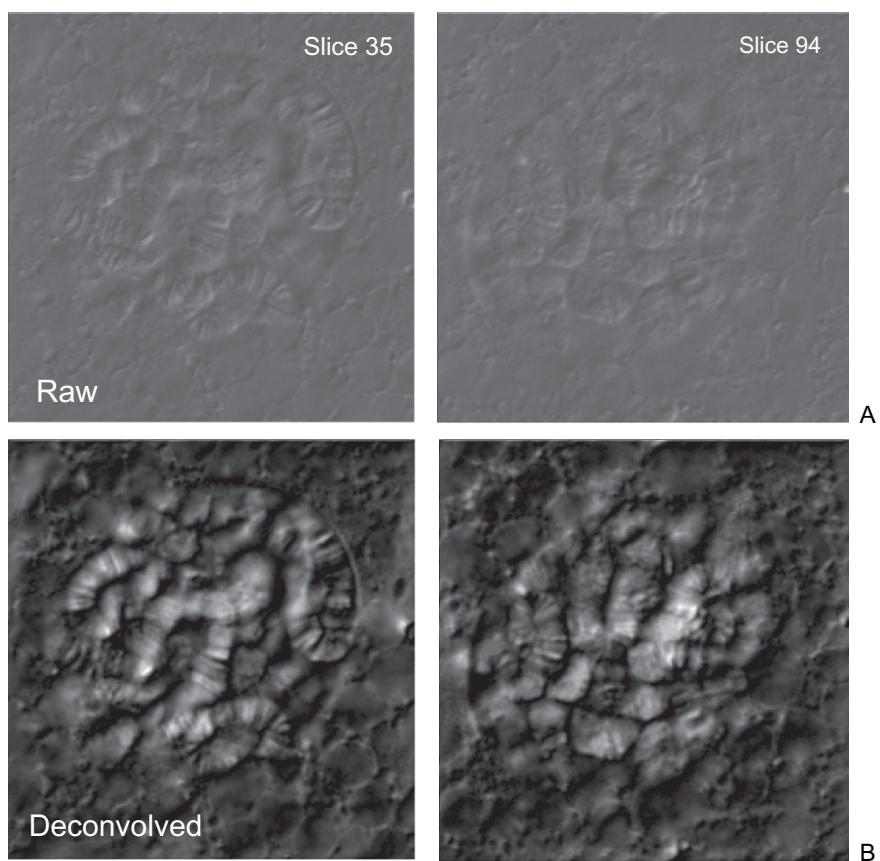
FIGURE 24.5. (A) Two optical sections from a DIC stack. (B) After deconvolution. The deconvolved images appear like CLSM fluorescent samples. No stain is used, but the brightness is proportional to the refractive index of the sample, resulting in a virtual refractive index map.
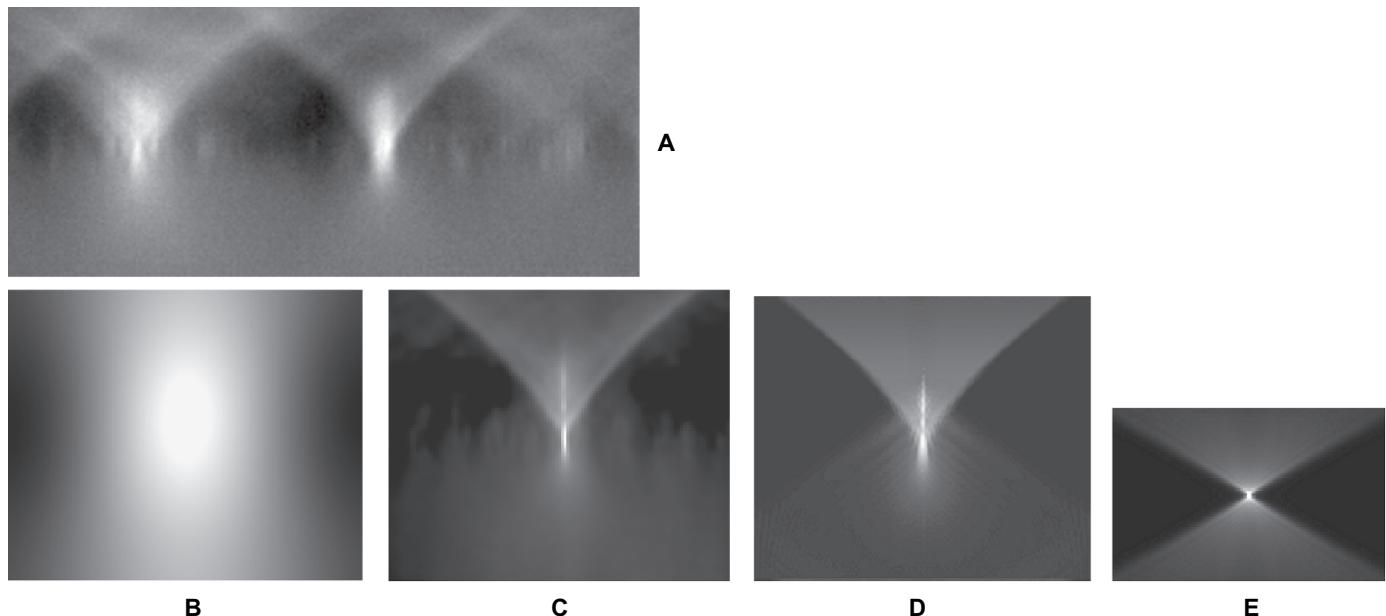
FIGURE 24.6. (A) *xz* slice through a 3D WF fluorescent data set collected with a water-immersion lens, showing spherical aberration. (Image courtesy of D. Kube, Case Western Reserve University.) (B) *xz* sum projection of the first guess of the PSF used to seed the first iteration of the blind deconvolution. (C) *xz* sum projection of the PSF produced by the blind deconvolution procedure. (D) Theoretically calculated PSF with a spherical aberration coefficient chosen to match the flare seen in Figure 24.6(A). (E) Theoretically calculated PSF without spherical aberration, for comparison.

that is characteristic of spherical aberration. This condition is especially evident with water-immersion lenses, when the collar on the objective has not been properly adjusted for the coverslip thickness. The PSF reconstruction is shown in Figure 24.6(B,C) and is discussed later in the section entitled "More Examples — Blind Deconvolution and Spherical Aberration."

A *non-blind* version of this algorithm (Holmes and Liu, 1991) is available for cases where the PSF can be accurately measured (Hiraoka *et al.*, 1990) or calculated (Gibson and Lanni, 1991).

The approach highlighted in this chapter uses a mathematical tool called maximum likelihood estimation (MLE). As this approach includes quantum photons as an underlying assumption, it results in a reduction in the noise that is caused by Poisson noise. Because of diffraction, any microscope represents a bandlimited system. As explained later, this means that any "good" signal lies inside the bandlimit. Much of the undesirable noise energy (i.e., noise signal) lies outside the bandlimit. The algorithm recognizes this and subsequently rejects the out-of-band noise automatically.
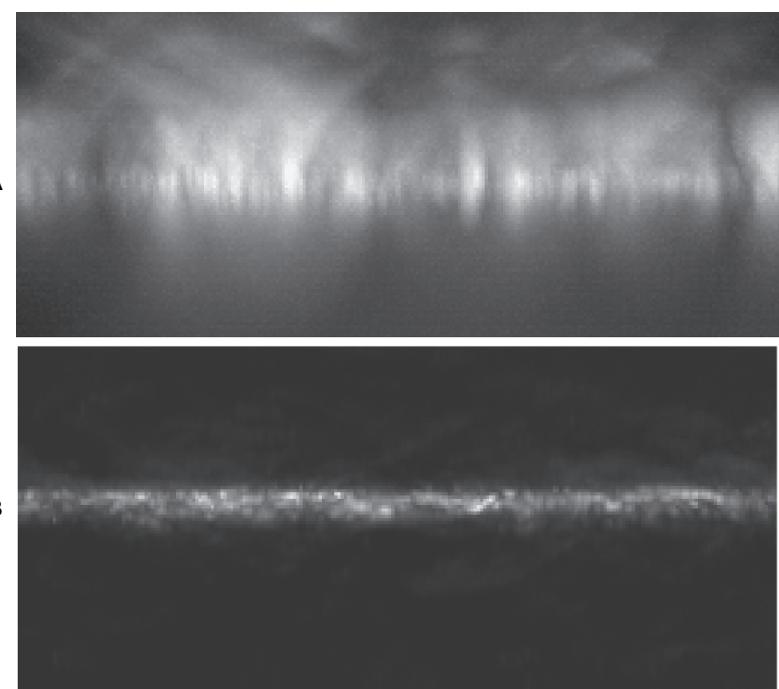
FIGURE 24.7. (A) *xz* maximum projection of the data set shown in Figure 24.6(A). Spherical aberration is apparent by the asymmetry in the amount of flare above and below the sample. (B) Blind deconvolution result, showing the elimination of spherical aberration. The PSF shown in Figure 24.6(B) was used as the PSF first guess, and the PSF shown in Figure 24.6(C) was produced by the blind deconvolution procedure.

It also constrains the deconvolved image to have only non-negative values, a process that cannot be implemented with a simple bandlimiting filter. Best results are achieved when the pixel or optical section spacing is finer than required by Nyquist sampling (Chapters 2 and 4, *this volume*).

PRINCIPLES

Data Collection Model

The data collection is modeled (Holmes, 1989, 1992; Holmes and Liu, 1989, 1991; Holmes *et al.*, 1991) according to the following equations:

$$\mu(x, y, z) = f(x, y, z) *h(x, y, z), \tag{1}$$

$$\mu_b(x, y, z) = \mu(x, y, z) + b(x, y, z), \tag{2}$$

and

$$\mu_n(x, y, z) = \mu_b(x, y, z) + N(x, y, z), \tag{3}$$

where x and y are the in-plane coordinates, z is the axial coordinate, $\mu()$ is the data collected under ideal conditions where there is no background level or noise, $f()$ is the fluorescent dye concentration as a function of position, * represents the convolution operation, $h()$ is the PSF, $\mu_b()$ is the ideal measured intensity with background and without noise, $b()$ is a background level mainly due to non-specific staining and the dark current in the image sensor (Janesick *et al.*, 1987; Snyder et al., 1993), $\mu_n()$ is the collected image data that contains noise, and $N()$ is noise, primarily due to quantum variations in the number of photons recorded.

Deconvolution produces an estimate, or restoration, of $f()$. The background level $b()$ is calibrated as discussed later and is addressed by precorrecting the raw data (Chapters 3 and 13, *this volume*).

MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood estimation is a mathematical optimization strategy (Van Trees, 1968) used for producing best estimates of quantities from data corrupted by random noise. Because of the randomness of quantum photon emissions, the collected image represents data having a certain likelihood of having been produced among a host of other data that might have been recorded. MLE produces the estimate of $f()$ that has the highest likelihood of having given rise to the specific data collected.

In mathematical terms, we are formulating a *log-likelihood function* that represents a measure of the likelihood that one would collect the specific noisy data that were actually collected. This function is actually a functional (i.e., a function of a function) of $f()$, $h()$, and $\mu_n()$. Its formulation is based on the random nature of the quantum photon noise. This functional is solved iteratively to produce its maximum value. In blind deconvolution, specific quantities of $f()$ and $h()$ are solved using an iterative search (Dempster *et al.*, 1977). The specific quantities of $f()$ and $h()$ that produce this maximum become the reconstructed image and the reconstructed PSF. With non-blind deconvolution, $h()$ is known from a prior measurement and the deconvolution algorithm solves only for $f()$.

Algorithms

WFF and CLSM Fluorescence

The flowchart for the blind deconvolution is shown in Figure 24.8. In Step 1, first guesses of $f^{(0)}()$ and $h^{(0)}()$ are made. Step 2 is designed from the expectation-maximization (EM) algorithm (Dempster *et al.*, 1977; Holmes, 1989, 1992; Holmes and Liu, 1989, 1991; Holmes *et al.*, 1991). Step 3 enforces the constraints placed on the solution of the PSF. The details of Step 3 are outlined in Figure 24.9.

The system may use a quantitative criterion for stopping the iterations. Several quantitative stopping criteria have been studied for applications in nuclear medicine (Veklerov and Llacer, 1987; Llacer and Veklerov, 1989). We prefer methods that monitor the progress in the likelihood value and then stop the iterations when the improvement in this functional between the last two iterations drops below a selected threshold.

The design of the flowchart in Figure 24.9 is modeled after the Gerchberg–Saxton approach (Gerchberg and Saxton, 1974). Step 1 enforces a unit summation constraint. Step 2 constrains the energy of the PSF to lie within an hourglass region b_h that is illustrated in Figure 24.10. This constraint prevents portions of the background intensity $b()$ from being mistakenly assigned to the PSF. Step 3 forces the 3D optical transfer function [OTF, the Fourier transform (FT) of the PSF] to be bandlimited. The 3D OTF has a frequency band (BW$_{xy}$) wherein it may be non-zero, with a missing cone region illustrated in Figure 24.11. Step 4 ensures non-negativity. This step is needed because Step 3 would otherwise cause negative numbers to appear in the result, but light intensity is never negative (these negative pixel values are set to zero). In the two-dimensional (2D) diffraction-limited case, the OTF frequency band is described by the interior of a circle having a radius of BW$_{xy}$, following the same formula as in the caption of Figure 24.11.

The CLSM fluorescence algorithm is similar to the WFF algorithm. The hourglass region of Figure 24.10 is replaced by a cylinder of radius r, and the bandlimit region of Figure 24.11 is replaced by a cylinder with twice the bandwidth along x,y and z and with no missing cone region. The bandwidths are doubled because the confocal PSF, when theoretically determined according to Fourier optics (Conchello *et al.*, 1994), is approximated by a multiplication of two widefield PSFs. By the convolution theorem, when any two functions are multiplied their bandwidths are added. In real life, however, the truly confocal condition is only met if the detector pinhole is infinitely small, and the signal level is so high that Poisson noise can be ignored. Clearly this can never happen and a smaller diameter of BW$_{xy}$ is appropriate.
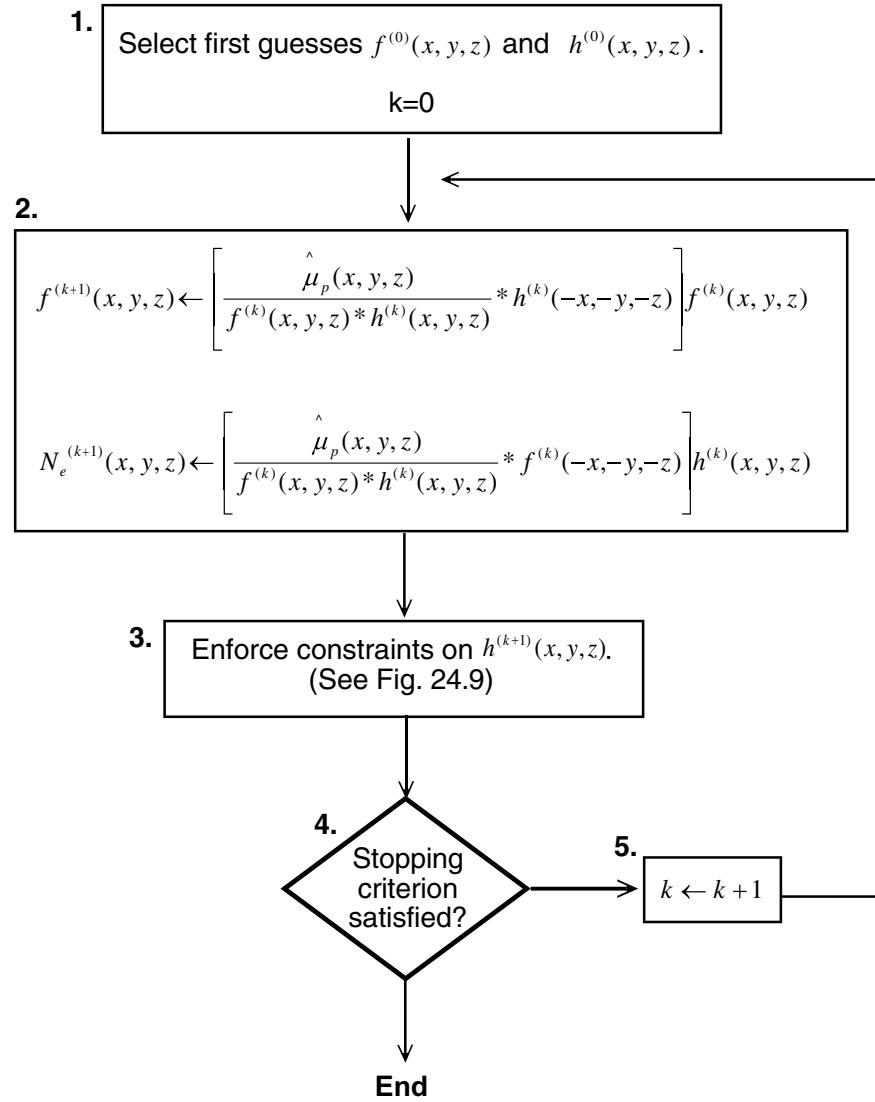
Transmitted Light, Bright-field (TLB)

The deconvolution of TLB images of absorbing stains are produced using a variation on the WFF algorithm (Holmes *et al.*, 1995; Holmes and O'Connor, 2000). The main requirement is that the condenser aperture must be larger than the objective lens aperture. This variation of the algorithms depends on the approximation that the raw TLB data $\mu()$ [(with noise $N()$ and background $b()$ ignored] follows a simple linear-systems model according to:

$$\mu(x, y, z) = B - h(x, y, z)*f(x, y, z), \tag{4}$$

where B is the bright background intensity, and $f()$ is the absorption coefficient. This approximation follows from Macias-Garza and colleagues (1989) when geometrical optics are considered, from Streibl (1984; Eq. 28) and when the phase contribution to the intensity in the image is ignored (i.e., the refractive index of the

FIGURE 24.8. Flowchart of the MLE blind deconvolution algorithm: $f^{(k)}(x, y, z)$ is the reconstructed dye concentration of the kth iteration, $h^{(k)}(x, y, z)$ is the reconstructed PSF of the kth iteration, $\hat{\mu}_p$ is the optically sectioned image data, after being corrected for non-uniform sensitivity and shutter-speed instability, $N_e^{(k)}(x, y, z)$ is the temporary data array that stores an unconstrained form of the PSF estimate. Figure 24.9 shows details of the constraints in Step 3.

1. Select first guesses $f^{(0)}(x, y, z)$ and $h^{(0)}(x, y, z)$.

   $k=0$

2. $$f^{(k+1)}(x, y, z) \leftarrow \left[\frac{\hat{\mu}_p(x, y, z)}{f^{(k)}(x, y, z) * h^{(k)}(x, y, z)} * h^{(k)}(-x, -y, -z) \right] f^{(k)}(x, y, z)$$

   $$N_e^{(k+1)}(x, y, z) \leftarrow \left[\frac{\hat{\mu}_p(x, y, z)}{f^{(k)}(x, y, z) * h^{(k)}(x, y, z)} * f^{(k)}(-x, -y, -z) \right] h^{(k)}(x, y, z)$$

3. Enforce constraints on $h^{(k+1)}(x, y, z)$. (See Fig. 24.9)

4. Stopping criterion satisfied?

5. $k \leftarrow k+1$

   End

specimen is the same as that of the mounting medium), from Erhardt and colleagues (1985) when light emanating from the specimen is approximated as being fully incoherent, and from Sheppard and Gu (1994) when the specimen is considered to be a weakly scattering object (i.e., structures not stained enough to make them opaque). We noticed that by preprocessing this noise-free raw data according to:

$$\mu_p(x, y, z) = B - \mu(x, y, z) \tag{5}$$

we have a model expressed according to:

$$\mu_p(x, y, z) = f(x, y, z) * h(x, y, z), \tag{6}$$

which is identical to Eq. 1.

The deconvolution is carried out as follows: First, the raw data are precorrected for background, flat-field non-uniformities and shutter-speed fluctuations (Cooper *et al.*, 1993; to arrive at $\hat{\mu}(x, y, z)$, which is an estimate of the $\mu(x, y, z)$ shown in Eq. 4. Then, $\hat{\mu}(x, y, z)$ is preprocessed according to Eq. 5 to produce an estimate of $\mu_p(x, y, z)$, which is denoted in Figure 24.8 as $\hat{\mu}_p(x, y, z)$. Finally, the flowchart of Figure 24.8 is executed on $\hat{\mu}_p(x, y, z)$.

Figure 24.12 shows stereo-pairs both before and after a TLB deconvolution of bright-field data.

Differential Interference Contrast (DIC)

The objective of DIC deconvolution is to produce a 3D image of the refractive index distribution. The DIC image formation model is nonlinear and uses Fourier optics principles (Holmes and Levy, 1987; Hansen and Conchello, 1988; Preza *et al.*, 1998). A rough conceptual schematic of the components in a DIC microscope is shown in Figure 24.13. Two orthogonally polarized beams are produced by Wollaston prisms and pass through the specimen as separate ray bundles slightly displaced from each other in the shear direction. The image represents a coherent-light interference pattern that is formed at the image sensor when the two beams are combined. Image contrast is produced by inhomogeneity of the sample's refractive index (RI) distribution near the plane of focus.

The observed image g is modeled by a linearized approximation according to the following equation:

$$g = s(K \cdot f * h_1 * h_2) + n \tag{7}$$

where f is the original refractive index distribution, h_1 is the theoretical DIC PSF that reduces to two delta functions (van Munster *et al.*, 1997), h_2 is the PSF that accounts for the resolution limit of the optics, n is Gaussian noise, and * is the convolution operator. K is a constant that converts the change in RI to a phase differ-

From Fig. 24.8 Step 2

1. Unit summation constraint

$$h^{(k+1)}(x, y, z) \leftarrow N_e^{(k+1)}(x, y, z) / N^{(k+1)}$$

2. Hourglass constraint

$$\text{For all}(x, y, z)\text{within}\,b_h^c: \; h^{(k+1)}(x, y, z) \leftarrow 0.$$

3. Bandlimit and missing cone constraint

$$H^{(k+1)} \leftarrow FFT\big[h^{(k+1)}(x, y, z)\big]$$

$$\text{For all }(u, v, w)\text{within}\,Z: H^{(k+1)}(u, v, w) \leftarrow 0 .$$

$$h^{(k+1)}(x, y, z) \leftarrow FFT^{-1}\big[H^{(k+1)}(u, v, w)\big]$$

4. Nonnegativity constraint

$$\text{For all}(x, y, z): \text{If } h^{(k+1)}(x, y, z) < 0, \text{ then } h^{(k+1)}(x, y, z) \leftarrow 0 .$$
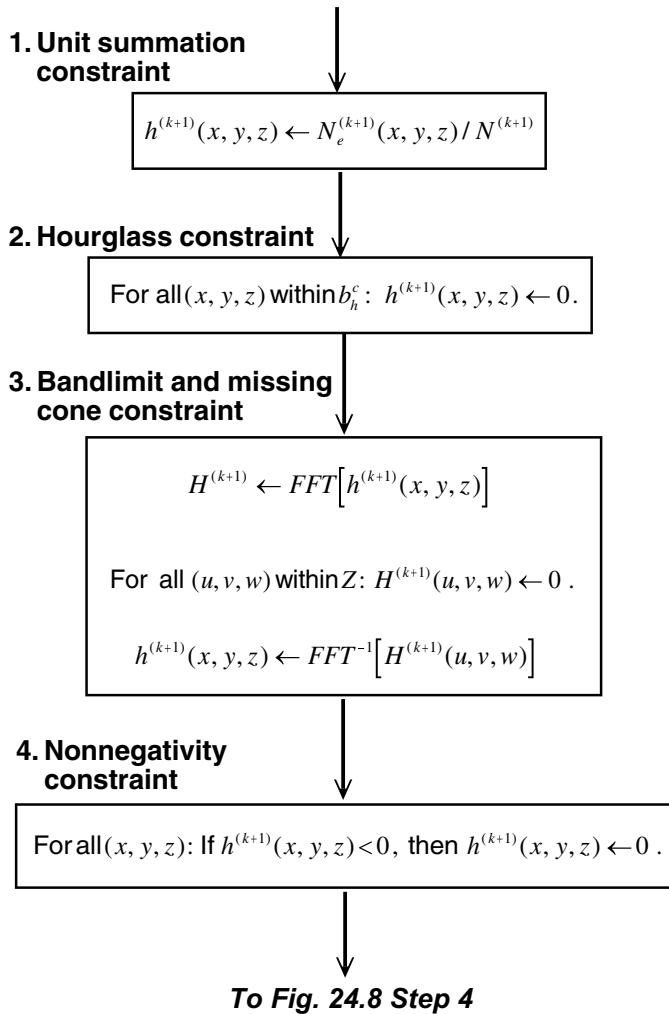
To Fig. 24.8 Step 4

FIGURE 24.9. Flowchart of constraints applied in WFF and TLB. These blocks make up Step 3 in Figure 24.8. The term $b_h{}^c$ is the complement of b_h (see Fig. 24.10) and is the region wherein $h()$ is near zero. $H^{(k+1)}(u, v, w)$ is the Fourier transform of $h^{(k+1)}()$. Z is the region of $H()$ which is zero. FFT and FFT^{-1} indicate the forward and inverse fast Fourier transforms, respectively. These constraints are carried out iteratively.

FIGURE 24.10. Hourglass region used in constraining the widefield PSF. The terms b_h and b_h^c denote the regions inside and outside the hourglass, respectively. The radius r is set to be a number of Airy disk widths according to $r = 0.61\,c\lambda/\text{NA}$, where λ and NA are the wavelength and numerical aperture, respectively. We use $c = 1$. The acceptance angle θ is found from $\theta = \sin^{-1}(\text{NA}/\eta)$, where η is the refractive index.

medium, the data are processed using a penalty on positive values. Finally, for proper reconstruction, the intensity data should be converted to phase data by reversing the effect of the nonlinear function $s()$.

The result of the reconstruction is a map of RI, or optical path length, from a DIC image, which is similar in nature to that produced by a transmitted light interference microscope (Inoué and Spring, 1997, p. 86; Dunn, 1988). The technique also requires only one image per focal plane, compared to the several images per focal plane that are required by the rotational diversity technique (Preza *et al.*, 1998). While reconstructed DIC data bear a superficial resemblance to darkfield (DF) images, it's important to remember that it is impossible to know if a bright feature in a DF image represents a structure having an RI that is higher or lower than its surroundings.

ence. The function $s()$ represents the transformation of the phase differences to intensity values that occurs when the two waves are combined at the image sensor.

The deconvolution process uses an iterative optimization algorithm that is complicated by several factors. First, DIC produces contrast only from changes in RI that occur in the shear direction. As a result the differential PSF has a significant null spectral component in the direction perpendicular to the shear, and this must be accounted for. Otherwise, streaking artifacts will appear parallel to the orientation in which the beams are separated (Kam, 1998). Second, one cannot assume nonnegativity in the RI change, relative to the background. Instead, negative values are only penalized. This approach provides a soft constraint where the negative values are, in a sense, minimized but not prevented altogether. This is a reasonable assumption because, typically, biological structures have a higher RI than the aqueous embedding medium. In cases where the specimen refractive index is less than the embedding

FIGURE 24.11. Illustration of the region wherein the widefield 3D OTF may be non-zero. The 3D coordinates u, v, and w are the frequencies with respect to x, y, and z, respectively. The radial bandlimit follows the formula: $\text{BW}_{xy} = 2\,\text{NA}/\lambda$. The axial bandlimit follows the formula $\text{BW}_z = (\text{NA})^2/(2\eta\lambda)$. θ follows the same formula as shown in the caption of Figure 24.10. The biconic region is the "missing cone." The region of the OTF where it is known to have all zero values, Z, is the union of this missing cone with the region outside of the cylinder.
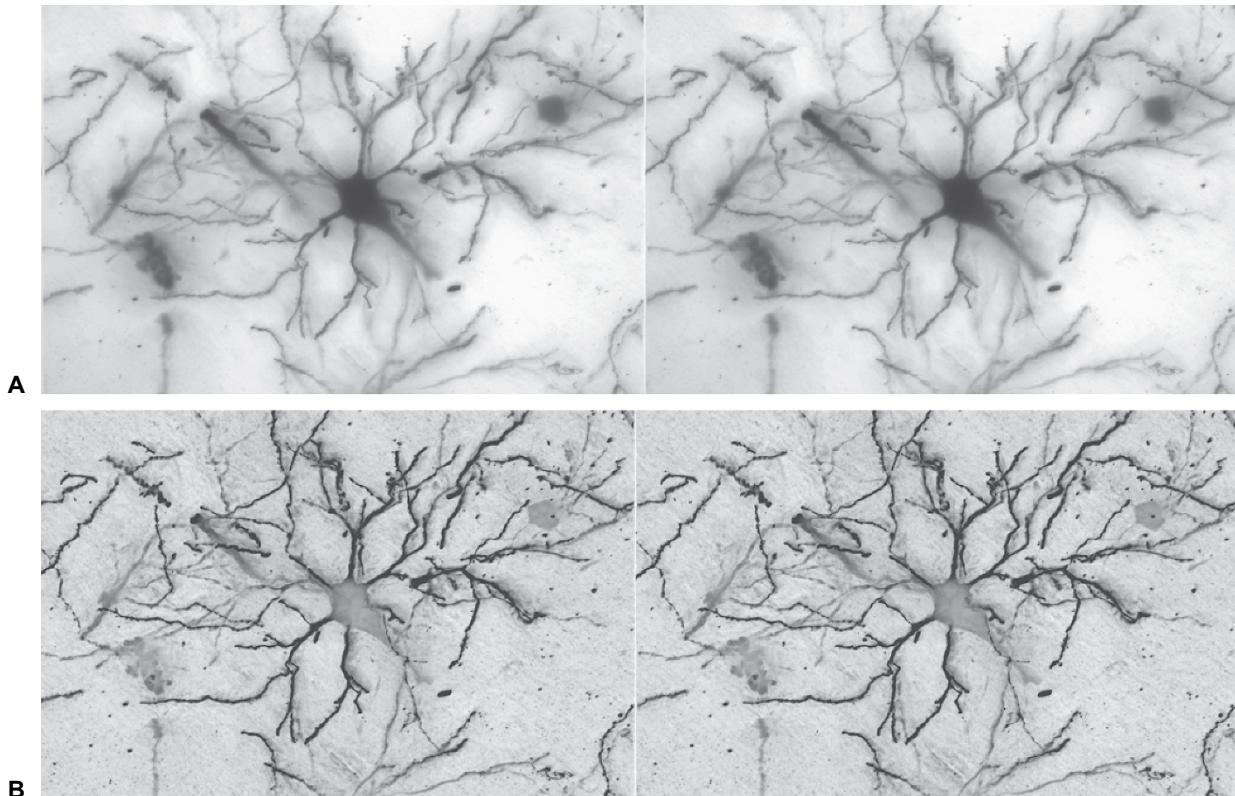
FIGURE 24.12. (A) Stereo-pair of a TLB stack. (B) After deconvolution.
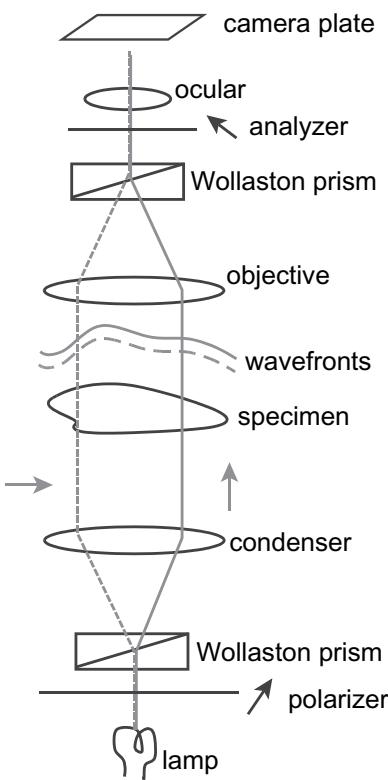

FIGURE 24.13. Rough schematic of the DIC microscope.

DIFFERENT APPROACHES

3D

The MLE algorithms, both *blind* and *non-blind*, are in a category of algorithms known as the *iterative constrained* algorithms (Wallace *et al.*, 2002). They are called *iterative* because they repeat the same steps to update $f^{(k)}()$, as shown in Figure 24.8. They are called *constrained* because the bandwidth is limited and the values of $f^{(k)}()$ are allowed only to be nonnegative. Among other approaches, described below, these iterative constrained approaches are considered to be the most quantitatively accurate; that is, the summation of the intensity in all the pixels of the image is preserved. The total number of photons in the final reconstruction is the same as was collected from all the charged-coupled device (CCD) images. As, in some sense, this is a normalization, one might expect that any deconvolution approach can yield a similar result, simply by following the deconvolution with a normalization operation. However, what is different here is that the normalization is inherent within the deconvolution operation itself. It is an intrinsic fallout of the MLE process and not something that must be done afterwards.

One reason these iterative constrained algorithms are more accurate is that they are capable of recovering Fourier transform information, such as the missing cone information, which was removed by the optics. Details about the missing Fourier components and how they are recovered may be found in Holmes and Liu (1989). The inverse filter and nearest neighbor approaches

Table 24.1. Processing Times and Memory Requirements for Implementation on a Windows 2000/NT/XP 3 GHz Computer

Image Size	10 Iterations, Blind Deconvolution	Inverse Filter
$256 \times 256 \times 10$	13 s	1.3 s
$1024 \times 1024 \times 50$	20 min	2 min

Requirements for other data-set sizes can be approximated. Processing time will be proportional to *M log M* (*M* = number of voxels) which is the proportionality of an FFT (Oppenheim and Schafer, 1975). Most of the processing time is taken by calculating the FFTs. Extra memory is not needed for large image sizes because the image may be decomposed into subregions that are deconvolved separately. For future computers with faster chip sets, processing speed is roughly proportional to the Pentium computer chip speed (e.g., 3 GHz).

described below do not recover this information. Other iterative constrained approaches include the *Jansson–van Cittert* algorithm (Holmes and Liu, 1992b) and *least squares* (Carrington, 1990) algorithms. Iterative-constrained algorithms are generally much slower to compute than the *inverse filter* or *nearest-neighbor* algorithms. However, as computer speed has increased, this issue has been mitigated significantly in recent years (Table 24.1). Today, typical iterative deconvolutions require only a few seconds to a few minutes on a personal computer (PC).

A variant on *Gold's* ratio method (Agard *et al.*, 1989) is another iterative-constrained algorithm. It is similar to the algorithm shown in Figure 24.8, except that the convolution of $h(-x, -y, -z)$ is removed. It has been developed by Sibarita and colleagues (2002) to provide high-throughput of hundreds of volumes in a short time. The advantage of this approach is that of speed, while some noise reduction is sacrificed.

The *nearest-neighbor* algorithm is the fastest algorithm. It is a 2D algorithm that works by sharpening each image in the stack one at a time. Mathematical details are in Holmes (1992) and Agard and colleagues (1989). The two adjacent images in the stack (above and below) are reblurred by convolving them with an approximation to the PSF. These two reblurred images are then multiplied by a coefficient (<1) and subtracted from the image being sharpened. This operation is known generically as unsharp masking (Russ, 1995). This algorithm is not quantitative. Image intensities are not preserved. A summation of the processed image is just a fraction of the summation of the original image. This algorithm is really an *image enhancement* method, although it is sometimes controversially called a deconvolution. A more generally

accepted term is *deblurring* algorithm (Wallace *et al.*, 2002). Its main advantage is that it runs in a fraction of a second on a PC. Figure 24.14 shows a nearest-neighbor processed image of the optical section shown in Figure 24.3.

The inverse filter is another fast approach. The Fourier transform of Eq. 1 is expressed as:

$$M_p(u, v, w) = F(u, v, w)H(u, v, w) \tag{8}$$

where $F()$ and $H()$ are the Fourier transforms of $f()$ and $h()$, respectively. Ideally, $f()$ ought to be recoverable from $M_p()$ by taking the inverse Fourier transform of the following operation:

$$\Lambda(u, v, w) = M_p(u, v, w)H_r(u, v, w), \tag{9}$$

where

$$H_r(u,v,w) = \frac{1}{H(u,v,w)}. \tag{10}$$

In reality, $H(u, v, w)$ contains zeros, so Eq. 10 is mathematically unstable and this approach does not work. Instead, a regularized variant of $H_r(u, v, w)$ is used, according to the expression:

$$H_r(u,v,w) = \frac{H^*(u,v,w)}{|H(u,v,w)|^2 + k}, \tag{11}$$

where the * superscript denotes complex conjugate and the term k is a very small number that is empirically chosen. Equation 11 is a regularized approximation to Eq. 10 that is based on the principles of the Wiener filter (Castleman, 1979). It approximately equates to Eq. 10, except where $H(u, v, w)$ is near zero.

The inverse filter is fast to execute (Table 24.1), yet preserves reasonable quantitative accuracy. Figure 24.15 shows the optical slice shown in Figure 24.3 after application of an inverse filter.

The inverse filter works well with phase objects imaged under a TLB microscope with a small condenser aperture. Iterative constrained methods will not work well to deconvolve such an image because it does not follow the mathematical model where $\mu_p()$ must be nonnegative and constructive interference causes areas in the specimen to be brighter than the background. An inverse filter works because it does not rely on $\mu_p()$ being nonnegative. Figure 24.16 shows summed intensity projections of an unprocessed stack and a processed stack.

2D Image Filtering

The nearest-neighbor algorithm has a variant, called the no-neighbors algorithm, which operates only on a single 2D plane of
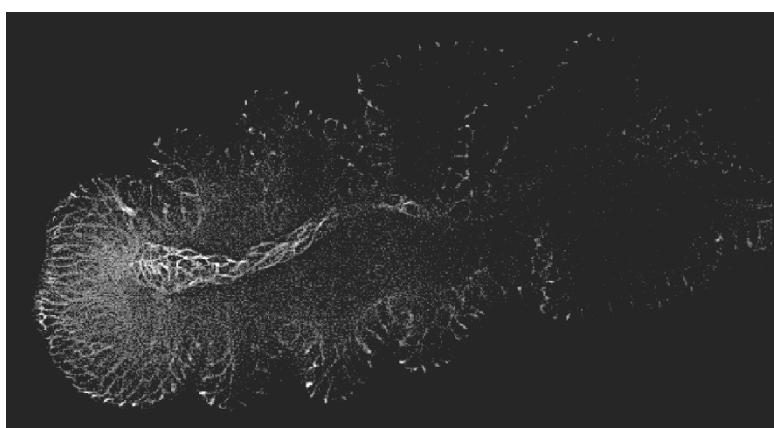
FIGURE 24.14. Optical section shown in Figure 24.3(A) after the nearest-neighbor algorithm.

the image being sharpened. Instead of reblurring adjacent images, the image being sharpened is reblurred, multiplied by the coefficient, and subtracted from itself. Because this no-neighbors variant operates on 2D data alone, it does not require 3D data, so it is a much simpler technique to use and requires inexpensive equipment. The complete system shown in Figure 24.2 is no longer needed to take advantage of 2D deblurring. Any microscope with a digital camera will benefit.

There is also a purely 2D variant of the MLE blind deconvolution algorithm (Biggs, 1998) which shares some of these 2D advantages. As with the no-neighbors algorithm, any microscope with a digital camera will benefit. This 2D MLE algorithm works well when the sample is inherently 2D, as with flat samples that fit within the depth of field, with objectives having a large depth of field or with total internal reflection (TIRF) microscopy. It works surprisingly well even in many cases where the sample is inherently 3D, as long as high-NA optics are used to reduce depth of field. However, in these inherently 3D cases, it is more accurate to think of this algorithm as a nonlinear sharpening filter than as a deconvolution because the algorithm assumes that the imaging process is 2D and this is an inaccurate assumption.

DATA CORRECTIONS

Images recorded using common intensified and standard CCD digital cameras can be deconvolved. Peltier-cooled sensors (−70° or lower) are preferred because they have the lowest noise level, the best linearity, the highest dynamic range (16 bits), and the lowest background levels. Image data collected from any cooled-CCD camera must be calibrated for the background level $b(x, y, z)$, non-uniform pixel sensitivity, $\beta(x, y)$, and the camera exposure-time instability $T(z)$ such that the raw collected data $\mu_r(x, y, z)$ are composed of:

$$\hat{\mu}_r(x, y, z) = \beta(x, y, z)T(z)\mu(x, y, z) + b(x, y, z). \quad (12)$$

The exposure time $T(z)$ varies with z as the shutter may fluctuate by several milliseconds between frames. When non-mechanical shutters are used, the exposure time does not vary, but the illumination varies because arc lamps wander and power variations affect incandescent lamps. The $T(z)$ term accommodates these variations. The correction method is outlined in Cooper and colleagues (1993).

The corrected data set $\hat{\mu}()$ is an estimate of $\hat{\mu}()$. For TLB, this estimate is further processed by Eq. 5, and for WFF it is substituted directly for $\mu_p()$ in Figure 24.8.

In CLSM fluorescent data, $\beta()$ and $T(z)$ are irrelevant, and only $b()$ must be determined. To estimate b, a background region of interest is chosen and the average background level is calculated from this data. Alternatively, a histogram of intensity levels is calculated on the image, and the background level is detected automatically by identifying the peak in the histogram that is close to zero.

LIGHT SOURCE AND OPTICS ALIGNMENT

Deconvolution is forgiving about microscope misalignment, but only to a limit. Misalignment causes degradation in the PSF. When the optics are misaligned, the PSF is less symmetric, and deviates from ideal theory in other ways. The non-blind algorithms and the inverse filter algorithm that use a PSF measurement, adapt to this assymetry, to a degree, because the assymetry occurs in both the measured and actual PSF. The blind deconvolution will adapt to misalignments too, much in the way that it adapts to spherical aberration as shown in Figure 24.6(C).

On the other hand, deconvolution follows the proverbial "garbage-in garbage-out" principle. Circular averaging is often performed to reduce the noisiness of the measured PSF and this causes the calibrated PSF to deviate from the true, assymetric PSF. The best images are obtained with proper Köhler alignment (see Chapter 36, *this volume*). For non-blind deconvolution, this limit is partly because the PSF measured with a fluorescent bead is not the same PSF that exists in the biological sample. As the sample itself is a heterogeneous refracting medium that is part of the optics, when the sample changes, the PSF changes. Although blind deconvolution adapts to PSF changes, the closer the true PSF is to the theoretical PSF, the better the results.

Another consideration in obtaining a well-behaving PSF, which works well for immersion lenses, is to adjust the refractive index of the medium that is between the objective lens and the coverslip. The technique described in Hiraoka and colleagues (1990) uses finely adjusted immersion media to minimize the spherical aberration (see Chapter 20, *this volume*). Oils having a different RI are placed onto and removed from the coverslip, one-by-one, with RI differences of 0.002. The refractive index that provides the most symmetrical PSF along the z axis is chosen. It is straightforward to see the symmetry by focusing up and down on a bead and noticing if the out-of-focus bead seen while focusing up looks the same as the out-of-focus bead seen while focusing down (see also Fig. 20.3, *this volume*). It is necessary to match the oil in this same way with the specimen. A fine structure needs to be identi-
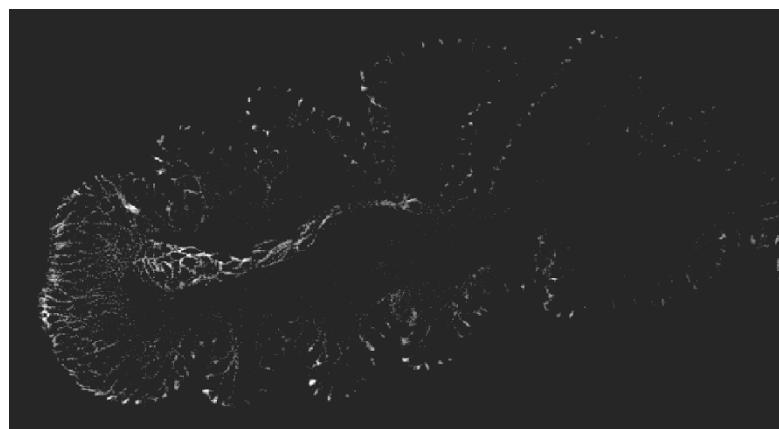

FIGURE 24.15. Optical section shown in Figure 24.3(A) after the inverse filter.
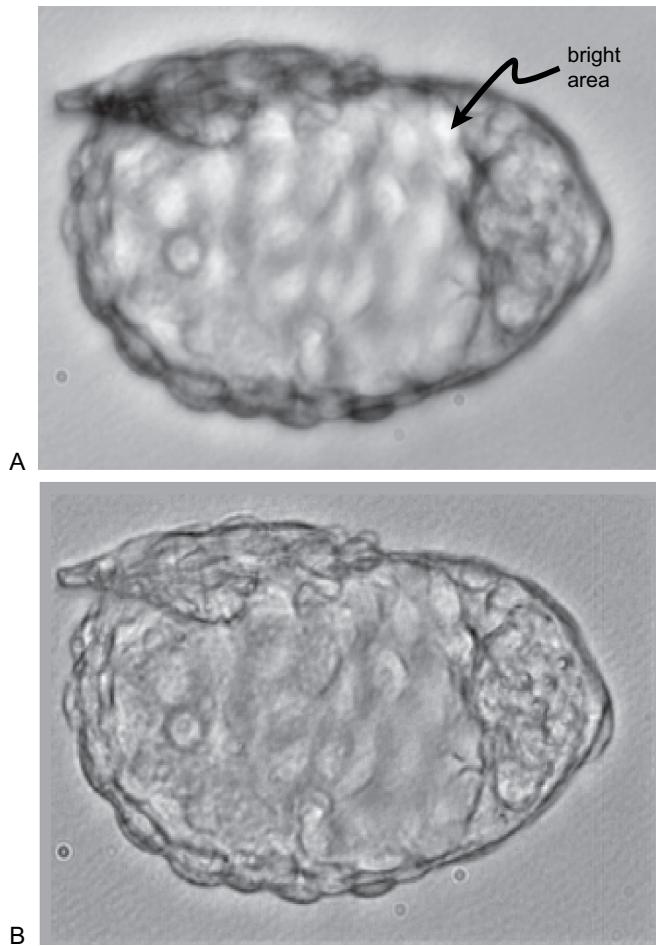
FIGURE 24.16. Summed projections of a phase object obtained by TLB. (A) Unprocessed; (B) deconvolved. (Image courtesy of G. Eddlestone, Empire Imaging Systems.)

fied in the sample, such as a microtubule, or anything that is small, and then focusing up and down while noticing if the out-of-focus object looks the same while focusing up compared to focusing down. This procedure minimizes the spherical aberration, which is the most severe aberration. Alternatively, the microscope manufacturers produce special lenses with adjustable "correction collars" that affect a correction for spherical aberration.

In polarizing and TLB imaging, non-uniformity in the light source can also cause problems (see Chapter 6, *this volume*). The theoretical PSF is cylindrically symmetric because it is formulated under the assumption that light enters the sample with equal intensity at every angle within the numerical aperture. In fact, light enters the sample at different intensities depending upon the angle because the lamp is not uniform. A point in the plane of the lamp corresponds to an angle of parallel light rays in the specimen. The intensity at the center of the lamp is proportional to the intensity of parallel rays passing through the sample at zero angle. The intensity at a distance from the center of the lamp enters the sample at an angle that is proportional to this distance. A mercury-arc lamp is brightest within its arc of light, so angles that correspond to points in this arc have the brightest intensity impinging on the sample. As a result the PSF is not circularly symmetric. Although deconvolution is forgiving of an asymmetric PSF, it still works best with a symmetric one. Ellis designed an illuminator using a vibrating fiber-optic that provides a uniform light source and thereby

provides a symmetric PSF (Ellis, 1979; Inoué and Spring, 1997, p. 160). These non-uniform lamp effects do not exist with the WFF microscope. In this case, the symmetric PSF relies on having spatially uniform excitation light, which is guaranteed by Köhler illumination, and on having well-aligned emission optics. The WFF PSF does not rely on having uniform intensity with respect to the angle of illumination.

The partial confocal effect also causes PSF assymetry (Hiraoka *et al.*, 1990). The field diaphragm behaves like a confocal detector aperture of sorts because it is conjugate to the specimen plane. As it is stopped down, it causes the WFF PSF to look ever more confocal-like and the intensity of the PSF along the optic axis decays more quickly than WFF theory predicts. Furthermore, this decay is assymetric and makes it more difficult to match the measured PSF (with the bead present) to the true PSF (with the specimen present). This effect makes it more difficult to match the assumptions that are used in blind deconvolution. Therefore, it is best, for deconvolution, to keep the field aperture wide open. The partial confocal effect still exists, but it is minimal.

NEWEST DEVELOPMENTS

Subpixel

Images may be deconvolved to produce results that have a pixel resolution different from the original data (Carrington *et al.*, 1995; Conchello and McNally, 1997; Wang, 2002). The idea is illustrated in Figure 24.17. Unobserved pixels are inserted into the mathematical model of the image. They are recovered by treating them as "unknown data" regions as outlined in Holmes (1989).

Figure 24.18(A) shows an image slice from a data stack (Wang, 2002). This is an image of a centrosome labeled with green fluorescent protein (GFP)–gamma-tubulin. The length of the centrosome, along its long axis, is ~1 μm. Figure 24.18(B) shows a deconvolved result made by increasing the pixel resolution by a factor of 3 along *x* and *y*. A verification of the structure is shown in Figure 24.18(C). This figure shows a TEM image of the same sample taken afterwards and verifies the recovered structure.

Khodjakov and colleagues (2002) discovered that these centrosomes are sometimes formed *de novo*, without association with a pre-existing organelle. Deconvolution was used to observe changes in the shape of the pericentriolar material cloud in living samples. Only TEM can record such shapes directly, but not in living cells.
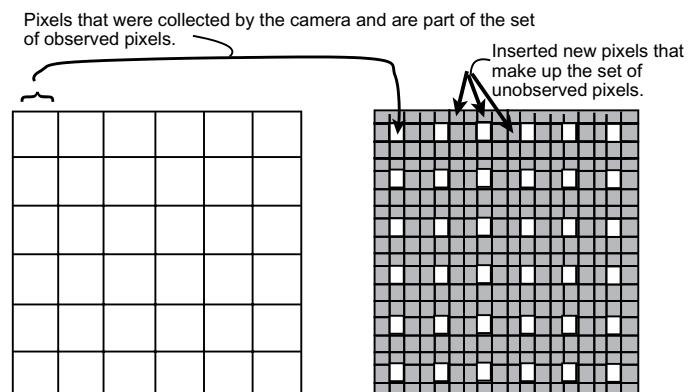
**FIGURE 24.17.** Schematic of subpixel deconvolution.
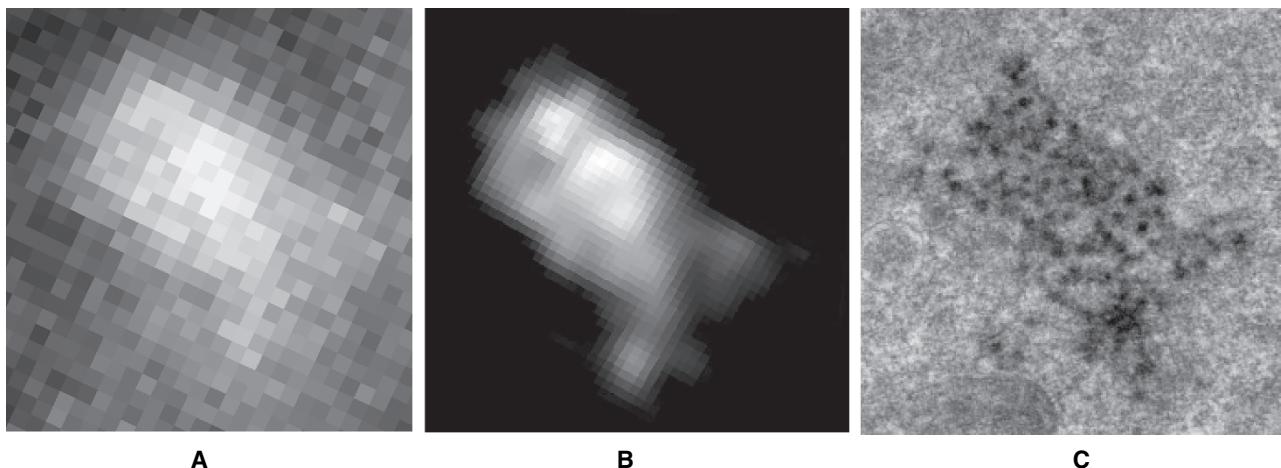
FIGURE 24.18. A subpixel deconvolution. (A) Optical section taken from a stack. The sample is a centrosome labeled with GFP–gamma-tubulin. It is just over 1 μm long. (B) Subpixel deconvolution. The pixel size was reduced by 3 in both directions. (C) EM micrograph of the same sample, verifying the structure shown in (B). (Image courtesy of A. Khodjakov, Wadsworth Laboratories.)

Polarized Light

It is possible to deconvolve 3D optical section data of birefringent samples made using a polarized light microscope. Strictly speaking, the image formation model for a polarized-light microscope is nonlinear (Hansen and Conchello, 1988), so from a pure standpoint, a PSF model such as the one used for WFF ought not to apply. However, from a practical standpoint, it is common engineering practice (Close and Frederick, 1993) to introduce a linear approximation to make any nonlinear model tractable. From this perspective, Eqs. 1 to 3 may apply, and can be considered to be a linearized approximation for the polarized-light microscope. Furthermore, optically sectioned images of birefringent samples sometimes appear similar to WFF samples as is the case shown in Figure 24.19(A), which shows a maximum-value projection stereo-pair of the raw data. Figure 24.19(B) shows the stereo-pair after deconvolution. This deconvolution should work as well with the Pol-Scope developed by Oldenbourg and Mei (1995), which is
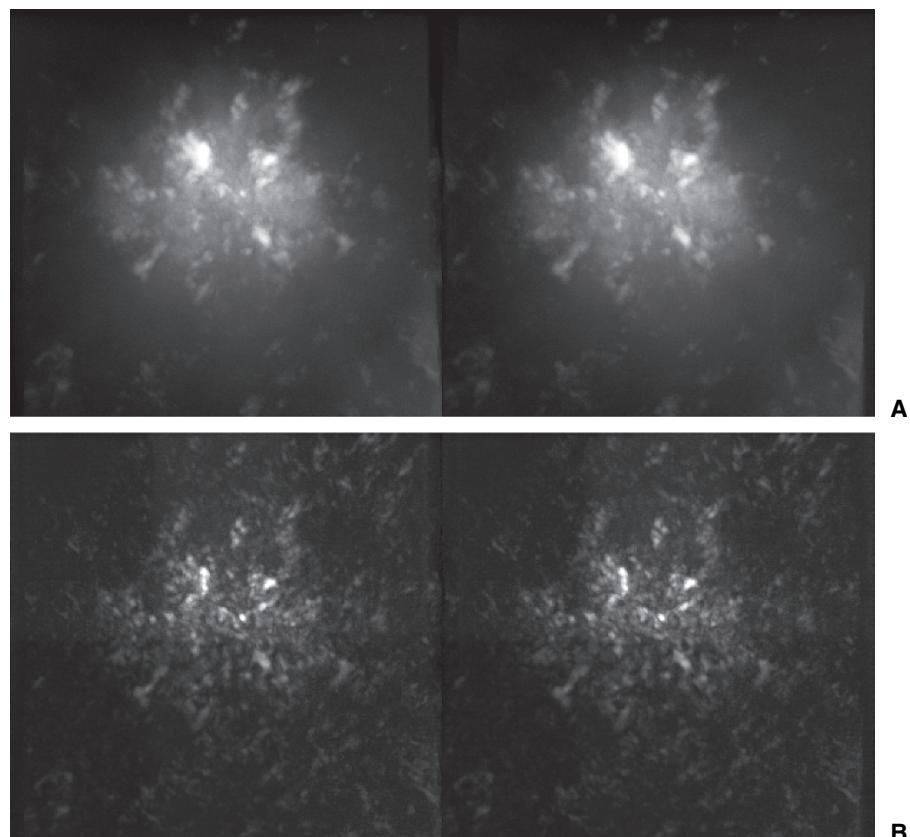
FIGURE 24.19. (A) Stereo-pair of a polarized light stack. This stack was acquired from an ice-cream sample as part of a study of food production quality. The object is a fat crystal. (B) After blind deconvolution. (Image courtesy of K. Baker, University of Guelph.)
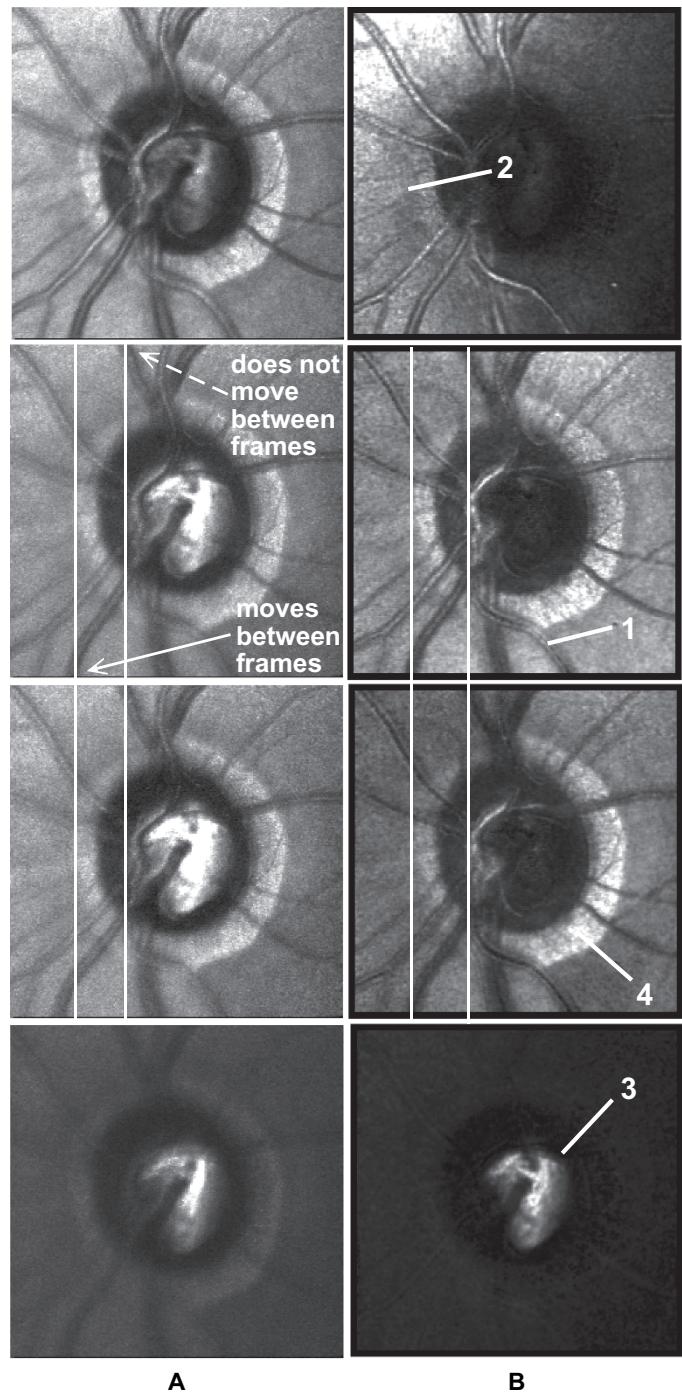
FIGURE 24.20. Sections from an SLO in the reflected light mode. The optic nerve region is shown with large vasculature (1), capillaries (2), the lamina cibrosa (3), and the scleral ring (4). These structures are clearer after deconvolution. Structures are isolated better to their proper location along *z*, as seen by the improved discrimination between the slices. Before deconvolution, the scleral ring and *lamina cibrosa* appear in most slices. After deconvolution, they are properly isolated to their appropriate locations. (Data collected using the Heidelberg Retinal Tomograph (HRT), Heidelberg Engineering, Heidelberg, Germany. Image courtesy of D. Garway-Heath, Moorfields Eye Hospital.)

a special type of polarized light microscope designed for sensitivity and quantitative accuracy. It yields an image in which the intensity is, to a first approximation, proportional to the magnitude of the birefringence. Strictly speaking, there will be a positive, nonlinear dependence on the scalar birefringence.

Live Imaging

Two important issues arise with live imaging: (1) when tissue is monitored over time, many deconvolutions (sometimes hundreds) must be performed. Such applications benefit from processing times on the order of seconds. (2) The sample moves during the image acquisition. The first issue is addressed in "Future Directions."

An example of the second issue is shown in Figure 24.20, which shows confocal sections from a scanning laser ophthalmoscope (SLO; Bartsch and Freeman, 1994; shown schematically in Fig. 24.21). Reflected light confocal sections were acquired during an outpatient procedure. Although the motion of the eye is mitigated by having the patient fixate on a blinking light, the eye still moves between frames and during the 1/20th of a second that each frame is acquired. There is an *x,y* translation and a rotation as well as geometric distortion between frames (O'Connor *et al.*, 1998). The two lines overlaid on Figure 24.20(A) show that there is motion of the blood vessel shown by the solid arrow and, as a result, there is geometric distortion between frames.

The alignment, rotation, and geometric distortion are corrected by preprocessing the optical sections with a software motion-detection algorithm that is based on a mathematical cross-correlation. Figure 24.20(B) shows the deconvolved optical sections. The algorithm described in the section "Algorithms — WFF and CLSM Fluorescence" was used. The geometric distortion is removed. Improved optical sections and better discrimination between slices are produced. The scleral ring, and other known anatomical structures are better isolated to a slice, while blood vessels and capillaries are clearer. Figure 24.22 shows a stereo-pair of the deconvolved stack.

MORE EXAMPLES

Blind Deconvolution and Spherical Aberration

An example of a blind deconvolution was presented earlier in the section entitled "Advantages and Limitations." The first guess of the PSF [Fig. 24.6(B)], used to seed the first iteration of the blind deconvolution of Step 2 in Figure 24.8, was calculated as the autocorrelation of the raw data. The autocorrelation operation may be found in the book by Gaskill (1978). It is, in essence (although not precisely), the convolution (see the Appendix for this chapter) of the PSF with itself. More precisely, it is expressed according to

$$h^{(0)}(x, y, z) = N(x, y, z) * N(-x, -y, -z), \qquad (13)$$

where $h^{(0)}()$ is the first guess of the PSF.
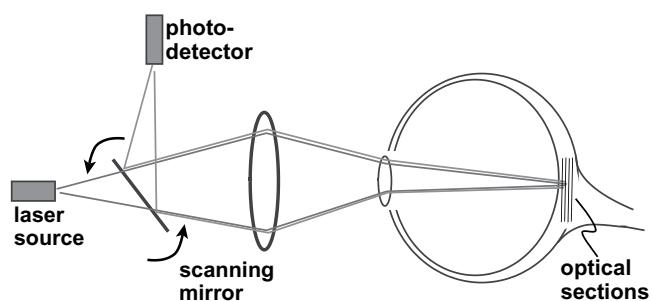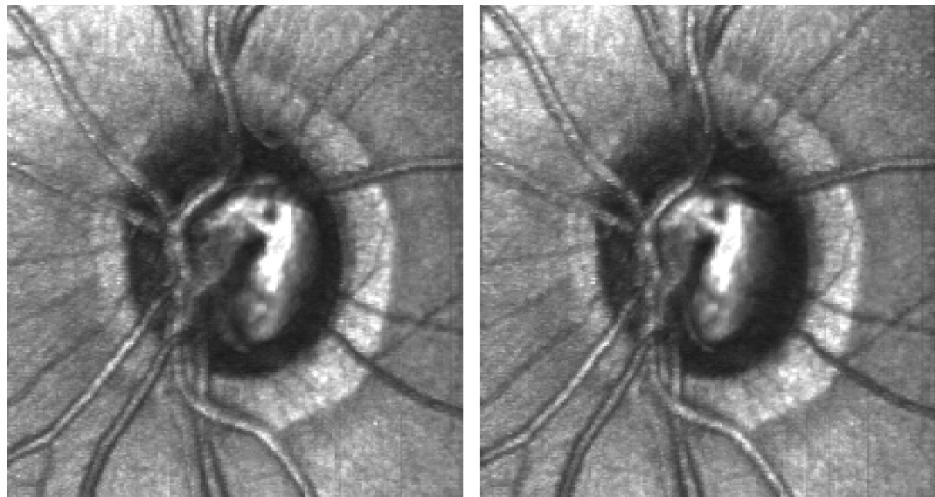
FIGURE 24.21. Schematic of the SLO.

FIGURE 24.22. Stereo-pair of the optic nerve data set shown in Figure 24.20.

The PSF that resulted after running the blind deconvolution starting from this first guess is shown in Figure 24.6(C). Note that it contains the same flare seen in the raw data of Figure 24.6(A). For comparison, Figure 24.6(D) shows a theoretically calculated PSF with spherical aberration that matches the flare seen in Figure 24.6(A). Figure 24.6(E) shows a theoretically calculated PSF without spherical aberration. From these comparisons it is obvious that the predominant flare character seen in raw data of Figure 24.6(A) and in the recovered PSF of Figure 24.6(C) is due to spherical aberration.

Widefield Fluorescence Simulation

Simulated deconvolutions have been carried out to guide development of deconvolution programs. Though not exhaustive, these simulations are instructive because, unlike a biological specimen, the original "test object" is known. Simulations provide insight into the advantages and limitations of different deconvolution approaches. Figures 24.23 and 24.24 show a WFF simulation.

A computer program created the object shown in Figures 24.23 and 24.24(A). The theoretical widefield PSF, like the one shown in Figure 24.6(E), was created by a computer program that used the diffraction theory as explained by Streibl (1984). This 3D PSF was convolved with the 3D object using Fourier transforms (Gaskill, 1978), and the Fast-Fourier-Transform algorithm (Oppenheim and Schafer, 1975). Once the diffraction-limited 3D data set was created in this way, quantum-photon noise was simulated by software written to create random numbers that follow Poisson statistics (Mortensen, 1987).

Figure 24.23 shows a schematic of the simulated test object (Markham and Conchello, 1998). It has the shape of wagon wheel pasta. When the out-of-focus haze is removed by the deconvolution, the spokes on the wagon wheel and the surface of the wheel are restored, except for those parts that are horizontally planar. The inability of deconvolution to restore horizontal planes is a well-known limitation of widefield deconvolution and is caused by the missing cone region (Holmes, 1992; Holmes *et al.*, 1995; Streibl, 1984) (Fig. 24.11) which removes all of the information necessary to restore a horizontal plane. Fortunately, very few biological structures have perfectly horizontal planar components labeled by

the dye. Those that do are better imaged with a confocal microscope. Note that the same horizontal spokes that are not restored in Figure 24.24(C) are restored in Figure 24.24(D), where the horizontal parts are at the edge of the cylinder, so they are not planar. Here, the necessary Fourier components needed to restore the horizontal parts reside outside of the missing cone, and therefore the deconvolution algorithm has sufficient information to recover them.

Spinning-Disk Confocal

Figure 24.25 shows a deconvolved data set from a spinning-disk confocal microscope. An advantage of the blind deconvolution is that it is capable of adapting to many modalities because the PSF does not need to be known perfectly.

Two Photon

Figure 24.26 shows optical sections made with a two-photon microscope, before and after deconvolution. The algorithm used is identical to that used with the confocal microscope, because the PSF of the two-photon microscope is similar to that of the confocal microscope. The signal-to-noise level and optical sectioning discrimination are improved by the deconvolution. Although not clear from this example, the in-plane resolving power may also be improved (see also Chapter 25, *this volume*).
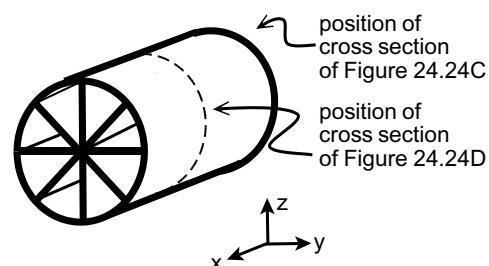
FIGURE 24.23. Schematic of the simulated test object.
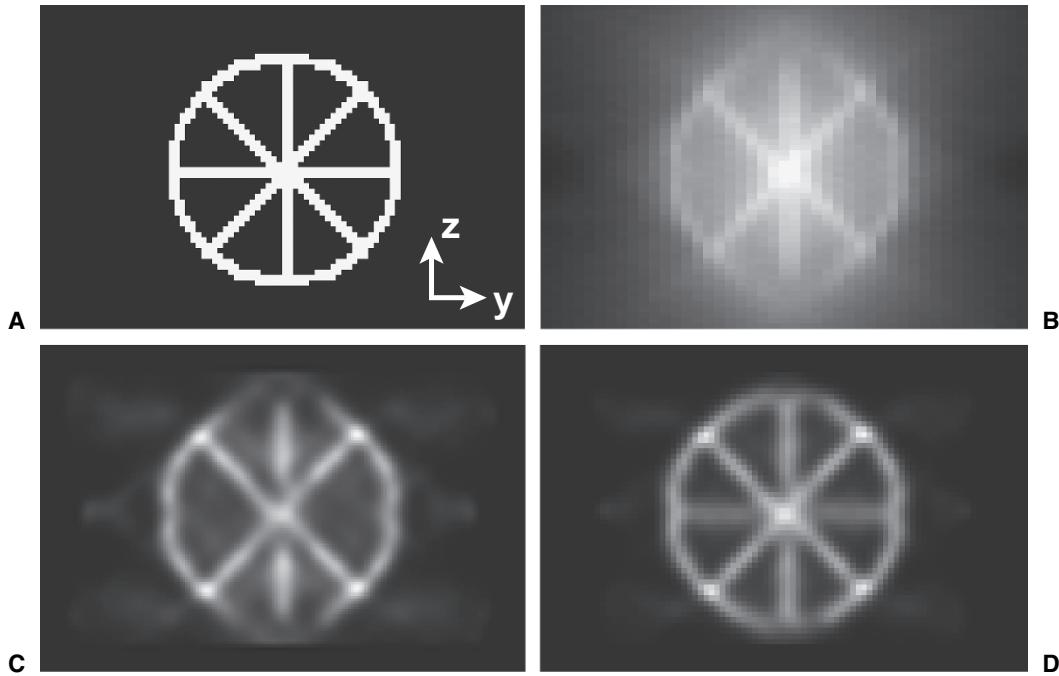
FIGURE 24.24. (A) *yz* slice of the test object (Fig. 24.23). (B) *yz* slice of the simulated WFF data. (C) Deconvolution of the central cross-section. (D) Deconvolution of the cross-section taken at the end of the cylinder.

SPEED

Prior speed limitations (Holmes *et al.*, 1995) are now less of a problem. Twelve years ago deconvolutions were performed on a mainframe supercomputer and now they are performed on a PC. Numbers of iterations used to be in the 100s and now they are on the order of 10. Deconvolutions are performed in seconds rather than hours. Even so, there will always be demand for higher speed. The operator needs fast visual feedback during four-dimensional (4D) live imaging, especially now that the electron multiplier CCD (EM-CCD) makes it possible to see faint signals at high frame rates with little readout noise. Current deconvolution speeds are listed in Table 24.1, see p. 476.
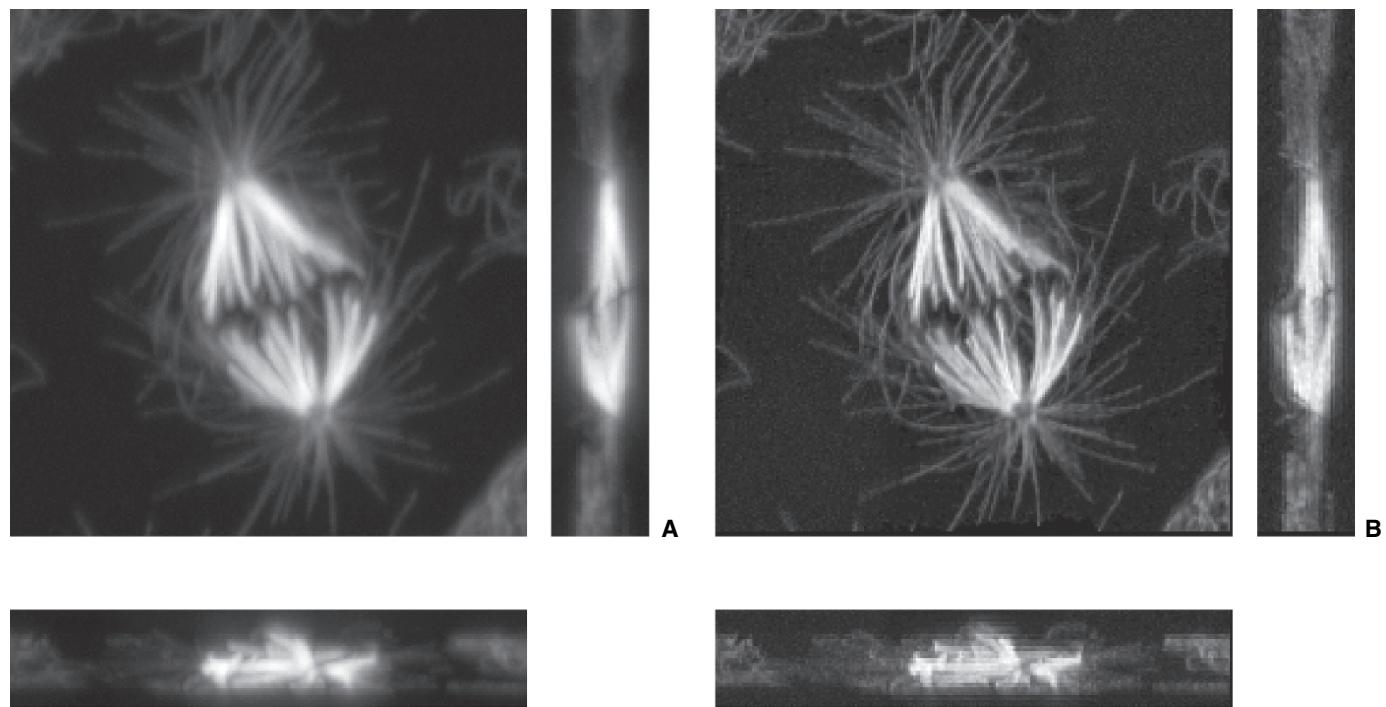
FIGURE 24.25. (A) Orthogonal maximum projections of a fluorescent sample taken from a Yokogawa spinning-disk microscope (Ichihara *et al.*, 1996). (Image courtesy of Baggi Somusundaram, PerkinElmer Life Sciences.) (B) Deconvolved result.
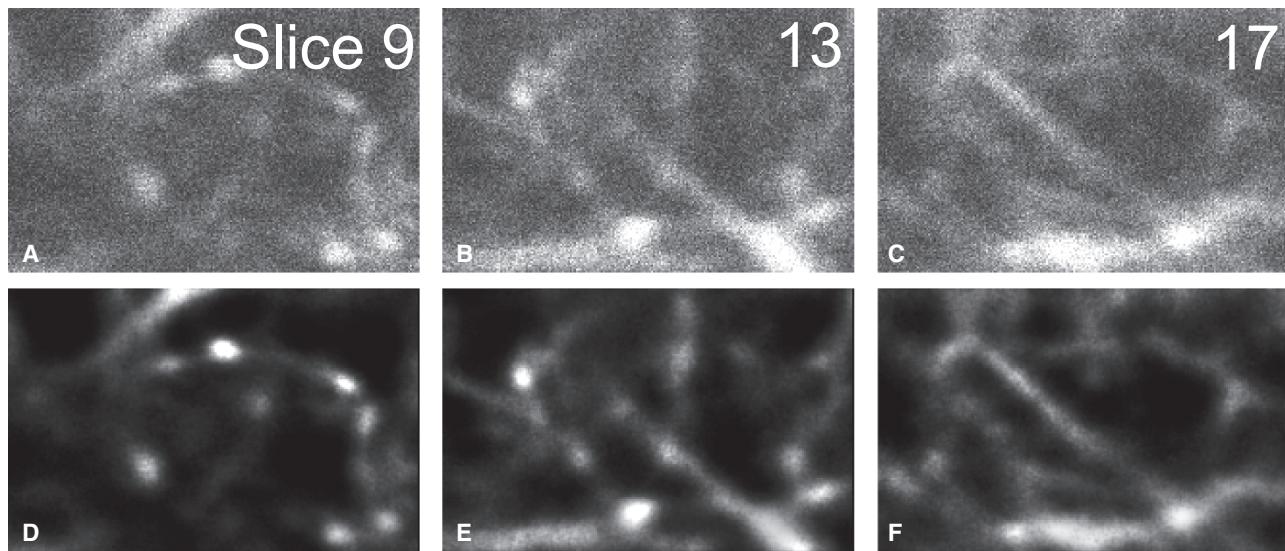
FIGURE 24.26. Sections of a two-photon data set. (A–C) Unprocessed. (D–F) Deconvolved. (Image courtesy of M. Ellisman, University of California, San Diego.)

FUTURE DIRECTIONS

Parametric modeling of more aberrations for the calculation of the PSF during the blind deconvolution by Markham and Conchello, (1999) will make the deconvolution more robust against aberrations. These aberrations include astigmatism, coma, and various types of chromatic aberration.

The blind deconvolution approach provides many efficient ways to utilize the reconstructed PSF obtained from one deconvolution to help make the next deconvolution converge more rapidly. One way is to use the PSF arrived at from a given data set as the first-guess PSF when processing other data collected under similar conditions. Another way is to use the PSF that was arrived at from a given subfield (or sub-volume) in a data set as a first guess when deconvolving a neighboring subfield. The actual PSF varies across a large field because the wavefront aberrations are dependent upon the x-, y-, and z-locations (see Fig. 23.5, *this volume*). Presently, blind deconvolution adapts to these positional changes by breaking up the data set into smaller sub-volumes and reconstructing an independent/adaptive PSF for each sub-volume.

Although, in a strict sense, every pixel ought to have its own PSF, the amount of computation necessary makes doing this extremely impractical. Having subfields with constant PSFs within each field is a compromise that works because, even though the PSF varies, it varies slowly across the field.

Because camera sizes and the number of channels will increase, and live, time-series imaging will become ever more common, processing speed improvements will always be a requirement of advancing work. Now that we can achieve deconvolutions in seconds instead of minutes, we would like to have them done in milliseconds. High-throughput pharmeceutical screening requires deconvolving large 3D data sets in real time (see Chapter 46, *this volume*). To measure ion concentrations in 3D and in real time requires millisecond deconvolutions.

There are two main approaches to meet these speed demands. The algorithmic approach involves refining the mathematical models and designing trade-offs so that the algorithm converges mathematically at a faster rate. A second approach is to utilize distributed processing and distributed computing. Distributed processing is a term used to describe the decomposition of computations into independent modules and sending each of these modules to a different CPU within the same computer. The Windows platform uses a programming technique called multi-threading, which provides the tools that allow programmers to carry out distributed processing in the Windows environment. Many of the Apple and Unix-based computers also provide distributed processing capabilities. In distributed computing, which is different from distributed processing, these self-contained parts are sent over the local network to different computers (rather than to different CPUs on the same computer) and then returned to the host computer after the tasks are completed. Deconvolution procedures can easily be decomposed by breaking up the data set into subvolumes (Bhattacharyya *et al.*, 1996).

SUMMARY OF MAIN POINTS

- The MLE blind deconvolution approach to deconvolution has a number of important advantages — no measurement of PSF is needed and inherent noise suppression.
- There are many other types of deconvolution, each one having different speed and performance advantages. Generally, speed is traded for image quality (resolution and/or noise).
- Deconvolution algorithms exist for many microscopic contrast modes: WFF, CLSM fluorescence, SLO, spinning-disk confocal, TLB, DIC, polarized light, two-photon.
- Although past limitations in speed have been surpassed, the need to process data from larger cameras more rapidly means that the processor is still a limit. Distributed processing and algorithm improvements will meet this increasing demand.

ACKNOWLEDGMENTS

REFERENCES

Agard, D.A., 1984, Optical sectioning microscopy: Cellular architecture in three dimensions, *Annu. Rev. Biophys. Bioeng.* 13:191–219.

Agard, D.A., Hiraoka, Y., and Sedat, J.W., 1989, Three-dimensional microscopy: Image processing for high resolution subcellular imaging, *SPIE* 1161:24–30.

Ayers, G.R., and Dainty, J.C., 1988, Iterative blind deconvolution method and its applications, *Opt. Lett.* 13:547–549.

Bartsch, D., and Freeman, W.R., 1994, Axial instensity distribution of the human retina with a confocal scanning laser tomograph, *Exp. Eye Res.* 58:161–173.

Bhattacharyya, S., Szarowski, D.H., Turner, J.N., O'Connor, N.J., and Holmes, T.J., 1996, The ML blind deconvolution algorithm: Recent developments, *SPIE* 2655:175–186.

Biggs, D.S.C., 1998, Accelerated iterative blind deconvolution, PhD thesis, University of Auckland, New Zealand.

Carrington, W.A., 1990, Image restoration in 3D microscopy with limited data, bioimaging and two-dimensional spectroscopy, *SPIE* 1205:72–83.

Carrington, W.A., Lynch, R.M., Moore, E., Isenberg, G., Fogarty, K.E., and Fay, F.S., 1995, Superresolution three-dimensional images of fluorescence in cells with minimal light exposure, *Science* 268:1483–1487.

Castleman, K.R., 1979, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

Close, C.M., and Frederick, D.K., 1993, Modeling and Analysis of Dynamic Systems, 2nd ed., Houghton Mifflin, Boston.

Conchello, J.A., and McNally, J.G., 1997, Subpixel resolution in maximum likelihood image restoration, *SPIE* 2984:158–168.

Conchello, J., Kim, J., and Hansen, E.W., 1994, Enhanced three-dimensional reconstruction from confocal scanning microscope images. II. Depth discrimination versus signal-to-noise ratio in partially confocal images, *Appl. Opt.* 33:3740–3750.

Cooper, J.A., Bhattacharyya, S., Turner, J.N., and Holmes, T.J., 1993, Three-dimensional transmitted light brightfield imaging: Pragmatic data collection and preprocessing considerations, *MSA Annu.* 51:276–277.

Dempster, A.P., Laird, N.M., and Rubin, D.B., 1977, Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statistic. Soc. B.* 39:1–37.

Dunn, G.A., 1988, Transmitted-light interference microscopy: A technique born before its time, *Proc. Roy. Microsc. Soc.* 33:189–196.

Ellis, G.W., 1979, A fiber-optic phase randomizer for microscope illumination by laser, *J. Cell. Biol.* 83:303a.

Erhardt, A., Zinser, G., Komitowski, D., and Bille, J., 1985, Reconstructing 3-D light-microscopic images by digital image processing, *Appl. Opt.* 24:194–200.

Gaskill, J.D., 1978, *Linear Systems, Fourier Transforms and Optics*, John Wiley and Sons, New York.

Gerchberg, R.W., and Saxton, W.O., 1974, Super-resolution through error energy reduction, *Opt. Acta* 21:709–720.

Gibson, S.F., and Lanni, F., 1991, Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy, *J. Opt. Soc. Am. A* 8:1601–1613.

Hansen, E., and Conchello, J., 1988, Restoring image quality in the polarizing microscope: Analysis of the Allen video-enhanced contrast method, *J. Opt. Soc. Am. A* 5:1836–1847.

Hiraoka, Y., Sedat, J.W., and Agard, D.A., 1990, Determination of three-dimensional imaging properties of a light microscope system: Partial confocal behavior in epifluorescence microscopy, *Biophys. J.* 57:325–333.

Holmes, T.J., 1989, Expectation-maximization restoration of band-limited, truncated point-process intensities with application in microscopy, *J. Opt. Soc. Am. A* 6:1006–1014.

Holmes, T.J., 1992, Blind deconvolution of quantum-limited incoherent imagery, *J. Opt. Soc. Am. A* 9:1052–1061.

Holmes, T.J., and Levy, W.J., 1987, Signal-processing characteristics of differential-interference-contrast microscopy, *Appl. Opt.* 26:3929–3939.

Holmes, T.J., and Liu, Y., 1989, Richardson-Lucy/maximum likelihood image restoration algorithm for fluorescence microscopy: Further testing, *Appl. Opt.* 28:4930–4938.

Holmes, T.J., and Liu, Y.H., 1991, Acceleration of maximum-likelihood image-restoration for fluorescence microscopy and other noncoherent imagery, *J. Opt. Soc. Am. A* 8:893–907.

Holmes, T.J., and Liu, Y., 1992, Image restoration for 2-D and 3-D fluorescence microscopy, In: *Visualization in Biomedical Microscopies: 3-D Imaging and Computer Applications*, (A. Kriete, ed.), VCH, Weinheim, Germany, pp. 282–327.

Holmes, T.J., and O'Connor, N.J., 2000, Blind deconvolution of 3D transmitted light brightfield micrographs, *J. Microsc.* 200:114–127.

Holmes, T.J., Bhattacharyya, S., Cooper, J.A., Hanzel, D., Krishnamurthi, V., Lin, W., Roysam, B., Szarowski, D.H., and Turner, J.N., 1995, Light microscopic images reconstructed by maximum likelihood deconvolution, In: *Handbook of Biological Confocal Microscopy* (J.B. Pawley, ed.), 2nd ed., Plenum Press, New York, pp. 389–402.

Holmes, T.J., Liu, Y.H., Khosla, D., and Agard, D.A., 1991, Increased depth-of-field and stereo pairs of fluorescence micrographs via inverse filtering and maximum likelihood estimation, *J. Microsc.* 164:217–237.

Ichihara, A., Tanaami, T., Isozaki, K., Sugiyama, Y., Kosugi, Y., Mikuriya, Abe, M., and Uemura, I., 1996, high-speed confocal fluorescence microscopy using a Nipkow scanner with microlenses for 3-D imaging of single fluorescent molecule in real time, *BioImage* 4:57–62.

Inoué, S., and Spring, K.R., 1997, *Video Microscopy*, 2nd ed., Plenum Press, New York.

Janesick, J.R., Elliott, T., and Collins, S., 1987, Scientific charge-coupled devices, *Opt. Eng.* 26:692–714.

Kam, Z., 1998, Microscopic differential interference contrast image processing by line integration (LID) and deconvolution, *Bioimaging* 6:166–176.

Khodjakov, A., Rieder, C.L., Cassels, G., Sibon, O., and Wang, C.-L., 2002, De novo formation of centrosomes in culture cells arrested during s phase. *J. Cell Biol.* 158:1171–1181.

Krishnamurthi, V., Turner, J.N., Liu, Y., and Holmes, T.J., 1995, Blind deconvolution of fluorescence microscographs by maximum likelihood estimation, *Appl. Opt.* 34:6633–6647.

Llacer, J., and Veklerov, E., 1989, Feasible images and practical stopping rules for iterative algorithms in emission tomography [errata published in *IEEE Trans. Med. Imaging* 1990;9:112], *IEEE Trans. Med. Imaging* 8:186–193.

Macias-Garza, F., Diller, K.R., Bovik, A.C., Aggarwal, S.J., and Aggarwal, J.K., 1989, Improvement in the resolution of three-dimensional data sets collected using optical serial sectioning, *J. Microsc.* 153:205–221.

Markham, J., and Conchello, J.A., 1998. Parametric blind deconvolution of microscopic images: Further results, *SPIE Proc.* 3261:38–49.

Markham, J., and Conchello, J.A., 1999. Parametric blind deconvolution: A robust method for the simultaneous estimation of image and blur, *J. Opt. Soc. Am. A* 16:2377–2391.

Mortensen, R.E., 1987. *Random Signals and Systems*, John Wiley and Sons, New York.

O'Connor, N.J., Bartsch, D., Freeman, W.J., Mueller, A.J., and Holmes, T.J., 1998, Fluorescent infrared scanning-laser ophthalmoscope for three-dimensional visualization: Automatic random-eye motion correction and deconvolution, *Appl. Opt.* 37:2021–2033.

Oldenbourg, R., and Mei, G., 1995, New polarized light microscope with precision universal compensator, *J. Microsc.* 180:140–147.

Oppenheim, A.V., and Schafer, R.W., 1975, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

Preza, C., van Munster, E.B., Aten, J.A., Snyder, D.L., and Rosenberger, F.U., 1998, Determination of direction-independent optical path-length distribution of cells using rotational diversity transmitted-light differential interference contrast (DIC) images, *SPIE Proc.* 3261:60–70.

Russ, J.C., 1995, *The Image Processing Handbook*, 2nd ed., CRC Press, Boca Raton, Florida.

Sheppard, C.J.R., and Gu, M., 1994, Three-dimensional transfer functions for high-aperture systems, *J. Opt. Soc. Am. A.* 11(2):593–598.

Sibarita, J.-B., Magnin, H., and De Mey, J.R., 2002, Ultra-fast 4D microscopy and high throughput distributed deconvolution, *Proc. 2002 Int. Symp. Biomed. Imaging* 2002:769–772.

Snyder, D.L., Hammoud, A.M., and White, R.L., 1993, Image recovery from data acquired with a charge-coupled-device camera, *J. Opt. Soc. Am. A* 10:1014–1023.

Streibl, N., 1984, Depth transfer by an imaging system, *Opt. Acta.* 31:1233–1241.

van Munster, E.B., van Vliet, L.J., and Aten, J.A., 1997, Reconstruction of optical pathlength distributions from images obtained by a widefield

differential interference contrast microscope, *J. Microsc.* 188:149–157.

Van Trees, H.L., 1968, *Detection, Estimation, and Modulation Theory*, John Wiley and Sons, New York.

Veklerov, E., and Llacer, J., 1987, Stopping rule for the MLE algorithm based on statistical hypothesis testing, *IEEE Trans. Biomed. Imaging* 6:313–319.

Visser, T.D., Oud, J.L., and Brakenhoff, G.J., 1992, Refractive index and axial distance measurements in 3-D microscopy, *Optik* 90:17–19.

Wallace, W., Schaefer, L.H., and Swedlow, J.R., 2002, A workingperson's guide to deconvolution in light microscopy, *Biotechniques* 31:1–16.

Wang, C., 2002, Three-dimensional blind deconvolution for light microscopy: Fundamental studies and practical implementations, PhD thesis, Rensselaer Polytechnic Institute, Troy, New York.

APPENDIX

Convolution: An Introduction For Those Who Don't Remember Second-Year Calculus

James B. Pawley and Timothy J. Holmes

The word "deconvolution" has achieved a mantra-like status among many light microscopists. It is often translated as the magic process that starts with blurry, hazy 3D data and produces sharp, contrasty 3D structures. While this is a powerful capability, and, as is stated throughout this volume, 3D data sets are almost always improved by being deconvolved, it is important to recognize that deconvolution is just the reverse of a mathematical operation called convolution. In the context of 3D light microscopy, convolution can be thought of as a function that mimics the effect of the blurring caused by diffraction and the addition of out-of-focus light.

The purpose of this Appendix is to demystify deconvolution by explaining (albeit in a simplified, heuristic way) the essential features of the mathematical process called convolution.

Let us start with the mathematical process of multiplication. We all understand multiplication: You take one number (a constant) and add it to itself as many times as the values of some second number. Integration is similar, except that one of the things you are "multiplying" is a variable, usually called the function and plotted in the *y*-direction. The integral of a function is the area under the plot of the function between some upper and lower "limits" along the *x* axis. So if you imagine dividing this space up into vertical strips, all of the same width (the constant), but of variable length (where the length is the value of the function for this value of *x*), then the total area can be thought of as multiplying a constant by a variable.

Convolution takes this process a step further. It is a special type of integration in which one multiplies one function (or variable) by another function (or variable), and to make things even more fun, the function can be a variable in more than one dimension (i.e., in *x*, *y*, and *z*, rather than just *x* as above).

I can hear your eyes glazing over!

Before you panic about functions, you don't have to think about equations here. Any set of paired numbers will do. For example, the pairing of a set of intensity numbers with locations in an *xy* raster is a function. In other words, a digital image is a function of *x* and *y*.

Perhaps some pictures can help show how this works, at least for a 2D function (i.e., a sort of image) as shown in Figure 24.A1.

We will start with two very simple 2D functions:

- A point object (shown in green): all the pixels equal 0 except the one in the middle, which equals 1.
- A simple blurring function (shown in red): a 2D blob about the shape of the central peak of an Airy disk, sampled with Nyquist-sized pixels.

To convolve the red function with the green one, first imagine setting the center pixel of the red function over the pixel at the top-left corner of the green one and then multiplying the value in each red pixel by the value in the green pixel that it overlaps. Then add up all these products and enter the sum in the top left pixel of a new "result" raster (in black). As initially all the multiplications involve a zero, all the products are zero and so is their sum, so you enter a zero in the top-left pixel (long arrow). Now step the red mask one step to the right and repeat. Another zero (short arrow). Eventually, however, as you proceed to step the red mask across and down, a time will come when the non-zero pixel in the green mask will coincide with a non-zero pixel in the red mask and a non-zero sum of products will be entered in the black raster (long, pink arrow, Fig. 24.A2).

It now becomes obvious that this slow, boring process will make the result look just like the red function. In fact, we have just mimicked the process of making a blurry image of a point object and our black result is, therefore, the infamous point-spread function, or PSF. To see that this process is a little more useful than just copying the red function, suppose now that the object is composed of two point objects, separated by five "zero pixels," as
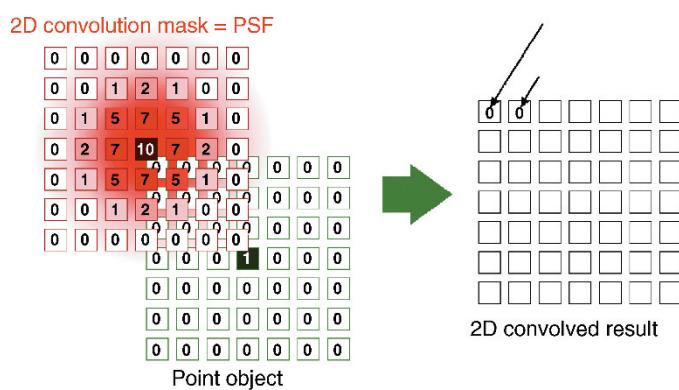

FIGURE 24.A1. First two steps of convolution operation between a point and a mask.

2D convolution mask

Point object

2D convolved result
(the image)
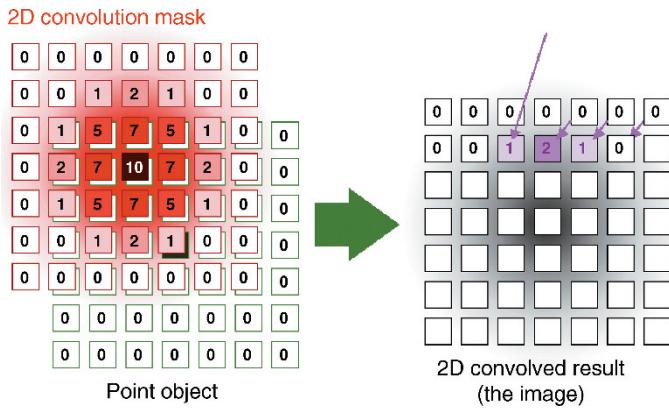
FIGURE 24.A2. Convolution of a point with a mask, continued.

focus plane at
point object

focus plane above
point object

focus plane farther
above point object

Images of point object, recorded at each focus position

FIGURE 24.A4. Three-dimensional blurring function applied to a point object in widefield fluorescence.

shown in the green raster of Figure 24.A3. When this object is convolved with the mask, the result (black raster) resembles a blurred image of two points.

To the extent that the mask mimics a 2D, diffraction-limited, point-spread function, the result in the black raster will look like a diffraction-limited image of an object composed of two dark points on a white background. As we have chosen a pixel size close to half of the Rayleigh resolution, two points separated by six pixels will be represented by two discernible features in the black raster. You may be more convinced of this if you look at Figure 24.A3 from the other side of the room.

The discussion so far has implied that convolution is a process that occurs in just two dimensions. Of course, in the case of microscopy, this is not true. Because the Airy figure is a 3D structure, the blurring it causes occurs in three dimensions. The 3D blurring in the near-focus region is difficult to diagram in 2D. However, as one moves away from this plane, it is a fair approximation to assume that the light passing through the objective appears as a circle. This circle is the cross-section of a cone centered at the plane of the object. The point object defines the apex of the cone and the objective NA defines its angle (Fig. 24.A4).

Three-dimensional convolution requires considering the out-of-focus light in the vertical direction as well as blurred light from the in-focus plane. The problem becomes more complex when two points are located at different heights in the specimen (see Fig. 24.A5).

It is important to remember that the total number of photons leaving the point and passing through the objective remains virtually constant no matter where the objective is focused. Therefore, the image of an in-focus point appears very intense, while the image of an out-of-focus point is much dimmer because now the same number of photons are being spread over a larger, blurred circle. It can be instructive to see just how important this effect can be.

About Out-of-Focus Light

Suppose that under given widefield illumination conditions, a CCD with Nyquist-sized pixels (0.1 μm referred to the sample) records a total of 10,000 photons from a particular point object in each image in the stack. When the point object is located at the focus plane, as illustrated in Figure 24.A6, about 80% of the 10,000 photons are distributed between the 12 bright pixels centered on the object or about 650 photons/pixel. Assuming a high-NA lens

Convolution mask

Convolved result
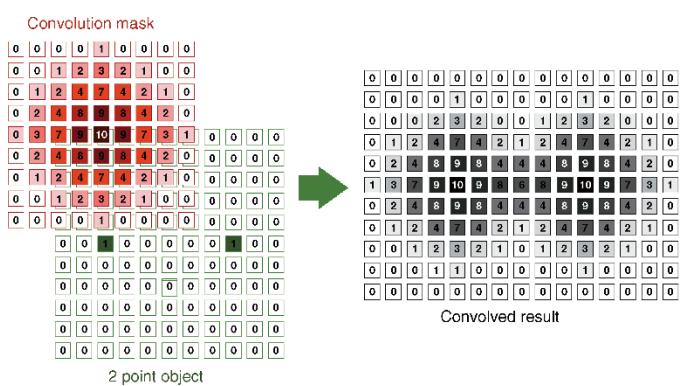
2 point object

FIGURE 24.A3. Convolution of a mask and a two-point object.

images recorded from each focus plane.
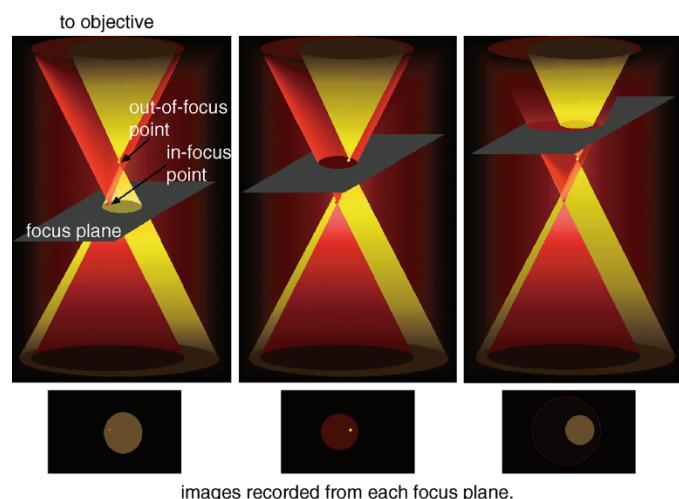
FIGURE 24.A5. Three-dimensional blurring function applied to two point objects in different planes in widefield fluorescence.

Area of focused
spot ~0.126 μm²
or about 12 pixels

Just 1 μm above the focus
plane, the beam area is
9.4 μm² or about 940 pixels.
As this is ~80x more pixels
than at the focus plane, the
light is ~80x less intense.

1.73 μm

0.1 μm

Focus
plane

60°   1 μm

If diam. of Airy dark ring is 0.4 μm,
CCD pixels should be 0.1 μm square,
Pixel area = 0.01 μm².
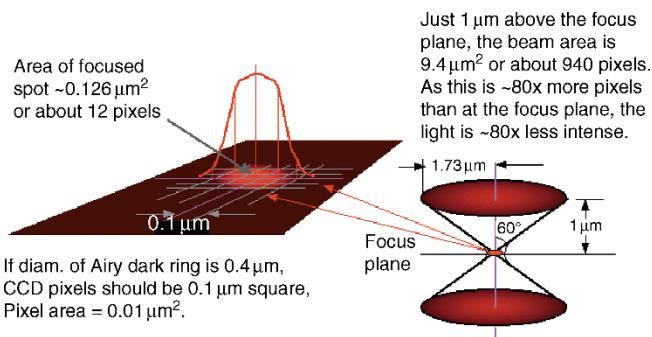
FIGURE 24.A6. Geometrical optics used to calculate the rapidity with which the signal from a point object diminishes with distance from the focus plane in widefield microscopy.

($\alpha = 60°$) and simple geometrical optics,[2] by the time the feature is only 1 μm out of focus, these photons will be spread over a circle having an area of 9.4 μm² or 940 pixels. If the intensity is spread evenly over this circle, then the detected signal level from the now out-of-focus feature drops to <7 photons/pixel. As this is only slightly more than the RMS noise level of a good, slow-scan CCD, it is hard to see how the detection of this out-of-focus light can provide much useful structural information.

The result is that, when the NA is high, the signal intensity recorded per pixel from a point source drops off very fast as the point goes out of focus, even in widefield fluorescence microscopy. As noted in Chapter 23, signals emanating from small features located much more than 2 μm from the focus plane are more likely to add Poisson noise than to add useful structural information.

Convolution, Fourier Transforms, and the FFT

The tedious process of overlapping, multiplying, adding, and replacing would not be of much use were it not for the fact that the process can be greatly simplified if we first apply a mathe-

matical operation called the Fourier transform (FT) to the 3D data sets representing the PSF and the object. The FT of an object looks a lot like its diffraction pattern. Like the "real-space" object, the "Fourier-space object" is a 3D distribution of values but, whereas the real object has only a single value at each 3D location (the intensity), the FT of the object has both an intensity and a phase angle. The phase angle information is used by the inverse FT process to convert the FT data back into a real-space object. To perform the process of convolution using FTs, we must:

1. Calculate the 3D FT of the object, creating a 3D array of intensity and phase values.
2. Calculate the 3D FT of the PSF, creating a second 3D array of intensity and phase values.
3. Simply **multiply** these two arrays, point by point, creating a new 3D intensity/phase array.
4. Calculate the inverse Fourier transform of this product array.

This may all sound very obscure, but fortunately there is a very fast and helpful computer algorithm called the fast Fourier transform that can rapidly perform both the 3D transform and its inverse!

Coincidentally, calculating the FT of an object exactly duplicates the effect that a perfect lens has on light that diffracts at its focus plane and then proceeds to its back-focal plane (BFP). That is, given a perfect lens, the phase and intensity pattern of the light passing through the BFP *is*[3] the 2D FT of the object at the focus plane. Unfortunately, we cannot just record the image of the BPF and then apply the inverse FT to reconstruct the 2D object because we can only record the intensity, not the phase, of this pattern.

X-ray diffraction studies confront the same problem but in this case, the ordered, crystalline nature of the specimen makes it possible to make shrewd guesses about the phases. As a result, success in discerning molecular structure depends on first measuring the intensity of the diffraction pattern and then on guessing and iterating ever-more-accurate estimates of the phases.

---

[2] An admitted oversimplification that works better as the distance from the focus plane increases.

[3] Strictly speaking, it is the **magnitude** and the phase of E-field in the BFP that is the FT of the E-field and phase in the object plane. The light **intensity** in each plane, is the square of these magnitudes. But, this is getting us into technical details that require their own chapter.