4

Points, Pixels, and Gray Levels: Digitizing Image Data

James B. Pawley

CONTRAST TRANSFER FUNCTION, POINTS, AND PIXELS

Microscopical images are now almost always recorded digitally. To accomplish this, the flux of photons that forms the final image must be divided into small geometrical subunits called pixels. The light intensity in each pixel will be stored as a single number. Changing the objective magnification, the zoom magnification on your confocal control panel, or choosing another coupling tube magnification for your charge-coupled device (CCD) camera changes the size of the area on the object that is represented by one pixel. If you can arrange matters so that the smallest feature recorded in your image data is at least 4 to 5 pixels wide in each direction, then all is well.

This process is diagrammed for a laser-scanning confocal in Figure 4.1, where the diameter of the scanning beam is shown to be at least four times the interline spacing of the scanning raster. This means that any individual fluorescent molecule should be excited by at least four overlapping, adjacent scan lines and that, along each scan line, it will contribute signal to at least four sequential pixels. Finally, it is important to remember that information stored following these rules will only properly resemble the original light pattern if it is first spatially filtered to remove noise signals that are beyond the spatial bandwidth of the imaging system. Image deconvolution is the most accurate way of imposing this reconstruction condition and this applies equally to data that have been collected by widefield or scanning techniques. If you do this right, your image should look like that in Figure 4.2.

If you are already convinced of this, jump to page 71 for the second half of this chapter, on gray levels. But if it all seems to be irrelevant mumbo-jumbo, read on. Incorrect digitization can destroy data.

Pixels, Images, and the Contrast Transfer Function

If microscopy is the science of making magnified images, a proper discussion of the process of digitizing these images must involve some consideration of the images themselves. Unfortunately, microscopic images are a very diverse breed and it is hard to say much about them that is both useful and specific. For the purposes of discussion, we assume that any microscopic image is just the sum of the blurred images the individual "point objects" that make up the object.

But what is a point object? How big is it? Is it the size of a cell, an organelle, or a molecule? Fortunately, we don't have to answer this question directly because we aren't so much interested in a point on the object itself as in the image of such an object. As should be clear from Chapters 1 and 2, our ability to image small features in a microscope is limited at the very least by the action of diffraction.[1]

So point objects can be thought of as features smaller than the smallest details that can be transmitted by the optical system. The final image is merely the sum of all the point images. Although the images themselves may be varied in the extreme, all are composed of mini-images of points on the object.

By accepting this simplification, we can limit our discussion to how best to record the data in images of points. Of course, we need more than the ability to divide the image flux into point measurements: the intensity so recorded must tell us something about microscopical **structure**. In order for an image to be perceivable by the human eye and mind, the array of point images must display **contrast**. Something about the specimen must produce changes in the intensity recorded at different image points. At its simplest, transmission contrast may be due to structures that are partially or fully opaque. More often in biology, structural features merely affect the phase of the light passing through them, or become self-luminous under fluorescent excitation. No matter what the mechanism, no contrast, no image. And the amount of contrast present in the image determines the accuracy with which we must know the intensity value at each pixel.

Contrast can be defined in many ways but usually it involves a measure of the variation of image signal intensity divided by its average value:

$$C = \frac{\Delta I}{\bar{I}}$$

Contrast is just as essential to the production of an image as "resolution." Indeed, the two concepts can only be thought of in terms of each other. They are linked by a concept called the contrast transfer function (CTF), an example of which is shown in Figure 4.3.

The CTF (or power spectrum) is the most fundamental and useful measure for characterizing the information transmission capability of any optical imaging system. Quite simply, it is a graph that plots the contrast that features produce in the image as a function of their size, or rather of the inverse of their size: their spatial frequency. Periodic features spaced 1 mm apart can also be thought of as having a spatial frequency of 1000 periods/m, or 1 period/mm or 1/1000 of a period/μm. Although we don't often view periodic objects in biological microscopy (diatom frustules, bluebird feathers or butterfly wing scales might be exceptions), any image can be thought of not just as an array of points having different intensities, but also as a collection of "spacings" and orientations.

---

[1] It is usually limited even more severely by the presence of aberrations.

James B. Pawley • University of Wisconsin, Madison, Wisconsin 53706

The geometry of the beam scannning the specimen

seen as pixels

or, as seen after re-oversampling, then smoothing

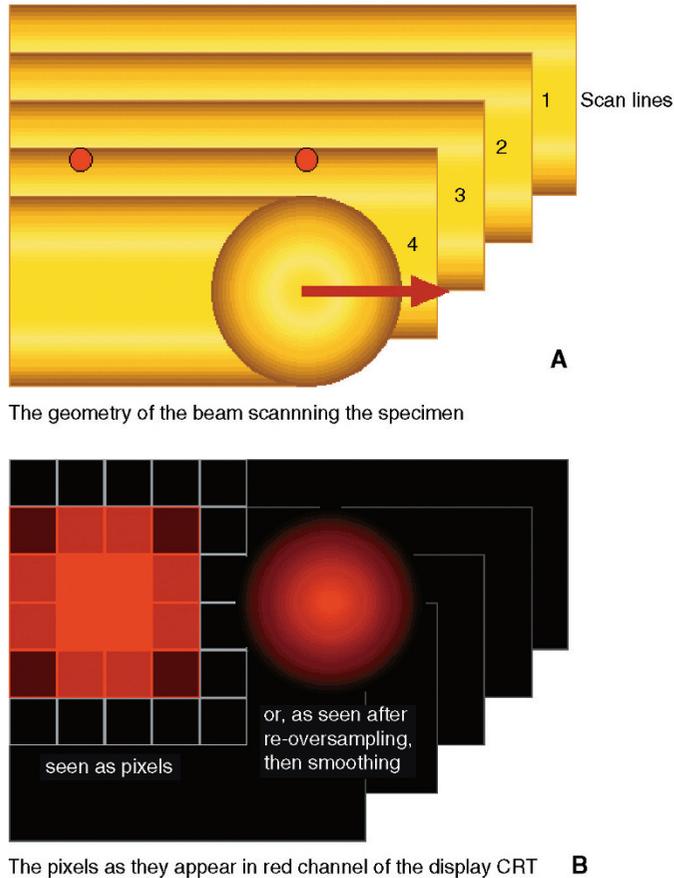The pixels as they appear in red channel of the display CRT **B**

FIGURE 4.1. What Nyquist sampling really means: the smallest feature should be at least 4 pixels wide. In (A), a yellow beam scans over two red point features. Because the "resolution" of the micrograph is defined, by the Abbe/Rayleigh criterion, as the distance from the center to the edge of the beam while Nyquist sampling says that pixels should be one half this size, pixels are one quarter of the beam diameter. From this, it follows that, at the end of each line, the beam will move down the raster only 25% of its diameter (i.e., it will scan over each feature at least four times). In (B) we see how such a signal will be displayed as a "blocky" blob on the screen, about 4 pixels in diameter. Because our eyes are designed to concentrate on the edges of each pixel, the screen representation doesn't look like an Airy disk (and would look even worse were we to add the normal amount of Poisson noise). We can get an "accurate" impression of the image of a point object only if we resample the 4 × 4 array into a much larger array and apportion the detected signal among these smaller, less distinct, pixels to form an image that looks like the circular blob on the lower right.

An image of a preparation of circular nuclei 7 μm in diameter has spacings of all possible orientations that are equal to the diameter of the nuclei in micrometers. The inverse of this diameter, in features/μm, would be the spatial frequency of the nuclei (in this case, about 150/mm).

The intensity of the CTF at zero spatial frequency is a measure of the average brightness of the entire image. The CTF graphs the image contrast assuming that the object itself has 100% contrast (i.e., that it is composed of alternating black and white bars having a variety of different periodicities; as few biological specimens have contrast this high, contrast in microscope images will be correspondingly lower). Because of the limitations imposed by diffraction, the contrast of the widest bars (spatial frequency near zero) will be almost 100% while bars that are closer together (i.e., have a spatial frequency nearer the diffraction limit) will be recorded with lower contrast in the image.
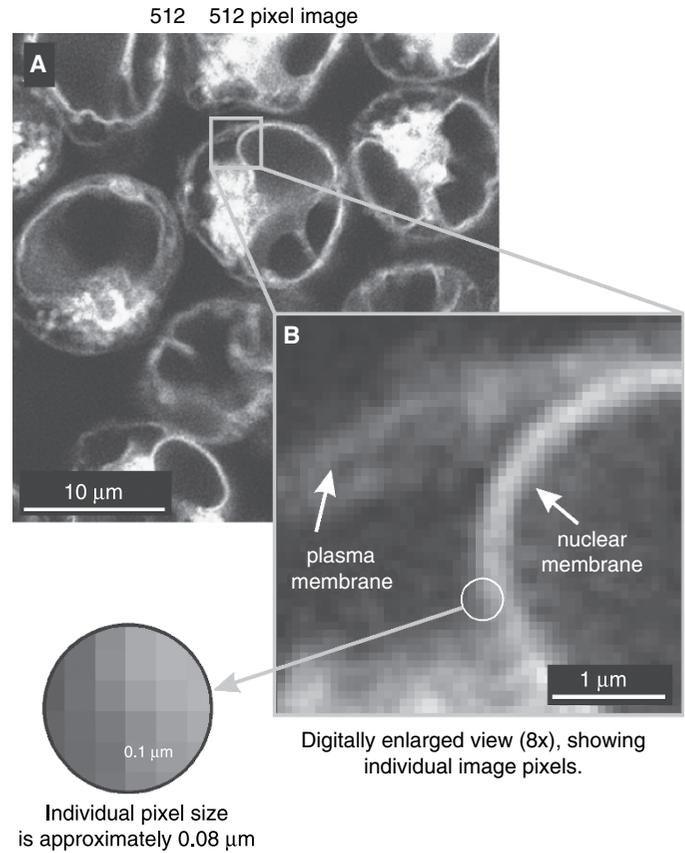
512   512 pixel image

FIGURE 4.2. A properly sampled 2D image. When your image is recorded with Nyquist-sized pixels then the smallest features will be 4 to 5 pixels across. (This figure kindly provided by Dr. Alan Hibbs of BioCon, Melbourne, Australia.)
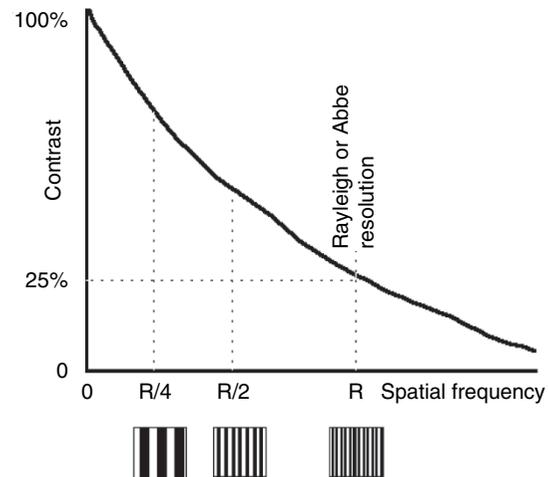

FIGURE 4.3. Contrast transfer function (CTF). This graph relates how the contrast of a feature in the image is inversely related to its size. Smaller "spacings" (see boxes below graph) have higher "spatial frequency" and will appear in the image with much lower contrast than they had in the object. Although the Rayleigh/Abbe resolution is conventionally set at the point where the CTF has dropped to 25%, even features that are twice this large (i.e., have one half the spatial frequency, R/2) are still represented in the image with only about half of their original contrast.

From Figure 4.3, one can see that the Rayleigh-criterion resolution is not really a hard-and-fast resolution limit but merely the spatial frequency at which the CTF of the optical system has dropped to about 25%. In general, features twice as big as the Rayleigh limit (i.e., R/2, half the spatial frequency) will be transmitted with a bit less than twice this contrast (i.e., ~50%), and so on for progressively larger features (although the image contrast can never be more than 100%).

One of the reasons that the CTF is such a useful guide to optical performance is that it emphasizes the performance for imaging small features. If we assume for a moment that we are using a high numerical aperture (NA) objective (NA 1.4) producing a Rayleigh resolution (R, in a microscope, this is often called the Abbe limit) of ~0.25 µm, then the part of the graph to the left of the R/4 marking describes the way that the optical system will transmit all the features larger than 1.0 µm (or R/4). All of the plot to the right of the R/4 mark refers to its transmission of features smaller than 1.0 µm. This is the part of the plot where problems are likely to occur. In addition, it reminds us that diffraction affects the appearance of features that are larger than the Abbe limit. In the end, resolution can only be defined in terms of contrast. It is NOT the case that everything works perfectly up to the Abbe limit and then nothing works at all.

The reason that the CTF is particularly useful in microscopy is that, if everything goes right (i.e., proper illumination, optically uniform specimen, no lens aberrations), its shape is entirely determined by the process of diffraction. If this is true, then the curve is directly analogous to what we can see in the back-focal plane (BFP) of the objective lens. You may recall that, when illuminated by axial illumination, large features (which have low spatial frequencies) diffract light near the axis while smaller features diffract light at larger angles. If you imagine that the left axis of the CTF plot (zero spatial frequency) is located at the exact center of the BFP, then the sloping part of the CFT curve can be thought of as representing a radial plot of the light intensity passing through the rest of the BFP.[2]

Light passing near the axis has been diffracted by large features. As many diffraction orders from these features will be accepted by the NA of the objective, they will be represented in the image with high contrast (Fig. 4.4).[3] Light out at the edge of the BFP consists of high-order diffraction from large features plus low-order diffraction from smaller features. The smallest features visible at this NA will diffract light at an angle that is almost equal to the NA of the objective, as defined by the outer border of the BFP. As only one diffraction order from these features will be accepted by the objective, the features that diffract at this angle will be represented in the image with low contrast.

As a result, one can "see" important aspects of the CTF, simply by viewing the BFP, for example, using a phase telescope or Bertrand lens. For example, when using a phase lens for fluorescent imaging, the phase ring present in the BFP of the objective partially obscures (50%–90% opacity) and shifts the phase of any rays passing through it. Therefore, features in the object that are

the correct size to diffract at the angles obscured by the ring will be less well represented in the image data recorded.

Finally, the CTF is useful because it is universal. Assuming that you normalize the spatial frequency axis of the CTF plot in Figure 4.3 for the NA and λ in use (i.e., the spatial frequency under the 25% contrast point on the curve should be the reciprocal of the Abbe resolution), it is a reasonable approximation of the CTF of any diffraction-limited optical system. As such it defines the best we can hope for in terms of direct imaging (i.e., without non-linear image processing such as deconvolution to be discussed later, or the use of clever tricks like STED as discussed in Chapter 31, *this volume*).

The CTF can be used to characterize the performance of every part of the imaging system: not only the optical system but also the image detector (film or video camera), the image storage system (film or digital storage), the system used to display or make hardcopy of the stored result, even the performance of your eyes/glasses!

The performance of the entire imaging chain is merely the product of the CTF curves defining all the individual processes. Because the CTF always drops at higher spatial frequencies, the CTF of an image having passed two processes will always be lower than that for either process by itself (Fig. 4.5). In other words, small features that have low contrast become even less apparent as they pass through each successive stage from structures in the object to an image on the retina of the viewer.

As can be seen from Figure 4.5, the steps with the lowest CTF are usually the objective and the video camera. A digital CCD camera (i.e., a CCD camera in which each square pixel reads out directly into a specific memory location) would produce better results than the video-rate television camera/digitizer combination shown in Figure 4.5 because the latter digitizes the data twice, a process that can reduce the contrast of fine, vertical lines that are sampled in the horizontal direction by a factor of 2. The performance of all of the steps past the ocular can be "improved" by
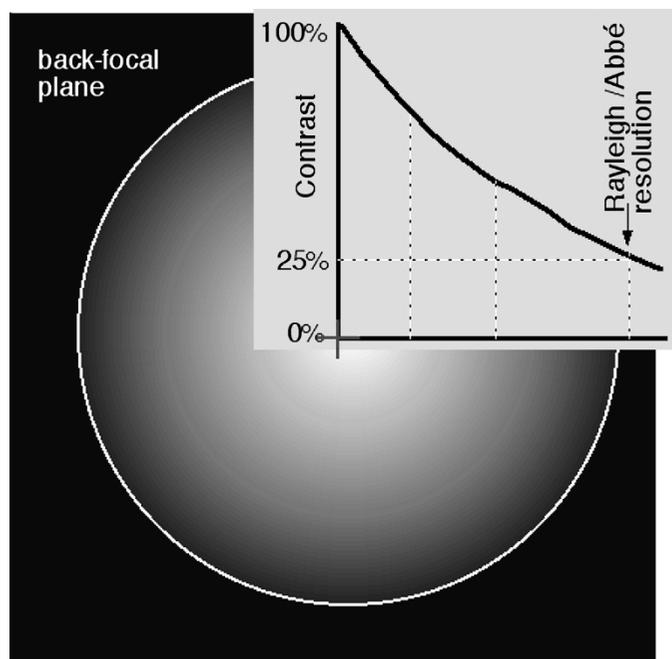

FIGURE 4.4. Relationship between the CTF and the position in the back-focal plane of the objective lens that axial light will diffract from features of different spatial frequencies.

---

[2] It is uncommon to image using only axial illumination, at least in part because filling the condenser BFP increases the number of diffraction orders that can pass through the objective, thereby doubling the resolution. It is assumed here only for illustrative purposes.

[3] Strictly speaking, the following analysis is only accurate for axial illumination. However, even for the convergent illumination used to get the highest resolution in transmission imaging, the general point is correct: light rays carrying information about smaller features are more likely to be represented by rays that pass near the edges of the back-focal plane.
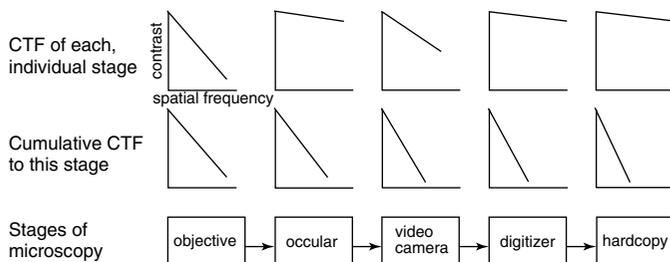
FIGURE 4.5. The CTF of each component in the microscope system affects the final result. Every optical component and digital process can be characterized by a CTF. The effects of a series of steps can be determined by multiplying together the CTFs of all the steps.

working at higher magnification: if the pattern of light presented to the camera (or eye) contains larger features, their contrast will be reduced less by imperfections in the camera (or eye) itself. However, this approach also has limitations. Working at higher magnification requires either a larger image sensor or a smaller field of view. Much of the remainder of this chapter is concerned with making the most appropriate choice of "magnification," although the discussion is usually in terms of "How large should a pixel be, referred to the object?"

Once the information is digitally encoded, further CTF degradation can be minimized as long as certain rules are obeyed (as discussed below and in Chapter 48, *this volume*).

The lessons so far are

- No matter how high the contrast of the optical process defining a feature in the object, smaller features are always depicted in the final image with less contrast than larger features.
- Features that have low intrinsic contrast in the object will have even lower contrast in the image.

On the other hand, remember that Figure 4.3 shows the best for which we can hope. It is not at all hard to end up with system performance that is substantially (~50%) worse than that described by Figure 4.3. This means that while one can no longer see the smallest features, one now might just as well use larger pixels.

In this chapter, we will assume that Figure 4.3 really does describe optical system performance, and go on to consider the other factors important to ensure that image data is digitally recorded in an optimal manner.

DIGITIZATION AND PIXELS

Image digitization refers to the process whereby an apparently continuous analog image is recorded as discrete intensity values at equally spaced locations on an *xy*-grid over the image field. This grid is called a raster.

Typically the image area is divided into an array of rows and columns in much the same way as a television image. In North and South America and Japan, the television image is composed of 483 lines covering a rectangular area having proportions that are 3 units high by 4 units wide. If each line in such an image is divided into about 640 equal picture elements or pixels, then each pixel will be square if you discard three lines and record a raster of 640 × 480 pixels.

Newer computer-based CCD image digitization systems do not rely on any broadcast television standard, and are more likely to

use rasters of 512 × 512 or 1024 × 1024 pixels, although other dimensions are not uncommon. In scientific imaging, it is advisable to avoid digitizing schemes involving pixels that do not represent **square** subunits of the image plane (for example, those produced by digitizing each line from a television image into only 512 pixels rather than 640 pixels) as there is little support for displaying or printing such images directly.

Digitization of Images

The actual process by which the signal from the image detector is converted into the intensity values stored in the computer memory for each pixel depends on the type of microscope involved.

CCD cameras: Typically, a widefield or disk-scanning confocal microscope uses a camera incorporating a CCD image sensor. Although we will not describe in detail the operation of these sensors (see Chapter 12 and Appendix 3, *this volume*), the camera operates by reading out a voltage proportional to the number of photons absorbed within a small square area of the sensor surface during the exposure time. As long as the intensity value readout is stored directly into the computer, this small area on the CCD defines the pixel size for the remainder of the imaging system.[4]

As far as the user is concerned, the most important parameters involved in attaching the CCD to the microscope are the NA of the objective, the wavelength, and the total magnification up to the surface of the sensor. Together these parameters determine both the proper size of a pixel referred to the plane imaged in the specimen, and also the optimal pixel size for the CCD. For example, if a CCD camera with $8 \times 8\,\mu m$ pixels is coupled to a microscope with a 40×1.3 NA objective via a 1× coupling tube, each sensor pixel will cover $8/40 = 0.2\,\mu m$. The same camera and coupling will produce "$0.08\,\mu m$ pixels" when used with a 100× objective, but the number of photons striking each pixel during a given exposure time will now be $2.5 \times 2.5 = 6.25\times$ less because signal intensity goes down with the square of the magnification.

Photomultiplier tubes (PMTs): On a laser confocal microscope, signal photons strike the photocathode of a PMT where some small fraction of them each produce a single photoelectron (PE). These PE are then amplified about a million times by charge multiplication. The signal current emerging from the PMT is digitized under the control of a pixel clock which also controls how the scanning mirrors sweep over a rectangular raster on the specimen. This clock divides the time taken to scan one line into the appropriate number of intervals, so that each time interval represents a square area of the image (i.e., each time interval represents the same distance along the scan line as the spacing between adjacent lines). As the PMT signal is digitized for each interval, or pixel, the pixel value represents the signal intensity of a small square area of the final image.

Because the shape of the raster in a laser confocal microscope is defined by the size of the electronic signals sent to the scan mirrors (Fig. 4.6) rather than by the fixed array of electrodes on the surface of the CCD, there is much more flexibility in terms of the size and shape of the rasters that can be scanned.

---

[4] This is not true if the CCD is read out to form an analog "composite video" signal which is then redigitized into the computer. Such uncorrelated redigitization can reduce the effective horizontal resolution of the data by almost a factor of 2 and should be avoided. Likewise, one should be careful when "resizing" images using image processing programs because, unless it is done in integer multiples, this process also involves resampling, a process that reduces image contrast.

FIGURE 4.6. Mirror scan angle and magnification. The galvanometer mirror scans the laser beam across the focus plane of the objective by effectively changing the angle at which the laser beam passes through the back-focal point of the objective lens. A larger deflection of the mirror scans the light over a longer line on the specimen (B). As the data from this longer line are finally displayed on the same sized computer monitor, the effect is to lower the overall magnification of the image. If the number of pixels digitized along each line remains constant, a longer line on the specimen implies larger pixels.
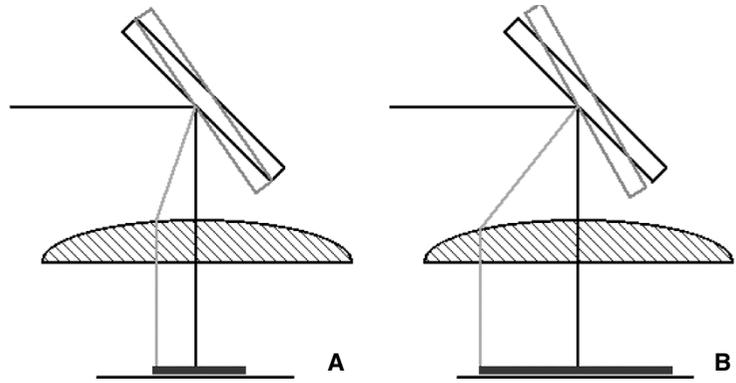
In particular, a combination of the magnification of the objective and the zoom magnification on the scan control panel defines the dimensions of the raster at the object plane in the specimen. If more current is sent to the scanning mirrors (low zoom magnification), they will drive the scanning beam over a larger area of the specimen and, assuming a fixed raster size (e.g., 512×512 pixels), this means that each pixel will now represent a larger area of the specimen (Fig. 4.7, darkest square). Conversely, higher zoom magnification will send smaller currents to the scan mirrors. This will make the raster scan over a smaller area on the specimen, and make the area represented by a single pixel proportionally smaller (Fig. 4.7, lightest square). As a result, and unlike the CCD case, pixel size is under continuous control as the user changes raster shape/size and zoom magnification settings. However, your control panel should constantly display the current pixel dimensions.

HOW BIG SHOULD A PIXEL BE? SAMPLING AND QUANTUM NOISE

Clearly, it is not possible to represent features spaced, say, $1\,\mu m$ apart if the pixel dimensions are $2 \times 2\,\mu m$. Having smaller pixels will increase the chance that small features of the specimen are adequately sampled. However, having smaller pixels also has disadvantages. It means either imaging a smaller area of the specimen or using a larger raster size [1024×1024 rather than 512×512; Fig. 4.8(A)]. If you choose a larger raster, you must store and analyze more data. You must also either collect fewer signal photons from each pixel [Fig. 4.8(B)] or take longer to scan the larger raster. Longer counts require you to expose the specimen to more light [Fig. 4.8(C)], a process that may be deleterious, especially to living specimens.

Settling for less signal in each pixel is also not without problems. The signal that is being counted is not continuous but is composed of photons, sometimes quite small numbers of photons. In fact, it is not uncommon for the signal from a single pixel in the bright area of a fluorescent confocal specimen to represent the detection of only 9 to 16 photons.

As the detection of a photon is a quantum mechanical event, there is an intrinsic uncertainty in the number actually detected on any given trial. This uncertainty is referred to as Poisson, or statistical, noise and is equal to the square root of the number of events (photons) detected. Therefore, reading 16 photons really means detecting 16 ± 4 events.[5] **Like diffraction, Poisson noise is a rigid physical limitation**. The **only** way to reduce the relative uncertainty that it causes is to count more events.

If we increase the zoom magnification by a factor of 2, there will be $4\times$ as many pixels covering any given scanned area of a two-dimensional (2D) specimen. If, at the same time, we also reduce the laser power by a factor of 4, the same total amount of signal/μm^2 will emerge from the reduced area now being scanned, producing the same bleaching or other photo damage but the average signal level in each bright pixel will now be not 16 photons, but only 4 ± 2 photons. The uncertainty of each measurement is now 50%. In other words, when photons are scarce, one seldom wants to use pixels smaller than are absolutely necessary to record the information in the image.

It is simply a case of "winning on the swings what you lose on the roundabouts." Either scenario has advantages and dis-
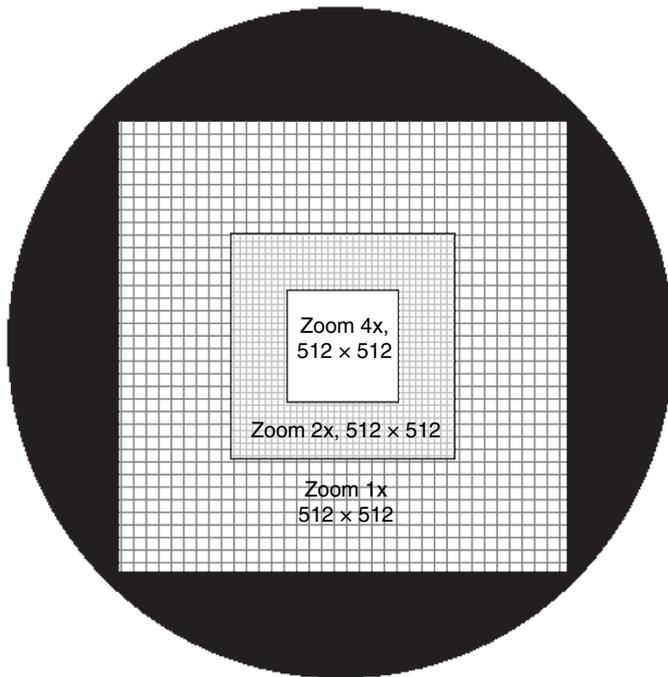

FIGURE 4.7. Relationship between zoom setting and area scanned on the specimen. A higher zoom magnification setting scans the beam over a smaller area of the sample. As each pixel now represents a smaller area on the specimen, we say that the pixels are "smaller." The important thing is to adjust the zoom magnification so that the pixel size is about 50% of the Abbe resolution for the NA and wavelength in use.

Zoom 4x, 512 × 512

Zoom 2x, 512 × 512

Zoom 1x 512 × 512

---

[5] That is, 67% of a series of measurements of this intensity would be in the range of 12 to 20 photons and 33% of such measurements will be outside even this range. In other words, if you detect 10 photons you really have very little idea about what the signal intensity really "should have been."
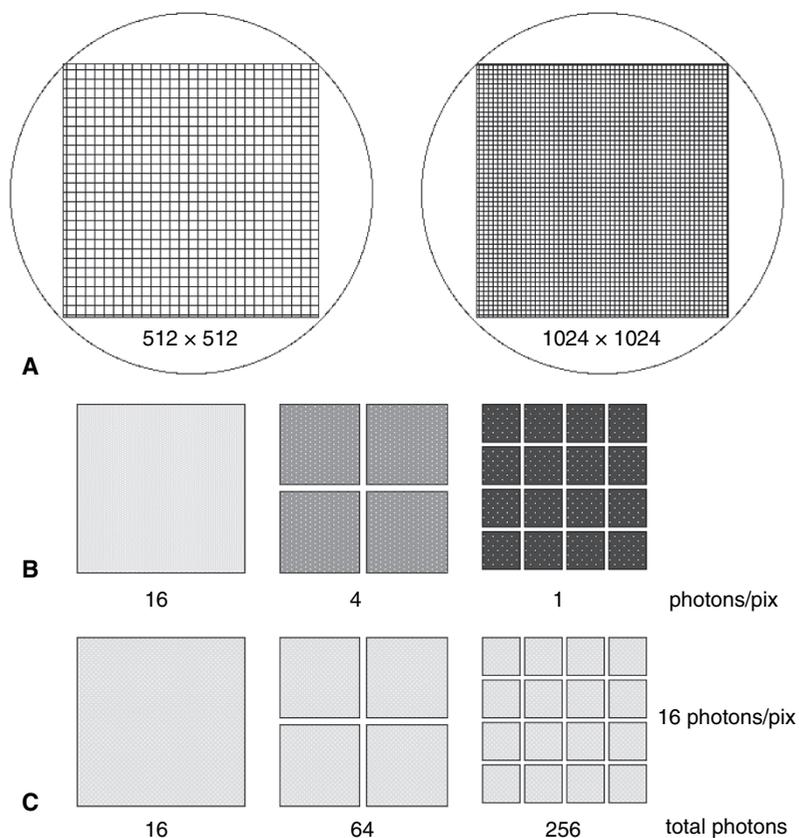
FIGURE 4.8. Relationship between raster size, pixel size, and "light dose." (A) If a microscope scanning at a zoom setting of 1 is switched from a 512 × 512 raster to one of 1024 × 1024 pixels, the dimensions of each pixel on the specimen will drop by 50%. (B) If the same amount of signal is split between more pixels, the signal level from each one goes down (and the Poisson noise goes up), but if the beam scans more slowly or is made more intense so that the same amount of signal is still collected from each pixel, (C), then the amount of damage/pixel increases. There is no free lunch!

advantages. Surely there must be a "best" strategy for setting the zoom correctly to produce the best pixel size.

Fortunately there is!

THE NYQUIST CRITERION

It was not until 1929 that Harry Nyquist, who worked for the telegraph company, gave much thought to the optimal strategy for digitally sampling an analog signal (Nyquist, 1928). When such sampling first became technically possible, the signal in question was electronic, perhaps the audio signal of a radio program. The process envisioned, as diagramed in Figure 4.9, requires six components: a **pre-amplifier** feeding the signal to the **analog-to-digital converter** (ADC), a **digital memory** system for storing the digital data from the ADC, **a digital-to-analog converter** (DAC) that reassembles the digital information into a continuous analog signal that can be passed to the **output amplifier,** and, finally, a **clock** to synchronize the whole process. The clock determines the time interval between samples (i.e., the sampling frequency, in samples/s).

The information content of any electronic signal is limited by the electronic bandwidth of the amplifier used to transmit it.[6] In

1949, Claude Shannon was able to prove Nyquist's theorem and show that, if the interval between the intensity measurements is less than half the period of the highest frequency in the signal, it will then be possible to faithfully reconstruct the original signal from the digital values recorded (Shannon, 1949). **The Shannon sampling frequency,** which is the inverse of the Shannon sampling interval, is also known as the **Nyquist frequency, especially in the imaging community**.

It is often forgotten that there is a second part of the Shannon/Nyquist theorem: the part about reconstructing the original data. The theorem states that the output "amplifier" through which you play back the reconstructed signal from the DAC must have the same bandwidth as the pre-amplifier that originally fed the signal to the ADC. This is an important condition, one that is often not satisfied in current confocal microscopes unless their images are deconvolved before presentation (as will be discussed later).

Attempting to apply Nyquist sampling to 2D or three-dimensional (3D) images gives rise to the question: How do we measure the "bandwidth" of the "amplifiers" when faced with the problem of digitizing 2D or 3D microscopical image data?

Electronic bandwidth is not a simple concept. The frequency response of any real amplifier does not remain flat until some frequency and then go abruptly to zero at any higher frequency. Rather, limitations imposed by the components of which the circuit is made cause its power response to decrease gradually as the frequency increases, usually dropping to one half or one quarter the original output power as the frequency goes up each octave above

---

[6] Think of this as the frequency response of your stereo system. Good high frequency response will let you hear your music more accurately. The frequency response of your stereo is usually plotted in decibels (a measure of relative power) on the *y*-axis against the log of the frequency on the *x*-axis. Note the similarities to Figure 4.1.

FIGURE 4.9. The components needed to digitize and reconstruct an analog signal. The "post-amp" is essential to remove the "single-pixel noise" that is added to the original analog signal by Poisson statistics. Because real, Nyquist-sampled data can never have features smaller than 4 pixels across, single-pixel noise can be removed by limiting the bandwidth of the post-amplifier. In microscopy, this limiting function is implemented by either Gaussian filtering or deconvolving the raw 3D data.
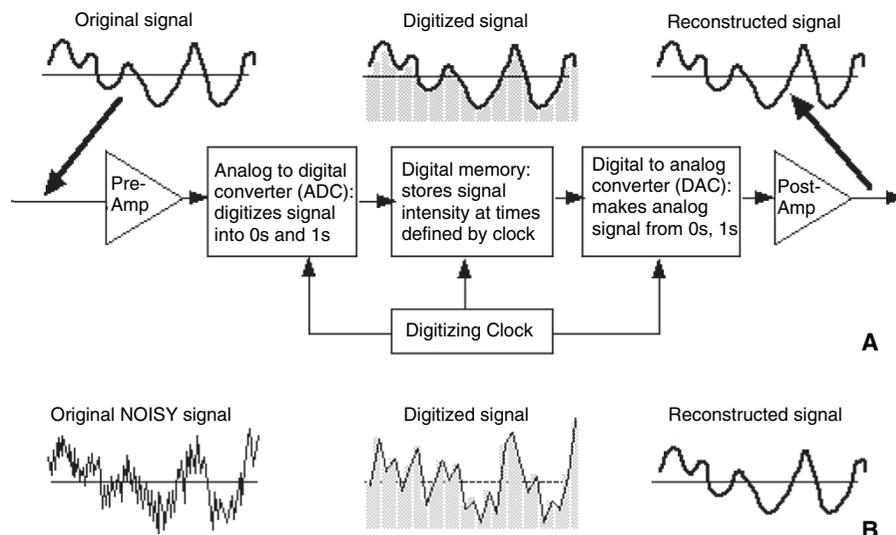
some "cut-off frequency."[7] As in optical systems, higher electronic frequencies are still transmitted, but at lower intensity. In electronics, the convention is to define the bandwidth by the frequency at which the power response drops to 50% of the linear response, a frequency called the "3 dB point." This defines the bandwidth Shannon used. In optical terms, we usually think of the image being useful until it drops to about 25% of its peak contrast (i.e., the Abbe criterion noted above), although this too is an arbitrary choice.

If we think of an analog electronic signal as a one-dimensional image, it is not hard to think of an image as a 2D (or 3D) version. Except that image data varies in space rather than time, the rest of the analysis applies. The "bandwidth" of an image must be somehow related to its "sharpness," and this is related to the highest spatial frequencies it contains.

Now if we were applying this analysis to the CCD sensor used in a consumer snapshot camera, we would have a problem. Although the "world" out there may be composed of objects of every size, we really have little knowledge of the CTF of the lens, let alone whether or not it is focused correctly or whether you are capable of holding it motionless during the exposure period. As a result, we really don't know the bandwidth of the data and consequently we don't know whether or not the pixels are small enough to meet the Nyquist criterion. "More is better" is the slogan that sells.

Fortunately, this is not the case in microscopy. Here we do know that, at the very least, diffraction limits the maximum sharpness of the data that can be recorded, and that the "spatial frequency response" of the microscope can be defined by a suitably calibrated version of Figure 4.3.

Therefore, the convention is to choose the size of the pixel to be equal to one half of the Abbe criterion resolution of the optical system.

There are some caveats. The structural features of a 1D image can only vary in that dimension. The structural features of a 2D image can vary in more than two possible directions. Although signals defining features such as vertical or horizontal lines, vary

in only the x- or y-directions, respectively, what about a set of lines oriented at 45° to these axes? It would seem that sampling points along a 45° line would be spaced 1.41× as far apart as sampling points along features that vary along the x- or y-axes. Pixels just small enough to sample a given small spacing when it is oriented vertically or horizontally would be 1.41× too big to sample this same structure were it to be oriented at 45°. However, this analysis neglects the fact that all image "features" extend in 2D. As a result, lines running at 45° will also be sampled by other pixels in the array and if we count all the pixels that sample the blurred features along a line at 45°, one finds that the sampling interval isn't 1.41× larger but in fact only 0.707 as large as the sampling interval in the x- or y-directions (Fig. 4.10). Clearly we want to be able to see structures oriented in any direction. To be on the safe side, it may be better to use pixels ~2.8× smaller than the finest spacing you expect to record in your image.[8]

Estimating the Expected Resolution of an Image

Assuming that the shape of the CTF curve describing the optics of the microscope depends only on the NA and the wavelength, it is also a plot of power level versus the logarithm of the frequency, just like the frequency response curve of a stereo. Although the CTF defines the best that one can hope for, it does not guarantee it.

Performance can be worse, and if, in fact, it is worse, does it make sense to use smaller pixels than we need?

Let us take some concrete examples. The calculation of the Abbe criterion resolution assumes that **two point objects of similar intensity** are represented in the image as Airy disks, spaced so that the peak of each is located over the first dark ring of the other. If we sum the light intensity of these two Airy disks, there will be a valley between the two peaks in the summed image. At the exact mathematical bottom of this valley, the intensity is

---

[7] As in music, an octave represents a factor of 2 in signal frequency.

[8] A similar line of argument could be used to suggest that one use even smaller pixels when sampling 3D data because the diagonal of a cube is 1.732x longer than its side. However, we will soon see that, as the z-resolution of the confocal microscope is always at least 3x lower than the xy-resolution, ignoring this factor does not cause any problem in practice.
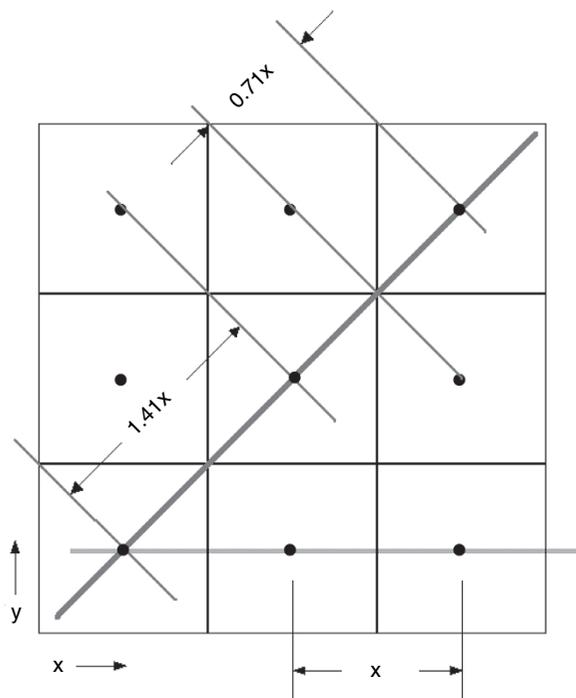
FIGURE 4.10. Spatial frequency and geometry. The 3 × 3 array of squares represents a small raster and the dots in the center of each represent the sampling points. Although one might be tempted to think that these sampling points would be too far apart along the diagonal to be able to properly sample any signal that just meets the Nyquist sampling criterion when oriented either horizontally or vertically, this is not so because the sampling points of the adjacent diagonal rows of pixels actually sample at 0.71 of the *x*- or *y*-raster pitch.

about 25% lower than the intensity of either peak. This is the basis of the idea that 25% contrast is equal to the Abbe criterion resolution (Fig. 4.11).

Under these circumstances, the smallest resolvable spacing is defined as the distance between the center of an Airy disk and the center of its first dark ring. To be properly sampled, pixels should be less than one half this distance in size.[9]

Suppose that, along a line joining centers of the images of the two points, one pixel just happens to be centered on the brightest part of one Airy disk. The adjacent pixel would then be centered over the valley between the peaks and the third pixel will be over the second Airy peak. If we sample the brightness at the center of these three pixels, the digital data will reflect the trough in intensity between them.

On the other hand, if the "valley pixel" has a value proportional not to the intensity at the exact center of the pixel but to the average intensity over the whole pixel,[10] the value stored for the center pixel will be much more than 75% of the peak intensity:

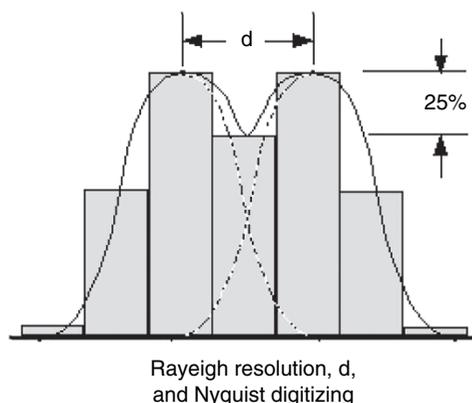Rayeigh resolution, d,
and Nyquist digitizing

FIGURE 4.11. Nyquist sampling of an image of two points separated by the Rayleigh resolution.

that is, the contrast recorded between the three pixels will now be much lower than 25% (Fig. 4.12).

If the two features that produced the two Airy disk images are not of equal brightness (surely the more likely occurrence) then the contrast along a line joining the peaks will again be much less than 25%.

Worse still, what if the peaks are uncooperative and are not squarely centered on two pixels, nicely spaced on either side of the central, darker pixel? If the value recorded at each pixel is the average of the intensity across the pixel, the contrast along a line between the features can be substantially reduced or even eliminated (Fig. 4.13).

Now it is fair to say that while these considerations are problems, to some extent, they only represent a serious problem if we ignore the second part of the Nyquist sampling theorem, the part having to do with reconstruction. If the image is properly reconstructed (deconvolved), in most cases, information from adjoining pixels (those in the rows in front or behind the printed page in Fig. 4.13) will allow one to smooth the image to form a good estimate of the structure of the original object as is discussed later in the chapter.[11]

Deconvolving or filtering the image data eliminates high spatial frequencies. Effectively, such filtering causes the signal to overshoot the contrast present in the digital signal. This process substantially reverses the apparent reduction in contrast that occurs on digitization.

---

[9] Or perhaps a bit less if we use the 2.3 or 2.8 samples/resolvable element (resel) suggested above. For simplicity, I will stick to 2 samples/resel in this discussion, because, as discussed below, in the case of the fluorescent images of most interest, lack of signal usually prevents one from realizing Abbe criterion resolution and consequently the "actual" resolution is lower than Abbe and using somewhat fewer/larger pixels is appropriate.

[10] In microscopy terms, the CCD samples the average value of a pixel while the ADC sampling the PMT signal in most single-beam confocals acts more as the center-sampling device.

---

[11] Periodic structures having a size near the resolution limit also present sampling problems. Suppose that the object is a sinusoidally varying structure with a period equal to the Abbe spacing. If the two samples required by Nyquist coincide with the plus and minus excursions of the sine wave, then we will have some measure of its magnitude and the position of its zero-crossing [Fig. 4.5(B)]. However, if the two samples happen to be taken as the sine wave crosses its origin, all the sampled values will be zero and hence can contain no information about the sine wave [Fig. 4.5(C)]. This apparent exception to Nyquist sampling success is not actually an exception in terms of the original application of the theorem: information theory. According to information theory, a sine wave contains no information beyond its frequency and magnitude. As long as you have slightly more than two samples/period, the sampling intervals will "beat" with the data to create a sort of moiré effect, from which one can estimate the magnitude and period of the sine-wave object. All this does not change the fact that an image of a periodic object must be at least 2x over-sampled if it is to be recorded with reasonable fidelity [Fig. 4.5(A)]. This is particularly important when imaging the regular patterns found on resolution test targets.

Nyquist digitizing of two point objects at Abbe separation.

FIGURE 4.12. Two methods of sampling: at the center point and as the average value.

Center sample   A

Average sample   B

25%

12%

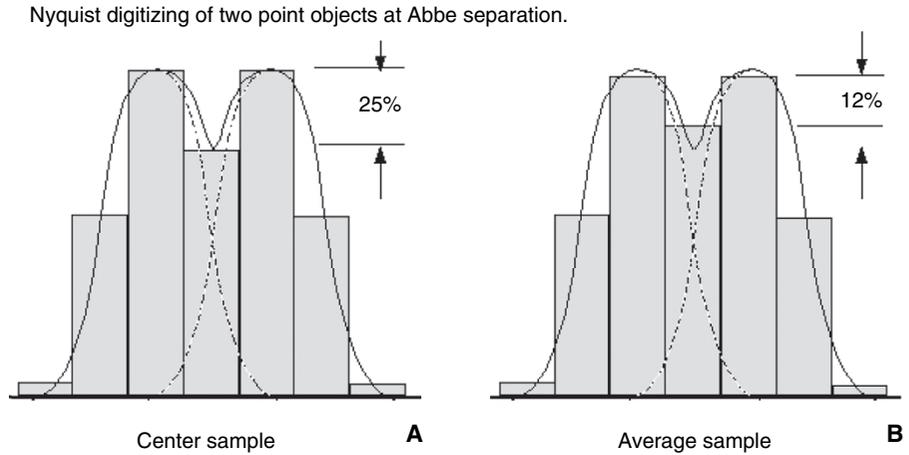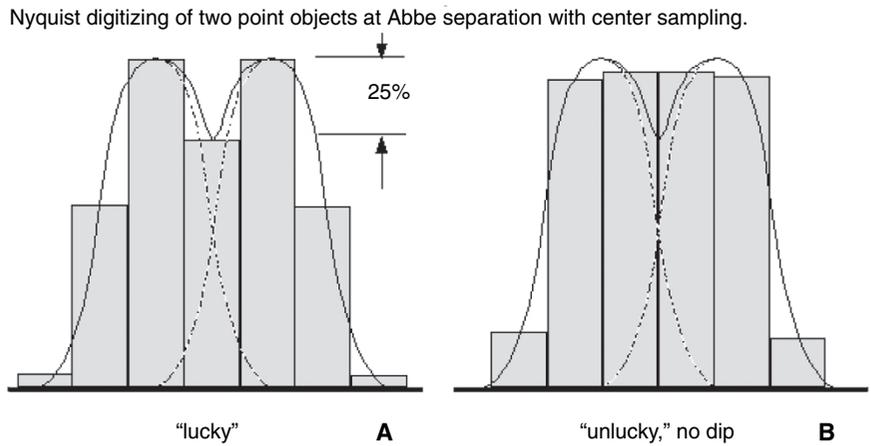Nyquist digitizing of two point objects at Abbe separation with center sampling.

FIGURE 4.13. How can this possibly work? "Lucky" and "unlucky" Nyquist sampling of the image of two points separated by one Rayleigh resolution.

25%

"lucky"   A

"unlucky," no dip   B

On the other hand, this reduction in contrast is entirely relevant if one tries to assess raw digital image data from a Nyquist-sampled confocal microscope directly from the cathode-ray tube (CRT) or liquid-crystal display (LCD) screen or when viewing unprocessed data as a hardcopy from a digital printer.

There is another problem that even proper reconstruction will not solve. Recall the example above in which a "bright" signal (perhaps the peak of an Airy disk?) was only 16 ± 4 photons. Clearly the ± 4 represents a 25% average error, that is the same order of uncertainty as the maximum expected contrast we hope to see between the peaks (Fig. 4.14). In other words, even though diffraction theory says that we should record a lower signal in the

pixel between two peaks of equal intensity, at these low signal levels, Poisson statistics says that, about 30% of the time, the intervening pixel will actually be measured as brighter than at least one of the two peaks. [As each peak pixel is subject to its own independent statistical variations, in a given image, it is unlikely that all 3 pixels (or 9 pixels if we consider the 2D image) will be recorded as the same brightness.]

Artifactual "features" such as those diagramed in Figure 4.14(B) and produced by "single-pixel" Poisson noise, will be removed if the dataset is deconvolved or even 3D-Gaussian smoothed as discussed below.

FIGURE 4.14. The effect of Poisson noise. While a Nyquist-sampled signal of Rayleigh-separated features seems to work well when the signal is composed of many photons and has little noise (A), when the number of photons counted drops by a factor of 100, and the signal-to-noise ratio (S/N) drops by a factor of 10, then random variations in the signal can play havoc with the data (B) allowing "single-pixel" noise features to masquerade as very small features.

Abbe spacing

500

Many counts   A

5

Few counts   B
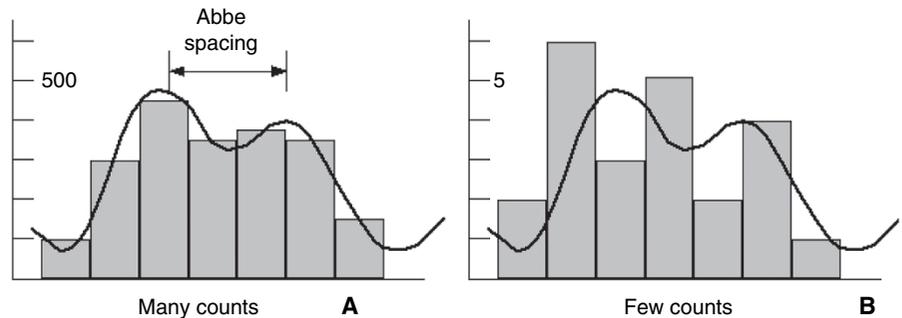
The Story So Far

Once we know the size of the smallest data we hope to record, we can adjust the zoom magnification on a confocal microscope or the CCD camera coupling tube magnification on a widefield microscope to make the pixels the right size.

But is Figure 4.3 really a good way to estimate this maximum spatial frequency?

REALITY CHECK?

Are we kidding ourselves in thinking we will be able to **see individual point features separated by Abbe criterion resolution** when viewing faint, fluorescent specimens? In fact, under these conditions, we may be lucky to separate features that are even twice this far apart and we now recognize that we could record such data using pixels that were twice as big and 4× less numerous (in a 2D image; 8× fewer in a 3D image).

On the other hand, our human ability to "see" (recognize?) extended features, such as fibers or membranes, is enhanced by the ability of our mind to extract structural information from noisy data. We do this "magic" by integrating our visual analysis over many more pixels (100×?). While viewing noisy, extended objects doesn't improve the quality of the data, it allows the mind the illusion of averaging out the statistical noise over more pixels because each is an independent measurement. In this case, Nyquist/Abbe sampling may be more worthwhile after all.

Is Over-Sampling Ever Wise?

Yes! When viewing a specimen that is not damaged by interacting with light, over-sampling can improve visibility by recording more data and hence reducing the effect of Poisson noise. Video-enhanced contrast microscopy has been utilized to image isolated features much smaller than the Abbe limit. When imaging structures such as isolated microtubules, one often employs "empty magnification," sampling much more finely than is required by Nyquist. This is effective because such structures produce only a very small amount of image contrast.

As a simplified example, assume that the signal from the feature is only 1% greater than that from the gray background. Turning the light signal into an electronic signal permits one to adjust the contrast arbitrarily. However, if the electronic signal is too noisy, the result will just be more contrasty noise.

To detect a 1% difference using photons, we must ensure that the contrast produced by Poisson noise variations in the background gray are less than that between the background and the feature. At the minimum, this involves counting at least 10,000 photons/pixel because the Poisson noise is $\sqrt{10,000}$ and $100/10,000 = 1\%$. One could produce an even more easily interpretable image if the intensity of the feature differs from the background by more than one standard deviation. Recording 100,000 photons/pixel would make the 1% signal become 3× more than the Poisson noise.

As most image sensors saturate (become non-linear) when exposed to more than 100,000 photons/pixel, the only way to "see" such a low contrast feature is to make many different measurements (i.e., use more pixels). A single pixel might be bright because of statistics but it is less likely that four adjacent pixels will all be recorded as bright. Using more pixels produces even greater visibility by further separating the signal representing the feature from that representing the background.[12]

Under-Sampling?

In some cases, the useful resolution of the image is set by non-optical limits. An example might be a fluorescence image of a cell containing a dye that changes its properties in response to the concentration of certain ions. If the diffusion of ions and dye molecules precludes the existence of small-scale variations in the fluorescence signal from such a cell (i.e., no small features), there is no need to divide the data into small pixels. Measuring each of fewer, larger pixels for a longer time may give more accurate results, especially when the expected changes in ion concentration produce only small changes in the fluorescent properties of the dye used (i.e., a low-contrast image) or when two noisy images must be ratioed to obtain the final result.

In such specimens, high **spatial** resolution is impossible because of diffusion, while high **intensity** resolution is required to make small changes visible. In this case, it is particularly important to spatially filter the raw digital data before attempting to display or ratio the data (see Chapter 42, *this volume*).

DIGITIZING TRADE-OFFS

We have now discussed how most of the relevant factors: pixel size, optical resolution, and photon signal strength all interact. The best choice will almost always depend primarily on the robustness of your sample: Assuming careful adjustment of the optics, more counted photons will always give a better estimate of the distribution of fluorescent molecules within the specimen.

You must decide when the need for better spatial or intensity resolution justifies increasing the signal level and when it cannot be tolerated because to do so would reduce the "biological reliability" of the data (i.e., kill or injure the cells, see Chapters 38 and 39, *this volume*). Data with higher spatial resolution may not be useful if they represent structural features of a cell that is dead or dying.

NYQUIST RECONSTRUCTION: "DECONVOLUTION LITE"

Elsewhere in this volume the technique for recording 3D data sets of both point objects and fluorescent specimens using a widefield microscope and a CCD camera and then computer-processing the resulting data to produce 3D images much like those produced by the confocal microscope are discussed in detail (Chapters 23, 24, and 25). The most advanced form of this processing is called iterative, constrained 3D deconvolution and uses the image of the point object to determine the 3D point-spread function (PSF) for the imaging system. Here, I will discuss only one part of this process, a process that can be thought of as filtering or smoothing.

---

[12] It is important to state here that I am not talking about limitations in the image that could be overcome by resetting the "contrast" and "brightness" of the image display system in order to make any image contrast more visible to the observer. These are assumed to be set in the best possible manner for the individual concerned. The limitation on visibility discussed here relates solely to the fact that the data in the recorded image is insufficiently precise for any observer (or even a computer!) to determine the presence or absence of the structure. For more about visibility and the Rose criterion, see Chapters 2 and 8, *this volume*.

As noted above, sampling the analog data to produce the digital record was only half of the process. The second part involves "passing the reconstructed signal through an amplifier having the same bandwidth as that from which the original data was received."

To see why this is necessary it may help if we imagine a reconstruction of the digital data as being sort of a bar graph, in which each bar represents the intensity value stored for this pixel [Fig. 4.15(A)]. Clearly a "signal" represented by the boxy contour line going along the tops of the bars will generally change much more abruptly than the original data. As a result, it is not a faithful reconstruction of the original signal.

How can it be made more similar? In terms of Fourier optics a "square-wave object," such as a bar, can be thought of as being composed of the sum of a number of sine-wave objects, each having a periodicity that is an integer-multiple (harmonic) of the square wave frequency. The first sine term in this series converts each "square" of the square wave into a rounded curve. As subsequent terms are added, they add the "ears" to the hump that make the sum resemble the original boxy square wave ever more accurately (Fig. 4.16).

If we apply this logic to the top line of our bar graph, we can think of it as the sum of a lot of sine waves. If we leave out the higher harmonic terms before reconstructing the original line, the boxy corners will be rounded. Passing the boxy reconstruction through an amplifier of limited bandwidth prevents the higher order terms (higher frequencies) in the sine-wave series from being included in the reconstructed signal [Fig. 4.15(C)].

This is important when viewing a digital image because our eye/brain system is designed to emphasize the sharp edges that define the boundary of each pixel on the liquid-crystal display (LCD) screen and this is more likely to happen when a single noisy pixel stands out from a darker background.

The same thing is true when we reconstruct an image from digital data. However, in the case of fluorescence or other low-intensity data, there is an additional complication. The Nyquist theorem assumes that the signal digitized is continuous, that is, that
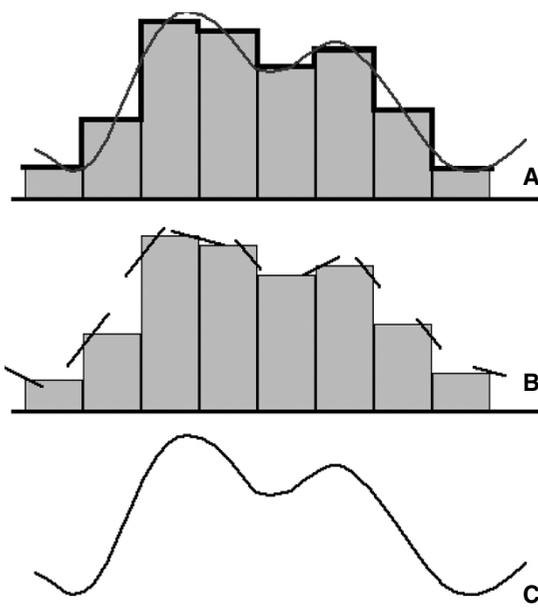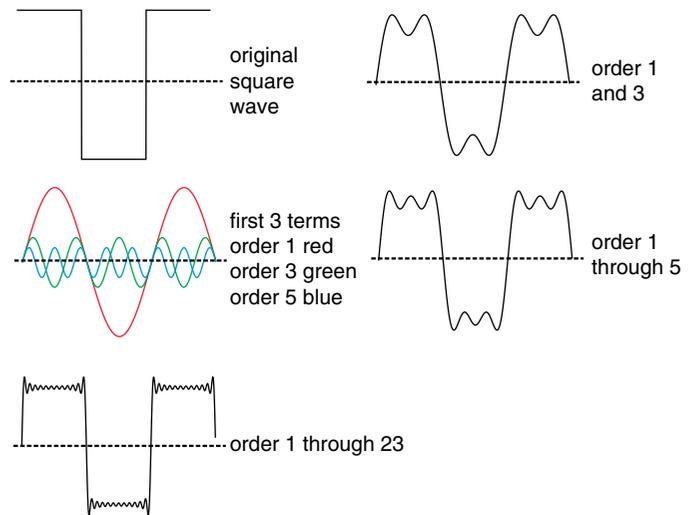

FIGURE 4.16. Fourier analysis. The Fourier theorem says that any periodic structure (such as the "square-wave" in the top left) can be represented as the sum of a number of sine waves, each of which is harmonic of the frequency of the structure. Think of these frequencies as the spatial frequencies introduced in Figure 4.2. As more components are added to the sum, the result looks more and more like the original. The same thing happens in microscopy, where using a lens with a higher NA allows more "terms" that are carrying high-frequency information (and therefore diffract at higher angles) to contribute to the image.

determining the intensity to be stored for each pixel does not involve measuring small numbers of quantum-mechanical events. A continuous signal is not capable of changing by a large amount from one pixel to the next because the pre-amplifier bandwidth was too narrow to permit such a rapid change.

In the microscopic case, the Abbe bandwidth limits the amount of change possible between adjacent Nyquist pixels. However, in the confocal, Poisson noise can effectively sneak past the "pre-amp"[13] and get digitized as part of the signal. As a result, such abrupt changes **can** be recorded.

Consider the following example: Suppose we record the image of a bright point object on a black background using Nyquist sampling. A one-dimensional (1D) transect across the center of this feature might include 5 pixels. If sampled many times, the **average** recorded intensities in the central pixels might represent 10 photons, 8 in the pixels on either sides, and 3 for the two pixels next farther out.

Had we recorded these averaged values, we would only have to worry about the "boxy corners" artifact noted above. However, if we only record a single set of values, Poisson noise introduces another factor. On any particular sampling of this line of data, we will generally not get the average values but something else. Were we to record not 3, 8, 10, 8, 3 but 2, 7, 13, 10, 4, the resulting reconstruction would now be very different. In particular, the center of the feature would have moved right and it would now appear narrower. The transients caused by the statistical nature of the signal have made a proper reconstruction more difficult.

In fact, one would be correct in saying that, as the accuracy of the values stored in the computer are always limited by the statistics involved in counting quantum mechanical events, we can never know their "true" intensity of any pixel and our efforts

FIGURE 4.15. How "limiting the bandwidth of the output amplifier" smooths off the rough corners (B) and improves the reconstruction (C).

[13] This is possible because, in this case, it is the microscope optics that limits the bandwidth rather than an electronic pre-amplifier.

to make a reconstruction of the object are doomed to be only approximate.

While this dismal analysis is correct, we would like at least to make this approximation as accurate as possible. We can do this by applying the second Nyquist constraint: treating the data stored in the image memory to make sure that they do not contain spatial frequencies that are higher than the optical system could have transmitted. Although the best way to do this is to subject the 3D data set to iterative 3D deconvolution, much benefit can be gained by applying a simple 2D or 3D Gaussian smoothing filter. The effect of such a filter is to make the intensity of every pixel depend to some extent on the intensity of 63 or 124 neighboring voxels (depending on whether a $4 \times 4 \times 4$ or a $5 \times 5 \times 5$ smoothing kernel is used). This filtering effect averages out much of the statistical noise, reducing it by an amount proportional to the number of voxels in the convolution kernel.

If we apply a smoothing filter that simply suppresses "impossible" spatial frequencies (i.e., those higher than the optical system is capable of producing), the contrast of small features that owe their (apparent) existence solely to the presence of quantum noise in the data will be greatly reduced.

It is important to note that applying such a filter will reduce the apparent contrast of the image data. Digital look-up tables can be used to increase the apparent contrast on the viewing screen and the resulting images will be just as contrasty and will show less statistical noise than the raw data.

MORAL: Your image is not suitable for viewing until it has been at least filtered to remove features that are smaller than the PSF, or, thought of the other way, to remove data having spatial frequencies beyond the maximum bandwidth of the optical system in use.

Some Special Cases

In classic sampling theory, the time (or space) taken to measure or sample the intensity in each pixel is very small compared to the inter-pixel sampling interval. Although this condition is met for the ADCs used in commercial confocal microscopes,[14] it is not met for the CCD camera, where it is common for the sensitive area of 1 pixel to be almost equal to the area it covers on the silicon surface. Clearly, this means that the space taken to sample the light flux is almost the same as the pixel size. The system works well enough as long as we stick to Nyquist sampling of a signal of known bandwidth (4–5 pixels/blob). In fact, some CCD manufacturers have gone even further and made an effort to increase the effective spatial resolution of a CCD by making a series of four or nine exposures, in which for each member of the series the sensor is offset from the previous one by one half or one third of a pixel in x and y, respectively.[15] While reasonable results can be obtained in this way, problems can develop when it is applied to color imaging systems that employ a color pattern mask on each pixel (Fig. 4.17).

Even in the confocal microscope, problems can occur because the zoom magnification control creates a variable relationship between the optical bandwidth of the signal and the electronic bandwidth which is set by the "time-constant" of the pre-amplifier just before the ADC. Not only that, but the optical bandwidth creates blurring in 2D while the pre-amplifier time constant only
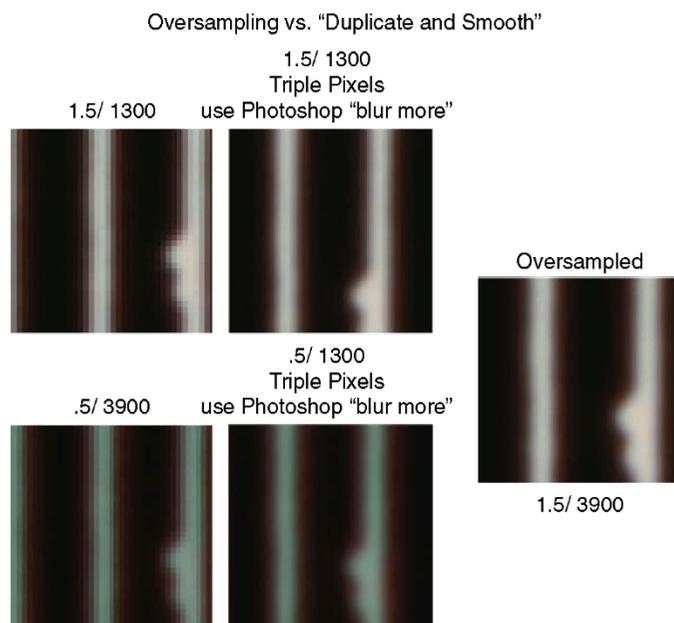

Oversampling vs. "Duplicate and Smooth"

FIGURE 4.17. Wiggling the chip to increase the resolution. A number of CCD cameras increase their effective spatial resolution by adding together the information from a set of four (or nine) images, each of which is collected with the CCD displaced horizontally by one half (or one third) of a pixel, first in x, then in y, compared to the other members of the set. The image in the top left was made at a tube magnification of 1.5× with the camera readout normally, while that below it was made with 3× less optical magnification and 3× more "resolution" by moving the chip. Both look quite similar when they have been up-sampled and then blurred in Photoshop except that the "wiggled" result has a color caste caused by the fact that the color mask filters on the chip have a "pitch" that is twice the nominal pixel spacing. The up-sampled and blurred picture on the right is analogous to the round, red feature in Figure 4.1.

limits the signal bandwidth in the fast scan direction (usually horizontal). If the zoom is set to under-sample high-contrast optical data, then very large pixel-to-pixel variations are possible and the bandwidth should be wide. The reverse is true for over-sampling. In response to this problem, some commercial instruments estimate the optical resolution from the NA of the objective and the wavelength of the laser and then use this information to set the pre-amplifier to the optimal time constant (bandwidth) on the basis of the zoom setting. When such a system is functioning properly, the "apparent noisiness" of the signal recorded from a bright but relatively featureless object will become less as the zoom is increased: the signal/pixel remains the same but the longer time constants effectively averages this noisy signal over more pixels in the horizontal direction.

Starting with the MRC-600, all Bio-Rad scanners used full-integration digitizers. These were composed of three separate sections. At any instant, one is integrating the total DC signal current from the PMT during a pixel, the second is being read out, and the third is being set back to zero. This system effectively emulates the image digitizing system of the CCD. This approach works well for under-sampled data and was a great improvement on earlier systems that used a time constant that was fixed at (pixel time/4) and therefore let a lot of high-frequency noise through to the ADC.

If you don't want to worry about any of this, stick to Nyquist!

---

[14] Common values might be pixel period = $1 \, \mu s$, sampling interval = 3 ns.
[15] For example, the Zeiss Axiocam.

GRAY LEVELS, "NOISE," AND PHOTODETECTOR PERFORMANCE

When an image is digitized it must be quantized in intensity as well as in location. The term **gray level** is the general term referring to the intensity of a particular pixel in the image. Beyond this general definition, things become more complicated. What kind of a measure? Linear? Logarithmic? What is **a gray level**? Let us begin at the beginning with a discussion of how these matters were handled by the first reliable method for recording image intensity: photography.

Optical Density

Early work on the quantification of image intensity was related to the performance of photographic materials. Developed photographic emulsions are darker where they have been exposed to more light. However, this darkening process is not linear because the individual grains of silver halide that make up the emulsion only become developable after absorbing, not one, but two light photons within a short space of time (~1s).[16] As a result, at low exposures the number of grains exposed is proportional to the square of the **light intensity**, a term we will use here to represent the number of photons per unit area per unit time at the detector.

The **photometric response** of photographic emulsions is quantified in terms of so-called HD curves. These plot the log of the light intensity (H) against the log of the darkening (D). Figure 4.18 shows the important features of such a curve. The darkening is measured as a ratio compared to a totally clear film, using logarithmic optical density (OD) units: OD = 0 implies no darkening and all the light is transmitted; OD = 1 means that the emulsion transmits 10% of the incident light; OD = 2 implies that it trans-
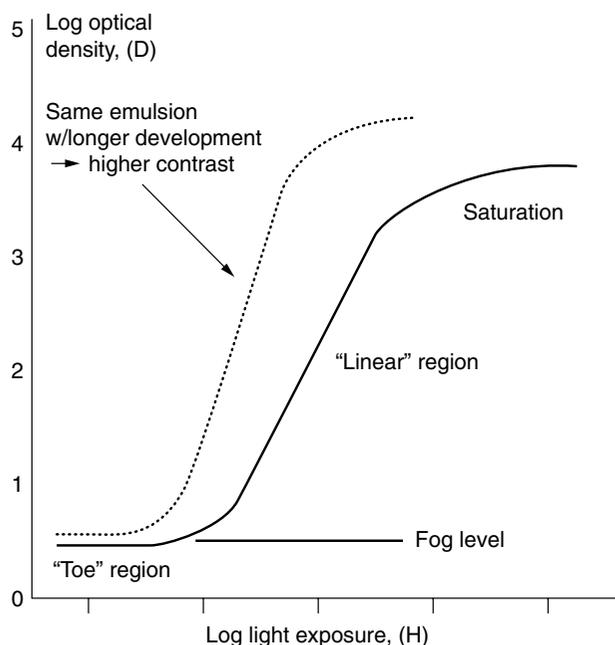
mits 1% of the incident light, etc. The use of a log/log scale allows one to describe the HD response over 4 to 5 orders of magnitude on a single plot. However, it also obscures much of the quantitative complexity of the plot and parts of it that seem "linear" would not seem so on a linear plot.

Because there is always some background exposure of the emulsion, **D** is never "zero" but starts at the "fog" level. Small exposures produce almost no additional darkening because few grains receive two hits. Eventually, however, the log of darkening seems to become proportional to the log of exposure and the response curve enters its "linear" region. At high intensity, the response saturates for two reasons: as there are only a finite number of grains in each pixel, one cannot do more than develop all of them. In addition, as more grains are developed, they are more likely to be "behind" other developed grains and so each new grain contributes relatively less to the darkening of the emulsion. The presence of a background or noise-level signal and some sort of saturation effect at high exposure is not unique to photographic emulsions, but characterizes all types of photodetectors.

The response of a given emulsion will depend on the development conditions (type and concentration of developer, time, temperature) as well as the exposure level (light intensity × exposure time). The "linear" part of the curve becomes steeper (higher contrast) and starts at a lower exposure level if the development time or temperature is increased.

As the best photographic negatives are recorded using exposures representing H values below the center of the linear portion of the H–D curve, the transition region from the fog level to the linear region (called the "Toe" response) is of prime importance to the final result. In this region, the density is roughly proportional to the exposure squared. Of course, the photographic paper used to print the final result also has a photographic emulsion. Although print development conditions are more standardized, printing papers can be purchased in different contrasts.

In principle, one might suppose that the ideal situation would be for the paper to have an H–D response that just complemented that of the negative. The resulting print would represent a close approximation of the intensity values of the various parts of the image originally passing through the camera lens. In practice, a perfect match of these two square-law curves is very hard to achieve but this sort of compensation still occurs to some extent. For example, every camera lens transmits more light/unit area (and hence produces more darkening of the negative) in the center of the field than at the periphery. However, as this is also true of the optical system used to print the negative, the two non-linearities partially cancel out because the denser center of the negative serves as a sort of local neutral density filter.

The Zone System: Quantified Photography

Ansel Adams is justly famous not only for the beautiful images he recorded of nature but also for inventing **The Zone System** for quantifying photographic exposures and the response of various different emulsions. Each zone represents a brightness in the image being recorded that differs in intensity from neighboring zones by a factor of 2.

Adams believed that a good print could transmit information over a range of seven zones[17] and that it was important to match

FIGURE 4.18. Photographic H–D curve.

---

[16] The response of photographic emulsions exposed to X rays, electrons, or other ionizing particles is quite linear.

[17] That is, that the brightest part of the print would reflect about $(2)^6 = 64$ times more than the darkest. This was a bit optimistic as a black print emulsion still reflects ~2%–3% and white only ~97%, a ratio of only about 1:30.

the range of brightness in the scene (which might be either more or less than seven zones) to the 64 : 1 range of brightness levels that could be seen in the print. This could be done by making a judicious choice of emulsion, exposure, and development conditions. While it is not appropriate here to go into all the details of this process, two aspects of this system deserve mention:

- The size of each inter-zone intensity steps relates to its neighbor logarithmically,[18] much like the eye/brain system (see below).
- The system is non-linear like the square-law response of a photographic emulsion exposed to light.[19]

Although this logarithmic response served well in photography, modern scientific imaging tends to prefer image recording systems with a linear response.

Linearity: Do We Need It?

There is obvious appeal to the idea that the intensity value detected in a given pixel should be linearly related to both the numerical value stored in the image memory and to the brightness of the same pixel when the image is finally displayed. It seems that this should be easy: most electronic photodetectors and ADCs are linear. It is also logical: how else could one represent what has been measured?

Although linearity does indeed have these advantages, there are some practical complications when applying it to electronic imaging, especially when viewing the sort of image data often encountered in fluoresecence microscopy. These complications have two sources:

1. Non-linearity is inherent in all the common methods whereby one can **view** digital image data: computer displays and grayscale or color hardcopy. In addition, there is the problem of how, or even if, one should try to account for the fact that the retinal/brain response of the eye is more logarithmic than linear.
2. Because of Poisson statistics, intensity values representing only a small number of photons are inherently imprecise; displaying as different tones, intensity steps that are smaller than this imprecision is pointless and can even be misleading. Worse still, the absolute imprecision is not constant but increases with the square-root of the intensity: the errors are greatest in the brightest parts of the image, "where the dye is."

The Goal

We start with the idea that the over-riding purpose of microscopy is to create in the mind of the observer the best possible estimate of the spatial distribution of the light intensities representing the structural features of the specimen. The question then arises as to whether or not one should bias the digitization or display processes away from "linearity" to compensate for the inherent statistical and physiological factors. We will try to answer this question with a **(very!)** quick review of some relevant aspects of human vision.

Problems Posed by Non-Linearity of the Visual System and Image Display Devices

Studies of the photometric response of the eye to light concentrate on the just noticeable difference (JND). It has been found that most people can recognize a feature that differs in brightness from its background by a fixed fraction of this background light intensity level, for example, 10%. Although there are some limitations on this response,[20] it can easily be seen to be inherently logarithmic. In addition, the number of "gray-levels" that a human eye can see is fixed by the size of the JND and the dynamic range of the image it is viewing.

Suppose that the background has 10 units of intensity, then if the feature is to be visible it will have to have either >9 or <11 units, a change of 10% or 1 unit. However, if the background is 100 units then the 10% JND step will be 10 units: 10 times bigger. No smaller increment or decrement will be visible to the eye.

How might you go about displaying the intensities 9, 10, 11, 90, 100, 110 units on a computer screen? Most computer video memories are 8 bits deep. This means that (notionally, at least) they can store and display a maximum of 2^8 or 256 different signal intensities.[21]

Suppose that we load our six intensity values into the video memory without change as the numbers 9, 10, 11, 90, 100, and 110. This will mean that the brightest part of our image uses less than half (110/256) of the numeric display range of which the display is capable. It also means that we do not utilize any of the memory levels between 11 and 89, 91 and 99 or 101 and 109, etc. Does this mean that we now have an image of 256 gray levels, of only the six levels of our object, or some other number?

Alternatively, to better utilize the dynamic range of the video memory, we might multiply the original numbers by ~2.3 before putting them into the video memory. The brightest patch would then be $2.3 \times 110 = 253$, almost at the top of the possible range. Do we now have an "8-bit" image?

What of the computer display itself? Is there a linear relationship between the number stored in the video memory and the number of photons/second emitted from the corresponding pixel of the CRT screen or LCD image display?

In a word: No! The exact relationship between these two is determined by the display manufacturer and the settings of the contrast, brightness, gamma, hue, and saturation controls. Although display standardization is possible for those working in the color printing industry, it is a very complex process seldom attempted by working microscopists, at least in part because it requires standardized room lighting. The fundamental problem for CRT displays is that, while the brightness of a pixel on the screen is directly

---

[18] The increment in darkening present between zones 6 and 7 represents the effect of recording 32 times more "additional photons/area" than the increment between zones 1 and 2.

[19] The steps in the brighter parts of the final image represent a larger increment of exposure than in the darker parts.

[20] We are very good at interpreting small changes in brightness at edges, but see the uniform areas on either side of the edge as shaded, even though they are not shaded. In fact, given the large number of recognized optical illusions, one must treat the eye/brain as more suitable for measuring contrast than intensity.

[21] In this discussion we will ignore the inconvenient fact that the performance of most display systems is itself limited by Poisson statistics. For instance, each "pixel" on the CRT contains only a small number of phosphor crystals, each of which may be more or less efficient at converting energy from the three beams of electrons into light. Only a very small fraction of these photons will pass through the pupil and be detected by the retina. How many actually do is subject to statistical variations. In addition, each of these three electron beams deposits only a small number (1000s?) of quantum-mechanical particles (electrons) into a small area of the tube surface during the "pixel-dwell" time. The exact number deposited is limited by Poisson statistics. Just because ignoring these complexities makes analysis easier does not mean that they are unimportant.

proportional to the amount of current in the electron beam during the time it is illuminated (i.e., the total charge deposited), this current is in turn proportional to the "3/2" power of the voltage applied to the control grid of the tube. Therefore, even if the digital number in the video memory is turned into a linearly proportional voltage at the grid of the CRT, changes in this value will produce a "more than proportional" brightness change from the phosphor.[22]

Cathode-ray tube manufacturers are aware of this problem and have developed a variety of countermeasures. The umbrella term for these efforts to introduce a compensating non-linearity in image brightness is **gamma**. In electronic imaging, gamma is a measure of the non-linearity of the relationship between stored intensity information (Fig. 4.19) and displayed brightness. It can be more or less than unity. Positive gamma stretches out changes in intensity occurring at the lower end of the intensity scale and compresses those occurring at the top of the scale. Negative gamma does the reverse. If nothing else, the presence of a gamma other than one means that shifting the average signal level up or down (using the black-level or "brightness" control) will also change the "gain" (or contrast) of the final result.

Given the uncertainty regarding the correction software used on a particular CRT, the setup adjustment of the individual electron guns themselves, not to mention differences introduced by user settings of the controls or the use of different phosphors on the face of the CRT, variations in the level and color of ambient lighting, the average user can have little confidence in the intensity linearity of most CRT displays.

The same situation is even more true for displays that incorporate LCDs, where viewing angle to the screen is an additional and important variable. Non-linearities also abound in all of the types of hard-copy renderings made using digital image printers:
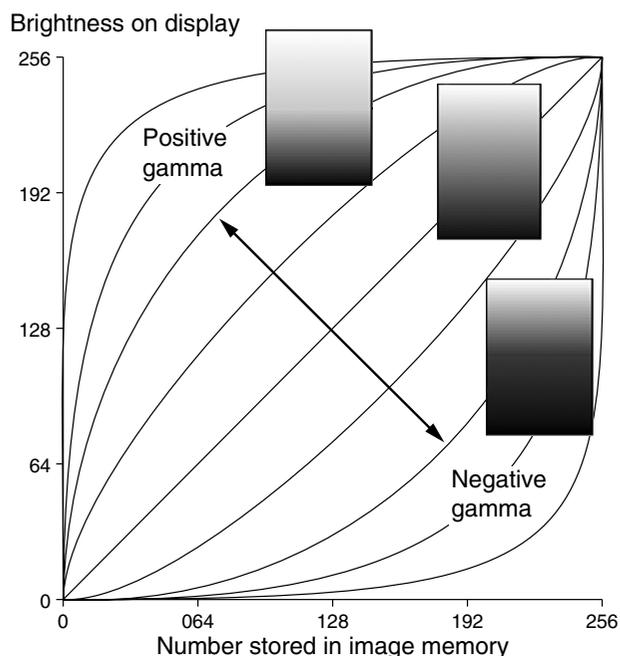

FIGURE 4.19. How the gamma control varies the relationship between input signal level and output brightness on the screen or the printed page.

---

[22] There are also other variables that affect pixel brightness: beam voltage (this may dip if the whole screen is bright vs. dark) or blooming (if the beam contains too much current, it will become larger, i.e., "bloom"). When this happens, central pixel brightness remains almost constant while adjacent pixels become brighter. This is not a complete list.

spectral properties of dyes, dither patterns, paper reflectance and dye absorption, etc. This topic is covered in more detail in Chapter 32.

Once one realizes that strict linearity is neither possible nor perhaps even desirable, one can move on to "distorting the gamma in a way that allows the observer to see the biological information that the image contains" while trying to be careful not to introduce irresponsible or misleading artifacts. Clearly, this is a hazy area in which much discretion is needed. The topic of responsibility when processing images is discussed in Chapter 14.

Matching the Gray Levels to the Information Content of the Image Data

When Ansel Adams developed the Zone System, no consideration was given to limitations on the recorded intensity other than the exposure conditions (exposure time and lens aperture), and the intrinsic illumination and reflectivity of object. This attitude was justified because the exposure of a single pixel on the large-format negatives he used involves the absorption of thousands of photons by a small fraction of the millions of silver-halide grains located there. As a result, statistical variations in this number (i.e., square root of the number of developed grains) were likely to be small compared to the 10% JND. The same was true of the printing papers.

In live-cell microscopy generally, and confocal fluorescence microscopy in particular, this condition is often not met. The fluorescence signal is inherently weak—about a million times less intense than the excitation light used to produce it. Although this limits the **rate** at which data can be produced, bleaching and phototoxicity may impose even more stringent limits to the total recorded intensity. **In other words, the imaging modality imposes absolute limits on the total number of photons that can be detected.** As a result, in biological fluorescence microscopy, we are usually starved for photons. In laser confocal microscopy, it is not uncommon to collect only 10 to 20 photons in the brightest pixels and zero or one photon in the unstained regions that often constitute a large majority (<99%) of the pixels in a particular scan.

Suppose that the signal in the **brightest pixel** of a confocal fluorescence image represents only **16 photons** (not an unusual figure). As we do not have negative photons, and even though we are collecting these data into an "8- or 12-bit" image memory having 256 or 4096 possible intensity intervals, respectively, one cannot imagine that an image in which the highest intensity was only 16 detected photons could possibly have more than 16 meaningful gray levels corresponding to 1, 2, 3, . . . 16 photons.

However, because the counting of photons is a quantum-mechanical event and hence limited by Poisson statistics, the number of "meaningful" intensity steps in this signal is even less. The brightest recorded signal is really 16 ± 4. The next dimmer signal level that can be discriminated from it by at least one standard deviation (σ), is 9 ± 3. With a 16-photon peak signal, we can discriminate only four "real" signal levels. These correspond to the levels 1 ± 1, 4 ± 2, 9 ± 3, and 16 ± 4.

This is really quite inconvenient. What can be done if the staining levels of our specimen, as modified by the CTF of our microscope, do not coincide with this square-law of statistical detectability? There is only one option: to collect more signal (more dye, longer exposure, etc.) or average the data in space over the <64 voxels that represent the whole, Nyquist-sampled, 3D PSF by **deconvolving** it as discussed above and, in more detail, in Chapter 25.

Beyond this, the only strategy is humility: Don't base great claims on the detected brightness of one or even a few pixels but on patterns visible in a number of images from many specimens.

GRAY LEVELS IN IMAGES RECORDED USING CHARGE-COUPLED DEVICES: THE INTENSITY SPREAD FUNCTION

The data recorded using CCD detectors for widefield /deconvolution are subject to similar limitations. Conventional CCDs have higher quantum efficiency (QE), but much higher readout noise, than the photomultiplier tube (PMT) detectors used in most single-beam confocals. The higher QE increases the number of detected photons (thereby reducing the effect of Poisson noise) but the presence of substantial read noise reduces the number of useful gray levels substantially below what one would estimate from Poisson noise alone. Read noise becomes even more important when excitation levels are lowered to reduce phototoxicity. Because both the sources and the form of the noise signals are quite different in each detector, it has been difficult to make comparisons that are both simple and informative and the recent introduction of the "electron multiplier" CCD readout amplifier (EM-CCD) has made comparisons even more complex. The discussion that follows describes an effort to define a suitable measure of overall photodetector performance.

What Counts as Noise?

Just what counts as "noise" in the context of fluorescence microscopy is far from settled. Should one count as noise the signal from non-specific staining? From stray or reflected light in the microscope? Fixed-pattern noise traceable to stray magnetic fields or electronic interference? Even among practicing microscopists, it is not uncommon for "noise" to become an umbrella term for anything that makes an image resemble the "snowy" output of a television receiver displaying the signal from a distant station. Although a variety of very different physical processes can produce such a "noisy" signal, only some of these can be related to defects in the performance of the detector/digitizer system. For example, it is common to hear that turning up the gain of the PMT makes the confocal image "noisier." It would be more proper to say that the fact that the PMT needs to be so high is an indication that the signal itself must be very weak and hence must contain a very high level of Poisson noise.[23]

In the discussion that follows, three types of noise will be considered:

- Poisson noise: the irreducible minimum uncertainty due to quantum mechanics.
- Readout noise: assumed to be random fluctuations in the electronics and virtually absent in the PMT and the EM-CCD.
- Quantum efficiency: Although many people think of QE as totally separate from noise, because it reduces the number of quantum mechanical events sensed, it increases the effect of Poisson noise.

One can define noise in any imaging system as that part of the electronic output of the detector that is not related to the number
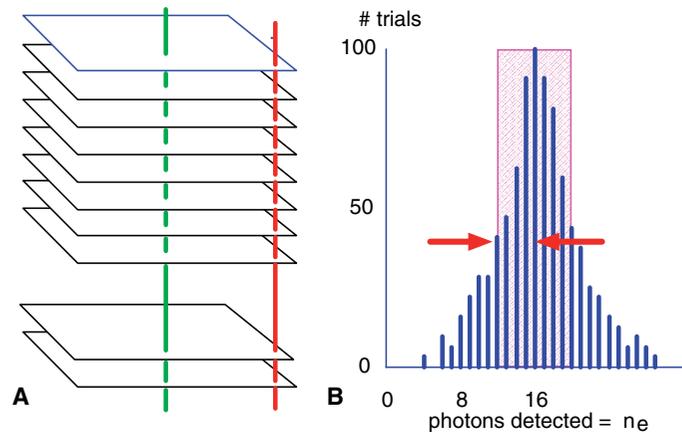

FIGURE 4.20. The ISF for a signal of 16 ± 4 photons/measurement. (A) About 700 identical exposures are acquired. The values collected at few particular pixels (red and green lines) are then converted from ADU units in the computer, to photoelectrons (n_e) and assembled to form a histogram (B). This example shows the number of photons absorbed at pixel p, assuming that, n_e, was 16 electrons/pixel and there is no read noise. Approximately 63% of the trials yield a value for n_e in the range of 16 ± 4 or between 12 to 20 (pink-shaded box). The halfwidth of this distribution (red arrows) equals the RMS noise of this measurement. The remaining 37% of trials yields a value outside this band.

of photons detected per unit time and/or space. However, as the electronic signal from the PMT is digitized in a very different way from that of the CCD, it is much more useful to compare the performance of these detector systems not in terms of the signal from the detector but in terms of the number that is measured by the ADC and stored in the image memory to represent the brightness of a single pixel.

Suppose that the average number of photons striking pixel, p, of a CCD during a number of identical exposure periods is n_p. This exposure will excite a number of electrons, n_e, into the valence band in the location of p, where n_e is smaller that than n_p because the QE is always less than 1. In fact:

$$\frac{n_e}{n_p} = QE \qquad (1)$$

One might imagine that the best we can do is to measure n_e. However, as noted above, even this is impossible because the absorption of a photon is a quantum mechanical event and therefore the number absorbed on any given trial will not be constant but will vary according to Poisson statistics. If the average number of photons is 16, the histogram of numbers of photons actually absorbed on a given trial versus the number of times when this number was detected will look like Figure 4.20.

The hatched area denotes the ± 4 electrons band of values that corresponds to the $\pm\sqrt{16}$ imposed by Poisson statistics. On average, 63% of the trials should give values that lie within this band. Although only a small fraction of these trials (about 100) will yield what we have defined to be the average value (16), it is important to recognize that even a perfect photodetector (i.e., one with a QE = 1 and no measurement noise), could never record data any more accurate than this.[24]

---

[23] Unless, of course, the PMT is faulty and actually generates noise when high voltages are applied.

[24] The analysis also holds for a PMT used in the photon-counting mode and assuming that the number 16 refers to the average number of photons actually counted. As there is essentially no readout noise in such a PMT, the analysis stops here.

Other characteristics of the photodetector, such as the presence of measurement noise or imperfect digitization, can only move the distribution to the left and also widen it compared to its mean value. For example, if the QE were only 25% rather than 100%, the recorded values would cluster about four detected photons rather than 16 and the error bar would be ±2 photons — a 50% likely error that is twice that of the perfect detector (16 ± 4 represents a 25% error). Indeed, because the ratio of the peak value of this histogram to its width is a function of both the QE and the measurement noise, and also because it measures directly the accuracy of the detector in determining the number of photons associated with pixel p, this ratio of peak (also the mean) to its standard deviation (SD) provides a perfect metric for comparing the performance of the different types of photodetectors used in fluorescence light microscopy.

In analogy to the term point spread function (PSF), this metric is called the intensity spread function (ISF). Both concepts have an ideal result: the ideal PSF is the 3D Airy figure for a given NA and wavelength. The ideal ISF is the Poisson distribution for a given number of quantum events. In each case, it is easy to compare the actual result with the ideal.

The ISF is the ratio of the halfwidth at half maximum of the histogram of the intensities recorded from one pixel, Δn_p, on sequential "reads" of a constant signal, to the mean value of this signal, n_p, all calibrated in photoelectrons. The ratio of number of electrons actually counted is converted to photons using published QE curves.

$$\text{ISF} = \frac{n_p}{\Delta n_p} \qquad (2)$$

MEASURING THE INTENSITY SPREAD FUNCTION

It is important to understand that the ISF is only a meaningful measure of detector performance if the graph is calibrated properly in terms of photoelectrons rather than arbitrary computer units. Only quantum mechanical events follow the Poisson distribution. This next section discusses how such calibration can be carried out.

Calibrating a Charge-Coupled Device to Measure the ISF

Because the readout noise of the conventional scientific CCDs used in microscopy is in the range of ± 3 electrons RMS to ± 15 electrons RMS, there is no way to discriminate the signal from a single real photoelectron from that of none. As a result, the gain of the amplifiers leading up to the ADC is usually adjusted so that the smallest digitizing interval (analog-digital unit or ADU) is equal to somewhere between half and all of the RMS noise value (sort of Nyquist sampling in "intensity space").[25] The specification defining #electron/ADU is called the gain setting. In other words, if the read noise is quoted as ±6 electrons RMS, then the gain setting should be in the range of 3 to 6 electrons/ADU. On the best cameras, this gain setting is measured quite accurately at the factory as part of the initial CCD setup process and is usually written on the inside cover of the user manual.

If this is not the case, a fair approximation of the gain setting can be calculated if one knows the full-well (maximum

signal/pixel) of the CCD and the dynamic range, in bits, of the camera system as a whole. Suppose that the full-well signal is 40,000 electrons and the camera uses a 12-bit digitizing system. As 12-bits implies 4096 digitizing intervals, and, assuming that the pre-ADC, electronic, gain has been adjusted so that a 40,000 electron/ pixel signal will be stored as a value slightly less than 4096, one can see that the an increment of 1 ADU corresponds to ~10 electrons/ pixel (Fig. 4.21).[26]
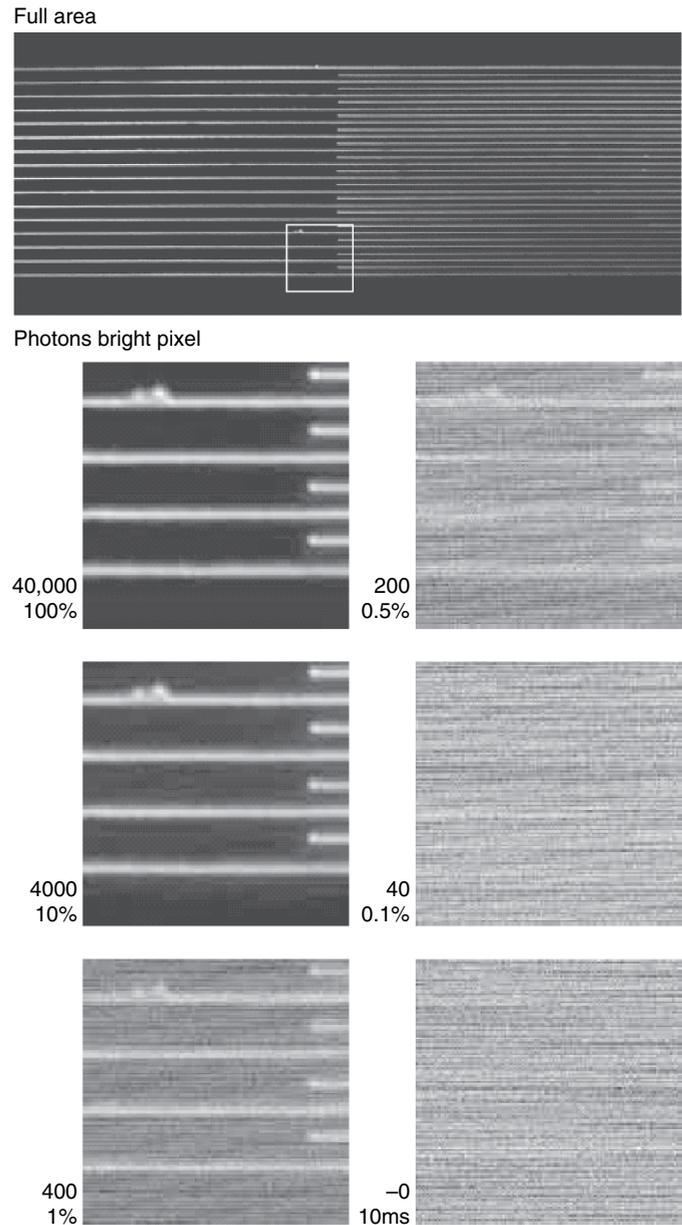
Full area

Photons bright pixel

FIGURE 4.21. Bit depth and CCD camera performance. The top image was recorded using a "12-bit" CCD camera with a "full-well" (brightest) signal level of 40 k electrons/pixel. Subsequent images were recorded with the same light level but steadily shorter exposure times. Although one might expect the camera to have a S/N of about 4000 : 1 (i.e., 12-bits), the image disappears into the noise when the peak signal is reduced by a factor of only 1000 (10-bits).

---

[25] Another factor is that ADCs tend to be made with certain fixed levels of resolution, 12-bit, 14-bit, etc., and as this feature can be "sold," it is sometimes the case that the CCD noise level spans 8 or even 16 ADU.

[26] The uncertainty is due to the practice of setting up the system so that a zero-photon signal is recorded in the image memory not as zero but as some small positive value. This prevents lost data in the event that the zero signal level drifts downwards.

Once one knows the scaling factor between digitally recorded intensities and electrons, one merely has to acquire some image data and plot an intensity histogram, as shown in Figure 4.20. Continuing the example above, one can convert the digital image data to "electron data" by multiplying the value in every pixel by 10.

"Fixed-Pattern" Noise

Although one might assume that one could measure the read noise by calculating a histogram of all the pixel intensities (converted to electrons) in a single image made with no light striking the CCD, this is not so. All the pixels on a CCD are not equal. Charge packets from pixels on the corner of the array farthest from the read node must be transferred many more times than charge packets from pixels near the read node. More transfers take a longer time and this increases the dark current in far pixels versus near pixels. Variation in dark current is just one of a number of factors that can produce the fixed-pattern noise signal. Fixed-pattern noise differs from read noise in that its form is somewhat predictable over time and consequently it can be virtually eliminated by "flat-fielding." Flat-fielding is a digital normalization process that compensates for variations in the dark current (offset or dark level) and sensitivity on a pixel-by-pixel basis. For weak images, in which pixel-to-pixel sensitivity variations are small compared to Poisson noise, this can be accomplished by subtracting a dark image, made by averaging many dark fields, from each recorded image.

To generate the data needed to plot an ISF histogram similar to that in Figure 4.20, one should not use intensity values from all the pixels of a CCD exposed to a "uniform" light signal for two reasons. Not only is it very difficult to create such a uniform signal, but the fixed-pattern noise will usually produce much more spread in the ISF than the read noise. Therefore, one must make the ISF histogram using data from **sequential** brightness measurements obtained by reading a specific pixel many times in succession.

On a 3D deconvolution system, this can be accomplished by recording as a 3D data stack a long time-series (~100 frames) of "black" images, each using the shortest possible exposure time. To obtain the set of intensities recorded from a single pixel, extract the values along a "line" oriented in the time direction in the X, Y, T data stack. As no signal photons reached the detector during these exposures and the short exposure time keeps the dark current low, the signal is shown centered on zero intensity.[27] The horizontal scale has been calibrated using the value of 2.5 electrons/ACU provided by the manufacturer. The read noise specification for this camera is ±5 electrons and this tallies well with the width of the peak in the histogram. The bottom horizontal scale is calibrated in photons, using the quoted QE of this chip at this wavelength: 40%.

The histogram should vary only slightly as the x- and y-coordinates of the vertical line are varied. In general, the noise measured in this way will be lowest in the corner of the image closest to the read node (usually the top left corner of the image on the screen), and highest in the opposite corner.

As the ISF includes the effect of errors caused by both read noise and Poisson noise, it can be used to estimate the error at any signal level. Those interested in live-cell light microscopy often operate their system so that the maximum recorded signal is less that 10% of the CCD full-well capacity (i.e., only 4000 ± 63 electrons/pixel). Under these conditions, ISF performance at low signal levels becomes important. Those using a CCD camera to record the image from a disk-scanning confocal may have a peak signal of only 10 to 100 photoelectrons.

To simulate this situation, expose the camera to uniform illumination, sufficient to fill all the pixels with a signal that is (on average!) just 10 times the read noise level (50 photons or 20 ADU, in our example). This can be done by setting up a microscope for transmitted light illumination with a clear specimen and using ND filters and the CCD camera shutter control to reduce the recorded intensity sufficiently. This may require a darkened room and it is also important to use a regulated DC supply on the transmitted light source so that all the exposures in the series are exposed to a constant light level.

This second histogram should cluster around an intensity that is 50 electrons. At 2.5 electron/channel, this would be 20 channels to the right of the center of the zero-signal histogram. It should be wider than the first histogram because now it contains ±7 electrons of Poisson noise in addition to the read noise. To the extent that the read noise signal really is random or "white" (i.e., that it does **not** have some fixed pattern caused, for example, by flickering in the light source), these two RMS noise signals should be added as the square-root of the sum of their squares. Therefore, when scaled in electrons, the width of the distribution will now be $\sqrt{(25+49)}$ or about ±8.1 electrons (or ±3.2 channels).

Although it is not possible to use this setup to measure the QE directly, comparative measurements between different cameras are possible. As QE is a strong function of wavelength, one should use a narrow bandpass filter [such as the interference green filter used for critical differential-interference-contrast (DIC) imaging] in addition to the ND filters.[28] It is also necessary to compensate for the pixel dimensions: larger CCD pixels collect more photons, have higher full-well capacities, and somewhat higher read noise. However, as the CCD must be coupled to the focus plane in the object by an optical system having the magnification required to make the CCD pixels satisfy Nyquist, pixel size *per se* is not of fundamental importance. In general the QE curves published by the chip manufacturers are hard to improve on.

GAIN-REGISTER CHARGE-COUPLED DEVICES

In January 2001, Marconi introduced the CAM 65, a CCD camera incorporating a new type of charge amplifier (Lamontagne, 2004; Robbins and Hadwin, 2003).[29] The heart of this device is a new solid-state electron-multiplier amplifier. The electron multiplier is essentially a second horizontal register, located between the output of the normal serial (horizontal) register and the input of the charge amplifier. Because it has an extra, grounded phase between charge-transfer phases 1 and 2, and a higher and variable voltage (35–45 volts vs. the normal 5–15 volts) on phase 2, electrons must pass through a high field region when passing from phase 1 to phase 2. As a result, there is a small (~1%) chance that collision amplifi-

---

[27] CCD electronics are usually set up so that the "no-light" signal is centered on a value somewhat above zero, usually about 50 ADU. This assures that signal will not be lost even if amplifier drift were to cause the peak of this distribution to drift to a lower value.

[28] All these instructions assume use of a monochrome CCD. Because of the complex way that signals from adjacent pixels, each equipped with a different colored filter, are interpolated to produce the final color image, noise in single-chip color CCDs is quite complex. As a first approximation, one could expose a color CCD to one color at a time, and calculate the ISF for each one.

[29] More info at: http://e2vtechnologies.com/technologies/13vision.htm

cation will occur. Therefore, the electron multiplier operates as a series of 500 to 700 very-low-gain avalanche diodes.[30]

When the light signal is low, the phase 2 voltage can be increased to multiply each charge packet by from 200× to 4000×. This process increases the signal from even a single PE well above the noise floor of the normal FET read amplifier, and does so even this amplifier is operating at a very high readout speed (10 to 35 MHz).

Because the EM-CCD reads out fast without high read noise, it makes focusing and searching quick and easy. Because each charge packet always remains separated from all others, EM-CCDs have full spatial resolution, and as the gain of the register can be varied, the device "can be made as sensitive as needed." Of course, if one reads out fast, there may not be enough time to accumulate much signal but this cannot be blamed on the camera.

This system breaks the old rule that scientific CCDs have more readout noise when they read out faster.

There **is** one snag.

Multiplicative Noise

The charge amplification of the gain register is not quite "noise free" because the exact amount that each electron in the packet is amplified depends on probability.

The gain-per-charge-transfer is very low: about 1%/stage. A charge packet containing 100 electrons, would, on average, contain 101 electrons after one transfer. Big deal! Why bother? Because, after 500 or 600 transfers, the average total gain can reach more than 1000×. However, some electrons do not get amplified very much while others get amplified a lot (more like the stock market than a bank!!). This gives rise to "multiplicative noise." Given a series of charge packets, each containing only a single electron, the spectrum of pulse heights at the gain register output follows a decreasing exponential with many small pulses and fewer large ones (several useful, online references are available at: http://e2vtechnologies.com/technologies/l3vision.htm). (See Figure 4.22.)

The multiplicative noise generated by this type of pulse-height distribution has the same form and magnitude as Poisson noise. As RMS noise terms are added as the square-root of the sum of the squares, the output of the gain register has 1.41× the uncertainty present in the input PE signal. Because the only way to overcome this loss in precision is to count twice as much signal, one can think of the entire system **acting as if** the amplifier were noise-free but the sensor QE is only half its actual QE.[31]

The back-illuminated sensors now available have an intrinsic QE of about 85%, or effectively 42% when used in the EM-mode. This is still very good compared to the PMT, especially in the red end of the spectrum.

The 50% reduction in effective QE only occurs when the gain-register amplifier is used. If it is turned off, the CCD QE operates as advertised, but the read noise now depends on the FET amplifier and will increase with readout speed.

Multiplicative noise also occurs in the PMT because the gain experienced by each electron at each dynode is also subject to statistical noise. Suppose that an electron from the photocathode (PE)
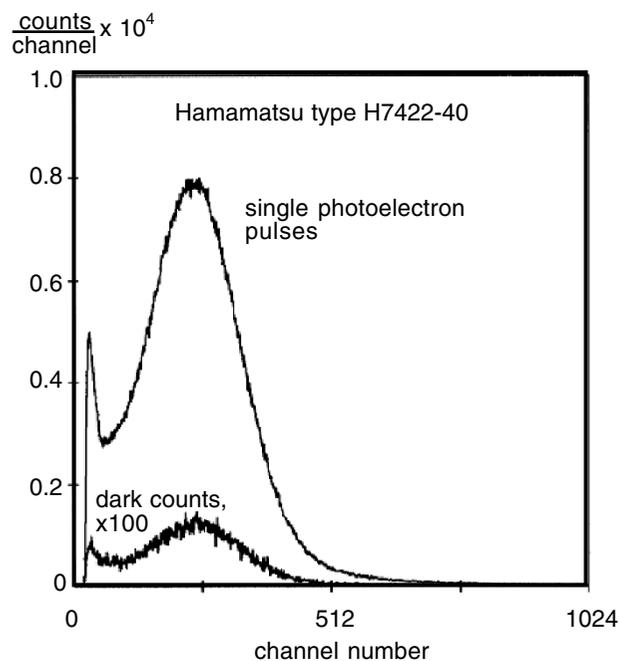

FIGURE 4.22. Single PE pulse-height spectrum.

produces an average of 16 secondary electrons (SE)/pixel on striking the first dynode. In fact, the number of SE arriving at the second dynode for a particular PE is governed by Poisson statistics: that is, it will be 16 ± 4. This uncertainty occurs at every dynode but it is most important at the first dynode because the number of quantum events is smallest there. Not surprisingly, after the pulse has propagated down the rest of the multiplier section, a PE that sent 20 SE to the second dynode is likely to produce a larger final pulse height than a PE that only produced 12 SE. The result is that single-PE pulses from the PMT may vary in size over a range of 10:1, although they will cluster around a mean value.[32] Because the gain/stage is much higher in a good PMT than in the gain register (more like 5×–10× than 1.01×), the single-PE pulse-height distribution has a distinct peak (Fig. 4.6). As a result, multiplicative noise adds only about 20% to the Poisson noise and the **effective QE** of a PMT is consequently reduced to about 70% (1/1.41) of that claimed by the manufacturer's QE curves.[33]

As used in the laser confocal, PMT multiplicative noise has another, more insidious, essentially psychological, effect. Although the histogram of the pixel intensities in a "normal" stored confocal image may show some pixels recorded at virtually every ADU level from 0 to 255, the image commonly represents a peak

---

[30] See Chapter 6 and Appendix 3 for definitions of phase I, etc.

[31] A more complete analysis of the noise produced by this detector can be found at http://www.marconitech.com/ccds/lllccd/technology.html. Try the third article: "Sub-Electron Read Noise at MHz Pixel Rates: University of Cambridge • Date: Jan 2001 • Filesize: 650kb.

[32] PMT suppliers can provide a single PE pulse-height distribution graph for any type of tube. A good tube will have a "valley" between the single-PE peak and the electronic noise peak (near zero size) that is less than 33% of the peak height.

[33] Actually, it is worse than this, because the published QE curves are "best performance" not average performance, and they refer only to photoelectrons leaving the photocathode, per incoming photon, not photoelectrons that actually reach the first dynode and result in charge multiplication. About 30% of the electrons leaving the photocathode fail to reach the first dynode and propagate. This loss further reduces the effective QE.

TABLE 4.1. Theoretical Intensity Function (ISF) Values for Various Photon Signals to Compare Performance of a "Conventional" Scientific CCD, a "Gain-Register" CCD, and a PMT, Calculated as Per Discussion in Text. Bold Numbers Show Best Performance

Signal (Photons)	Slow Scientific CCD		Charge-Multiplier CCD (Optimal)		Charge-Multiplier CCD (Available)		PMT	
	Signal (PE)	ISF	Signal (PE)	ISF	Signal (PE)	ISF	Signal (PE)	ISF
5	4	.63	2		1	1	(0.5)	
10	8	1.2	4	**2**	2	1.41	1	1
20	16	2.2	8	**2.8**	4	2	2	1.41
45	36	**4.2**	18	**4.2**	9	3	4.5	2.1
80	64	**6.4**	32	5.7	16	4	8	2.8
100	80	**7.4**	40	6.3	20	4.5	10	3.1

signal of only about 10 photoelectrons. Were it not for the multiplicative noise in the PMT,[34] the operator would be alerted to the low signal level by the "posterizing" effect of having at most 10 gray levels displayed. The presence of multiplicative noise gives the illusion of a full gamut of "gray levels" when in fact the data are insufficient to define more than three or four levels at one-standard-deviation precision.

In large part, the ISF was developed to compare the photodetector performance of confocal microscopes employing EM-CCDs with those using PMTs.

At zero light level, the halfwidth of the ISF is equal to the read noise level. Although the ISF itself will vary with the signal level, a few standard signal levels could be chosen (say 10 or 100 photons/pixel) to specify detector performance.

Alternatively, as there seems to be a general trade-off between QE and read noise, one could calculate or measure the "crossover" signal level at which equal ISF performance should occur on two detectors. Given a 100 photon/pixel input, a perfect detector would have an ISF of 10 (i.e., QE = 1 and only Poisson noise, and ISF = $\sqrt{100}$). A real gain-register CCD with an effective QE of 40% and no read noise, would have an ISF of about $40/\sqrt{40}$, or 6.3. A conventional scientific CCD with a QE of 80% and a "white" read noise of ±6 electrons would have a Poisson noise of ±9 electrons, and an ISF of ($80/\sqrt{(36+80)} = 80/11$, or about 7.3. Other values are listed in Table 4.1, above.

Unless high read speed is important, Table 4.1 shows that a normal CCD with the parameters stated will yield more accurate data (higher ISF) for any signal greater than about 45 photons/pixel.

The situation is graphed in Figure 4.23 which shows a "mean-variance" plot of the noise of an EM-CCD camera operated with no gain, low-gain, and high gain (Janesick, 2001). It shows how the "noise" (as measured by the variance) varies with the signal level (in PE). The low-gain is the noisiest overall. The no-gain is least noisy above a signal level of ~20 PE (or ~45 photons) because the effective QE is higher with no EM amplification. However, below this level, the high-gain camera becomes increasingly superior as the signal level approaches zero. Because the

EM-CCD effective QE is about three times higher than that of a PMT, this represents the high end (!) of the signal range of a confocal microscope.

Although this may sound like bad news for the EM-CCD, in fact, for a CCD camera on a disk-scanning confocal, the peak signal level may indeed be <45 photons/pixel/frame time (or indeed, less than 4!). Although 45 photons/pixel may not sound like much, remember that as long as the pixels (or voxels) are the correct size to meet the Nyquist sampling criterion (see Chapter 4, *this volume*), the 3D image of a single point object will produce significant counts in at least 64 to 125 voxels. Assuming proper digital filtering of the stored data before it is displayed, this means that at least $45 \times 64 = 2880$ photons will reach the photodetector from each "bright" point. Between 1152 and 2304 of these will be "detected," giving statistical accuracy in the 3% to 4% range.

In this simulation, the PMT never provides the highest ISF. In fact, it will do better than a gain-register CCD at very low **average** photon fluxes (~1 PE/ pixel) because its dark current and read noise is actually about 10× lower than the EM-CCD. Both the PMT and the EM-CCD are well-suited for confocal imaging because they measure zero light very accurately and this ability matches well with the fact that most voxels contain no stain.
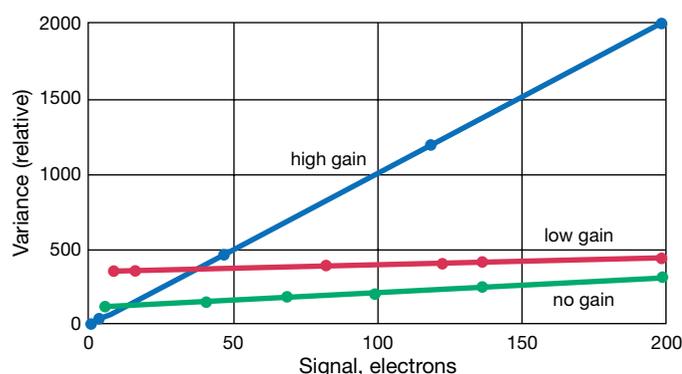

FIGURE 4.23. Mean-variance curves of an EM-CCD camera operated with no gain (green), low gain (red) and high gain (blue). The variance is a measure of the uncertaintly of the brightness measurement, obtained by subtracting two sequential CCD images of a flat "white" scene. The mean signal is the average value of the scene, calibrated in PE. One can easily see the greater accuracy provided by high-gain operation when detecting signal lower than ~20 PE. These plots kindly collected and provided by Colin Monks, Brian Bodensteiner and Pavani Korada (Intelligent Imaging Innovations, Denver, CO).

---

[34] Or if it were used in the photon-counting mode, which eliminates multiplicative noise by counting all single-PE pulses the same. Unfortunately, pulse-counting systems are usually by pulse-pileup limited to peak count rates of ~30MHz. This means that in a 1μs pixel time, 50% of a 15 counts/pixel signal would be lost to pileup.

TRADE-OFFS

Digital microscopy performance is dependent on three factors that have nothing directly to do with the specimen: optics/diffraction, photon statistics, and digitization. Unfortunately these three factors always interact. Although it is possible to make very poor digitization choices that substantially reduce the quality of the data collected (e.g., using less than Nyquist sampling or failing to count as much of the photon signal produced as possible), it is unfortunately the case that, even when one does everything right, there are still trade-offs. Assuming that the resolution is determined by the Abbe equation, this determines the pixel size needed to record data up to the Abbe resolution level. However, one will still not be able to visualize features separated by this distance unless the sample can provide enough photons. The intensities recorded in the pixels defining any small structure can only be determined approximately. If the false "contrast" produced by Poisson noise and read noise becomes comparable to the contrast of the image feature, the feature will become invisible. The only possible escape from this conundrum is ALWAYS to deconvolve 3D data and filter 2D data as this will eliminate artifactual "out-of-bandwidth" features and effectively average out the Poisson noise over the number of Nyquist voxels needed to define the central peak of the 3D Airy figure (see Chapter 25, *this volume*).

If low signal levels make seeing small details impossible, use bigger pixels (lower the zoom magnification, "bin" the CCD, or change its coupling tube magnification) and count more photons in each, bigger pixel. This will give you better statistical accuracy for the same damage level and it won't really reduce the spatial resolution because resolution is now limited more by counting statistics than by diffraction, or under-sampling. It will also have the added bonus of either a shorter scan time or a larger recorded field of view.

In live-cell microscopy, you can't have it all!!

ACKNOWLEDGMENTS

REFERENCES

Janesick, J., 2001, Scientific Charge Coupled Devices, SPIE Press, Bellingham, WA. Vol. PM83 1–920 pages.

Lamontagne, N.D., 2004, Is it right for you? *Biophotonics International*, 11(1):38–42.

Nyquist, H., 1928, Certain topics in telegraph transmission theory. *Trans. AIEE* 47:617–644.

Robbins, M.S., and Hadwen, B.J., 2003, The Noise Performance of Electron Multiplying Charge Coupled Devices, *IEEE Trans. Elect. Dev.* 50(5): 1227–1232.

Shannon, C.E., 1949, Communication in the presence of noise, *Proc. Inst. Radio Eng.* 37:10–21.