Chapter 7
Confocal Digital Image Capture

W. Gray (Jay) Jerome

7.1 Introduction

Good microscopy includes optimizing the information obtainable from your confocal image with the minimum amount of effort and time. A key to achieving this goal is to match, as closely as possible, the parameters of all hardware and software in the image collection train. For reasons outlined later in this chapter, it is particularly important to match the resolution of your digital image to the output resolution from the microscope. In Chap. 6 we discussed the resolution parameters of the digital image. In this chapter we cover the optical resolution obtainable from fluorescence confocal light microscopes, we review the important parameters of the image acquisition hardware that influence image information, and we illustrate how to match these two parameters to optimize confocal image acquisition.

For images, two related types of resolution are important: spatial resolution and contrast resolution. Spatial resolution defines the size (or fineness of detail) of objects that can be reliably detected. In addition to spatial resolution, to "see" an individual feature, there must be a sufficient contrast difference between the feature and its surroundings. Blue writing on blue paper will not be visible, even if the letters are quite large, because there is not a suitable difference in contrast. The contrast difference can be a difference in color or in brightness. Thus, bright yellow letters on a bright blue background will be easily distinguishable as will light gray letters on a dark gray background. Since an RGB color image is really three grayscale images combined (see Chap. 6 for details), we will limit our contrast discussion to grayscale values (densities), but it should be remembered that in color images, differences in color are just as valid as differences in brightness. Chapter 6

W. G. Jerome (✉)
Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA
e-mail: Jay.Jerome@Vanderbilt.edu

W. G. Jerome, R. L. Price (eds.), *Basic Confocal Microscopy*,
https://doi.org/10.1007/978-3-319-97454-5_7

discussed the dynamic range of a digital image. A larger dynamic range allows for smaller shading gradations and thus higher digital contrast resolution. Of course, our eye cannot distinguish the small differences between two adjacent gray values in the 4096 value dynamic range of a 12-bit image. However, the computer can, and so post-image processing can be used to selectively highlight specific values as long as the values were accurately captured during initial image acquisition.

The contrast resolution of a digital image collected with a fluorescent or confocal microscope is dependent upon multiple factors, (1) the inherent contrast differences in areas of the specimen, (2) the ability of the microscope to maintain those differences through the light path, (3) the camera's ability to capture those inherent differences, and (4) the image storage and display system's ability to accurately render those differences. In other words, even if a specimen generates a bright fluorescent signal that is easily distinguished from the background noise, that information will not be resolved in the final image if those contrast differences are degraded as the light passes through the specimen and microscope components, or if the camera system cannot detect the signal differences or represent those differences as discretely different values during the analog to digital conversion process. This is true even if the structures are well resolved spatially. A strength of the confocal microscope is that by reducing or eliminating fluorescence signal from out of focus planes, the contrast difference between in-focus fluorescent objects and the background is increased.

7.2 Basics of Optical Microscopy Resolution

A highly magnified image is not very useful if it is not sharp. A fuzzy image lacks fine (high-frequency) detail. Once lost, this detail cannot be regained no matter how high the image is magnified ("Resolution is a One Way Street!"). Increased magnification will only make the fuzziness more apparent. This excess magnification is referred to as "empty magnification" because it achieves no gain in image information. Since you cannot gain back information, it is imperative during image capture to obtain as much information as possible, or at least as much as is necessary for your experimental purpose. We can define the maximum image information obtainable under specific conditions as the highest resolution. If we could obtain infinitely fine resolution, we could produce an infinitely sharp image of infinitely small objects. Unfortunately, the physical properties of light constrain the size of objects that can be sharply imaged using photons with wavelengths within the visible spectrum; objects smaller than this limit will be fuzzy. Cameras and the analog to digital conversion process introduce additional constraints that potentially can reduce the resolution of the final image. For this reason, the good microscopist must understand how the microscope hardware, the digital capture device, and the image storage method can contribute to the loss of image information. The goal is to limit loss as much as possible.

7.2.1 Lateral Spatial Resolution

The limitation in sharpness of a two-dimensional microscopic image is termed the lateral (or angular) spatial resolution. The lateral resolution is usually defined in terms of how close two objects can be and still be recognized as two distinct objects. As the objects get closer together, it becomes harder to see the space between the spots. At some point, the space is no longer visible, and the two spots appear to merge into a single object. At this distance, the space is no longer resolved. The corollary is that structures with dimensions smaller than this distance will also not be resolved because these fine details will blur together.

This definition works very well for a bright-field image where multiple points are illuminated at the same time and light from adjacent objects can interfere. However, in fluorescence confocal microscopy, we generally project a focused point (or series of discrete points in the case of multipoint scanning instruments) into a sample and then image only that small illuminated area. Thus, each point in the final image is independently collected. For this type of imaging scheme, the important questions are: (1) what is the smallest excitation spot size (probe diameter) that can be produced by focusing a beam of photons, (2) how does the imaging system affect the information arising from that illuminated volume as it is projected to the detector, and (3) how faithfully does the analog to digital conversion system (camera) collect the information projected by the microscope? This section deals with the first two of these questions.

Confocal microscopes illuminate a three-dimensional volume in the specimen, but, since the constraints on the size of that volume are slightly different in the lateral (X,Y direction) and the axial (Z) direction, we will deal with the lateral and axial dimensions separately.

The key reason for the lateral image blurring in a microscope is diffraction of light waves. If the photons from a small point are focused through a lens (including the lens of our eyes), the two-dimensional image of the original point will not appear as a point in the lateral plane but rather will be seen as an Airy pattern (Fig. 7.1a). This pattern is named for the astronomer George Airy, who first described the phenomenon. The Airy pattern has a central bright circle (Airy disk) surrounded by a series of halos (maxima). The halos are separated from the central disk by dark rings (minima). As shown in Fig. 7.1b, which plots the intensity of light along the line x traversing the Airy pattern, most of the light from a point is focused into the Airy disk (about 97% of the light). The first maxima (first halo) contains most of the remaining light (about 1.7%). Each successive maxima after the first has progressively less light. In most situations we can ignore the first maxima, but you should always remember it is there. In some low light situations, it can contribute to the image.

It is obvious then that the Airy disk places a finite limit on how small an area we can illuminate in our specimen. This phenomenon also limits the size of the image of the illuminated area sent to a detector since this path also involves passage of light waves through a lens. The brightness at any particular point away from the

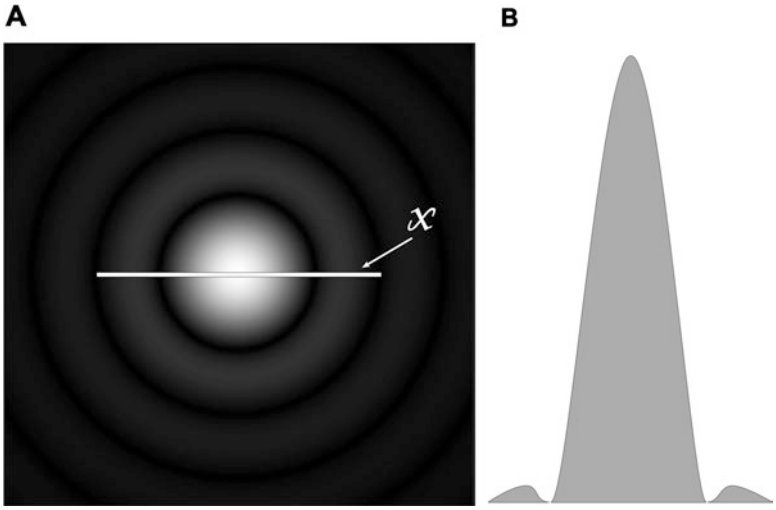

Fig. 7.1 (**a**) Image of an Airy pattern. The central spot is the Airy disk. (**b**) Plot of intensity along line X in 7.1A

center of the Airy pattern is dependent upon the strength of the light waves as they merge and the mixture of phases of all the merged wave fronts. The graph in Fig. 7.1b illustrates that there is not a sharp demarcation between very bright and very dark regions of the Airy pattern but rather a continuous change in intensity from the very brightest point of the Airy disk to the very darkest region of the first minimum.

Although it is not critical to understand all the details of why the Airy pattern is formed, it is worth a brief discussion of the general theory behind the effect. An Airy pattern results from the diffraction of light. When electromagnetic waves, such as light, meet a solid edge, the waves bend around the edge to form a new wave front. According to Huygen's principle, each point on the wave front emits vibrations in multiple directions (Huygen's wavelets). This concept is illustrated in Fig. 7.2a. At the point of focus of a lens, depicted by A in Fig. 7.2b, these wavelets arrive in phase, and so their amplitudes are additive (Fig. 7.2c). However, wavelets also arrive at other points in the same image plane (S_i) away from the focus point A. In regions where wavelets arrive exactly out of phase, such as point B in Fig. 7.2b, their amplitudes will cancel each other out as illustrated in Fig. 7.2c. This subtraction will form the first dark ring (minima). Waves converging even further from the focal point will again arrive in phase and form the first halo (maxima). Most of the light will fall within the central (Airy) disk, and this will be the brightest point. By comparison, the first maxima will be considerably dimmer than the Airy disk.

The size of the Airy disk and spacing of the halos are dependent upon the wavelength of the light and the proportion of the total light coming from the specimen that is collected by the lens (measured as the collecting angle). The radius of the disk (r) can be defined by Eq. 7.1.

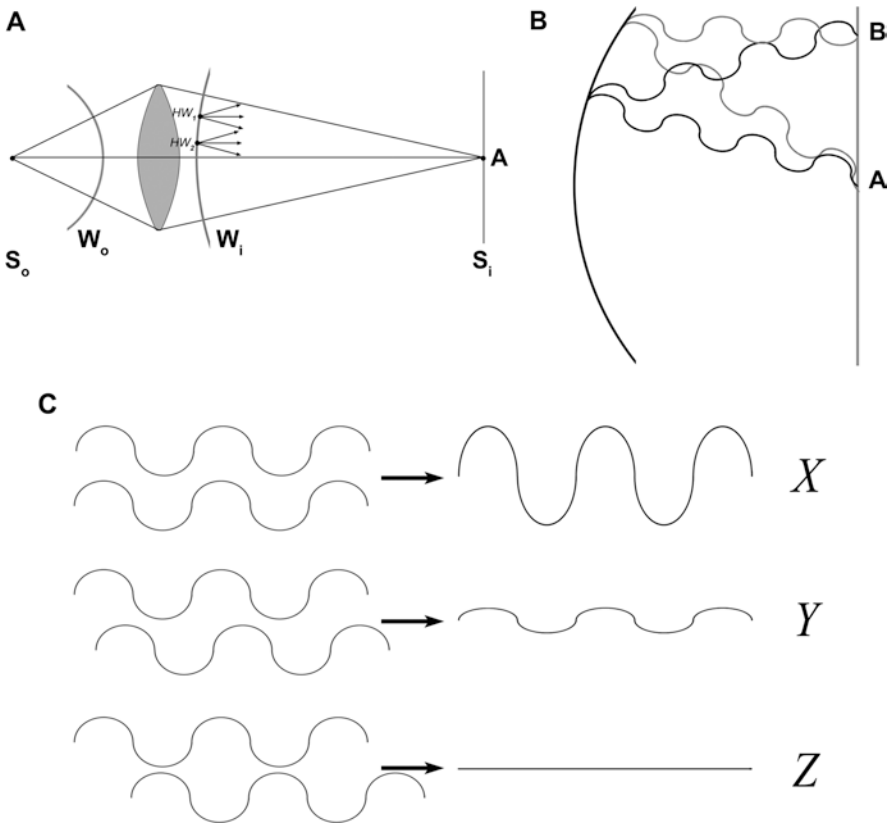

Fig. 7.2 Theory of Airy pattern formation. (**a**) Huygen's wavelets (HW1 and HW2) are formed all along the wave front (W1). (**b**) At the focal point of the lens (point **a**), the Huygen's wavelets arrive in phase and so are added together. At other points in the same plane (such as point B), the wavelets arrive out of phase and so are subtractive. (**c**) Waves that are exactly in phase (example X) will add together, and the resulting amplitude will be the sum of the individual amplitudes; waves that are exactly out of phase (example Z) will cancel each other out, and the resulting amplitude will be zero; waves that are not totally in or out of phase (example Y) will produce a signal with an amplitude in between that of X and Z

$$r = 0.61\lambda / \sin\theta \tag{7.1}$$

where λ = wavelength

θ = half the collecting angle of the lens

The equation tells us that the radius of the Airy disk is equal to a constant times the wavelength of the illumination divided by the sine of half the collecting angle. Half the collecting angle of the lens is the angle between the optical axis and the edge of the lens (Fig. 7.3).

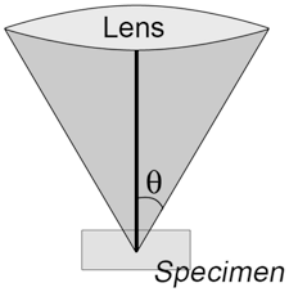

Fig. 7.3 Diagram of the cone of light collected by a lens. θ represents one half of the collecting angle of the lens

This equation defines an absolute value for the area of the specimen illuminated by visible photons. However, since we illuminate discrete volumes separately, in confocal microscopy we have the potential to localize the presence of individual small molecules that are smaller than the diameter of an Airy disk. Nevertheless, because we have illuminated an area larger than the molecule, we cannot pinpoint the precise location. We can only determine that the emitter was contained within the area illuminated. Likewise, when the object we are viewing is smaller than an Airy disk, we have no information about its size and shape. In order to get this information, the object must be larger than an Airy disk. So the Airy disk provides practical limits to the fineness of the detail we can obtain from the optical images of an object.

In recent years some newer, specialized imaging techniques have been developed that use multiple lenses or multiple illuminations to independently collect data below the diffraction limited level (Demmerle et al. 2017; Heddleston and Chew 2016; Lippincott-Schwartz and Manley 2009; Sahl et al. 2017). These increased resolution fluorescent microscopy techniques (discussed in Chap. 8) are essentially methods for gaining further information in order to pinpoint the location of an emitter smaller than an Airy diameter. As such, they present very exciting extensions of our imaging capabilities for very specialized situations. However, each of these methods comes with tradeoffs that restrict their uses to specific situations. This is why, for the majority of confocal microscopy studies, resolution is limited by the size of the Airy disk, and we will use this limit for discussion throughout the remainder of this chapter.

7.2.2 Lateral Contrast Resolution

As discussed earlier, our ability to "see" two points as separate and distinct is dependent upon both the lateral and contrast resolution. So, the observed resolution greatly depends on the detection method. In the early days of microscopy, the eye was the detection method, and so early definitions of resolution were based on the ability of our eyes to discriminate contrast differences. In the nineteenth century, John William Strutt (Lord Rayleigh) examined the resolution problem and came up

with a method of approximating the resolution limit in standard fluorescence microscopy. He proposed that there would be sufficient contrast between the centers of the Airy disks of two objects if the center of one Airy disk lies at or further away than the first minimum of the other object's Airy disk (Fig. 7.4a). In other words, we can determine the resolution limit with Airy disk in Eq. 7.1. This is true in air or in a vacuum. However, in a material with a higher refractive index, the velocity of light slows down. Thus, when using immersion oil or water, we need to account for this change in velocity because wavelength is related to velocity. Equation 7.2 shows this relationship.

$$\lambda = {}^{v}\!/\!_{f} \tag{7.2}$$

where: λ = wavelength
v = velocity
f = frequency

The frequency of light waves does not change. So, as velocity decreases so does wavelength. We can account for this in Eq. 7.1 by adding a refractive index term to the equation. The refractive index of a medium is the ratio of the velocity of light in a vacuum compared to the velocity of light in the medium of interest.

$$n = {}^{v_v}\!/\!_{v_m} \tag{7.3}$$

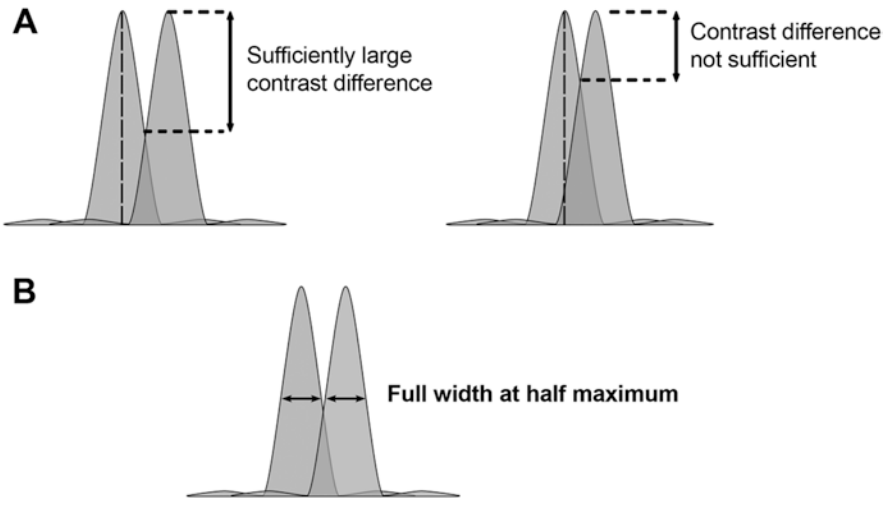

Fig. 7.4 (**a**) The Rayleigh criterion for contrast resolution. Two objects can be resolved only if the contrast difference is sufficiently large enough. (**b**) Illustration of full width at half max value for the Airy disks of two adjacent objects

where: n = refractive index
v_v = velocity of light in a vacuum
v_m = velocity of light in a medium

We can define a further term, the numerical aperture of a lens, as the sine of ½ the collecting angle times the refractive index of the medium used with the lens. This value is provided for all good quality lenses and is usually designated as the numerical aperture. Taking into consideration numerical aperture, we can formalize the Rayleigh criterion as the following:

$$r = 0.61\lambda \Big/ (\sin\theta)(n) = 0.61\lambda \Big/ \mathrm{NA} \tag{7.4}$$

where: r = minimum resolved distance
λ = wavelength
θ = half of the collecting angle of the lens
n = refractive index
NA = numerical aperture

Some typical values for microscopy imaging would be light at a wavelength of 450nms, $\sin\theta = 0.92$ (for $\theta = 67°$), $n = 1.515$. The NA would be 1.515×0.92 which equals 1.4. Given this, Eq. 7.4 allows us to determine that r (minimum resolved distance) would be 196 nm.

The Rayleigh resolution criterion is a relative one and based on the ability of our eye to discriminate contrast. If we have a detector that exceeds our eyes' capacity, the value for r can be decreased further. A related criterion partially takes detector efficiency into consideration. In this criterion, usually referred to as "full width at half max," it is assumed that objects can be discriminated if they are at least as far apart as the full width of their image at the half-maximal brightness value (Fig. 7.4 b). The double arrows in Fig. 7.4b indicate the full width of each curve at half the maximal intensity value (FWHM). For the eye, the FWHM distance will be only a little less than that determined by the Raleigh criterion. It is important to remember that neither the Rayleigh criterion nor FWHM represent absolutes. Rather, they are general rules of thumb. Methods that can increase contrast further or detectors that can discriminate smaller levels of contrast can resolve even smaller distances. One strength of the confocal microscope is the increase in image contrast. Based, at least partly, on this, the confocal provides a gain in lateral resolution. This improvement in lateral resolution is defined in Eq. 7.5.

$$r(\mathrm{confocal}) = 0.4\lambda \Big/ \mathrm{NA} \tag{7.5}$$

where: r = minimum resolved distance
NA = numerical aperture
λ = wavelength

No matter what resolution criteria we choose, the equation for the Airy disk describes the minimum area that can be illuminated by an electromagnetic wave

such as light. Since the image of that illuminated area is also the product of electromagnetic waves traversing a lens, the equation for the Airy disk also describes how the microscope affects the image of that illuminated volume. As the prominent early microscopist Ernst Abbe pointed out, we can think of a large object as a series of points that either emit or diffract light. Thus, the Rayleigh criterion defines the level of fine detail we can discern within a larger object. Resolution is, thus, the key determinant of how sharp detail will appear in our image. Just as we describe the pixel as the smallest resolvable unit of a digital image, we can define a resolved element (resel) as the smallest resolvable unit of a microscopic image. Of course, there is also a role for magnification, the other major component of a microscope. For instance, our eye cannot discriminate distances much smaller than about 0.2 mm. Thus, we need to magnify the resel in order to see it. We will have more to say on this later in this chapter.

It is worth reiterating that all criteria for resolution are intimately dependent upon the Airy disk size. We can circumvent some of its limitations, but if we use a lens for imaging, we cannot completely get rid of the consequences of the diffraction of light.

7.2.3 Lateral Contrast Transfer Function

Just as a point of light is altered as it traverses the optical path, the contrast inherent in an object is also altered by the microscope and other components of the imaging system. The effect of the imaging system on contrast can be described by the contrast transfer function (CTF). Since resolution is intimately dependent on contrast, the CTF is just as critical as the size of the Airy disk for understanding how an optical system affects resolution. Figure 7.5 depicts a theoretical CTF for a microscope. A full discussion of the CTF is more appropriate for an advanced optics book.

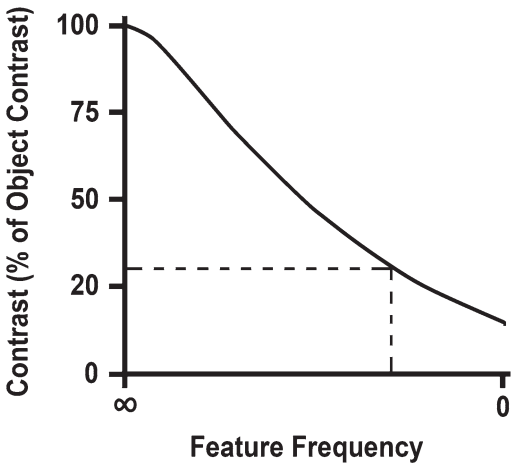

Fig. 7.5 Typical contrast transfer function (CTF) for a lens. The *dotted line* represents the Rayleigh criterion cut off point for contrast resolution

However, a few key points can be easily seen from the graph. First, there is always a loss of contrast during the imaging process. Thus, the microscope does not faithfully reproduce an object down to the Abbe-defined diffraction limit and then the image rapidly degrades; there is degradation in contrast at all levels. Second, the CTF is inversely related to feature size, e.g., small features will be reduced more in contrast than larger features. Throughout this chapter, we have reiterated the critical interdependence of spatial and contrast resolution in determining our ability to "see" an object. In fact, the Rayleigh criterion sets the resolution limit at that point where the image contrast is reduced to 25% (dotted line on graph in Fig. 7.5).

Figure 7.5 only depicts the CTF of a theoretical optical system. However, all aspects of the imaging train can also affect the CTF. Thus, we also must concern ourselves with the ability of our eye or camera to detect contrast differences. In other words, we need to know something about the CTF of each of these components. Since confocal detectors are generally grayscale instruments, the contrast resolution of a detector is related to how many separate gray tones between black and white can be discriminated. For a digital image, this is the bit depth discussed in Chap. 6. A minimum requirement for most microscopy is 8 bits (256 gray levels), but as discussed later in this chapter, higher bit depth (12-bit) detectors provide some significant advantages, particularly for quantitation.

7.3 Image Production and Resolution in a Scanning Confocal Microscope

The confocal microscope sequentially scans a focused point of light through the specimen and measures the fluorescence generated at each point. Details of this process are described in Chaps. 8 and 9. An image of the specimen's fluorescent intensity at each spot is then sequentially determined and stored as pixel information. The brightness of individual pixels in the resulting image is a representation of the fluorescence intensity at the corresponding spot in the specimen. The laser scanning microscope usually does this with a single point of light, while the spinning disk uses multiple, but still independent, points at once.

To convert the signal from an excited area in the specimen into a pixel gray value in the image requires the collection and measurement of the photons generated from the excited area by a detector. The resulting image pixel, representing a defined spot in the specimen, will have three critical values: the X and Y position of the pixel, the X and Y dimensions of the pixel (pixel size), and the gray value assigned to the pixel. Since the digital image of the scanned area is formed by summing all of the photons coming from the area illuminated into a single value that becomes the brightness value (grayscale value) for that pixel, it is clear that the resolution of an image cannot be any smaller than the size of the illumination spot projected into the specimen or onto the detector. Since, except for specialized situations, the radius of the illumination spot cannot be any smaller than the radius of the Airy disk, the Airy

disk defines the smallest area we can illuminate and thus the smallest object we can visualize correctly or, for objects larger than an Airy disk, the finest detail we can visualize. This is depicted in Fig. 7.6. Thus, except for very specialized situations, this places a theoretical limit on the lateral (X-Y) resolution that can be detected.

However, we must have a detector that is capable of detecting these closely spaced signals. This means that the detector's spatial resolution must match or exceed the spatial resolution of the microscope image (Fig. 7.6). In addition, the detector must be sensitive enough to discriminate the signal differences (e.g., have sufficient contrast resolution). Detectors for light signals run the gamut from film to diode arrays. However, most commercial confocal microscopes utilize one of two types of detector schemes. For point scanning instruments, the detectors usually quantitatively analyze the photons coming from the specimen without discriminating from where the photons were generated. The location in X, Y, and Z dimensions is inferred from the location of the excitation bean. Spinning disk confocal microscopes, on the other hand, generally employ an array of detectors that discriminate the X and Y location of the signal (the Z location is inferred from the position along the Z axis of the focused excitation light). As explained in the next sections, each of these approaches brings with it certain assumptions that must be recognized and considered in collecting and interpreting the resultant confocal image.

Throughout this chapter, the term spinning disk will refer to disks that contain an array of pinholes. However, for completeness, it should be pointed out that another variety of spinning disk microscope (often referred to as a slit-scan confocal) uses a disk with a series of narrow slits rather than pinholes. As the disk spins, the slits

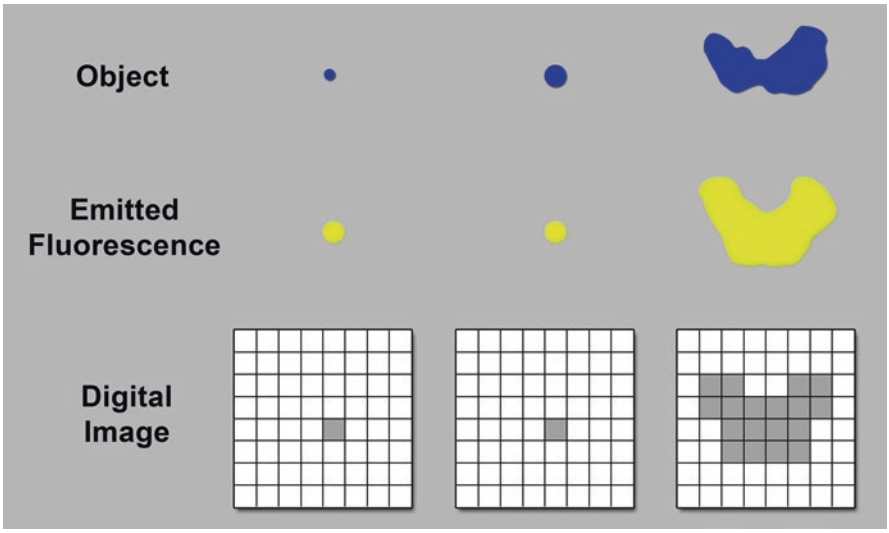

Fig. 7.6 Comparison of detection (yellow) and digital display (gray) of objects (blue) smaller than the Airy excitation area (diffraction limited fluorescence), just as large as the Airy excitation area, or much larger than the Airy excitation area illuminated by focused beam of photons

scan the excitation light over the surface of the specimen and also restrict the emitted light returning to the detector from out of focus planes. These instruments increase the light transmittance onto the specimen but at the cost of a decrease in the axial spatial resolution. Spinning disk confocals are discussed more thoroughly in Chap. 8.

7.3.1 Image Production in a Single Point Scanning Confocal Microscope

Sensitivity is a key to determining the appropriate detection device for a microscope setup. Fluorescence emission occurs over all angles, not just in the direction of the collecting (objective) lens. Thus, only a small percentage of the emitted photons enter the optical path of the microscope. Higher NA lenses will collect from a wider angle and so will gather more of the photons, but even high NA lenses collect only about 30% of the emitted light. Loss of photons within the optical train due to absorption, reflection, and deflection lowers the number of photons that reach the detector even further. In a well-maintained system with high-quality optics, it is still doubtful that more than about 20–25% of emitted photons will reach the detector. Thus, highly efficient detectors are necessary to avoid attenuating the signal any further. Besides sensitivity, a detector needs to have low noise to maintain a high SNR (good contrast) even when the signal strength is low. A good detector also should be equally sensitive to all wavelengths and should have a linear response and broad dynamic range to accommodate both faint and bright signals.

Single-point scanning confocal microscopes, such as laser scanning confocals, usually use a detector system to collect light emitted from the sample that does not discriminate from where in the sample that light originated. This simplifies signal collection. The detector does not need to acquire positional information because the fluorescence signal arriving at the detector should only arise from that area where the excitation beam is focused. When the microscope is set correctly, the confocal pinhole removes all other extraneous signals. Thus, the location of the fluorescence origin can be inferred from the location of the focused beam. The size of the pixel is assumed to be equivalent in size to the excitation area in the specimen. Historically, photomultiplier tubes (PMT) have been the detection device of choice. With a PMT, incoming photons strike a photocathode. This liberates electrons that are then accelerated by an electric potential in a vacuum tube. The accelerated electrons are directed successively to a series of electrodes (dynodes). This is depicted in Fig. 7.7. The collision of incoming electrons with the dynode surface dissipates the kinetic energy of the incoming electrons, and this energy liberates multiple electrons from the dynode surface. These electrons are directed to another dynode. By having multiple dynodes, the initial signal can be greatly amplified. The final generated electrons interact with an anode, and the consequent current is measured. The strength of the output signal is proportional to the input signal. The software then translates

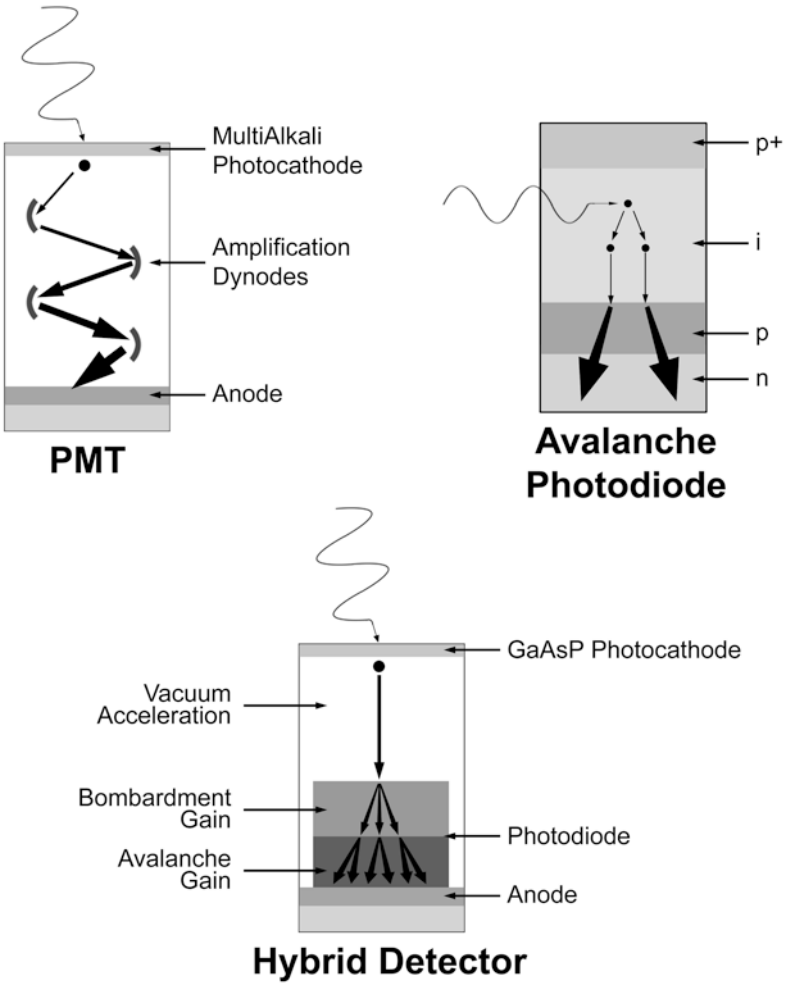

Fig. 7.7 Photon detectors for point scanning confocal microscopes. Wavy arrow indicates photon interacting with detector. Increasing size and number of arrows indicates multiplication of electrons generated from initial photon detection event

the detected signal into a grayscale value for each individual pixel of the digital image. The grayscale value of the pixel is assigned based on the signal strength, which is a function of the number of photons absorbed by the photocathode. Thus, the conversion of the analog information into a digital signal is relatively direct and straightforward with PMTs.

Standard PMTs with a multi-alkali cathode are a good robust solution for general point scanning confocal microscopy. They are relatively inexpensive compared to other types of photon detectors, and they are robust and have a high dynamic range, excellent signal-to-noise ratio, and a fast response time. However, multi-alkali

PMTs have a low quantum efficiency (QE). Quantum efficiency (QE) is a measure of the effectiveness of the conversion of photons to electrons. If every photon hitting the cathode was converted to an electron, the QE would be 1 (100%). Also, the QE of a detector usually is not uniform across the spectrum of wavelengths so will vary with wavelength of light detected. Multi-alkali PMTs have a QE of about 25% in the range from 350 nm to 800 nm. The QE of PMTs can be increased to between 40% and 50% by using a gallium arsenide phosphide (GaAsP) cathode. However, GaAsP PMTs are more expensive and more susceptible to damage compared to multi-alkali PMTs. The QE of GaAsP detectors also falls off dramatically at wavelengths below 400 nm or above 650 nm. For this reason, confocals with GaAsP detectors usually also provide at least one multi-alkali PMT as well.

A second type of photon detector is the avalanche photodiode (Fig. 7.7). An avalanche photodiode (APD) is essentially a PIN diode with an additional layer to allow signal amplification. As in a standard PIN diode, there is a positively doped P layer (p+) and a negatively doped N layer (n+) with a silicon Intrinsic absorption layer (i) between them. However, in avalanche photodiodes a thin p layer is added adjacent to the n + layer. The pn + region creates a signal amplification. During operation, a high reverse bias voltage is applied. When an incident photon creates an electron-hole pair in the i region, the hole moves toward the p + layer and the free electron migrates toward the n + layer. Successive impact ionizations of the initial electron and those generated from it in the pn + layer create additional electron-hole pairs, thus amplifying the initial free electron into multiple free electrons. This is known as the avalanche effect. APDs have high sensitivity and high speed and can have a high quantum yield (>50%) and relatively low noise, although noise level increases as reverse bias (and thus amplification) increases. APDs are very good in situations where the specimen signal is dim and quantitative single photon detection is needed. However, APDs have very low dynamic range, so they are not suited for general types of confocal microscopy or any operations where signal strength from the sample will be bright.

Recently, a new detector design has evolved that takes advantage of the high dynamic range of PMTs and the high sensitivity of APDs. These hybrid detectors (Fig. 7.7) are vacuum tubes, like PMTs, and use a GaAsP photocathode for high QE of initial photon detection. The generated electrons are then accelerated in a high voltage within the vacuum tube. However, the hybrid detectors use an APD to amplify the signal. Hybrid detectors have an overall gain that is less than a standard PMT and thus not as good a dynamic range, but they have significantly better dynamic range than APDs alone.

Hybrid detectors have many advantages for use in specific situations. Hybrid detectors have an active region that is smaller than a PMT but significantly greater than a standard APD. This compromise decreases thermal noise, but the larger active region still allows precise focusing of the sample photons onto the target region. Besides better SNR, hybrid detectors have fast response time, better pulse resolution than PMTs, and very limited after pulse. After pulse is a false signal generated within the detector by stray radiation. Hybrid detectors are excellent choices in situations where linearity of detection, wide dynamic range, and fast response

time are critical such as for two photon microscopy, fluorescence lifetime imaging, or fluorescence correlation spectroscopy.

Since the first edition of this book, the use of spectral detection has become common in confocal microscopy. Spectral detection is not really a new type of detector but rather a different approach to photon detection that allows collection of more information about the photons being directed toward the detector. In a spectral detection system, light coming from the sample passes through a dispersive element that spreads the light out spatially along its different energies (wavelengths). In this way, the various components of the signal are spectrally resolved before detection. Detection is usually by PMTs or hybrid detectors. In this way the system can analyze the number of photons generated at a precise region of the sample and determine the wavelength of the photons (usually to within 5–10 nm resolution) of those photons. Different types of spectral systems employ different methods for spreading and detecting the energies, but the result remains the same; the information is combined into a high-resolution map of the component energies emitted from the specimen. Figure 7.8 depicts an array of detectors, but other solutions to capturing spectral information, including the acousto-optical beam splitter (AOBS) Leica spectral system, have been developed. The Leica detector will be discussed further in Chap. 9. Each type of spectral system has their advantages and drawbacks. If considering purchasing a spectral detection system, the best approach is to test the different types on relevant samples to make sure the performance is consistent with your needs.

One benefit of spectral detection is the ability to estimate the contribution of two or more fluorochromes with partially overlapping emission spectra to the multifluorochrome-produced final signal. This is depicted in Fig. 7.9. Figure 7.9a shows a potential complicated emission spectrum from two fluorochromes with overlapping emission spectra. By knowing the emission spectrum for each fluorochrome (Fig. 7.9b) derived from ideal samples, one can mathematically estimate how much each fluorochrome contributed to the intensity is measured at each wavelength (Fig. 7.9c). In this way, some bleed through can be removed to get a more accurate analysis of the intensity of each fluorochrome in the sample. Different confocal manufacturers use different algorithms to unmix the collected spectrum, but all do a reasonable job of separating signals when the overlap of emission spectra is not too great.

7.3.2 Image Production in a Multipoint Scanning Confocal Microscope

In contrast to single-point scanners, detectors for multipoint scanners such as spinning disk confocals must be able to collect spatial information. Although these microscopes do excite individual points, multiple points are excited at once, and it is critical to discriminate photons coming from one point in the image from those

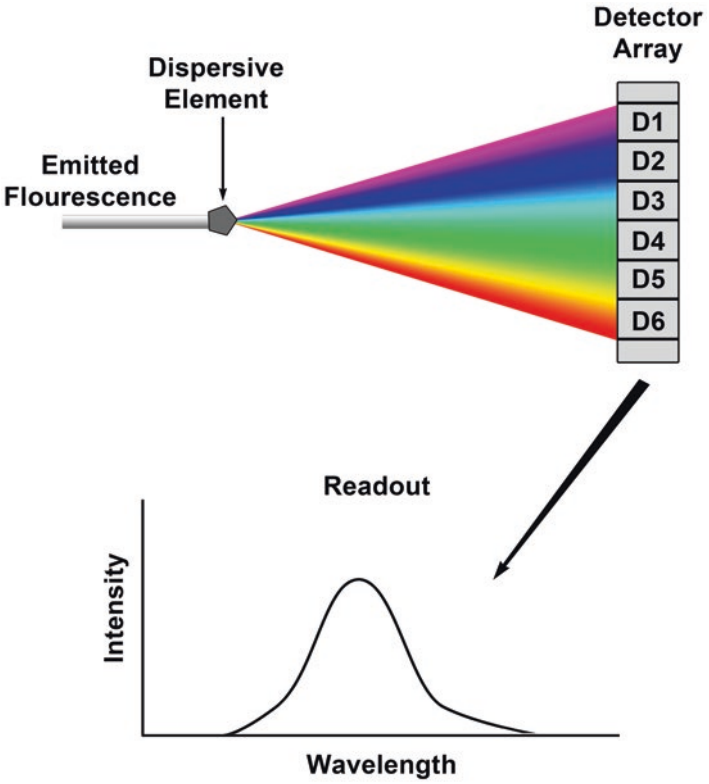

Fig. 7.8 Typical spectral detection scheme. Emitted fluorescence is divided into component wavelengths and the components sampled individually to produce a graph of intensity values at different wavelengths

photons generated in a different location. The bases for the modern simultaneous detectors required for multipoint confocals are solid-state electronic devices composed of arrays of discrete photodetectors. Each detector captures photons coming from their respective area of the specimen and converts this information to an electrical signal that is proportional to the amount of light sensed by the detector. The electrical signal is converted to digital information, and that digital information can be stored or displayed. Thus, unlike detectors for single-point scanning instruments, the capture device not only measures the signal from the specimen but also its location along the X-Y axis. This means that the size of the individual discrete photodetectors can limit the size of an object that can be visualized. This key point will be discussed later in this chapter.

The most widely used microscope photodetector systems have some form of charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) for sensors (Fig. 7.10). Although there are a number of similarities, there are also dramatic differences between these two approaches (Magnan 2003). The

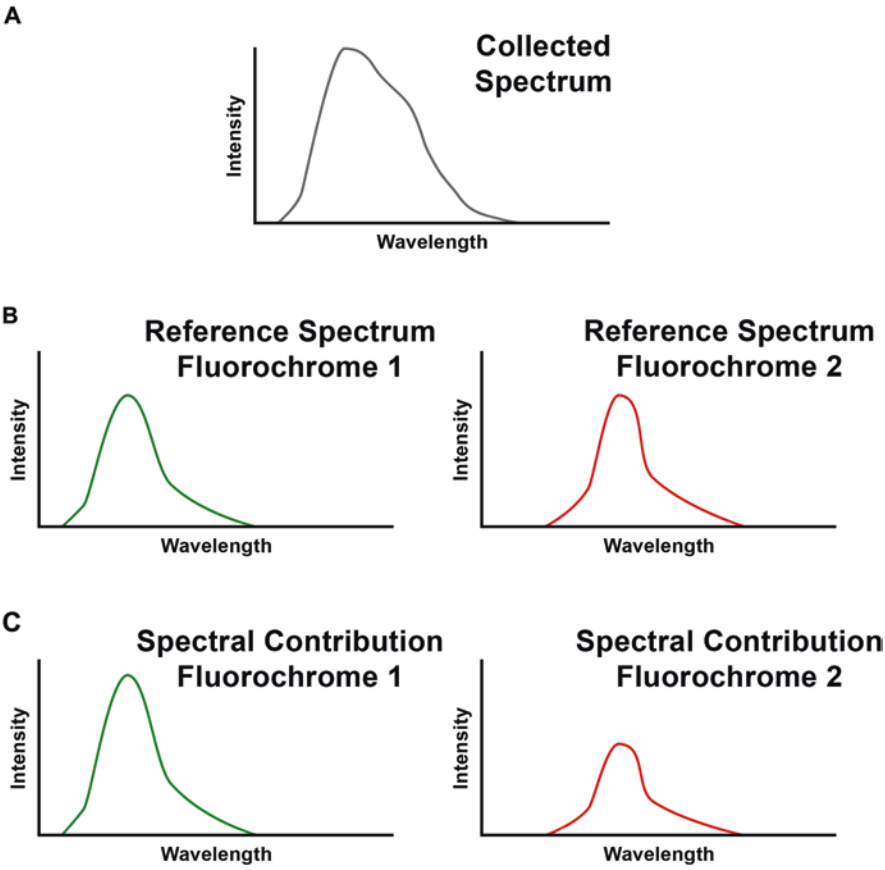

Fig. 7.9 Spectral detection allows the discrimination of the contribution to the total fluorescence of two fluorochromes with somewhat overlapping emission spectra

cooled CCD[1] (Fig. 7.10a) contains a two-dimensional array of photodiodes capable of sensing photons. The charge-sensing region of the photodiode is coupled to a charge storage region. In Fig. 7.10a, each gray square represents an individual photodiode, and each photodiode defines a single pixel. As photons from the specimen hit the photodiode, they are converted to charge. Thus, the charge strength is related to the number of photons hitting the detector during a short exposure time period. When the image capture period is ended, the charge information is read out sequentially from each photodiode. There are various schemes for reading out, but essentially the charge cloud of each photodiode is transferred one at time to an amplifier which measures the value of the charge and converts it into a voltage. Associated electronics then reduce noise, digitize the values, and output the digital value for

[1] All modern scientific-grade CCD and CMOS cameras are cooled to reduce noise and increase signal to noise levels.

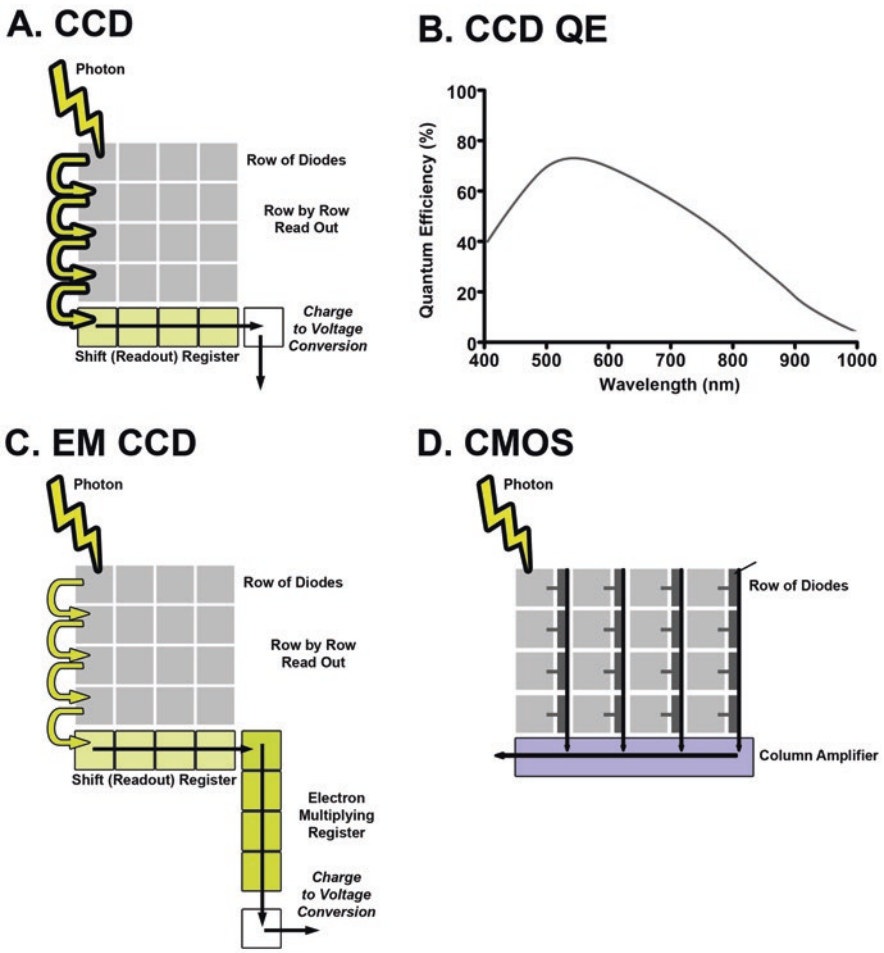

Fig. 7.10 CCD and CMOS design. (**a**) Design of a CCD camera. (**b**) Typical QE curve for scientific-grade cooled CCD. (**c**) Design of an EM-CCD camera. (**d**) Design of a CMOS camera. (*Based on Jerome. 2017. Microscopy Today doi:*https://doi.org/10.1017/S1551929511700061X. *Reprinted with permission*)

storage, usually in computer memory. CCD detectors have high QE across the visible spectrum (Fig. 7.10 b) and large dynamic range. Noise levels, however, can be high in very low light situations. For these situations, electron multiplying CCDs (EM-CCD) have been developed. They increase signal output without substantially increasing noise by passing the output charge through an electron multiplier consisting of multiple electrodes before analog to digital conversion (Fig. 7.10 c). The electrodes generate additional electrons by impact ionization, increasing the signal from 100- to 1,000 fold. However, the individual photodiode size in EM-CCD cameras is usually larger than that for CCDs, and so lateral (X-Y) resolution is less for

most EM-CCDs. CCDs are very serviceable and have been the backbone of microscope cameras for years. The main drawback of the CCD design, however, is that the sequential transfer takes time and thus limits how fast individual images can be acquired because all photodiode charge transfer processes must be completed before a new round of image collection can begin.

CMOS cameras also have a two-dimensional array of photodetectors that converts photons to electrical charge in a manner proportional to the photon energy absorbed. However, on a CMOS chip, the amplification and digitization steps happen at each individual sensor in parallel before the information is passed off the chip for storage (Fig. 7.10 d). This parallel processing approach dramatically reduces the time it takes to relay the image information and prepare the sensor array to capture another image. CMOS cameras have traditionally been the camera of choice for live cell imaging because their fast speed limits exposure of the specimen to the potentially damaging excitation beam. One of the tradeoffs for this increased speed, however, is that there is increased pixel-to-pixel variability since each pixel has its own separate amplifier circuit.

Even just 5 years ago, choosing between a CCD and CMOS camera was fairly simple. One chose a CCD for high resolution or a CMOS system if you needed fast acquisition to capture rapid events or limit exposure. With current developments, particularly advances in CMOS systems, the choices are not as clear cut. One needs to delve deeper into the specifics of each camera system in order to match a particular camera to your needs. Among the many parameters of camera design that can be investigated, we feel a few of these provide the most information regarding whether a system will meet your needs.

One of the key components to consider is the size of the individual photodiode sensors (pixels). This determines the spatial resolution limit of the sensor. For a given magnification, smaller photodiode size allows higher frequency information to be captured. The effect of pixel (photodiode) size is discussed in detail in the next section. Photodiode size, however, is not the only determinant of spatial resolution, so specific tests should be run to determine the highest spatial resolution the sensor is capable of obtaining. As individual photodiode size gets smaller, its photon capacity (how many photons before it becomes saturated) decreases. Thus, smaller pixels have lower potential SNRs and lower potential dynamic range (Chen et al. 2000). Today, most high resolution microscope cameras (both CCD and CMOS) with a photodiode size of ~6.5 um will provide full microscope resolution with good dynamic range and signal to noise characteristics. Table 7.1 compares some critical parameters of modern cooled CCD and scientific CMOS (sCMOS) cameras with pixel size of 6.5 um^2. CMOS systems have a little more noise than equivalent size CCD systems, but with the latest round of scientific CMOS (sCMOS) cameras, I have not found the difference to be very noticeable, except at very low light levels.

Resolution is often stated as the number of megapixels or the pixels per inch of a system. I find these parameters much less useful than the individual pixel dimensions. Megapixels and pixels per inch do not provide critical information regarding the smallest object in a sample that can be effectively identified and captured with the camera system. The number of megapixels is not totally useless as it does

Table 7.1 cCCD and sCMOS cameras: comparison of key parameters

Parameter	cCCD	sCMOS
Pixel size	6.5 um × 6.5 um	6.5 um × 6.5 um
Number of pixels	1392 × 1040	2048 × 2048
Peak QE	70% at 600 nm	72% at 560 nm
Readout speed	Up to 17 fps	Up to 100 fps
Readout noise	2e	2.5e
Dynamic Range	3,000:1	30,000:1

indicate how large a field of view can be captured, but it provides no information about the spatial resolution within that field of view. Pixels per inch is derived from the pixel size, so sensor size from pixels per inch can be calculated, but why not just look at the sensor size directly for comparing maximum achievable spatial resolution between different camera systems?

The spectral response of the camera refers to how well specific wavelengths are detected by the sensor. Both cooled CCD and high-end (scientific) CMOS sensors are sensitive to wavelengths between 400 nm and 1000 nm. However, one should always check the spectral response curve of the specific system being considered. There is usually some sensitivity in the near IR range, and CMOS sensors tend to be better at detecting IR compared to CCD sensors.

The frame rate of a camera refers to the number of full frames that can be captured per second. Digital camera systems allow adjustment of the frame rate up to some maximum. This is where the parallel processing design of CMOS cameras clearly has the edge. Maximum frame rates for CMOS systems are faster than for CCD systems, although some clever tricks have improved CCD frame capture rates. However, for both systems, since faster frame rates mean reduced exposure time, the SNR decreases as frame rate increases.

Quantum Efficiency (QE) is a measure of the efficiency of conversion of photons to charge. As with PMTs and hybrid detectors, the QE of photodiode sensors is not uniform across the spectrum. Both CCD and CMOS detectors tend to have less QE at both ends of their usable wavelengths. Several years ago, the QE of CMOS cameras were significantly worse than for CCD systems, but the most recent sCMOS designs rival the QE of CCD sensors. At their maximum, QE values range around 60%.

The dynamic range of a camera indicates how many separate and distinct gray levels the camera is capable of effectively detecting and storing. In other words, what is the lowest light level the sensor can detect and how many photons can a sensor collect before it is saturated. If there is a large difference between the lowest light detectable and saturation, there is the possibility to discriminate more subtle differences in the light coming from different areas of the specimen. For instance, if the lowest detection possible above noise is 5 and the sensor saturates at 50, then 5 defines what the sensor determines as black and 50 as white. That leaves the values of 6–49 to define shades of gray between black and white. So this sensor can discriminate 44 different charge levels and assign them to 44 different gray values. However, if the range is 2–200, more subtle differences in the number of photons

falling on one photodiode versus an adjacent one can be discriminated. Therefore, the greater the dynamic range the better the camera is at detecting subtle differences in shading within a specimen. The dynamic range is the full well capacity of the sensor divided by the read noise. The dynamic range is usually reported as how many bits of information can be discriminated. The dynamic range is different at different frame capture rates, so camera systems should be compared at the frame rate needed in requisite experiments. Dynamic range is another area where CCD systems used to excel, but sCMOS technology is rapidly catching up.

Read noise is the inherent electronic noise in the system. The camera noise level is important because it helps determine the dynamic range of the sensor and how efficiently the camera can capture very dim light. Read noise increases as frame-rate increases. Thus, even though CMOS cameras are capable of very high frame rates (100 frames per second), these rates are really only useful when a very strong signal is present so that the increased noise does not interfere with detection.

Dark noise is noise that is generated by temperature excitation of electrons within the photodetectors. Cooling a camera reduces the dark noise. For most situations, the dark noise of cooled CCDs is negligible, not so for cooled sCMOS cameras. However, dark noise in sCMOS cameras has been steadily going down as sCMOS manufacturers improve performance. It may come about that soon cooled sCMOS cameras will rival cooled CCDs in terms of dark noise.

Camera manufacturers provide a long list of specifications for their systems, especially for those parameters for which their products excel. However, we have found that the few parameters described above are sufficient to narrow down the potential list of cameras that are useful for a specific purpose. Then, despite what the specifications sheet indicates, one should always compare systems on their own microscopes with their specimens to make sure that the camera system is appropriate for their needs. This practical test is necessary because there is chip-to-chip variability and also because each manufacturer includes hardware and software in their camera that alters the initial photon signal collected by the photodiode. This signal processing is an integral part of the camera, but this means that two cameras from different manufacturers with the same CCD or CMOS chip can have a different level of performance.

7.3.3 Matching Microscope Image Lateral Resolution to Photodiode Size for cCCD and sCMOS

CCDs and CMOS systems have high QE, and their dynamic range and linearity are excellent. However, their sensitivity is not as good as PMTs, and the time to collect an image and transfer it to computer memory is slower than that of PMTs. However, because multiple photodiodes are arrayed together in a single camera, the overall frame capture rate can be faster than PMTs. The most important feature of photodiode array cameras that must be considered in microscopy is that, unlike PMTs, array photodiodes do have a physical size. This must be taken into account in order to optimize image resolution.

Table 7.2 Magnification of Airy disk radius by lenses with different NA and magnifications

Lens	NA	Airy disk diameter (micrometers)	Magnified Diameter (micrometers)
40× dry	0.93	0.3	12.0
63× water	1.3	0.2	12.6
100× water	1.3	0.2	20

In addition to considering the size of the photodiode in a CCD or CMOS system, matching the microscope image resolution to the photodiode size also involves knowledge of the resolving power of the microscope lens. The photodiode sensor size should be available from the camera manufacturer. A sensor size of 6.45 square micrometers is a popular size for high resolution microscope cameras. This turns out to be a very convenient size for microscopy.

From Eq. 7.5 we can determine that the resel obtained from an emitted beam of 550 nm focused with a lens having a NA of 1.3 (a pretty standard NA for the type of high resolution water immersion lens often used in confocal microscopy) will be a little less than 0.2 micrometers in diameter. We will round this number to 0.2 micrometers for the purposes of this discussion. If the lens had a magnifying power of 63 times, the smallest resolved objects (0.2 micrometers) would be projected onto the chip at a size of 12.6 micrometers (Table 7.2), and this is roughly twice the size of the sensor. A 40× dry lens with NA = 0.93 will resolve structures of approximately 0.3 micrometers which would then be magnified onto the chip at a size of 12.0 micrometers; Again, roughly twice the size of a single sensor.

The projection to twice the size of the sensor is important because of a key concept in sampling statistics called the Nyquist-Shannon theorem (Nyquist theorem, Nyquist criterion). This theorem is a fundamental principal in sampling theory. Obviously, the fidelity of the analog to digital conversion of signals, such as the irregular analog signal depicted in Fig. 7.11a, is dependent upon how small we set our sampling interval. In the case of analog to digital conversion, the sampling interval sets the quanta size. Figure 7.11c has a smaller sampling interval (digital quanta) than Fig. 7.11b and, thus, produces a better reproduction of the original analog signal. The Nyquist-Shannon theorem shows that an analog signal can be reasonably well reproduced if the smallest element (frequency) of the signal is sampled at least twice. This means we can only have confidence in structures that are at least twice as large as our sampling interval. In the case of a digital image reconstruction, Nyquist-Shannon tells us that our sampling probe (the photodiode elements of the camera) must have a diameter of half the size of the object in both the X and Y directions. Thus, a sensor of 6.45 micrometers is very close to the Nyquist criterion of 6.3 micrometers to sample a resel of 12.6 micrometers. In practice, the sampling interval is often set to sample the smallest detail 2.3× to be sure of capturing adequate information.

Table 7.2 also describes the magnified resel for a 100×, NA 1.3 water lens. You can see that a 6.45 micrometer sensor far exceeds the Nyquist criterion. We can take advantage of this fact and combine, or bin, the information from four adjacent sensors to make a sensor with a size 12.9 × 12.9 micrometers. This is again close to the Nyquist criterion (10 micrometers), although it slightly under samples the resel.

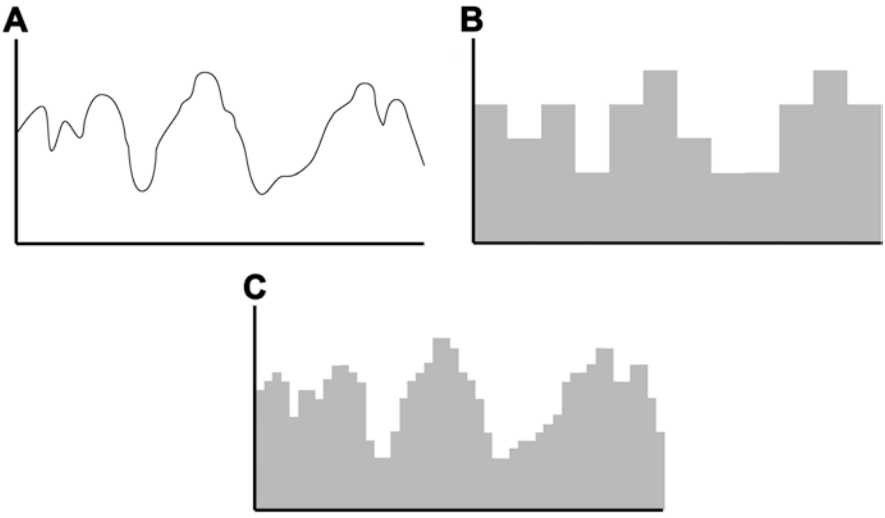

Fig. 7.11 Effect of digital sampling interval on fidelity of signal reproduction

Binning of CCD cameras has the advantage that the image acquisition time can be enhanced and the SNR decreased since we are collecting 4 times the signal with relative little decrease in image resolution (there would be no decrease in image resolution if we exactly matched the Nyquist criterion). Since CMOS cameras do processing at each photodiode, binning pixels of a CMOS camera does not increase readout time, but it does reduce signal to noise. We will have more on binning in Sect. 7.4.

Of course, all of these calculations do not take into consideration that in confocal microscopy, we are generally imaging a thick specimen. Interactions of both excitation and emitted photons with the specimen can reduce the resolution from that which is theoretically calculated. Moreover, there are other factors at play when using high NA lenses and laser beams are polarized, which can distort the point spread function (PSF) in one direction (see Sect. 7.5 for more information on the PSF). Despite these practical issues, it is always better to match as close as possible those parameters you can determine than to ignore physics and hope for the best.

7.3.4 Effect of Missampling

Incorrect sampling frequency can have significant effects on image quality. These are best illustrated by imaging a periodic test specimen, such as a grating with uniform spacing. Comparing the original grating to the image produced as depicted in Fig. 7.12, it is clear that sampling at the Nyquist frequency produces an accurate representation of the original. Oversampling adds no additional spatial information to the final image but does increase the exposure of the sample to a potentially

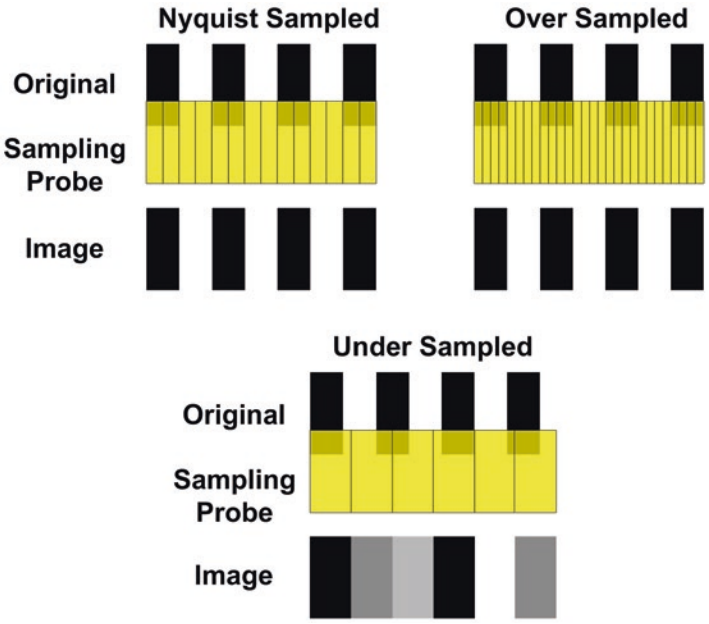

Fig. 7.12 Effect on image fidelity of under- and oversampling a specific frequency

harmful beam. This can exacerbate photobleaching and damage to the sample. In contrast, undersampling significantly alters the representation such that the original pattern is obscured in the image. If this were a real biological sample, this obscuring would blur the detail of structures smaller than twice the sampling frequency. A potentially more problematic effect of incorrect sampling frequency, however, is spatial aliasing. Spatial aliasing is the occurrence in an image of structural information that was not in the original object. This occurs when periodic structures are undersampled at a frequency that leads to misinterpretation of the structural periodicity. Moiré patterns (Fig. 7.13) are an example of spatial aliasing. Figure 7.13b shows a Moiré pattern formed when an image of a brick wall is collected at too low a sampling frequency. It is easy to spot such artefacts when viewing a known structure such as a brick wall. However, when these patterns arise in a biological sample of unknown structure, they can be misinterpreted as specimen fine detail. Spuriously jagged outlines are another example of spatial aliasing.

7.4 CDD Dynamic Range and Array Size

It should be clear by now that the diode size and correct sampling are very important in determining the resolution and quality of the final digital image produced by a CCD or CMOS camera connected to a confocal microscope. However, image

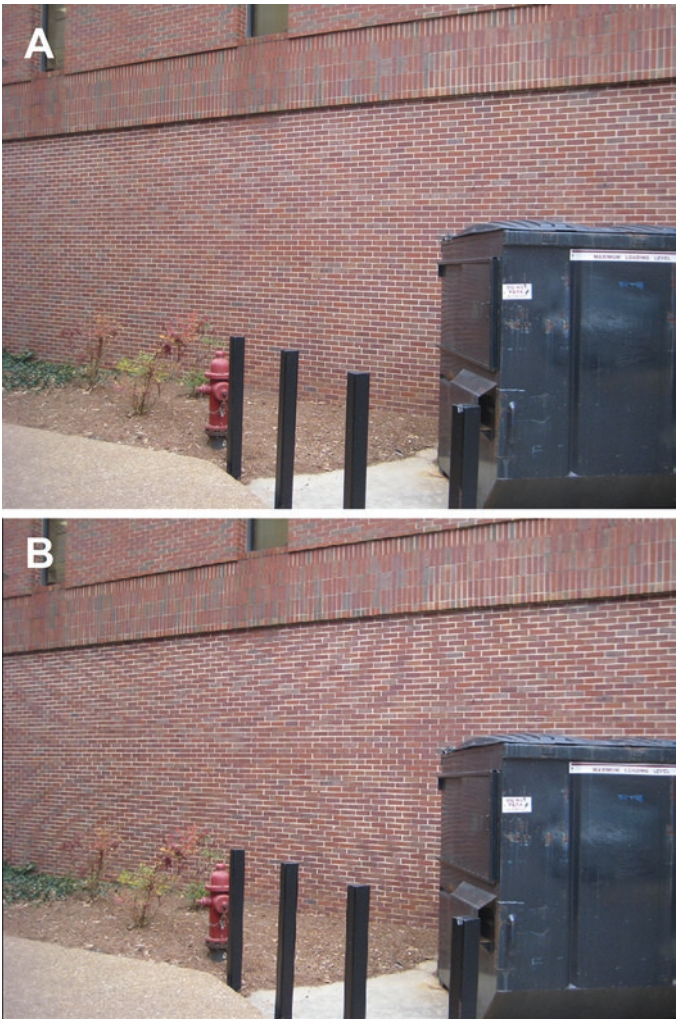

Fig. 7.13 Aliasing artifact caused by undersampling

contrast is just as important to the ability to recognize and record information from closely spaced objects. Thus, the dynamic range of the camera (number of gray values between black and white that the camera can accurately detect) is also critical. The dynamic range is affected primarily by the inherent noise level and the full well capacity of the sensor.

When photons hit the photodiode sensors of the camera, electrons are generated. The well capacity, or well depth, is the number of electrons that can be generated before the capacity of the sensor is exceeded. Obviously, large capacities are desired. However, the smaller the sensor size the smaller will be the capacity. As introduced above, the other major influence on dynamic range is readout noise. Scientific-grade

array cameras for microscopy are now cooled, and these, coupled with electronic advances and reproducibility of signal readout, bring the noise in these cameras to very low levels of <10 electrons per sensor at reasonable readout rates.

The dynamic range of a photodiode array camera can be determined as the full well capacity divided by the noise (SNR). A photodiode with a full well capacity of 41,000 electrons and a noise level per well of 10 electrons would be able to detect 4,100 different gray values. In other words, it could produce a 12-bit image (2^{12} = 4096 gray levels).

$$\text{Dynamic range} = \frac{41{,}000}{10} = 4{,}100\,\text{gray values}$$

Thus, even if the A to D conversion produced a 16-bit gray value (16 binary digits long), 4 of those digits would be spurious information because the sensor was only capable of an accuracy of 12 bits.

There is, however, a way of actually increasing the dynamic range of a photodiode array system. This is by binning several sensors together into a single virtual sensor. Since read out noise is not additive, binning of sensors produces huge increases in dynamic range. For instance, binning 4 pixels (2 sensors along the X direction and 2 along the Y direction) for the sensors in the example above makes the camera capable of detection of 16,400 gray levels.

$$\text{Dynamic range} = \frac{41{,}000 \times 2 \times 2}{10} = 16{,}400\,\text{gray values}$$

The array size (number of pixels in the X and Y direction) determines the size of the image that can be collected and thus the area of sample (field size) from which information can be collected for each image. Currently available CCD and CMOS cameras for scientific use have array sizes in the range of about 1,000 × 1,000 pixels to upwards of 5,000 × 5,000 pixels and larger. A 1,000 × 1,000 array of sensors with an individual sensor size of 6.45 micrometer would produce a camera that can capture an area 6,450 micrometers × 6,450 micrometers. With a 63× lens in place, this means the confocal image would contain information from 102 micrometers × 102 micrometers (10,404 square micrometers) of sample area. Larger cameras collect more sample information, but for CCD cameras, this comes at the expense of slower readout rates.

7.5 Image Capture in 3-D

The most important aspect of a confocal microscope for most biological studies is the ability to acquire high-resolution information about the three-dimensional architecture of the specimen. As discussed throughout this book, confocal microscopy isolates a defined planer thickness of the sample. In the case of a single slice, the

third dimension is the thickness of the optical slice. In cases where a larger volume is built up by adding together multiple slices, the thickness of an individual slice remains a critical component of the 3-D data set. With this in mind, we need to consider the characteristics of all three dimensions of a point source projected into a sample. Just as the lateral dimensions of the image of a point source are blurred to form an Airy pattern, the axial dimension of the image is also blurred. The degree of blurring is referred to as the axial PSF. The PSF describes how the optical system affects the image of a point source along the X, Y, and Z axis. The Airy disk is the lateral component of the PSF. In this section, we will consider the remaining axis: the optical axis (Z axis).

Confocal microscopy of fluorescent objects is an incoherent imaging system. This means that each point of information is collected independently, and if the system is designed and aligned well, there should be no interference between the light waves coming from one point and those from a different point. In other words, the light waves are independent. In the single-point scanning confocal, incoherence is assured because each point is temporally separated. In a multipoint scanning system, each point collected at a single time point is sufficiently spaced as to be optically spatially separate.

The image of a point source along the axial plane (the axial PSF) is an hourglass shape with the smallest point being at the focal point of the lens (Fig. 7.14a). By removing out of focus light, the confocal microscope converts the hourglass to an oblate spheroid (ellipsoid), as shown in Fig. 7.14b. Note that this means the PSF is not a sphere; the axial dimension is larger than the lateral dimension. Thus, the limit of resolution in the axial (Z) direction will not be as good as that in the lateral (X-Y) plane.

Like the Airy disk, the radius of the PSF along the optical axis is dependent on the illumination wavelength and the NA of the lens. Several equations have been generated to define confocal PSF resolution. Equation 7.6 describes the most often used equation and highlights the relationship between wavelength and NA when the pinhole aperture is set to exactly capture the full diameter of the Airy disk focused onto the aperture.

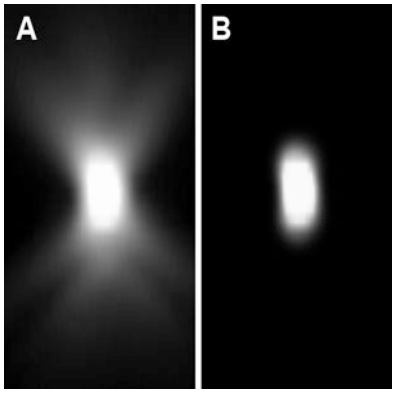

Fig. 7.14 Z axis of point spread function of a 0.15 um bead imaged by a microscope without a pinhole (**a**) and one with a pinhole set to 1 Airy unit (**b**)

$$d_{\text{axial}} = \frac{1.4\left(\lambda^{*} n\right)}{\text{NA}^{2}} \tag{7.6}$$

where: λ = wavelength
n = refractive index
NA = numerical aperture

Just as the Airy disk determines the lateral resolution, the PSF defines the volume resolution (3-D resel) and, therefore, determines the minimum digital voxel size that can be accurately represented in a 3-D digital image collected by a confocal microscope. Table 7.3 presents some practical dimensions for the lateral and axial radii of the PSF for a lens with NA 1.4 in a system using immersion oil with refractive index 1.515. It is worth producing similar tables for the lenses and wavelengths most often used for a confocal microscope and having the data close by the microscope for quick reference. Of course, the resel dimensions listed are the minimum dimensions obtainable. Unlike the lateral resolution, reducing the pinhole below one Airy unit produces little gain in axial resolution but does significantly decrease the number of photons passing through the aperture. In contrast, if one opens the pinhole of the microscope greater than one Airy unit, then more photons pass to the detector. This is because the photons passing are from the out-of-focus regions. This will blur the edges of objects and so will degrade both the lateral and axial resolution. However, opening the pinhole is useful when the signal is too low for efficient detection.

Since the PSF determines the volume being illuminated and imaged, an understanding of the PSF is important for determining how to correctly sample an object along the Z axis. The Z axis sampling rate is set by instructing the microscope how far to move the focal point along the Z axis between collection of each image plane (slice thickness). Taking into consideration the Nyquist criterion, in the optical system defined in Table 7.3 using light of wavelength 510 nm, the slice thickness should be set to 0.276 nm in order to correctly collect the details of the smallest resolvable structures in the sample volume. Luckily, the software that is provided with most modern confocal microscopes can automatically determine the volume being excited and detected based on the pinhole setting.

Often constraints such as acquisition time, fluorophore bleaching, and specimen beam sensitivity preclude one from sampling an object at theoretical resolution limit. For instance, it would take about 72 scans to collect the full volume of a 20-micrometer-thick specimen at a rate of 0.276 micrometers per slice (the frequency required for Nyquist sampling of the smallest resolvable structures). This could well be impractical, particularly if the fluorophore was prone to photobleach-

Table 7.3 Calculated X, Y, and Z dimensions for PSF at different wavelengths of an optical system with refractive index 1.515 and NA 1.4

Color	Wavelength	X and Y	Z
Blue	442 nm	0.129 micrometers	0.478 micrometers
Green	541 nm	0.158 micrometers	0.585 micrometers
Red	690 nm	0.202 micrometers	0.747 micrometers

ing. It is perfectly permissible to use a different sampling frequency, but the pinhole diameter should be adjusted accordingly. The ease of changing the pinhole diameter to match the situation is an advantage of laser scanning microscopes not shared by multipoint scanners.

There are multiple reasons why it is important to match the optical slice thickness to that which is actually being collected by the instrument during image formation. The first, and most obvious, is that if you image two very thin planes that are widely spaced apart, the resultant 3-D image will either be flatter than the specimen (thus distorting 3-D relationships) or the computer will have to fill in the empty space with interpolated information. A second reason to match the optical slice thickness to the collection frequency is that it avoids the possibility of aliasing artifacts. Since detail smaller than the optical slice thickness will not be resolved, very little axial resolution is gained by maintaining a smaller pinhole (although an improvement in lateral resolution is gained by increased contrast). Finally, by opening up the pinhole, a larger signal is collected which can dramatically improve the SNR of the image. All of the compromises made when selecting the pinhole size are addressed further in Chap. 9.

A critical point to recognize from Table 7.3 is that the slice thickness collected using blue excitation light (when the pinhole is set to one Airy unit for blue light) is not the same as when red light (with pinhole set to one Airy unit for red light) is used to excite the specimen. Thus, the axial information in the 3-D resel will not be equivalent. This is illustrated in Fig. 7.15. It compares the axial slice of a cell illuminated with blue light (A) or red light (B). The bright region at the tip of the excitation beam arrow indicates the expected axial PSF. If the resulting 3-D images of sequential blue scans and red scans through the cell are simply combined without consideration of the different depth of each, the software will distort one of the images to match the thickness of the other. This can introduce artifacts into the image. If you want to combine two images collected with different excitation wavelengths, the easiest solution is to set the pinhole to 1 Airy unit for the longer wavelength and then open the pinhole for the shorter wavelength so that the slice thickness matches that of the longer wavelength scan. Most modern point scanning confocal microscopes allow you to do this easily. Chapter 9 provides more details on this process.

The discussion above outlined the factors determining the theoretic minimum PSF. However, alignment errors, lens aberrations, refractive index mismatches or discrepancies, and contamination can increase the PSF or introduce asymmetries in the PSF. The true PSF will also include the interaction of the light waves with the

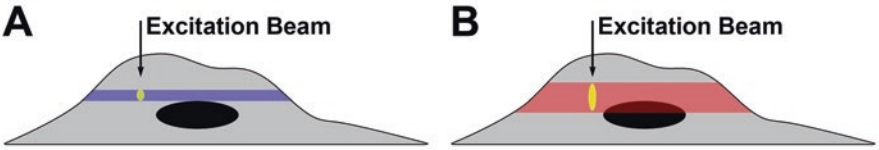

Fig. 7.15 Depiction of difference in axial (Z axis) resolution when a sample is scanned with a beam of blue light versus scanning with red light

sample. However, the effect of interaction with the biological sample is very difficult to predict. This is because the specimen content is rarely homogenous and the composition is usually changing with time. The specimen interaction generally causes further blurring of the image, and this effect becomes more pronounced as thickness of the specimen increases. Luckily, for quick calculations of optimal sampling frequency, we can normally ignore the effect of the sample and use the theoretical calculations based mostly on what we know about the objective lens. Our lack of knowledge about the true PSF does, however, indicate we must be aware that our image may contain artefactual information.

7.6 A Word About Color Image Capture: Don't!

The cameras for confocal microscopy are almost universally black and white cameras. These convert signal strength to a grayscale value within the dynamic range of the instrument. The image does not contain color information. Color information is provided separately by the use of specific lookup tables (LUTs) selected by the user. The color assignment is usually based on knowledge of the filters used to restrict the wavelengths reaching the detector. However, other colors could also be assigned. The green fluorescent protein signal could just as easily be displayed as yellow or red. Specifics of how this is done are provided in Chap. 9.

PMTs are inherently gray scaled and so color must always be assigned. On the other hand, color diode array cameras are available. At present, though, it is generally best to avoid these for high resolution work. The reason for this is that current color cameras acquire color information either at the expense of image lateral resolution, image dynamic range (because of degradation of the signal to noise ratio) or by greatly lengthening exposure time. As discussed earlier, any increase in the exposure of the sample to the beam increases the potential for specimen degradation.

To understand why current color diode array cameras are problematic for high resolution work, we need to discuss the various ways that CCD cameras can capture color information. The most common type of color CCD chip places an array of color filters in front of the photodiodes of the CCD. In this way, only narrow wavelengths of light get through to be detected by the photodiode. The most common type of color filter array is the Bayer filter. This places a red, green, or blue color filter in a specific pattern (Fig. 7.16) in front of each individual sensor. Each unit cell of the filter encompasses four sensors: two collecting green light and one each for red and blue. The color assigned to each pixel in the image is a mixture of red, green, and blue based on the levels of red, green, and blue detected by the sensors. Depending upon the algorithm used for interpolation and anti-aliasing, this binning of pixels reduces the lateral resolution by a factor of two to four. It should be noted that the hue and brightness value assigned each pixel is a blend of the information collected for that photodiode blended with the information collected for surrounding pixels. Another approach is to divide up the output signal into red, green, and blue components and detect them separately with three different detectors and to

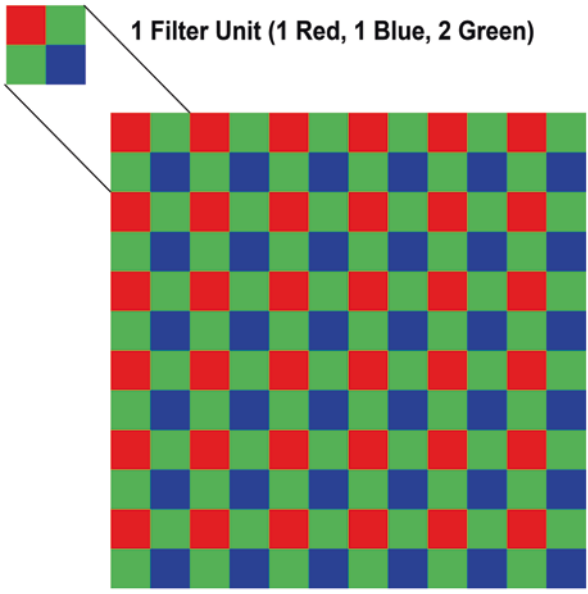

Fig. 7.16 Diagram of typical Bayer filter pattern present on a color CCD array

then recombine the image. This introduces additional elements into the light path that can degrade the signal.

A new development in color sensing arrays has been designed by Foveon (Foveon Inc., Santa Clara, CA). Their design uses three arrays stacked sequentially on top of each other with filters in between each array. Since red light penetrates deeper into silicon than blue light, the first layer captures the blue information. A filter prevents the blue light from traversing through to lower layers. In the same manner, the second array captures the green information, and the bottom layer captures the red information. This scheme allows the full resolution of the chip and avoids problems of post-collection alignment of the red, green, and blue images. So far this approach has only been implemented on photographic cameras and is not commercially available for confocal microscope cameras. However, the technology shows promise if it can be adapted for high resolution (small diode size) confocal microscopy work. For the reasons above, the best approach for confocal microscopy using diode array cameras remains sequentially placing filters in front of the CCD array. This allows the sequential collection of the red, then green, and then blue information separately. The three images can then be combined. This approach uses the full resolution of the camera. It is not without drawbacks, however. One must remember that full color information requires at least three separate exposures of the specimen and, thus, increases the possibility of photodamage. The exposure time for image collection could be reduced to reduce damage, but this increases the noise in the image. Sequential imaging also introduces issues of correct alignment of the three images that can reduce image resolution. Despite these drawbacks, the sequential approach remains the most useful that retains full resolution if full color imaging is required. Thankfully, most of the time, color information can be obtained or implied in other ways.

7.7 Conclusions

It should be clear, now, that there are numerous user selectable options that allow optimization of the final confocal image. This is extremely useful, but it does require diligence on the part of the user to assure image quality and fidelity. Optimization of the information in an image requires maximizing the spatial and contrast resolution in your image. Both the optical and digital restrictions on resolution must be considered as part of the optimization process. Failure to recognize the constraints and adapt the imaging procedure to these constraints at the least can result in loss of image information. Even more problematic, however, is that failure to recognize the constraints and adapt accordingly can lead to the inadvertent presentation of artefactual information as structural detail. Experimental and sample limitations may dictate that the theoretical optimal resolution is not feasible and that compromises must be made. It is imperative that any compromises are made rationally and with full knowledge of the consequences. For instance, it makes no sense to collect optical slices (movement of the focal plane between collecting individual image planes) at 1 micrometer intervals but have the pinhole set to resolve 0.5 micrometer slices. Rather, the pinhole should be set to a resolution along the Z axis which assures that the sampling frequency along the Z axis meets the Nyquist criterion. Without proper consideration of the various resolution constraints imposed by the sample, microscope, and detector, the data contained in an image could be incorrectly interpreted.

Literature Cited

Chen T, Catrysse P, Gamal A, Wandell B (2000) How small should a pixel size be? In: Blouke M, Sampat N, Williams G (eds) Proceedings SPIE, vol. 3965, Sensors and camera systems for scientific, industrial and digital photography applications. SPIE, pp 451–459

Demmerle J, Innocent C, North AJ, Ball G, Muller M, Miron E, Matsuda A, Dobbie IM, Markaki Y, Schermelleh L (2017) Strategy and practical guidelines for successful structured illumination microscopy. Nat Protoc 12:988–110

Heddleston JM, Chew T-L (2016) Light Sheet microscopes: novel imaging toolbox for visualizing life's processes. Int J Biochem Cell Biol 80:119–123

Lippincott-Schwartz J, Manley S (2009) Putting super-resolution fluorescence microscopy to work. Nat Methods 6:21–23

Magnan P (2003) Detection of visible photons on CCD and CMOS: a comparative view. Nucl Inst Methods Phys Res A 504:199–212

Sahl SJ, Hell SW, Jakobs S (2017) Flourescence nanoscopy in cell biology. Nat Rev Mol Cell Biol 18:685–701