Basics of Digital Microscopy

Optical microscopy is a well established tool in biological research. In modern microscopy, analog light detectors like the human eye or photographic film are replaced with a digital camera. Digital microscopy comprises image formation by optics, image registration by a digital camera, and saving of image data in a computer file. Making optimal use of digital microscopy requires taking into account limitations that are particular to each of these processes. The discussion presented here is limited to wide-field light microscopy. Confocal, multiphoton, and other specialized microscopy techniques are described in other units (see Chapter 2).

OPTICAL IMAGING

Resolution in the Spatial Domain

The laws of geometrical optics predict no limits on the size of the object that may be observed in a microscope. However, when light passes through an objective, diffraction occurs at the pupil owing to the wave nature of light. Thus, the image formed by a point light source at the focal plane is not a single point. Instead, the light intensity is distributed (spread) over an area on the image plane. This distribution is given by the point spread function (PSF). The pattern (Airy disc) features a central high-intensity spot and many concentric rings. The actual shape of the PSF (degree of spreading) depends on the wavelength of light used for imaging and on the numerical aperture (NA) of the microscope objective (Fig. 12.2.1).

Spatial Resolution Criteria

Spatial resolution refers to the ability to distinguish two small or point-size objects separated by a given distance. Resolution in optical microscopy is often assessed by means of the Rayleigh criterion (Rayleigh, 1879), which was originally formulated for determining the resolution of two-dimensional telescope images observed with the human eye. According to this criterion, two closely spaced Airy disks are considered distinct if the distance d between them is greater than the radius of the Airy disk (Fig. 12.2.1). This critical distance can be estimated using the formula:

$$d = 0.6100\ \lambda/\mathrm{NA} = 0.6100\lambda/n\sin(\alpha/2)$$

where λ is the imaging wavelength, α the full aperture angle, and n the light refraction

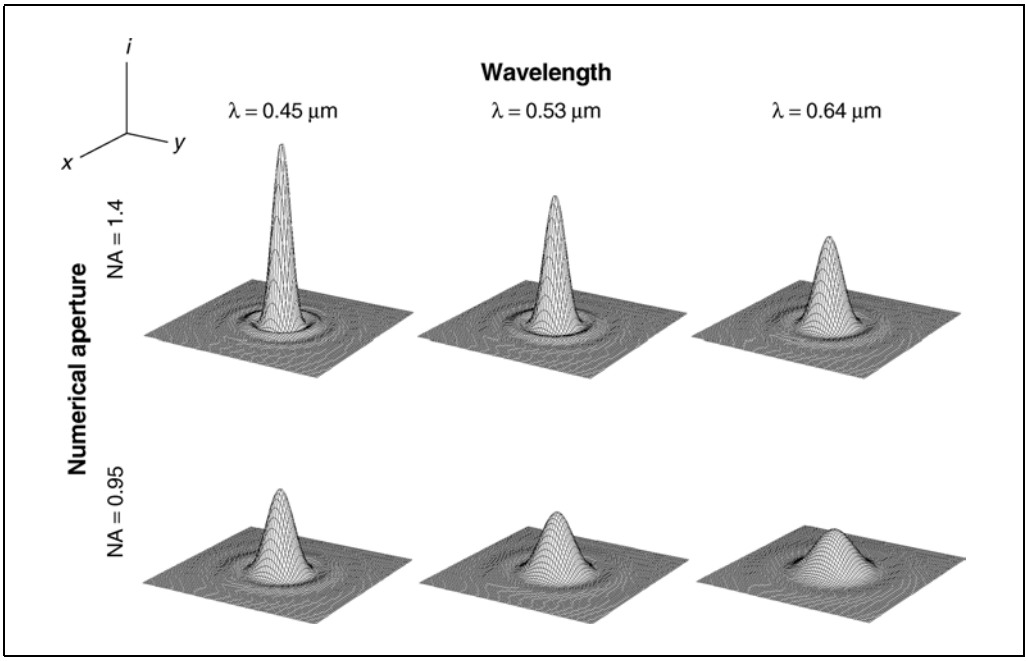

Figure 12.2.1 Dependence of microscope PSF on numerical aperture (NA, in rows) and wavelength (λ, in columns). Plot coordinates are indicated in the top left corner: x,y = position (from -0.75 to 0.75 μm); i = intensity. Width of the PSF decreases with increasing numerical aperture and increases with wavelength.

Contributed by Tytus Bernas

Current Protocols in Cytometry (2005) 12.2.1-12.2.14

Copyright © 2005 by John Wiley & Sons, Inc.

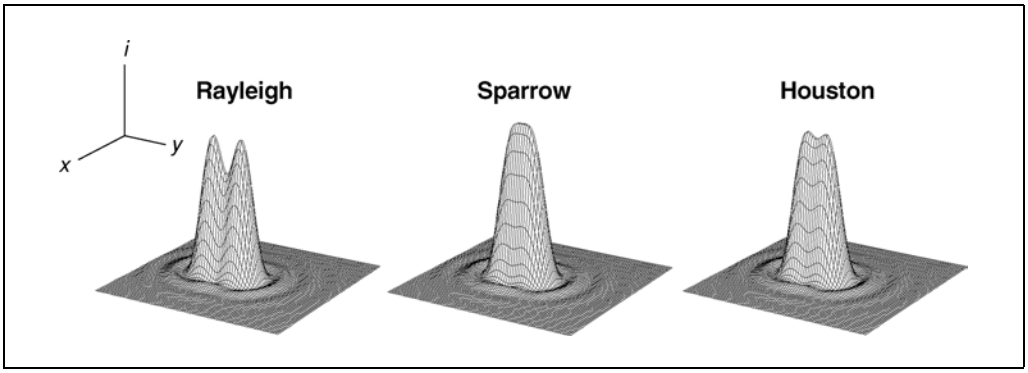

Figure 12.2.2 Summed image intensities (PSFs) of two points resolved according to (from left to right) Rayleigh, Sparrow, and FWHM (Houston) criteria. Plot coordinates are indicated in the top left corner: x,y = position (from −0.75 to 0.75 μm); i = intensity. Resolution (critical distance) depends on what intensity contrast is considered sufficient.

coefficient. It is assumed that the numerical apertures of condenser and objective are equal. Note that the smaller the value of d, the better the resolution.

Sparrow (1916) suggested an alternative, which depends on the property of the summed intensity profiles. As one moves two initially overlapping images of points apart, an intensity minimum develops. Sparrow suggested that resolution occurs at the line separation where this saddle point first develops (i.e., the gradient at the peak of the summed profile is zero). The critical distance for the Sparrow criterion is:

$$d = 0.4750\lambda/\mathrm{NA}$$

Another method is to adopt the size of a point image to define resolution. Houston (1926) suggested that two points may be resolved if the distance between them is equal to the width of the Airy disk measured at half the maximum intensity (full width half maximum, FWHM). Thus, the distance for the Houston criterion is given by:

$$d = 0.5015\lambda/\mathrm{NA}$$

Figure 12.2.2 shows a comparison of the three criteria. It is clear that the choice of resolution criterion depends on what intensity difference (between maxima and the separating minimum) is considered sufficient. These resolution criteria were based on the sensitivity of the eye of a human observer. In modern digital instruments, image-processing techniques are used to improve contrast (and thus perceived resolution) and to analyze images in an automated way. Hence, spatial resolution is usually described in the frequency domain using the Fourier transform of the PSF, the optical transfer function (OTF).

Resolution in the Frequency Domain

A description based on the OTF states that any "optical object" can be described as a weighted sum of sinusoidal distributions of light, where the sinusoids have varying spatial frequencies. The way an object "propagates" through an optical system depends on the OTF of that system. Thus, by knowing the OTF, one can also predict how the image of an object will look. Formally, the spatial frequency amplitudes of the image will be given by the spatial frequency amplitudes of the object multiplied by the OTF, which is itself a function of the spatial frequencies. The OTF is a two-dimensional Fourier transform of the PSF:

$$\begin{aligned}\mathrm{OTF}(f_x, f_y) &= \boldsymbol{F}\{\mathrm{PSF}(x,y)\} = \boldsymbol{F}\{\mathrm{PSF}(r)\} \\ &= \mathrm{OTF}(f_r)\end{aligned}$$

where $\boldsymbol{F}(x)$ is the two-dimensional Fourier transform operation, x,y are the spatial coordinates, r is the radial coordinate, f_x,f_y are the spatial frequencies in the x and y direction, respectively, and f_r is the radial spatial frequency. The OTF (shown in Fig. 12.2.3) is 0 for frequencies greater than the cutoff frequency given by the formula $f_c = 2\mathrm{NA}/\lambda$, where NA is the numerical aperture, λ the imaging wavelength, and, f_c the cutoff frequency.

In other words, spatial frequencies greater than f_c can not be transferred by a given optical system, which therefore is said to be band-limited. More detailed discussion of OTF and PSF can be found in *UNIT 2.6*.

DIGITAL REGISTRATION OF AN OPTICAL IMAGE

Architecture of the Light Sensor

The majority of light sensors (cameras) used in microscopy are based on silicon. This

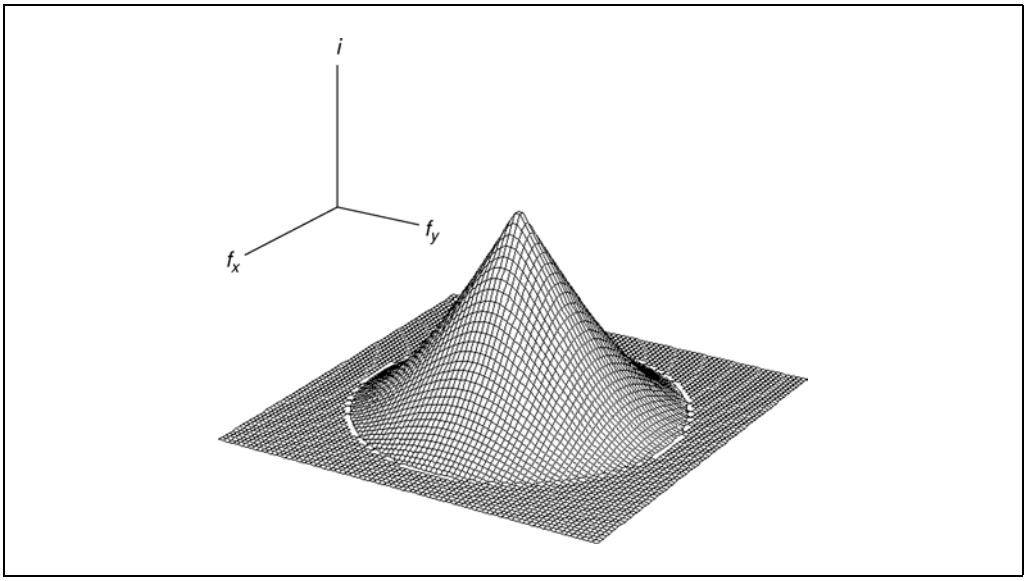

Figure 12.2.3 Optical transfer function (OTF) corresponding to PSF from Figure 12.2.1 (lower part, NA = 0.95; $\lambda = 0.53$ μm). Plot coordinates are indicated in the top left corner: f_x, f_y = spatial frequencies in x and y direction (from -4.4 to 4.4 μm^{-1}); i = intensity. OTF is 0 above cutoff frequency (i.e., is band-limited).

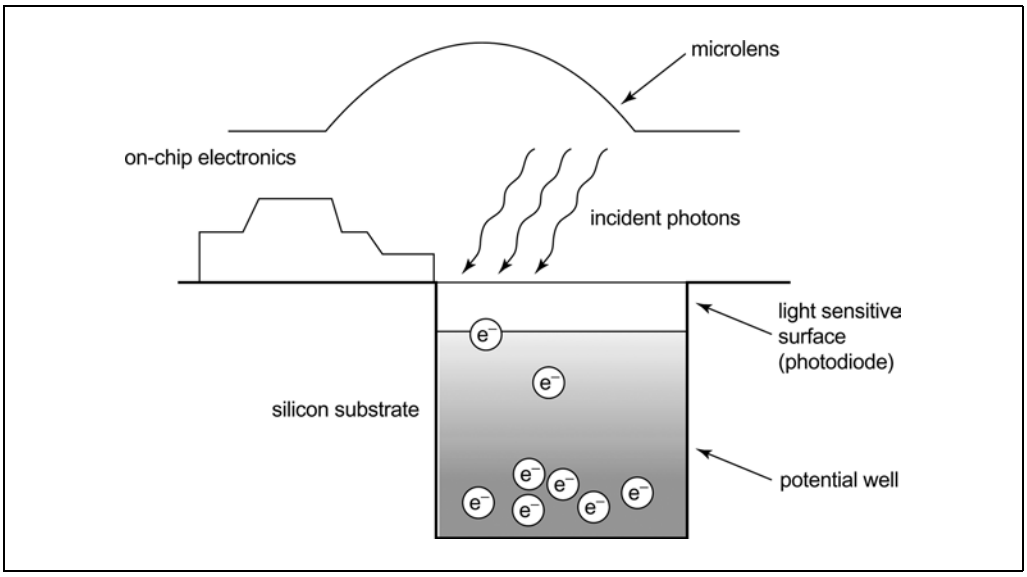

Figure 12.2.4 Schematic representation of a light-sensitive element (pixel) of a CMOS camera. The photons incident on the light-sensitive area and photoelectrons (e^-) trapped in the potential well are indicated.

element can form large crystals in which each atom is bound to its six neighbors, forming a rectangular, three-dimensional lattice structure. An incident photon may break one of these bonds, releasing an electron (photoelectric effect). A thin metal layer deposited on the surface of the silicon and charged with a positive voltage creates a potential well that collects and stores the electrons (Fig. 12.2.4). Each potential well corresponds to a light-sensitive element of the camera. Depending on the element design, several types of cameras are distinguished: CCD (charge-coupled device), CID (charge-injection device), CMOS APS (complementary metal-oxide semiconductor active pixel sensor), CMOS PPS (complementary metal-oxide semiconductor passive pixel sensor), and others. Nonetheless, all cameras considered here operate as a rectangular array of light-sensitive elements (chip). The performance of such a device is characterized by the signal-to-noise ratio, SNR, which determines the quality of measurement. SNR is usually expressed in decibels, dB, using the formula SNR = 20log{signal/noise}. The following section briefly discusses the sensor

parameters, which influence both terms of this quotient.

The amount of signal is limited by the well depth, which is a measure of how much charge an individual light-sensitive element on a camera chip can accumulate. This parameter is generally measured in electrons and depends on the fraction of the chip that is light sensitive (fill factor). Often, half or more of the available sensor area is covered by opaque charge-transfer circuitry, leaving gaps between the pixels and reducing the fill factor below the ideal of 100%. The chip can be coated with a thin layer of microlenses, each of which focuses the incoming light it receives onto the sensitive areas of one pixel. Well depth is proportional to the size of the element (pixel). For example, a Kodak KAF-3200E CCD chip (6.8-μm pixels) has a well depth of 55,000 electrons, whereas a Kodak KAF-6303E chip (9.0-μm pixels) features a well depth of 100,000. When the well depth is exceeded, electrons penetrate to adjacent wells, resulting in a bright streak extending vertically from a saturated spot. This effect, called blooming, is prevented if wells in a camera chip have drains to remove excess electrons. However, chips equipped with this antiblooming protection have much lower well depths, and thus are less sensitive than their counterparts without this feature. An area of adjacent pixels can be combined into one larger pixel in a process called binning. For instance, 2×2 binning means that the electric charge from 4 adjacent pixels is pooled together. This increases the sensitivity to light by a factor of 4. However, the effective width and height (in pixels) of the chip, the resolution, is correspondingly halved.

Only a fraction of incident photons are converted into electrons (photoelectrons), which are stored and then read out at the end of the exposure. The number of photons converted depends on the camera's quantum efficiency (QE), which is a function of wavelength. Standard cameras are most sensitive to green and red wavelengths in the region between 550 and 900 nm. Most midrange cameras have maximum QE in the range of 25% to 50%, whereas high-grade scientific CCDs may have a QE close to 98%.

All acquired images are contaminated by noise, a stochastic phenomenon that can neither be compensated for nor eliminated. The noise sources that play a role in scientific cameras are photon noise, thermal noise (dark current and hot pixels), readout noise (amplifier noise and on-chip electronic noise), and quantization noise.

Photon noise is a result of inherent variation in the arrival rate of photons incident on a camera chip owing to the quantum nature of light. The number of photoelectrons fluctuates randomly with photon incidence at each element (pixel) on a camera. Since the interval between photon arrivals is governed by Poisson statistics, the photon noise is equivalent to the square root of the signal. Thus, even if the photon noise were the only noise source, SNR would still be finite.

Dark noise arises from electrons thermally released from the silicon structure of a camera chip and subsequently accumulated by light-sensitive elements. Dark noise is not affected by incident light, but is highly dependent on device temperature. Cooling radically reduces the dark noise. The generation rate of thermal electrons at a given temperature is referred to as dark current. Like photon noise, dark noise follows Poisson statistics and is equivalent to the square root of the number of thermal electrons.

Readout noise originates in the process of reading the signal from the sensor and is caused by the on-chip electronics. This noise depends on the readout rate: it is inversely proportional to very low readout rates, approximately constant (and minimal) for moderate readout rates (20 to 500 kHz), and increases again for high readout rates. The readout noise is additive, Gaussian distributed, and independent of the signal. It is therefore expressed by its standard deviation (root mean square, rms) in number of electrons.

Quantization noise is a result of round-off errors caused by conversion of continuous values of electric charge accumulated by light-sensitive elements to a finite number of discrete intensity levels. This noise is additive, uniformly distributed, and independent of the signal. Quantization noise increases as the number of levels (bit depth) of the analog-to-digital converter (ADC) decreases. However, even for an 8-bit ADC (minimum for a scientific camera) the noise does not exceed 0.5 electrons per pixel. Thus, quantization noise is usually ignored.

Overall system signal-to-noise is commonly calculated using the following formula for camera system signal-to-noise ratio:

$$\mathrm{SNR} = PQ_e t/[PQ_e t + Dt + N_r^2\ ]^{1/2}$$

where P is the incident photon flux (photons/pixel/sec), Q_e represents the CCD quantum efficiency, t is the integration time (sec), D is the dark current value (electrons/pixel/sec), and N_r represents readout noise

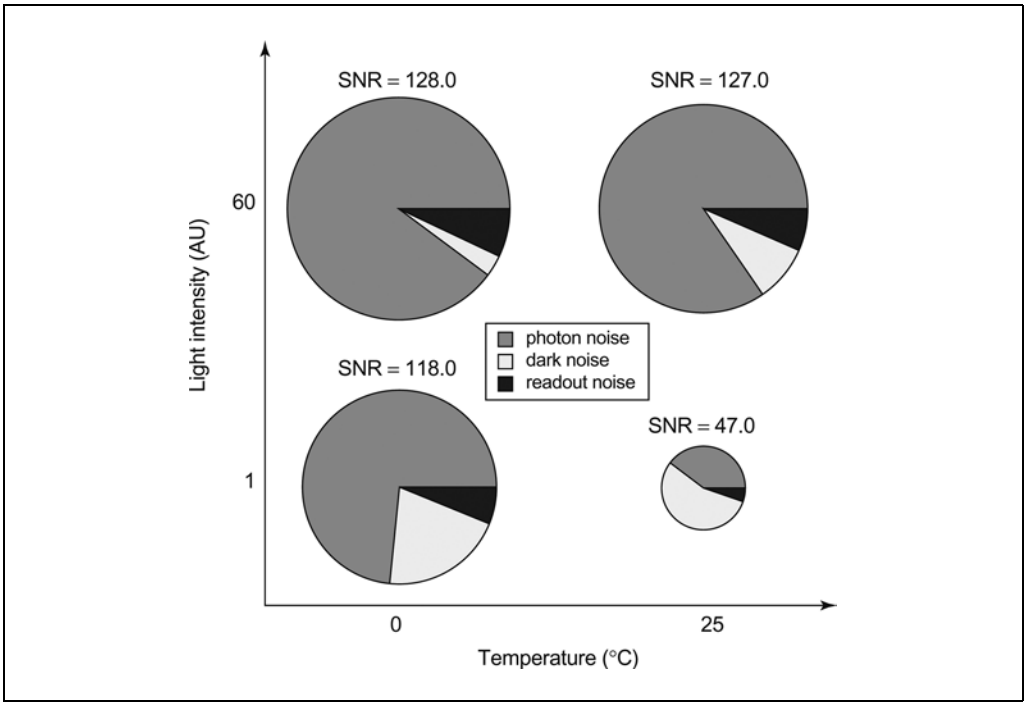

Figure 12.2.5 Influence of camera chip temperature (*x* axis, °C) and incident light intensity (*y* axis, arbitrary units) on different components of camera noise. Signal-to-noise ratio (SNR) is indicated by pie size, whereas relative contributions of photon, readout, and dark noise are depicted by pie slices. The parameters were estimated for a Sony ICX085 chip and full well capacity.

[rms, $(\text{electrons/pixel})^{1/2}$]. Influence of the photon flux and the dark current on SNR is illustrated in Figure 12.2.5.

Color Versus Monochrome Cameras

The operation of camera sensors discussed so far relies on measurement of total light intensity, but not wavelength spectrum. Thus, only monochrome images are obtained. Registration of color images requires measurement of light intensity at several wavelengths. Typically red, green, and blue ranges are chosen, to mimic human vision. Two main types of color cameras are used: either a single sensor with a wavelength-selection filter or a three-sensor system. Single-sensor cameras utilize a set of three filters, a single liquid-crystal tunable filter, or an adherent filter matrix to register the red, green, and blue images. When the set of filters (placed in a slide or filter wheel) or the tunable filter is used, the three images are obtained in sequence. High transmission (>95% in the selected range) is an advantage of using three separate filters. However, changing the filters is a relatively slow process and may be a source of vibrations. Tunable filters (Fig. 12.2.6), while free from these disadvantages, are characterized by relatively low maximum transmittance, which does not exceed 50% with unpolarized incident light (Reichman, 2000). The adherent filter matrix differs from systems described previously in that each light-sensitive element is coupled with its own filter (red, green, or blue; see Fig. 12.2.7). This design combines high transmittance with speed, as the color image is acquired in one pass. However, actual sampling intervals and resolution are decreased in comparison with the same sensor without the filter matrix. Software interpolation is used to match component images and increase the number of pixels. Some color cameras employ a piezo-controlled translocation mechanism to increase the sampling frequency, at the cost of reduced sensitivity or increased acquisition time. An interesting multilayer color sensor (the Foveon X3; Fig. 12.2.8) was recently introduced to ameliorate the problem of decrease in spatial frequency. This sensor features three layers of light-sensitive material (blue, green, and red), a construction which permits simultaneous registration of three colors with full resolution. However, its quantum efficiency is lower (<50%) than that of the majority of scientific-grade cameras of classic construction. Simultaneous acquisition of three colors with high sensitivity and resolution is possible using a three-sensor camera, which has a beam-splitting prism and trim

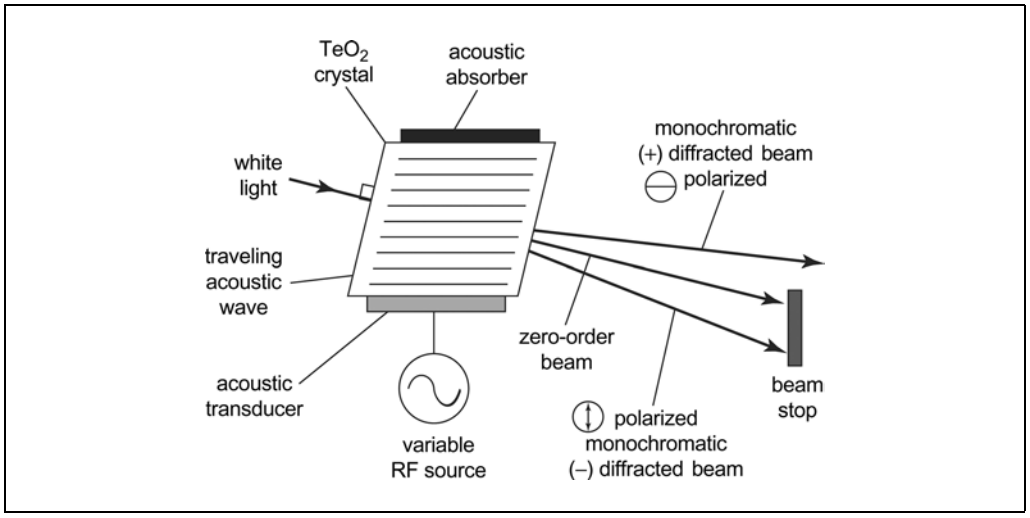

Figure 12.2.6 Schematic representation of acousto-optical tunable filter (from Brimrose Corp.).

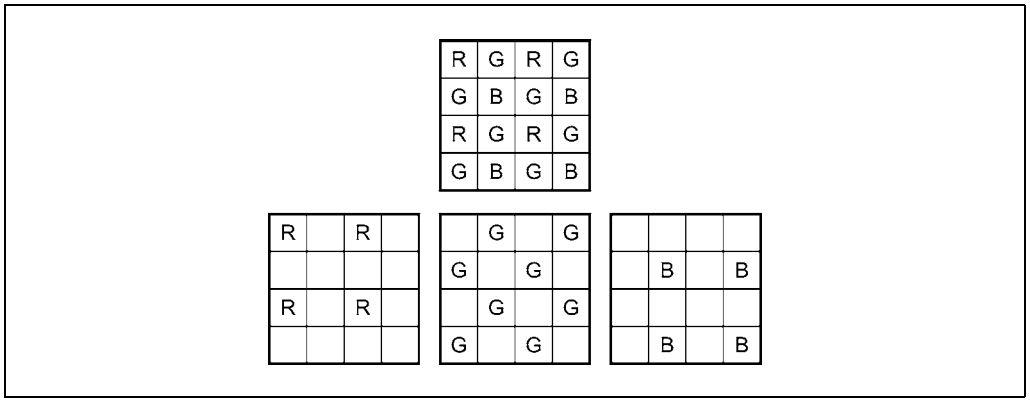

Figure 12.2.7 Adherent filter matrix (bayer pattern) comprising red (R), green (G), and blue (B) filters. Top row: the whole matrix. Bottom row: decomposition into color components. Increased sampling intervals in the component images result in lower resolution.

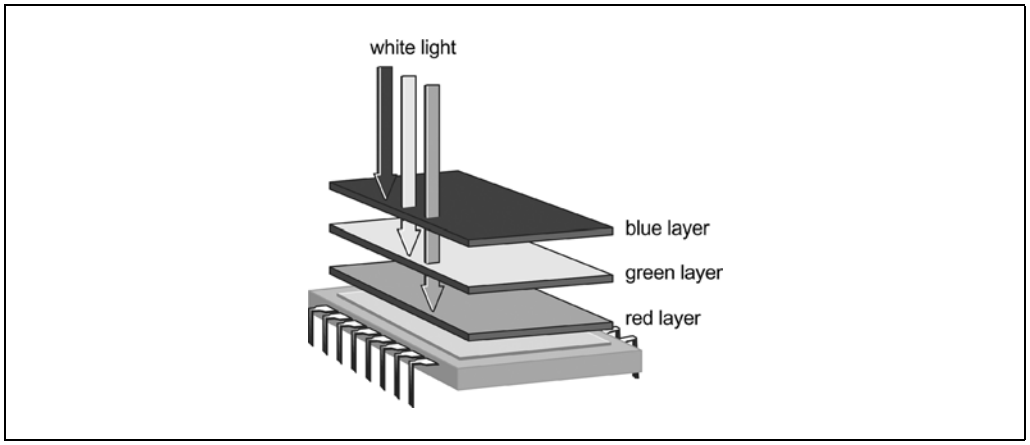

Figure 12.2.8 Diagram of a Foveon X3 chip (From Foveon Inc.; *http://www.foveon.com*), which filters the color components (RGB) by wavelength-dependent absorption via silicon layers.

filters that enable each sensor to image the appropriate color. However, these cameras are far more expensive than single-sensor ones. In general, color cameras are less sensitive than their monochrome counterparts because of the additional beam-splitting and wavelength selection components. Nonetheless, this disadvantage may be offset by the ability to image a specimen at multiple wavelengths simultaneously.

Table 12.2.1 Microscope Magnification Requirements for Maximum Optical Resolution ($\lambda = 0.515$ μm)

Objective (NA)	Nyquist sampling interval in the object plane (μm)	Sampling interval in the image plane when no relay lens is used	Relay lens needed (total magnification), 6.8-μm pixels	Relay lens needed (total magnification), 9-μm pixels
10× (0.45)	0.286	2.86	2.4× (24×)	3.2× (32×)
20× (0.75)	0.172	3.44	2.0× (40×)	2.6× (52×)
40× (0.75)	0.172	6.88	1.0× (40×)	1.3× (52×)
40× (1.15)	0.112	4.48	1.5× (60×)	2.0× (80×)
60× (0.95)	0.136	8.16	0.8× (50×)	1.1× (66×)
60× (1.40)	0.092	5.52	1.2× (74×)	1.6× (97×)
100× (1.25)	0.103	10.3	0.7× (66×)	0.9× (87×)
100× (1.40)	0.092	9.2	0.8× (74×)	1.0 (97×)

Sampling and Quantization

The microscopic image of a specimen is continuous with respect to x and y coordinates (space) and intensity (amplitude). Digital imaging relies on the projection of such an image onto a 2-D array of light-sensitive elements. The output of each element (electric charge accumulated in the potential well) is proportional to the incident light intensity. Hence, the intensity is measured (probed) at a finite number of points (pixels) in the process called sampling. Similarly, the intensity is converted to a discrete scale (usually integer-based) during quantization. Accurate transition between optical and digital images requires that sensor characteristics be properly matched to the optical system performance. Rules for such optimization are discussed briefly in the following section.

Spatial Sampling

A fundamental principle of sampling is expressed by the so-called Whittaker-Shannon sampling theorem, which is useful for understanding the information loss resulting from discrete sampling. The theorem states that if a two-dimensional function $f(x,y)$ is band-limited to spatial frequencies below f_{cx} cycles per unit value of x (in the x direction), and to frequencies below f_{cy} cycles per unit value of y (in the y direction), then the function can be completely reconstructed by taking four $f_{cx}f_{cy}$ samples per unit area on the x,y plane. In the absence of noise, f_{cx} and f_{cy} are equal to the OTF cutoff frequency. Thus, the maximum sampling interval (in the x and y directions of the object plane) is given by the Nyquist criterion, $D_{so} \leq 0.25\lambda/\text{NA}$.

Since the optical system (microscope) projects a magnified image of the object onto the sensor array, D_s in the image plane is given by $D_{si} \leq (0.25\lambda/\text{NA}) \times M$ where M is the total magnification.

An alternative approach involves digitizing with a sampling interval that is no greater than one-half the size of the smallest resolvable feature of the optical image. If this method is adopted, $D_{so} = 0.5d$ and depends on the resolution criterion. Hence, D_s in the image plane is given by $D_{so} \leq 0.5d \times M$.

The pixel size of the sensor should not exceed D_{so} in order to achieve optimum sampling. This optimal sampling is usually obtained by varying the total magnification (M) of the microscope using a relay lens (see examples in Table 12.2.1). When the sampling frequency is lower than the Nyquist limit, then frequencies in the original signal that are higher than half the sampling rate are aliased (i.e., appear in the image as lower spatial frequencies). Conversely, if too many pixels are gathered by the camera, no additional spatial information is afforded, and the image is oversampled (see Fig. 12.2.9). The extra pixels theoretically do not contribute to the spatial resolution, but can often help improve the accuracy of feature measurements taken from a digital image (Young, 1996).

Quantization

After an object has been imaged and sampled, each of the continuous intensities represented within the specimen is converted into a digital brightness value (level). The accuracy of the digital value is directly proportional to the bit depth of the digitizing device. For

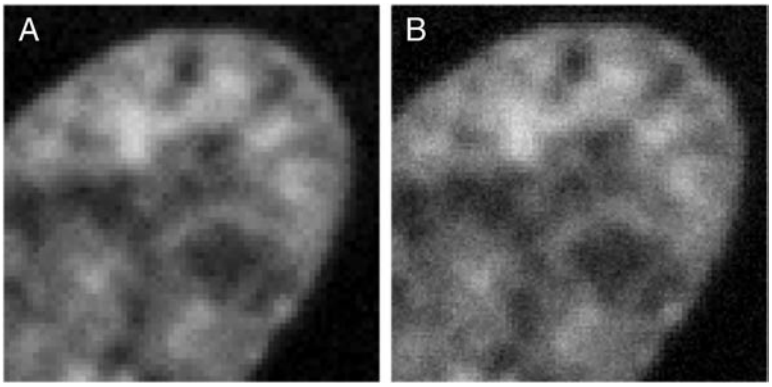

Figure 12.2.9 Image of AO-stained fibroblast nucleus sampled optimally at the Nyquist rate (2x binning, **A**) and oversampled two times (no binning, **B**).

instance, if one bit is utilized, the image can represent only two brightness levels. Four or eight bits are sufficient to express 16 or 256 levels, respectively. The number of bits needed to represent a signal accurately depends on the SNR of the sensor:

$$20\text{Log}(N_{\text{bit}}) \geq \text{SNR}.$$

The SNR of the sensor depends on the imaging conditions (see the equation above for calculation of camera system overall SNR). The bit depth should be chosen so that the maximum SNR can be accurately represented. The upper limit of precision may be estimated by $\text{SNR}_{\text{max}} =$ well depth/readout noise. However, owing to the presence of dark and photon noise, this maximum SNR is not achieved in practice. On the other hand, obtaining real *N*-bit data requires more than simply the use of an *N*-bit ADC: illumination stability, freedom from digital electronic interference, and proper bandwidth for sampling rate are additional considerations. Hence the above equation gives a safe estimate of necessary bit depth.

The quantization scheme discussed so far is linear; it assumes that all the brightness levels are separated by the same value (minimal noise). However, in the case of scientific-grade CCD signals, degradation comes mainly through photon noise, which is proportional to $\sqrt{N}$. In such a system, if one demands that brightness levels be separated by an amount at least equal to their standard deviation, the difference between the first two such levels (1 events/pixel and 4 events/pixel) is 3 events/pixel, whereas that between the 15th and 16th levels is 31 events/pixel. In other words, the number of "meaningful" gray levels is proportional to $\sqrt{N}$. Hence, only $0.5N_{\text{bit}}$ is necessary to preserve the information.

The resolution limit of an optical microscope can be extended by deconvolution (Holmes et al., 1995). However, the presence of noise results in a decrease of cutoff frequency (Stelzer, 1997) and imposes a limit on deconvolution efficiency (Holmes and Liu, 1991; Holmes, 1992). Several methods for estimating the achievable resolution have been developed (van Dekker and Bos, 1997). As SNR is affected by pixel size, larger pixels (and sampling below optimum rate) may bring an increase in actual resolution when SNR is very low (Neifeld, 1998).

STORAGE OF DIGITAL IMAGE DATA

Following sampling and quantization, digital images are stored as files in permanent memory (hard drives, CDs, DVDs, Flash ROM modules, and others). This section contains an overview of standards of color coding and compression, as these factors have significant impact on the fidelity and efficiency of digital image storage. A description of the most popular file formats is included as well.

Color Spaces

Color spaces are schemes to code colors using components. Although several are found in digital applications (Gonzalez and Wood, 2002), images are usually stored using RGB, CMYK, or YUV.

RGB (Red-Green-Blue) is a method of generating colors in a video system (monitors and cameras) that uses the additive primaries method. Percentages of red (R), blue (B), and green (G) are mixed to form the colors. Zero percent of the colors creates black, 100 percent of the colors creates white.

CMYK (or sometimes YMCK) is a color model used in color printing and is based on mixing pigments of the following colors: cyan (C), magenta (M), yellow (Y), and black (K, Key). The mixture of ideal CMY color is subtractive (cyan, magenta, and yellow printed together on white result in black). CMYK works through light absorption. The colors that are seen are from the wavelengths of light that are not absorbed. In CMYK, magenta plus yellow produces red, magenta plus cyan makes blue, cyan plus yellow generates green, and the combination of cyan, magenta, and yellow forms black. Because the "black" generated by mixing the subtractive primaries is not as dense as that of a genuine black ink (one that absorbs throughout the visible spectrum), four-color printing uses black in addition to the three primaries.

YUV (also known as YCbCr and YPbPr) is a color space in which the Y stands for the luminance (brightness) component and U (Cb, Pb) and V (Cr, Pr) are chrominance (color) components. It is commonly used in video applications, where it is also referred to as component video. YUV signals are created from an original RGB source. The weighted values of R, G, and B are added together to produce a single Y signal, representing the overall brightness, or luminance, of that spot. The U signal is then created by subtracting the Y from the blue signal of the original RGB, and V by subtracting the Y from the red.

Compression Algorithms

An array of algorithms have been developed in order to decrease the space occupied by a stored image file. These algorithms differ with respect to compression efficiency, speed, and complexity. Lossless algorithms introduce no distortion, in contrast to lossy compression schemes. However, the former offer smaller compression ratios than the latter.

LZW is a lossless compression algorithm named after its developers, Lempel and Ziv, with later modifications by Welch. Typically, LZW compresses image files to about one-half their original size. Compression ratios as high as 5:1 are also obtainable when the image has long runs or a lot of solid-color areas. LZW is referred to as a substitutional or dictionary-based encoding algorithm. The algorithm builds a data dictionary (also called a translation table or string table) of data occurring in an uncompressed data stream. Patterns of data (substrings) are identified in the data stream and are matched to entries in the dictionary. If the substring is not present in the dictionary, a code phrase is created based on the data content of the substring, and is stored in the dictionary. The phrase is then written to the compressed output stream. When a reoccurrence of a substring is identified in the data, the phrase of the substring already stored in the dictionary is written to the output. Because the phrase value has a physical size that is smaller than the substring it represents, data compression is achieved.

RLE (run-length encoding) is a lossless data compression algorithm that is supported by most image bitmap file formats. RLE works by reducing the physical size of a repeating string of characters. This repeating string, called a run, is typically encoded into two bytes. The first byte represents the number of characters in the run and is called the run count. In practice, an encoded run may contain 1 to 128 or 256 characters; the run count is usually expressed as the number of characters minus one (a value in the range of 0 to 127 or 255). The second byte is the value of the character in the run, which is in the range of 0 to 255, and is called the run value. RLE schemes are simple and fast, but their compression efficiency depends on the type of image data being encoded. A black-and-white image (two intensity levels) that is mostly white, such as the page of a book, will encode well (with ratios of 6:1 or better), owing to the large amount of contiguous identical data. An image with multiple intensity levels will not encode efficiently (ratios lower than 2:1 may be expected).

Huffman encoding is a simple compression algorithm introduced by David Huffman in 1952 and supported today by several imaging file formats. This lossless algorithm uses a predefined dictionary of commonly occurring image byte patterns (strings), which are given low (short) indices (codes) in the dictionary. Data are encoded by replacing each image string that occurs in the dictionary with its index number. The dictionary is not part of the compressed file. Compression efficiency is dependent on whether the dictionary entries encode common byte patterns occurring in a particular image. Therefore, ratios between 1.3 and 2.5 are most common in practice.

Deflation is a lossless compression algorithm that uses a combination of LZ77 (which is the basis of LZW) and Huffman coding to achieve better compression than is possible with either one alone. It was originally defined by Phillip W. Katz for his version of ZIP (PKZIP).

JPEG is a lossy compression scheme introduced by the Joint Photographic Experts Group to compress photographs. This scheme is based on the discrete cosine transform (DCT) and divided into the following stages:

1. Transformation of the image into an optimal color space. The best compression ratios are achieved when a luminance/chrominance color space, such as YUV, is used.
2. Down-sampling (up to 2× horizontal and vertical) of chrominance components (U and V) by averaging groups of pixels together.
3. Transformation with DCT performed on 8×8 blocks of pixels to give 64 frequency coefficients for a block.
4. Quantization of DCT coefficients in each block using weighting functions optimized for the human eye. Each of the 64 positions of the DCT block has its own coefficient, with the higher-order (frequency) terms being quantized more heavily than the lower-order (frequency) terms (that is, the higher-order terms have larger quantization coefficients). Furthermore, separate quantization tables are employed for luminance and chrominance data, with the chrominance data being quantized more heavily than the luminance data. This allows JPEG to discard data exploiting further differing sensitivity of the human eye to luminance and chrominance. It is this step that is controlled by the "quality" setting of most JPEG compressors.
5. Encoding the resulting coefficients (image data) using a Huffman variable-word-length algorithm to remove redundancies in the coefficients.

JPEG may achieve high (15:1 or better) compression ratios. The majority of JPEG encoders allow adjustment of the ratios via a setting for desired "image quality." It should be emphasized that artifacts (distortions) are introduced even with the lowest possible compression (highest quality).

JPEG2000 is a new image compression standard being developed by the Joint Photographic Experts Group. It is designed for different types of still images (bi-level, gray-level, color, multicomponent) and provides high-quality images at low bit rates, overcoming many of the limitations of the original JPEG standard. This scheme is based on the discrete wavelet transform (DWT) and divided into the following stages:

1. Splitting of the image into a set of tiles (a tile can occupy the whole image). Optional transformation into YUV (luminance/chrominance) color space.
2. Decomposition of the image with DWT into the low-resolution (low-pass) and high-resolution (high-pass) component images (sub-bands).
3. Scalar quantization of the DWT coefficients. This is a lossy step unless integer DWT was used.
4. Grouping of DWT coefficients from corresponding sub-bands. Decomposition of the grouped coefficients into bit planes.
5. Truncation of the resulting bit streams to achieve demanded quality (distortion). This is a lossy step.
6. Compression with a lossless algorithm (e.g., Huffman coding).

JPEG2000 may compress images in a lossless manner (ratios up to 2:1). Better compression (ratios comparable to JPEG) may be obtained at the cost of introducing distortions. Nonetheless the distortions are less severe compared with those produced in JPEG.

File Formats

Implementation of various color-coding and compression schemes has led to several image file formats. The choice of format depends on the type and intended use of the image data. This section includes a short overview of the most common formats.

BMP (basic multilingual plane, bitmap) is the standard Microsoft Windows raster file format, which is not recognized by other computer systems (with the exception of a few programs). The format supports pixels represented by 1 bit (monochrome image), 4 bits (16 possible values), or 8 bits (256 values). In this case, a color table is used to assign displayed colors to these pixel values. The colors are encoded in an RGB scheme. Alternatively, pixels may be represented by 24 bits. In this case, each RGB component is represented by three 8-bit segments (true color). Thus, no color table is used. The format supports lossless (run-length) compression for 4- or 8-bit variants.

GIF (graphics interchange format) is a service mark of CompuServe Incorporated introduced in 1987 (GIF87a). In 1989 a revised specification (GIF89a) added some features to the format, namely the capacity to store animations and textual comments. Pixels are encoded using 8 bits (256 possible values). A color

table is used to assign displayed color (RGB scheme) to a pixel value. The colors include binary transparency. Every pixel in the graphic that has the value assigned as transparent becomes invisible. The GIF file format also supports interlacing, i.e., saving a file using four passes instead of just one. On each pass, only certain lines of the image are saved to the file. To display an interlaced image, GIF loads progressively so that a lower-resolution preview can be seen before the full image is shown. The format supports lossless compression with the LZW algorithm. However, since the algorithm was patented (by Unisys) several programs write only uncompressed files (which nonetheless are compliant with LZW specification).

JPEG implements the lossy JPEG compression algorithm. The format supports pixel encoding using only 24 bits divided into three color components (8 bits each). These are luminance (Y) and two chrominance components (U and V, respectively). A JPEG file can be saved as a progressive JPEG, which is very similar to the interlaced GIF. As with GIF, this presents a low-quality image first, and over several passes improves the quality. JPEG does not support transparency, however.

JPEG2000 was created by the Joint Photographic Experts Group committee with the intention of superseding their original JPEG standard. The JPEG2000 algorithm may be used to obtain either lossy or lossless compression. The specification supports up to 16,384 components with 38-bit precision each. However, currently pixels are encoded as 1-, 4-, or 8-bit indexed color (with color table), 8- or 16-bit grayscale, 24-bit RGB (3 color components), or 32-bit CMYK (4 color components). Some implementations offer the alpha (transparency) channel (component).

PNG (portable network graphics) is a relatively new image format that is becoming popular on the World Wide Web and elsewhere. The format was developed largely to deal with some of the shortcomings of GIF. PNG stores image data using 1,2, 4, or 8 (PNG8) bits per pixel and a color table to assign displayed color (RGB scheme) to a pixel value. The table may be omitted in 8-bit images, which are saved as grayscale (i.e., with pixel values representing intensities only but not colors). Alternatively, pixels may be encoded using 24 bits to represent the three components of RGB color (PNG24). Both grayscale and RGB PNGs can have 16-bit precision, that is, 16-bit and 48-bit pixels, respectively. Furthermore, both can also have an 8-bit alpha channel to represent 256 levels of transparency for a pixel. Other image attributes that can be stored in PNG files include gamma values (see Gonzalez and Woods, 2002), background color, and textual information. PNG uses a lossless compression algorithm known as deflation.

Tagged Image File Format (abbreviated TIFF) was created by Aldus for use with PostScript printing. Now controlled by Adobe, TIFF has become the standard graphics format for high-bit-depth (32-bit) graphics, and can be directly manipulated using PostScript. TIFF features an array of options that can be used to include all sorts of image formats in the file. The actual characteristics of a given TIFF file, including simple geometry of the image, data arrangement, and compression type, are defined using specific tags in the file header. Pixels represent indexed colors (with color table) or grayscale values (without the table) with 1, 4, or 8 bits. The format also supports full-color images in which color is stored using components encoded with 8- or 16-bit precision. The components may belong to RGB, YUV, or CMYK color spaces. A separate alpha-component value may be assigned to pixels to describe transparency. Therefore, TIFF permits using up to 64 bits per pixel. The format may be used as a container for images compressed in a lossless manner with Huffman, RLE, LZW, or ZIP (deflate) algorithms. Lossy compression is achieved with JPEG. Nonetheless, despite the flexibility of this format, the vast majority of TIFF files, and the code that reads them, are based on a simple 32-bit (RGB and alpha) uncompressed image.

Table 12.2.2 summarizes the main features of file formats discussed in this unit. File formats are compared in Figure 12.2.10 as well. One may note that most efficient lossless compression algorithms are implemented in JPEG2000, TIFF, and PNG. The first format also permits lossy compression with relatively few distortions. However, JPEG2000 is not yet widely supported, having been introduced only recently. The basic standard of TIFF, on the other hand, is widely supported and highly versatile. However, owing to their complex structure, TIFFs tend to be rather large. PNG lies in between as far as size and support are concerned.

CONCLUSIONS

The usefulness of digital microscopy depends on whether or not image registration and transfer into digital form are performed in an optimal way. The first step requires finding a proper balance between maximizing sampling

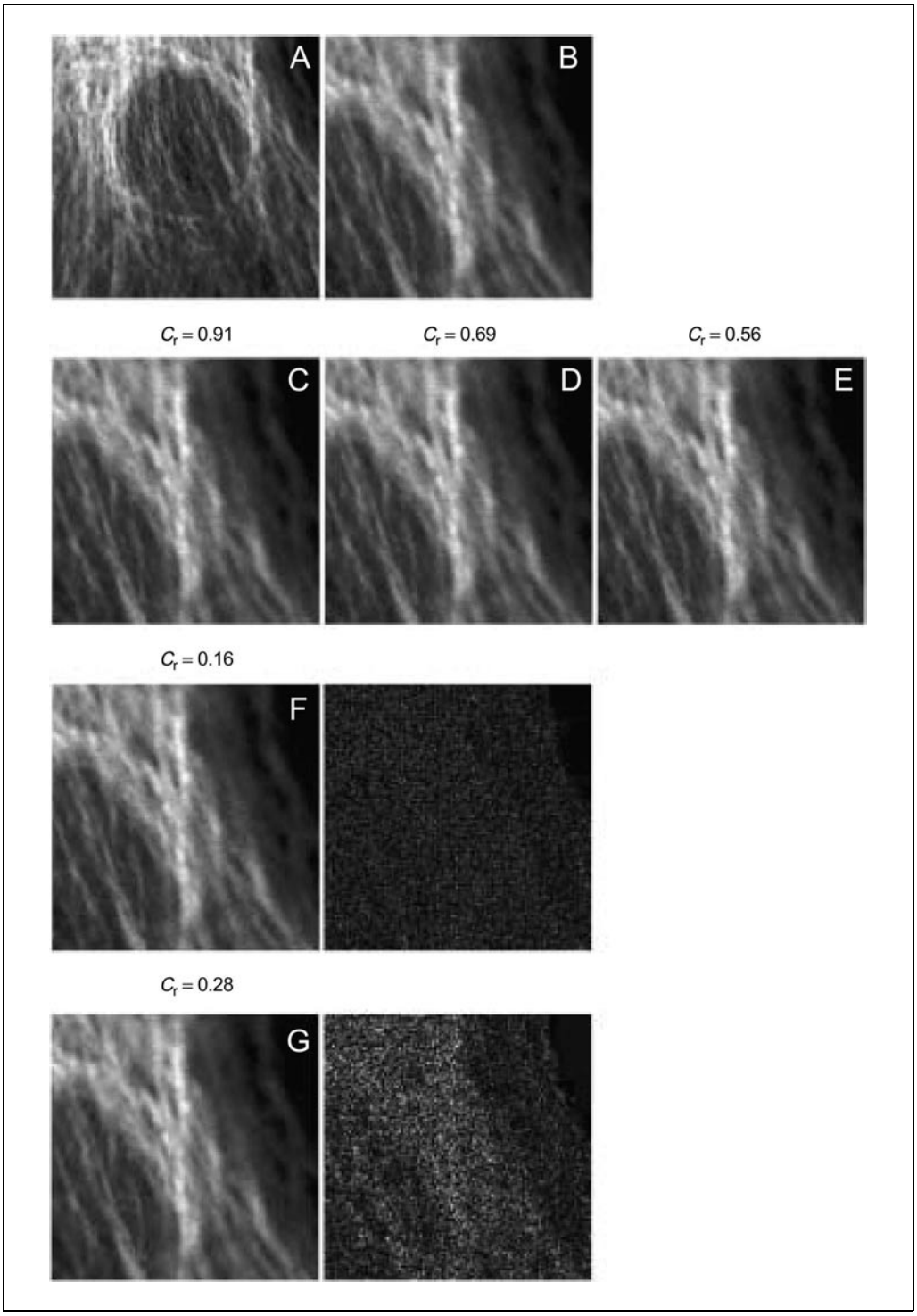

Figure 12.2.10 Comparison of image file formats. (**A**) Original image of tubulin fibers in FITC-stained fibroblast and (**B**) its magnified fragment. This image (stored as an uncompressed TIFF, 8-bit grayscale) was used as the standard for calculation of the compression ratios (C_r). The image was compressed using: (**C**) lossless LZW (TIFF); (**D**) lossless deflation (PNG); JPEG2000 using either lossless algorithm (**E**) or lossy algorithm (**F**); and (**G**) lossy JPEG (JPG). Compression errors (lossy JPEG2000 and JPEG) are illustrated using difference images (pixel values were multiplied by 8).

Table 12.2.2 Comparison of Common Image File Formats

File type	Bit depth	Color coding	Transparency (bits)	Compression		Portability	Remarks
				Lossless	Lossy		
BMP	1/4/8	Grayscale	—	RLE	—	Poor (Microsoft format)	Inefficient compression[a]
	24	RGB	—	—	—	—	—
GIF	8	RGB color table	1	LZW	—	Excellent	No grayscale support
JPEG	24	YUV	—	—	JPEG	Good	Very efficient compression[a] (with distortion). No RGB and grayscale support.
JPEG2000	1/4/8	Color table	38[b]	JPEG200	JPEG200	Fair	Very efficient compression.[a] Relatively new format intended to replace JPEG.
	8/16	Grayscale	—	—	—	—	—
	24/48	RGB	—	—	—	—	—
	36/64	CMYK	—	—	—	—	—
PNG	1/2/4/8	Grayscale or RGB color table	8	Deflation	—	Good	Efficient compression. Format developed to replace GIF on the World Wide Web.
	24/48	RGB	—	—	—	—	—
TIFF	1/4/8	RGB color table	8	Deflation (ZIP), RLE, Huffman, LZW	JPEG	Good	Format offers large array of features, which may be extended further using specific tags. Only a limited number of these are supported universally. Efficient compression.[a]
	1/4/8/16	Grayscale	—	—	—	—	—
	24/48	RGB, YUV	—	—	—	—	—
	32/64	CMYK	—	—	—	—	—

[a]Microscopic images.

[b]As specified by the standard. However, this feature is currently not supported by any program.

density (and thus precision) and SNR (and accuracy). If the balance is found, then imaging is performed with maximum attainable resolution. The second step requires matching characteristics of the image file with generated image data so that the file occupies minimum space and can be easily transferred with no loss of necessary information.

LITERATURE CITED

Gonzalez, R.C. and Woods, R.F. 2002. Digital Image Processing, 2nd ed. Prentice Hall, Upper Saddle River, N.J.

Holmes, T.J. 1992. Blind deconvolution of quantum-limited incoherent imagery: Maximum-likelihood approach. *J. Opt. Soc. Amer.* A-9:1052-1061.

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

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

Houston, W. V. 1926. The fine structure and wavelength of the Balmer Lines. *Astrophys. J.* 64:81.

Neifeld, M.A. 1998. Information, resolution and space-bandwidth product. *Opt. Lett.* 15:1477-1479.

Rayleigh, J.W.S. 1879. Investigations in Optics with special reference to the Spectroscope. *Phil. Mag.* 8:261-274.

Reichman, J. 2000. Handbook of Optical Filters for Fluorescence Microscopy. Chroma Technology Corp., Rockingham, Vt.

Sparrow, C. M. 1916. On spectroscopic resolving power. *Astrophys. J.* 44:76.

Stelzer, E.H.K. 1997. Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: Fundamental limits of resolution in fluorescence light microscopy. *J. Microsc.* 189:15-24.

Van Dekker, A.J. and Bos, A. 1997. Resolution: A survey. *J. Opt. Soc. Amer.* A-14:547-557.

Young, I.T. 1996. Quantitative microscopy. *IEEE Eng. Med. Biol.* 15:59-66.

Contributed by Tytus Bernas
Purdue University
West Lafayette, Indiana