Choosing a digital camera for your microscope

John C. Russ, Materials Science and Engineering Dept., North Carolina State Univ., Raleigh, NC

One vital step is to choose a transfer lens matched to your sensor or chip size. Digital cameras come with chip sizes that range from the size of a 35 mm film negative (36 mm x 24 mm) down to consumer cameras with stated sizes such as "1/3" or "1/1.8" inch. Note that these descriptions are not actual dimensions! The "1/3" chips are actually 4.8 x 3.6 mm (and there are smaller sizes out there). Matching a high-quality single lens reflex body such as the Nikon D80 or Canon EOS 400 to a camera may be an option, given its other potential uses in the lab, and reasonable cost, but it is important to understand a few things about all digital cameras before making a choice. The D80, as an example, has a 10 million photosites (called individual sensors below, since the word "pixel" has too many different meanings in various contexts) each of which is 6 μm square. By comparison, in a 1/1.8 chip with 4 million photosites each would be 3 μm square. (These values also allow for a space of about one half μm between the individual sensors to isolate them electrically, and this applies to CCD detectors, not CMOS chips which also require two or three transistors for each photosite, which also take up space). Figure 1 illustrates the relative sizes of different detectors.

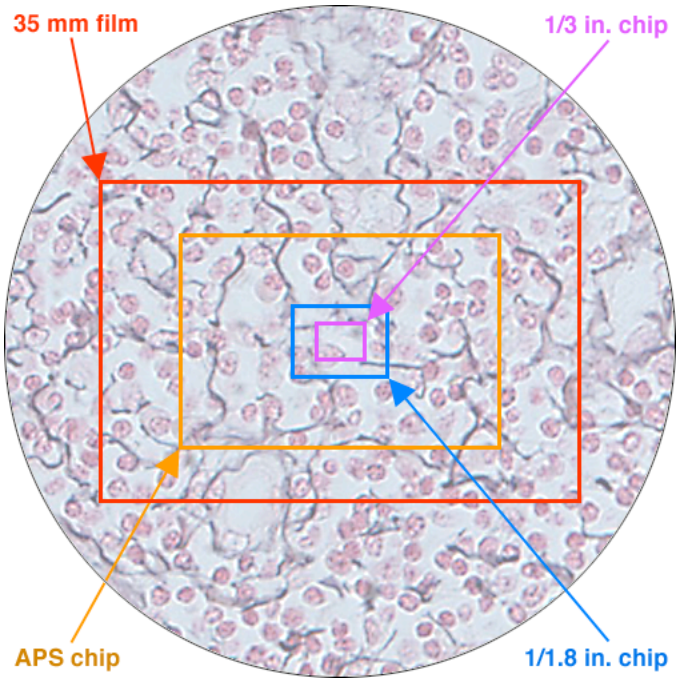

Figure 1. Relative sizes of a 25 mm negative, APS chip (used in digital SLR's such as the Nikon D80), and the "1/1.8" and "1.3" inch chips used in typical consumer cameras.

The role of the transfer lens is to project the image onto the chip. Normally the rectangular area that is captured is set well inside the circular field of the microscope, to avoid focus and illumination variations near the edge. In my rather typical setup, with a 10x objective the captured image field is about 1600 μm wide, and with a 100x objective it is about 160 μm wide. We'll use those numbers again shortly.

The maximum resolution of a digitized image is defined by the Nyquist limit as the spacing corresponding to two pixels (not one, since there must be a dark pixel separating two light ones, or vice versa, for them to be distinguished). For a 3600x2400 sensor chip, typical of a high end 10 Megapixel single-lens-reflex camera, this corresponds to 13 μm on the chip, and for a 10x objective represents 0.9 μm on the specimen. This is, of course, better than the optical resolution of that objective lens. With a 100x objective lens, assuming an optical resolution of 0.5 μm on the specimen, the same chip would represent that 0.5 μm distance with 45 digitized pixels, which is a vast oversampling.

There is no reason to use a camera with that many sensors. The Polaroid DMC camera, no longer supported but one of the good early digital microscope cameras that I used for years, had 1600x1200 photosites (each 7 μm wide) on a 12.15 mm wide chip. With the same 100x objective and an appropriate transfer lens, this still covers 0.5 μm on the specimen with a more than adequate 5 pixels, and with the 10x objective the camera's resolution limit would be 2 μm, better than the optics.

So, given that there is no need for huge numbers of photosites ("pixels" in camera advertising) for a camera to mount on the microscope, is there a disadvantage to using something like a Nikon D80? Not necessarily. The digitized images are large, but storage, computer memory and processing speed have advanced so quickly that these are not important factors. But there IS a major reason not to use one of the cameras with a small chip size. As the size of the individual sensors is reduced, each photodiode has a smaller capacity to capture photons and hold electrons. The larger detector sizes in an SLR chip or specialized scientific camera hold more electrons and produce a greater dynamic range for the image.

Digitization of the image produces a minimum amount of noise (from statistical, thermal and electronic sources) that corresponds to about 3-5 electrons (Peltier cooled cameras represent the lower end of this range). As the light intensity increases, it is the statistics of producing electrons as photons are absorbed that controls the amount of random noise. The signal to noise ratio in the final image is dependent on the maximum number of electrons that the diode can hold without saturation, which is directly proportional to size. Photodiodes in silicon have maximum capacities (the “well size”) of about 1000-1500 electrons per square micrometer, so doubling the width produces a 4x increase in area, collects more photons to produce more electrons, for a resulting improvement in maximum signal. Since the minimum remains the same, the dynamic range in increased and with it the signal to noise ratio.

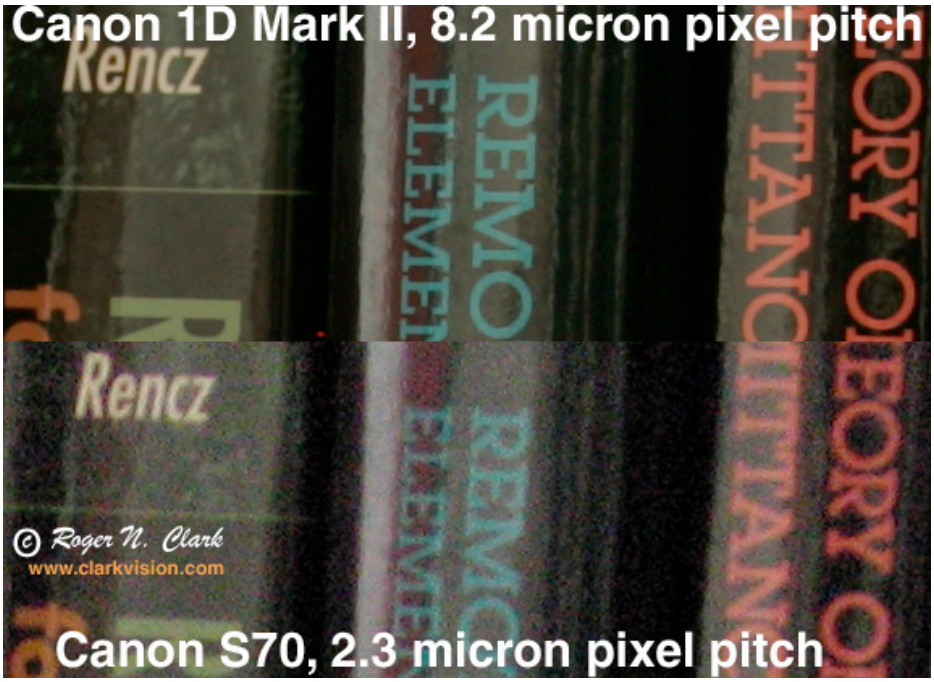

Figure 2. Comparison of the noise level in otherwise identical images acquired with chips having large and small photosites.

The small chip cameras struggle to produce 8 bits (representing 256 linear intensity values - this value is somewhat optimistic and is based on the organization of memory into 8-bit bytes rather than on detector perform-

ance) whereas a large chip camera can easily capture 12 bits (representing more than 4000 values). This large dynamic range is important for images with bright and dark areas, particularly fluorescence microscopy, astronomy, and everyday photography (a typical scene can easily include bright and shadow areas that vary in absolute brightness by 1000:1). The dynamic range provided by 12 bits is roughly comparable to the performance of film cameras, which is why digital SLR cameras are now used by many professional photographers as good replacements for film cameras (the spatial resolution of a 10 MP camera is also similar to that of 35 mm film).

The examples shown in Figure 2 are two photos of exactly the same scene, with the same illumination and the same camera settings, using cameras with large and small sensor sizes. The difference in noise level is apparent.

The quality of the chips used in digital camera chips comprises many characteristics, one of which is variations in sensitivity of the individual sites (called fixed pattern noise). There are also likely to be a few dead (zero output) or locked (maximum output) diodes, whose signals are usually removed by in-camera processing (a median filter, for example, replaces extreme values with one from a neighboring sensor). CMOS chips have more non-uniformity, more noise, more locked pixels, lower sensitivity (because of the area used for the transistors) and other problems as compared to CCD chips. At least for the foreseeable future CCDs are the better choice for scientific applications, but CMOS development continues because the devices are less expensive to produce and require less power, which is important in some applications such as cell phone cameras.

Color detection in digital cameras may be accomplished in several different ways. Early microscope cameras used a single chip with a filter wheel that captured three images, one each for red, green and blue. This method allowed, among other things, for balancing the time for each exposure to compensate for the variation in sensitivity of silicon photodiodes with wavelength. But such cameras were slow to operate and did not provide a live color image.

Another method uses prisms and filters that direct the red, green and blue light to three separate detector chips. These cameras are expensive and delicate (the alignment of the chips is critical), are less sensitive (because of the light lost in the optics), and produce color shading with wide angle lenses.

Still another technique developed by Foveon fabricates a single chip in which there are three diodes at each photosite, one above the other. Red light penetrates deeper into silicon than blue, so the diode at the surface measures the amount of blue light while filtering it out. The next diode below it does the same for green, while the red light penetrates to the deepest layer. This method is still somewhat experimental, has thus far been accomplished only using CMOS technology, and has some problems with the processing needed to obtain accurate colors, but has the advantage that every photosite receives the full color information.

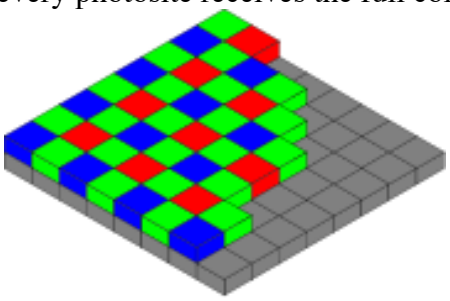

Figure 3. The Bayer color filter array in which one-half of the photosites have a green filter, one quarter have a red filter, and one quarter a blue filter.

That is not the case with the most common method of color detection, a color filter array (CFA). The most commonly used arrangement is the Bayer filter which places red, green and blue filters on individual photodiodes in a pattern as shown in Figure 3. Since with this scheme each sensor measures only the corresponding

color of light, the camera electronics must perform some rather clever interpolation (using proprietary "demosaicing" algorithms) to estimate the color values at each location. This process is made somewhat more complicated by the fact that the filters each pass broad and somewhat overlapping ranges of wavelengths. The interpolation, along with the use of an "antialiasing" filter placed in front of the detector chip to scatter the light and spread the photons out to several sensors, reduces the resolution of the digitized image to about 50% of the value that might be expected from the number of "pixels" advertised in the camera specification! (Note - there are some camera manufacturers who carry this interpolation much farther, and advertise a number of pixels in the resulting stored image that is many times the number of sensors on the chip.)

Silicon photodiodes are inherently linear, so that the output signal from the chip is a direct linear response to the intensity of the incident light. Professional cameras that save the “raw” information from the chip (which because of the demoisaicing and removal of shot noise is usually not quite the direct photodiode output) do indeed record this linear response. But programs that interpret the raw file data to produce viewable images, usually saved as tiff or jpeg formats, apply nonlinear conversions. One reason for this practice is to make the images more like those produced by traditional film cameras, since film responds to light intensity logarithmically rather than linearly.

It is also done because of the nature of human vision, which responds to changes in intensity more-or-less logarithmically. In a typical scene, a brightness change of a few percent is noticeable. A difference between 20 and 40 (on the 0=black to 255=white scale used by many programs) represents a 100 percent change, while the same absolute difference between 220 and 240 is less than a 10% change.

In film photography, images may be recorded on either “hard” or “soft” films and paper. As shown in Figure 4, these are characterized by plots of the density against the log of the incident light intensity that are very steep, producing a high contrast image, or more gradual, covering a greater range of intensity. The slope of the central linear portion of the curve is called the “gamma” of the film or paper.

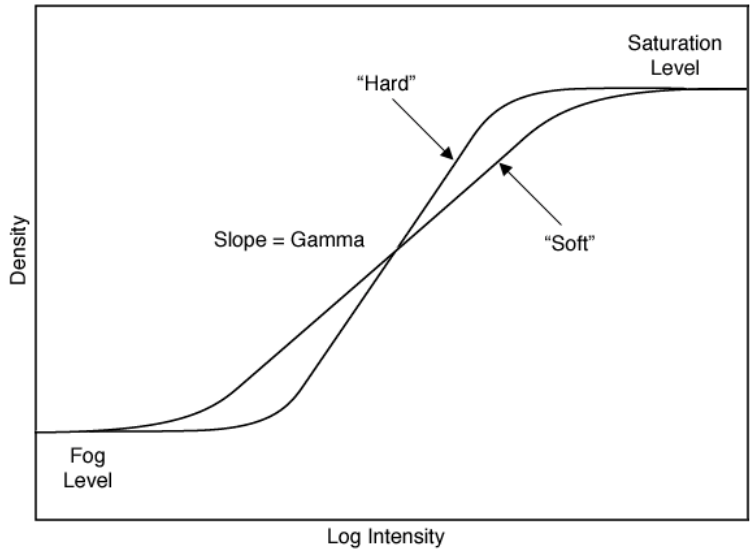

Figure 4. H & D curves for hard and soft photographic film.

The same name is now applied to the same characteristic of digitized images in computer software. Computer displays are nonlinear, with a typical value of gamma in the range from 1.8 (the Macintosh standard) to 2.2 (the Windows standard). The mathematical relationship is Output = Input ^ Gamma (where input and output are normalized to the 0..1 range). To compensate for this nonlinearity, the images may be processed from the camera with a gamma value of about 0.5 (=1/2.0) as shown in Figure 5. But most image processing and display

software allows the gamma value to be adjusted for optimum viewing (or printing, which may require different settings).

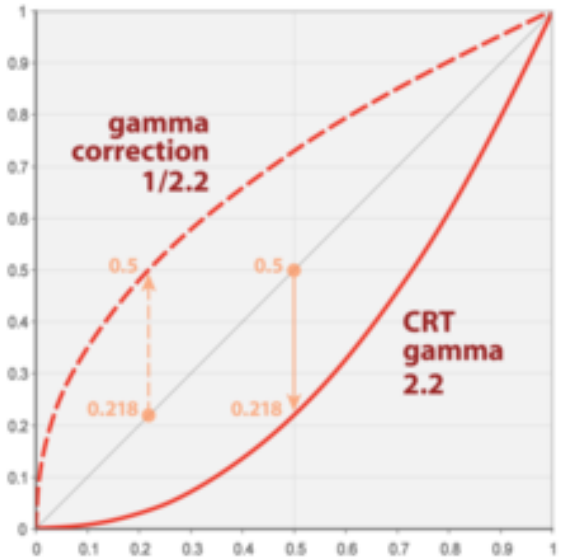

Figure 5. Example of the gamma curve for a computer display, with the compensating gamma value applied to the raw data.

A common problem experienced by users of digital images is that they "look different" on different monitors, or in different programs (particularly when going from gamma-aware programs like Photoshop to presentation tools such as Powerpoint or Word, which may not make the same adjustments). Hardcopy prints often present a different appearance than the on-screen image, and in any case cannot represent as great a range of contrast. Because adjusting overall contrast and gamma was commonly done in traditional film photography, it is widely assumed that making arbitrary adjustments to similar parameters for digital images is a permissible operation, and tools for accomplishing this are part of most imaging software. But it is important to keep in mind that such adjustments can alter the appearance of images so that some details may become either more or less visible, and that all such operations (including the selection of the area to be recorded by the camera) represent opportunities to bias the results, and are necessarily subject to the usual caveats about scientific responsibility.

Some cameras do not give the user access to the raw sensor data, but only save or transfer to the computer one that has already been converted. Even more serious is the use of lossy compression to reduce the file size, either to save storage space or reduce the time needed to transmit or save the result. Images are large: a 10 Megapixel image using 2 bytes each (necessary if the dynamic range exceeds 256 values) for red, green and blue is 60 MBytes. JPEG compression can easily reduce that to 10 MBytes, and in many cases to much less. The problem with all lossy compression schemes (and even the ones that call themselves "lossless" can truncate or round off values) is that there is no way to predict just what details may be eliminated from the original image. Fine lines may be erased, feature boundaries moved, texture eliminated, and colors altered, and this may vary from one image to another and from one location to another within an image. There is no way to recover the lost information, and images that have been compressed using JPEG, fractal compression, or other lossy schemes must never be used for any scientific or forensic application. Given the availability of large amounts of inexpensive storage, and increasingly fast networks for image transmission, there is simply no valid reason to ever use these procedures.