High Throughput Microscopy
Image Sensor Basics – The Pixel

By Kevin McCarthy, Chief Technology Officer

The compound optical microscope was invented by three Dutch spectacle makers around the year 1590. For the next three centuries, advances in optical design and available glass types yielded continuous performance improvements, but the "image sensor", as it was, remained the human eye. Discoveries using microscopes could be sketched and reported, but no image data was available. Beginning at the end of the 19th century, photography based on silver halide emulsions was adapted to microscopy, allowing permanent image capture and sharing. Film based microscopy suffers from a number of non-linear effects, but some quantitative data was eventually coaxed from the resulting images via slow point scanning of film, using a device known as a microdensitometer. All this changed in the last decades of the twentieth century, when new digital image sensors were developed, rapidly displacing film in both scientific and consumer photography.

The job of a digital image sensor is on one level pretty basic: it is to turn as many as possible of the photons striking the image sensor into electrons, which can then in turn be accurately measured. This conversion of photons to electrons must preserve the full X-Y spatial information in the original image, and digital sensors do this by consisting of a very large number of individual photosensitive sites, each called a "pixel" (shorthand for "**pic**ture **el**ement"). Typical pixel counts for the image sensors used in modern life science applications range from several hundred thousand, to tens of millions. But at the heart of modern digital image sensors lies the individual pixel, that discrete element by which light is sensed. In this white paper, we will examine a number of aspects of digital image sensor pixels, with an emphasis on their relevance for microscopy. While image sensors come in two distinct semiconductor technology flavors, CCD and CMOS, those distinctions will be covered in a future white paper. In this white paper, we will examine aspects common to the pixels of both sensor technologies.

Pixel Basics: The Silicon Photodiode

The heart of both CCD and CMOS image sensor pixels is the same: a silicon photodiode. This is a feature fabricated in a single crystal of silicon, wherein differential doping produces p and n channel regions (aka, a "p-n junction"). Photons above a certain energy threshold that enter this junction deposit their energy there and generate an electron-hole pair, in a bit of physics known as the photoelectric effect. As light from a stationary object continues to fall upon the image sensor, the number of electrons produced, and available to measure, increases. The brighter the illumination intensity, and the longer this "integration time", the more electrons will be produced, in an extremely linear fashion. At the end of the integration period, the amount of charge is then measured (digitized), providing a permanent digital record of the light intensity at that pixel.

A Useful Analogy

A useful analogy for digital image sensors is that of an open field, filled with a large array of empty buckets (pixels). A passing rain cloud briefly pauses over the field, delivers some rain (photons), and then, after a short while (the integration time), stops raining, and moves on, leaving different amounts of water (photoelectrons) in each bucket. The job is now to measure the water level (number of electrons) in each bucket, as this is directly proportional to the amount of rain (photons) each bucket (pixel) recorded.

Key Sensor Pixel Parameters

In this section, we will examine a number of aspects of image sensor pixels as they relate to microscopy.

- **Pixel size:** This is a very important choice to make in any digital microscopy imaging system; pixel sizes range at the extremes between 1 and 25 micrometers (symbol µm, aka "microns"). One micron is one millionth of a meter, and one thousandth of a millimeter. In microscopy applications, the physical pixel on the imaging sensor corresponds to an equivalent (but considerably smaller) pixel down at the sample, with the ratio between the sensor and sample pixels being the magnification of the microscope. For a popular (if expensive) 4 megapixel sensor in a microscope with a magnification of 20X, the relationship between sensor and sample pixels is shown in Fig 1:

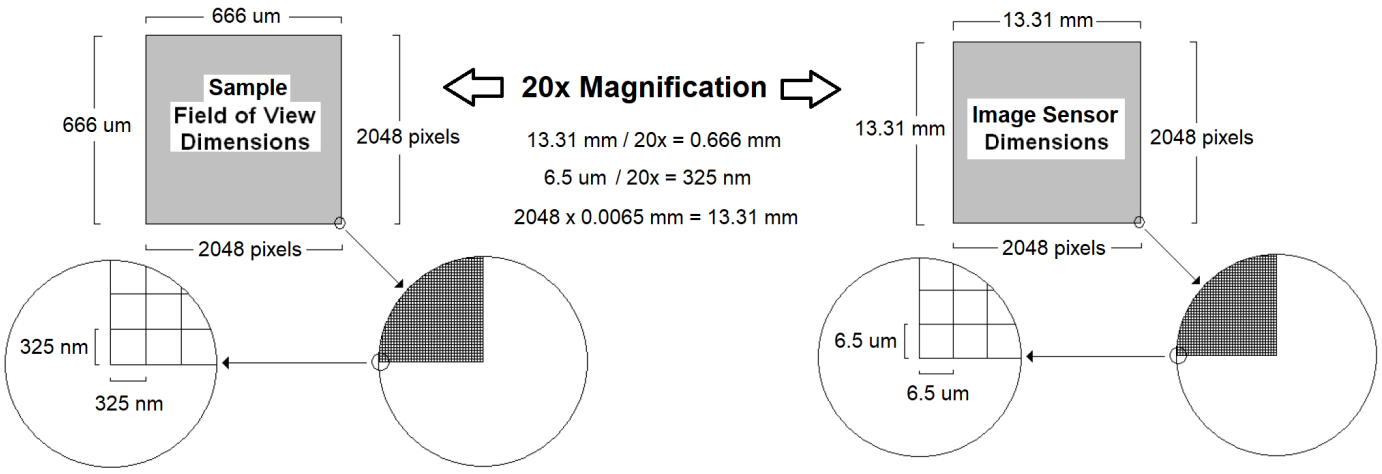

Fig. 1: Pixel dimensions at the sensor and the sample

At first glance, smaller pixels would appear to have an advantage, since they provide finer resolution. Alas, such "geometric" resolution can be misleading. Light behaves as both a particle and a wave, and the wave nature of light means that diffraction sets a limit to the achievable resolution. In microscopes, the diffraction limited resolution is equal to 0.61 * λ / NA, where λ is the wavelength, and NA is the objective's Numerical Aperture. This very important parameter, which controls both the diffraction limited resolution and the depth of field, is directly related to how wide the cone of light from a point on the sample is when it enters the objective. Numerically, NA is equal to the inverse sine of the half angle of that light cone. For green light at 550 nanometer (nm) wavelength, and with a microscope objective Numerical Aperture of 0.80, the corresponding diffraction limited resolution is 419 nm.

In general, we don't want to design an automated digital microscope in which the sample pixels are equal in size to the system's diffraction limited resolution. Rather, we want to "oversample", with multiple sample pixels for each diffraction limited pixel. The optimal value for image sensor pixels is one in which the optical system's diffraction limited resolution is between two and three times larger than the equivalent pixel size down at the sample. We normally pick a ratio of 2.3X; in the above example, with 550 nm light and an objective NA of 0.80, that corresponds to an optimum sample pixel size of 419 nm/2.3, or 182 nm. For the image sensor in Figure 1, with 6.5 µm sensor pixels, the microscope

magnification is inadequate at 20X, since the corresponding sample pixel is too large at 325 nm (oversampling of only 1.3X). In this case, a 40X objective with 162 nm sample pixels would be a better match, as would be a 20X objective paired with a sensor with 3.64 µm pixels.

- **Pixel and image sensor aspect ratio:** Unlike consumer cameras, where a rectangular sensor with a 4:3 or 16:9 sensor aspect ratio is popular, microscope optics have cylindrical symmetry, and hence a square sensor aspect ratio is strongly preferred. Similarly, while sensors with rectangular pixels are commercially available, they provide a needless complication, with square pixels almost universal in digital microscopy.

- **Number of Pixels (Sensor size):** Under the assumption of square pixels and sensors (see above), the sensor size is simply the square root of the number of pixels times the pixel size; for example, in a 1 megapixel sensor with 10 µm pixels, the sensor will be 10 mm square. The number of pixels (and hence the sensor size) entails a compromise between cost, throughput, and image download speed; for the same magnification, a larger sensor will have higher imaging throughput, as it will have a larger field of view. However, field sizes are limited by optical considerations; to preserve image quality in the corners of the sensor, it is important that the sensor field of view diagonal not exceed the field number for which the objective is corrected. This value is specified for any given model of microscope objective, and typically ranges from 22 to 27 mm. The relationship between field number and the dimensions of the field of view is again simply the magnification; for a 20X objective with a field number of 25 mm, the allowable field of view should not exceed 25 mm / 20X, or 1.25 mm diagonal (0.884 mm square). But while the field of view is strictly limited by the field number and nominal magnification of the objective, the same limitation does not apply to the image sensor. By varying the total magnification, which is a function of both the objective and the tube lens, sensors that exceed the field number can be fully utilized, although issues of sensor cost and image download times argue against the use of physically large image sensors. This use of a custom focal length tube lens provides a valuable knob to turn (the tube lens focal length) to vary the total magnification, and in so doing, help optimize both the oversampling ratio and the field of view for any given objective. The sweet spot for pixel and image sensor size in many microscopy applications is about a 12 mm square sensor (17 mm diagonal), with around 10 million pixels, each about 4 µm square.

- **Full Well Capacity**: The region of the photosite that stores the charge is limited in how many photoelectrons it can hold, much as the buckets in our analogy don't hold any more water once they become full. At some point, as photons continue to stream into the photosite and generate photoelectrons, the photodiode will saturate, and reach what is called its "full well capacity". All imaging applications go to lengths to avoid saturation, and so remain in the strictly linear regime. The full well capacity varies according to pixel area and sensor design, but typically range from 5k to 100k electrons. In applications with very long integration times, such as in astronomy, substantial fractions of the full well capacity are regularly used. In biological imaging, the signals are often faint, and fill only a small fraction of the available full well capacity.

- **Quantum Efficiency**: Image sensors vary widely in the efficiency by which they convert incoming photons to photoelectrons. This is a function of both the intrinsic properties of silicon, whose

sensitivity varies with wavelength, and each sensor's physical implementation of pixel structure. Inevitably, pixels must include adjacent circuit wires and transistors that take up space; these prevent all of the pixel area from contributing to useful signal. To evaluate image sensor performance, the term "Quantum Efficiency" is used. This is a dimensionless number, expressed in percent, which varies from 0 to 100. At a QE of 100%, every single photon entering the pixel will produce one photoelectron; at 50%, only one photon out of two will produce a photoelectron. For what it's worth, the human eye has a QE of about 10%. QE varies with wavelength, and so is typically shown as a plot of QE versus wavelength. Older generation CCDs had QE's of as low as 30%, but thanks to a couple of tricks (terms for which follow), cutting edge image sensors today can boast of peak QE's (typically in the yellow-green portion of the spectrum) that can hit 95%! Such an impressive QE chart, with its trick described in the following section, is shown in Fig 2:

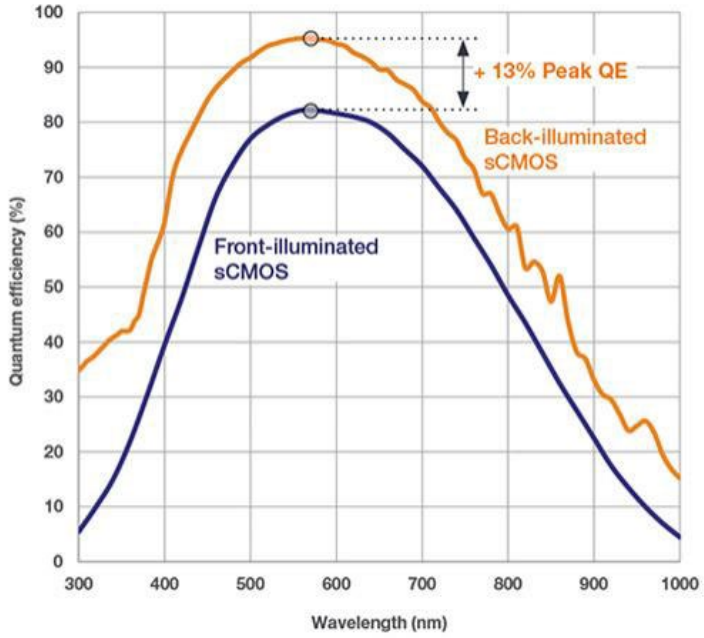

Fig. 2: Graph of Quantum Efficiency versus wavelength[1]

- **Back Side Illumination (BSI)**: As mentioned above, one key limitation on sensor QE is that not all of the photosite area contributes to image formation; photons falling on features such as circuit traces and transistors of each pixel don't make it to the p-n junction, and are lost. Semiconductor engineers have devised a neat trick to get around this. An entire wafer of sensors is bonded upside-down onto to what is called a "carrier wafer". All of the circuitry and transistors required to operate each pixel, which were blocking a fraction of the incident photons, are now hidden ***beneath*** the active photodiode. The only problem is that there is now an entire wafer thickness worth of silicon between the back of the photodiode, and the light that is trying to get there. The solution is to precisely grind off nearly all of the silicon of the original wafer, exposing the back of the photodiode to light, using a process called "CMP" (Chemical Mechanical Polishing). The carrier wafer is what provides the structural support, as the newly thinned sensor wafer would

[1] Image source: https://andor.oxinst.com/products/fast-and-sensitive-scmos-cameras

be far too flimsy on its own. In Figure 2, switching from traditional to thinned, backside illumination delivers a 13% increase in QE, which (all other things being equal) equates to 13% higher imaging throughput. Backside illumination comes at a price, but for cutting edge applications such as DNA sequencing, this increased performance is worth its weight in gold.

- **Microlenses**: Before there was backside illumination, sensor designers came up with a different solution to the problem that not all parts of a pixel can generate photoelectrons. That solution is an array of molded plastic square microlenses that is applied to the front face of the sensor. Photons that would have struck inactive circuit elements are instead focused to fall on the active photodiode, increasing the QE. In our field full of buckets analogy, a lot of rain would fall in the gaps between circular buckets. Adding square faced funnels on the top of each bucket eliminates those gaps, and makes sure that every raindrop finds its way into a bucket. Microlenses continue to be used for both CCD and CMOS sensors as a simple way to boost QE.

- **Dark Current**: Photoelectrons aren't solely produced by incoming photons. Thermal jostling of atoms within the photosite will also contribute photoelectrons, which is referred to as "dark current", the units of which are electrons per pixel per second. Dark current is typically proportional to the pixel area, with larger pixels generating more dark current. Since this is not true signal, it can distort precision image measurements. Since dark current is thermal in nature, the most common way to reduce it is to cool the sensor below ambient temperature. For every 6° to 9° C of cooling (the so-called "doubling temperature"), the dark current will drop by a factor of two, so maintaining the sensor at -40° C will reduce the dark current by a factor of ~ 500X. In modern, high performance, deep-cooled cameras, dark current can be as low as 0.001 electrons per pixel per second. In astronomy, where exposures can be very long, such extremely low dark current is valuable. In life science and diagnostic applications, where the typical exposures are far shorter (in many cases considerably less than one second), such levels of dark current suppression are often overkill, and cameras can be operated with little to no sensor cooling. This is especially true of modern CMOS image sensors, which have dramatically lower dark current than CCD based devices. In the latest CMOS image sensors, pixel dark current at room temperature can be as low as 0.03 electrons per second (Fig 3).

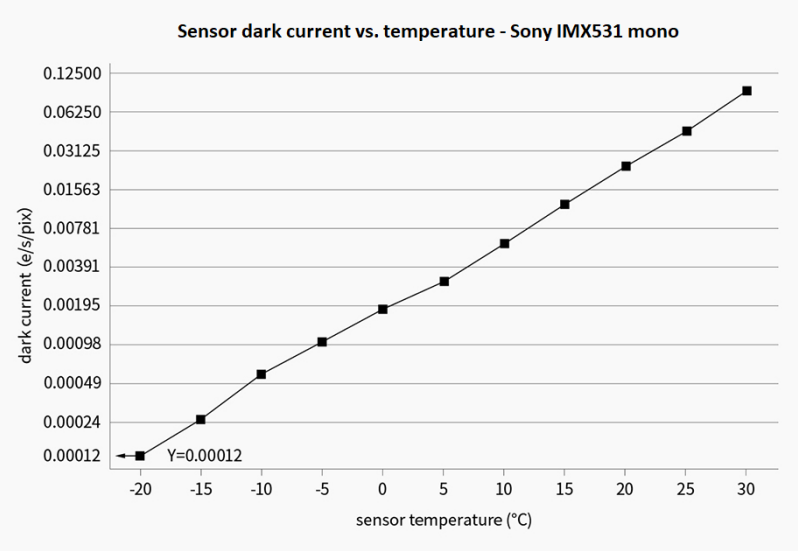

Fig. 3: Sensor (Sony IMX531 mono) dark current vs temperature[2]

- **Read Noise**: The readout electronics in digital image sensors keeps getting better, but the mere act of measuring (digitizing) the analog signal of the accumulated photoelectrons inevitably injects a certain amount of noise. In principle, successive images of an unchanging illuminated field should all return exactly the same pixel signal values, but the difference between these images is largely due to read noise. This noise sets a limit to the signal to noise ratio, and in recent years, CMOS sensors have pulled well ahead of their CCD counterparts in this regard. CCD sensors seem to have plateaued at between 7 and 20 electrons of read noise, whereas the latest CMOS sensors can achieve read noise levels of a mere 1 to 2 electrons. One exception to the above rule of thumb occurs with EMCCD's (electron multiplying CCDs), which employ an additional exponential gain section (a horizontal register with impact ionization at each transfer) to boost signal. EMCCDs can drive read noise below one electron, but do so with a number of additional trade-offs.

- **Bit depth**: This parameter is determined by the resolution of the A:D converter that digitizes the image. A:D converters range from 10 bits on the low end, to 16 bits on the high end. In applications with very faint signals, such as single molecule or faint fluorophore imaging, the bit depth may not matter much, since the range of faintest to brightest objects may fit entirely within a ten-bit range.

- **Dynamic Range**: A digital imaging system will have a maximum linear signal level, which occurs a bit below hitting the full well capacity. At the other end, for very faint signals, measurement precision is capped by the read noise. The ratio of the maximum linear signal to the read noise accordingly defines the camera's dynamic range. Using the example of one high-performance camera, the full well capacity is 85,000 electrons, and the read noise is 1.6 electrons. This camera's dynamic range is accordingly 53,000:1. Such a camera, one would ideally have a 16 bit A:D converter, which can resolve a bit finer than the dynamic range, with 65,536 output levels.

[2] Image source: https://astronomy-imaging-camera.com/product/asi533mm-pro

In practice, the usable dynamic range would be slightly lower, since it is a bit dangerous to data quality to operate right up to the edge of saturation (full well capacity).

- **Hot and cold pixels**: In an ideal world, every one of the millions of pixels in an image sensor will provide identical signal at all illumination levels, including complete darkness (which should produce zero signal). In the real world, individual pixels vary in their response to both dark and light. In some cases, individual pixels are essentially broken, and are "stuck on", producing substantially higher false signal than normal pixels. These are referred to as "hot pixels", and are best removed from consideration through the creation of a "bad pixel map". Another class of broken pixels that should be captured in a bad pixel map are so-called "cold pixels", which either produce substantially lower signal than expected, or none at all. While these too should be captured in the bad pixel map, it takes an illuminated field to find them. If you are willing to pay enough money, nearly perfect image sensors are available, but the creation and use of a bad pixel map is an economical way to use less than perfect sensors. There are a variety of algorithms that can provide a good guess at the correct value of a bad pixel, by interpolation using the values of its eight good adjacent neighbors.

- **Pixel to pixel variation**: In addition to very discordant (broken) hot and cold pixels described above, there is also inevitably some variation from pixel to pixel in good ones that are otherwise working. This is specified in two way: DSNU (Dark Signal Non-Uniformity), which measures pixel to pixel variation in the absence of light, and PRNU (Photo Response Non-Uniformity), which measures pixel to pixel variation in the presence of light.

- **Vignetting and related issues:** The optical system will inevitably have some level of vignetting (light falloff) towards the corners of the field, in both the illumination side uniformity and the subsequent light collection. It may also suffer from dust motes and other non-symmetrical issues.

The effects of the above imperfections can be subtle, but the following steps can help to ensure that their effects are minimized in the final image data:

  - **Dark Frames**: When no light at all is allowed to fall on an image sensor, its 2D array of pixels should all read "zero". Alas, a number of physical effects, including but not limited to read noise and dark current, mean that an image with no light at all will tend to have non-zero pixel values. Moreover, this false signal can vary from pixel to pixel, and tends to increase over time. Fortunately, for any given exposure duration (integration period), this dark signal is very repeatable, both from exposure to exposure, and from pixel to pixel. An excellent way to deal with dark signal is to take a separate exposure with no light (a "dark frame"), and then subtract all of those pixel values from those of your corresponding (and importantly, equal exposure duration) "light frames". Better yet is to take a lot of dark frames, and create a "master dark" frame that is the average of the group.

  - **Flat frames**: Nearly all optical systems exhibit some degree of what is called "vignetting", which simply means that the brightness falls off as you get farther from the optical axis.

The issue is a combination of illumination uniformity and light collection uniformity across the field of view. Both issues are worst in the four corners of the sensor, but are generally symmetrical around the optical axis. In addition to this symmetrical light fall-off to and from the sample, there can be other, localized effects, which are caused by a number of factors. These include dust motes on the sensor window or lenses, reflections off internal surfaces along the optical path, fingerprints on filters, you name it. All contribute to the measured brightness across the sensor failing to accurately reflect the corresponding brightness across the sample, which is, after all, the whole point of automated digital imaging. To address this issue, the sample field is uniformly illuminated, and an image is acquired. Better yet, as with dark frames, a set of uniformly illuminated images is acquired, and averaged to produce a master "flat frame". Unlike with dark frames, which are subtracted from the light frames, the light frames are divided by the flat frame to produce the corrected image frame. For example, if the brightness of a uniformly illuminated field at a pixel in the sensor corner was only half that of one in the field center, we would divide the value of that pixel by 0.5, which is to say we would double its value, bringing it to parity with the field center. Optical designers do their best to minimize illumination and light collection falloff towards the corners, since when we increase gain in the dimmer corners, we also increase the image noise.

- **Bias Frames**: In this last take on image perfection, each image acquisition has a repeatable pattern from pixel to pixel of bias injected by the read process. This is distinct from the non-repeatable read noise discussed above. This bias, unlike what was addressed with the application of dark frames, is present in even the shortest exposure period. Accordingly, we fix it by turning off all light to the sensor, but then acquiring and averaging a number of individual "bias frames", acquired at the shortest possible exposure period. The resulting master bias frame is then subtracted from the master dark frame before it is used with the light and flat frames.

- **All together now**: While not every life science or diagnostic imaging application requires a full set of image corrections, for those that need it, the above steps of creating a bad pixel map, as well as acquiring and using dark, flat, and bias frames, goes a long way to making the millions of pixels behave in a nearly identical fashion, ensuring very high image accuracy.

ABOUT THE AUTHOR...

Kevin McCarthy is the Chief Technology Officer of Dover Motion, and recently celebrated his 40th year at the company. He holds a B.S. degree in Physics from The Massachusetts Institute of Technology. Kevin can be reached at: kevin.mccarthy@dovermotion.com.