High Speed Digital Cameras

SNR & Dynamic Range,
What does it really mean?

Jan 2009

Commentary on Overstating the Dynamic Range & Bit Depth

High-speed video cameras are widely used to capture fast moving objects. The captured video is played back at a speed much slower than the real-time record rate. Real-time to most people is 30 frames per second (fps) but in our applications, 1000 to +1,000,000 fps is what we mean as real-time, the @ speed recording. Steadily, the image resolution has been improving. In fact, Photron has introduced the world's fastest 4-megapixel camera, Photron FASTCAM SA2. The FASTCAM SA2 2k x 2k resolution at 1000 fps is a technology breakthrough. As the resolution increase, so does the volume of information recorded. In fact, high-speed digital cameras typically push over 3 gigapixels per second. That is a lot of information to transfer and the amount of information only increases as the dynamic range/bit-depth increases. This is one of the main reasons you should know what is real when a manufacture speaks about their cameras having 10, 12 or even 16-bits.

Many knowledgably experts on imaging often dance back and forth when speaking about dynamic range and bit depth as if they are the same. Some technical people talk also about the camera has (n) bits. The Number of Bits a camera reads out, the Bit Depth and the Dynamic Range seem to be the same but they are truly not the same.

Camera has (n) Bits refers to the number of digital bits that are captured or read or written out from the camera. This does not accurately reflect how many of the bits are signal and how many of the bits are noise.

Bit-Depth refers to the number of tonal steps (grayscale) in the density range defined by the Dynamic Range.

Dynamic Range refers to the difference in density between the top of the range (light) to the bottom of the range (dark).

OK. So what do these definitions mean when applied to a high-speed camera? Simply said, dynamic range is the range of values from the lightest to darkest detectable, expressed in bits, generated within a pixel. Dynamic range has improved in high-speed cameras, but unfortunately, some camera manufactures have not been clear about claims of high-speed color images with 14-bit or 16-bit dynamic range. The bit-depth (8, 10, 12, 14 or 16 bits) represents how many tonal steps exist between the lightest to the darkest detectable level within a pixel. For a manufacture to say they have a bit-depth of 12-bits would most certainly mean that 4096 tonal steps should be discernable, but is that really the situation? To help clarify the subject of bit-depth, we need to explain a few simply concepts.

The majority of high-speed cameras are using CMOS sensor technology. CMOS sensors have many advantages for high-speed cameras. Most scientific cameras used for applications that require 14-bits or 16-bits of dynamic range are CCD based sensors that operate at much slower clock speeds, at slower frame rates and often use cooling techniques not found in most high-speed cameras. Frame averaging and binning techniques have been used with CCD sensors to reduce the noise level while greatly increasing the signal level, for a higher SNR and dynamic range. Binning in a CMOS sensor does not reduce the noise or increase the SNR or dynamic range. Frame averaging would have the same results as found with a CCD sensor. However, for high-speed applications, frame averaging will leave undesirable image artifacts such as blur and edge displacement. CMOS sensor technology is not used in most scientific imaging applications requiring 14-bit or 16-bit linear dynamic range. The reason is the linearity and noise produced in a CMOS Active Pixel Sensor (APS) is not sufficient for these demanding applications. Wait a minute! If CMOS is not used in cameras needing a bit-depth of 14 or 16-bits, how is it that some camera manufactures claim a 14 or 16 bit high-speed camera? You may ask how can this be true when the high-speed cameras are running at a higher clock rate and generating more internal noise than the slow scan cameras used in scientific applications. Let's try to understand what is real when bit depth is discussed.

Getting More Than 12-bits

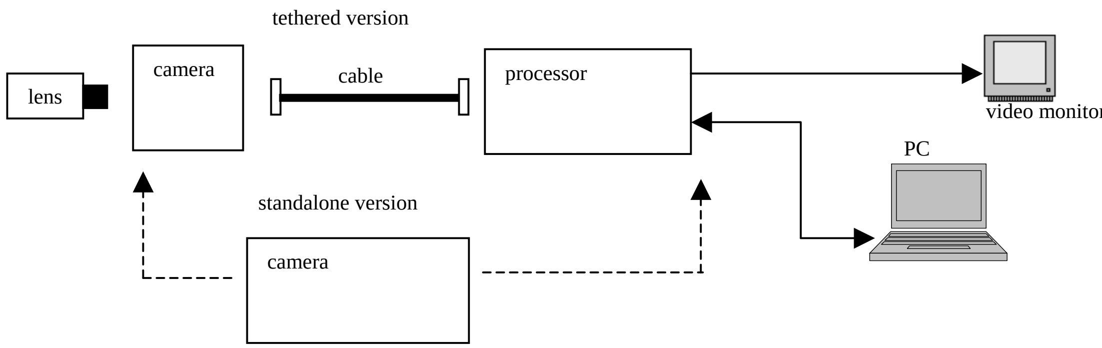
CMOS APS is becoming more mature, producing 10 & in some cases, nearly 12-bit HDR (high dynamic range) images that are linear. Some CMOS sensors have the capability to reset the level of each pixel to a preset level during integration. This technique original called WDR (wide dynamic range) was developed out of JPL years ago. The resulting image may have a wider range of values, as the name implies, but the range is not linear. This makes it very difficult to create accurate color as well as good stop motion photography due to the variability of the reset from pixel to pixel. The WDR technique also known as Extended or Extreme Dynamic Range (EDR) or Dual Slope is used in high-speed cameras to achieve a greater bit-depth than 12-bits. They all can produce this non-linear wide dynamic range with a huge bit-depth but not all manufactures report the bit-depth as 14 or 16 bits.

But the Camera has a 12-bit ADC?

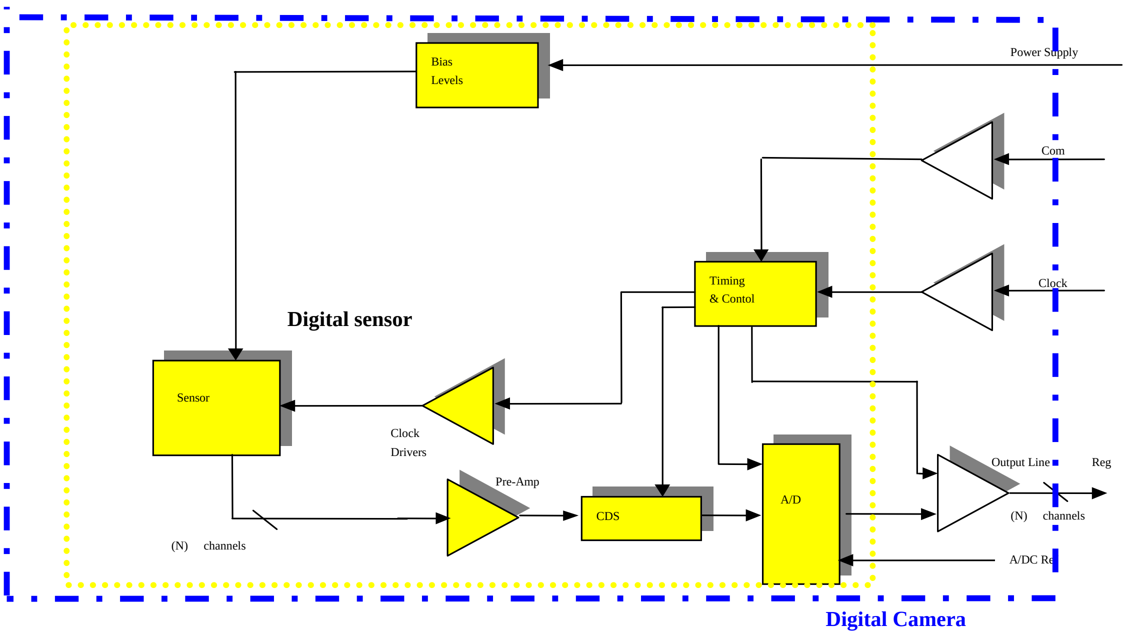
Majority of the high-speed sensors come from the two semiconductor companies. Therefore, the differences in dynamic range or their bit-depth have more to do with pixel size, full well capacity, output conversion and the various noise sources. Claiming a camera has a 12-bit dynamic range simply because the camera has an ADC (Analog-to-Digital-Converter) of 12 or 16-bits is misleading because noise and the capacity of the pixel well to produce such a dynamic range has not been considered. You should ask yourself, "Can I see the range of tonal scale expected for 12-bits"? My advice would be to closely look at what is being claimed as far as the dynamic range & bit-depth. Then make judgments on the image quality that you actual get from the camera.

Let's have a closer look at the process of capturing a digital high-speed image sequence. From this we will see what it takes to get a bit depth of 12, 14 or 16-bits.

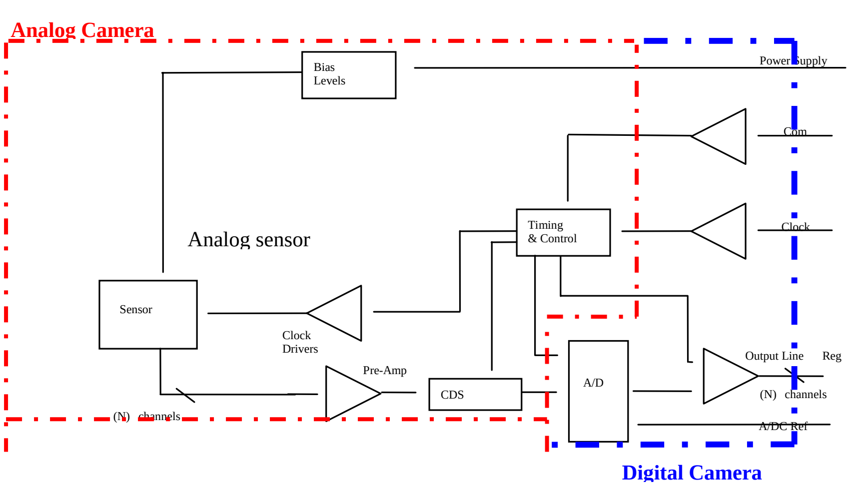
The building blocks required for a high speed imaging system are shown below. An important distinction can be seen in the two configurations, the tethered & standalone versions. A cable with analog video signals is missing in the standalone version. In fact, there is a third version not shown, a tethered version where the image output is digital from the back of the camera. The front end of the camera begins with the conversion of an image from photons to electrons. This is an analog process. Some cameras have what is called a digital sensor. This simply means that much



of the signal processing electronics has been integrated onto the imaging sensor (camera-on-a-chip) with images in digital format are readout. Other cameras have a sensor that is not integrated as much and the output is still in the analog format.   What is not contained on the sensor still needs to exist at some point in the camera architecture.   It is just a question on where the conversion from the analog domain to the digital domain occurs. The reason this is important to realize is that the analog domain is more susceptible to noise then the digital domain. Therefore, all high-speed cameras have both an analog and digital domain. The analog portion is where the level of noise will have the greatest detrimental effect on the quality of the image. A camera with a digital output still has at some point in the architecture a conversion from analog to digital and it is subjected to the same noise sources as that of an analog sensor.



The blue outline represents a tethered digital camera that has a sensor with a highly integrated digital sensor. If digital memory, display controller and controller were added to the block diagram, we would have a standalone digital camera. The blocks represented in yellow are ones that you typically would find inside a digital sensor. The block marked A/D is where the analog signal is converted into the digital domain. A/D converters will digitize the analog signal into (n) bits, depending on the resolution of the A/D (i.e. number of bits, 8, 10, 12, 14 & 16). Currently, all known high-speed digital sensors (camera-on-a-chip type) have A/D resolution of 10 bits maximum due to the limitations in ADC size to fit within two pixel column's width. Typically, there are as many A/D converters as there are horizontal pixels in these high-speed digital sensors. Claiming the sensor has a 10-bit ADC and the camera digitizes the image to 10-bits is correct. However, saying all 10-bits represent the actual image signal without noise is not correct. The analog signal has noise mixed in and this signal is converted into the digital domain. In fact, the process of converting the analog signal to a digital signal introduces additional errors called quantization noise. Quantization noise is reduced to negagiable levels when a 12-bit or greater ADC is used. Unfortunately, reducing quantization noise is only a small portion of the possible noise that you will find in a high



speed image.

Above is the same sensor block diagram as before except this block diagram shows what would be included in an analog camera (red dash line). Also, shown are the blocks that would be added to make a tethered digital camera. As discussed previously, adding digital memory, a display controller and a controller, we would have a standalone digital camera. This configuration has an analog sensor, meaning the output of the sensor is still in the analog domain. Signal processing and conversion of the analog image into a digital image by the A/D converter is still required, as we previously shown.

The point of showing these configurations is to show the same steps are performed to produce a digital image, from a camera that has a digital or analog sensor. There are advantages of having a digital sensor. Most obvious is that the camera size can be much smaller, simpler in design and most likely less expensive to make. Arguably, the digital sensor can be optimize to have less noise structures within the sensor but the A/D conversion is often compromised in resolution or the conversion method (folded A/D). Further reading on this subject may be needed to clarify specifics about the pro's & con's of a highly integrated digital sensor vs. an analog sensor. However, this is not the focus of our paper. Let's continue our discussion on bit-depth and dynamic range of a camera.

Signal to Noise Ration (SNR)

Let's begin with the discussion of signal to noise (SNR) as it applies to a camera, what is called the system SNR. After all, the image you see represents the system SNR not just the sensor SNR. As we have discussed before, a luminous scene imaged by an electronic camera converts light from photons to electrons. The electrons are signal processed, amplified and either read or stored in a camera system. Noise is added to the image signal each time it is detected, processed, and converted. In the classical mathematical representation, SNR for an image is defined as the ratio of the light detected on the sensor to the sum of the noise in reading the signal. SNR is expressed in units of power or decibels (dB).

$$\mathbf{SNR\ (dB) = 20\log\ (Signal\ e\text{-}\ /\ Noise\ e\text{-})}$$

The maximum number of electrons (e-) that a sensor can collect within a pixel is the full well capacity. The larger the pixel, the more electrons (e-) a pixel can hold. As an example, the typical full well capacity of a high-speed sensor is usually under 100,000 electrons (e-) and more often than not it is around 60,000 (e-). Cooled scientific cameras operating at 1 MHz pixel clock frequency could expect a noise floor around 7 (e-). Most high-speed cameras are not cooled and their pixel clock frequency, depending on their frame rate can range from 20MHz to 95 MHz, an average of 50x greater than scientific cameras. It is common to see a noise floor in the hundreds of (e-). If we use 150 (e-) as our noise floor, just for the sensor, the very best SNR expected would be:

$$\mathrm{SNR} = 20\log(60{,}000/150) = 52.04\ \mathrm{dB}$$

Assuming the image could be read from the sensor with no other noise, your A/D converter's resolution can be determine from the following table.

Number Bits	8	10	12	14	16
Ratio	256:1	1024:1	4096:1	16384:1	65536:1
Infor Increase	------	4x	16x	64x	256x
Max. Dynamic Range (dB)	48	60	72	84	96

Therefore, a 10-bit A/D would be required. Please note that the image you get will not be a true 10-bit image simply because your camera has a 10-bit A/D converter. In fact, at 52 dB, your image pixel depth is not even 9 bits or 54 dB.

Some camera manufactures say their cameras produce 14 bits, however, when you look at their own application notes (see reference below), the performance is shown to have a dynamic range of 1320:1. That is amazing since an 11-bit dynamic range should be 2048:1 and a 10-bit dynamic range would be 1024:1. Therefore, claiming a 14-bit camera with a dynamic range of 1320:1, you can get the same performance out of a camera with an 11-bit dynamic range. Some caution clearly needs to be exercised when selecting a camera based on a 14-bit claimed performance does not match reality. A 14-bit ADC (analog digital converter) does not mean you will get a 14-bit dynamic range image.

Figure 1: Response and noise of a monochrome v10 vs focal plane exposure levels.

at different illumination levels, we will find that noise also changes with illumination.

A typical plot of average response and noise for a Phantom V10 mono is shown in Figure 1. We see that the response saturates a bit above 0.06 lux·s.[2] The dark level is at 400 counts, so the saturation signal is 3696 counts. Noise in the darkness, also called *readout noise* is 2.8 counts.[3] Noise increases as the square root of light level, which is characteristic of *photon shot noise* limited cameras.

3.2 Signal to Noise Ratio

A related plot is shown in Figure 2. This plot shows the ratio of signal (camera response minus black level) to noise for different illumination levels. Using this graph, we can determine the minimum exposure we need in order to capture a given object with the desired signal to noise ratio.

[2] This translates into a saturation-based ISO at 1300.

[3] The dynamic range of this particular camera is 1320:1, or 62.4 dB

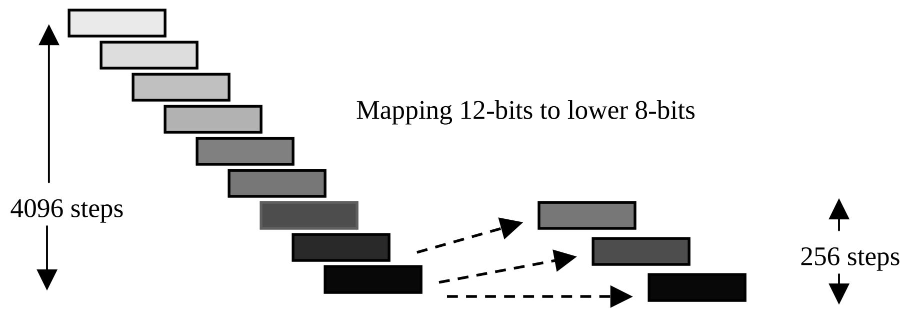
5

Reference - Vision Research App Note: Digital Camera Exposure Indices, by Radu Corlan, 2006

Displaying an Image with a Bit Depth Greater than 8 bits

Many image file formats (JPG, BMP, TIFF8 etc...) support only 8-bits per image plane, meaning a RGB image has three color planes or 24-bit color. Millions of colors can be encoded, but today's 24-bit PC video card technology is incapable of displaying more than 256 shades per plane on the monitor. There is no great incentive to display more than 256 shades per image plane since the human vision can only resolve about 32 to 64 shades of gray.

Therefore, high dynamic range (HDR) images must be mapped in sections for display on PC monitors. Another method would be to compress the HDR images such that the image can be shown on the PC monitor. The mapping is a better method since more detail within the 256 steps can be viewed. Below is a graphic example of mapping 12-bit image to an 8-bit display.



The advantage of having the ability to map HDR images is that more detail can be viewed when looking at dark areas or bright areas within the image. Having additional dynamic range also eases the lighting requirements since the image can be underexposed. As an example, one-fourth the light level used to capture a 12-bit image will still have a shade range of 1024 steps or 10 bits. However, if you have a camera that has a 14-bt ADC but the lower 3 bits are noise, then your underexposed image with one-fourth the light may be clipped, since the lower bits are important for low light images. Again, it is important to know what your real dynamic range is on a camera. Belo is an example of a 10-bit image where a bit-depth of 10 bits exist but only the lower 8 are displayed on the left and the upper 8 are displayed on the right.

Summary

The distention of having a camera with a dynamic range of 12-bits is not the same as having a camera with a 12-bit ADC. We have noted and shown the difference. Assuming you have 12-bits of dynamic range, displaying 12 bits on a PC requires a technique we call windowing or bracketing 8 bits at a time out of the 12 bits. We have described how this is easily done with the Photron PFV program. Don't let a Sales person claim performance results without the ability to backup the claim with test results such as the ISO 12232 Standard for camera performance.