High Speed Digital Cameras

16, 14, 12, 10 & 8-bits cameras,
What does it really mean?

March 2007

Commentary About Overstating the Dynamic Range

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 250,000 fps is what we mean as real-time, the @ speed recording. Steadily, the image resolution has been improving and soon we will see 4-megapixel high-speed cameras. As the resolution increase, so does the volume of information recorded. In fact, high-speed digital cameras typically push over 3 gigapixels per second.

The subject of this paper is a discussion on what to expect in dynamic range from high-speed cameras. Simply said, dynamic range is the range of values, 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. Therefore, we feel that it is important to help the end user understand more about dynamic range so judgments can be made on what it all really means.

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 high-speed cameras.   Frame averaging and binning techniques have been used with CCDs 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 often used in scientific imaging requiring 14-bit or 16-bit dynamic range.   The reason is the linearity and noise produced in a CMOS Active Pixel Sensor (APS) is not sufficient for these demanding applications, CMOS APS is becoming more mature, producing 10 & in some case 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. Some companies have named WDR as Extended, Extreme Dynamic Range (EDR), or Dual Slope.   All use the same technique. Majority of the high-speed sensors come from the two semiconductor companies. Therefore, the differences in dynamic range have more to do with pixel size, full well capacity, output conversion and the various noise sources. Claiming a camera has a 14-bit dynamic range simply because the camera has an ADC (Analog-to-Digital-Converter) of 14-bits is misleading because noise and the capacity of the pixel well to produce such a dynamic range has not be considered. My advise would be to closely look at what is being claimed as far as the dynamic range and make judgments on the image quality that you actual get from the camera.

The first area of discussion will be to explain the signal processing steps in converting light into digital numbers for both digital and analog imaging sensors. An explanation will be given on signal to noise ratio that determines the range of values a pixel can produce, also known as the dynamic range.   The next area of discussion will focus on CMOS sensor technology that is used in most high-speed megapixel cameras. An example will be given on what would be expected from a CMOS sensor to get a dynamic range to yield an SNR for 16, 14, 12, 10 and 8-bits.

This paper will provide end users of high-speed cameras with enough information to make informed decisions on camera specs and what to expect. Let's begin with a bird's eye view of a classical high-speed digital imaging system's components.

Imaging Chain – Digital & Analog Imaging Sensors

The building blocks required for a high speed imaging system are shown below. For our purposes, we will describe this imaging system as an imaging chain. The analogy is that a chain has many links and each link must work together to bind, as chain and the weakest link will affect the overall performance. This same analogy is true with a high-speed video system.
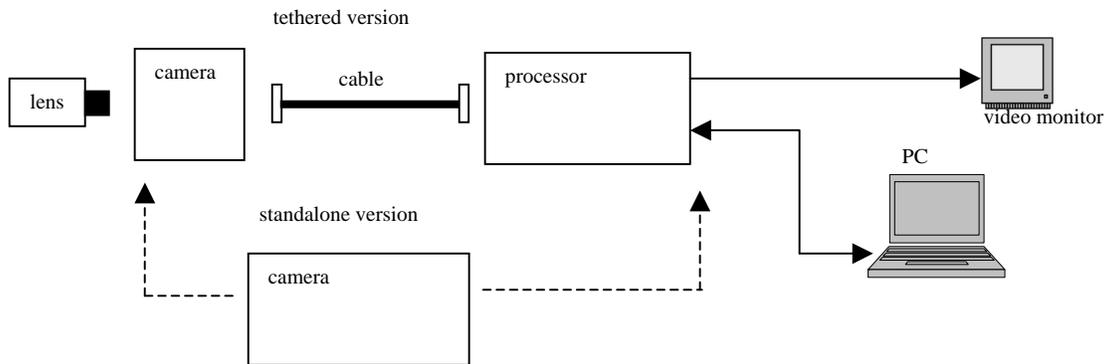

An important distinction can be seen in the two configurations, the tethered & standalone versions. A cable with analog video signals is missing with the standalone version. In fact, there is a third version not shown, a tethered version where the image output is digital from the back of the camera. 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 imaging chain. 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 sensor that has a digital output is not any better than a camera with a sensor that has an analog format. Let's take a look at why this is the situation.
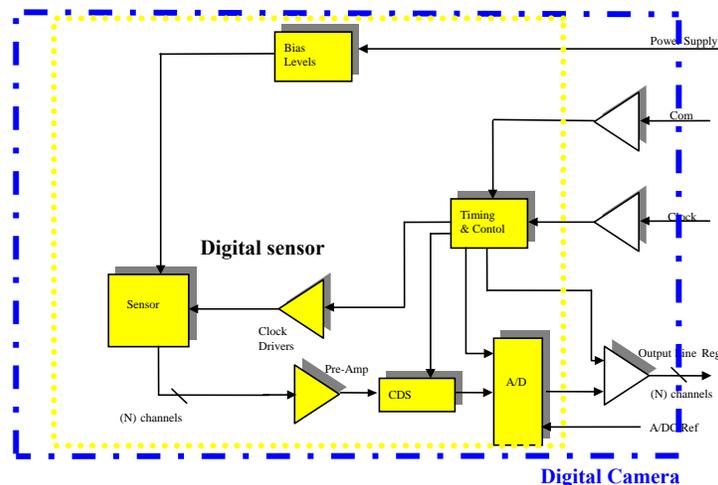
The blue outline represents a tethered digital camera that has a sensor, which is highly integrated digital sensor with a digital output. If digital memory, display controller and controller were added to the block diagram, we would have a standalone digital camera. The blocks represented in yellow are 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 outputs images digitized 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 or digital noise called quantization noise. Quantization noise is due to the finite resolution of the ADC, and is an unavoidable imperfection in all types of ADC. The magnitude of the quantization error at the sampling instant is between zero and half of one LSB. The signal is much larger than one LSB. The quantization error is not correlated with the signal, and has a uniform distribution. Its RMS value is the standard deviation of this distribution, given by

$$\frac{1}{\sqrt{12}}\text{LSB} \approx 0.289 \text{ LSB}$$
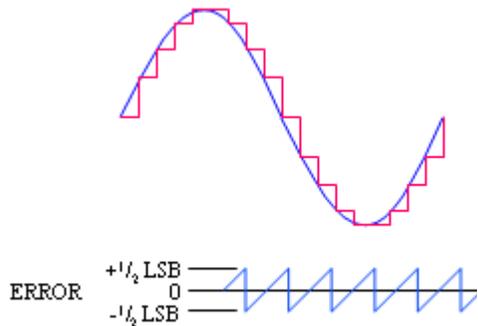
| A/D Converter Resolution (Binary Code)		(*) Value of 1LSB		(*)Value of 1/2LSB	
Number of Bits (n)	Number Of Increments (2ⁿ)	0 to +10V Range (mV)	+10V Range (mV)	0 to +10V Range (mV)	+10V Range (mV)
16	65536	0.152	0.305	0.076	0.152
12	4096	2.44	4.88	1.22	2.44
11	2048	4.88	9.77	2.44	4.88
10	1024	9.77	19.5	4.88	9.77
9	512	19.5	39.1	9.77	19.5
8	256	39.1	78.2	19.5	39.1

ERROR +¹/₂ LSB 0 -¹/₂ LSB

(*) Note: scale 0 to 10 volt range to 0 to 1 volt for video

And the percentage of quantization error for various ADCs can be seen in the following table.

ADC Res	8	9	10	12	14	16
Q error (%)	0.113	0.056	0.028	0.007	0.0017	0.0015

The above table assumes that there are no missing codes in the ADC. Missing codes sometimes happens due to design errors or processing errors. If a voltage ramp is applied to the input of an ADC, the output of the ADC should increment up as if it was a counter. However, if there are missing codes, the sequence of counting would be broken by having a count repeat or no output at all. This is what you would see with a missing code. This type of error is far greater than a quanitization error since it is at least a full bit.

It is true that the quantization error becomes negogable after 12-bit, however, this is not the major noise source found in your image.
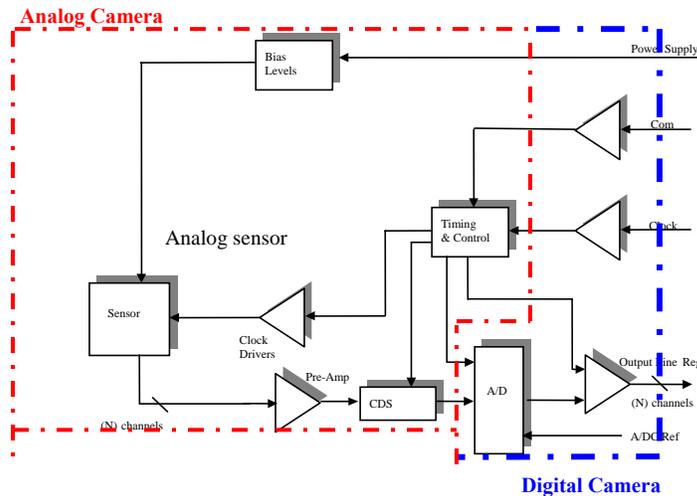
Above is the same sensor block diagram as before except this block diagram shows what would be included in an analog camera (red dash line). Also, shown are the blocks that would be added to make a tethered digital camera. As discussed previously, adding digital memory, a display controller and a controller, 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 optimized 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 on to the meaning of SNR and the resulting 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. 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).

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

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-). Now, 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:

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

Assuming the image could be read from the sensor with no other noise, a quantum limited design, 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. And, to assume that the camera signal processing channel is quantum limited is a big leap of faith.

Previously, the noise floor was characterized as dominated by read noise and/or shot noise. There are other noise sources. To illustrate the source and operational contribution, the following table will describe the noise in a concise manner. Regardless if it is a digital sensor or analog sensor, the sources are the same for the sensor.

	Dark	Illuminated	
		Below Saturation	Above Saturation
Fixed Pattern Noise (FPN)	Dark Signal Non-Unif. / Pixel Random Noise / Shading	Photo Response Non-Uniformity / Pixel Random Noise / Shading	
	Dark Current Non-Uniformity / Pixel Wise FPN / Row Wise FPN / Column Wise FPN		
	Defects		
Temporal Noise	Dark Current Shot Noise	Photon Shot Noise	
	Read Noise (noise floor) / Amplifier Noise / (Reset Noise)		
			Smear, Blooming
	Image lag		

Table Source (1)

Another way to look at the relationship of signal and noise can be seen in the following graphs illustrating signal, noise, dynamic range & electrons.
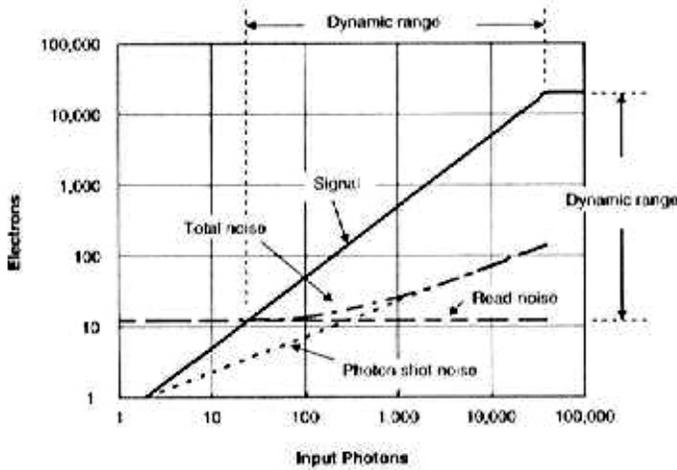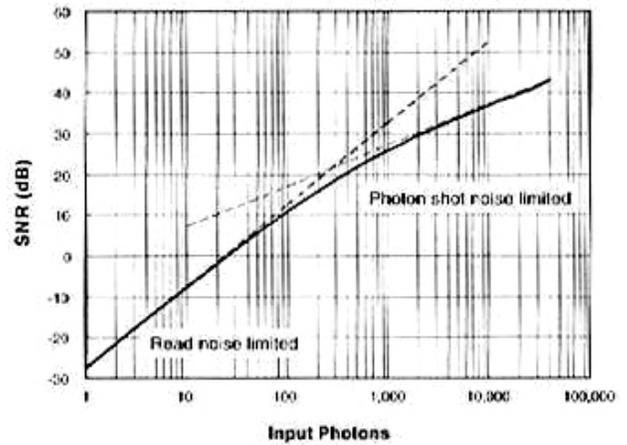
Graph Source (1)   Graph Source (1)

You can see in the right graph the read noise is dominant when the light level is very low (100 < photons) and the shot noise is the main contributor at higher light levels above several hundred photons. The graph on the left is an excellent representation of the conversion process, relating dominant noise sources, linear signal levels to saturation, the dark level and the dynamic range in electrons or photons. You may recall that the quantization error after 12-bits we discussed previously is considered small, yet the dominant noise will still be digitized as shown in these graphs and represented as the lower ADC bits. This is noise data and not image signal data. So, these lower bits represent erroneous information. So, why would you think 14-bits is so much better than 12 bits?

What we have shown up to now has to do with just the noise associated with the sensor in the analog domain. Upon reading the analog signal representing the image, other noise sources come into the equation for SNR. The various sources for additional noise are shown in the block diagram below.

Analog sensor   Digital sensor
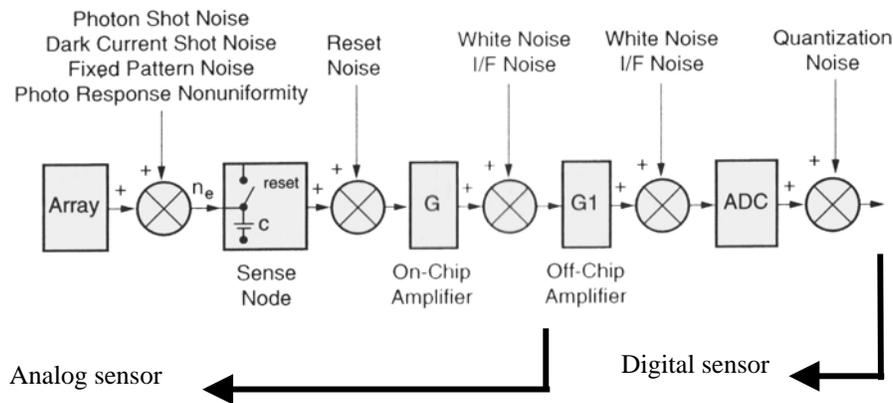
Graphic Source (1)

Therefore, the system noise would be expressed as:

$$(n_{sys}) = \sqrt{(n_{shot})^2 + (n_{pattern})^2 + (n_{reset})^2 + (n_{on\text{-}chip})^2 + (n_{off\text{-}chip})^2 + (n_{ADC})^2}$$

(1) Referance -"Image Sensors & Signal Processing for Digital Still Cameras" CRC Taylor & Francis, pg 67.

Some noise sources can be minimized, such as the reset noise using a correlated double sampler (CDS) in the signal chain. The simplified model for sensor noise can be written as:

$$(n_{sys}) = \sqrt{(n_{shot})^2 + (n_{pattern})^2 + (n_{floor})^2}$$

The system noise is expressed in units of root mean square, (RMS) and measured as the standard deviation. And the system SNR is expressed as:

$$SNR_{sys} = signal \ (e\text{-}) \ / \ \sqrt{(n_{shot})^2 + (n_{pattern})^2 + (n_{floor})^2}$$

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 a 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 images on a PC

Many image file formats (JPG, BMP, TIFF8 etc...) support only 8-bit grayscale. Millions of colors can be encoded, but each color plane supports only 256 steps. Today's 24-bit PC video card technology is incapable of displaying more than 256 shades of grays on the monitor. There is no great incentive to display more than 256 shades of gray since 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.

Mapping 12-bits to lower 8-bits

4096 steps

256 steps
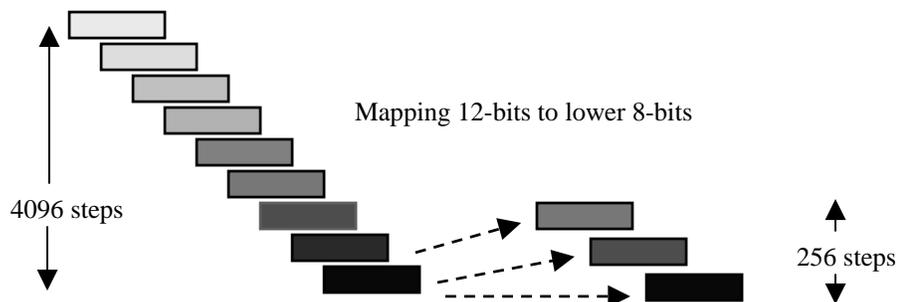
The advantage of having the ability to map HDR images is that more detail can be viewed when looking at dark areas or bright areas within the image. Having additional dynamic range also eases the lighting requirements since the image can be underexposed. As an example, ¼ the light level used to capture 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 ¼ 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.

Display FAQs

- LCDs will have a better contrast than Plasma.

- RGB settings are typically 0.55 or the inverse of 1.8

- Contrast ratios are very important for good looking color images on LCD displays. As an example, a LCD monitor with 200:1 contrast can not display an 8-bit image. Use LCD monitors with 400 to 512:1 contrast ratio.

- LCD displays should have a brightness setting of at least 400 NITS for outdoor use (800 NITS is better). LCDs are brighter than Plasma displays

- Video (analog) displays have a different gamma setting than RGB computer monitors (typically 0.45 or the inverse is 2.2)

Comparing two camera's responsivity

In order to compare two cameras responsivity, you need to have data on how the cameras respond under exposure. You also need to have data about the camera's imaging sensor to have a complete understanding.   Typically, a camera manufacture will provide a spectral response curve that shows the output response plotted against wavelength. The unit for camera response is DN/(nJ/cm2). The unit DN stands for Digital Number, which is the digital pixel value and (nJ/cm2) is the level of illumination coupled with the exposure time in the camera. Another common unit used for camera response is lux-sec. Illumination can be expressed in lux or candelas. However, to state correctly the responsivity of a camera, the incident light on the sensor has to be related to exposure time, hence lux-sec. To compare cameras using different units you must convert one system of units to the other.

The conversion for 1 lux-sec = 0.159 nJ/cm2.   Most often, a camera's responsivity with be expressed at a given wavelength of light, such as 20 DN/(nJ/cm2) @600 nm). In some cases, color temperature or the type of illumination will be specified. An example would be 20 DN/(nJ/cm2) using a halogen light source with a color temperature 3200K. Once again, to compare camera's responsivity, data from the camera must be known, and the units used must be the same for both cameras.

Responsivity data (full well, sensor responsivity, charge conversion efficiency, and dynamic range) alone may not be enough to correctly compare two cameras. Sensor data is required to make valid comparisons. As an example, the sensor in one camera quotes a responsivity specification of 32 DN(nJ/cm2).   Another manufacturer's camera states a responsivity specification of 60 LSB(nJ/cm2) where LSB (Least Significant Bits) is the same unit as DN. Clearly, the camera with 60 DN should be better than the one that has 32 DN.   Which camera would perform best? This is when we need to look at the sensor data to fully understand which camera would perform best. The responsivity of the sensor used in the first camera is specified as 20 Volts/(microJoule/cm2). The 2nd camera is specified as 8 Volts/(microJoule/cm2).   Even though the 2nd camera appears to have a responsivity of 2X that of the first camera, actually the first camera's sensor has a higher responsivity than the 2nd camera.   How is it that a better sensor appears to have a worse responsivity than the 2nd camera?   The answer is that the first camera has a sensor with a full well capacity of 200,000 electrons, dynamic range of 5000 (~ 86 dB), and a VSAT level of 3 volts. The 2nd camera has a smaller pixel, which has a lower full well capacity of 70,000 electrons, a dynamic range of 2000, and VSAT level of 0.7 volts. Therefore, drawing conclusions that the 2nd camera is a far better camera is not correct, since the first camera has a huge dynamic range with the well capacity and VSAT output.   To produce a higher responsivity number, the first camera's internal gain would be increased resulting in a lower dynamic range but there is so much dynamic range, this is not a factor. The results would be the first camera could be gained up by a factor of 4 times, providing a responsivity of 128 DN/(nJ/cm2).   When comparing camera responsivity based only on the camera's responsivity numbers, you need to also have the sensor specification to reach a final conclusion.