White Paper

High Dynamic Range Imaging

High Dynamic Range Imaging for Machine Vision

What is Dynamic Range?

Dynamic Range is the term used to describe the difference between the brightest part of a scene and the darkest part of a scene at a given moment in time – in other words, the amount of contrast within a single image. Dynamic range is typically measured in bits or dB. For example, a scene where the brightest point is roughly 1000 times brighter than the darkest part of the scene would be said to have approximately 10-bits of dynamic range or ~60 dB.

Note that this instantaneous measurement of contrast is different than brightness that changes over time, which can typically be handled by some sort of auto-exposure capability.

An evenly-lit scene with no shiny objects or light sources in the field of view would be considered to have low dynamic range. In such a scene, the brightest point would generally be less than 255 times brighter than the darkest part of the scene, thus enabling all parts of the scene to be accurately represented in an 8-bit digital image. Cameras with up to 12-bit imaging capabilities are fairly common in the machine vision world, enabling scenes to be captured with dynamic ranges approaching 72 dB (4096:1).

How is Dynamic Range Handled by a Digital Camera?

The pixel values associated with a digital image are determined by the amount of light (photons) that strikes each physical pixel on the camera's imager during the exposure period. Photons can come from light that is reflected off objects in the scene (reflected light), or can come directly from light sources that are within the camera's field of view (emitted light).

The photons cause the pixel well to fill with some amount of electrons which are then measured as an electrical voltage to determine what value to assign to that pixel. For a scene with a low amount of contrast, an exposure time can be selected such that none of the pixel wells are completely empty and none of the pixel wells are completely full at the end of the exposure period. This enables each pixel in the image to be assigned a value that accurately represents its brightness relative to other pixels in the image.

What is High Dynamic Range Imaging?

In scenes with high levels of contrast (brightness/darkness) it may be difficult or impossible to find an exposure time that properly accommodates the range of photons striking the physical pixels. Some pixels that are receiving photons from light sources or other highly reflective areas in the scene may completely fill up ("saturate") before the end of the exposure period. All of these pixels will thus be assigned a pure white value, regardless of their relative brightness to one another. Reducing the exposure time to avoid this may cause pixels in the darker portions of the image to not receive any photons, thus all of these would be assigned a zero (black) value regardless of their relative brightness.

High Dynamic Range (HDR) imaging refers to the ability to capture such high contrast scenes in a way that maintains the relative order of brightness between all pixels, with few or no bright pixels saturating or dark pixels being assigned a zero value.

<u>Example of a high dynamic range scene</u>

The four images below provide an example of a high dynamic range scene. Note how there are details in each exposure that may not be visible in a different exposure (such as the clouds in the sky, or the shadows behind the columns on the building façade). None of the exposures can capture the entire scene without significant portions of it either under-saturating (turning black) or over-saturating (turning white). This indicates that the contrast of the scene is beyond the dynamic range of the camera.

Figure 1 - "StLouisArchMultExpEV-4.72" by Kevin McCoy - Own work. Licensed under CC BY-SA 3.0 via Commons - https://commons.wikimedia.org

High Dynamic Range Imaging Methods

There are many different ways to handle scenes with High Dynamic Range. The three most common approaches are:

- Sequential image fusion
- Multi-slope pixel integration
- Dual-sensor image fusion

Each of these methods will be summarized in subsequent sections.

Linear vs. Non-linear High Dynamic Range

Before we describe the different high dynamic range methods, it is helpful to understand the two fundamental ways that high dynamic range data is captured and presented.

For example, in a scene with 14-bits of dynamic range, the darkest point in the scene would reflect enough light to warrant a pixel value of 1 and the brightest areas would emit or reflect enough light to have a value approaching 16,384 (2^{14}). To accurately assess the true intensities of every pixel in the image, a camera would have to be able to output any pixel value between these two extremes. A graph representing the relationship between possible pixel values recorded on the camera's imager and the values that are output to the host PC is shown below.

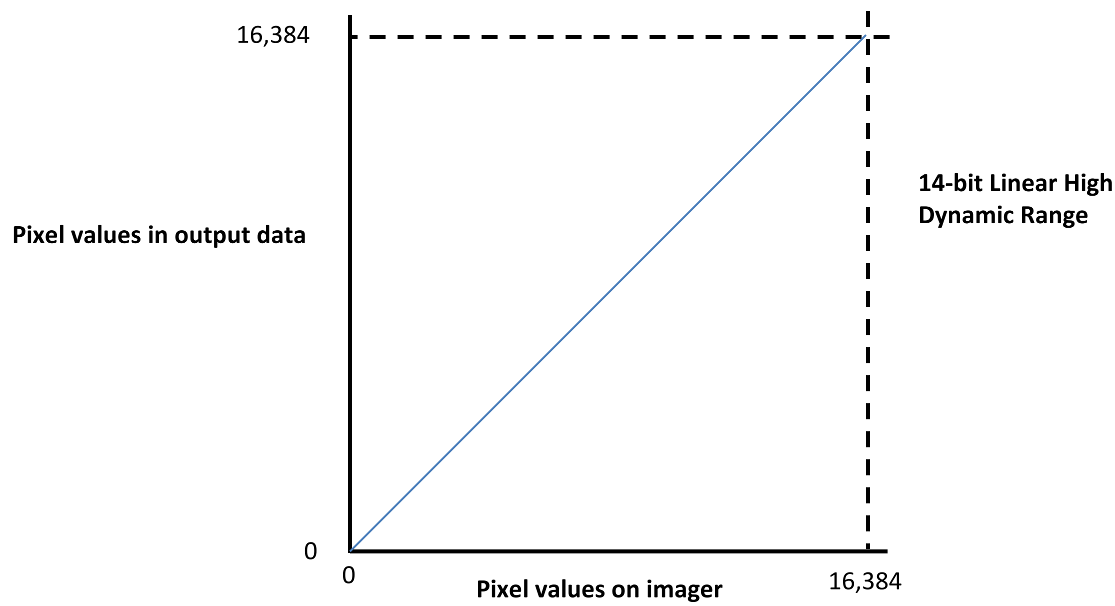

Figure 2 - Linear dynamic range

Such a completely "linear" relationship can be extremely important in applications where the exact brightness values of specific points in the scene are needed. Astronomy and other scientific applications often require this type of precision.

For many applications, however, exact values are not necessary or even desirable. Consider the fact that a standard computer monitor has a contrast ratio of 1000:1. That equates to a 10-bit (60 dB) dynamic range. A top-notch monitor might have a contrast ratio close to 2000:1. Even so, that is only 11-bits (66 dB) of dynamic range. That means there is no way to precisely display a 14-bit HDR image. Instead, a technique called "piece-wise linear" compression is typically used. This approach takes advantage of the fact that most of the pixels in a given HDR scene fall within a much narrower range.

In our 14-bit linear example above, it is quite likely that most of the scene could fit within, say, 8-bits, or the first 256 values of the graph. A small number of bright areas in the image are then sparsely scattered across the remaining 16,128 values at the "bright" end of the graph (see below).

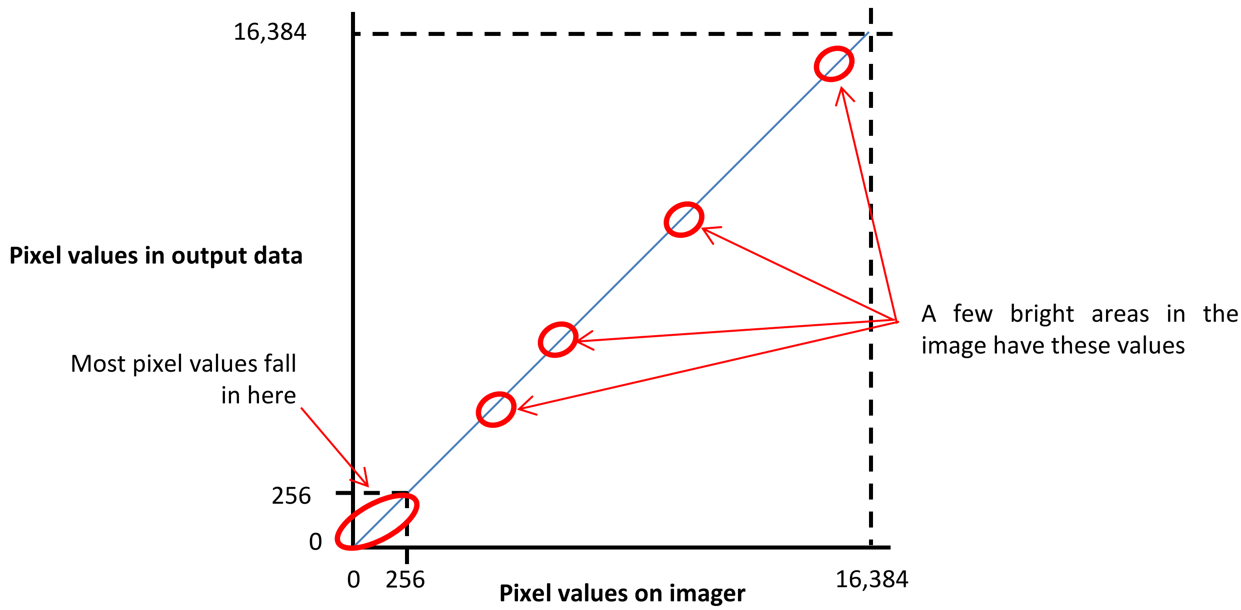

Figure 3 - Distribution of brightness values

To display this image in the 10-bits of dynamic range available on our standard monitor, we can "break" the portion of the line above the first 8-bits of data into several "pieces" that we then "bend" to fit into the top two bits of the dynamic range we have available. In the graph below we have made this "piece-wise" adjustment.

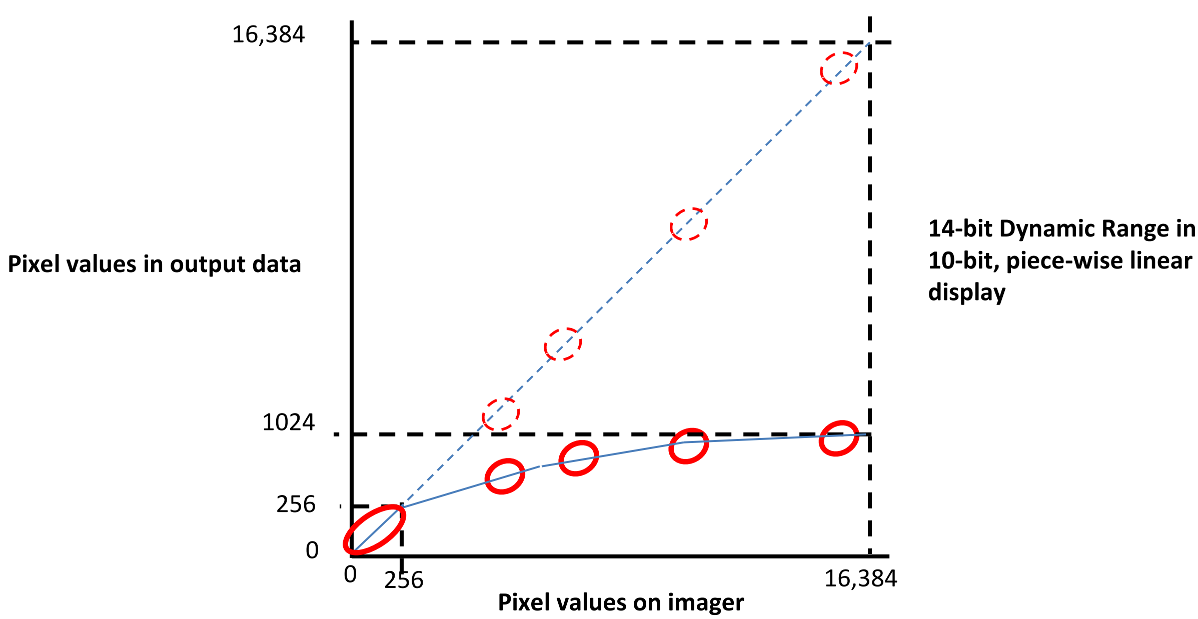

Figure 4 - Piece-wise linear compression

A piece-wise linear approach such as this maintains the hierarchy of brightness between pixels – the brightest parts of the image still have the highest output values. But it is considered a "non-linear" approach because these bright pixels are "compressed" into a much narrower set of total values than they would be if they were handled in a strictly linear fashion.

Now, our 14-bit image can be displayed on a standard monitor, and detail can be seen in both the bright and dark areas of the image. But if two bright areas are side-by-side, it can be difficult for the eye (or the PC) to distinguish sharp edges or transitions, and it is certainly not possible to do any accurate measurement of the true brightness values for these "compressed" or "non-linear" pixels.

The graph below shows the same piece-wise linear treatment, but with the axes adjusted to a logarithmic scale so that the compression of values can be more easily portrayed

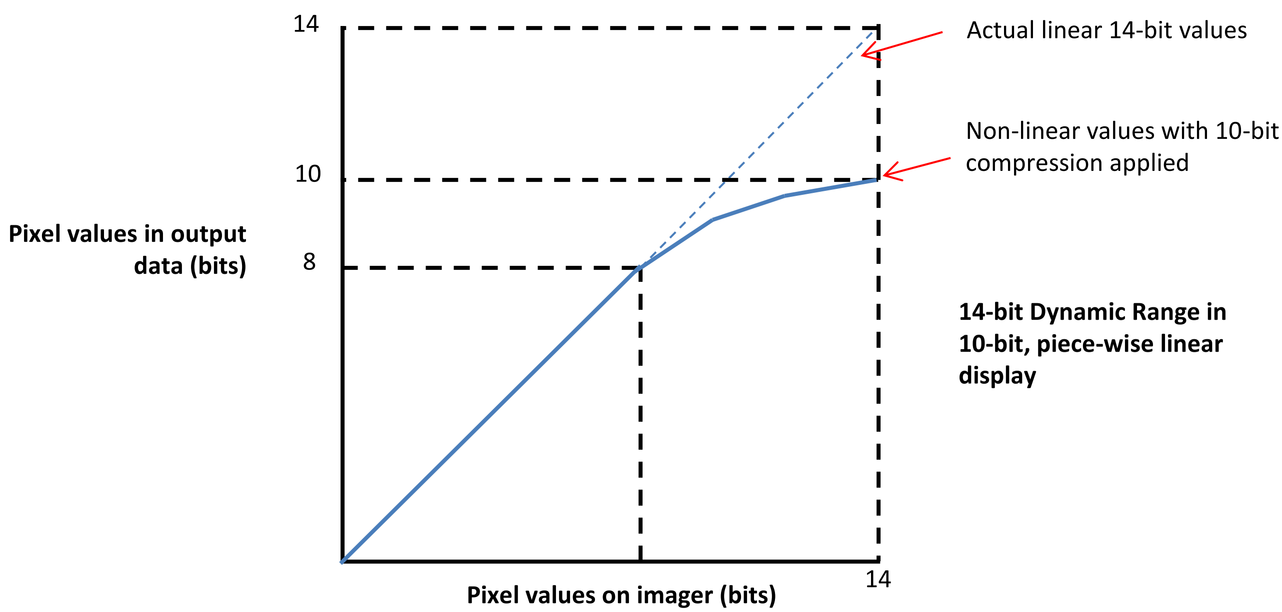

Figure 5 - Piece-wise linear compression on a logarithmic scale

Depending on the HDR method used, you may be able to achieve linear results, piece-wise linear, or both.

HDR Method 1 – Sequential image fusion

This method has been used by digital still photographers for many years. The photographer takes two consecutive images of the same high-contrast scene, quickly switching to a different exposure for the second image. The exposures are selected in the following way:

For exposure 1, a relatively slow shutter speed is selected such that the darker areas of the scene produce non-zero pixel values, ideally values that can be easily distinguished from the camera's own "dark noise" values.

For the second exposure, a shutter speed is selected such that the brightest *non-saturated* pixels in the first image produce small, but non-zero, values in the second image. These pixels are used to determine the "overlap" point between the two images.

Software is then used to "fuse" together the two images. Using the overlap point as a reference, saturated pixels in the first image are replaced with the corresponding non-saturated pixel values from the second image. Using this technique, two 8-bit images can easily handle the previously-described 14-bit scene and generate accurate linear data. A compression function can be applied after the fact in order to display the result in a non-linear/piece-wise linear fashion on screens with lower contrast ratios.

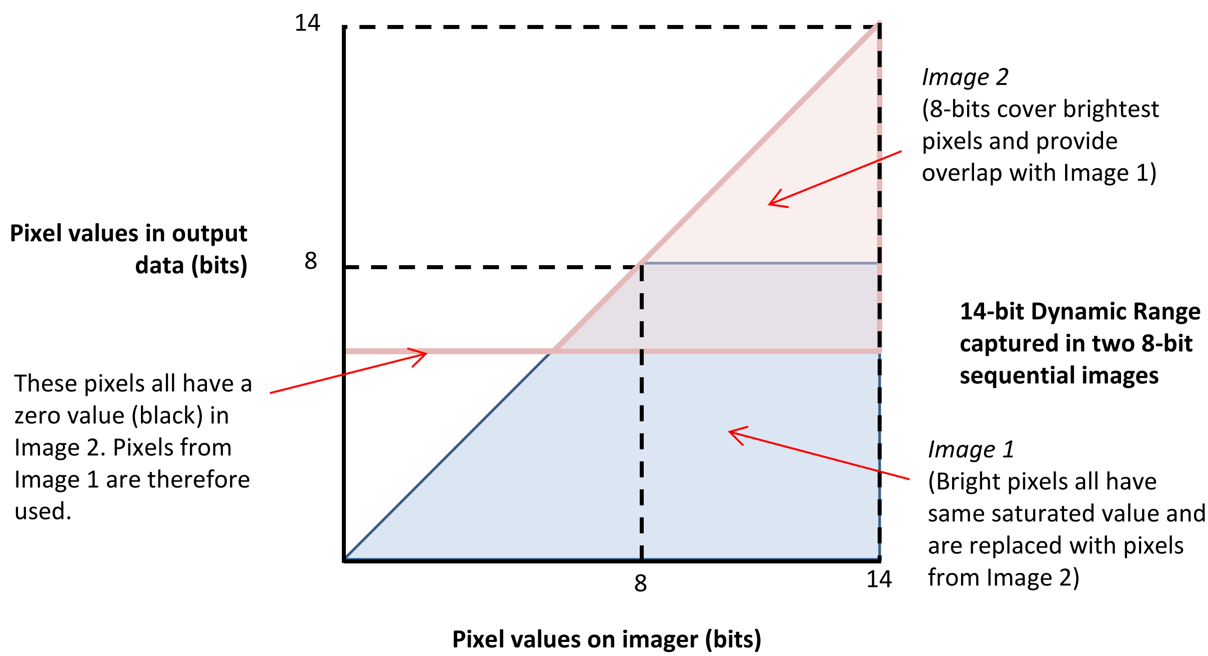

Figure 6 - Dual-image overlap and fusion

For still photographers, this technique is typically used with landscapes or other scenes with little or no movement. Thus, even if it takes a second or two to change the exposure, the two images can be fused together without any spatial offset occurring between the two images.

At first glance, this method appears to be unusable for machine vision applications, especially those involving rapid movement. But the development of a camera function called "Sequence Trigger" along with continued advances in camera frame rates, has made this method a possibility for some machine vision uses.

The Sequence Trigger function, which is available in most JAI cameras, allows users to predefine two different exposure settings that can be alternated at every trigger point. This means that a camera such as JAI's SP-5000-CXP4, which boasts the ability to provide 5-megapixel images at over 250 fps, can accept consecutive triggers at an interval of roughly 1/250 of a second. Putting this another way, the camera can produce pairs of bright and dark exposures at an effective rate of roughly 125 times per second, provided enough CPU power is available, a speed that is fast enough to avoid a 1-pixel spatial shift between exposures in many slower moving machine vision applications.

To support such uses, the JAI SDK and Control Tool software is equipped with a built-in function for automatically “fusing” together the image pairs produced by an alternating Sequence Trigger. This enables customers to utilize a wide range of cameras, with different speeds and resolutions, to quickly and easily generate high dynamic range images.

Summary – Sequential Image Fusion

Advantages: Easy to use; wide range of cameras; linear and non-linear data output possible.

Disadvantages: Consecutive images may produce spatial distortion (and poor fusion results) in fast-moving applications.

HDR Method 2 – Multi-slope pixel integration

Improvements in shutter technology and image quality have made CMOS imagers increasingly popular for machine vision applications. This has also brought with it a second option for high dynamic range imaging. This is because the transistor-based pixel architecture of CMOS imagers enables a higher degree of on-chip functionality to be built into the imager than in the traditional CCD architecture.

One of these added capabilities involves the ability to selectively “reset” individual pixels during the exposure period based on whether or not they have already reached the saturation point. This enables a piece-wise linear (non-linear) high dynamic range image to be captured in a single exposure and at the full frame rate of the camera, provided the imager has sufficient circuitry to handle the pixel reset process and image readout simultaneously (see note at end of section).

The method is typically called multi-slope or dual-slope HDR, which refers to the way a graph of a single pixel’s “integration” chart looks when this method is applied. For example, consider the case of a very bright pixel in an image as it fills with electrons during the exposure/integration period. In the graph on the following page, the Y axis represents the fill level of the pixel, while the X axis represents the exposure time. (Note: while this graph looks similar to the preceding graphs, it is showing the behavior of a pixel *over time* rather than how multiple pixels react to different light intensities.)

The dashed horizontal line shows the maximum fill level (or voltage) for the pixel, which translates to a digital value of 255 in 8-bit mode, 1023 in 10-bit mode, or 4095 in 12-bit mode. If the pixel reaches this point before the end of the exposure (as shown here), any additional electrons that are generated “overflow” the pixel and are discarded. The pixel value is simply presented as the maximum number regardless of how much brighter than the maximum the pixel should really be.

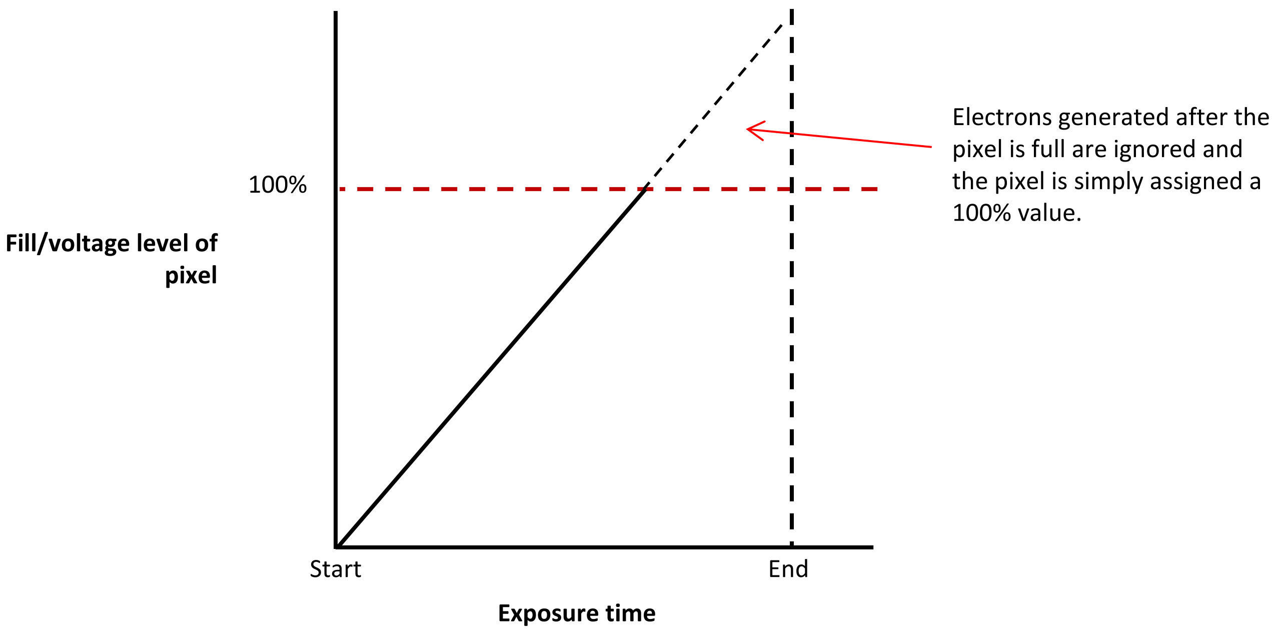

Figure 7 - Fill level of bright (saturated) pixel during exposure period

For HDR imaging, the objective is to keep bright pixels from saturating, without "darkening" the rest of the image, thus increasing the overall contrast ratio. With the multi-slope method, this is achieved by establishing a reset point during the exposure time and also a reset level (sometimes called a "skimming" level). The following graph shows how this would work for the hypothetical bright pixel in our example.

This time, we add a reset point at approximately 90% of the full exposure period. We set our reset level at around 80% of a full pixel.

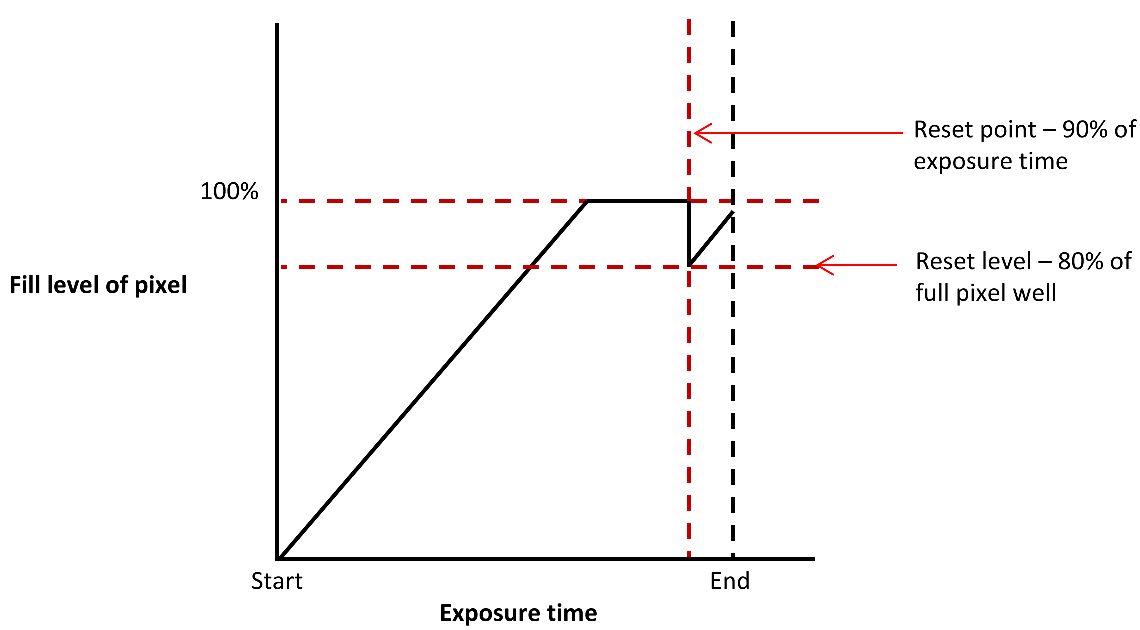

Figure 8 - Exposure period with reset point and reset level

At the reset point, the voltages (fill levels) of all pixels are analyzed by the HDR function of the imager/camera. Any pixels with a level higher than the reset level (such as the one in our example) are instantaneously “reset” to the 80% level so that they can continue to accumulate electrons during the remainder of the exposure period. The effect is as if the pixel wells that were too full have been “skimmed” down to a lower level. In our example above, the bright pixel continues to receive photons and generate electrons at the same rate throughout the entire exposure period, but the reset process creates a second slope that finishes just below the 100% fill level, thus avoiding saturation.

To see how this can create a piece-wise linear HDR image, we need to look at multiple pixels in the same graph. The different color lines in the following graph represent 6 different pixels, each subject to a different level of brightness within the scene. The slopes of the lines indicate the brightness of the pixels. Bright areas in the scene cause pixels to fill with electrons faster. These pixels have steeper slopes. Darker areas cause pixels to pixels fill slowly, creating slopes that are less steep.

Each pixel maintains its slope during the exposure period. Like the previous example, the bright pixels are reset once during the exposure period to keep them from saturating.

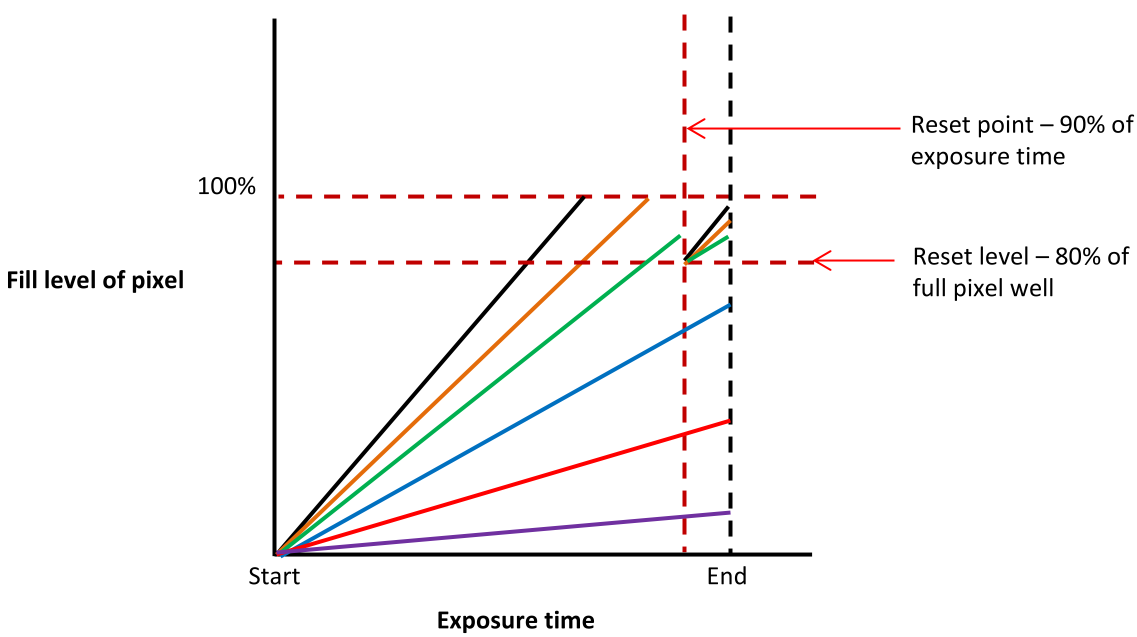

Figure 9 - Multi-slope HDR compression applied to six pixels of varying intensities

Notice how all the darker pixels accumulate electrons normally, with no reset required since they are not above the reset level at the reset point. The three brighter pixels are all reset to the same value (80%) at the reset point. But because they accumulate electrons at different rates (slopes) based on their brightness within the scene, by the end of the exposure they have reestablished their relative relationship, albeit in a much more compressed fashion. More importantly, this has been accomplished in a single exposure, eliminating any concern about pixels shifting between two consecutive images.

Changing the reset point and/or reset level will produce different results which might be needed to best accommodate a wider or narrower range of pixel slopes. Cameras that support this method of HDR let users control the process in one of several ways. For example, they may let users simply select the overall amount of compression they want in the final image. Choosing "200%" compression would mean that the user wants the brightest pixels in the image to handle 2X the normal number of photons/electrons without saturating – thanks to the reset process. Choosing "400%" would imply a 4X increase, and so on. JAI's SP-5000 and GO-5000 cameras, for example, let users choose from one of four possible HDR levels representing 200%, 400%, 800%, and 1600% compression (see graph below).

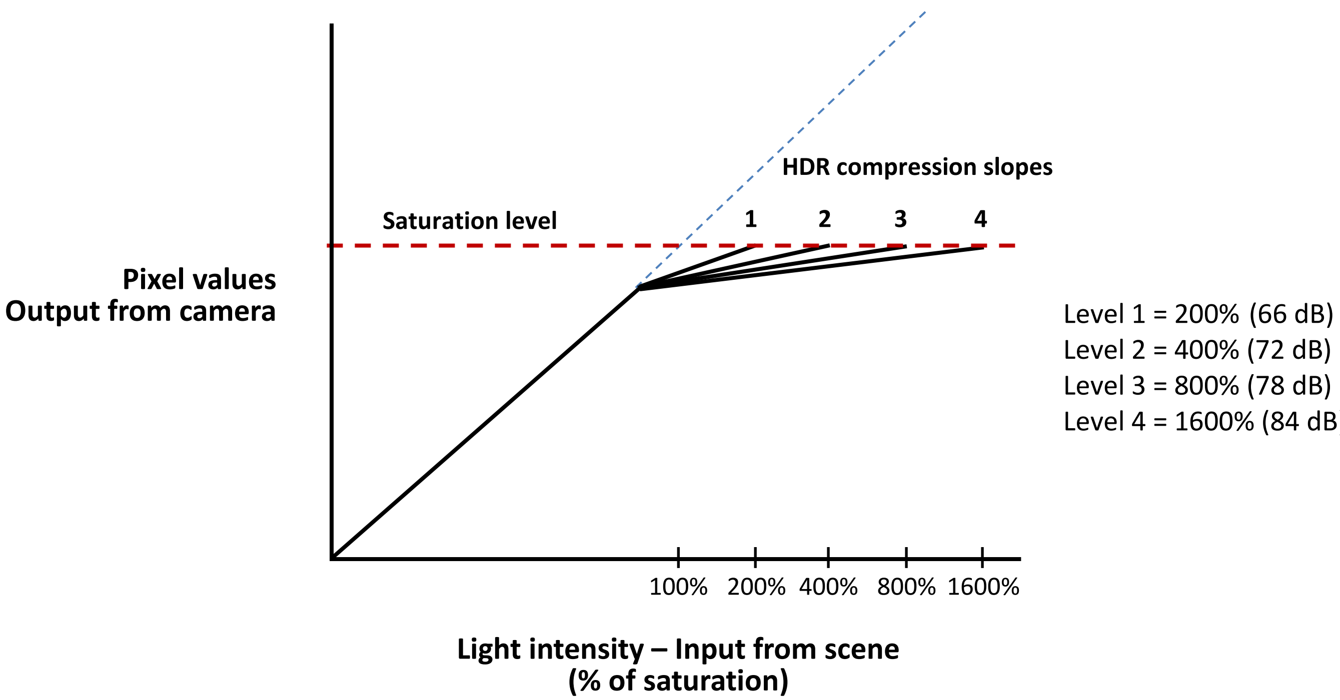

Figure 10 - HDR compression choices in SP-5000 and GO-5000 cameras

Other cameras give the user the ability to define one or more "knee points" defined by the percentage of exposure time and percentage of saturation voltage (fill level) they represent. While this gives the user even more control over the amount of HDR compression to apply, it is more complex for the user. The HDR mode in JAI's SP-12000 camera lets users define two knee points as shown on the following page.

For more details about HDR settings, see the camera manuals for these products.

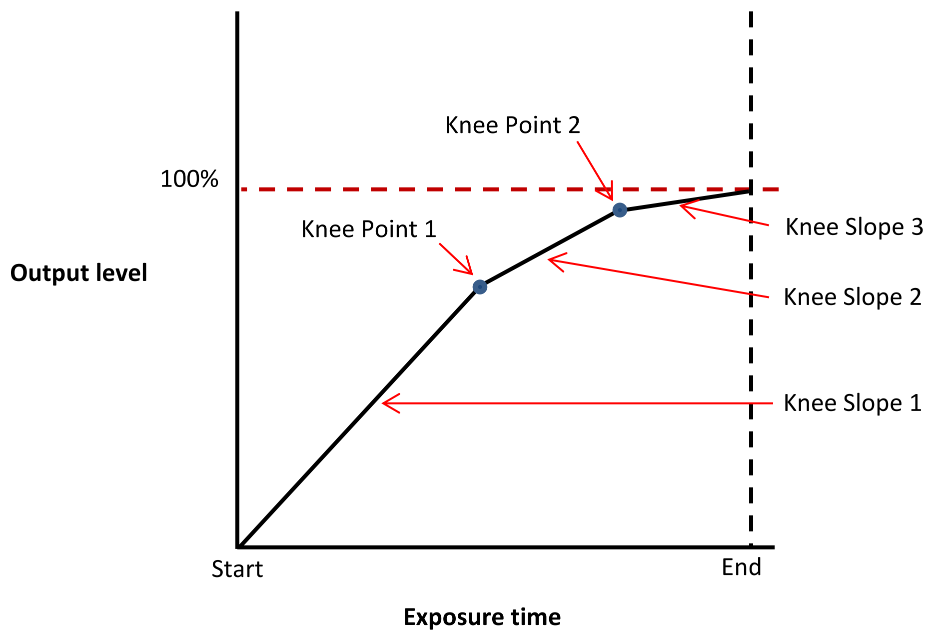

Figure 11 - HDR setting process for SP-2000 and SP-12000 cameras

As noted at the start of this section, the multi-slope pixel integration method offers the potential for capturing an HDR scene in a single exposure at the full frame rate of the camera. But the pixel reset process described above requires the use of pixel-level circuitry during the exposure period to check saturation levels and reset the appropriate pixels at the appropriate time. This extra processing necessitates a performance compromise in some CMOS imagers whereby readout of a previous frame cannot be overlapped with exposure of the next frame. Instead, these two processes must be handled consecutively, resulting in a reduced frame rate similar to the image fusion method.

However, even if the frame rate is reduced, the multi-slope integration method is still often a better choice for applications involving movement, especially those involving triggering. That is because this method captures the scene in a single triggered image, meaning there is no concern over spatial shifting of pixels between two consecutive triggered images. Thus, provided there is some spacing between objects, the actual frame rate becomes less important.

JAI offers cameras with both types of CMOS imagers – those that can maintain their full frame rate while performing multi-slope integration, and those that perform multi-slope pixel integration at reduced speeds. See Appendix for a listing of cameras.

Summary – Multi-slope pixel integration
Advantages: Single exposure applies HDR to moving objects without errors caused by pixel shifts between multiple images. Function is performed on imager – no post-processing required.

Disadvantages: Limited number of cameras available with this function. Piece-wise linear output only – true linear HDR data not available.

HDR Method 3 – Dual sensor image fusion

One final method for capturing HDR images utilizes the unique capabilities of dual-sensor, prism-based cameras. In these cameras a beam-splitter prism divides the incoming light to two precisely-aligned imagers mounted to the back faces of the prism. Thus, both imagers are able to simultaneously capture identical views of the high contrast scene. The prism architecture of such a camera is shown below.

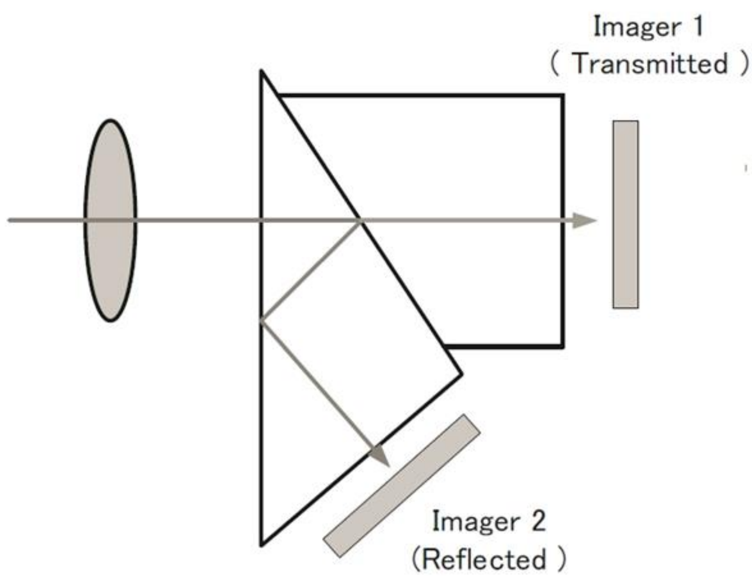

Figure 12 – Dual-sensor prism architecture

Separate electronics are provided enabling the user to apply different shutter and gain settings to each of the imagers. In this way, bright and dark areas in the image can be handled by two separate images – the same as in Method 1. The difference is that the prism arrangement enables both images to be captured at the exact same time. This eliminates any concerns over movement between exposures.

By reading out the two images and identifying the overlap point, true linear HDR values can be assigned to the pixels with a range that is nearly double the bit depth of the individual imagers. In other words, if the imagers are operating in 8-bit mode, linear HDR values of 15-bits or greater are possible. In 10-bit mode, values can span an 18- or 19-bit range.

As in Method 1, software can also be used to “compress” the linear values for displaying on a monitor or in a standard image format. This can be done on the host PC, however JAI’s line of prism cameras also feature built-in functions for performing the image fusion and compression inside the camera.

JAI's AD-080, AD-131, and AD-132 models provide users with several pre-defined options for how to fuse together and output their HDR images. All of these models utilize CCD imagers for maximum image quality. Frame rates are significantly lower than CMOS-based cameras.

Summary – Dual sensor image fusion

Advantages: Simultaneous exposure of both images supports applications involving movement. Linear and/or piece-wise linear output is available for processing.

Disadvantages: Limited number of cameras available with this function. Prism-technology is more expensive than single-imager cameras. Currently not available with faster CMOS imagers.

Conclusion – multiple HDR options support many applications

When the contrast ratio of a scene is large, it can be difficult to capture all the brightness levels in a single image. High dynamic range methods allow cameras to apply different exposures to the bright and dark areas of the scene, either at the image level or at the individual pixel level, to capture details while minimizing or eliminating the oversaturation or under-saturation of pixels.

Depending on the requirements of the application, HDR data can be provided in linear form with the exact relationship between all pixels maintained, or can be provided in a piece-wise non-linear form where the hierarchy of brightness is maintained, but in a compressed value range for the brighter pixels.

Slow shutter

Fast shutter

HDR image

Figure 13 - "StLouisArchMultExpEV-4.72" by Kevin McCoy - Own work. Licensed under CC BY-SA 3.0 via Commons - https://commons.wikimedia.org

Appendix

JAI camera models supporting Sequential Image Fusion

The following JAI cameras include a Sequence Trigger function, which can be integrated with the JAI SDK sample application for Sequential Image Fusion. This approach can be used with monochrome, Bayer color, and RGB/YUV interpolated color output.

(As of 15 March 2016)

Model name	Output type(s)
AT-200GE	3CCD RGB
AT-140GE	3CCD RGB
CV-M9GE	3CCD RGB
BB-500GE / BM-500GE	Bayer color / Monochrome
BB-141GE / BM-141GE	Bayer color / Monochrome
CB-200GE / CM-200GE	Bayer color / Monochrome
CB-140GE / CM-140GE	Bayer color / Monochrome
CB-080GE / CM-080GE	Bayer color / Monochrome
CB-040GE / CM-040GE	Bayer color / Monochrome
CB-030GE / CM-030GE	Bayer color / Monochrome
CM-030GE-RH	Monochrome
GO-5000C-PGE / GO-5000M-PGE	Bayer color / Monochrome
GO-5000C-PMCL / GO-5000M-PMCL	Bayer color / Monochrome
GO-5000C-USB / GO-5000M-USB	Bayer color / Monochrome
GO-2400C-PGE / GO-2400M-PGE	Bayer color / Monochrome
GO-2400C-PMCL / GO-2400M-PMCL	Bayer color / Monochrome
GO-2401C-PMCL / GO-2401M-PMCL	Bayer color / Monochrome
SP-20000C-CXP2 / SP-20000M-CXP2	Bayer color, RGB / Monochrome
SP-20000C-PMCL / SP-20000M-PMCL	Bayer color / Monochrome
SP-20000C-USB / SP-20000M-USB	Bayer color / Monochrome
SP-12000C-CXP4 / SP-12000M-CXP4	Bayer color / Monochrome
SP-5000C-CXP4 / SP-5000M-CXP4	Bayer color / Monochrome
SP-5000C-CXP2 / SP-5000M-CXP2	Bayer color, RGB / Monochrome
SP-5000C-GE2 / SP-5000M-GE2	Bayer color, RGB, YUV / Monochrome
SP-5000C-PMCL / SP-5000M-PMCL	Bayer color / Monochrome
SP-5000C-USB / SP-5000M-USB	Bayer color / Monochrome

Note: AD-081GE, AD-131GE, and AD-132GE also have Sequence Trigger functions however these are dual-sensor prism-based cameras so dual-sensor image fusion is always recommended for HDR.

JAI camera models supporting Multi-slope Pixel Integration

The following JAI cameras feature CMOS imagers with a multi-slope HDR function.

(As of 15 March 2016)

Model name	Notes
GO-5000M-PGE	Monochrome. 4 preset slopes. Frame rate reduced.
GO-5000M-PMCL	Monochrome. 4 preset slopes. Frame rate reduced.
GO-5000M-USB	Monochrome. 4 preset slopes. Frame rate reduced.
SP-12000M-CXP4	Monochrome. 2 knee points.
SP-5000M-CXP4	Monochrome. 4 preset slopes. Frame rate reduced.
SP-5000M-CXP2	Monochrome. 4 preset slopes. Frame rate reduced.
SP-5000M-GE2	Monochrome. 4 preset slopes. Frame rate reduced.
SP-5000M-PMCL	Monochrome. 4 preset slopes. Frame rate reduced.
SP-5000M-USB	Monochrome. 4 preset slopes. Frame rate reduced.

JAI camera models supporting Dual-Sensor Image Fusion

The following camera models from JAI's Fusion Series contain beam-splitter prisms with two separate, co-site aligned sensors to enable simultaneous capture of the same scene at two different exposures.

(As of 15 March 2016)

Model name	Output Type
AD-081GE	Monochrome (0.8 megapixels)
AD-131GE	Monochrome (1.3 megapixels)
AD-132GE	Color (1.3 megapixels)