White Paper

Wide Dynamic Range Imaging

Abstract

Wide dynamic range cameras are quickly becoming mainstream in video surveillance. However, there are many myths and misconceptions about the very concept of Wide Dynamic Range (WDR). This paper explains what dynamic range is, how it affects image quality and discusses the advantages and limitations of WDR cameras.

Wide Dynamic Range Imaging

MICHAEL KORKIN, PH.D.
Vice President of Engineering

WHAT IS DYNAMIC RANGE?

By standard definition, dynamic range is a ratio between the largest and the smallest value of a variable quantity, such as light or sound. Difficulties in understanding and applying the dynamic range concept revolve around the question of its measurement.

Consider a task of measuring precipitation rate using a bucket. In heavy rain, the bucket would quickly overflow making it impossible to determine the largest value of precipitation: the results would be clipped at the bucket capacity level. In light rain, the bucket would sporadically receive a drop during one


Needs Shorter Capture Time   **Needs Longer Capture Time**

measurement interval and two during another, making the smallest value uncertain, or noisy. To increase the certainty of small value readings, a longer collection time is needed, but this does not work for the large values, it causes the overflow.

This basic example illustrates how measurement is in fact an information channel: it can convey, lose or distort information about the variable quantity, either at the top of the range, or at the bottom, or at both ends at the same time.

Consider a video camera as a measurement instrument. It measures the amount of light falling onto each of its millions of light-sensitive elements, or pixels, arranged in a two-dimensional array. Each pixel integrates the photon flow it
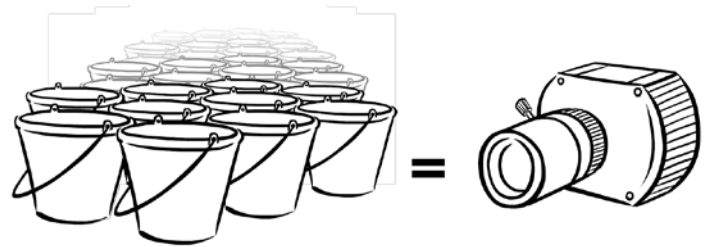
Camera as a Measurement Instrument

receives over a period of time, then converts it into an electrical signal to be read out. If the photon flow arriving from the scene is strong, or if the integration time is long, the signal may hit the limit and saturate (clip). Then, all luminosity variations that correspond to fine details of the bright areas of the scene, will be lost. Similarly, if the photon flow from the scene is weak, or if the integration time is short, the signal will produce uncertain, noisy readings, and all details of the scene will be lost.

As in any information channel, the quality of the video camera can be judged by how well it conveys the information, that is the luminosity variations present in the scene. In particular, is the camera able to capture subtle variations in the highlights of the scene without clipping? Is it able to capture subtle variations in the shadows without drowning them in noise? What about capturing scene details at both ends of the dynamic range at the same time? The answer to these questions depends both on the dynamic range of the scene itself and how it compares to the dynamic range capability of the camera as a measurement instrument. Generally, if the dynamic range of the scene is the same or narrower than that of the camera, the resultant images would faithfully convey scene details both in the shadows and in the highlights of the scene with no noise or clipping. If the dynamic range of the scene is wider, the camera would either clip details in the highlights, or mask details in the shadows by excessive noise, or do both at the same time.

Under Exposed   **Over Exposed**

Clipped shadows, details in highlights are visible   How to achieve the perfect balance?   Clipped highlights, details in shadows are visible

How can the dynamic range of the scene be evaluated independently of the camera's own capability? This can be done piecewise by taking multiple scene captures at different exposure times. Very long exposures reveal details in the shadows while saturating the highlights. Very short exposures reveal details in the highlights while drowning the shadows in the noise floor. The ratio of the longest exposure time required to reveal details in the shadows to the shortest exposure time to reveal details in the highlights provides a reliable estimate of the dynamic range of the scene.
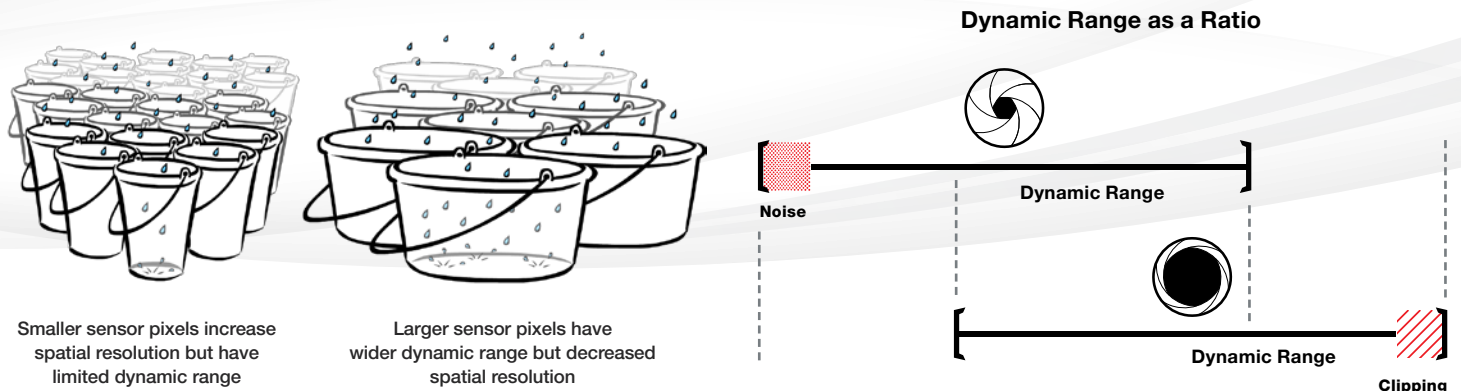
WHY IS DYNAMIC RANGE LIMITED?

Why are cameras limited in their dynamic range capability? The dynamic range limitations come from multiple causes, primarily from the physical properties of the light-sensitive pixels of the image sensors. Consider two identical image sensors that are different only in the area size of their light-sensitive pixels, and their well capacity related to the size. Which sensor would have a wider dynamic range? The answer is: the one with the larger pixels. On top of the range where the light flow is very strong they will not saturate as readily as the smaller ones, due to larger well capacity. On the bottom of the range where few photons are present, the larger surface area of sensitive pixels will gather more photons and reduce the uncertainty (noise) of the small value readings.

A conflict between the need to increase the size of sensitive pixels for achieving a wider dynamic range, and the need to make them smaller to increase spatial resolution, is always present. Often, spatial resolution which is nominally high due to small size of pixels is seriously compromised by their narrow dynamic range which causes either noise, or clipping, or both. Noise masks delicate signal variations containing fine image
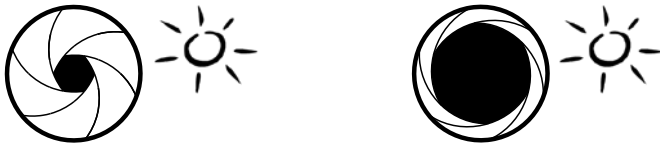
details in the shadows, while clipping erases the highlight details. Contributing to the conflict between the dynamic range and the spatial resolution is the temporal aspect of imaging – objects in the field of view may be moving, causing motion blur. Motion blur is a reducer of spatial resolution because it smears the moving objects over multiple adjacent pixels. A need to reduce motion blur demands shorter integration time, but that, in turn, raises measurement noise. Noise also destroys spatial resolution by masking subtle signal variations. To summarize, the effective sensor resolution may be lower than the nominal due to narrow dynamic range of small pixels, due to motion blur, and due to measures intended to reduce motion blur.

IT'S THE RATIO, NOT THE ABSOLUTE RANGE

An important attribute of the dynamic range concept is that it is defined as a ratio, and not as an absolute value. This means that different scenes with an identical dynamic range may have entirely different average luminosity, or the absolute range. The same effect of shifting the range up or down is produced by the camera's aperture which limits the amount of light reaching the image sensor. So, a camera specified to support a particular dynamic range will produce different results for different absolute ranges and different apertures because its own capabilities do have absolute limits, primarily due to physical size of its pixels. In particular, a scene with a given dynamic range but a lower average luminosity (or a smaller aperture) would produce more noise even if the dynamic range is fully supported by the camera, while the scene where the range is shifted upwards (or the aperture is larger) may produce clipping, and potentially additional noise in the mid-range, specific to WDR cameras.

Smaller sensor pixels increase spatial resolution but have limited dynamic range

Larger sensor pixels have wider dynamic range but decreased spatial resolution

Dynamic Range as a Ratio

Noise

Dynamic Range

Dynamic Range

Clipping

Aperture Setting

Less Light Enters
- Increases depth-of-field
- Worsens trade-off between motion blur and image noise

More Light Enters
- Reduces depth-of-field
- Improves trade-off between motion blur and image noise

It should be noted that the size of the aperture which limits the amount of light falling onto the sensor also affects the depth of field, the distance between the nearest and the farthest objects in the scene that appear acceptably sharp. So, there is a conflict between the desire to allow more light onto the sensor which helps to reduce image noise and motion blur, and the need to get a sharper image. Less known is the fact that not only the aperture size is important, but so is its shape. The conventional circular shape is not the best in terms of the power spectrum of spatial frequencies that it passes through. The power spectrum of the circular aperture contains multiple zero-crossings that essentially filter out some content from the scene. There are less conventionally shaped apertures that produce a much better spatial response, and may even allow more light onto the sensor, however their usage requires a great deal of additional computation in image processing.

Aperture Shape

Conventional Aperture
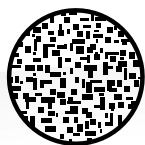
Unconventional Aperture with Improved Spatial Performance

MULTI-EXPOSURE METHOD

Is there any way to go beyond the physical limitations of the image sensor that constrain the dynamic range of the camera? Fortunately, the answer is yes: the most popular approach is a multi-exposure method: two or more snapshots are captured at different exposure times (shutter speeds), and then composed into a single wide dynamic range image. Shorter

exposures reveal scene details in the highlights while losing the details in the shadows, the longer exposures reveals details in the shadows while overexposing the highlights. The resultant WDR image is composed of pixels coming from both the short and the long shutter images: pixels from the short shutter image come primarily from the brighter parts of the scene, while pixels from the long shutter image(s) come primarily from the darker parts. Most WDR cameras use dual-shutter method, while some use triple and quadruple shutter technology to achieve a wider dynamic range.

**Short Shutter Exposure**

Long Shutter Exposure

Dual-shutter method combines pixels from short shutter and long shutter images to achieve wide dynamic range.

Selecting the individual exposure durations and combining multiple individual snapshots into a single WDR image is where the technical difficulties are. The general goal in selecting the short shutter duration is to avoid clipping in the highlights, and to maximize the signal-to-noise ratio in the brighter parts of the scene. Clipping can be automatically determined by the on-camera auto-exposure algorithm from the image histogram, as the camera adjusts the short shutter speed. Similarly, when selecting the long shutter speed the goal is to prevent dark clipping which causes noise in the shadows.

Even though this approach appears to be straightforward, consider the less obvious implications. For example, depending on the scene itself, the brighter parts of the scene captured by the short shutter exposure may also contain less bright areas that are not dark enough to be ultimately replaced with pixels from the long shutter image in the final WDR image. Similarly, the long shutter image capturing the darker areas may contain less dark scene fragments that are not bright enough to be replaced by the short shutter pixels in the composed image. As a result, these in-between areas of the scene may receive inadequate exposure – too long in the long shutter image, potentially causing overexposure of these areas, and too short in the short shutter image, potentially

causing excessive noise in the affected areas. The effect of these exposure artifacts on the final WDR image manifests itself as signal-to-noise ratio discontinuity, and appears as areas of excessive noise around the middle of the dynamic range.


Noise
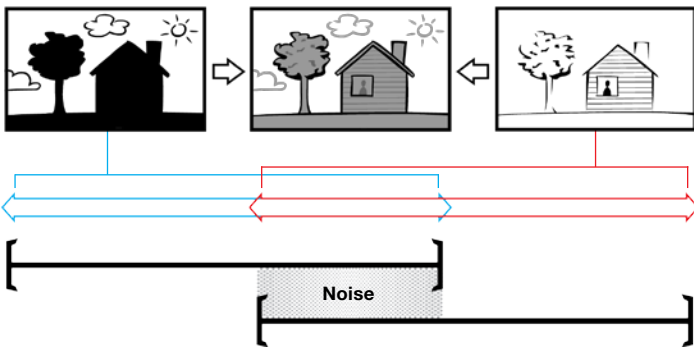
LIMITATIONS OF WDR IMAGING

It should be noted that in some of the mass produced WDR technologies, the choice of available shutter ratios is very limited due to certain technological and cost constraints. As a result, problems caused by selecting the inadequate shutter ratios become difficult to resolve, making noise and other artifacts highly dependent on the scene itself.

A further limitation of WDR imaging is related to motion artifacts. Due to different shutter speeds, parts of the scene that were moving at the time of capture will appear at different locations in the two (or more) individual snapshots that make up the composed WDR image. For example, consider a dark object that moves across a bright background scene. If the object was stationary, the dark pixels in the composed WDR image would come from the long shutter image, while the surrounding outline would come from the short shutter image. However, because the object is moving, its outline in the short shutter image does not match its outline in the long shutter image. A decision has to be made as to how to fill-in the mismatch. This becomes especially difficult when motion is accompanied by motion blur, which produces brightness values in-between the brightness of the object and the background. In the case of color WDR imaging, motion artifacts may produce bright color contours around the moving objects unless special computational measures are taken in image processing. Moreover, when excessive temporal noise

is present, it could produce motion-like artifacts resulting in colorful noise patterns. Measures to suppress motion artifacts in WDR are known as motion compensation.

Motion compensation may produce visual artifacts in a special case of capturing WDR scene containing very bright artificially illuminated areas. Bright areas of the scene are captured primarily with the short shutter. If the duration of the short shutter is shorter than the duration of the half-cycle of artificial illumination (8.33ms in 60Hz countries, or 10ms in 50Hz countries), there could be significant local brightness variations between the short and the long shutter images. These variations may be mistaken by the on-camera motion compensation as legitimate motion, causing visual artifacts. In this scenario, disabling motion compensation may be a reasonable solution.

Motion artifacts from combining different shutter images
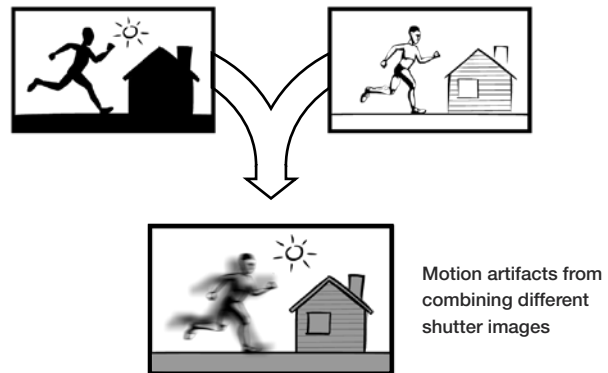
DISPLAYING WDR VIDEO ON A TV MONITOR

Are there any special considerations in displaying the wide dynamic range video on a television monitor? The answer is yes: monitors are generally incapable of displaying wide dynamic range imagery, they are limited to a dynamic range which is 200-300 times narrower than that of a typical WDR camera. Fortunately, there is a solution: the problem is solved by taking advantage of one of the many idiosyncrasies of human visual perception which places much more importance on local contrast variations, and much less on the global variations between the large areas making up the overall scene. To take advantage of this, the WDR image is put through a nonlinear image processing process known as tone mapping, which reassigns pixel brightness values to achieve the reduction of global contrast while preserving the local one. Consider a typical
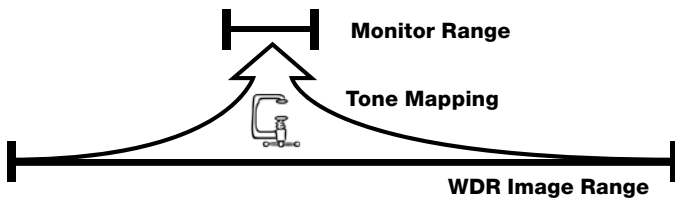
Typical monitor range is much narrower than that of the WDR image.
Tone mapping reassigns pixel brightness to make the image displayable.

WDR scene in which the highlights correspond to the bright outdoor scenery while the shadows correspond to the indoor space. If the overall brightness range of the outdoor areas is significantly reduced to bring it closer to the overall brightness range of the dark indoor areas while preserving the local brightness variations in both areas, the overall appearance of the scene will remain perfectly acceptable to the human eye, while the image becomes displayable on a monitor.

COMPARING PERFORMANCE OF WDR CAMERAS

It should be noted that tone mapping combined with the multi-shutter image capture is a highly nonlinear process which may produce significant differences in the WDR image appearance even with relatively small changes in the scene. This makes it especially challenging, and requires a great deal of technical expertise, when comparing the performance of different WDR cameras. In particular, care must be taken to make sure that the fields of view, including their aspect ratios, are identical, and the apertures are the same. Further, it is important to always look at the entire captured WDR image, and not at a cropped fragment of it post-capture, for example, showing only the bright areas. This is because the appearance of the bright area in terms of its overall contrast, noise content, and its brightness may be entirely different depending on the content of the rest of the WDR scene. If there was a large dark area in the full scene, the appearance of the bright area will be significantly modified via tone mapping and via multi-exposure image capture in terms of its final contrast, brightness, and noise, as compared to a scene without a large dark area. Unfortunately, it has become a common practice to present WDR images cropped post-capture, which makes it impossible to judge camera performance.
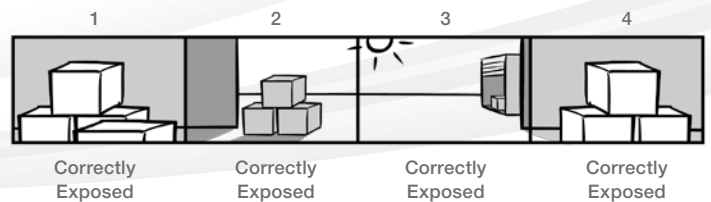
PANORAMIC WDR

The combined effect of tone mapping and multi-exposure image capture becomes especially challenging when it comes to achieving a uniform panoramic image appearance in a multi-sensor panoramic WDR camera. Adjacent parts of the scene captured by individual image sensors may appear entirely different in terms of brightness, noise, and contrast, just because there is a slight change in scene content from sensor to sensor. In order to produce a more uniform panoramic image, it is often necessary to sacrifice the dynamic range in some of the channels of the multi-sensor camera or perform an additional elaborate post-processing to equalize the appearance of the resultant multiple WDR images.

Typical WDR Scene

1	2	3	4
Incorrectly Exposed	Reference Channel	Correctly Exposed	Incorrectly Exposed
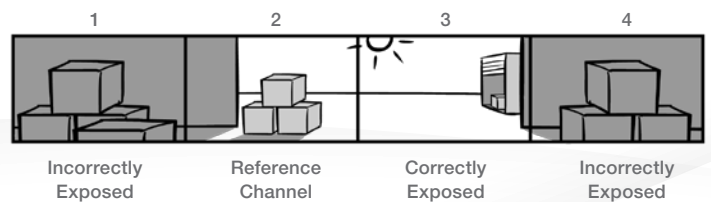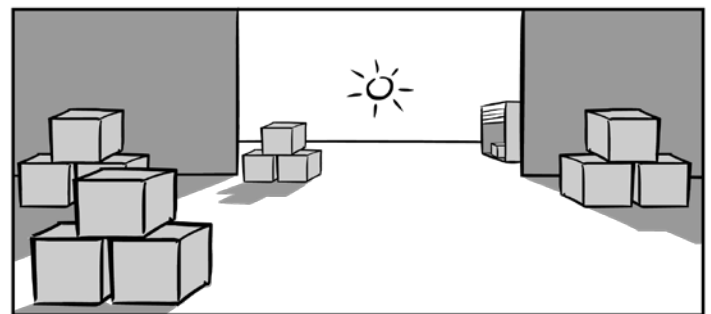
Uniformity Achieved
Reference channel properly exposed
but other channels may not be.

1	2	3	4
Correctly Exposed	Correctly Exposed	Correctly Exposed	Correctly Exposed

Non-Uniform (No Reference Channel)
Independent multi-sensor WDR sacrifices uniformity,
but achieves proper WDR capture.

Arecont Vision

Arecont Vision is the leading manufacturer of high-performance megapixel IP cameras and associated software. Arecont Vision products are made in the USA and feature low-cost massively parallel image processing architectures MegaVideo® and SurroundVideo® that represent a drastic departure from traditional analog and network camera designs. All-in-one products such as the MegaDome® 2, MicroDome™, MegaBall®, MegaView® 2, SurroundVideo® and D4F/D4S/D4SO series provide installer-friendly solutions. Compact JPEG and H.264 series cameras address cost-sensitive applications. These innovative technologies enable Arecont Vision to deliver multi-megapixel digital video at IP VGA camera price points. Arecont Vision is **Leading the Way in Megapixel Video** with cutting-edge innovation, superior performance, and the broadest selection of megapixel IP cameras.

10 Years
2003 - 2013

North America

Los Angeles, USA
Arecont Vision, LLC. Headquarters
425 E. Colorado St., 7th Floor
Glendale, CA 91205, USA

Phone: +1.818.937.0700
Toll Free: 877.CAMERA.8
Fax: +1.818.937.0464

Sales
Phone: +1.818.937.0700 x 3
Email: avsales@arecontvision.com

Support
Phone: +1.818.937.0700 x 1
Email: support@arecontvision.com

South America

Latin America
Email: latam.sales@arecontvision.com

EMEA

Frankfurt, Germany
Arecont Vision Europe GmbH
Lyoner Straße 38, 5th Floor
60528 Frankfurt, Germany

Phone: +49.69.6953.4600
Email: europe.sales@arecontvision.com

Moscow, Russia
CJSC Arecont Vision
121351, Russian Federation, Moscow, 58/1,
Molodogvardeiskaya str.

Phone: +7.495.649.0577
Fax: +7.495.221.5227
Email: info@arecontvision.ru

Middle East
Email: mena.sales@arecontvision.com

Africa
Email: africa.sales@arecontvision.com

APAC

Shanghai, China
Resident Representative Office of
Arecont Vision, LLC in Shanghai
Taolin Road No. 18, Building A
World Plaza, Room #1603
Pudong New District, Shanghai, 200135
P.R.China

Phone: +86.21.6176.2596
Alt. Phone: +86.21.6176.2597
Fax: +86.21.6085.2520
Email: apac.sales@arecontvision.com