Demystifying the HDR Ecosystem

Few terms in display technology are more widely used, less well defined, and more often misunderstood than HDR or high dynamic range. This confusion stems from history, context, and the time immemorial conflict between scientists, engineers, and marketing departments. Here, the authors will provide some clarity.

by **Karl Lang and Neil Robinson**

HDR in Image Capture

The concept of high dynamic range (HDR) image capture has existed in the photography and videography industry for a significant amount of time. For image capture, the functional use of the term "dynamic range" is identical to the classic metrological definition (see **Sidebar 1**). The first HDR photography methods were developed in the Apollo era to capture images with sunlight and shadow on the surface of the moon. The goal of HDR capture is to record the complete range of luminance values that exist in a scene, for example, documenting the contents of the deep shadows and resolving subtle detail in sunlit highlights. Camera systems and sensors continue to extend their dynamic range, with typical professional film or digital cameras able to capture 14 stops of dynamic range, a ratio of 16,000:1, while some state-of-the-art cinema cameras are capable of 16 stops, for a total acquired luminance range exceeding 64,000:1.

HDR in Image Processing

Adding confusion to the term's use, a large portion of the population associates HDR with a particular style or look of an image. The context is that of perceptual dynamic range compression or rendering (a type of image processing used in advanced photography and video to reproduce HDR content on more limited output media).

When we view a real scene, what we perceive is the result of complex processing by the human vision system (HVS). The HVS is capable of incredible performance, from detecting single photons in absolute darkness, to the evaluation of subtle color under 100,000 lux of solar illumination.

Final perception (what we see) is assembled from many individual pieces—tiny glances of the eye scanning the overall scene. For each of these local captures, the eye and HVS adapts to the light levels in the region to acquire the most information. Because of this local adaptation, human perception is quite different than a photographic image.

Advanced computer models of the HVS have been used for some time to process images captured with HDR cameras into images that can be viewed on normal displays and as photographic prints. The goal of such models is to reproduce our perception given the limitations of these display devices and prints. These images should be more realistic than traditional reproduction methods using simple tone curve corrections.

Many people are familiar with such processed images and collectively refer to them as HDR images. Some have negative connotations associated with this term based on early applied models that created less than pleasing image artifacts. Many are often unaware that their own cell phone is using state-of-the-art versions of these HVS models to create pleasing and seamless renderings of their photographs.

HDR image processing, which may include these kinds of local adaptation models, is often referred to as "tone mapping," an important component of the HDR display ecosystem.

HDR in Display Technology

"High" is a relative term; without context, it is ambiguous. Initially, in the early 2000s, the term HDR display was applied to new technologies that extended the maximum luminance range of

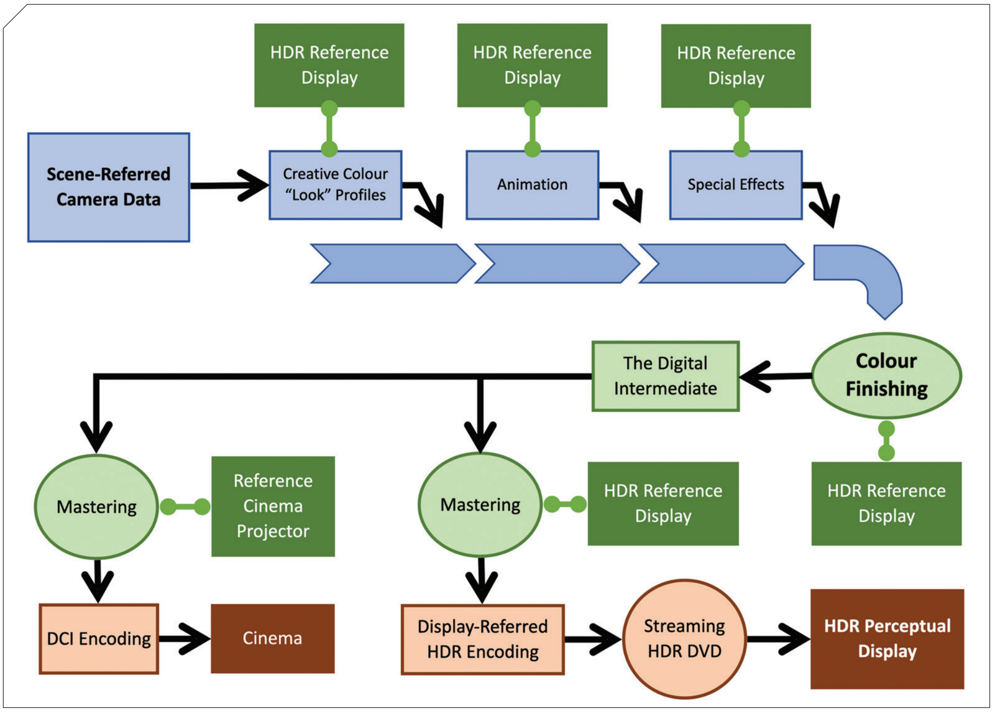

Fig. 1.

Flow diagram.

traditional displays. LCD backlights with 2D addressable arrays were an early exemplar. At that time, it was clear that an LCD display that could extend its maximum luminance to 2,000 nits was in a different category than the more typical 200 nits. However, because display technology is undergoing constant evolution, many technologies have since achieved high luminance ranges and other capabilities, such as very dark blacks. The nature of the HVS is such that being able to produce dark detail closer to absolute black also extends a display's perceptual "dynamic range" and its ability to achieve the perceptual intent.

At the time of their invention, HDR displays were a novelty; they had no available HDR content, nor did methods exist to deliver such content. Using a traditional video signal to cover a much wider luminance range did not reflect the original perceptual intent of the signal. An HDR ecosystem was needed to combine and facilitate HDR scene capture, HDR artistic intent, and HDR display technology. HDR display technology is not one thing, and, most importantly, it is not a performance threshold. Instead, it should be viewed as an ecosystem with a specific goal. That goal is based on human perception, not fixed display capabilities or specifications.

Many display technologies are now assembled under the banner of HDR. Recently, the authors had to confront the absolute question, "what is an HDR display?"' for the purposes of display metrology. As members of the International Conference on Display Metrology (ICDM) HDR working group, we were tasked with creating standard methods of measurement for HDR displays. We therefore needed to examine all of the products and technologies that stake a claim to this title and determine a way to provide unbiased evaluation.

We arrived at a method that, at first, might seem counterintuitive. Instead of defining a functional performance threshold to categorize displays, we defined an HDR display as one that accepts a specific HDR-encoded signal. The HDR signal we chose was specified for a key reason; the signal is display referred. This decision was derived from the overall purpose of the HDR ecosystem, which is to deliver a perceptual sensory experience from a content creator to a viewer. Thus, a display can be considered HDR if it fits within the HDR ecosystem we describe. We should note here that while the words "high dynamic range" apply specifically to luminance, the basis of the HDR ecosystem—the HDR signal, also includes wide color gamut encoding, as such dynamic range and color gamut are intertwined.

The ultimate goal of the HDR display is to create a perception in the mind of the viewer, to stimulate a response in the brain that correlates to the intended experience. It stands to reason that a more capable display can create a more realistic perception; however, the nature of the HVS is such that reproducing actual scene luminance levels is not likely needed to create the original or intended perception.

Our aim is to help disambiguate HDR by defining two classes of

SIDEBAR 1

DYNAMIC RANGE

THE HISTORICAL CONCEPT OF "DYNAMIC RANGE" ARISES from the science of metrology. A detector of some type creates a signal in response to a stimulus. The stimulus to be measured might be pressure, light, sound, or radiation. The detector has a capability envelope limited by its method, composition, and construction. The capability envelope is defined by the noise floor (the lowest detectable energy) and the saturation point (the highest measurable energy). This capability envelope, expressed as a ratio, is the detector's dynamic range.

As technology evolved, devices were created not only to measure signals but to reproduce them. At that moment, the concept of dynamic range expanded to involve the capability of reproduction, not only measurement. Also at that time, advertising began to get involved. If a device is designed to record a signal and reproduce it, what is its dynamic range? If the signal input to a device has a limited dynamic range, what does it mean if the output range is much larger or smaller? These issues are not well addressed in published product specifications.

SIDEBAR 2

SPATIAL, TEMPORAL, AND COLOR CONSIDERATIONS IN PERCEPTUAL HDR DISPLAY

ACHIEVING THE LEVEL OF PERFORMANCE REQUIRED for an HDR reference display is technically challenging and expensive. For this reason, mass market HDR displays are more likely to fall into the perceptual HDR display category. Manufacturers can design their displays to achieve a still large capability envelope under a more limited set of circumstances, something we might call the performance envelope. Displays can be optimized to render typical HDR content while not achieving the complete performance envelope of an HDR reference display. The methods used and the rendering achieved in the perceptual HDR display category vary widely.

Higher luminance might be achieved by globally adjusting the brightness of a backlight, light engine, or drive system in response to an analysis of the incoming signal content. Such global luminance adjustment combined with signal processing can significantly improve the rendering of very dark or very bright scenes. However, such global dynamic luminance technology is limited in its ability to accurately reproduce HDR content within the same frame, for example, a sunset scene with the foreground in silhouette. Such technologies may also have time limitations at which they can maintain high brightness or the speed with which they can change the global response.

More complex systems may divide the display into segments, adjusting brightness regionally. These segmented displays overcome some of the shortcomings found in the global dynamic system by allowing for a greater dynamic range within a single frame. However, in these systems, various visual artifacts are present that correlate to the spatial frequency of the segmentation. A greater number of segments increases the quality of the rendering and the complexity of the system. While the capability of an individual segment may be great, driving many or all segments at their maximum performance may not be possible due to limitations of the power supply architecture or the cooling system. These same limitations also can result in temporal anomalies of performance during sustained high-drive levels.

In many architectures, achieving high peak luminance over the full display screen area consumes a large amount of power. A common behavior in perceptual HDR displays is a change in the performance envelope in response to average picture level (APL). This characteristic is sometimes fundamental to the technology and sometimes is a characteristic that has been implemented by design to protect systems or reduce power. HDR video content typically has a low APL (<20%); therefore, a manufacturer may optimize their display to perform best in this low APL region.

Achieving bright flashes of light, short duration specular highlights, and even bright sunny scenes also can be accomplished with technology that more efficiently generates light with a wider spectrum. This can mean a change in the color performance of a display system in the high luminance region, where the display sacrifices color saturation for luminance.

Sustained high-peak luminance consumes power and likely generates a significant amount of heat. HDR video content typically, though not always, has bright highlights of short duration. A display maker, therefore, may optimize their display to achieve very high luminance for relatively short durations.

All of these technology considerations are taken into account by the display tone mapping processor, and decisions on tone mapping will be bound to the capabilities of the particular design.

HDR display and to provide a simplified functional block diagram (**Fig. 1**) to understand the HDR ecosystem.

Defining the HDR Ecosystem

We need to carefully define some terms to specifically understand the HDR ecosystem. The definitions outlined here are unique; these same terms may have slightly different meanings as used by other people, industries, or standards organizations. We have tried to create precise meanings that will facilitate a concise interpretation and clear categorization.

SCENE-REFERRED

Scene-referred is a concept that describes a type of content encoding, whereby the encoded signal values represent real color and luminance values in an actual recorded scene. The encoded signal values can be directly transformed to the absolute CIE XYZ (the CIE1931 XYZ tristimulus response) values originally present. These types of encodings are created from calibrated raw camera data. The perception of a real scene is very different from the perception of pixels on a display, and reproducing scene-referred values on a display is rarely a pleasing or intended result. Instead, scene-referred information is the input to the process of editing and color grading by content creators. Finished content rarely bears any resemblance to scene-referred information; it must first be creatively interpreted. As an example, a night street scene for a movie may actually be shot during the day with significant lighting equipment to create distinct shadows and provide exposure range to capture color detail. Finishing the "look" of a street at night is applied to the data, changing the light levels and the colors in complex ways.

DISPLAY-REFERRED

Display-referred is a concept that describes a type of content encoding, whereby the encoded signal values exist within the capability envelope of an ideal display. The display-referred signal represents the intent of the content creator. A display-referred signal includes all the information needed to decode the values precisely into actual pixel light. The encoded signal values can be transformed directly to absolute CIE XYZ.

HDR SIGNAL

For the purposes of describing the HDR ecosystem, we are limiting the discussion to the most widely used HDR content-mastering format. The HDR signal comprises RGB component values of ITU-R BT.2020 colorimetry with SMPTE ST 2084 PQ luminance encoding and SMPTE ST 2086 metadata. This is an example of a pure display-referred signal. It is important to understand that the signal does not describe luminance values from a real scene. The signal describes the intent of the content creator; it describes what is intended to be reproduced from the signal.

Here we will not discuss the HLG broadcast signal. While it can be part of HDR content delivery, its implementation is imprecise. HLG is not display-referred, and counter to some interpretations, HLG is not scene-referred. HLG does not have a single correct transform to display light, and for this reason, we feel it can only create confusion and obscure our understanding of the subject.

COLOR CAPABILITY ENVELOPE

The color capability envelope (or capability envelope) is the complete color volume of a display. This can be defined as a minimum and maximum luminance, in conjunction with a set of display colorimetry specifications. However, some display systems do not follow the additivity assumed by such a description. Therefore, it is better to describe the complete envelope (a 3D hull) in the CIE L*a*b* color space combined with a maximum and minimum luminance. A displays actual performance envelope can differ from its capability (see **Sidebar 2**).

IDEAL DISPLAY

Ideal display refers to the concept of a hypothetical display, which perfectly reproduces HDR content as encoded, with a direct translation to CIE XYZ display light for every pixel. The Ideal display can directly realize the full HDR display-referred signal, with each pixel of every frame correctly rendered in color, position, and time. The display-referred signal describes the content in the context of this ideal display.

The full encoding volume of the HDR signal currently exceeds the real-world capability of display technology, providing room for technology to evolve and for a class of displays to create the perceptual experience of the ideal display from displays with more limited capability envelopes.

HDR REFERENCE DISPLAY

An HDR reference display attempts to accurately reproduce the HDR display-referred signal within a specified capability envelope. A reference display may have a more limited color capability than the ideal display of the HDR signal. The reference display will attempt to accurately reproduce all of the signal values within its capability envelope. Signal values that result in colors outside the capability envelope, in chromaticity or luminance, will be clipped at the envelope boundary and will result in the reproduction of the closest possible match. Reference displays typically have large capability envelopes and precise calibration. Reference displays should have performance envelopes that closely match their capability envelopes. The aim of a reference display is to act like an ideal display within the capability envelope. However, cost is a factor, and some performance variation may exist.

Fig. 2.

An example of an early sunrise river scene interpreted using a colorimetric transform from scene light (top) versus a creative cinema grading by a colorist (bottom). Preserving the creative intent of the content creator requires a display-referred signal. The creator must have a reference display with a fixed response to define the intent they wish to deliver, as well as the ability to collaborate with other creatives on the project.

Content creators use the reference display to perform final finishing, color grading, and mastering of the content and is the foundation for the intended perceptual reproduction of the content (**Fig. 2**). The capability envelope of the reference display used for mastering is embedded in the content itself as metadata.

CONTENT CREATOR

The concept of a content creator is an important part of the HDR ecosystem. HDR content, such as movies and television programs, is not a faithful reproduction of scene colorimetry. Instead, it is a form of artistic expression. Even for a live performance, artistic color grading choices occur. The directors, editors, cinematographers, colorists, and special effects artists all work together to create a final artistic vision.

MASTERING DISPLAY AND THE MDM

Mastering display is an HDR reference display used for mastering by the content creator. An HDR reference display that matches the capability of the mastering display can reproduce the content as it was viewed by the content creator at the time of mastering.

Mastering display metadata (MDM) is a static metadata information block embedded in the HDR signal. Static metadata does not change on a frame-by-frame or scene-by-scene basis. It applies to a piece of content in its entirety. MDM describes the color capability envelope of the mastering display. This information can be used to guide a perceptual HDR display's tone mapping.

HDR DYNAMIC METADATA

Dynamic metadata provides auxiliary frame-by-frame or scene-by-scene information that may be used for perceptual display rendering. Dynamic metadata is not required or used for a reference display. Dynamic metadata is typically part of proprietary perceptual rendering methods. Typically, it provides direction from the content creator on how to render HDR content on displays with smaller capability envelopes than that used to master the content.

PERCEPTUAL HDR DISPLAY

This term is the authors' creation. It describes a large segment of HDR displays that does not behave like a reference display. A perceptual HDR display typically has a more limited capability envelope than that of the HDR signal or typical HDR reference displays. Instead, it attempts to deliver the intended perceptual experience of the display-referred HDR signal. To do so, it performs advanced algorithmic processing to realize the perceptual intent of the HDR content within the display's capability envelope.

A perceptual HDR display performs perceptual image rendering, or "tone mapping," changing encoded signal values to other values in an attempt to perform a perceptual match. There is no single correct method for perceptual rendering, and methods can differ significantly; research in this area is actively evolving. Each system prioritizes different perceptual components, such as luminance, color saturation, hue, and local contrast. More complex proprietary systems rely on additional dynamic metadata encoded into the content to optimize such rendering on a scene-by-scene basis.

It is important to understand that unlike the precise CIE XYZ rendering of a signal value on a reference display, a color that is within the capability envelope of the perceptual display may not render accurately to the signal; this is by design. The same signal value may render differently from scene to scene or frame to frame. Obviously, this can pose an obstacle for metrology and calibration; perceptual displays do not necessarily have a predictable colorimetric response.

The vast majority of consumer products will be of the perceptual HDR display type. It should also be noted that perceptual HDR displays may have multiple modes that apply different types of tone mapping. One mode might provide a simple roll-off of the luminance curve with more predictable results, while another may provide advanced local contrast and spatial processing. Some perceptual displays may also offer a mode that simulates a reference display within its performance envelope; depending on the envelope, such a mode may have limited utility.

A Simplified HDR Workflow

CONTENT CREATION AND FINISHING

HDR content begins with HDR capture. In a modern digital workflow, this may be HDR cameras recording in high bit-depth RAW/log formats or traditional motion picture film that is then scanned. HDR components also may comprise special effects, animation, and 3D rendering. The details of the creative color workflow after capture can be enormously complex. Creative professionals craft the "look" of scenes by changing color in sometimes drastic ways to achieve their artistic vision.

At the end of the creative process is finishing, a penultimate final color grading. This involves creating a "digital intermediate" (DI) that contains the final cut of the film. The DI is an enormous amount of digital information. This DI includes all the precision of the original sources combined with the processing and color transforms applied by the creatives. The finishing is performed on the best, high-capability reference displays, state-of-the-art types with wide color capability and peak luminance that may reach 4,000 cd/m^2.

The DI will be archived as the source for all future media distribution. While the finishing is performed on a reference display, the DI itself is not fully display-referred. The DI likely contains colors and information beyond the reference display's capability and can even exceed the HDR signal encoding itself.

MASTERING

The final step to create HDR content is called mastering. Mastering is the final color grading for distribution and is sometimes called the color trim pass. Mastering is performed specifically for a single type of distribution media. The mastering for digital cinema is unique, an interpretation of the DI mastered to the DCI signal encoding. For the content to be distributed in HDR, separate mastering is performed from the original DI to the HDR signal encoding. The mastering of HDR content is performed using an HDR reference display. The result of HDR mastering is a creator-approved display-referred encoding of the content. While the HDR content is mastered on a reference display and that reference display is described by the MDM, the HDR content itself may contain colors outside the envelope of the mastering display. The decision to include or exclude colors outside the MDM is not currently standardized within the industry. This may cause problems if a perceptual HDR display performs tone mapping with the assumption that color in the content is limited to the MDM. Some content may be mastered with additional proprietary dynamic metadata, such as Dolby Vision or HDR10+. This may allow the content creator some additional control over how a perceptual HDR display that supports those systems will perform tone mapping on display systems that are less capable than the mastering display.

HDR CONTENT

The final mastered HDR content is encoded in high precision in the full display-referred HDR encoding volume, 0–10,000 cd/m^2, PQ EOTF, and BT.2020 colorimetry. This, however, does not mean that the content itself uses this complete volume. Current HDR content is likely to have been mastered using a reference display with the P3 color space and a peak luminance of 1,000 cd/m^2. However, significant variability exists; some content is mastered at 2,000 or 4,000 cd/m^2, and MDM information on some content indicates mastering using the complete BT.2020 color space.

HDR DISPLAY RENDERING

The final reproduction of the display-referred HDR signal is defined and verifiable for an HDR reference display. However, the vast majority of content ultimately will be viewed on perceptual HDR displays; these displays do not have a single defined rendering. No rules or standards exist for reproduction accuracy within this class of HDR displays. These displays may use many different image processing techniques to achieve their perceptual goals. These displays may have other limitations that are not directly evident using standard metrics, such as maximum luminance and color capability (see **Sidebar 2**).

Conclusion

As the HDR industry matures, we will need to develop new methods to evaluate the perceptual performance of this display class. For the time being, marketing departments have the upper hand. The good news is we now have an established standard methodology and a clearly defined HDR test signal for metrology. With this foundation, experimental work can begin to evaluate methods for efficacious comparison.

Karl Lang is a color scientist/engineer and the president of Lumita, Inc. With 30 years' experience in the digital imaging industry, Lang has designed, developed, and brought to market numerous products and provided consultation services to a wide range of companies, including Apple, Adobe, Radius, Epson, and Sony. He has extensive experience with integrated color management systems and is an active member of various display metrology standards groups. Lang is the chair of the ICDM Color Volume and Accuracy Workgroup and is an expert in the ICDM HDR Workgroup.

Neil Robinson is a display engineer at LG Electronics. He worked on the first HDR algorithms used in LG TVs and implemented the advanced calibration process recognized by the Hollywood Professional Association and won a Technical Emmy. Robinson also works on the optimization of display pipelines for gaming in partnership with NVIDIA, resulting in the first GSync-compatible television. He is an active member of the ICC and various display standards organizations, chair of the UHD Alliance Technical Working Group, and an expert in the ICDM HDR Workgroup.