White Paper

Dolby Vision and HDR10

September 2016

By

Chris Chinnock
Insight Media
3 Morgan Ave.
Norwalk, CT 06851 USA
203-831-8464
www.insightmedia.info
Table of Contents

Introduction ................................................................................................................ 3
Executive Summary ................................................................................................... 4
SDR vs. HDR ............................................................................................................. 9
Production ................................................................................................................12
   Acquisition ........................................................................................................12
      Visual Capture ................................................................................................12
      HDR Metadata: Static and Dynamic .................................................................15
      Live Production ...............................................................................................16
   Post Production ..................................................................................................16
      Editing, VFX and Color Grading Tools ............................................................16
      Digital Intermediate .......................................................................................17
      Mastering ........................................................................................................17
Distribution ..............................................................................................................18
   Encoding Options ...............................................................................................18
   Distribution Channel Options ...........................................................................19
Playback ...................................................................................................................20
   HDR/WCG Cable Options ...................................................................................20
   Image Optimization ...........................................................................................20
Summary ..................................................................................................................23

Table of Figures

Figure 1: HDR10 is a subset of Dolby Vision ............................................................. 3
Figure 2: Reduction in Dynamic Range Through the Capture-to-Display Pipeline (Source: Dolby) ......................... 9
Figure 3: Comparison of color primaries for ITU-R BT.709 versus all color humans can see and the resulting SDR Color Volume (Source: Dolby) ................................................................. 10
Figure 4: Comparison of how a larger color volume allows brighter, more saturated colors to be displayed versus SDR (Source: Dolby) ........................................................................ 11
Figure 5: BT.1886 SDR(L) vs. PQ HDR(R)OETF (top) and EOTF (bottom) Luminance vs. Code Value ............ 13
Figure 6: Banding, Posterization or Quantization Errors ........................................................................................ 14
Figure 7: ΔE2000 Visible Errors vs. Absolute Luminance ........................................ 14
Figure 8: “Smarter Pixels” Real time scene by scene content analysis to optimize image rendering (Source: Dolby) .................................................................................................................. 16

Table of Tables

Table 1: Comparison of HDR10 and Dolby Vision ..................................................... 8
Table 2: Matrix of Possible Signal Conversions (Source: SMPTE Study on HDR) .......... 21
Introduction

There has been much discussion about HDR10, Hybrid Log Gamma (HLG) and Dolby Vision as a way to master, deliver and display a High Dynamic Range (HDR) video image. HDR has elicited such high interest and excitement because it is truly a remarkable improvement toward lifelike visual experiences. It has taken over a decade of science, technology, product development and creative learning to get to this point, with many companies contributing to its rise.

But this is a complex and rapidly changing topic area with sometimes incorrect or misleading information being published – not necessarily with malice. While much has been done already in the development of HDR images and standards, much work remains to have this become the new “normal.” In this paper, we will focus on Dolby Vision and its related format, HDR10, to try to help clarify the similarities and differences.

In reality, the HDR10 format is a subset of the Dolby Vision approach, as shown in Figure 1. Some have characterized these two formats as competitors – and they are to a degree. But it is wrong to call it a format war as there will not be a winner and a loser. Dolby Vision is a more comprehensive approach that has value in the market, while HDR10 is more like a special ‘light’ case of Dolby Vision – again offering value in the market. Both formats, and others, will coexist in the market with no winners or losers.

Figure 1: HDR10 is a subset of Dolby Vision

We have made a lot of progress so far and few would dispute the fact that Dolby was an early supporter and developer of HDR technology and has worked extensively throughout the entire ecosystem to develop the Dolby Vision solution and to enable HDR10. In my conversations with Dolby, I learned of an important internal mantra: “Lead Generously.” This means support a growing segment of the industry, contribute where you can and don’t expect to make a profit on every development you offer to the market. That is noble and of course must

Insight Media
3 Morgan Ave.
Norwalk, CT 06851 USA

www.insightmedia.info
Copyright 2016
All rights reserved
Dolby Vision and HDR10

also have a commercial element, which is the licensing of Dolby Vision playback devices, but their leadership has been instrumental.

Dolby has contributed at all levels, from working with content creatives to understand the power and possibilities of HDR, to working with software and hardware companies to develop the editing, grading, and display tools to create content. Dolby has developed encoding solutions, image processing and color conversion science and technology. They have conducted human perception studies and worked to develop standards and chip sets to power the ecosystem and deliver an end-to-end HDR solution that is an entire reimagining of the traditional standard dynamic range (SDR) pipeline.

Perhaps the most significant “lead generously” contribution was the Perceptual Quantizer (PQ) concept for digitizing the capture and display of HDR images. This is the foundation for HDR10, Dolby Vision and the Dolby Cinema experience. It has been standardized as SMPTE ST-2084 in 2014 and is an open standard.

Think about it: suppose Dolby had not done any of the activities described above to help with HDR. There is no doubt the industry would not be as far along right now.

To date, only file-based or non-real time HDR content production has been done on a non-trial basis for movies and episodic content. This includes content in the HDR10 and Dolby Vision formats. For live TV production, the Hybrid Log Gamma (HLG) technology is being considered in addition to producing content using the inverse PQ curve and distributing in an HDR10 or Dolby Vision format. However, standards and technology are much less mature here and there is much more work to be done. As a result, we will focus mainly on file-based/episodic content workflows in this paper – which can be part of the broadcast workflow as well.

The HDR market is new but it is now poised for growth. All the key ingredients are in place: content, tools, commitment, displays, pipelines, standards and business models. Since HDR is an entire ecosystem covering the capture, processing, distribution and display of HDR video, it is a multi-layered topic. In this white paper we will try to cover some of the key points of commonality and differentiation as an approach to discussing HDR10 and Dolby Vision. We will organize the white paper along the acquisition-to-display pipeline, highlighting key aspects in each area.

Executive Summary

Table 1 provides a summary of the details we will talk about in this white paper. Key commonalities and differences are summarized in the bullets below.

- HDR10 and Dolby Vision both make use of the PQ transfer function, enabling high dynamic range, wide color signals to be carried efficiently within the capabilities of typical production and distribution pipelines.
- HDR10 and Dolby Vision both make use of standard HEVC encoding (Dolby Vision can be adapted to other video codecs as required).
Dolby Vision and HDR10

- HDR10 and Dolby Vision both use static metadata (according to the SMPTE ST-2086 standard) to signal information about how the content was mastered, that can be used in suitably equipped playback devices to improve reproduction.
- Dolby Vision has “dynamic metadata,” additional real time information about the image that enables the content to be mapped intelligently to exploit the capabilities of each display.
- Dolby Vision supports up to 12 bit precision versus the 10 bit precision that HDR10 employs, meaning no visible contouring even in computer generated images and subtle gradients. Dolby Vision provides options to enable 12 bit delivery even where video coding is limited to 10 or 8 bit precision.
- Dolby Vision content is mastered at 4,000 nits of peak brightness (with a 10,000 nit mastering potential), versus a 1,000 nit master for HDR10. More contrast means less “blowout” effect.
- Dolby VS10 is a universal playback solution, implemented in TVs and other devices, that can playback Dolby Vision, HDR10 and legacy SDR content – and optimize replay of each. In comparison, HDR10-only TVs cannot deliver the full experience of Dolby Vision content.
- Dolby Vision equipped displays use a consistent color mapping engine that offers better color and tone accuracy. For HDR10-only TVs, vendors implement their own solutions, resulting in inconsistent experiences across platforms and brands.
- All the major Hollywood studios are supporting Dolby Vision mastering for theatrical. Home masters are either mastered in Dolby Vision for library content or derived from the theatrical master. The HDR10 version is then derived as needed.
- Over 60 feature film titles as well as several seasons of premium episodic drama are now available in Dolby Vision for streaming to the home with HDR10 versions available for TVs limited to HDR10 support.
### Dolby Vision and HDR10

<table>
<thead>
<tr>
<th>Content</th>
<th>HDR10</th>
<th>Dolby Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studio Support for Theatrical</td>
<td>None</td>
<td>All 7 major Studios</td>
</tr>
<tr>
<td>Theatrical Presentation</td>
<td>None</td>
<td>More than 200 Dolby Cinemas have been installed or committed to globally</td>
</tr>
<tr>
<td># Titles for Streaming</td>
<td>15</td>
<td>&gt;60</td>
</tr>
<tr>
<td>Blu-ray authoring tools</td>
<td>Available</td>
<td>Fall 2016</td>
</tr>
<tr>
<td># Blu-ray Titles</td>
<td>57</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Capture</th>
<th>HDR10</th>
<th>Dolby Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>OETF bit depth</td>
<td>10</td>
<td>10/12</td>
</tr>
<tr>
<td>EOTF bit depth</td>
<td>10</td>
<td>10/12</td>
</tr>
<tr>
<td>EOTF peak luminance (nits)</td>
<td>1,000</td>
<td>4,000</td>
</tr>
</tbody>
</table>

### Encoding

<table>
<thead>
<tr>
<th>Parameter</th>
<th>HDR10</th>
<th>Dolby Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Streams</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Codec</td>
<td>HEVC</td>
<td>HEVC/AVC</td>
</tr>
<tr>
<td>Profile</td>
<td>Main 10</td>
<td>Main 10</td>
</tr>
<tr>
<td>Bit Depth</td>
<td>10</td>
<td>10/12</td>
</tr>
<tr>
<td>Encoders Available?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Layer(s)</td>
<td>HDR Base Layer</td>
<td>HDR Base Layer</td>
</tr>
<tr>
<td>Metadata</td>
<td>Static</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Resolution</td>
<td>UHD</td>
<td>UHD/HD</td>
</tr>
<tr>
<td>Color representation model</td>
<td>YCbCr, 4:2:0, 10-bit</td>
<td>YC, 4:2:0</td>
</tr>
<tr>
<td>Independent white point?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>To TV</td>
<td>YC, 4:2:0, 10-bit</td>
<td>YC, 4:2:0, 10/12-bit</td>
</tr>
<tr>
<td>Able to adapt to different viewing ambients?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Backward Compatible?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Playback</td>
<td>Blu-ray, Streaming, Set-top Box, Download</td>
<td>Streaming, Set-top Box, Download</td>
</tr>
</tbody>
</table>

### OETF/EOTF

<table>
<thead>
<tr>
<th>Bit Depth</th>
<th>OETF</th>
<th>EOTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoder?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>EOTF peak luminance (nits)</td>
<td>1,000</td>
<td>4,000</td>
</tr>
<tr>
<td>Capture</td>
<td>HDR10</td>
<td>Dolby Vision</td>
</tr>
<tr>
<td>OETF bit depth</td>
<td>10</td>
<td>10/12</td>
</tr>
<tr>
<td>EOTF bit depth</td>
<td>10</td>
<td>10/12</td>
</tr>
<tr>
<td>EOTF peak luminance (nits)</td>
<td>1,000</td>
<td>4,000</td>
</tr>
</tbody>
</table>
## Dolby Vision and HDR10

<table>
<thead>
<tr>
<th></th>
<th>HDR10</th>
<th>Dolby Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mastering</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastering Static Metadata</td>
<td>Yes per ST-2086</td>
<td>Yes per ST-2086</td>
</tr>
<tr>
<td>Dynamic scene-by-scene metadata</td>
<td>No</td>
<td>Yes per ST-2094-1/10</td>
</tr>
<tr>
<td>Maximum Monitor Luminance (nits)</td>
<td>10,000 possible per ST-2084</td>
<td>10,000 possible per ST-2084</td>
</tr>
<tr>
<td></td>
<td>1,000 typical target</td>
<td>4,000 typical target</td>
</tr>
<tr>
<td>Color Gamut</td>
<td>DCI-P3 in a BT.2020 container</td>
<td>P3 and BT.2020 in a BT.2020 container</td>
</tr>
<tr>
<td>Metadata-based Regrading</td>
<td>Separate SDR grade required</td>
<td>Automatic SDR grade with option for trim pass</td>
</tr>
<tr>
<td><strong>Connectivity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDMI 1.4a; HDMI 2.0</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>HDMI 2.0a</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DisplayPort 1.3</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DisplayPort 1.4</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>USB 3 Type-C</td>
<td>Yes using DP 1.4 signaling</td>
<td>Yes with DP 1.4 or HDMI 1.4b signaling</td>
</tr>
<tr>
<td><strong>Playback</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reads display EDID to know capabilities</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uses metadata to do tone mapping and color volume remapping</td>
<td>Static applied to entire piece of content</td>
<td>Dynamic applied on a scene-by-scene basis</td>
</tr>
<tr>
<td>Tone mapping and color volume remapping guided by</td>
<td>Product brand's algorithm</td>
<td>ST-2094-10 reference implementation or optional Dolby optimized solution</td>
</tr>
<tr>
<td>Color representation model</td>
<td>YC_bC_r</td>
<td>IC_rC_p or YC_bC_r</td>
</tr>
<tr>
<td>Processing bit depth</td>
<td>10-bit</td>
<td>12-14 bit</td>
</tr>
<tr>
<td>HDR Playback Options</td>
<td>HDR10 only</td>
<td>HDR10 and Dolby Vision</td>
</tr>
<tr>
<td><strong>Business Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standards-based</td>
<td>Standards based</td>
<td>Standards based</td>
</tr>
<tr>
<td>No Royalties</td>
<td>Royalties on playback devices</td>
<td>Royalties on playback devices</td>
</tr>
<tr>
<td>Requires HEVC</td>
<td>Requires HEVC or AVC</td>
<td>Require HEVC or AVC</td>
</tr>
</tbody>
</table>

---

**Copyright 2016**
All rights reserved
## Dolby Vision and HDR10

### Table 1: Comparison of HDR10 and Dolby Vision

<table>
<thead>
<tr>
<th>Distribution Channel</th>
<th>UHD TV Channels</th>
<th>today ~ 30 by 2025 ~ 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ultra HD Blu-ray</td>
<td>HDR 10 base layer with static metadata in SEI part of HEVC encode Dolby Vision and other optional enhancement layers 100 Mbps data rate Players available from a few brands</td>
</tr>
<tr>
<td></td>
<td>Streaming Services</td>
<td>Amazon Prime, VUDU, Ultraflix, YouTube, FandangoNOW, Comcast and Netflix are already streaming HDR/WCG UHD resolution content 12-20 Mbps data rates Available via broadband, Blu-ray players, set-top boxes and streaming sticks Can support HDR10, single- or dual-stream Dolby Vision</td>
</tr>
<tr>
<td></td>
<td>Over-the-air</td>
<td>ATSC 3.0 next generation system supports HDR Unclear what flavor of HDR will be supported IP-based, not backward compatible Very flexible to mix OTA and broadband services 8-25 Mbps data rates Trials on-going now: regular service ~2018</td>
</tr>
<tr>
<td></td>
<td>Cable</td>
<td>A few set-top boxes support HDR10 Dolby Vision supported set-top boxes expected soon HDR may be tied to both 4K and HD content HDR production/delivery trials underway More HDR capability expected by 2017/2018</td>
</tr>
<tr>
<td></td>
<td>Satellite</td>
<td>Many satellites support 4K broadcasting Several 4K HDR services available (HDR10) Data rates to 30 Mbps expected More HDR capability expected by 2017/2018</td>
</tr>
<tr>
<td></td>
<td>IPTV</td>
<td>A few IPTV channels being offered No HDR services yet Data rates to 30 Mbps expected</td>
</tr>
<tr>
<td></td>
<td>Game Consoles</td>
<td>Xbox One S supports HDR10 but not Dolby Vision Playstation Neo expected by end of year - HDR support TBA Can playback 4K HDR Blu-rays or stream HDR content</td>
</tr>
<tr>
<td></td>
<td>Download Services</td>
<td>Download instead of stream 4K content Vidity may be only service supporting HDR (HDR10)</td>
</tr>
</tbody>
</table>
**SDR vs. HDR**

Dynamic range describes what can be seen or displayed simultaneously in terms of the peak brightness and the dimmest black levels for both white as well as colors.

The top graphic in Figure 2 shows the ~ 4 to 5 orders of magnitude of dynamic range that the human eye can see at any particular ambient light. The middle graphic represents traditional Standard Dynamic Range (SDR) production, which severely restricts contrast and colors. While modern professional cameras can capture 11-14 f-stops of luminance, this wide dynamic range is not preserved. In the production process and subsequent processing, this range is “squeezed down” to a smaller dynamic range. That is why a TV or projector doesn’t look like the real world – the colors and range of luminance are greatly reduced.

High Dynamic Range or HDR changes the way content is mastered so that it now has a greater range in luminance and color values, as shown in the bottom graphic of Figure 2. Images captured, produced and displayed in the HDR format will be brighter, darker and more colorful than the same content mastered in a SDR workflow.

The other key attribute of HDR content is Wide Color Gamut (WCG) as specified by ITU-R BT.2100, one of the key documents describing HDR production for international program
Dolby Vision and HDR10

exchange. While the BT.2100 document calls for single wavelength primaries, no displays can achieve this. As a practical matter, most HDR displays aim to achieve at least 90% of the DCI-P3 cinema color recommendation. The combination of all possible colors at all allowable intensities is known as the “color volume” and is an important concept as we compare HDR to SDR. An example of the color volume for BT.709 SDR is shown in Figure 3. Note the RGB primaries which are based on the capabilities of CRT phosphors allow reproduction of all the colors that can be reproduced by adding those 3 primaries. White will always be the brightest since it is the sum of all three. The allowable intensities for each color are shown in the color volume plot on the right side of Figure 3.

Color Volumes are defined by their color primaries e.g. BT.709 (CRT primaries), P3 (cinema primaries) or BT.2020/2100 (laser primaries) and their peak white value in candelas per meter squared (or “nits”). So an SDR color volume for example is defined as Rec 709/100 nits.

In the case of HDR, with both a higher peak white and more colors, the resulting color volume is larger than SDR. There are many examples of HDR color volumes such as Rec 2020 primaries at 1000 nits peak white (Rec 2020/1000 Nits). Another example is P3 primaries at 4000 nits (P3/4000 Nits) or P3 at 100 Nits (P3/100 Nits). Sometimes the minimum luminance is also included as in the case of the UHD Alliance certification program e.g. Rec 2020/1000/.05 Nits (LCD) or Rec 2020/540/.0005 Nits (OLED) indicating both a brighter and darker resulting color volume than SDR. The result of a larger color volume is that brighter whites as well as brighter colors can be displayed.
Figure 4: Comparison of how a larger color volume allows brighter, more saturated colors to be displayed versus SDR (Source: Dolby)

Figure 4 shows an example of the result. Note that on an SDR display, today we are limited to stay within the 100 nit peak white color volume and the Rec 709 color primaries. The upper picture shows an example of the resulting image. If the same content is mastered to HDR standards, for example Rec 2020 colors and 1000 nits peak white color volume, we are able to show brighter, more saturated colors closer to the real image. This is represented in the lower image of Figure 4, which simulates the greater contrast with more saturated colors.
The exact definition of HDR from an ecosystem perspective is still emerging but in general any color volume which is larger than SDR can be considered HDR. Note that the SMPTE ST-2084 standard specifies luminance up to 10,000 Nits but is agnostic about color primaries.

In May 2016, SMPTE issued the ST-2094 standard that included the following definition: “A High Dynamic Range System (HDR System) is specified and designed for capturing, processing, and reproducing a scene, conveying the full range of perceptible shadow and highlight detail, with sufficient precision and acceptable artifacts, including sufficient separation of diffuse white and specular highlights”.

We think this is a very good definition of the intent of HDR, but it does not specify parameters to define an HDR system compared to an SDR system. Some groups like the UHD Alliance specify some top level specs for Premium certification such as Rec 2020/1000 Nits for LCD and Rec 2020/540 nits for OLED, but they don’t disclose how these tests are done. How can one then independently validate the performance? In addition, the metrology or measurement techniques for HDR and WCG are still being developed. More work is needed.

**Production**

**Acquisition**

**Visual Capture**

The first part of the acquisition stage refers to the capture of light in the camera and any processing that is done to the image prior to exiting the camera. Camera companies often offer their own camera gamma curves that are well matched to their particular sensor (Slog 3, Clog, etc). The idea is to digitize the signal and apply the gamma curve to be very efficient in the use of the bits allocated. As HDR has been commercialized, the camera gamma has been standardized for both HDR and SDR content acquisition. This is now referred to as the Opto-Electric Transfer Function (OETF).

At the display side, the inverse of the OETF is applied to regain linearized light. At the display, this transform is called the Electro-Optic Transfer Function (EOTF). Figure 5 shows the OETF and EOTF for both SDR and HDR (HDR10 and Dolby Vision). Note that the product of the OETF and EOTF should yield a linear line. Note however, that creating linear light is not always desirable. Colorists or broadcasters can apply tone mapping that effectively alters the Optical to Optical Transfer Function (OOTF) for “artistic intent.”

The PQ curves shown in Figure 5 allocate many more code values to the darker parts of the picture and extend the luminance range significantly – from 100 nits to 10,000 nits. The shape was chosen to better model the human visual system response for contrast sensitivity vs. absolute ambient light level. PQ is utilized in both HDR10 and Dolby Vision.
As summarized in the content section of Table 1, the HDR10 specifications are really a subset of the Dolby Vision specification. That is, the Dolby Vision pipeline can be up to 12 bits instead of 10 bits and the mastering and/or end display luminance can go up to 10,000 (4,000 today) instead of 1,000 nits. SDR content uses the “legacy” specification of 8-bits per color, a 2.2 gamma and a 100 nit peak luminance for a mastering display.

So why does Dolby Vision support up to 12 bits per color when other approaches are limited to 10 bits per color? The simple answer is to avoid seeing any discrete changes in subtle color differences. This is also known as banding, posterization or quantization errors.

Dolby recognized early on that the steps or code values available in 8-bit coding were not sufficient to cover the expanded range of luminance and color values of HDR content without creating artifacts. These artifacts will be most evident in slowly changing shades of a color. You may have noticed this in representations of the sky in a video where there are clear bands of different blue shades (Figure 6). It is a particular problem in darker regions too. A quantitative way to evaluate the impact of bit depth on visible artifacts is to use a parameter called $\Delta E_{2000}$.

It is generally accepted that a $\Delta E_{2000}$ of 1 is the threshold for people to see a difference in a subtle color shift. Values above this become more and more visible.

Figure 7 shows $\Delta E_{2000}$ vs. luminance for 10-bit and 12-bit encoding. Notice that the 10-bit encoding never goes below a $\Delta E_{2000}$ of 3, whereas the 12-bit encoding is always below 2. Also notice that at low light levels the $\Delta E_{2000}$ rises so we are more sensitive to slight changes in luminance here. This is often evident when looking at the shadows in a scene where banding can be quite visible, especially with 8-bit SDR. Sensor noise or dithering can often mask this banding effect, but produces a different kind of artifact, so is not desirable either.
going to 12-bit avoids these trade-offs, which is why the Dolby Vision system adopts this level of quantization.

Figure 6: Banding, Posterization or Quantization Errors

Figure 7: ΔE2000 Visible Errors vs. Absolute Luminance

Note that for video sources, 10-bit quantization, which is the standard for HDR10, is generally viewed as acceptable due to camera noise. It is when one needs to show computer generated content or subtle gradients where 12 bits of quantization are really needed to avoid any banding issues.

The PQ range was developed to go all the way up to 10,000 nits of peak luminance. While displays can’t produce this level, scenes can, so it is important to capture to this level. On the
display side, while no display can reach this level today, they might someday – and maybe they should.

In developing the PQ curve, Dolby did some extensive studies showing content with a range of brightness level to people to gage their satisfaction. These studies showed that 84% of people would be happy with a display that had a range of 0.005 to 10,000 nits. To be clear, the 10,000 nits would apply to specular highlights to give them added realism as well as allow more saturated colors as described above. Given technology’s evolutionary path in the display space, it seems prudent to plan for displays to eventually have this capability. This adds even more credence for moving to 12 bits as this will be needed to cover this wider range without artifacts especially given that humans are much more sensitive to contrast changes at higher light levels. At a practical level, the PQ curve is designed in a way such that quantization precision isn’t lost when signals aren’t using the full 10,000nit range.

**HDR Metadata: Static and Dynamic**

HDR metadata comes in two flavors: static and dynamic. Static metadata has been standardized in SMPTE ST-2086 and provides information about the display capabilities used to master the content for use downstream. Static metadata is required for HDR10 and Dolby Vision content.

Dolby Vision content additionally includes dynamic metadata, carrying content mastering information on a scene-by-scene basis. Where content has been post-produced, it captures details of how the colorist tone mapped the scene and adjusted colors to get the “look” that was wanted. Where content is produced live, dynamic metadata can be generated automatically at the point of delivery encoding.

Dynamic metadata offers advantages for optimizing the content at the end user’s display. That’s because there is and will be a wide variety of display performance capabilities – and they will rarely match the display on which the content was mastered. As a result, every end device needs to “adjust” the delivered HDR content fit the specific capabilities of that display. This optimization means performing tone mapping to fit the luminance range of the TV and adjusting the colors to fit the TV’s color capabilities. The more data that is provided from the mastering session, the better the TV or display device can decide what changes to make.

Static metadata is used to optimize image quality for the whole piece of content but dynamic metadata can be used to optimize the image on a scene-by-scene basis, which should better reproduce what the content creators had in mind. Below in Figure 8 is an example of the real time analysis that occurs. All the pixels of the image on the right are shown in the color volume of mastering display on the left (white dots) and maximum, mean, minimum values are computed in real time on a scene by scene basis. This information is sent with a Dolby Vision bit stream to provide guidance to the downstream processing to enable more accurate color volume mapping for the particular playback display. Dolby calls this concept “Smarter Pixels.” Dolby’s approach has been standardized in SMPTE as ST-2094-10.
Live Production

While no HDR technology is in use today for commercial live production, two major OETFs are being evaluated for this: Hybrid Log Gamma (HLG) and inverse PQ. The advantage of inverse PQ is that it enables a wider luminance range and the full Rec2020 color range, whereas in typical use HLG has a narrower luminance range and does not support full Rec2020. Content captured with inverse PQ OETF can be used to create Dolby Vision, HDR10 or HLG-based content. In fact, the recent ITU-R BT.2100 includes the formulas for converting between inverse PQ and HLG OETFs and between inverse HLG and PQ EOTFs. Even content captured with HLG can be used to create Dolby Vision content, albeit in a compromised manner due to the lower fidelity of HLG compared to PQ.

But even if other camera gamma curves are used like Slog3 or Canon log, these can be transformed into PQ for post-production.

The way that PQ-based and HLG-based acquisition and production is going to be handled in the near term is to NOT carry metadata in the production pipeline. Instead, metadata generation will have to be automated, most likely at the point of final transmission encoding. The metadata is then included with the final HEVC encoding as a deliverable.

Demonstrations of live PQ and HLG-based production are being done now, but it is too early to see how this will shake out yet.

Post Production

Editing, VFX and Color Grading Tools

The post production stage consists primarily of editing plus the generation and compositing of special effects (VFX). Here, these tasks remain nearly identical to those performed with an
SDR workflow, but with the requirement that the tools need to be updated to support either an HDR10 or Dolby Vision workflow.

Post production professionals can cover both the Dolby Vision and HDR10 workflows by simply implementing a Dolby Vision solution. The creation of the HDR10 version can then be easily accomplished with a trim pass or automatically generated. However, the inverse is not necessarily true. If content is edited and composited for a 10-bit HDR10 workflow, upgrading to the 12-bit will require more than a trim pass.

Many popular editing and VFX tools have been upgraded to support HDR10 or Dolby Vision, so this is not an impediment to adoption. This can be accomplished with a version upgrade or through the addition of a plug-in.

**Digital Intermediate**

The Digital Intermediate or DI stage of production encompasses the final color grading and tone mapping process, mastering for specific deliverable formats and encoding to facilitate delivery. HDR10 and Dolby Vision both use the same BT-2020 color primaries, white point and encoding coefficients. However, Dolby Vision uses a 12-bit PQ curve whereas HDR10 is using a 10-bit curve. The relative benefits of this are explained above.

Color grading for Dolby Vision is done on a reference monitor that can reach 4,000 nits of peak luminance compared to the 1,000 nits used for HDR10 grading (and the 100 nit monitor for SDR). This means that any light captured above 1000 nits on the HDR10 grade needs to be clipped or tone mapped to fit within the 1000 nit color volume. This can result in a loss of details in bright objects in this luminance range, like texture and subtle details in bright clouds, for example.

By using a 4000 nit monitor along with the 12-bit PQ curve, the colorists can see and preserve more of these high brightness details. To be clear, there are no TVs today that can support a 4000 nit playback, so these details will have to be color volume mapped to the capabilities of the particular display. But by providing the extra bit depth and metadata to the TV, one can provide a much better reproduction of the original without relying on the judgment of the chip set in the TV to make these tone mapping and color volume remapping decisions. This is extremely appealing to the content creation community as it represents the highest probability that the creative intent will be preserved all the way to the playback device.

**Mastering**

Once the color grade has been approved, the production enters the mastering stage to create the required deliverables. For HDR mastering, static metadata can be provided per SMPTE ST-2086. This standard specifies the mastering display environment parameters including details such as chromaticity of the primaries, white point, and max and min luminance to help downstream processing.

Beyond the SMPTE ST-2086 standard, for Ultra HD Blu-ray mastering, some details about the content itself are recorded as metadata such as the maximum content light level (MaxCLL) and maximum frame-average light level (MaxFALL) for the entire piece of content.
Dolby Vision and HDR10

For HDR10 content, the MaxCLL will generally be 1000 nits, but for Dolby Vision content, MaxCLL is 4000 nits at present.

The MaxFALL number is also important as it gives an idea of the average brightness level of the HDR content. Colorists need to be aware of this when doing the grading so as not to create content that is too dark or too bright. Unfortunately, the Ultra HD Blu-ray standard does not specify a diffuse white luminance level, which some believe would help create more consistent image quality.

In reality, HDR content has been captured at a range of white diffuse levels from 100 to 180 to maybe as high as 300 nits (the SDR standard is 100 nits). This is an area of discussion within the community.

The Dolby Vision process goes beyond ST-2086 by adding dynamic metadata that includes color remapping and tone mapping data on a scene-by-scene basis. This dynamic metadata is generated from the grading session. There are actually two additional mastering steps needed to generate this dynamic metadata. For the Ultra HD Blu-ray deliverable, an HDR10 grade is generated from the Dolby Vision Master. The changes in the tone mapping done by the colorist to generate the HDR10 grade are captured on a scene-by-scene basis. On the Blu-ray disc, the HDR10 grade is the base layer and the optional enhancement layer contains the metadata to convert the HDR10 grade back to the Dolby Vision version.

The second way that dynamic metadata is generated is to start with the Dolby Vision master and generate an SDR grade. Again, the changes the colorist makes in tone mapping and color volume are captured as metadata. This is a tougher trim pass as the peak luminance changes from 4000 nits to 100 nits and the color primaries change from BT 2020 to BT 709. In this case, the base layer is the SDR grade and the enhancement layer contains the metadata to restore the SDR grade to the full Dolby Vision grade or any intermediate display color volume.

*With the HDR10 grading workflow, an SDR grade is usually created as well, but the knowledge of the colorist in doing this pass is not required to be captured in metadata. However, such knowledge can be captured by third party providers of the grading and mastering service.*

The mastering section of Table 1 summarizes the differences between HDR10 grading and Dolby Vision grading. Once content is finalized, it can be delivered for creation of the various distribution deliverables. To do this, content creators typically lightly compress the content (mezzanine encoding) in a JPEG2000 or ProRes format. HDR10 and Dolby Vision content are compatible with these mezzanine encoding schemes.

## Distribution

### Encoding Options

The encoding of HDR10 or Dolby Vision content for delivery to consumers does not require new codecs. Both typically utilize standard HEVC coding; Dolby Vision also supports additional coding options.
HDR10 is encoded using the HEVC Main-10 profile (Main 10 profile at 10 bits per color). The static metadata is encapsulated in the SEI part of the MPEG stream and is added at the distribution encoder.

Dolby Vision also typically uses standard HEVC coding for delivery, although other formats such as AVC can be supported. Various configurations are supported that enable 12-bit image quality, even where 10-bit or 8-bit video encoding is used. These options are summarized in the encoding section of Table 1.

For the single-layer OTT implementation, the Dolby Vision content is encoded in an HEVC MPEG stream including dynamic metadata at the distribution encoder. This signal is similar to an HDR10 signal, but the Dolby Vision chip sets in the TV offer some advantages in the signal processing. Dolby Vision also offers two dual-stream approaches. These are designed for more backward compatibility including the use of AVC encoders. In each case, there is a base layer and an enhancement layer. In one case, aimed at more conventional video delivery platforms, the base layer is the SDR grade and the enhancement layer allows for the reconstruction of the Dolby Vision grade. The other case is for Blu-ray disc delivery where the base layer is the HDR10 grade and the enhancement layer is used to produce the Dolby Vision version.

Distribution Channel Options

Today, HDR content is being distributed Over the Top (OTT) via a broadband connection or played back using a 4K Blu-ray player or upgraded game consoles. Soon, we expect other distribution platforms like cable, satellite, IPTV and over-the-air options to become available.

The distribution section of Table 1 summarizes the status of each of these options including the support for HDR10 and Dolby Vision.

Over the top solutions for example, are already delivering 4K HDR content in either HDR10 or Dolby Vision formats. 4K Blu-ray discs are available in an HDR10 format with Dolby Vision versions expected soon as Mediatek has an SoC playback in the player and Sony DADC, Scenarist, and Ateme are all onboard for disc authoring. The Xbox One S supports HDR10 but not Dolby Vision today. The upcoming PlayStation Neo is expected soon but specific HDR support has not been announced. These platforms allow playback of HDR UHD Blu-ray discs or streaming HDR content.

A number of 4K channels are on air today over cable, satellite and streaming platforms and we are just starting to see the first HDR content becoming available. Trials are underway for the new over the air standard, ATSC 3.0. Whether these options will support HDR10, HLG and/or Dolby Vision is still being decided.

It is still early in the roll out of 4K and HDR content and services and the landscape will evolve, but enough content and distribution methods are available now to serve this growing market.
Playback

HDR/WCG Cable Options

Once HDR content is in the home via one of the delivery methods above, it needs to be transported to the TV. The options are summarized in the connectivity section of Table 1.

There are three options: HDMI, DisplayPort and USB 3.0 Type-C. HDMI 1.4a and 2.0 can support Dolby Vision dynamic metadata via a proprietary signaling scheme at present but will be supplanted by a standardized implementation. HDMI 2.0a was developed to allow delivery of static metadata per ST-2084 and is the solution that most are using in new implementations and one that works well.

Other solutions are still developing. For example, DisplayPort 1.4 now offers support for HDR10 and the new USB 3.1 Type C connector can also be used to deliver HDR content. The Alt Mode of the USB 3.1 specification allows for other signaling methods to be used. In particular, it supports DisplayPort 1.4 signaling, but not HDMI 2.0a, so the DisplayPort Alt mode is needed for HDR10 delivery.

At the end of September, we also expect HDMI, LLC to announce that that the USB 3.1 Alt Mode will now support HDMI 1.4b signaling. That will be good enough to support Dolby Vision HDR, but not HDR10, which needs HDMI 2.0a capabilities.

Image Optimization

This is perhaps the most critical part of the ecosystem and an area where the differences between HDR10 and Dolby Vision may be most apparent. Most would agree that the goal is to try to have the most consistent image quality across platforms, brands and types of content and to preserve the creative intent to the extent possible. This is and will continue to be a big challenge and one the industry needs to work on.

The Dolby Vision approach does offer advantages here over HDR10 and alternatives, but does require an incremental licensing cost for playback devices beyond HDR10. Each brand and platform will have to decide if the performance is worth the cost and many brands including gRAP, LG, TCL, Hisense, etc. already have.

Since Dolby Vision provides its processing algorithms in the SoCs offered by many TV, set top box and other platform chip set makers, brands have the ability to integrate these SoCs into products. If one company provides these algorithms across multiple brands and platforms, one could argue you might get a more consistent image rendering across the devices. That sounds good to the content creative. A key challenge will be to enable this while also preserving the flexibility of brands to deliver a differentiated image from their competitors.

Today there is a wide range in HDR TV capabilities. OLED HDR TVs offer up to 540 nits of peak luminance and around 90% of P3 color volume. LCD HDR TVs vary from ones that are “HDR capable” meaning they will read the HDR data and try to do something with it, but they have a limited luminance range (usually up to 350 nits) and a limited color palette (perhaps
Dolby Vision and HDR10

around Rec. 709). Other HDR TVs increase the luminance and color volume reaching around 1000 nits of peak luminance and 90% or more of the P3 color gamut. Backlight technology varies from edge lit to direct type with the number of zones varying widely too.

Table 2 from a SMPTE report on HDR shows some of the combinations of HDR masters and configurations of consumer TVs. In general, any input can show up on any TV and what the TV does with this input varies greatly – depending upon what the content is and the capabilities of the TV.

Table 2 illustrates one of the key needs for dynamic metadata and the advantage of the Dolby Vision solution over HDR10. With the Dolby Vision solution, there is a much more consistent visual experience on the wide range of display devices as display-based optimization using the dynamic metadata is all within the same set of processing algorithms.

HDR10 systems are much more mix and match with third party providers doing the mastering, metadata generation and tone mapping/color volume remapping at the display device. This can result in a less consistent visual experience.

Table 2: Matrix of Possible Signal Conversions (Source: SMPTE Study on HDR)

<table>
<thead>
<tr>
<th>From / To Input Format</th>
<th>HD_{1080}Rec709</th>
<th>UHD_{1080}Rec2020</th>
<th>HDR_{1080}Rec709</th>
<th>HDR_{1080}Rec2020</th>
<th>HDR_{1080}Rec2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDR_{1080}Rec709</td>
<td>x</td>
<td>matrix only</td>
<td>SDR-&gt;HDR</td>
<td>SDR-&gt;HDR</td>
<td>SDR-&gt;HDR</td>
</tr>
<tr>
<td>UHD_{1080}Rec2020</td>
<td>chroma compresse</td>
<td>x</td>
<td>SDR-&gt;HDR</td>
<td>SDR-&gt;HDR</td>
<td>SDR-&gt;HDR</td>
</tr>
<tr>
<td>HDR_{1080}Rec709</td>
<td>tone map</td>
<td>tone map &amp; matrix</td>
<td>fits</td>
<td>matrix only</td>
<td>fits</td>
</tr>
<tr>
<td>HDR_{1080}Rec709</td>
<td>tone map</td>
<td>tone map &amp; matrix</td>
<td>tone map</td>
<td>x</td>
<td>tone map</td>
</tr>
<tr>
<td>HDR_{1080}Rec2020</td>
<td>tone map &amp; chroma compresse</td>
<td>tone map &amp; chroma compresse</td>
<td>matrix &amp; chroma compresse</td>
<td>x</td>
<td>fits</td>
</tr>
<tr>
<td>HDR_{1080}Rec2020</td>
<td>tone map &amp; chroma compresse</td>
<td>tone map &amp; chroma compresse</td>
<td>matrix &amp; chroma compresse</td>
<td>x</td>
<td>fits</td>
</tr>
</tbody>
</table>

Note: Subscripts indicate the peak luminance intended to be conveyed by the source or target format.

Table 2: Matrix of Possible Signal Conversions (Source: SMPTE Study on HDR)

Image optimization at the playback device is quite complex as Table 2 suggests. A comparison of the approaches used by HDR10 and Dolby Vision is summarized in the playback section of Table 1.

Both HDR10 and Dolby Vision implementations in TVs must perform the same tasks. This starts with understanding what the capabilities of the display are (peak luminance, black level, color primaries, number of zones, etc.) and optimizing the HDR image for that particular TV. This is done in TV’s (or other device’s) SoC to perform two key functions: tone mapping and color volume remapping. As noted earlier, HDR10 content contains static metadata so the optimization is done for the entire piece of content.
With Dolby Vision content, static and dynamic metadata is included in the content allowing optimization on a scene-by-scene basis. This is important as it captures more details of what the colorist intended and allows those subtle changes to be identified and applied to the content when played back.

The color volume remapping methodology is also different with HDR10 and Dolby Vision. Color volume remapping is needed whenever the color gamut of the content does not match the color gamut of the display. With HDR content, most is being mastered to the P3 color gamut. Many HDR TVs now offer 90% or more of this color gamut, so color volume remapping needs will be modest.

Since many HDR TVs offer or are approaching 1000 nits of luminance and are at or near the P3 color gamut, these TVs are well matched to the mastering environment. As a result, HDR10 content on these TV should come close to faithfully reproducing the image the colorist saw.

Where Dolby Vision has an advantage is when the final display is not well matched to the mastering display and where content is coming in a variety of formats.

For example, there are displays that use quantum dots with a color gamut going to 90% or higher of the BT.2020 gamut and display that have a color gamut closer to 709 and can decode the Dolby Vision signal. Such products could well exist in the future as brands differentiate their line offering good, better and best HDR displays, but use a standard SoC in them.

To serve this diversity of HDR displays you need more sophisticated color volume remapping and tone mapping algorithms. For color remapping, Dolby uses the IC\textsubscript{Cp} color representation scheme which helps preserve the hue and intensity, or the more conventional YC\textsubscript{b}C\textsubscript{r}. This IC\textsubscript{Cp} color representation scheme has been standardized now in ITU-R BT.2100. HDR 10 uses YC\textsubscript{b}C\textsubscript{r} color representation method for color volume remapping, which is known to have some hue and intensity shifts, especially for blues and reds.

For tone mapping, the dynamic metadata will definitely help on a scene-by-scene basis. Plus, Dolby Vision can do this with 12-14 bits per color compared to 10 bits for HDR10 – again, potentially providing more accurate tone mapping and color volume remapping.

And one final point that has not yet been discussed much in the industry. What happens when you start to send content in a variety of formats: HDR10, Dolby Vision, HLG and even SDR. This could happen when on a single channel or more likely, as the consumer is switching channels. Source devices like set-top boxes will have the additional complexity of understanding the capabilities of the connected display, and serving up the best possible output.

To address these challenges, Dolby offers device manufacturers Dolby VS10 “a universal decoder” that can decode SDR, HDR10 and Dolby Vision content. Having a consistent image quality experience from channel to channel and device to device is something we might take for granted, but that may not be the case as we add HDR channels and mix and match with SDR ones. HDR10-only devices may have an issue with this so it is certainly something to consider and for the industry to address.
Summary

In this paper we have explained the differences and similarities between HDR10 and Dolby Vision from camera to display. As the executive summary and Table 1 details, both offer a different value proposition for various players in the market.

We hope it is now clear that HDR10 is really a subset or special case of the full Dolby Vision approach.

Dolby Vision is an end-to-end solution that provides consistent image reproduction across a wide range of viewing devices today and is scalable as production, distribution, and display technology improves tomorrow. It provides 12 bit precision, a variety of implementation scenarios, real time scene-based content optimization which HDR10 does not.

It is simply up to each player to evaluate these approaches to decide which makes sense for their particular situation. Hopefully, this white paper will facilitate this evaluation and decision process.