Efficiently distribute live HDR/WCG contents

By Julien Le Tanou and Michael Ropert (November 2018)

The HDR/WCG evolution

Today, the media distribution industry is undergoing an important evolution. **The adoption of High Dynamic Range (HDR) and Wide Color Gamut (WCG) distribution in live ecosystems imposes a specific set of technical challenges to broadcasters and live content providers.** Our clients were asking relevant questions such as, how were they going to stream SDR content to HDR systems? Or on the contrary, how were they going to be able to stream HDR content to legacy SDR systems?

We understand that for many in the industry, the move from SDR to HDR may seem like a cumbersome task. While, working alongside our industry partners and clients, MediaKind has gained extensive experience regarding efficient HDR/WCG distribution. Playing a part in our client's journey to HDR/WCG provided us with a unique opportunity to acquire an in-depth perspective of how to resolve issues stemming from the SDR to HDR shift. It has helped inspire new developments in our products and solutions, and more specifically to the MediaKind SW encoding product line to meet the needs of running an HDR/WCG live distribution chain. We invite you to explore the context of HDR/WCG with us, and to **take a closer look at how transitioning to HDR/WCG can benefit you**.

The HDR/WCG context

End-users have high expectations and demand the same level of quality across multiple systems and devices. Yes, the transition to HDR/WCG can pose a certain number of challenges to broadcasters and providers who are expected to maintain legacy systems, and provide content for new displays, while ensuring the same level of quality. However, **there are now effective solutions available to help face those challenges and meet the requirements for HDR/WCG evolutions**. Before we can fully explain the benefits of transitioning, let's take a moment to define HDR and WCG.

Evolution to HDR

What is HDR?

High Dynamic Range (HDR) imaging aims to capture, and reproduce, the entirety of the dynamic range of visible light perceivable by the human eye, and thus, **improve the perceived image quality.**

As illustrated in *Figure 1,* the visible light range varies between 10^{-6} to 10^{8} nits. The visible light range thus offers 14 orders of luminance magnitude. The human vision system (HVS), or more simply put, the human eye, uses a certain number of mechanisms to adapt our eyesight to the range of perceived light, allowing us to perceive the entirety of the dynamic. Without these adaptations, the human eye can only instantaneously perceive 5 orders of luminance magnitude, and only if those orders range between 10^{-6} to 10^{8} nits.

Today, HDR devices can capture or render about 5 orders of luminance magnitude. In contrast to HDR, and as depicted in *Figure 1*, the Standard Dynamic Range (SDR) displays, and capture devices, are only able to process a total of 3 orders of luminance magnitude.

The main benefit of HDR systems is the ability to process a much larger scale of luminance than SDR systems (5 vs 3 orders) to improve the range of light perception.

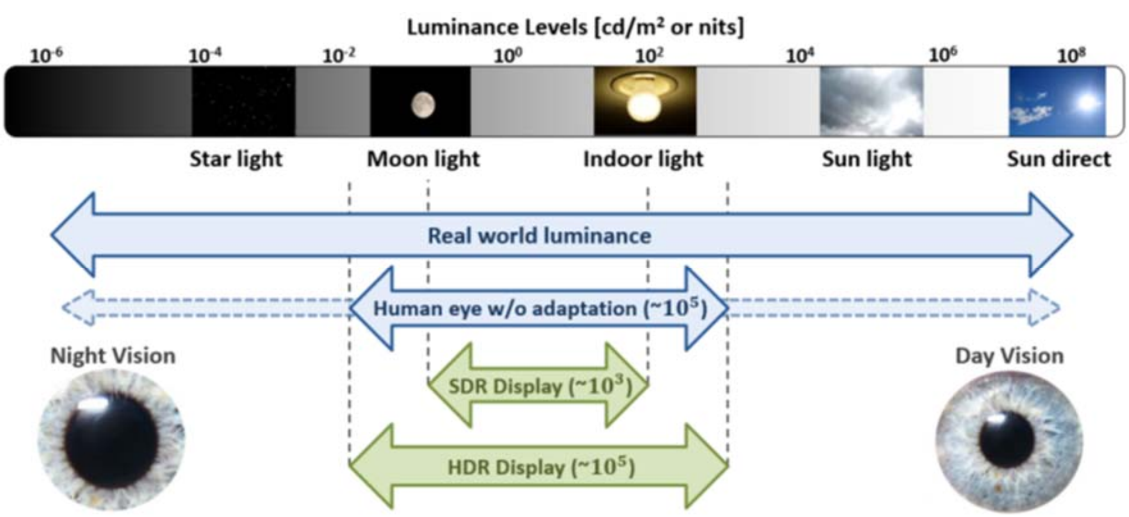


Figure 1 : Light sources and associated luminance levels showing the dynamic ranges of the human eye in relation to SDR and HDR systems

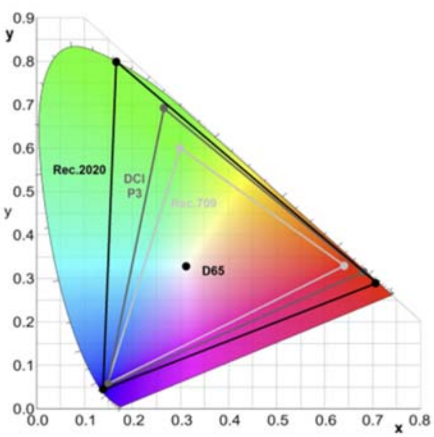
Wide Color Gamut (WCG) and HDR

The mere range of perceived light alone is not the only factor to consider when estimating perceived quality.

Color gamut has a significant role to play in HDR systems. Color gamut is specified by specific color primaries (red, green, blue) and white point coordinates.

While HDR improves the perception of light, in relation to luminance, **Wide Color Gamut (WCG) improves how color is rendered.** WCG processes a broader range, or gamut, of color than the standard color gamut for SDR.

As display capabilities continue to evolve, we can now render more colors, allowing us to make color restitution seem closer to real-world scenes. Historically, the gamut for digital TV was aligned with Rec. 709 (or BT. 709)[1]. The next step was to move to DCI-P3, which offered a wider color domain that was adopted to develop digital cinema. The most recent evolution was adopting Rec. 2020 (or BT. 2020) [2], which has become the reference for digital television and offers even more colors (*Figure 2*). Today, when talking about WCG, we most often imply the Rec. 2020 color gamut.



Color Gamuts	Rec.709	DCI P3	Rec.2020
x_W	0.3127	0.3127	0.3127
y_W	0.3290	0.3290	0.3290
x_R	0.6400	0.6800	0.7080
y_R	0.3300	0.3200	0.2920
x_G	0.3000	0.2650	0.1700
y_G	0.6000	0.6900	0.7970
x_B	0.1500	0.1500	0.1310
y_B	0.0600	0.0600	0.0460
K_R	0.2126	~	0.2627
K_B	0.0722	~	0.0593

Figure 2: Color Gamuts and corresponding coordinates (in CIE 1931 diagram)

Why transitioning from SDR to HDR ?

From glass-to-glass, **HDR systems process a broader scale of luminance, or light, in comparison to SDR systems.** Improving the dynamic range requires increasing the range between the minimum and maximum luminance values, meaning much brighter and darker light in images captured by a camera. These images are then processed and transmitted over a network or rendered by an end-display. The result is better image perception, and that is what we want for viewers.

The combination of dynamic range (luminance) and color gamut defines the color volume. This means that HDR/WCG (Rec.2020) systems, defined as *HDR-TV*, introduce a wider and greater color volume than SDR/Rec.709 systems, defined as *SDR-TV*. The difference is presented in *Figure 3* below.

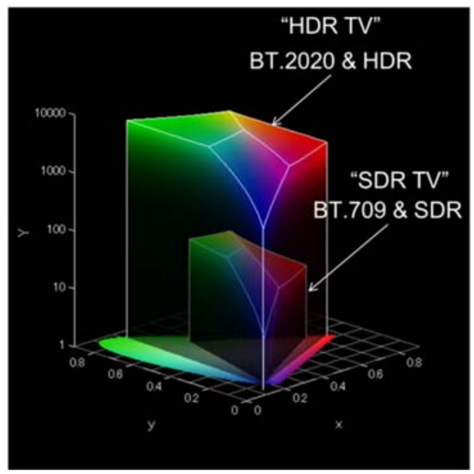


Figure 3 : Differences in color volumes for SDR/Rec.709 (SDR TV) and HDR/Rec.2020 (HDR TV) systems (in CIE1931 an xyY color space). Source: Sony

HDR-TV systems enlarge color volume, rendering the color colors, brightness, and details to their full potential. In comparison to SDR-TV, the broader and more complete color volume in HDR-TV improves the perceived quality of experience (QoE) for the end-user. *Figure 4* extrapolates (i.e. since viewed on SDR monitor) benefits in perceived quality of HDR-TV system over SDR-TV system which explains the push to transition to HDR/WCG to maximize QoE for end-users.

Figure 4 : QoE benefits of "HDR TV" over "SDR TV" (source: Sony)

Change for TV/IPTV transmission ecosystems in live HDR distribution

We know that HDR and WCG increase the perceived range of luminance and the color gamut. However, we also know that [6][7] **luminance perception is not linear**. This implies that HDR and WCG are relative to non-linear luminance and chrominance signals, and bit-depth precision, which impacts perceived quality. To benefit from improvements that evolutions for HDR and WCG offer, TV/IPTV systems are thus required to evolve as well. Although the context is technical, the explanation is simple: the evolutions to the color volume perception pose very real, and specific challenges to TV/IPTV systems. HDR and WCG improvements require change to attain the level of QoE that end-users are expecting.

Studies show that differences in luminance are not perceived with the same intensity in low or high lights. HVS is much sensible to difference in low lights. This phenomenon can be described using a Threshold vs Intensity (TVI) non-linear function [7]. It describes the just noticeable difference of luminance and chrominance that can be detected by the HVS. In the context of digital TV transmission, this TVI function is referred to as an OETF (Opto-Electronic Transfer Function), and must be applied prior to the digitization or quantization of the video signal.

For any video systems, the digitization or quantization of the signal impacts the perceived quality, and the related bit-depth precision must also be optimized accordingly to prevent from visible artifacts.

The power-law, also called the gamma-correction (or inverse gamma function), as specified in ITU-R BT.709 or in ITU-R BT.1886 [5], has been defined as the OETF function for the luminance range of SDR system. *Figure 5* illustrates the differences between quantization with the inverse gamma OETF and without. These diagrams show that linear uniform quantization does not result in uniform perceptual error. Quantization errors are more visible in the dark areas. However, with the inverse-gamma OETF, we obtain a uniform perceptual error for the SDR luminance range. It is noted that, at the display side, we apply an EOTF (Electro Optical Transfer Function = OETF $^{-1}$) to retrieve and display the relative luminance values, which means having to turn-back digital code values into visible display light.

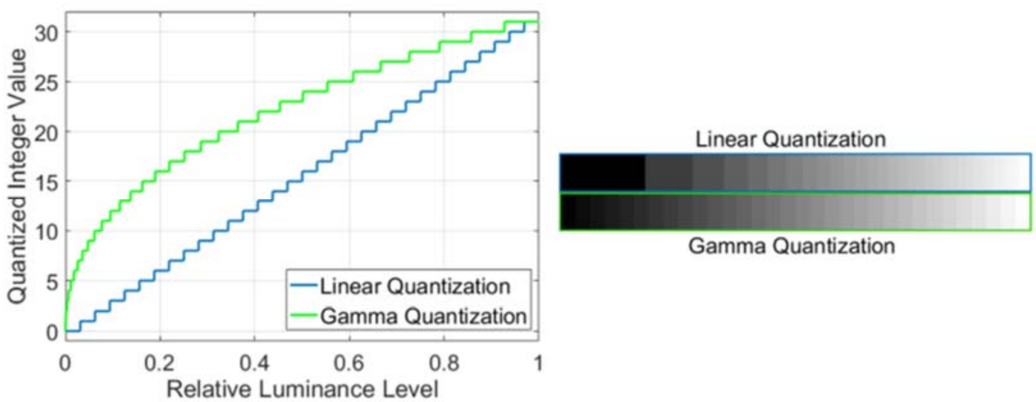


Figure 5 : Differences in luminance quantization in Linear and Gamma domain (Rec.709)

Since signal digitalization/quantization inevitably includes loss of data, the best choice is to keep quantization errors below the visible threshold of the HVS. As shown in *Figure 6,* for the SDR this is achieved for approximately 8 bits precision, as recommended in ITU-R BT.709. For lower bit depth, quantization artifacts known as “banding” or “contouring” appear.

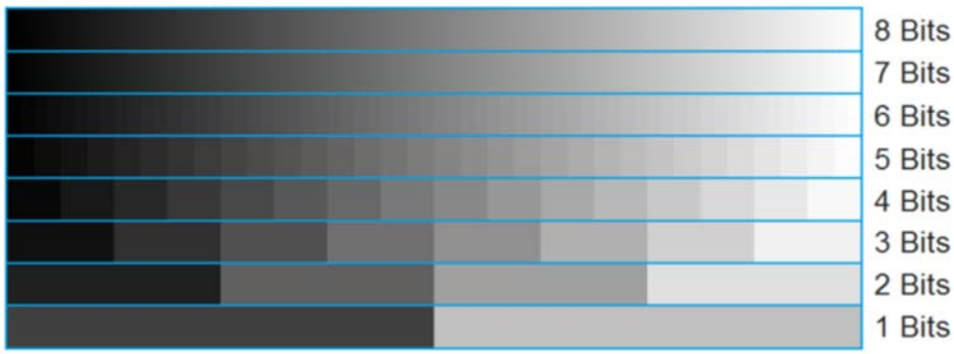


Figure 6 : SDR Luminance range quantization with varying bits (or bit-depth). Based on your display capacity in luminance range, the luminance steps have varying visibility.

The same applies for color components, where bit-depth needs to properly span the color gamut in order to remain below the visible threshold of the HVS.

HDR system requires changing both OETF/EOTF and bit-depth to optimize the perceived quality of the HDR digitalized signal to transmit, as further detailed in ITU-R BT.2390 [3] and ITU-R BT.2100 [4].

ITU-R BT.2100 specifies a bit-depth of 10 or 12 bits along with the introduction of two new OETF/EOTFs: Perceptual Quantizer and Hybrid-Log Gamma.

- ### Perceptual Quantizer (PQ)

This is the Transfer Function that defines OETF/EOTFs introduced by Dolby and initially standardized in SMPTE ST.2084 [8]. The PQ HDR system is defined as a “display-referred” system, and is not backward-compatible with SDR displays. It generates content that is optimum for viewing on a reference monitor installed in a reference viewing environment. The reference monitor would ideally be capable of accurately rendering black levels down to or below 0.005 cd/m2, and highlights up to 10 000 cd/m^2. Also, the ideal monitor would be capable of showing the entire color gamut within BT.2020 triangle. It means that additional static metadata are required for mapping to different display capabilities as specified in SMPTE ST.2086 [9]. These metadata must be conveyed using SEI. This design was specialized for films and post produced contents.

- ## Hybrid-Log Gamma (HLG)

This is the Transfer Function that defines OETF/EOTFs introduced by BBC and NHK, and was initially standardized in ARIB STD-B67 [10]. It is a "scene-referred" system, and is backward-compatible with an SDR BT.2020 10bit display (not with SDR BT.709 8-bit display). "Scene-referred" means that the system leaves the conversion of the digital code values to the display output light to the display. Theoretically, the same signal can be optimally displayed on a 1000 cd/m^2 display or, a 4000 cd/m^2 without the need of any metadata. The design was intended specifically for live broadcast.

For sake of comparison, shapes of each transfer function (and corresponding OETF/EOTF) are plotted in *Figure 7*, showing the mapping between the relative scene, or displayed light from/to 10-bit digital code values.

The data in *Figure 7* helps to define OETF and EOTF:

- **OETF:** The opto-electronic transfer function, which converts linear scene light into the video digital signal (code values), typically within a camera.
- **EOTF:** The electro-optical transfer function, which converts the video digital signal (code value) into the linear light output of the display.

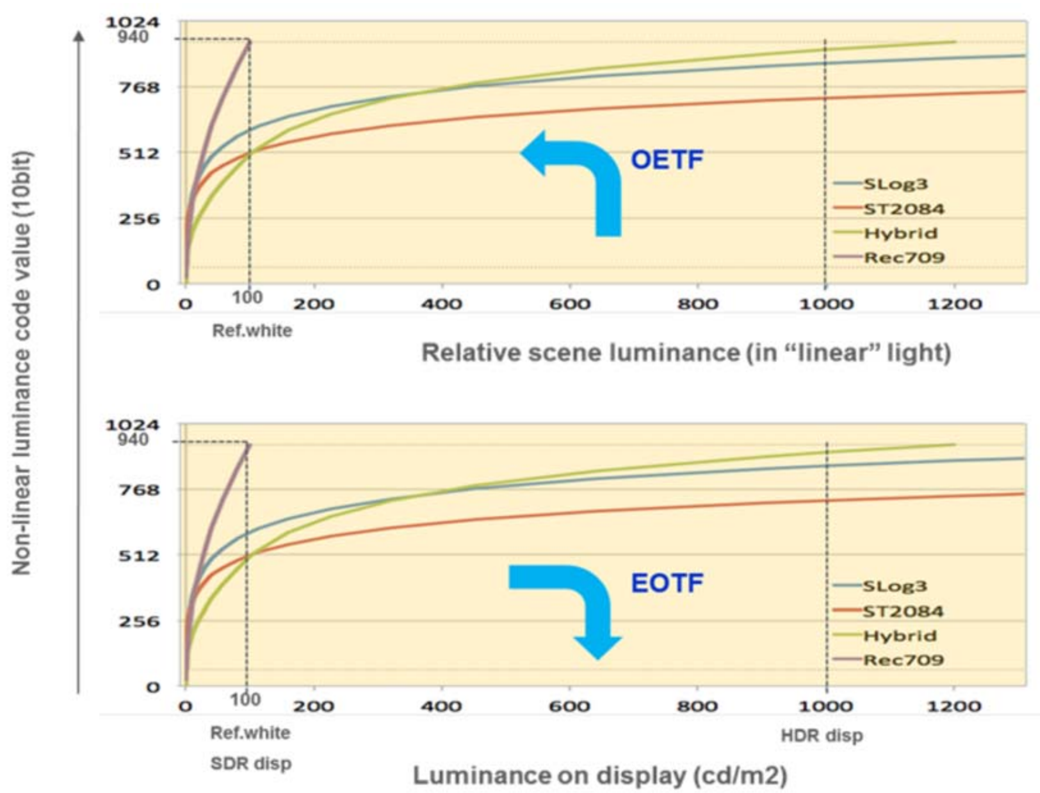


Figure 7 : Transfer functions for SDR (Rec.709/purple), HLG-HDR (Hybrid/green) and PQ-HDR (ST2084/red) systems. SLog3 (in blue) is a proprietary transfer function from Sony used in Camera or for production (not related to distribution)

HDR/WCG in relation to UHD

It is worth mentioning that HDR and WCG are a subset of the Ultra-High Definition (UHD), and only expand the luminance range and the color space in comparison to SDR video. This difference was clarified by the Ultra HD Forum [11] and further specified by DVB in ETSI 101 154 v2.3.1 [12]or ATSC [A/421]. The main point here is that **UHD-TV HDR doesn't necessary imply 4K and above resolutions**. It's important to note that **full-HD resolution and lower, down to 960x540p, are also included in UHD-TV HDR**.

	UHD	**HD**	**Legacy DVT**
Resolution	Up to 8K	Up to HD	480i/576i
Frame rate	Up to 120p fps	Up to 60p fps	50i/60i fps
Bit depth	Up to 16 bits	Up to 10 bits	8 bits
Brightness range	Up to 10000 cd/m2	Up to 100 cd/m2	100 cd/m2
Color range	Up to BT.2020 color gamut	Up to BT.2020 color gamut	Up to BT.709 color Gamut

Table 1 : HDR/WCG as a subset of UHD.

HDR and distribution formats

One of the primary challenges to face in live HDR distribution is as far as possible maintaining quality regardless of the distribution format, or the conversion of formats between SDR and HDR. To simplify the concept, we can say that video content is a mere representation of light captured by a camera. Overall the formats we use store pixel colors and luminance: red (R'), green (G') and blue (B') or luma (Y') and chroma (Cb and Cr). We are going refer to these as *containers*.

Broadly speaking there are two format types: scene referred, and display referred. A **scene referred** format describes how the light, captured by the camera (or output from the production process) is translated into the values stored in the container. A **display referred** format describes how the values in the container may be converted into light to be emitted from the display.

Today, the three main formats for HDR coding and distribution are: PQ10, HLG10 and HDR10. They are all relying on a Y'CbCr container, 4:2:0 sub-sampling and 10 bit-depth precisions.

Main formats:

PQ10: display referred signal, using PQ transfer function, Rec.2020 color gamut, Y'CbCr container, 4:2:0 sub-sampling and 10 bit-depth precisions.

HLG10: scene referred signal, using HLG transfer function, Rec.2020 color gamut, Y'CbCr container, 4:2:0 sub-sampling and 10 bit-depth precisions.

HDR10: PQ10 format + static metadata based on:

- SMPTE ST.2086 conveying Mastering Display Color Volume information. It includes information of the reference/mastering display used at production:
  - ◊ color primaries and white point coordinates of mastering display
  - ◊ minimum and maximum mastering display luminance
- MaxFALL for Maximum Frame Average Light Level information
- MaxCLL for Maximum Content Light Level information

All these static (set once per production) metadata are used for adapting (mapping) signal (referred to a display used at the production) to the end-user display. HDR10 media profile has been initially specified by CTA for Blue-ray support.

Recent formats using dynamic metadata

More recent formats introduce dynamic metadata related to color volume transforms (or mapping) in order to optimize the rendering at the end-display based on scene or frame level content information (e.g. optimize mapping for night scenes or daylight scenes). These formats also help to better preserve the creative intent of the original HDR/WCG content across variety of available displays. The related dynamic metadata describe the management (how, when and where) of the color volume processing at the end-display. In opposition, to the previous main formats, these formats rely on dynamic metadata meaning that signaled information may change down to the frame level, which can be an issue for Live operations. However, both formats remain backward-compatible with HDR10 format.

HDR10+: HDR10 as baseline + dynamic metadata defined in the SMPTE 2094-1 and SMPTE 2094-40 describing scene-based color volume mapping technology from Samsung.

Dolby Vision (*ST. 2094-based*): HDR10 as baseline + dynamic metadata defined in SMPTE 2094-1 and the SMPTE 2094-10 describing a parametric Tone Mapping for display adaptation from Dolby.

Codec and signaling

HDR signal coding is mainly supported by HEVC/H.265 codec [13] using either VUI and/or SEI signaling; we point out that all DVB and ATSC 3.0 UHD HDR formats use HEVC/H.265 as the codec.

VUI messages are mandatory; three messages are relevant to HDR signaling:

- **transfer_characteristics:** specifies the transfer function (i.e. OETF/EOTF info)
- **colour_primaries:** specifies the color primaries (R,G,B) coordinates of the color gamut
- **matrix_coeffs:** specifies matrix coefficients from RGB to Y'CbCr container conversion for a given color gamut

SEI messages are optional and may be discarded. Regarding the main formats (i.e. PQ10, HLG10, HDR10) listed above, there are three messages relevant to HDR signaling:

1. **Mastering Display Color Volume information** (i.e. ST.2086 info)
2. **Content Light Level information** (i.e. MaxFALL/MaxCLL info)
3. **Alternative transfer characteristics information**: this SEI specifies a preferred transfer characteristic to be used as an alternative to the transfer characteristic specified in VUI. It is mainly used for signaling of HLG10 with backward compatibility to SDR BT.2020 display. Typically, this HLG10 backward compatible mode would signal SDR BT.2020 information in VUI and HLG as preferred characteristic in SEI. Such the output stream would be interpretable by both SDR BT.2020 display and HDR HLG10 display.

Table 2 summarizes the signaling for the main UHD-TV HDR and SDR formats:

Format	VUI			SEI
	Colour Primaries	Transfer Characteristics	Matrix Coefficients	
UHD-TV SDR	1 (BT. 709)	1 (Rec. 709)	1 (BT. 709)	None
UHD-TV SDR WCG	9 (BT. 2020)	14 (Rec. 2020)	9 (BT. 2020)	None
UHD-TV HDR HLG10	9 (BT. 2020)	14 (Rec. 2020)	9 (BT. 2020)	preferred_transfer_characteristic = 18
UHD-TV HDR HLG10	9 (BT. 2020)	18 (HLG)	9 (BT. 2020)	None
UHD HDR PQ10	9 (BT. 2020)	16 (PQ)	9 (BT. 2020)	None
UHD HDR HDR10	9 (BT. 2020)	16 (PQ)	9 (BT. 2020)	SMPTE ST 2086 + MaxFALL + MaxCLL

Table 2 : HDR/WCG formats with bitstream signaling (static metadata).

Table 3 summarizes the signaling for the more recent UHD-TV HDR formats using dynamic metadata:

Format	VUI			SEI
	Colour Primaries	Transfer Characteristics	Matrix Coefficients	
UHD HDR HDR10+	9 (BT. 2020)	16 (PQ)	9 (BT. 2020)	SMPTE ST 2086 + MaxFALL + MaxCLL + SMPTE 2094-1 + SMPTE 2094-40
UHD HDR Dolby Vision	9 (BT. 2020)	16 (PQ)	9 (BT. 2020)	SMPTE ST 2086 **+** MaxFALL + MaxCLL + SMPTE 2094-1 + SMPTE 2094-10

Table 3 : HDR/WCG formats with bitstream signaling (dynamic metadata)**.**

Interfaces

- **HDMI**:
  - ◊ HDMI 2.0a introduced support for signaling PQ formats and the BT.2020 color space and static metadata. (i.e. PQ10 and HDR10 support)
  - ◊ HDMI 2.0b added support for the signaling of Hybrid Log-Gamma formats.CTA-861-G [18] specifies how ETSI TS 103 433 [19][20] metadata can be carried on HDMI.
- **SDI**: includes the carriage of the signaling and metadata required for the HDR formats [14].

Issues related to efficient HDR/WCG distribution

Our clients have asked how to deliver HDR/WCG contents in various transmission cases: cable, satellite, or over IP (internet).

Figure 8 presents an overview of the processing pipeline to broadcast, or to stream, an HDR/WCG channel (without lacking the generalities in an SDR channel). The first step in the pipeline is "ingest", where input/source characteristics (the transfer function, color gamut, etc.) must be extracted and forwarded along with the raw pixel data at each step of the encoding process. Either the input is uncompressed (SDI) or compressed (TS IP/HEVC).

For SDI, the specific fields are reserved for carrying HDR information [14]. For IP/HEVC, HDR related information are extracted from VUI/SEI when decoding the elementary stream. The "convert" module is responsible for applying the proper mapping (w.r.t. to dynamic range and color space information) from input format to the requested output format. After the possible luminance and color conversions, a pre-processing for input characteristics may be applied to ease and optimize the final encoding step. Finally, encoding is processed as per the dynamic range and color gamut.

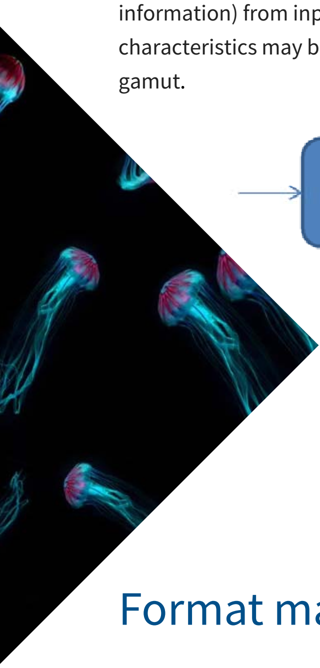


Figure 8 : Main transcoding elements in the transmission process to consider**.**

Format management (mapping and conversion)

Mixed format management

The first major issue operators and broadcasters that are anticipating the launch of an HDR/WCG channel face is how to manage the variety of input media formats. In particular how to map, or convert SDR BT.709, SDR BT.2020, HLG10, PQ10, HDR10, etc. to the broadcasted or streamed output format (PQ10, HLG10 or HDR10). The simple resolution to this specific challenge is avoiding mixing multiple HDR or SDR formats in a single channel to account for the unknown display capabilities and for simplified operations. Here are two examples of the contexts that our customers faced in relation to mixed format management:

- The first example of mixed format management is related to premium channels. The premium "HDR-TV" channel requires end-users to have a UHD HDR capable display at home. However, the reality is that content for the premium channel not only includes native premium HDR/WCG content but also several legacy SDR BT.709 content (or possibly BT.2020) that requires to be up-converted. Examples would be advertisements, movie and program trailers, etc. This requires mixing legacy content with HDR/WCG content.

- The most common example of mixed format management is wanting to provide a level of backward-compatibility with legacy SDR devices or systems. This is normally the case when addressing a second screen at home (e.g. tablet or smartphone), or for some end-users that are not yet eligible for HDR/WCG.

As we have experienced with our customers, there is a need for seamless mapping (for the end-user). Each case requires a specific approach and design. *Figure 9* displays an example of the most common use-cases.

Let's say an operator wants to launch a UHD HDR PQ10 channel, and provide backward-compatibility with legacy second screens that only support formats up to HD-SDR. Let's also say that we are operating from a single feed that contains both HDR and SDR segments.

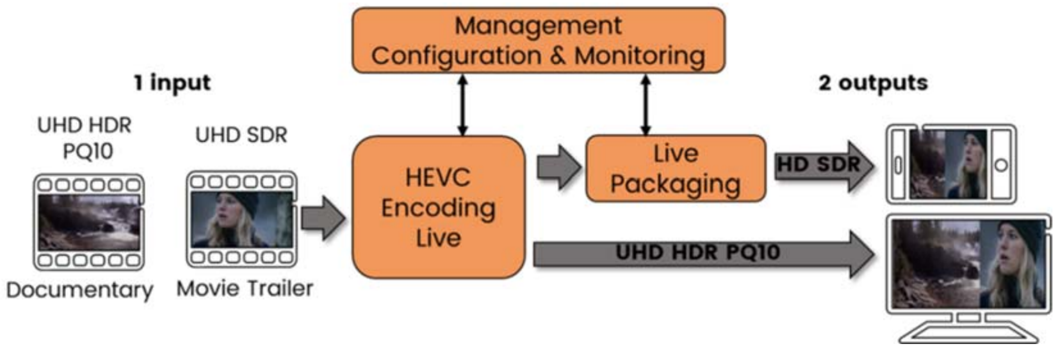


Figure 9 : Example of use cases for HDR/SDR broadcast or streaming**.**

The format conversion (or mapping) has to allow for "up-conversion", for any SDR feed segment toward a given HDR output format, as well as for "down-conversion", for any HDR feed segment toward SDR output format. It also must allow for converting one HDR format to another, for example, from PQ to HLG.

Example of conversion scenarios

Table 4 lists the different conversions that could be required depending on the input vs. output combination: :

- The "down-conversion" of the Luminance range or dynamic is named Tone Mapping (TM)
- The "up-conversion" of the Luminance range or dynamic is named Inverse Tone Mapping (ITM or TM^{-1})
- The reduction of the color gamut is named Color gamut Reduction (CR)
- The extension (or extrapolation) of the color gamut is named Color gamut Extension (CE)
- The non-linear transformation from one given transfer function to another for the same Luminance range/dynamic is named Luminance Conversion (LC)

Output signal / Input signal		HLG10	HDR10/PQ10	SDR BT709
HLG10			LC	TM+CR
HDR10/PQ10		LC		TM+CR
SDR	BT709	TM^{-1}+CE	TM^{-1}+CE	
	BT2020	TM^{-1}	TM^{-1}	CR

CR: Color gamut reduction

CE: Color gamut extension

LC: Luminance conversion

TM: Tone mapping

TM^{-1}: Inverse tone mapping

Table 4: List of possible input/output mapping and related conversions .

Luminance component conversions

The essential operations related to each type of conversions are described in the next sections.

HDR to HDR (Luminance Conversion)

The main difference between HDR10/PQ10 and HLG10 is the transfer function used. The transfer function (signaled) defines the way numerical code values (decoded pixels) must be rendered on the display, meaning how code values are processed to display light. If the signaling is changed, then numerical values must also be changed to keep the rendering on the display unchanged.

Schematically, the basic idea of these conversions is to return to the linear light domain, meaning that we apply the $OETF^{-1}$, which is the signal domain where luminance values are comparable, and to reapply the requested output transfer function (OETF) and related signaling. These transformations are described in the ITU-R BT.2100 recommendation. More details and explanations can also be found in the ITU-R BT.2390 recommendation.

SDR to HDR (Inverse Tone Mapping)

Legacy SDR contents are still being made available to end-users. As previously explained, the simplest way to convey the content onto an HDR channel is by converting the SDR input signal to HDR right after the content ingest and decoding process. The goal here is to preserve the SDR-look when rendered on HDR capable displays, which preserves the artistic intent that was initially optimized for SDR displays. Depending on the PQ or HLG characteristic of the requested HDR channel, there are two specific use-cases related to SDR to HDR inverse tone mapping:

- HLG has been designed to provide backward-compatibility with an SDR BT.2020 10 bit signal. If compared with the same color gamut and bit-depth, then SDR and HLG share the same transfer function (converting light to numerical values in 10 bits). This means that HLG can natively interpret SDR values. Consequently, only a signaling change would be required to identify SDR values as HLG ones.

- SDR to PQ conversion requires a nonlinear operation. Similarly, to the HLG to PQ conversion, SDR is brought back to the linear light domain ($OETF^{-1}$), assuming the pick luminance is 100 cd/m2. Then, an ad-hoc transfer function (OETF) and associated signaling are applied to produce a PQ signal.

In addition to the SDR and HLG, or SDR to PQ conversions, a light (luminance) expansion can be applied to provide an HDR-look to the legacy content. This means that a non-linear transformation is applied to original values. However, we suggest avoiding a light or luminance expansion, from a content provider perspective, because it modifies the artistic intent of the content creator.

HDR to SDR (Tone Mapping)

For reasons related to eligibility, broadcasters or streaming operators may offer HDR content for legacy devices or SDR-only displays. If so, then native HDR content must be downgraded to an SDR flavor. This is not an easy task. Once projected in the linear light domain, meaning after $OETF^{-1}$, a part of the signal must be discarded to remain in the SDR light range. There is no one, or unique solution that can apply to all cases, and the SDR version acceptability is subjective. On one hand, Tone Mapping designed to maximize the conservation of initial information does not produce a realistic-looking content. On the other hand, Tone Mapping based on linear remapping crushes dynamics or contrast, resulting in a loss of details in high or low light. Our MediaKind product line offers an intelligent intermediate solution using a nonlinear light remapping.

Luminance component conversions

Color Gamut remapping

These conversions are largely motivated because of needs related to remapping the color gamut, which means going from one color triangle to another, and vice versa.

Color gamut extension

Fortunately, the 3 main color domains are inclusive, as described in *Figure 2*. This makes the conversions simple. As a first step, color mapping from a smaller gamut, like BT.709 to a wider gamut like BT.2020, does not require any advanced conversion, such as color extension or extrapolation. It only requires a linear operation/mapping to change the color container, which means a change to the color referential . The absolute color sample positions in the CIE 1931 xy diagram remain the same (*Figure 10*– left). This first solution, preserves the artistic intent of the SDR/BT.709 source content.

In a more advanced conversion, color extension consists in extrapolating missing colors, which is a delicate operation in terms of color perception and intent preservation. This solution aims to give a WCG-look to the SDR/BT.709 source content; WCG-look which remains very subjective. The less impacting manner to artificially extend the initial gamut is by up-scaling while centered on the white point (*Figure 10* - right). This color extension technique would typically shift all the absolute color sample positions in the CIE xy diagram, without changing their relative positions into the color gamut container (BT.2020 or BT.709). This would result in more intensely rendering the colors in the video. It's important to state that we try to discourage color extension because it modifies the artistic intent.

Note: Both options are included in the MediaKind SW encoding product line.

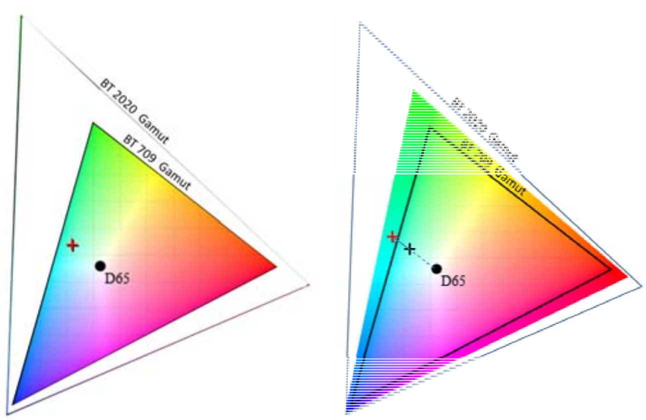


Figure 10: Color conversions - unchanged values (left) or gamut extension (right) in CIE 1931 xy diagram.

Color Gamut Reduction

Reducing the color space (in the case of BT.2020 to BT.709 gamut conversion) requires discarding some colors that are unable to be represented. Several strategies can be applied to replace one color with a different color depending on their relative distance in the reduced color space. If the replacement value is on the border of the destination gamut triangle, as shown *Figure 11*, then internal colors that were already in the destination triangle are kept unchanged. The two usual techniques are the (centripetal) projection on the border in the (D65) white point direction (left), and the other is the orthogonal projection (right) [15]. Both render similar results, which leads to color saturation and a loss of color tones, compared to the initial wider color space. An alternative technique would be trying to shrink BT.2020 into a BT.709 triangle, which generates HUE shifts artefacts and is far more of a disturbance than color saturation that projection methods generate.

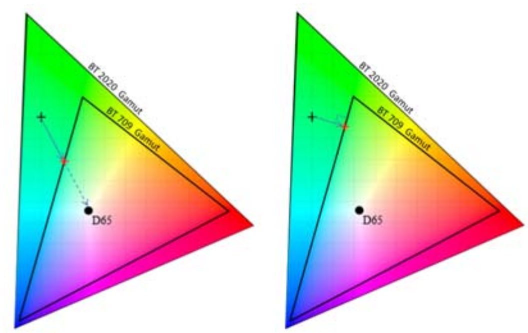


Figure 11: Color conversions (gamut reduction): projection towards a white point, or orthogonal projection.

Optimizing HDR/WCG content compression

Another major issue related to HDR/WCG transmission is estimating the overhead in bitrate when comparing HDR vs SDR cost, and estimating the compression efficiency of HDR. We have seen that the HEVC/H.265 codec has dedicated signaling for HDR/WCG coding. Even if the HEVC standard explains how to decode the elementary stream, recommendations about how to optimize the coding efficiency are not described, as they are out of scope. Information related to how to preprocess the signal to ease rate distortion optimization, or how to minimize psycho-visual artefacts do not belong to the any standard. Some insights can be found about (luma and chroma) [16] quantization adaptation or content preparation prior to encoding, but overall, HDR video compression is still a topic under study when it comes to optimizing the perceived quality.

Pre-processing and the local adaptive quantization are the two main levers to improve the video quality of HDR compressed images, and they must be advantageously utilized in any industrial implementation.

HDR video pre-processing

Most of the captured images made with devices at the edge of emerging technology are full of noise. This is not for artistic purposes. For example, very dark scenes require camera sensor amplification, which also concurrently augments the capture noise, making dark areas noisy.

Two side effects occur in encoders due to noise:

- The internal motion estimation is impacted, causing an inefficient use of temporal redundancy (motion vectors are inaccurate).
- Preference is given to Intra prediction modes, which are very efficient in predicting signals with low temporal correlation but high spatial correlation such as captured noise. As a result, Intra prediction modes are more often selected after a Rate-Distortion (R-D) optimization in dark areas. It tends to increase bitrate in dark areas to minimize the distance to the source.

In a pure R&D sense, it may seem logical to recommend keeping this behavior, as it may be desirable for objective metric scores, such as PSNR. From a psycho-visual stand-point, the opposite is also valid: Why waste bits for encoding noise?

We have taken great care in researching and testing these different scenarios in order to best adapt our product line to meet the requirements for efficient and cost-effective transmission and compression. **Today, we can confidently say that we have developed a de-noising filter, adapted to the local HDR relative linear light value.** The filter simultaneously reduces artifacts and lowers the bit consumption. However, it may reduce video quality objective scores, such as PSNR. This filter must be deactivated if running a clean source.

Adaptive local quantization based on Transfer function

Video codecs, such as HEVC, adapt the quantization parameter. This parameter controls how much data is discarded or lost from the input signal to code at the block level, or coding unit (CU) for HEVC. Each video is coded per group of images (GOP) then per block (or CU in HEVC) within an image.

As for pre-processing, when optimizing the adaptive local quantization, video quality objective measurements and subjective perceived quality come into conflict. Whereas the pure Rate (L2-) Distortion optimization (RDO) tends to favor high luminance values, meaning that the quantization step is lowered, on the contrary subjective evaluations suggest that it is more relevant to better encode low luminance values; assuming the noise has been removed. We have opted for more accuracy on darker areas, meaning that we spend more bits for dark areas compared to bright ones.

Bitrate allocation between low and high light is balanced thanks to the in house Spatio-Temporal Adapted Quantization (STAQ) algorithm. It aims to find the optimal CU quantizers that minimize the total GOP weighted distortion for a given bitrate. The weights, depending of the transfer function in use, are both defined in accordance to the local luminance and the local spatial complexity, and this while considering the temporal distortion propagation.

The MediaKind SW encoding product line for live HDR/WCG distribution

Encoding formats and conversions

All the input/output formats and related conversions described in Table 2 and Table 4 are supported in the MediaKind encoding line product. From a customer perspective, the HDR/WCG-related features available across the MediaKind encoding product versions are listed below in Table 5.

		Encoding Live v10.x	Encoding Live v11.0*	Encoding On-demand v11.0*
Ingest metadata (extraction/convey)	Uncompressed SDI		√	
	Compressed IP	√	√	√
Format Conversion	HDR to HDR	√	√	√
	SDR to HDR	√	√	√
	HDR to SDR	√	√	√
Pre-Processing	HDR-aware denoising		Roadmap	√
Compression	HDR-aware adaptive quantization	Roadmap	√	√
	HEVC 10-bit encoding	√	√	√

Table 5: HDR/WCG-related features in the MediaKind encoding product line.
*Contact MediaKind representatives for availability

MediaKind encoding solutions for cost-efficient live HDR distribution

The problems we encountered in relation to efficient HDR/WCG distribution in a Live environment allowed us to meet these challenges head-on and develop our solutions to meet new technical requirements. We are confident that any broadcaster or streaming operator launching an HDR/WCG premium channel can rely on the MediaKind encoding product line and its optimized features for live HDR/WCG distribution.

Glossary of terms

HVS: Human Vision System

OETF: Opto-Electronic Transfer Function

EOTF: Electro Optical Transfer Function

SEI: Supplemental Enhancement Information

VUI: Video User Information

DVB: Digital Video Broadcast

ATSC: Advanced Television Systems Committee

SMPTE: Society of Motion Picture and Television Engineers

ITU: International Telecommunications Union

CTA: Consumer Technology Association (formerly Consumer Electronics Association)

References

1. Recommendation ITU-R BT.709, *"Parameter values for the HDTV standards for production and international programme exchange",* 2012
2. Recommendation ITU-R BT.2020, "*Parameter values for ultra-high definition television systems for production and international programme exchange*", 2014
3. Recommendation ITU-R BT.2390, "*High dynamic range television for production and international programme exchange*", 2016
4. Recommendation ITU-R BT.2100, "*Image parameter values for high dynamic range television for use in production and international programme exchange*", 2016
5. Recommendation ITU-R BT.1886, "*Reference electro-optical transfer function for flat panel displays used in HDTV studio production*", 2011
6. F. Banterle, A. Artusi, K. Debattista, and A. Chalmers. "*Advanced High Dynamic Range Imaging, Second Edition*". CRC Press, 2017.
7. D. Gommelet, "*Methods for Improving the Backward Compatible High Dynamic Range Compression*", PhD Thesis, Ericsson-INRIA, 2018
8. SMPTE, Standard ST 2084, "*High Dynamic Range Electro Optical Transfer Function of Mastering Reference Displays*" September 2014.
9. SMPTE, Standard ST 2086, "*Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images*", 2014.
10. ARIB STD B67, "*Essential Parameter Values for the Extended Image Dynamic Range Television System for Programme Production*." July 3, 2015, Available online at http://www.arib.or.jp/english/html/overview/doc/2- STD-B67v1_0.pdf
11. "*End-to-end Guidelines for UHD Phase A implementation*" Available online at https://ultrahdforum.org/resources/phasea-guidelines-description/
12. ETSI TS 101 154 v2.3.1 Digital Video Broadcasting (DVB), "*Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*", Feb. 2017.
13. Recommendation ITU-T H.265 and ISO/IEC 23008-2, "*High efficiency video coding*" 2013.
14. SMPTE, Standard ST 425-1, "Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface", 2012.
15. T-F Bronner et al. "*Evaluation of Color Mapping Algorithms in Different Color Spaces*", Proceedings Volume 9971, Applications of Digital Image Processing, SPIE, 2016.
16. Recommendation ITU-T Rec H Supplement 15, "*Conversion and coding practices for HDR/ WCG Y'CbCr 4:2:0 video with PQ transfer characteristics*", 2017.
17. SMPTE Standard, ST 2094-1, "*Dynamic Metadata for Color Volume Transform— Core Components*", jun. 2016.
18. CTA Standard CTA-861-G: "A DTV Profile for Uncompressed High Speed *Digital Interfaces"*, Nov. 2016.
19. ETSI TS 103 433-1 v1.2.1 Digital Video Broadcasting (DVB), "*High-Performance Single Layer High Dynamic Range (HDR) System for use in Consumer Electronics devices; Part 1: Directly Standard Dynamic Range (SDR)*", Aug. 2017.
20. ETSI TS 103 433-2 v1.1.1 Digital Video Broadcasting (DVB), "*High-Performance Single Layer High Dynamic Range (HDR) System for use in Consumer Electronics devices; Part 2: Enhancements for Perceptual Quantization (PQ) transfer function based High Dynamic Range (HDR) Systems (SL-HDR2)*", Jan. 2018.