SCTE · ISBE CABLE-TEC EXPO 2017
DENVER, CO
OCTOBER 17-20

2017 Fall Technical Forum
SCTE·ISBE • NCTA • CABLELABS

Single-Layer HDR Video Coding with SDR Backward Compatibility

A Technical Paper prepared for SCTE/ISBE by

David Touze
Research Engineer
Technicolor
975 avenue des Champs Blancs - ZAC des Champs Blancs - 35576 Cesson Sevigné, France
+33 2 99 27 31 08
david.touze@technicolor.com

Leon van de Kerkhof
Chief Architect
Philips
leon.van.de.kerkhof@philips.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Introduction

The arrival of the High Efficiency Video Coding (HEVC) standard enables the deployment of new video services with enhanced viewing experience, such as Ultra HD broadcast services. In addition to an increased spatial resolution, Ultra HD will bring a wider color gamut (WCG) and a higher dynamic range (HDR) than the standard dynamic range (SDR) HD-TV currently deployed. Increasing of dynamic range, i.e. the luminance ratio of bright over dark pixels, has been shown to dramatically improve the user experience. Increasing gamut and dynamic range are two faces of the same coin as they basically augment the color volume to which pixels belong. Furthermore, luminance and colors are intrinsically linked in legacy workflows that are non-constant luminance: the signal non-linearity is not applied directly to the luminance, but instead the non-linear luminance is a combination of non-linear quantities (typically RGB).

Different solutions for representing and coding HDR/WCG video have been proposed [1][2] [3] [4]. As stated in [5][6][7][8]. SDR backward compatibility with decoding and rendering devices is an important feature in video distribution systems, such as broadcasting or multicasting systems. The coming American broadcast standard ATSC 3.0 is expected to emit both SDR BT.709/2020 and HDR BT.2020 streams. The European DVB standard has already introduced SDR UHDTV in the BT.2020 color space and will extend it to HDR BT.2020 soon. Peak brightness is expected to migrate from legacy 100 nits to about 1000 nits, but compression solutions should be flexible enough to handle future higher brightness as well as non-broadcast applications that may take advantage of more nits.

Dual-layer coding, for instance using the scalable extension of HEVC (a.k.a. SHVC) is one solution to support SDR backward compatibility. However, due to its multi-layer design, this solution is not adapted to all distribution workflows. An alternative is to transmit HDR content and to apply at the receiving device an HDR-to-SDR adaptation process (tone mapping). One issue in this scenario is that the tone mapped content may be out of control of the content provider or creator. Another issue is that a new HDR-capable receiving device is needed to apply this tone mapping for existing SDR displays. Alternatively, the Hybrid Log Gamma (HLG) transfer function [2] has been designed as a straightforward solution to address the SDR backward compatibility, that is, an HDR video graded on a display using the HLG transfer function can be in principle directly displayed on an SDR display (using the BT.1886 transfer function [2]) without any adaptation. However, this solution may result in color shifting when the HLG-graded video is displayed on an SDR rendering device, especially when dealing with content with high dynamic range and peak luminance [10][11][12]. Also, there is no way to optimize the brightness and contrast of the SDR image.

The proposed Single Layer SDR backward compatible HDR video distribution solution detailed in this paper, named SL-HDR1, and standardized in ETSI TS 103 433 specification [13]**Error! Reference source not found.**, aims at addressing these issues. SL-HDR1 leverages SDR distribution networks and services already in place. It enables both high quality HDR rendering on HDR-enabled CE devices, while also offering high quality SDR rendering on SDR CE devices.

The main features of the HDR distribution system are as follows:

- Single layer with metadata: SL-HDR1 is based on a single layer coding process, with side metadata that can be used at post-processing stage. The metadata payload corresponds to a few bytes per picture, GOP or scene.

- Codec agnostic: SL-HDR1 does not impact the core codec technology and is codec independent. SL-HDR1 is based on an encoding pre-processing applied to the HDR input, and on a corresponding decoding post-processing (functional inverse of the pre-processing) applied to the reconstructed video from decoding. Use of a 10-bit codec is recommended, since an8-bit codec could introduce artefacts such as banding effects, due to having too few codewords available for the precision required for HDR content.
- Enable SDR backward compatibility: a decoded bitstream can be displayed as is on an SDR display. The color fidelity is preserved compared to the HDR version. An additional post-processing is applied to convert the decoded SDR version to HDR, thanks to the metadata, with preservation of the HDR artistic intent.
- Enable preserved quality of HDR content: there is no penalty due to the SDR backward compatibility feature; coding performance compared to HLG are improved, in particular in terms of color impairments.
- Enable adaptation of the HDR content to the HDR display capabilities: if the HDR content peak brightness is higher than the HDR display peak brightness, the post-processing adapts the HDR content to display peak brightness, preserving all details and HDR artistic intent.
- Limited additional complexity: the pre- and post-processing steps are of limited added complexity; in particular the involved operations are only sample-based, without inter-sample dependency.
- Independent from the input OETF: the pre- and post-processing operate in linear-light domain, and are therefore independent from the input OETF.

The document is organized as follows. The solution overview is presented in section 1. Section 2 describes the HDR-to-SDR decomposition and section 3 the HDR reconstruction process. Section 4 relates to the metadata signaling. Section 5 details the display adaptation feature. Section 6 presents tests results, assessing the SDR quality and the HDR compression performance of SL-HDR1 comparatively to distribution solutions based on PQ and HLG transfer functions. Conclusion section provides closing remarks.

Content

1. SL-HDR1 System Overview

Figure 1 shows an end-to-end workflow supporting content production and delivery to HDR and SDR rendering devices. The core of the HDR distribution solution SL-HDR1 corresponds to yellow and green boxes. SL-HDR1 involves a single-layer SDR/HDR encoding-decoding, with side dynamic metadata. At the distribution stage, an incoming HDR signal is decomposed in an SDR signal and content-dependent dynamic metadata. The SDR signal is encoded with any distribution codec (e.g. HEVC Main 10) and carried throughout the existing SDR distribution network with accompanying metadata conveyed on a specific channel or embedded in the SDR bitstream. The dynamic metadata are typically carried in an SEI message when used in conjunction with an HEVC codec. The post-processing stage is functionally the inverse of the pre-processing and performs the HDR reconstruction. It occurs just after SDR bitstream decoding. The post-processing takes as input an SDR video frame and associated dynamic metadata in order to reconstruct an HDR picture. Single-layer encoding/decoding requires only one encoder instance at HDR encoding side, and one decoder instance at player/display side. It supports the real-time workflow requirements of broadcast applications. The dynamic metadata are produced by the HDR decomposition process and remain internal to the distribution process. They do not need to be conveyed to the rendering device. Additional metadata, originated from the production/post-production, can optionally be distributed and conveyed to the rendering device.

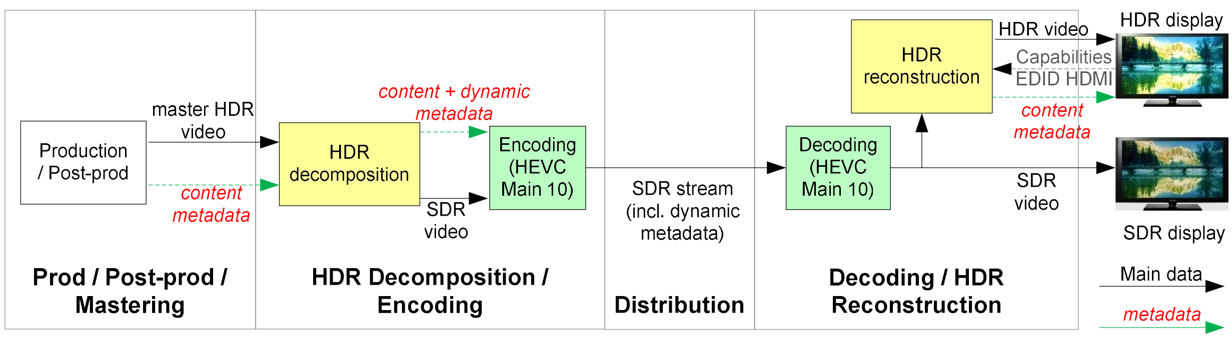

Figure 1 - Example of HDR end-to-end system.

The block diagram in Figure 2 depicts in more details the HDR decomposition and reconstruction processes. The center block included in red dashed box corresponds to the distribution encoding and decoding stages. The left and right grey-colored boxes respectively enable format adaptation to the input video signal of the HDR system and to the targeted system (e.g. a STB, a connected TV). The yellow boxes show the HDR specific processing. The first step of the HDR decomposition process linearizes the input HDR content, allowing the system to ingest every HDR production format such as PQ (display referred), HLG (scene referred) or any other production format. For the HLG case, as it is a relative format, the peak brightness of the HLG content needs to be provided to the system. The linearized content is then independent from the input format and allows the system to always work in the same consistent linear-light domain. The core component of the HDR decomposition stage is the HDR-to-SDR conversion that generates an SDR video from the linear-light HDR signal. Optionally, gamut mapping may be used when the input HDR and output SDR signals are represented in different color spaces. This optional gamut mapping may be introduced either before or after the HDR-to-SDR conversion. The decoder side implements the inverse processes, in particular the SDR-to-HDR reconstruction step that inputs the SDR video provided by the decoder and that transforms it back to an HDR video at a peak luminance adapted to the HDR display capabilities.

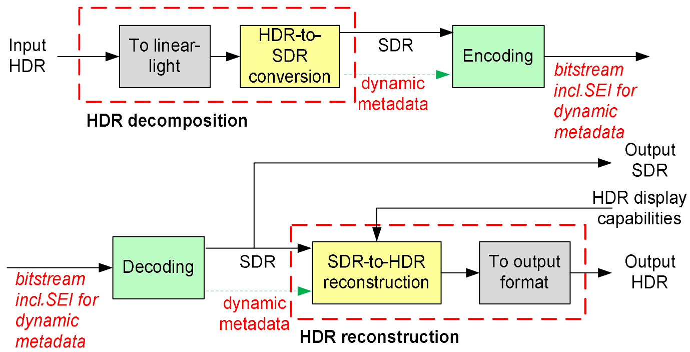

Figure 2 - HDR system architecture overview.

2. HDR-to-SDR Decomposition Process

The HDR-to-SDR decomposition process aims at converting the input linear-light 4:4:4 RGB HDR signal to an SDR Y'CbCr 4:2:0 compatible version. The process uses side information such as the color primaries and gamut of the container of the HDR and SDR pictures. The process operates without color gamut change: the HDR and SDR pictures are defined in the same color gamut. If needed, a gamut mapping processing may be applied either on the HDR pictures or on the SDR pictures. In the former case the HDR picture is converted from its native color gamut to the target color gamut before the HDR-to-SDR decomposition process. And in the latter case the SDR picture generated by the HDR-to-SDR decomposition process is converted from its native color space to the target SDR color gamut.

The HDR-to-SDR decomposition process is depicted in Figure 1 and Figure 3. It is primarily based on the HDR content analysis (picture per picture) in order to derive a set of mapping parameters that will be further used to convert the HDR signal into SDR (step 1). Once the mapping parameters have been derived, a luminance mapping function, noted *TM*, is obtained. In step 2, the luminance *L*, derived from the HDR linear-light RGB signal, is mapped to an SDR luma signal using the luminance mapping function *TM* (step 2). The chroma components are then derived (step 3). A final color correction is applied in order to match the SDR colors to the input HDR signal colors (step 4). Steps 2 to 4 are detailed in the following sub-sections.

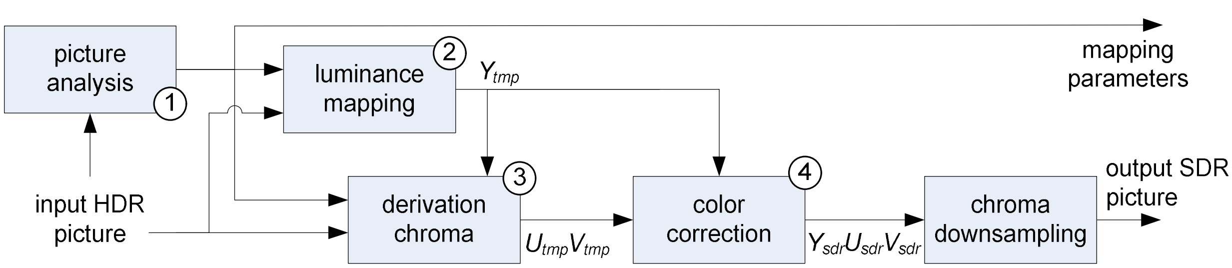

Figure 3 - Synoptic of HDR-to-SDR Decomposition Process

2.1. Luminance mapping

The luminance mapping (step 2) aims at converting the input linear-light luminance signal, derived from the HDR linear-light RGB signal, into an SDR luma signal using a luminance mapping function *TM*. This is done according to the following equations:

$$L = A_1 \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$Y_{tmp} = (LUT_{TM}(L))^{\frac{1}{2,4}} \quad (2)$$

where $A = [A_1 A_2 A_3]^T$ is the conventional 3x3 R'G'B'-to-Y'CbCr conversion matrix (e.g. BT.2020 or BT.709 depending on the color space), A_1, A_2, A_3 being 1x3 matrices.

The mapping function or look-up-table *TM* is built as follows. The mapping is based on a perceptual transfer function, and uses a limited set of control parameters, that have to be further conveyed to the post-processing in order to be able to invert the luminance mapping process. The input linear-light luminance signal *L* is first converted to the perceptually-uniform domain based on the mastering display peak luminance, using a perceptual transfer function illustrated in left picture of Figure 4. This process is controlled by the mastering display peak luminance parameter. To better control the black and white levels, a signal stretching between content-dependent black and white levels (parameters *blackLevelOffset* and *whiteLevelOffset*) is applied. Then the signal is tone mapped using a piece-wise curve constructed out of three parts, as illustrated in Figure 5. The lower and upper sections are linear, the steepness being determined by the *shadowGain* and *highlightGain* parameters. The mid-section is a parabola providing a smooth bridge between the two linear sections. The width of the cross-over is determined by the *midToneWidthAdjFactor* parameter. The curve can be further fine-tuned using a piece-wise linear corrective function. Then the signal is converted back to the linear light domain based on the targeted SDR display maximum luminance of 100 cd/m2, as illustrated in the right picture of Figure 4. The resulting signal is the SDR luma.

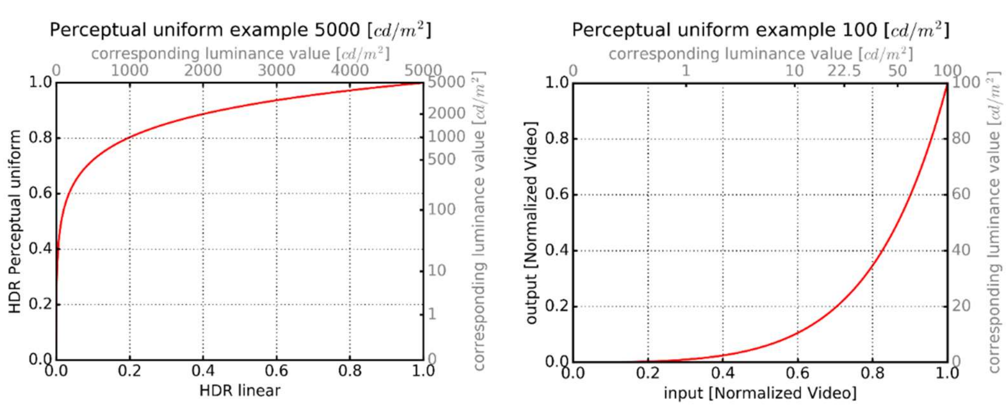

Figure 4 - Example conversion curves for converting from linear light to perceptual domain (left, with peak luminance 5000 cd/m2) and back to SDR linear light (right).

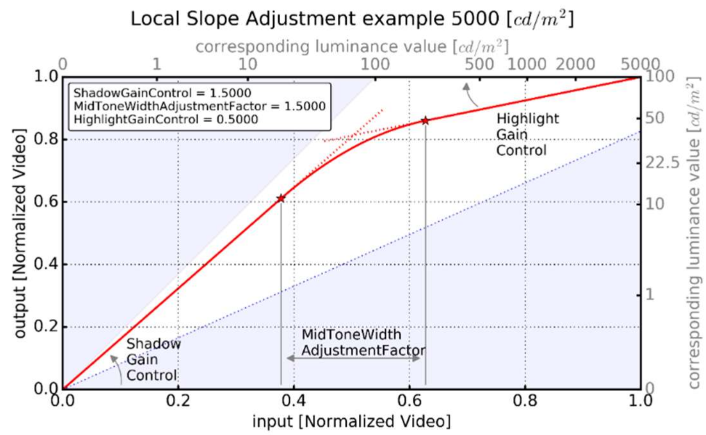

Figure 5 - Tone mapping curve shape example.

2.2. Chroma components derivation

The chroma components are derived as follows (step 3). First a square root is applied to input HDR linear-light R, G, B and L values to reproduce a transfer function close to the BT.709/BT.2020 OETF (the usage of a square root guarantees the reversibility of the process). Then the resulting squared-root R, G, B values are scaled by the squared-root L value, which results in a gamma-tized SDR version of the input R, G, B signals. The resulting R, G, B signal is converted to chroma components U_{tmp}, V_{tmp}:

$$\begin{bmatrix} U_{tmp} \\ V_{tmp} \end{bmatrix} = \frac{1}{\sqrt{L}} \times \begin{bmatrix} A_2 \\ A_3 \end{bmatrix} \times \begin{bmatrix} \sqrt{R} \\ \sqrt{G} \\ \sqrt{B} \end{bmatrix} \qquad (3)$$

2.3. Color correction

A final color correction is applied in order to match the SDR colors to the input HDR signal colors (step 4). First the chroma components are adjusted by a scaling factor $1/\beta(Y_{tmp})$, where $\beta(Y_{tmp})$ is a function that enables to control the color saturation and hue and that is constructed by matching primaries and white points between the SDR and the HDR gamut.

$$\begin{bmatrix} U_{sdr} \\ V_{sdr} \end{bmatrix} = \frac{1}{\beta(Y_{tmp})} \times \begin{bmatrix} U_{tmp} \\ V_{tmp} \end{bmatrix} \qquad (4)$$

Then the luma component is adjusted to further control the perceived saturation, as follows:

$$Y_{sdr} = Y_{tmp} - Max(0, a \times U_{sdr} + b \times V_{sdr}) \qquad (5)$$

where a and b are two control parameters. This luma adjustment step helps in recovering the color perception difference that occurs when a specific color is rendered at different luminance level.

As demonstrated in [13], this color correction step is fundamental to control the SDR colors and to guarantee their matching to the HDR colors. This is in general not possible when using a fixed transfer function.

3. HDR Reconstruction

The HDR reconstruction process is depicted in Figure 6. This section describes the revertible process without taking into account the display adaptation feature that is detailed in section 5. From the input dynamic metadata (detailed in section 4) a luma-related look-up table, *lutMapY*, and a color correction look-up table, *lutCC*, are derived. The next step consists in applying the SDR-to-HDR reconstruction from the input SDR picture, the derived luma-related look-up table and color correction look-up table. This process produces an output linear-light HDR picture. An optional gamut mapping can be applied when the color spaces of the SDR picture and of the HDR picture are different (either before or after the SDR-to-HDR reconstruction).

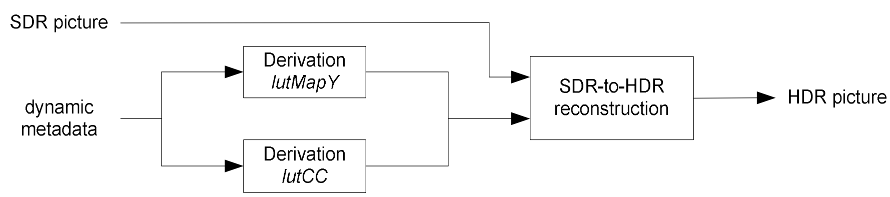

Figure 6 - Overview of the HDR reconstruction process.

The SDR-to-HDR reconstruction process is the functional inverse of the decomposition process. However, for implementation complexity reasons, some operations are concatenated or applied in a different order. This is the case for the final R, G, B reconstruction step described below. The operation reordering and concatenation allows this step to be implemented as a single “output” LUT and this method optionally allows the addition of an output EOTF (such as PQ or HLG) in the same “output” LUT. The LUT *lutMapY* actually corresponds to the inverse of the square-root of the mapping LUT *TM*. The post-processing color correction LUT *lutCC* is actually linked to the pre-processing color correction LUT β and the tone mapping LUT *lutMapY* by the following equation:

$$\beta[Y] = 2^B \times lutMapY[Y] \times lutCC[Y] \quad (6)$$

where B is the bit-depth of the luma signal. It can be demonstrated that the inverse of $lutCC$ is close to a linear function.

The HDR reconstruction process performs the following successive steps for each sample Y, U (Cb component), V (Cr component), of the SDR picture. First U and V are centered (by subtracting the chroma offset, e.g. 512 for a 10 bits signal). Then the variables Y_{post} U_{post} and V_{post} are derived as:

$$Y_{post} = Clamp(0, 2^B - 1, Y + Max(0, a \times U + b \times V)) \quad (7)$$

$$\begin{bmatrix} U_{post} \\ V_{post} \end{bmatrix} = lutCC[Y_{post}] \times \begin{bmatrix} U \\ V \end{bmatrix} \quad (8)$$

The reconstruction of the HDR linear-light *R*, *G*, *B* values is made up of the following steps. A parameter *T* is first computed as:

$$T = k0 \times U_{post} \times V_{post} + k1 \times U_{post} \times U_{post} + k2 \times V_{post} \times V_{post} \quad (9)$$

where *k0*, *k1*, *k2* are predefined parameters that depend on the coefficients of the R'G'B'-to-Y'CbCr conversion matrix *A*. The intermediate values R_{im}, G_{im}, B_{im} are derived as follows:

$$\begin{bmatrix} R_{im} \\ G_{im} \\ B_{im} \end{bmatrix} = A^{-1} \times \begin{bmatrix} \sqrt{1-T} \\ U_{post} \\ V_{post} \end{bmatrix} \quad (10)$$

A clamping is done to 0, $\sqrt{L_{HDR}}$, where L_{HDR} is the HDR mastering display peak luminance.

Then, linear-light *R*, *G*, *B* values are obtained by the following equation:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \left(lutMapY[Y_{post}]\right)^2 \times \begin{bmatrix} R_{im}^2 \\ G_{im}^2 \\ B_{im}^2 \end{bmatrix} \quad (11)$$

It can be demonstrated that equations (9) to (11) invert the pre-processing operation of (1) to (3), that is, the conversion of the HDR version *R*, *G*, *B* into chroma components. When *T* is larger than 1 (which is in principle not possible, but may happen because of quantization and compression), U_{post} and V_{post} are scaled by $1/\sqrt{T}$, the resulting *T* becoming equal to 1. As this scaling applies simultaneously on the two chroma components, the resulting hue remains stable.

4. Metadata Description

The post-processing uses the LUTs *lutMapY* and *lutCC*, and the parameters *a*, *b, k0, k1 and k2*, as dynamic data. These data enable to finely control the texture and colors of the SDR version, and to ensure a good fit to the HDR intent. The LUTs *lutMapY* and *lutCC* are conveyed either using a limited set of parameters (parameter-based mode), or explicitly coded (table-based mode). In both cases, the metadata payload corresponds to a few bytes per video frame or scene. The parameter-based mode may be of interest for distribution workflows which primary goal is to provide direct SDR backward compatible services with very low additional payload or bandwidth usage for carrying the dynamic metadata. The table-based mode may be of interest for workflows equipped with low-end terminals or when a higher level of adaptation is required for representing properly both HDR and SDR streams.

Next to the dynamic metadata, the system uses information that define the properties of the mastering display used when grading the HDR content, as defined in Mastering Display Colour Volume (MDCV) message. This is static information (typically fixed per program) required by the post-processing. It comprises the color gamut of the SDR/HDR signal and the mastering display peak luminance.

In the parameter-based mode, the metadata for reconstructing *lutMapY* consist of the parameters mentioned in section 2.1. For reconstructing *lutCC*, a default pre-defined LUT is used at the post-processing side, and a piece-wise linear table made of at most 6 points is used as a scaling function to adjust the default table. These parameters are conveyed using the parameters defined in the SMPTE ST 2094-20 specification. Typical payload is about 70 bytes per scene, including Mastering Display Colour Volume (MDCV) message. In the table-based mode, *lutMapY* and *lutCC* are explicitly coded using the parameters defined in the SMPTE ST 2094-30 specifications. Typical payload is about 186 bytes per scene, including Mastering Display Colour Volume (MDCV) message. In both cases, the metadata are limited to the codec space. They do not come from the production side, and do not need to be conveyed outside the decoding platform. They are conveyed using standardized metadata containers.

The usage of dynamic metadata allows a fine control of the SDR texture (using the tone mapping LUT *lutMapY*) and of colors (using the color correction LUT *lutCC* and the parameters *a, b, k0, k1* and *k2*). This guarantees the preservation of the HDR texture and intended colors in the SDR version, as illustrated in pictures in next section. High SDR and HDR video quality is obtained, without any strong limitation of the dynamic range and peak luminance (no limitation to peak luminance of around 1000-1500 nits). This also gives high flexibility which enables to easily adapt the system (for instance thanks to the easy control of the dynamic metadata payload) to the distribution workflow.

5. Display Adaptation

The display adaptation feature is only active with parameter-based metadata mode and is based on the Tone Mapping and Inverse Tone Mapping computation blocks of the system.

On the HDR decomposition side (see Figure 3), the luminance mapping block (step 2) computes a Tone Mapping curve based on the dynamic parameters described in section 2.1 and based on the mastering display peak luminance and the targeted SDR display maximum luminance of 100 cd/m2.

On the HDR reconstruction side (see Figure 6), the *lutMapY* derivation block computes an Inverse Tone Mapping curve based on the same dynamic parameters, the same SDR display maximum luminance of 100 cd/m2 and the same HDR mastering display peak luminance. The resulting Inverse Tone Mapping curve is the inverse of the Tone Mapping curve computed by the HDR decomposition block, as depicted in Figure 7.

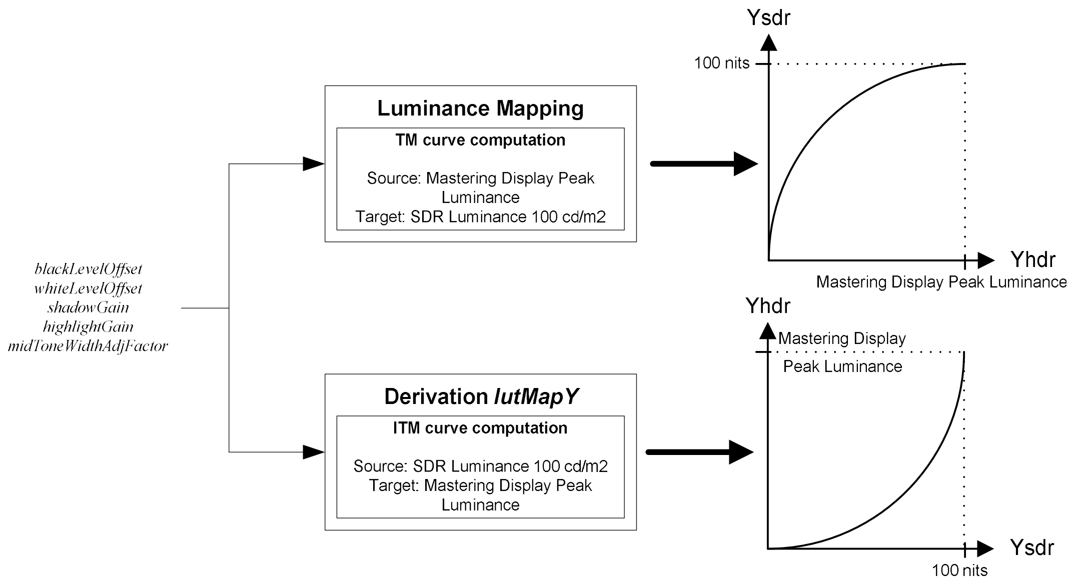

Figure 7 - Tone Mapping and Inverse Tone Mapping computation blocks

The display adaptation feature is an extension of the HDR reconstruction process and takes place in the *lutMapY* derivation and *lutCC* computation process, as shown in Figure 8.

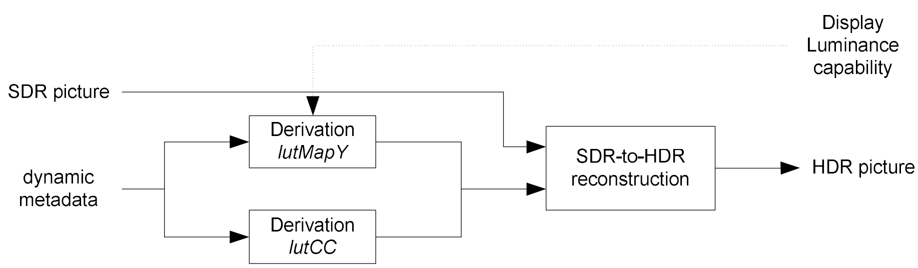

Figure 8 - Overview of the HDR reconstruction process with display adaptation

Given that the attached display provides its peak luminance capability (either through the EDID data of the HDMI connection to the display or inherently when the processing is integrated in the display device), the solution consists of a cascaded calculation of the previously described Inverse Tone Mapping curve and an adapted Tone Mapping curve. This added Tone Mapping block uses the same dynamic parameters and the same input mastering display peak luminance but now relies on the provided presentation display peak luminance capability as the target output peak luminance, as shown in Figure 9:

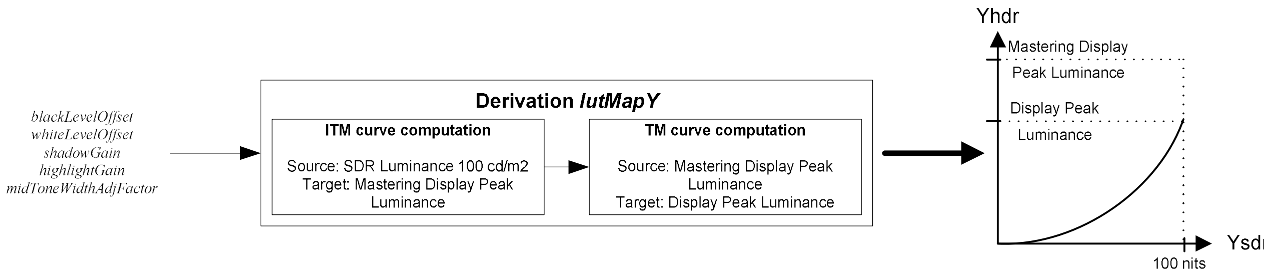

Figure 9 - Inverse Tone Mapping curve computation with display adaptation

The cascading of these two processes is easily implemented in the *lutMapY* LUT with no added complexity.

This process in essence only adapts the luminance of the signal and aims to preserve the artistic intent for a presentation display with a different peak luminance than the targeted SDR or the original HDR peak luminance by means of a reconstructed adapted signal.

6. Performance Evaluations

Two reference distribution solutions, PQ and HLG, can be considered for evaluating an HDR video distribution solution. PQ [14] is a non-SDR backward compatible transfer function, based on the human contrast sensitivity model developed by Barten [15] more adapted to HDR than the usual gamma. HLG [2] is a new transfer function aiming at offering some level of SDR backward compatibility. The two next sections report comparative results of SL-HDR1 vs. distribution solutions using these two transfer functions.

6.1. HDR compression comparison with PQ transfer function

This section reports HDR compression results of SL-HDR1, compared to the non-SDR backward compatible solution HDR10 based on the PQ transfer function. The considered test sequences, selected for the MPEG HDR and WCG Call for Evidence[5], are natively in RGB, 4:4:4, linear-light format, represented in BT.2020 color primaries container. The complete conversion and coding, decoding and back-conversion chain, for all tested solutions has been performed in BT.2020 color primaries container. All contents have been compressed using the MPEG reference HEVC encoder HM 16.2.

Several metrics to assess HDR image objective quality have been proposed at MPEG [5]. Unfortunately, none of them is very satisfactory in the sense that better metric values do not necessarily imply better subjective quality. For this reason, metric values have to be considered with care and are provided in Table 1 for both 4:2:0 subjectively optimized luminance mapping parameters (the automatic tuning is adjusted to get optimal visual SDR quality, left column) and 4:4:4 metric oriented parameter optimization (the automatic tuning is adjusted to get high objective metric DE100 values, right column). The gains are expressed in average bitrate savings compared to the HDR10 reference using the Bjøntegaard-Delta-rate measure [16] and do include the overhead due to the metadata (encoded in dedicated SEI message) associated to the proposed method. A negative value of x% indicates that SL-HDR1 requires x% less bitrate than the reference (PQ/HDR10) for a similar quality.

The metric DE100 is more color oriented and the metric L100 evaluates the luminance quality only. Both metrics are based on an extension of the well-known Lab 2000 color space [17]**Error! Reference source**

not found.. It is observed that high metric gains are possible, but subjectively optimized compressed videos show much less gains, proving once again the inconsistency of the objective metrics. In this test, it has been verified by subjective tests performed by non-naïve viewers that all visually optimized videos have a quality at least comparable to the best proponents of the MPEG HDR and WCG Call for Evidence [5]**Error! Reference source not found.**. The artifacts observed on the reconstructed HDR content include typical compression artefacts already observed in SDR compression such as blocking artefacts and also inconsistent color patches generation especially for color areas reaching the color gamut boundaries.

Table 1 - HDR compression performance on MPEG metrics compared to HDR10.

			420 subjective optim.		444 DE100 optim.
Sequence	Resolution	Peak luminance	DE100	L100	DE100
Market3	HD	4000 nits	-30.4%	-25.1%	-71.8%
AutoWelding	HD	4000 nits	-14.8%	-2.8%	-38.8%
ShowGirl2Teaser	HD	4000 nits	-34.1%	-8.9%	-56.8%
StEM_WarmNight	HD	4000 nits	-17.7%	-2.8%	-41.1%
BalloonFestival	HD	5000 nits	-3.1%	-13.6%	-61.3%
		Average	-20.0%	-10.6%	-53.9%

It should be noted that SDR backward compatibility imposes a natural balance, but not optimal for HDR compression, between chroma and luma. The balance may be compensated by an adequate adjustment of the chroma quantization parameter (controlled by syntax element QPChromaOffSet in HEVC). It has been observed that the tuning of this parameter also improves the quality of HDR10 anchors and it may as well improve the proposed solution. However this raises the issue of having to develop, for HDR content, particular encoders significantly different from encoders optimized for SDR video. An extra-advantage of SDR backward compatible solutions is that existing SDR encoders can be re-used.

6.2. HDR Compression Comparison with HLG Transfer Function

This section reports SDR quality evaluation and HDR compression results of SL-HDR1, compared to the SDR backward compatible solution based on the HLG transfer function. For these tests, visual evaluations made by an independent lab, under supervision of V. Baroncini (MPEG tests chair), have been performed. Two different tests were performed. First tests aimed at evaluating the SDR backward compatibility feature, by checking the quality of the SDR video generated from the HDR content. Second tests aimed at evaluating the HDR compression performance. For these two tests, comparative evaluation was performed between three solutions: HLG, SL-HDR1 with first tuning, SL-HDR1 with second tuning. First tuning tends to generate brighter pictures, i.e. picture with a higher average luminance level, than second tuning. In both cases, the tuning is fully automatic, but is performed according to one of these two different modes.

In the experiments, the HLG implementation from HDRTools software, version 0.12 (accessible at link https://gitlab.com/standards/HDRTools/tags/v0.12), made by the HLG designers, has been used to generate the HLG results. As recommended by the HLG designers, a system gamma correction was applied to the input linear-light RGB HDR content prior to converting it with HLG. The value of the system gamma γ in the HLG pre-processing depends on the peak luminance L_{peak} of the HDR mastering display, and is derived as follows:

$$\gamma = 1.2 + 0.42 \text{ x } \mathrm{Log10}(\ L_{peak} / 1000\ ) \tag{12}$$

6.2.1. Test Sequences

In order to have a future-proof evaluation, and to anticipate the evolution of HDR displays capabilities, sequences with various peak luminance have been used. The content color gamut is either BT.709 or P3D65, but all sequences are represented in BT.2020 color primaries container. All sequences are natively represented in EXR RGB 4:4:4 linear-light half-float format. The complete conversion and coding, decoding and back-conversion chain, for all tested solutions has been performed in BT.2020 color primaries container.

The test sequences are listed in Table 2.

Table 2 - Test Sequences for Comparative Tests with HLG.

Name	Sequence	Peak luminance	Content gamut	Container gamut	fps	Size	duration
S0	Market3	4000	709	2020	50	HD	8s
S1	EBU_04_Hurdles	3000	709	2020	50	HD	10s
S2	EBU_04_Starting	3000	709	2020	50	HD	10s
S3	EBU_13_LongJump	3000	709	2020	50	HD	10s
S4	HdM ShowGirl	5000	P3D65	2020	24	HD	10s
S7	CableLabs Rope	5000	709	2020	24	HD	10s

6.2.2. Evaluation of SDR Quality

The goal of these tests is to verify that colors and texture of the SDR generated by HLG and SL-HDR1 conform to those of the HDR content. The methodology described in [18] is used. For each solution, the viewers have to assess the conformity of the SDR displayed on an SDR monitor to the HDR displayed on an HDR monitor (Sim2). In particular, they have to check the conformity of colors and the texture preservation. This test set-up uses two displays that, when driven with a "black" input signal, become not visible to the viewing subjects. Furthermore, an opaque non reflective curtain is placed between the displays, in a way that, even if the viewer can still watch both displays, any visible interference due to reflections and indirect illumination among the displays is avoided.

Test results are depicted in the graphs of Figure 10 and Figure 11. The vertical axis depicts the Mean Opinion Score (MOS). A score of 10 corresponds to quality of reproduction that is perfectly faithful to the original. A score of 0 denotes a quality of reproduction that has no similarity to the original. A worse quality cannot be imagined. 5 corresponds to a fair quality. The average score value, with related confidence interval, is depicted for each tested solution. In all cases but one, SL-HDR1, with any tuning, was judged much better than HLG. Only for sequence S7, SL-HDR1 tuning1 and HLG are equivalent, and better than SL-HDR1 tuning2. The average scores and confidence intervals for the 6 sequences, and for the three methods, are depicted in Table 3. SL-HDR1 with both tuning outperforms HLG by around 1.7 MOS points.

Illustrative SDR pictures resulting from the HDR conversion by HLG (left) and by SL-HDR1, with first tuning (right) are depicted in Figure 12. Color hue and contrast issues can be observed in HLG versions. In general, texture losses are observed in HLG, especially in bright areas (wall in Market, ground track in Hurdles). And saturated colors (such as red and purple colors) suffer from noticeable hue shifts.

Table 3 - Average MOS and Confidence Intervals

	Average MOS	Confidence interval
HLG	4.82	0.35
SL-HDR1 tuning1	6.56	0.26
SL-HDR1 tuning2	6.54	0.22

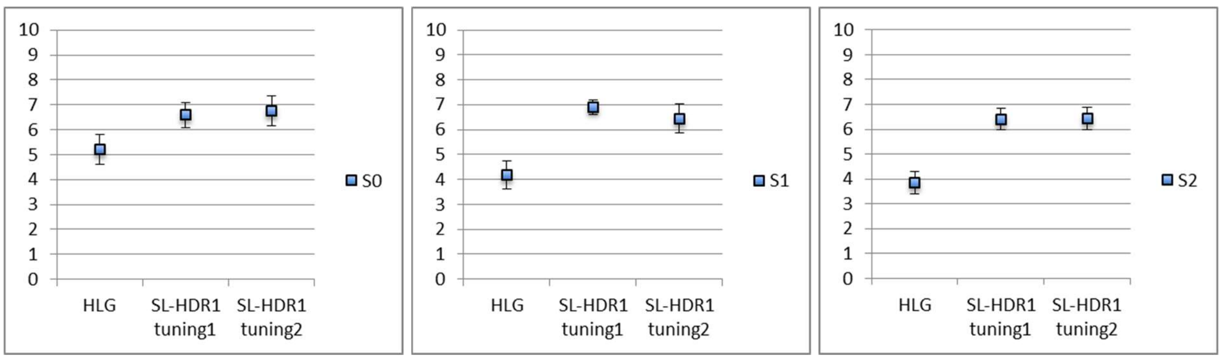

Figure 10 - Average MOS values per sequence, with corresponding confidence interval.

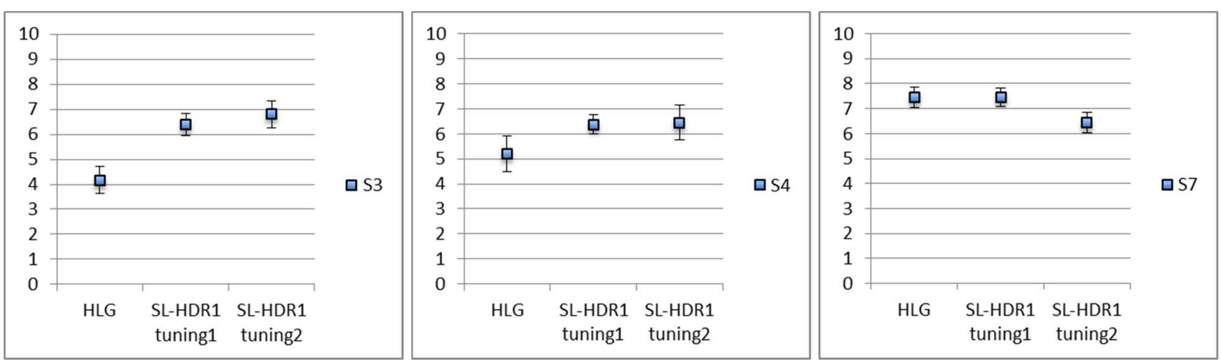

Figure 11 - Average MOS values per sequence, with corresponding confidence interval.

Figure 12 - SDR version from HLG (left) and SL-HDR1 (right) of Market (top) an EBU Hurdles (bottom) HDR first picture.

6.2.3. Evaluation of HDR compression performance

The goal of these tests is to evaluate the quality of the reconstructed HDR using SL-HDR1 compared to the reconstructed HDR using HLG (that is, HDR coded with HLG transfer function, compressed then decompressed with HEVC Main10). The test procedure for the formal subjective evaluation uses one of the test methods described in [19], specifically the Degradation Category Rating (DCR) method. Four bitrates were used, adapted to each test sequence, as listed in Table 4. The encoding was performed using a professional HEVC file encoder from Elemental, Main 10 profile, with same settings for the three tested solutions. The Intra period was set around 1s (24 for 24 fps content, 48 to 50 fps content). The hierarchical-B GOP structure was of size 8, with 4 reference pictures, and B-pictures used as reference.

The average bitrate savings of SL-HDR1 compared to HLG for each sequence have been computed from the MOS vs. bitrate data to quantify the achieved bitrates savings. Table 5 shows the MOS Bjøntegaard-Delta-rate (BD-rate) for each sequence. BD-rate measures as described in [16] were used with MOS scores taking the place of the peak signal-to-noise ratio (PSNR) values that have been typically used with BD-rate measurements, with negative numbers indicating percentage of rate reduction at the same MOS quality. The estimated rate saving by using SL-HDR1 in place of HLG is quite significant. Only in one case (sequence S7), the compression performance with SL-HDR1 tuning 1 is similar to HLG. The average MOS BD-rate gain (that gives an estimation of the bitrate saving) of SL-HDR1 compared to HLG is of 24.6% for tuning 1, and 26.7% for tuning 2.

Table 4 - Bitrates (kbps) per Sequence

Sequence	R1	R2	R3	R4
S0	8000	5000	3000	2000
S1	8000	5000	3000	2000
S2	8000	5000	3000	2000
S3	6000	4000	2000	1000
S4	6000	4000	3000	2000
S7	3000	1700	1000	0700

Table 5 - MOS BD-rate savings measurements.

	SL-HDR1 tuning1 vs HLG	SL-HDR1 tuning2 vs HLG
S0	-33.1%	-27.5%
S1	-13.9%	-9.6%
S2	-38.5%	-50.7%
S3	-37.1%	-17.9%
S4	-23.9%	-41.4%
S7	-1.2%	-13.1%
Avg	**-24.6%**	**-26.7%**

Conclusion

The co-existence of SDR and HDR on one hand, and BT.709 and BT.2020 contents on the other hand is likely to happen and last for at least a few years. Some applications, like broadcasting, will benefit from SDR backward compatible solutions that avoid simulcasting versions with various ranges and gamut. Backward compatibility would make the transition from SDR HDTV to HDR UHDTV smoother with increased interoperability.

The proposed solution, SL-HDR1, addresses the SDR/HDR backward compatibility by offering a new dynamic reducer with consistent SDR/HDR colors. This solution also adapts to the wide range of HDR display brightness capabilities by tuning the content to the display capabilities while preserving the artistic intent. It shows solid compression gain compared to the conservative non-backward compatible HDR10 approach. Tests results also show that SL-HDR1 outperforms HLG in a statistically significant way. This holds for both tests, i.e. both the visual quality of the SDR and that of the reconstructed HDR are better for SL-HDR1 compared to HLG. For SDR visual quality, SL-HDR1 outperforms HLG by 1.7 points in Mean Opinion Score (MOS). For HDR compression, a bitrate saving of around 25% is obtained by SL-HDR1 compared to HLG. SL-HDR1 is compliant with a 4:2:0 distribution workflow as well as with existing HDR color spaces (namely BT.2020 CL CbCr), non-linearity (PQ or HLG EOTF) and bit-depth (10 bits) used as input to rendering devices.

The solution has been designed with a particular focus on low complexity and high performance. The pre- and post-processing are of very low added complexity. The involved operations are pixel-based, without inter-sample or temporal dependency. The complexity increase is very reasonable (a few operations and

LUTs) relatively to the HDR coding gain and the new backward compatible feature that are provided. Associated metadata can be encapsulated in the compressed bit-stream and would not require their transmission from the production to the display. The solution has been standardized as ETSI TS 103 433.

Abbreviations

ATSC	Advanced Television Systems Committee
BD-rate	Bjøntegaard-delta-rate
CE	Consumer Electronics
DCR	degradation category rating
EBU	European Broadcasting Union
EDID	Extended Display Identification Data
ETSI	European Telecommunications Standards Institute
fps	frame per second
GOP	group of pictures
HD	high definition
HDR	high dynamic range
HEVC	high efficiency video coding
HLG	hybrid log gamma
ITM	inverse tone mapping
kbps	kilo bits per second
LUT	look-up table
MDCV	mastering display color volume
MPEG	Moving Picture Experts Group
MOS	mean opinion score
OETF	optical to electrical transfer function
PQ	perceptual quantization
PSNR	peak signal to noise ratio
SCTE	Society of Cable Telecommunications Engineers
SDR	standard dynamic range
SEI	supplemental enhancement information
SHVC	scalabilty extension of HEVC
STB	set-top box
TM	tone mapping
TS	technical specification
TV	television
UHD	ultra high definition
UHDTV	ultra high definition television
WCG	Wide color gamut

Bibliography & References

[1] SMPTE, "High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays", SMPTE ST 2084 (2014).

[2] Borer, T. and Cotton, A., “A display independent high dynamic range television system”, International Broadcasting Convention (IBC), http://www.bbc.co.uk/rd/publications/whitepaper309 (September 2015).
[3] EBU, “High-Performance Single Layer Directly Standard Dynamic Range (SDR) Compatible High Dynamic Range (HDR) System for use in Consumer Electronics devices (SL-HDR1)”, ETSI technical specification ETSI TS 103 433 (May 2016).
[4] Diaz, R., Blinstein, S. and Qu, S., “Integrating HEVC Video Compression with a High Dynamic Range Video Pipeline”, SMPTE Motion Imaging Journal, Vol. 125, Issue 1., pp 14-21 (February 2016).
[5] Luthra, A., François, E. and Husak, W., “Call for Evidence (CfE) for HDR and WCG Video Coding”, ISO/IEC JTC1/SC29/WG11 MPEG2015/N15083 (October 2015).
[6] Luthra, A., François, E. and Husak, W., “Requirements and Use Cases for HDR and WCG Content Coding”, ISO/IEC JTC1/SC29/WG11 MPEG2015/N15084 (October 2015).
[7] François, E., Fogg, C., He, Y., Li, X., Luthra, A. and Segall, A., “High Dynamic Range and Wide Color Gamut Video Coding in HEVC: Status and Potential Future Enhancements”, IEEE TCSVT, Vol. 26, Issue 1, pp 63-75 (July 2015).
[8] François, E., “Standardizing the landscape, High Dynamic Range & Wide Color Gamut Video”, Proceedings of DVB scene, Issue 45, pp. 7 (March 2015).
[9] ITU-R, “Reference electro-optical transfer function for flat panel displays used in HDTV studio production”, Recommendation ITU-R BT.1886 (March 2011).
[10] Luthra, A., François, E and van de Kerkhof, L. 2016, “Report of HDR Core Experiment 7: On the visual quality of HLG generated HDR and SDR video”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-W0027 (February 2016).
[11] Pindoria, M., Naccari, M., Borer T. Cotton, A., “Some considerations on hue shifts observed in HLG backward compatible video”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-W0119 (February 2016).
[12] Holm, J., “Information and Comments on Hybrid Log Gamma”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-W0132 (February, 2016).
[13] Lasserre, S., Le Léannec, F., Poirier, T. and Galpin, F. 2016, “Backward Compatible HDR Video Compression System”, In Proc. Data Compression Conference, pp 309-318 (March, 2016).
[14] Miller, S., Nezamabadi, M., Daly, S., “Perceptual signal coding for more efficient usage of bit codes”, SMPTE Motion Imaging Journal, vol. 122, no. 4, pp 52-59, (May-June 2013).
[15] Barten, P. G. J., Formula for the contrast sensitivity of the human eye, in Proc. SPIE – Image Quality and System Performance, vol. 5294, pp 231–238, San Jose, (December 2003).
[16] Bjøntegaard, G., “Improvements of the BD-PSNR model”, ITU-T SG16/Q6, 35th VCEG meeting document VCEG-AI11, Berlin, Germany (July 2008).
[17] Sharma, G., Wu, W., Dalal, E. N., “The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations”, Color Research & Applications 30 (1), pp 21–30, (2005)
[18] Baroncini, V., “Proposed test methodology of SDR quality evaluation with HDR reference”, MPEG document ISO/IEC JTC1/SC29/WG11 MPEG2015/M37407, Geneva, (October 2015).
[19] International Telecommunication Union – Telecommunication Standardization Sector; “Recommendation ITU-T P.910 Subjective video quality assessment methods for multimedia applications” (September 1999).

