|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  26th Meeting: Geneva, CH, 12–20 January 2017 | Document: JCTVC-Z0023 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Suggested new draft text of Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding** | | |
| *Status:* | Proposal | | |
| *Purpose:* | Draft text of technical report | | |
| *Author(s) or Contact(s):* | Edouard François Dmytro Rusanovskyy Peng Yin Pankaj Topiwala Gary J. Sullivan Matteo Naccari | Email: | [edouard.francois@technicolor.com](mailto:edouard.francois@technicolor.com) [dmytror@qti.qualcomm.com](mailto:dmytror@qti.qualcomm.com) [pyin@dolby.com](mailto:pyin@dolby.com) [pankaj@fastvdo.com](mailto:pankaj@fastvdo.com) [garysull@microsoft.com](mailto:garysull@microsoft.com) [matteo.naccari@bbc.co.uk](mailto:matteo.naccari@bbc.co.uk) |
| *Source:* | JCT-VC | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document is a proposed revised version of the draft technical report on Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding, based on document JCTVC-Y1012 / MPEG N16508.

This document is a draft of a technical report on Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding. This report complements and extends work in another report, “Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics”. Specifically, this report expands on the application of ICTCP, HLG and SEI messages in the coding of HDR/WCG video.

Summary

This document relates to high dynamic range (HDR) and wide colour gamut (WCG) video distribution, based on the High Efficiency Video Coding (HEVC) standard (Rec. ITU-T H.265 | ISO/IEC 23008-2), and when applicable on the Advanced Video Coding (AVC) standard (Rec. ITU-T H.264 | ISO/IEC 14496-10), using either single-layer or dual-layer coding. This document serves several purposes. It covers some use cases of HDR/WCG video content distribution. It reviews approaches, using supplemental enhancement information (SEI) messages, to enable backward compatibility with legacy decoding systems that do not have the ability to detect, decode, or properly display HDR/WCG video content. It also describes methods for adapting HDR/WCG video content for use with display technologies in which the dynamic range and colour gamut capability differs from the original content. The document surveys the signalling mechanisms for handling HDR/WCG video with HEVC, and when applicable, with AVC. Indicators are provided for usage of the ICTCP colour representation, the hybrid log-gamma (HLG) and Perceptual Quantizer (PQ) transfer functions, and for usage of different specific SEI messages for SDR backward compatibility and display adaptation functionalities. A dual-layer coding approach using the Scalable Main 10 profile of HEVC for backward compatibility with standard dynamic range (SDR) systems is also documented.

This report complements the material provided in another report, entitled “Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics”. In brief, it reviews the use of certain SEI messages to achieve the capabilities mentioned above.

# Scope

This document relates to high dynamic range (HDR) and wide colour gamut (WCG) video coding and distribution, using single-layer or dual-layer coding, with HEVC (and when applicable, AVC) signalling. It provides a review of identified VUI syntax elements and SEI messages applicable for HDR/WCG video distribution. Guidance for usage of ICTCP colour representation and usage of hybrid log-gamma (HLG) transfer function is provided. Usage of colour remapping information (CRI) and tone mapping information (TMI) SEI messages for support of SDR backward compatibility and display adaptation functionalities is also documented.

**SDR backward compatibility** relates to the ability of HDR/WCG video coding and distribution systems to produce a video signal suitable for SDR-only capable rendering devices (e.g. UHD SDR display with Rec. ITU-R BT.2020 colour primaries [3]). It can be defined in two modes:

* In HDR/WCG distribution systems that support **bitstream** SDR backward compatibility, the decoded video signal from a standard-compliant decoder (e.g. HEVC Main 10 decoder) can be directly displayed on an SDR-capable display without adaptation. Two categories of **bitstream** SDR backward compatibility can be considered:
  + In **“static” bitstream** SDR backward compatibility, the decoded video is an HDR signal (for instance, Y′CbCr 4:2:0 10-bits with the Rec. ITU-R BT.2100 HLG transfer function [3] and Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100 colour primaries) that can be directly displayed on an HDR-capable display or an SDR-capable display, without adaptation. In this context, the HDR processing chain is a fixed function, not dependent on the input video data.
  + In **“dynamic”** **bitstream** SDR backward compatibility, the decoded video is an SDR signal. A post-processing step can be further used to reconstruct an HDR signal, e.g. using information conveyed in CRI or TMI SEI messages. In this context, the HDR processing chain is dynamic, and adapts to the input video data.
* In HDR/WCG distribution systems that support **display** SDR backward compatibility, the decoded video signal from a standard-compliant decoder (e.g. HEVC Main 10 decoder) is an HDR signal (for instance, Y′CbCr 4:2:0 10-bits with Rec. ITU-R BT.2100 PQ transfer function [3] and Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100 colour primaries). A post-decoding adaptation process is applied to the decoded video signal to produce an SDR video signal that can be displayed on an SDR-capable display. The adaptation process can use metadata, conveyed for example in CRI SEI messages, to perform this conversion.

**Display adaptation** is a generic term covering techniques of video signal processing which adapts the decoded video signal to a target display. Techniques providing display SDR backward compatibility can be considered as a subset of display adaptation. Display adaptation techniques aim at converting an HDR/WCG video signal, originally produced for a reference display capable to display a certain colour volume (dynamic range and colour gamut), to a video signal suitable to a target rendering device of different (from reference display) colour volume capabilities. For instance, it can be used to convert a non-constant luminance (NCL) Y′CbCr 4:2:0 10-bits Rec. ITU-R BT.2100 PQ signal (denoted PQ10 in the present document) based on a master produced on a display with a given peak luminance, to a lower peak luminance capable display (e.g., 1000 candelas per square metre (cd/m2) to 100 cd/m2). Display adaptation could also increase the colour volume, if desired. Another term used in the industry for display adaptation is regrading. Display adaptation can be driven by metadata transmitted along with the video bitstream, typically using SEI messages.

NOTE – Conversion and coding practices related to production and compression of HDR/WCG video signal represented with NCL Y′CbCr 4:2:0 video with Rec. ITU-R BT.2100 PQ transfer characteristics are outside of scope of this document. These aspects are specifically addressed in [1][2].

# Definitions

This document defines the following terms:

* 1. **Dynamic range adaptation**: A mapping process of a content from one colour volume to another colour volume.
  2. **Hybrid log-gamma**: One of the two transfer functions specified in Rec. ITU-R BT.2100 [3].
  3. **Inverse transfer function:** The inverse of either the camera Opto-Electronic Transfer Function or the display Electro-Optical Transfer Function. Note, the Opto-Electronic Transfer Function and the Electro-Optical Transfer Function are not the inverse of one another.
  4. **Narrow range:** A range in a fixed-point (integer) representation that does not span the full range of values that could be expressed with that bit depth. In this document the range from 64 (black) to 940 (peak white) is used for Y′, and the range from 64 to 960 is used for Cb and Cr for 10 bit representations, which is aligned with [3]. Narrow range is in some applications called by synonyms such as: “limited range”, “video range”, “legal range”, “SMPTE range” or “standard range.”
  5. **Full range:** A range in a fixed-point (integer) representation that spans the full range of values that could be expressed with that bit depth. For full range 10-bit signals, black corresponds to code 0 and peak white corresponds to code 1023 [3].
  6. **Perceptual quantizer (PQ)**: One of the two transfer functions specified in Rec. ITU-R BT.2100 [3].
  7. **Transfer function:** The function used to convert between linear light (linear colour value) and non-linear electrical signals (or non-linear codeword values) in either the pre-encoding process or post-encoding process. For scene referred systems (e.g. hybrid log-gamma, Rec. ITU-R BT.709, Rec. ITU-R BT.601) the camera Opto-Electronic Transfer Function (OETF) has primacy. In display-referred systems (e.g. PQ systems) the display Electro-Optical Transfer Function (EOTF) has primacy.

# Abbreviations and acronyms

This document uses the following abbreviations and acronyms:

AVC Advanced Video Coding, specified in Rec. ITU-T H.264 | ISO/IEC 14496-10

CL Constant Luminance

CI Constant Intensity

CLVS Coded Layer-wise Video Sequence

CRI Colour Remapping Information SEI message

DRA Dynamic Range Adaptation

EOTF Electro-Optical Transfer Function

HDR High Dynamic Range

HEVC High Efficiency Video Coding, specified in Rec. ITU-T H.265 | ISO/IEC 23008-2

HLG Hybrid Log-Gamma

ICTCP Alternative colour space representation to Y′CbCr, specified in Rec. ITU-R BT.2100

LMS Long, Medium, and Short wavelength-based colour space, specified in Rec. ITU-R BT.2100

LUT Look-up Table

NCL Non-Constant Luminance

PQ Perceptual Quantizer, specified in Rec. ITU-R BT.2100

PQ10 HDR content representation that utilizes the 4:2:0 chroma sampling format, Rec. ITU-R BT.2100 colour primaries, the Rec. ITU-R BT.2100 PQ transfer function (PQ), and Rec. ITU-R BT.2100 non-constant luminance Y′CbCr encoding

QP Quantization Parameter

RAPAU Random Access Point Access Unit

RGB Red, Green, and Blue component colour system

SDR Standard Dynamic Range

SEI Supplemental Enhancement Information

OETF Opto-Electronic Transfer Function

OOTF Opto-to-Optical Transfer Function; in general, an OOTF is a concatenation of OETF, artistic adjustments and EOTF

SMPTE Society of Motion Picture and Television Engineers

TF Transfer Function

TMI Tone Mapping Information SEI message

VUI Video Usability Information

WCG Wide Colour Gamut

XYZ The CIE 1931 colour space. Y corresponds to the luminance signal.

Y′CbCr Colour space representation commonly used for video/image distribution as a way of encoding RGB information. Also commonly expressed as YCbCr or Y′C′BC′R. The relationship between Y′CbCr and RGB is dictated by certain signal parameters, such as colour primaries, transfer characteristics, and matrix coefficients. Unlike the (constant luminance) Y component in the XYZ representation, Y′ in this representation might not be representing the same quantity. Y′ is commonly referred to as “luma”. Cb and Cr are commonly referred to as “chroma”.

# Conventions

## General

NOTE – The mathematical operators used in this document are similar to those used in the C programming language. However, the results of integer division and arithmetic shift operations are defined more precisely, and additional operations are defined, such as exponentiation and real-valued division. Numbering and counting conventions generally begin from 0, e.g., "the first" is equivalent to the 0-th, "the second" is equivalent to the 1-th, etc.

## Arithmetic operators

The following arithmetic operators are defined as follows:

|  |  |
| --- | --- |
|  | Addition |
| − | Subtraction (as a two-argument operator) or negation (as a unary prefix operator) |
| \* | Multiplication, including matrix multiplication |
| xy | Exponentiation. Specifies x to the power of y. In other contexts, such notation is used for superscripting not intended for interpretation as exponentiation. |
| / | Integer division with truncation of the result toward zero. For example, 7 / 4 and (−7) / (−4) are truncated to 1 and (−7) / 4 and 7 / (−4) are truncated to −1. |
| ÷ | Used to denote division in mathematical equations where no truncation or rounding is intended. |
|  | Used to denote division in mathematical equations where no truncation or rounding is intended. |
|  | The summation of f( i ) with i taking all integer values from x up to and including y. |
| x % y | Modulus. Remainder of x divided by y, defined only for integers x and y with x >= 0 and y > 0. |

## Bit-wise operators

The following bit-wise operators are defined as follows:

|  |  |
| --- | --- |
| & | Bit-wise "and". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| | | Bit-wise "or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| ^ | Bit-wise "exclusive or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| x >> y | Arithmetic right shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y. Bits shifted into the MSBs as a result of the right shift have a value equal to the MSB of x prior to the shift operation. |
| x << y | Arithmetic left shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y. Bits shifted into the LSBs as a result of the left shift have a value equal to 0. |

## Assignment operators

The following arithmetic operators are defined as follows:

|  |  |
| --- | --- |
|  | Assignment operator |
|   | Increment, i.e., x  is equivalent to x  x  1; when used in an array index, evaluates to the value of the variable prior to the increment operation. |
| − − | Decrement, i.e., x− − is equivalent to x  x − 1; when used in an array index, evaluates to the value of the variable prior to the decrement operation. |
| += | Increment by amount specified, i.e., x += 3 is equivalent to x = x + 3, and x += (−3) is equivalent to x = x + (−3). |
| −= | Decrement by amount specified, i.e., x −= 3 is equivalent to x = x − 3, and x −= (−3) is equivalent to x = x − (−3). |

## Mathematical functions

EOTFPQ( x ) the PQ EOTF used to convert a non-linear light PQ representation to a linear light representation.

Floor( x ) the largest integer less than or equal to x.

a a representation

Ln( x ) the natural logarithm of x (the base-e logarithm, where e is natural logarithm base constant 2.718 281 828...).

Log10( x ) the base-10 logarithm of x.

OETFHLG( x ) the HLG OETF used to convert a scene-referred linear light representation to a non-linear light representation.

Round( x ) = Sign( x ) \* Floor( Abs( x ) + 0.5 )

Sqrt( x ) =

## Order of operations

When order of precedence in an expression is not indicated explicitly by use of parentheses, the following rules apply:

– Operations of a higher precedence are evaluated before any operation of a lower precedence.

– Operations of the same precedence are evaluated sequentially from left to right.

Table 4‑1 specifies the precedence of operations from highest to lowest; a higher position in the table indicates a higher precedence.

NOTE – For those operators that are also used in the C programming language, the order of precedence used in this document is the same as used in the C programming language.

Table 4‑1 – Operation precedence from highest (at top of table) to lowest (at bottom of table).

|  |
| --- |
| **Operations (with operands x, y, and z)** |
| "x++", "x− −" |
| "!x", "−x" (as a unary prefix operator) |
| "xy" |
| "x \* y", "x / y", "x  y""", "x % y" |
| "x + y", "x − y" (as a two-argument operator), "" |
| "x  <<  y", "x  >>  y" |
| "x < y", "x  <=  y", "x > y", "x  >=  y" |
| "x  = =  y", "x  !=  y" |
| "x & y" |
| "x | y" |
| "x  &&  y" |
| "x  | |  y" |
| "x ? y : z" |
| "x..y" |
| "x = y", "x  +=  y", "x  −=  y" |

# Overview

The document is structured as follows.

* Clause 6 reviews identified signalling mechanisms of the HEVC specification [8][9] and its draft new amendment [16], applicable to HDR/WCG video coding and distribution.
* Clause 7 describes usage of HLG transfer function and CRI and TMI SEI messages for the support of SDR backward compatibility with single-layer profile (e.g. HEVC Main 10).
* Clause 8 describes a dual-layer HDR/WCG video coding system with SDR backward compatibility implemented with SHVC Scalable Main 10 profile.
* Clause 9 addresses the display adaptation functionality.

# HEVC signalling mechanisms applicable to HDR/WCG video

In this clause, VUI syntax elements and SEI messages specified in HEVC [8][9] and its new draft amendment [16], applicable to HDR/WCG video and relevant to the scope of this document, are reviewed.

## VUI syntax elements

By design, metadata signalled in syntax elements of VUI in HEVC is not mandatory for constructing the luma or chroma samples by the decoding process, and may be ignored by the decoder. However, such syntax elements can provide useful parameters or attributes of an encoded signal and be utilized in the video system design. Examples of VUI parameters relevant to HDR/WCG video system design include colour primaries, transfer characteristics and matrix coefficients specified in Tables E.3, E.4 and E.5 of HEVC specification [8][9] respectively. Table 6‑1 and Table 6‑2 below provide values of VUI syntax elements that indicate usage of Rec. ITU-R BT.2100 representation of the video signal, including matrix coefficients associated with Rec. ITU-R BT.2020.

Rec. ITU-R BT.2100 specifies HDR-TV image parameters for use in production and international programme exchange [3]. It defines two transfer functions: Perceptual Quantizer (PQ) and Hybrid Log-Gamma (HLG). RGB colour primaries are specified identically as in Rec. ITU-R BT.2020 [4]. Rec. ITU-R BT.2100 describes two different luminance and colour difference signal representations: Non-Constant Luminance (NCL) Y′CbCr and Constant Intensity (CI) ICTCP. Syntax elements of VUI of HEVC can be used to convey the metadata describing such attributes of the coded signal. A transfer function is indicated using transfer\_characteristics syntax element, as described in Table 6‑1. RGB colour primaries are indicated using colour\_primaries syntax element (set equal to 9 for Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100 colour primaries). Colour representation is indicated using matrix\_coeffs syntax element, as described in Table 6‑2.

Conversion and coding practices for HDR/WCG video signals represented with NCL Y′CbCr 4:2:0 video with Rec. ITU-R BT.2100 PQ transfer characteristics are addressed in [1][2], and are outside of scope of this document.

Conversion and coding practices related to HDR/WCG video signals represented with ICTCP 4:2:0 video with Rec. ITU-R BT.2100 PQ transfer function are provided in clause 6.3.

Conversion and coding practices related to HDR/WCG video signals represented with NCL Y′CbCr 4:2:0 video with Rec. ITU-R BT.2100 HLG transfer characteristics are discussed in sub-clause 7.1.2.

Table 6‑1 – Values of transfer\_characteristics indication in VUI in HEVC.

|  |  |  |
| --- | --- | --- |
|  | **PQ** | **HLG** |
| **transfer\_characteristics** | 16 | 18 |

Table 6‑2 – Values of matrix\_coeffs indication in VUI in HEVC.

|  |  |  |
| --- | --- | --- |
|  | **NCL Y′CbCr** | **CI ICTCP** |
| **matrix\_coeffs** | 9 | 14 |

NOTE 1 – PQ is also specified in SMPTE ST 2084 [12] and HLG is also specified in ARIB STD-B67.

NOTE 2 – VUI syntax elements values for PQ are also defined in AVC [6][7].

### PQ EOTF

The inverse PQ EOTF is described in Table E.4 of the HEVC specification [8][9] for transfer\_characteristics equal to 16. More specifically, the non-linear light representation V of a linear light intensity signal Lo, which takes values normalized to the range [0, 1], can be computed as:

(6‑1)

where c1, c2, c3, m, and n are constants defined as follows:

c1 = c3 − c2 + 1 = 3424 ÷ 4096 = 0.835 937 5 (6‑2)

c2 = 2413 ÷ 128 = 18.851 562 5 (6‑3)

c3 = 299 ÷ 16 = 18.687 5 (6‑4)

m = 2523 ÷ 32 = 78.843 75 (6‑5)

n = 1305 ÷ 8192 = 0.159 301 757 812 5 (6‑6)

The peak value of 1 for Lo is ordinarily intended to correspond to an intensity level of 10 000 cd/m2, while the value of 0 for Lo is ordinarily intended to correspond to an intensity level of 0 cd/m2.

The PQ (EOTF) is described as the exact inverse of equation (6‑1). More specifically, the linear light intensity signal Lo can be computed from the non-linear representation V, which takes values in the range [0, 1], as follows:

(6‑7)

A plot of the inverse PQ EOTF is shown in Figure 6‑1.

Figure 6‑1 : Graph of the inverse PQ EOTF (iEOTFPPQ).

### HLG OETF

The HLG OETF, OETFHLG( ), as described in Table E.4 of the HEVC specification [8][9] for transfer\_characteristics equal to 18, applies as follows to normalized linear light scene referred R, G, B samples. For Lc = R, G, or B, the following applies:

V = OETFHLG( Lc ) = a \* Ln( 12 \* Lc − b ) + c for 1 >= Lc > ( 1 ÷ 12 ) (6‑8)

= Sqrt( 3 ) \* Lc0.5 for( 1 ÷ 12 ) >= Lc >= 0

where a, b and c are constants defined as follows:

a = 0.178 832 77 (6‑9)

b = 1 – 4 \* a = 0.284 668 92 (6‑10)

c = 0.5 – a \* Ln(4 \* a) = 0.559 910 73 (6‑11)

A plot of the OETF for HLG is shown in Figure 6‑2.

Figure 6‑2 : Graph of the HLG OETF (OETFHLG).

The inverse of the HLG OETF is formulated as follows:

Lc = iOETFHLG( V ) = Exp( ( V - c ) / a ) + b for 1 >= V > ( 1 ÷ 2 ) (6‑12)

= V2 / 3 for( 1 ÷ 2 ) >= V >= 0

Note that the inverse of the OETF does not equate to the EOTF. For more information on the HLG EOTF, refer to [3].

NOTE – PQ and HLG systems do not work similarly, as explained in details in [3][5]. In both cases, there is an end-to-end OOTF (opto-optical transfer function) which adjusts the displayed light to compensate for changes in the response of the human visual system as it adapts (however, the OOTF defined for HLG is different from the OOTF defined for PQ). Referring to BT.2100, the PQ OETF (scene linear light to non-linear representation) is defined based on the EOTF (non-linear to display linear light representation), as OETF = EOTF-1(OOTF). The HLG EOTF (non-linear to display linear light representation) is defined based on the OETF (scene linear light to non-linear representation), as EOTF = OOTF(OETF−1). This explains why the functions “iEOTF” (i.e. EOTF−1) and “iOETF” (i.e. OETF−1) are used in sub-clauses 6.1.1 and 6.1.2.

## SEI messages applicable for HDR

As in the case of VUI, metadata signalled in SEI messages of HEVC is not mandatory for constructing the luma or chroma samples by the decoding process, and may be ignored by the decoder. However, this signalling mechanism can be utilized in HDR/WCG video distribution systems to convey information providing additional description (in addition to VUI) of the coded content or can be used to convey information enabling a post-decoding processing. HEVC contains several SEI messages applicable to HDR/WCG video coding and distributing systems. A first set of SEI metadata convey descriptive information about the content, the way it was produced, the nominal conditions it is assumed to be rendered, how it should be interpreted by a display. These SEI messages are reviewed in clauses 6.2.1 to 6.2.4. A second set of SEI metadata is more specifically devoted to enabling specific post-processing of the decoded samples, e.g. signal adaptation processes. For example, some SEI metadata can be used for the conversion of decoded content from one colour volume to another colour volume. This approach can apply, for example, in bitstream SDR backward compatibility use cases for SDR-to-HDR conversion (the decoded signal is SDR, and the converted signal after post-processing using metadata conveyed in the SEI message is HDR). This approach can also apply in display adaptation use cases for converting a decoded HDR signal to an HDR version of lower colour volume, down to an SDR version for display SDR backward compatibility. These SEI messages are described in clauses 6.2.5 and 6.2.6.

### Mastering display colour volume SEI message

The mastering display colour volume (MDCV) SEI message specifies the colour gamut and dynamic range of a hypothetical monitor used for viewing while authoring the video content. It conveys the colour primaries and white point of the monitor, expressed in the CIE 1931 xyY colour space [11], and provides the minimum and maximum linear light luminance of the monitor (expressed in candelas per square metre, denoted cd/m2). The indicative information provided by MDCV may assist the receiving system in adapting the received video content for local display with characteristics that may differ from the assumed mastering display characteristics. The MDCV SEI message persists until the end of the CLVS. All MDCV SEI messages that apply to the same CLVS shall have the same content.

NOTE 1 – The mastering display colour volume SEI message is also defined in AVC [5][6].

NOTE 2 – Mastering display colour volume metadata are also specified in SMPTE ST 2086 [13].

### Content light level information SEI message

The content light level information (CLL) SEI message conveys the maximum light level and average light level, in the linear light domain (expressed in cd/m2), among the 4:4:4 R, G, B samples of the content pictures in the coded video sequence. Similarly to MDCV, the indicative information provided by CLL SEI message may assist the receiving system in adapting the received video content to local display capabilities. It can be used for instance to help better controlling the energy consumption for local display [23]. The CLL SEI message persists until the end of the CLVS. All CLL SEI messages that apply to the same CLVS shall have the same content.

NOTE 1 – The content light level information SEI message is also defined in AVC [5][6].

NOTE 2 – Corresponding metadata associated with CLL and example of derivation algorithms are specified in CEA-861.3 [15].

### Ambient viewing environment SEI message

The ambient viewing environment (AVE) SEI message aims at describing the ambient viewing environment assumed when mastering the associated video content. It conveys the environmental illuminance and chromaticity coordinates (in the CIE 1931 xyY colour space [11]) of the mastering nominal ambient viewing environment. This indicative information may assist the receiving system in adapting the received video content for local display in viewing environments that may differ from those assumed when mastering the video content. The AVE SEI message persists until the end of the CLVS. All AVE SEI messages that apply to the same CLVS shall have the same content.

### Alternative transfer characteristics SEI message

The alternative transfer characteristics (ATC) SEI message provides a preferred alternative value for the transfer\_characteristics syntax element that is indicated by the colour description syntax of VUI parameters. This SEI message is especially applicable in bitstream SDR backward compatibility use cases, where the VUI transfer\_characteristics syntax element is supposed to signal an SDR transfer function (transfer\_characteristics values 1, 6, 14 or 15). The ATC SEI message can be conveyed to indicate the transfer function used to produce an SDR-compatible HDR video signal representation. This is applicable to the HLG transfer function (transfer\_characteristics value 18), further discussed in clause 7.1. The ATC SEI message persists until the end of the CLVS. All ATC SEI messages that apply to the same CLVS shall have the same content.

### Tone mapping information SEI message

The tone mapping information (TMI) SEI message is designed to carry one or more tone mapping curves within a coded video sequence. The tone mapping curves are used to convert the decoded image to a mapped image, for instance to target a specific display. Four ways of implemented the tone mapping function are specified for the TMI SEI message, including by a linear function with a clip, a sigmoidal function, a piece-wise linear function, or an explicit 1D look-up table (LUT). In practice, for each case, the mapping curve can be represented by a 1D-LUT LUTTM. The tone mapping function applies either to the luma component, or simultaneously to the three RGB components of the decoded signal, for instance as indicated below.

(6‑13)

The TMI SEI message includes a syntax element tone\_map\_id that may be used to identify the purpose of the TMI. The TMI SEI message also includes a syntax elementtone\_map\_model\_id that specifies the model type used for mapping the coded data.

More than one tone mapping can be associated with a coded video sequence through the tone mapping information SEI message identifier tags. This enables simultaneous support for multiple dynamic range targets, including targets that have greater or less dynamic range than the decoded video data.

NOTE – The tone mapping information SEI message is also specified in AVC [5][6].

### Colour remapping information SEI message

The colour remapping information (CRI) SEI message conveys information used to remap decoded pictures from one colour volume to another one. The syntax of CRI remapping model includes three parts: a first piece-wise linear function applied to each colour component (“Pre-LUT”), followed by a three-by-three matrix applied to the three resulting colour components, and followed by a second piece-wise linear function applied to each resulting colour component (“Post-LUT”). Each one of these sets of data is optional (for instance, only the Pre-LUTs can apply, leading to the application of only one transfer function to each colour component of the input signal). A maximum of 33 pivot points per LUT may be coded to specify the piece-wise linear functions. When the three-by-three matrix is activated, the conversion process using CRI must apply in 4:4:4. When it is not activated, the mapping process can apply in 4:2:2 or 4:2:0 chroma sampling formats.

The following equations illustrate the application of the complete CRI model to the R, G, B values of a colour sample (CRI can also be applied to Y′CbCr samples):

(6‑14)

(6‑15)

(6‑16)

The CRI SEI message includes a syntax element colour\_remap\_id that may be used to identify the purpose of the colour remapping information. For instance, colour\_remap\_id value may be used to indicate that the input of the remapping is the result of a first conversion process, such as conversion to Y′CbCr, RGB or GBR colour representation, or to enable cascading of different remapping processes, or to support conversion for multiple dynamic range targets.

The CRI SEI message also includes syntax elements that convey information to describe of the resulting colour volume of post-processed signal, namely colour\_remap\_video\_signal\_info\_present\_flag, colour\_remap\_transfer\_function, colour\_remap\_full\_range\_flag, colour\_remap\_primaries, colour\_remap\_matrix\_coefficients, colour\_remap\_output\_bit\_depth. The purpose of these syntax elements are analogous to the purpose of colour description syntax of VUI.

## ICTCP colour representation

An overview of CI ICTCP colour representation is given in [3][5]. This clause mainly provides a set of guidelines on processing HDR/WCG video for consumer distribution, including conversion steps for converting from a linear light RGB representation with Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100 colour primaries to a 10-bits, narrow range, PQ transfer function (as defined in Rec. ITU-R BT.2100), 4:2:0, ICTCP representation.

The HDR/WCG system described in this clause follows the same workflow of 10-bits, narrow range, PQ transfer function, 4:2:0, non-constant luminance Y′CbCr representation described in [1][2]. It consists of four major stages:

* a pre-encoding stage consisting of several pre-processing processes (clause 6.3.1),
* an encoding stage (clause 6.3.2),
* a decoding stage (clause 6.3.3), and
* a post-decoding stage, also consisting of several post-processing processes (clause 6.3.4).

These four stages are applied sequentially, with the output of one stage being used as input to the next stage.

It is assumed that both the input to and the output of the HDR/WCG system are 4:4:4, linear light, floating-point signals, in an RGB colour representation using the same colour primaries. The output signal is targeted to resemble the input video signal as closely as possible. Other video formats can be used as input to the HDR/WCG system by first converting them to the above defined input signal representation.

NOTE 1 – The main goal of this clause is to highlight the conversion and coding differences between PQ ICTCP and PQ NCL Y′CbCr signal. More details on common parts of these two systems can refer specifically in [1][2].

NOTE 2 – The same workflow can be used for HLG ICTCP by substituting PQ with HLG.

### Pre-encoding process

The pre-encoding process described in this document includes the following components, as presented in Figure 6‑3 :

1. a conversion component from a linear RGB data representation to a linear LMS data representation, described in sub-clause 6.3.1.1
2. a conversion component from a linear LMS data representation to a non-linear data representation using the PQ transfer function, described in sub-clause 6.3.1.2,
3. a colour format conversion component that converts data to the ICTCP representation, described in clause 6.3.1.3,
4. a conversion component that converts a floating-point to a fixed-point representation (e.g. 10 bits), following the same process described in sub-clause 6.1.4 in [1][2], and
5. a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, following the same process described in clause 6.1.3 in [1][2] and further clarified in sub-clause 6.3.1.4.



Figure 6‑3 – Conventional PQ ICTCP pre-encoding process system diagram.

#### Conversion from a linear RGB representation to a linear LMS representation

Conversion from a linear RGB to a linear LMS representation is performed using a 3x3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100.

(6‑17)

#### Conversion from a linear to a non-linear representation: LMS to L'M'S'

The inverse PQ EOTF (equation (6‑1)) is applied to all L, M, and S linear light samples, where each component is a number between 0.0 (representing no light) and 1.0 (representing 10 000 cd/m2). This results in their non-linear light counterparts L′, M′, and S′ as follows.

L′ = iEOTFPQ( L ) (6‑18)

M′ = iEOTFPQ( M ) (6‑19)

S′ = iEOTFPQ( S ) (6‑20)

#### Colour representation conversion: L′M′S′ to ICTCP

Conversion from L′M′S′ to ICTCP representation is performed using a 3x3 matrix conversion process.

(6‑21)

#### Chroma down-conversion

Converting the ICTCP video data from a 4:4:4 representation to a 4:2:0 representation follows the same process described in clause 6.1.3 in [1][2].

It is noted that the closed loop luma adjustment method in clause 6.2 in [1][2], which aims at fixing NCL Y'CbCr chroma down-sampling artifacts, does not prove to be useful for ICTCP chroma down-conversion [25]. This can be explained by the constant luminance property of ICTCP. As shown in Figure 6‑4, the achromatic axis of Y′ deviates substantially from luminance especially when colour goes more saturated. On the other hand, the achromatic axis of ICTCP (I) corresponds very closely with luminance. The constant luminance property of ICTCP can reduce chroma leakage errors that can be introduced when spatially sub-sampling the chroma components (such as the 4:2:0 widely used for compression) compared with NCL Y′CbCr*.*

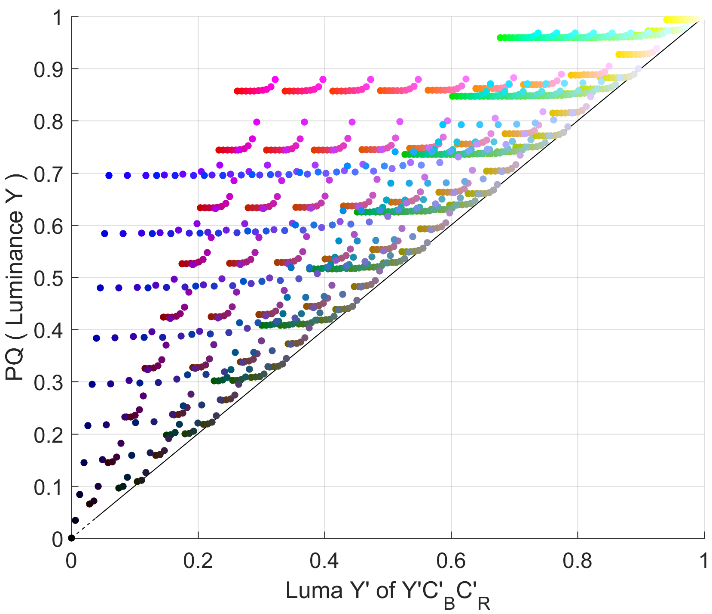
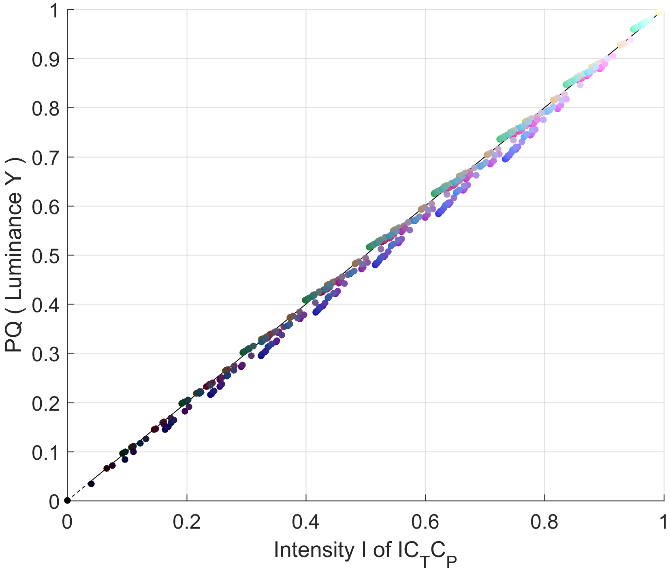


Figure 6‑4 – Luminance correlation with ICTCP and Y′CbCr colour representations.

### Encoding process

After pre-processing, the data is ready for compression. [1][2] presents some guidance on how an encoder may be configured for HDR/WCG 10-bits, narrow range, PQ, 4:2:0, NCL Y′CbCr data. Two quantization methodologies, perceptual luma quantization, in clause 7.1, and chroma QP offset in clause 7.2, are detailed [1][2]. For the HDR/WCG 10-bits, narrow range, PQ, 4:2:0, ICTCP data coming out from pre-processing step, those two methodologies are still valid. Study shows that ICTCP exhibits different characteristics than NCL Y′CbCr data for current test data [21], therefore the exact parameter settings for these two quantization methods might need to be reconfigured. One practice for ICTCP encoder is to configure the encoder settings to follow the bitrate allocation behavior of NCL Y'CbCr [21]. An example of encoder setting is described in clause 6.3.2.1.

Apart from modifying the QP allocation in the encoder, it may also be desirable for an encoder manufacturer to adjust other non-normative encoding processes in their encoders, such as the motion estimation, intra and inter mode decision, trellis quantization, and rate control among others, as described in clause 7.3 in [1][2].

#### Example of encoder setting

As observed in [21], 1) the Y' component in Y′CbCr and I component in ICTCP have very similar compression characteristics in terms of variance; this results in similar coding bits given same luma quantizer. It suggests that techniques designed to improve coding efficiency of Y′ component (e.g., perceptual luma quantization in [1][2]) can be used for I component directly. 2) for colour components, CT and CP have higher variance level than Cb and Cr, thus modifications have to be applied for technologies designed to improve chroma coding efficiency, such as the adaptive chroma QP offset model in [1][2].

In the following example on how to set chromaQP offset for ICTCP, it is assumed that the colour primaries of the mastering display/capture device are known.

Based on this knowledge, the following model is used to assign QP offsets for Cb, corresponding to CT, and Cr, corresponding to CP, based on the luma QP and a factor dependent on the capture and representation colour primaries. The model is expressed as:

(6‑22)

(6‑23)

where ccb = 1 and CbOffset = 6 if the capture colour primaries are the same as the representation colour primaries, ccb= 0.5 and CbOffset = 8 if the capture colour primaries are equal to the P3D65 primaries and the representation colour primaries are equal to the Rec. ITU-R BT.2020 primaries, and ccb=0.9 and CbOffset = 6 if the capture colour primaries are equal to the Rec. ITU-R BT.709 primaries and the representation primaries are equal to the Rec. ITU-R BT.2020 primaries.

Similarly, ccr = 1 and CrOffset = 6 if the capture colour primaries are the same as the representation colour primaries, ccr=0.9 and CrOffset = 7 if the capture colour primaries are equal to the P3D65 primaries and the representation colour primaries are equal to the Rec. ITU-R BT.2020 primaries, and ccr= 1.6 and CrOffset = 6 if the capture colour primaries are equal to the Rec. ITU-R BT.709 primaries and the representation primaries are equal to the Rec. ITU-R BT.2020 primaries.

Finally, and .

#### HEVC encoding

When creating the HEVC bitstream it is suggested to set syntax elements to the values listed in Table 6‑3 in the sequence parameter set (SPS) of the bitstream. The syntax elements in Table 6‑3 are conveyed in the Video Usability Information syntax branch of the SPS defined in Annex E of the HEVC specification. They may also be duplicated and carried in various application-layer headers.

Table 6‑3 – Suggested settings for HEVC encoding of ICTCP 4:2:0 PQ 10-bit signal.

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 |
| transfer\_characteristics | vui\_parameters( ) | 16 |
| matrix\_coeffs | vui\_parameters( ) | 14 |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

For HDR/WCG content represented with the colour primaries of Rec. ITU-R BT.2020, Rec. ITU-R BT.2100 PQ transfer function and ICTCP colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec. ITU-R BT.709 colour primaries and Rec. ITU-R BT.709 OETF (Rec. ITU-R BT.1886 EOTF) transfer function. Chroma QP adjustment, as described in clause 6.3.2.1 can be performed by adjusting and controlling the HEVC syntax elements pps\_cb\_qp\_offset, slice\_cb\_qp\_offset, pps\_cr\_qp\_offset and slice\_cr\_qp\_offset. Similarly, perceptual luma quantization as discussed in clause 6.3.2.1 could be achieved by adjusting the syntax elements cu\_qp\_delta\_abs and cu\_qp\_delta\_sign\_flag.

#### AVC encoding

When creating the AVC bitstream it is suggested to set syntax elements to the values listed in Table 6‑4 in the SPS of the bitstream. The syntax elements in Table 6‑4 are conveyed in the Video Usability Information syntax branch of the SPS defined in Annex E of the AVC specification. They may also be duplicated and carried in various application-layer headers.

Table 6‑4 – Suggested settings for AVC encoding of ICTCP 4:2:0 PQ 10-bit signal.

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value** |
| profile\_idc | seq\_parameter\_set\_data( ) | 110 |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_data( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 |
| transfer\_characteristics | vui\_parameters( ) | 16 |
| matrix\_coefficients | vui\_parameters( ) | 14 |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

For HDR/WCG content represented with the colour primaries of Rec. ITU-R BT.2020, Rec. ITU-R BT.2100 PQ transfer function and ICTCP colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec. ITU-R BT.709 colour primaries and Rec. ITU-R BT.709 transfer function. Chroma QP adjustment, as described in clause 6.3.2.1 can be performed by adjusting and controlling AVC syntax elements chroma\_qp\_index\_offset and second\_chroma\_qp\_index\_offset. Similarly, perceptual luma quantization as discussed in clause 6.3.2.1 could be achieved by adjusting the syntax element mb\_qp\_delta.

### Decoding process

When the bitstream is an HEVC bitstream, the decoding process as in the HEVC specification [8][9] is performed.

When the bitstream is an AVC bitstream, the decoding process in the AVC specification [5][6] is performed.

NOTE – The decoding process for HDR/WCG video is not different from the decoding process of SDR video.

### Post-decoding process

The post-decoding stage includes the following components:

1. a chroma up-conversion component that converts data from 4:2:0 to 4:4:4, following the same process described in sub-clause 9.1.2 in [1][2],
2. a conversion component that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, following the same process described in sub-clause 9.1.1 in [1][2],
3. a colour format conversion component that converts data from the ICTCP representation back to the non-linear L′M′S′ representation, described in sub-clause 6.3.4.1,
4. a conversion component from the non-linear L′M′S′ data representation back to a linear LMS data representation, described in sub-clause 6.3.4.2,
5. a conversion component from a linear LMS data representation to a linear RGB data representation, described in sub-clause 6.3.4.3.



Figure 6‑5 – Conventional post-decoding process system diagram.

#### Colour representation conversion: ICTCP to L′M′S′

Conversion from the ICTCP representation back to the L′M′S′ representation can be performed using a 3x3 matrix conversion process.

(6‑24)

#### Conversion from a non-linear to a linear light representation: L′M′S′ to LMS

This PQ EOTF (Equation (6‑7)) is applied to all L′, M′, and S′ non-linear representations, resulting in their linear light counterparts L, M, and S as follows:

L = EOTFPQ( L′ ) (6‑25)

M = EOTFPQ( M′ ) (6‑26)

S = EOTFPQ( S′ ) (6‑27)

#### Colour representation conversion: LMS to RGB

Conversion from the LMS representation back to the RGB representation can be performed using a 3x3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100 and each component, L, M, S, R, G, and B, is a number between 0.0 (representing no light) and 1.0 (representing 10 000 cd/m2).

(6‑28)

# Bitstream SDR backward compatibility with single-layer coding

This clause addresses the use cases of bitstream SDR compatibility. Two approaches are described: Approach 1 which uses a static HDR opto-electrical transfer function (OETF) to deliver “native” bitstream SDR backward compatibility; and Approach 2 which applies an “adaptive” HDR-to-SDR conversion process prior to the encoding, with dynamic metadata required to perform the inverse conversion after decoding. Approach 1 can be addressed with HEVC or AVC using the Rec. BT.2100 HLG transfer function. Approach 2 can be addressed with HEVC using the CRI or TMI SEI messages to convey the dynamic metadata.

## Approach 1: usage of HLG for "static" bitstream SDR backward compatibility

### Introduction

The HLG transfer function [3] has been designed to provide some level of backward compatibility with the legacy SDR systems. The legacy system is taken to be the SDR systems not needing colour gamut conversion (for example, not needing conversion from BT.2020 to BT.709 colour primaries) but needing only dynamic range conversion. HLG is suggested to be used with BT.2020 colour primaries.

Rec. ITU-R BT.2100 [3] specifies two transfer functions, PQ and HLG. The HLG transfer function provides a scene referred signal, while the PQ transfer functions provides a display referred signal. In the scene referred approach the HDR signal represents the light in the scene (with or without artistic adjustment), for example detected by a camera, while a display referred signal represents the light emitted by a reference or grading display (with or without artistic adjustment). Since the overall optical system is non-linear (typically characterized as a power or “gamma” law) scene and display referred signals are not linearly related. This non-linearity is modelled by the Opto-Optic Transfer Function (OOTF, [3][22]) which compensates for the changes in the tonal response of the human visual system as the eye adapts. With scene referred signals the primary transfer function is the OETF, which defines how the signal is related to linear scene light. For a display referred signal the primary transfer function is the EOTF which defines how the signal should be converted to be displayed on a reference or grading monitor. Display referred signals may not be used directly in scene referred systems, even if both correspond to “linear light”, without distorting the pictures. More details about the differences between scene and display referred signals can be found in Annex A of Rec. ITU-R BT.2100 [3].

NOTE – The conversion and coding workflow specifically described in [1][2] for PQ NCL Y′CbCr 4:2:0 10-bits signal can be used for HLG substituting PQ with HLG. The workflow described in Figure 7‑1 and Figure 7‑2 of clauses 7.1.2 and 7.1.4 is to highlight the conversion and coding differences between PQand HLG signals.

### HLG pre-encoding conversion process

HLG pre-encoding conversion process is illustrated in Figure 7‑1. The conversion chain converts an input linear RGB 4:4:4 signal to HLG 10-bits Y′CbCr 4:2:0 signal. It consists of the following successive steps:

1. a pre-conversion step (denoted “inverse system gamma” in Figure 7‑1 and described in clause 7.1.2.2) when the input HDR signal is display-light referred; when the input HDR signal is scene-light referred (e.g. from camera), this step is not required,
2. a conversion from an input scene referred linear RGB 4:4:4 data representation to a non-linear data representation R′G′B′ using the HLG OETF as described in sub-clause 7.1.2.1,
3. a colour format conversion from non-linear HLG R′G′B′ 4:4:4 signal to Y′CbCr 4:4:4 following the same process described in clause 6.1.2 in [1][2],
4. a conversion step that converts a floating-point to a fixed-point representation (e.g. 10 bits), narrow range, following the same process described in sub-clause 6.1.4 in [1][2],
5. a chroma down-conversion that converts data from 4:4:4 to 4:2:0, following the same process described in sub-clause 6.1.3 in [1][2].



Figure 7‑1 – Pre-processing for conversion with HLG.

#### HLG conversion from a linear RGB to a non-linear representation R′G′B′

For a scene referred linear light input, the HLG OETF as specified in equations (6-8) from sub-clause 6.1.2 is applied on the R, G and B components separately. When the input HDR video is display referred linear light, a preliminary conversion process (“inverse system gamma”) should be applied to convert the display referred video to a scene referred signal. This step removes the inherent OOTF within the display referred signal. Sub-clause 7.1.2.2 describes this process.

#### Conversion to linear light display referred to linear light scene referred

This step aims to remove the OOTF inherent in display referred signals. The process input is an RGB 4:4:4 display referred linear light signal considered to have been produced on a mastering display having a peak luminance equal to P cd/m2. The output is a scene referred linear light signal. The process applies the following steps to each R, G, B samples:

* Let Rdisplay, Gdisplay, Bdisplay be the display referred R, G, B values. For X = R, G, B, the following applies:

Xn = Max( 0, Xdisplay ÷ P ) (7‑3)

* The luminance Ydisplay is derived from Rn, Gn, Bn values:

Ydisplay = a0 \* Rn + a1 \* Gn + a2 \* Bn (7‑4)

where a0 = 0.2627, b0 = 0.6780 and c0 = 0.0593 are the conventional RGB-to-Y derivation coefficients for BT.2020 colour gamut container.

* The value of the system gamma g is estimated from the peak luminance P as follows:

g = 1.2 + 0.42 \* Log10( P ÷ 1000 ) (7‑5)

* The value gscale is derived as follows

(7‑6)

* The output mapped samples Rscene, Gscene, Bscene samples are eventually obtained by scaling Rn, Gn, Bn. For X = R, G or B, the following applies:

(7‑7)

NOTE – Transcoding between PQ and HLG signals is specified in Section 7 of ITU-R report BT.2390 [5]. When the peak brightness of the HLG and PQ displays are the same, the original and transcoded signals will look identical. Typically, however, for PQ content the brightness of low-lights and mid-tones remains the same, regardless of the peak brightness of the display. Brighter PQ displays offer increased headroom for specular highlights, but the overall image brightness remains unchanged. HLG, however, is based on relative brightness. Thus, as the display’s peak brightness increases so does the brightness of the entire image. The headroom for specular highlights is a constant number of stops, but the brighter image makes it suitable for viewing in brighter environments. When, for example, a PQ signal with 4000 cd/m2 peak luminance is transcoded to HLG using the method specified in ITU-R BT.2390, whilst it will appear identical on a 4000 cd/m2 HLG display, it may appear darker when shown on a 1000 cd/m2 HLG display in brighter environments. As the PQ and HLG systems have different characteristics, a format conversion rather than a simple transcode is required. Such format conversions are outside of the scope of this document.

### Encoding process

After pre-processing, the data is ready for compression. The quantization methodologies, perceptual luma quantization, in clause 7.1, and chroma QP offset in clause 7.2, presented in [1][2], can be used in case of HLG signal. The exact parameter settings for these two quantization methods might need to be reconfigured.

To allow the HEVC or AVC encoded bitstream to be viewed on an SDR display, the resulting Y′CbCr 4:2:0 10-bits signal can be signalled as being an SDR signal. In both cases, this is done by indicating in VUI a transfer\_characteristics value corresponding to an SDR transfer function, e.g. BT.2020 with code value 14. For HDR HLG compliant displays, the actual HLG transfer function can be indicated by means of the ATC SEI message, as described in clause 7.1.6.

**Table 7‑1 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 HLG 10-bits signal with SDR compatibility.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested  value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 |
| transfer\_characteristics | vui\_parameters( ) | 14 |
| matrix\_coeffs | vui\_parameters( ) | 9 |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

**Table 7‑2 – Suggested settings for AVC encoding of Y′CbCr 4:2:0 HLG 10-bits signal with SDR compatibility.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested**  **value** |
| profile\_idc | seq\_parameter\_set\_data( ) | 110 |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_data( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 |
| transfer\_characteristics | vui\_parameters( ) | 14 |
| matrix\_coefficients | vui\_parameters( ) | 9 |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

NOTE – For HDR-only systems, with no SDR backward compatibility, the VUI transfer\_characteristics element should indicate that the transfer function corresponds to HLG (transfer\_characteristics set equal to 18 in both HEVC and AVC cases).

### Decoding process

When the bitstream is an HEVC bitstream the decoding process as in the HEVC specification [8][9] is performed.

When the bitstream is an AVC bitstream the decoding process in the AVC specification [6][7] is performed.

NOTE – The decoding process for HDR/WCG video is not different from the decoding process of SDR video.

### HLG post-decoding conversion

HLG post-decoding inverse conversion process is illustrated in Figure 7‑2. The inverse conversion chain consists of the following successive steps:

1. a chroma up-conversion that converts data from Y′CbCr 4:2:0 to Y′CbCr 4:4:4, following the same process described in sub-clause 9.1.2 in [1][2],
2. a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, following the same process described in sub-clause 9.1.1 in [1][2],
3. a colour representation conversion from Y′CbCr 4:4:4 to R′G′B′ 4:4:4, following the same process described in sub-clause 9.1.3 in [1][2],
4. a conversion using the HLG EOTF of the input R′G′B′ 4:4:4; the HLG EOTF is the concatenation of the inverse OETF and of the OOTF; this last step aims at adapting this content to the rendering display, by applying a suitable end-to-end transfer function of the whole system (OOTF); this transfer function is dependent on the viewing environment (peak brightness of the display, brightness of the surround).



Figure 7‑2 – Post-processing for conversion with HLG.

### Signalling of alternative transfer characteristics SEI message

HEVC signalling allows the delivery of HLG HDR video signals to both HDR capable receivers and legacy SDR/BT.2020 ones. This is achieved by setting the VUI parameters and using the ATC SEI message as described below.

Settings in VUI for SDR rendering:

* transfer\_characteristics 14 (BT.2020) to satisfy legacy SDR receivers
* colour\_primaries 9 (BT.2020)
* matrix\_coeffs 9 (BT.2020 NCL)
* video\_full\_range\_flag 0 (narrow range)

Settings in ATC SEI message to override VUI transfer\_characteristics for HDR HLG transfer function signalling:

* preferred\_transfer\_characteristics 18

One ATC SEI message should be conveyed with each Random Access Point Access Unit (RAPAU) picture.

## Approach 2: “dynamic” bitstream SDR backward compatibility with dynamic metadata signalled in SEI messages

This clause is focused on adaptive bitstream SDR backward compatibility support using a single-layer coding framework. Figure 7‑3 provides an overview of a single-layer distribution system enabling bitstream SDR backward compatibility. The system typically uses HEVC Main 10 profile for the bitstream generation and decoding. It includes a pre-processing block, prior to encoding, that converts an input HDR signal into an SDR version. Metadata can be generated in this step. After encoding and decoding the SDR signal, such metadata can be used in a post-processing step to reconstruct an HDR version of the signal. The decoded SDR video can be directly rendered on an SDR display without adaptation.



Figure 7‑3 – Bitstream SDR backward compatibility using single layer coding system.

Different approaches to use HEVC SEI messages for this task are available; three are illustrated. Two use the CRI SEI message, and operates in the Y′CbCr 4:2:0 and 4:4:4 domains. Another one uses the TMI SEI message, and operates in the R′G′B′ 4:4:4 domain. Other configurations are also possible.

Examples of conversion and encoding pre-processing producing SDR/BT.2020 compatible HEVC bitstream are illustrated in Figure 7‑4 (applying in Y′CbCr 4:2:0 domain and using CRI SEI message), Figure 7‑5 (applying in Y′CbCr 4:4:4 domain and using CRI SEI message), and Figure 7‑6 (applying in R′G′B′ 4:4:4 domain and using TMI SEI message).

In addition, it should be noted that the bitstream SDR backward compatibility functionality can also be addressed using recommendation ETSI TS 103 433 [10]. This recommendation specifies dynamic metadata and related processes to perform SDR-to-HDR conversion as a post-processing in a bitstream SDR backward compatible HDR video distribution system. In brief, the SDR-to-HDR conversion process, applied to an input Y′CbCr SDR signal, is based on an inverse luma mapping function applied to the SDR luma component, and on a scaling of the SDR chroma components using a luma-dependent scaling function. The luma mapping and luma-dependent chroma scaling functions are modelled and conveyed as dynamic metadata. Two metadata formats are specified, either using a compact parametric modeling defined in a user-data registered SEI message, or using piece-wise linear representations as in a CRI SEI message.

### CRI applied in Y′CbCr 4:2:0 domain

In the first design, the pre-encoding conversion process (top of Figure 7‑4) converts an input linear RGB 4:4:4 signal to SDR 10-bits Y′CbCr 4:2:0 signal by applying the following successive steps:

1. a conversion from an input linear RGB 4:4:4 representation to a non-linear representation using the inverse PQ EOTF, following the same process described in sub-clause 6.1.1 in [1][2],
2. a colour format conversion from non-linear PQ R′G′B′ 4:4:4 signal to Y′CbCr 4:4:4 following the same process described in sub-clause 6.1.2 in [1][2],
3. a conversion step that converts a floating-point to a fixed-point representation (i.e. 10 bits), narrow range, following the same process described in sub-clause 6.1.4 in [1][2],
4. a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, following the same process described in sub-clause 6.1.3 in [1][2], resulting in a PQ 10-bits 4:2:0 Y′CbCr signal (PQ10),
5. a dynamic range adaptation (DRA) step that applies three different transfer functions to the 4:2:0 Y′, Cb, and Cr components of the PQ10 signal to generate a 10-bits SDR 4:2:0 Y′CbCr signal.

The resulting Y′CbCr signal, having BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries, is then encoded, using an HEVC Main 10 compliant encoder. The DRA transfer functions ca n be implemented in the shape of 1D-LUTs that directly apply to the PQ10 Y′, Cb and Cr components, in 4:2:0 format.

After HEVC Main 10 compliant decoding, the decoded 4:2:0 Y′CbCr signal has BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries. The post-decoding inverse conversion processing (bottom of Figure 7‑4) is the inverse of the pre-encoding processing. It is made of the following steps:

1. an inverse DRA process, converting the SDR 10-bits Y′CbCr 4:2:0 signal into a PQ10 compatible signal using the inverse DRA transfer functions,
2. a chroma up-conversion that converts data from Y′CbCr 4:2:0 to Y′CbCr 4:4:4, following the same process described in sub-clause 9.1.2 in [1][2],
3. a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, following the same process described in sub-clause 9.1.1 in [1][2],
4. a colour representation conversion from Y′CbCr 4:4:4 to R′G′B′ 4:4:4, following the same process described in sub-clause 9.1.3 in [1][2],
5. a conversion using the PQ EOTF from the input R′G′B′ 4:4:4 to linear RGB 4:4:4, following the same process described in sub-clause 9.1.4 in [1][2].



Figure 7‑4 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2020 with CRI applied in Y′CbCr 4:2:0 domain.

### CRI applied in Y′CbCr 4:4:4 domain

The second design (Figure 7‑5), similarly to the first one, makes use of the CRI SEI message to perform the inverse DRA process. However, DRA and inverse DRA apply in Y′CbCr 4:4:4 domain, instead of Y′CbCr 4:2:0 domain.

Pre-encoding processing (top block-diagram) applies the following successive steps:

1. a conversion step from an input linear RGB 4:4:4 representation to a non-linear representation using the inverse PQ EOTF, following the same process described in sub-clause 6.1.1 in [1][2],
2. a colour format conversion from non-linear PQ R′G′B′ 4:4:4 signal to Y′CbCr 4:4:4 following the same process described in sub-clause 6.1.2 in [1][2],
3. a conversion step that converts a floating-point to a fixed-point representation (i.e. 10 bits), narrow range, following the same process described in sub-clause 6.1.4 in [1][2],
4. a dynamic range adaptation (DRA) step that applies to the 4:4:4 Y′, Cb, and Cr components of the PQ10 signal to generate a 10-bits SDR Y′CbCr 4:4:4 signal,
5. a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, following the same process described in sub-clause 6.1.3 in [1][2], resulting in a 10-bits SDR 4:2:0 Y′CbCr signal.

Post-decoding processing (bottom block-diagram) applies the following steps:

1. a chroma up-conversion component that converts data from Y′CbCr 4:2:0 to Y′CbCr 4:4:4, following the same process described in sub-clause 9.1.2 in [1][2],
2. an inverse DRA process, converting the SDR Y′CbCr 4:4:4 signal into a PQ 10-bits Y′CbCr 4:4:4 compatible signal using the metadata conveyed in a CRI SEI message,
3. a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, following the same process described in sub-clause 9.1.1 in [1][2],
4. a colour representation conversion from Y′CbCr 4:4:4 to R′G′B′ 4:4:4, following the same process described in sub-clause 9.1.3 in [1][2],
5. a conversion using the PQ EOTF from the input R′G′B′ 4:4:4 to linear RGB 4:4:4, following the same process described in sub-clause 9.1.4 in [1][2].

Working in Y′CbCr 4:4:4 domain enables using the complete CRI model, that is, three Pre-LUTs, followed by a three-by-three matrix, followed by three Post-LUTs. This complete CRI model may be also preferably used for enabling an SDR backward compatibility to SDR BT.709 colour primaries while the input HDR colour primaries are BT.2020. In this case the three-by-three matrix can be of use for converting from BT.2020 colour gamut to BT.709 colour gamut.

NOTE – In case on non-backward compatible applications, the CRI SEI message can also be used to work in a colour space other than Y′CbCr decoding process for HDR/WCG video.



Figure 7‑5 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2020 with CRI applied in Y′CbCr 4:4:4 domain.

### TMI applied in R′G′B′ 4:4:4 domain

In the third design (Figure 7‑6), the TMI SEI message is used to perform the inverse DRA and it applies to a R′G′B′ 4:4:4 signal.

Pre-encoding processing (top block-diagram) applies the following successive steps:

1. a conversion step from an input linear RGB 4:4:4 representation to a non-linear representation using the inverse PQ EOTF, following the same process described in sub-clause 6.1.1 in [1][2],
2. a conversion step that converts a floating-point to a fixed-point M-bits representation (typically M being set to 14 or 16 bits), following a similar process as described in sub-clause 6.1.4 in [1][2],
3. a dynamic range adaptation (DRA) step applying to the M-bits PQ R′G′B′ 4:4:4 signal to produce an SDR R′G′B′ 4:4:4 signal (having BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries),
4. a colour format conversion from non-linear R′G′B′ 4:4:4 signal to Y′CbCr 4:4:4 following the same process described in sub-clause 6.1.2 in [1][2],
5. a chroma down-conversion that converts data from 4:4:4 to 4:2:0, following a similar process as described in sub-clause 6.1.3 in [1][2], resulting in an SDR 4:2:0 Y′CbCr signal.
6. a re-quantization to 10 bits.

Post-decoding processing (bottom block-diagram) applies the following steps:

1. a re-quantization to M-bits of the input decoded signal having BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries,
2. a chroma up-conversion that converts data from Y′CbCr 4:2:0 to Y′CbCr 4:4:4, following a similar process as described in sub-clause 9.1.2 in [1][2],
3. a colour representation conversion from Y′CbCr 4:4:4 to R′G′B′ 4:4:4, following the same process described in sub-clause 9.1.3 in [1][2],
4. an inverse DRA process, converting the SDR R′G′B′ 4:4:4 signal into a PQ M-bits R′G′B′ 4:4:4 compatible signal,
5. a conversion component that converts a M-bits fixed-point representation, to a floating-point representation, following the same process described in sub-clause 9.1.1 in [1][2],
6. a conversion using the PQ EOTF from the input R′G′B′ 4:4:4 to linear RGB 4:4:4, following the same process described in sub-clause 9.1.4 in [1][2].

The inverse DRA transfer function is signalled in a TMI SEI message. The function is modelled using one of the four model types available for the TMI SEI message, for example the piece-wise linear function. It is noted that one of the pre- or post-LUTs of the CRI SEI message can also be used to model this function.



Figure 7‑6 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2020 with TMI applied in RGB 4:4:4 domain.

In these three examples presented in sub-clauses 7.2.1, 7.2.2 and 7.2.3, the DRA functions can be derived from the analysis of input HDR signal properties, with the aim of producing an SDR approximation. Parameters of DRA can be generally derived from the HDR graded signal by using an HDR-to-SDR conversion algorithm, or directly from a graded SDR signal in case an SDR master is provided as input to the encoding system.

It is recommended to generate at least one CRI or TMI SEI message per scene. In scenes with large temporal changes, it can be required to generate several CRI or TMI SEI messages per scene (up to one per picture). At least, one CRI or TMI SEI message should be conveyed with each Random Access Point Access Unit (RAPAU) picture.

### Settings with colour remapping information SEI message

For the adaptive bitstream SDR backward compatibility use case described in clause 7.2, with the implementation shown in Figure 7‑4, the inverse DRA functions are coded using the three Pre-LUTs of a CRI SEI message. The conversion process directly applies to the decoded Y′CbCr 4:2:0 signal (Figure 7‑4), and the three-by-three matrix and the Post-LUTs are not activated. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma sample Y′SDR and each chroma samples USDR and VSDR as follows:

(7‑8)

(7‑9)

(7‑10)

When the conversion process applies to the decoded Y′CbCr 4:2:0 signal upsampled to 4:4:4 (implementation shown in Figure 7‑5), the full CRI model, made of the three Pre-LUTs, the intermediate three-by-three matrix, and the three Post-LUTs, can be activated. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma and chroma samples Y′SDR, USDR and VSDR as follows:

(7‑11)

(7‑12)

(7‑13)

Table 7‑3 indicates suggested parameters settings of HEVC Main 10 bitstreams to correctly render the decoded video on displays compatible with SDR/BT.2020 representation and to conduct the HDR reconstruction to PQ10 compatible representation.

**Table 7‑3 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 10-bits signal with bitstream SDR compatibility using CRI.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 (narrow range) |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 1 (BT.709) or 9 (BT.2020) depending on SDR target |
| transfer\_characteristics | vui\_parameters( ) | 14 (SDR transfer function) |
| matrix\_coeffs | vui\_parameters( ) | 1 (BT.709 NCL) or 9 (BT.2020 NCL) depending on SDR target |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |
| colour\_remap\_video\_signal\_info\_present\_flag | colour\_remapping\_info( ) | 1 |
| colour\_remap\_transfer\_function | colour\_remapping\_info( ) | 16 (BT.2100 PQ) |
| colour\_remap\_full\_range\_flag | colour\_remapping\_info( ) | 0 |
| colour\_remap\_primaries | colour\_remapping\_info( ) | 9 (BT.2020) |
| colour\_remap\_matrix\_coefficients | colour\_remapping\_info( ) | 9 (BT.2020 NCL) |
| colour\_remap\_output\_bit\_depth | colour\_remapping\_info( ) | 10 |

### Settings with tone mapping information SEI message

For the use case described in clause 7.2, with the implementation example shown in Figure 7‑6, the inverse DRA function is coded using one 1D-LUT of a TMI SEI message. Two implementation examples of the inverse DRA processes are provided below.

In a first implementation example, the SDR-to-HDR conversion achieved at post-processing stage applies to the SDR 4:4:4 R′G′B′ M-bits samples as follows:

(7‑14)

In a second implementation example, the tone mapping function is applied to the SDR luma component Y′, to derive a scaling ratio to be applied to the SDR 4:4:4 R′G′B′ M-bits samples as follows:

(7‑15)

For each sample value X = R, G or B:

(7‑16)

where M is the bit-depth of the output samples resulting from the inverse DRA application (corresponding to the TMI SEI message syntax element target\_bit\_depth). This second implementation can be beneficial to limit the colour hue shift resulting from the SDR-to-HDR conversion (see [24] for a fuller exposition).

Table 7‑4 indicates suggested parameters settings of HEVC Main 10 bitstreams to correctly render the decoded video on displays compatible with SDR/BT.2020 representation and to conduct the HDR reconstruction to PQ10 compatible representation.

**Table 7‑4 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 10-bits signal with bitstream SDR compatibility using TMI.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 (narrow range) |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 |
| transfer\_characteristics | vui\_parameters( ) | 14 (SDR transfer function) |
| matrix\_coeffs | vui\_parameters( ) | 9 (BT.2020 NCL) |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |
| coded\_data\_bit\_depth | tone\_mapping\_info ( ) | M (typically 14 bits) |
| target\_bit\_depth | tone\_mapping\_info ( ) | M (typically 14 bits) |
| tone\_map\_model\_id | tone\_mapping\_info ( ) | 3 (piece-wise linear model) |
| num\_pivots | tone\_mapping\_info ( ) | 32 |

# Bitstream SDR backward compatibility with dual-layer SHVC coding

## Introduction

SHVC, the scalable form of HEVC [8][9][17], is based on a multi-layer framework. Typically, a base layer and an enhancement layer are both coded. The enhancement layer coding uses inter-layer prediction mechanisms that produce an inter-layer picture used as reference picture. HEVC Scalable Main 10 profile includes colour mapping prediction specifically designed for inter-layer prediction from a given colour volume to another one. In an SDR-backward compatible system, this profile can be used to convey in a dual layer-bitstream an SDR video and an HDR video. The input of the coding system is an SDR content as base layer, and its HDR version as enhancement layer. The SDR version can result from an HDR-to-SDR conversion such as the one mentioned in clause 7.2, or can be provided as an input SDR master.



Figure 8‑1 – Example of SDR backward compatible solution implemented with SHVC.

## Encoding and decoding stages

When creating the dual-layer SHVC bitstream it is suggested to set syntax elements in the base layer and enhancement layer sequence parameter sets (SPS) of the bitstream to the values listed in Table 8‑1 (SDR base layer) and Table 8‑2 (HDR enhancement layer). The syntax elements in Table 8‑1 and Table 8‑2 are conveyed in the Video Usability Information syntax branch of the SPS defined in Annex E of the HEVC specification. They may also be duplicated and carried in various application-layer headers.

**Table 8‑1 – Suggested settings for base layer (Y′CbCr 4:2:0 SDR 10-bits) in a dual-layer SHVC bitstream.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 (narrow range) |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 1 (BT.709) or 9 (BT.2020) depending on SDR input |
| transfer\_characteristics | vui\_parameters( ) | 14 (SDR transfer function) |
| matrix\_coeffs | vui\_parameters( ) | 1 (BT.709 NCL) or 9 (BT.2020 NCL) depending on SDR input |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

**Table 8‑2 – Suggested settings for enhancement layer (Y′CbCr 4:2:0 PQ 10-bits) in a dual-layer SHVC bitstream.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 7 (Scalable Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 (narrow range) |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 (BT.2020) |
| transfer\_characteristics | vui\_parameters( ) | 16 (BT.2100 PQ) |
| matrix\_coeffs | vui\_parameters( ) | 9 (BT.2020 NCL) |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |

The SDR video is coded as the base layer of SHVC encoding conforming to the Main 10 profile, and the HDR video is coded as the enhancement layer conforming to the Scalable Main 10 profile. When the colour containers in which the SDR and HDR are represented are different, SHVC colour gamut scalability (CGS) tool [18] can apply to encode the content with good coding efficiency [18][19]. Even if the colour containers in which the SDR and HDR are coded are the same, the colour volume occupied by the SDR and HDR pictures is typically different. Even in such cases, the CGS tool can encode the video with good coding efficiency.

The dual-layer bitstream is then decoded using an SHVC decoder, with the output from the base layer resulting in an SDR video representation and the enhancement layer resulting in a HDR video representation.

# Display adaptation

## Introduction

Display adaptation aims at converting a decoded HDR video (typically PQ10) to a version adapted to the target rendering device. For instance, it may be used to convert a PQ10 video mastered on a display with a given peak luminance, to a display capable of displaying a lower peak luminance. The concept is illustrated in Figure 9‑1. The conversion process can be driven by metadata produced during the mastering stage, and conveyed with the main bitstream. CRI or TMI SEI messages can be used for this purpose.

Display SDR backward compatibility, using CRI or TMI SEI messages, is a specific example of display adaptation, when the rendering device has only SDR display capabilities. This specific case is further described in clause 9.2. More generally, the adaptation of an HDR content with a given colour volume, to a version with a different target colour volume, can be achieved using CRI or TMI SEI messages. If multiple target colour volumes are considered, multiple instances of CRI or TMI SEI messages would be required. Identifier tags signalled in these SEI messages can be used to identify the different target colour volumes.



Figure 9‑1 – Illustration of the concept of display adaptation.

It should be noted that the display adaptation functionality is also addressed by SMPTE ST 2094 [14]. SMPTE ST 2094 defines metadata for use in colour volume transforms of content. The metadata are content-dependent and can vary scene-by-scene or frame-by-frame. The metadata are intended to transform HDR/WCG source content for presentation on a display having a smaller colour volume than the source content’s mastering display. Multiple applications provide particular colour volume transforms. Currently there are four applications supported in ST 2094.

A metadata set in ST 2094 incorporates a time interval, a window, metadata describing the targeted system display and parameters controlling the colour volume transform. A receiver of the metadata can use the targeted system display metadata to select the metadata sets that are most applicable to the actual output device. A receiver of the metadata can use the targeted system display metadata to adjust the output of the colour volume transform from the specified output device to the actual output device. Two different approaches have been identified as transform methods in ST 2094 applications: parametric mapping and reference based numerical derivation. Application #1, #2 and #4 belong to the parametric model, and #3 belongs to the numerical data-fitting model. Part of the syntax in Application #3 is supported in colour remapping information SEI message.

Display adaptation for HLG systems is specified as part of the display EOTF in BT.2100.

## Case of display SDR backward compatibility

### Conversion and coding process example

Dynamic mapping from HDR to SDR is the reverse problem of bitstream SDR backward compatibility studied in clause 7.2. In the present document, this use case is named display SDR backward compatibility.

The system design can be straightforwardly deduced from Figure 7‑4 and Figure 7‑5 by placing a forward DRA process to the decoder side and signalling DRA parameters through SEI message such as CRI.

Two examples of encoding process producing HDR/BT.2020 compatible HEVC bitstream, which include DRA control parameters for decoder side guided conversion from HDR to SDR, are illustrated in top block-diagrams of Figure 9‑2 and Figure 9‑3. These parameters can be conveyed to the decoder in CRI SEI messages, and used for performing the HDR-to-SDR conversion at decoder side. A corresponding decoding process, which enables to conduct an optional mapping from HDR to SDR through forward DRA process, is illustrated in bottom block-diagrams of Figure 9‑2 and Figure 9‑3. After HEVC Main 10 compliant decoding, the decoded signal has BT.2100 PQ transfer characteristics and BT.2020 colour primaries. The DRA can be used as a post-decoding processing to convert the decoded HDR/BT.2020 video signal to an SDR representation. The resulting Y′CbCr 10-bits signal has BT.709/BT.2020 transfer characteristics and BT.2020 or BT.709 colour primaries (depending on the application use case).

In the first example (Figure 9‑2), DRA operates in Y′CbCr 4:2:0 domain. It typically applies three transfer functions to the Y′, Cb, and Cr components of the decoded PQ10 signal, for instance using the three Pre-LUTs of a CRI SEI message. The functions may be implemented in the shape of 1D-LUTs (LUTDRAk, for *k*=0, 1, 2), that directly apply to the PQ10 Y′, Cb and Cr components. The 3x3 matrix and Post-LUTs are not activated.



Figure 9‑2 – Example of pre-processing (top) and post-processing (down) for display SDR backward compatibility with CRI applied in Y′CbCr 4:2:0 domain.

In the second example (Figure 9‑3), DRA operates after the conversion of the decoded Y′CbCr 4:2:0 10-bits PQ BT.2020 video to 4:4:4. The complete CRI model may be used. This may be relevant when the target SDR colour primaries are BT.709. In that case, the three-by-three matrix can be beneficial to perform the conversion from BT.2020 to BT.709 primaries.



Figure 9‑3 – Example of pre-processing (top) and post-processing (down) for display SDR backward compatibility with CRI applied in Y′CbCr 4:4:4 domain.

### Using colour remapping information SEI message

In the implementation illustrated in Figure 9‑2, the DRA operates in Y′CbCr 4:2:0 domain. The three DRA functions can be implemented in 4:2:0 using the three Pre-LUTs of a CRI SEI message. The three-by-three matrix and the Post-LUTs are not activated. Therefore, the conversion process can directly apply to the decoded 4:2:0 signal. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma sample Y′HDR and each chroma sample UHDR and VHDR as follows:

(9‑1)

(9‑2)

(9‑3)

In the implementation illustrated in Figure 9‑3, the DRA operates in Y′CbCr 4:4:4 domain. The full CRI model can be used, involving three Pre-LUTs, three-by-three matrix, and three Post-LUTs. This may be applicable, when the use case requires an SDR backward compatibility to SDR BT.709 colour primaries while the input HDR colour primaries are BT.2020. The application of the matrix can be of use to help converting from one colour representation to the other one. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma and chroma samples Y′HDR, UHDR and VHDR as follows:

(9‑4)

(9‑5)

(9‑6)

Table 9‑1 indicates suggested parameters settings of HEVC Main 10 bitstreams to correctly display video on PQ10 compatible devices and to conduct a guided mapping from SDR to HDR with CRI post-processing.

**Table 9‑1 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 PQ 10-bits signal with display SDR compatibility using CRI.**

|  |  |  |
| --- | --- | --- |
| **Syntax element** | **Location** | **Suggested value in HEVC** |
| general\_profile\_space | profile\_tier\_level() | 0 |
| general\_profile\_idc | profile\_tier\_level() | 2 (Main 10) |
| vui\_parameters\_present\_flag | seq\_parameter\_set\_rbsp( ) | 1 |
| video\_signal\_type\_present\_flag | vui\_parameters( ) | 1 |
| video\_full\_range\_flag | vui\_parameters( ) | 0 (narrow range) |
| colour\_description\_present\_flag | vui\_parameters( ) | 1 |
| colour\_primaries | vui\_parameters( ) | 9 (BT.2020) |
| transfer\_characteristics | vui\_parameters( ) | 16 (BT.2100 PQ) |
| matrix\_coeffs | vui\_parameters( ) | 9 (BT.2020 NCL) |
| chroma\_loc\_info\_present\_flag | vui\_parameters( ) | 1 |
| chroma\_sample\_loc\_type\_top\_field | vui\_parameters( ) | 2 |
| chroma\_sample\_loc\_type\_bottom\_field | vui\_parameters( ) | 2 |
| colour\_remap\_video\_signal\_info\_present\_flag | colour\_remapping\_info( ) | 1 |
| colour\_remap\_transfer\_function | colour\_remapping\_info( ) | 14 (SDR transfer function) |
| colour\_remap\_full\_range\_flag | colour\_remapping\_info( ) | 0 |
| colour\_remap\_primaries | colour\_remapping\_info( ) | 1 (BT.709) or 9 (BT.2020) depending on target SDR |
| colour\_remap\_matrix\_coefficients | colour\_remapping\_info( ) | 1 (BT.709 NCL) or 9 (BT.2020 NCL) depending on target SDR |
| colour\_remap\_output\_bit\_depth | colour\_remapping\_info( ) | 10 |

# Bibliography

1. Supplement ITU-T HSTP-HDR10 (2017), *Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics*.
2. Technical Report ISO/IEC 23008-14 (2017), *Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics*.
3. Recommendation ITU-R BT.2100-0 (2016), *Image parameter values for high dynamic range television for use in production and international programme exchange*.
4. Recommendation ITU-R BT.2020-2 (2015), *Parameter values for ultra-high definition television systems for production and international programme exchange*.
5. Report ITU-R BT.2390-0 (2016), *High dynamic range television for production and international programme exchange*.
6. Recommendation ITU-T H.264 (V10) (2016), *Advanced video coding for generic audiovisual services*.
7. ISO/IEC 14496-10 (2014), *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*.
8. Recommendation ITU-T H.265 (V3) (2015), *High Efficiency Video Coding*.
9. ISO/IEC 23008-2 (2015), *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding.*
10. ETSI TS 103 433 (2016), *High-Performance Single Layer Directly Standard Dynamic Range (SDR) Compatible High Dynamic Range (HDR) System for use in Consumer Electronics devices (SL-HDR1)*.
11. ISO 11664-1 (2007), *Colorimetry – Part 1:* *CIE standard colorimetric observers*.
12. SMPTE ST 2084 (2014), *Electro-Optical Transfer Function for High Dynamic Range Reference Display*.
13. SMPTE ST 2086 (2014), *Mastering Display Color Volume Metadata for High Luminance and Wide Color Gamut Images*.
14. SMPTE ST 2094 (2016), *Dynamic Metadata for Color Volume Transform – Core Components*.
15. CEA-861.3 (2015), *HDR Static Metadata Extensions*, <https://standards.cta.tech/kwspub/published_docs/CEA-861.3-Preview.pdf>.
16. R. Joshi, S. Liu, G. J. Sullivan, Y.-K. Wang, J. Xu, Y. Ye (editors), “HEVC Screen Content Coding Draft Text 6”, 23th JCTVC meeting, San Diego, USA, February 2016.
17. J. M. Boyce, Y. Ye, A. K. Ramasubramonian (2015), “Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard”, *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 26, Issue 1, pp 20-34, July 2015.
18. P. Bordes, P. Andrivon, X. Li, Y. Ye (2016), “Overview of Color Gamut Scalability”, *IEEE Trans. Circuits and Systems for Video Technology*, Mar. 2016.
19. Y. Ye, V. Baroncini, Y.-K. Wang (2016), “Verification Test Report for Scalable HEVC (SHVC)”, JCTVC-W1004, 23rd JCT-VC meeting, San Diego, CA, USA, Feb. 2016*.*
20. V. Baroncini, W. Husak, A. K. Ramasubramonian, J. Sole, G. J. Sullivan, Y.-K. Wang, Y. Ye (2016), “SHVC Supplemental Verification Test Report”, JCTVC-X1004, 24th JCT-VC meeting, Geneva, CH, May 2016*.*
21. F. Pu, T. Lu, P. Yin, T. Chen, W. Husak (2016), “AHG13: ICTCP Compression Using HEVC Main 10”, JCTVC-X0050, 24th JCT-VC meeting, Geneva, CH, May 2016.
22. T. Borer, A. Cotton (2015), “A “Display Independent High Dynamic Range Television System”, BBC Research and Development white paper, WHP 309, Sep. 2015, <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP309.pdf>.
23. M. Zink, M. D. Smith (2015), “Managing HDR content production and display device capabilities”, *Proc. International Broadcasting Convention 2015 (IBC 2015)*, Amsterdam, NL, Sep. 2015,   
    <http://www.ibc.org/__media/PDF/michael_zink_managing_hdr_content_production_and_display_device_capabilities.pdf>.
24. P. Topiwala, W. Dai, M. Krishnan, “AHG14: Tone Mapping and Related SEIs for HDR Coding”, JCTVC-Y0042, 25th JCT-VC Meeting, Chengdu, CN, Oct. 2016.
25. F. Pu, T. Lu, P. Yin, T. Chen, W. Husak (2016), “AHG13: on Luma Adjustment”, JCTVC-X0043, 24th JCT-VC meeting, Geneva, CH, May 2016.