
**Information technology — High
efficiency coding and media delivery
in heterogeneous environments —**

**Part 15:
Signalling, backward compatibility and
display adaptation for HDR/WCG video**

*Technologies de l'information — Codage à haut rendement et
fourniture de supports dans les environnements hétérogènes —*

*Partie 15: Signalisation, compatibilité amont et adaptation de
l'affichage pour la vidéo HDR/WCG*

STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 23008-15:2018



STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 23008-15:2018



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	3
5 Conventions	4
5.1 General.....	4
5.2 Arithmetic operators.....	5
5.3 Bit-wise operators.....	5
5.4 Assignment operators.....	6
5.5 Relational, logical and other operators.....	6
5.6 Mathematical functions.....	6
5.7 Order of operations.....	7
6 Overview	8
7 HEVC signalling mechanisms applicable to HDR/WCG video	9
7.1 General.....	9
7.2 VUI syntax elements.....	9
7.3 SEI messages applicable for HDR/WCG video.....	10
7.3.1 General.....	10
7.3.2 Mastering display colour volume SEI message.....	10
7.3.3 Content light level information SEI message.....	11
7.3.4 Ambient viewing environment SEI message.....	11
7.3.5 Alternative transfer characteristics SEI message.....	11
7.3.6 Tone mapping information SEI message.....	11
7.3.7 Colour remapping information SEI message.....	12
7.4 Overview of PQ and HLG transfer functions.....	13
7.4.1 General.....	13
7.4.2 Reference PQ EOTF.....	14
7.4.3 Reference HLG OETF.....	15
7.5 IC _{TCP} colour representation.....	16
7.5.1 General.....	16
7.5.2 Pre-encoding process.....	17
7.5.3 Encoding process.....	20
7.5.4 Decoding process.....	22
7.5.5 Post-decoding process.....	22
8 Bitstream SDR backward compatibility with single-layer coding	24
8.1 General.....	24
8.2 Approach 1: usage of HLG for “static” bitstream SDR backward compatibility.....	24
8.2.1 General.....	24
8.2.2 HLG pre-encoding conversion process.....	25
8.2.3 Encoding process.....	27
8.2.4 Decoding process.....	29
8.2.5 HLG post-decoding conversion.....	29
8.2.6 Colour representation conversion: Y’CbCr to R’G’B’.....	30
8.3 Approach 2: usage of SEI messages for “dynamic” bitstream SDR backward compatibility.....	30
8.3.1 General.....	30
8.3.2 CRI applied in Y’CbCr 4:2:0 domain.....	31
8.3.3 CRI applied in Y’CbCr 4:4:4 domain.....	32
8.3.4 TMI applied in R’G’B’ 4:4:4 domain.....	33

8.3.5	Derivation of DRA functions	34
8.3.6	Settings with colour remapping information SEI message	35
8.3.7	Settings with tone mapping information SEI message	36
9	Bitstream SDR backward compatibility with dual-layer SHVC coding	37
9.1	General	37
9.2	Encoding and decoding stages	37
10	Display adaptation	39
10.1	General	39
10.2	Display SDR backward compatibility	39
10.2.1	Conversion and coding process example	39
10.2.2	Using colour remapping information SEI message	41
Bibliography		43

STANDARDSISO.COM : Click to view the full PDF of ISO/IEC TR 23008-15:2018

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. A technically aligned twin text is published as ITU-T H.Supp18.

A list of all parts in the ISO/IEC 23008 series can be found on the ISO website.

Introduction

High dynamic range (HDR) video is a type of video content in which the sample values represent a larger luminance range than conventional standard dynamic range (SDR) video. HDR video can provide an enhanced viewer experience and can more accurately reproduce scenes that include, within the same image, dark areas and bright highlights, such as emissive light sources and reflections. Wide colour gamut (WCG) video, on the other hand, is video characterized by a wider spectrum of colours compared to what has been commonly available in conventional video. Recent advances in capture and display technology have enabled consumer distribution of HDR and WCG content. However, given the characteristics of such content, special considerations may need to be made, in terms of both processing and compression, compared to conventional content.

This document relates to HDR/WCG video coding and distribution, using single-layer or dual-layer coding, with the signalling specified for Rec. ITU-T H.265 | ISO/IEC 23008-2 High efficiency video coding (HEVC), and when applicable, Rec. ITU-T H.264 | ISO/IEC 14496-10 Advanced video coding (AVC).

This document serves several purposes:

- It provides a survey of identified video usability information (VUI) syntax elements and supplemental enhancement information (SEI) messages specified in HEVC and AVC applicable for HDR/WCG video.
- It covers conversion and coding chains using the IC_{TCp} colour representation, and the hybrid log-gamma (HLG) transfer functions.
- Examples of using colour remapping information (CRI) and tone mapping information (TMI) SEI messages for the support of SDR backward compatibility and display adaptation functionalities are described.
- A dual-layer coding approach using the Scalable Main 10 profile of HEVC for backward compatibility with SDR systems is also documented.

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 15:

Signalling, backward compatibility and display adaptation for HDR/WCG video

1 Scope

This document reviews approaches for processing and coding of HDR/WCG video content. The purpose of this document is to provide a set of publicly-referenceable methods for the operation of AVC or HEVC video coding systems adapted for compressing HDR/WCG video for consumer distribution applications.

This document first includes a review of the video usability information (VUI) indicators and supplemental enhancement information (SEI) messages applicable for HDR/WCG video. It provides a description of processing steps for converting from 4:4:4 RGB linear light representation video signals into video signals with IC_{TCp} colour representation and perceptual quantizer (PQ) transfer function, or with $Y'CbCr$ colour representation and HLG transfer function (IC_{TCp} , PQ and HLG are defined in Rec. ITU-R BT.2100-1). Some high-level approaches for compressing these signals using either Rec. ITU-T H.264 | ISO/IEC 14496-10 or Rec. ITU-T H.265 | ISO/IEC 23008-2 are provided. A description of post-decoding processing steps is also included for converting back to a linear light, 4:4:4 RGB representation. The document also addresses the standard dynamic range (SDR) backward compatibility, that is, the compatibility with legacy decoding systems that are not able to detect and properly display HDR/WCG video content. It describes example implementations of this feature using three different solutions: using HLG as a backward compatible transfer function, using CRI and TMI SEI messages, using dual-layer approach with the Scalable Main 10 profile of HEVC and an SDR compatible base layer. Finally, the document illustrates the usage of CRI SEI messages to convey metadata enabling the dynamic range and colour gamut adaptation at the display side of the decoded video to the display capabilities.

NOTE The document complements the material provided in ITU-T H.Supp15 | ISO/IEC TR 23008-14, which is focused on conversion and coding practices for non-constant luminance (NCL) $Y'CbCr$ video signals using the PQ transfer function.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

Recommendation ITU-T H.264 | ISO/IEC 14496-10: 2014, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

Recommendation ITU-T H.265 | ISO/IEC 23008-2: 2017, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in Rec. ITU-T H.264 | ISO/IEC 14496-10, Rec. ITU-T H.265 | ISO/IEC 23008-2, and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at <http://www.electropedia.org/>
- ISO Online browsing platform: available at <https://www.iso.org/obp>

3.1
dynamic range adaptation

DRA

mapping process to convert content from one colour volume to another colour volume

3.2
electro-optical transfer function

EOTF

function which converts a non-linear video signal into a quantity of output linear light

Note 1 to entry: An example of output linear light is light emitted by a display.

3.3
full range

range in a fixed-point (integer) representation that spans the full range of values that could be expressed with that bit depth, such that, for 10-bit signals, black corresponds to code value 0 and peak white corresponds to code value 1023 for Y'

Note 1 to entry: As per the full range definition from Rec. ITU-R BT.2100-1

3.4
hybrid log-gamma

HLG

one set of transfer functions offering a degree of compatibility with legacy displays by more closely matching the previously established television transfer curves

Note 1 to entry: Sets of transfer functions related to HDR signals are specified in Rec. ITU-R BT.2100-1.

3.5
narrow range

range in a fixed-point (integer) representation that does not span the full range of values that could be expressed with that bit depth such that, for 10-bit representations, 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

Note 1 to entry: As per the narrow range definition from Rec. ITU-R BT.2100-1.

3.6
opto-electronic transfer function

OETF

function which converts a source input linear optical intensity into a non-linear video signal

Note 1 to entry: An example of input linear optical intensity is light input to a camera.

3.7
opto-optical transfer function

OOTF

function which has the role of applying the “rendering intent” on video signal

Note 1 to entry: In general, an OOTF is a concatenation of an OETF, artistic adjustments and an EOTF.

3.8**perceptual quantizer****PQ**

one set of transfer functions achieving a very wide range of brightness levels for a given bit depth using a non-linear transfer function that is finely tuned to match the human visual system

Note 1 to entry: Sets of transfer functions related to HDR signals are specified in Rec. ITU-R BT.2100-1.

3.9**random access point access unit****RAPAU**

access unit in the bitstream at which the initiation of the decoding process for some or all subsequent pictures in the bitstream is intended to be feasible

3.10**reference electro-optical transfer function****reference EOTF**

specified EOTF for use under specific viewing environment, named the reference viewing environment

3.11**reference opto-electronic transfer function****reference OETF**

specified OETF implemented within cameras, to ensure consistency of the image between cameras from different manufacturers

3.12**reference viewing environment**

parameters to establish a reproducible viewing environment for critical viewing of material that can provide repeatable results from one facility to another when viewing the same material

Note 1 to entry: Rec. ITU-R BT.2100-1:2017, Table 3 provides reference viewing environment parameters for HDR programme material.

4 Abbreviated terms

For the purposes of this document, the abbreviated terms given in Rec. ITU-T H.264 | ISO/IEC 14496-10, Rec. ITU-T H.265 | ISO/IEC 23008-2 and the following apply.

ATC	alternative transfer characteristics
AVC	advanced video coding, specified in Rec. ITU-T H.264 ISO/IEC 14496-10
AVE	ambient viewing environment
CGS	colour gamut scalability
CI	constant intensity
CL	constant luminance
CLL	content light level
CLVS	coded layer-wise video sequence
CRI	colour remapping information
FIR	finite impulse response
HDR	high dynamic range

HEVC	high efficiency video coding, specified in Rec. ITU-T H.265 ISO/IEC 23008-2
IC _T C _P	alternative colour space representation to Y'CbCr, specified in Rec. ITU-R BT.2100-1
LMS	long, medium, and short wavelength-based colour space, specified in Rec. ITU-R BT.2100-1
LUT	look-up table
MAD	mean absolute difference
MDCV	mastering display colour volume
NCL	non-constant luminance
PQ10	HDR content representation that utilizes the Rec. ITU-R BT.2100-1 colour primaries, the Rec. ITU-R BT.2100-1 reference PQ EOTF, and the Rec. ITU-R BT.2100-1 Y'CbCr colour space representation with 10 bits per sample in the 4:2:0 chroma sampling format
QP	quantization parameter
RGB	colour system using red, green, and blue components
SDR	standard dynamic range
SEI	supplemental enhancement information
SHVC	scalable high efficiency video coding
SPS	sequence parameter set
SSE	sum of squared errors
TMI	tone mapping information
UHD	ultra-high definition
VUI	video usability information
WCG	wide colour gamut
XYZ	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, Y' _B C _R , or Y' _B C' _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".]

5 Conventions

5.1 General

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.

5.2 Arithmetic operators

+	addition
-	subtraction (as a two-argument operator) or negation (as a unary prefix operator)
*	multiplication, including matrix multiplication
x^y	exponentiation (Denotes 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.]
÷	division in mathematical formulae where no truncation or rounding is intended
$\frac{x}{y}$	division in mathematical formulae where no truncation or rounding is intended
$\sum_{i=x}^y f(i)$	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 $x \geq 0$ and $y > 0$.)

5.3 Bit-wise operators

&	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.)
$x \gg y$	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.)
$x \ll y$	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.)

5.4 Assignment operators

- = 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 given, i.e., x += 3 is equivalent to x = x + 3, and x += (-3) is equivalent to x = x + (-3)
- = decrement by amount given, i.e., x -= 3 is equivalent to x = x - 3, and x -= (-3) is equivalent to x = x - (-3)

5.5 Relational, logical and other operators

- == equality operator
- != not equal to operator
- !x logical negation “not”
- > larger than operator
- < smaller than operator
- >= larger than or equal to operator
- <= smaller than or equal to operator
- && conditional/logical “and” operator
(Performs a logical “and” of its Boolean operators, but only evaluates the second operand if necessary.)
- || conditional/logical “or” operator
(Performs a logical “or” of its Boolean operators, but only evaluates the second operand if necessary.)
- a ? b : c ternary conditional
(If condition a is true, then the result is equal to b; otherwise the result is equal to c.)

5.6 Mathematical functions

$$\text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$$

Ceil(x) smallest integer greater than or equal to x

$$\text{Clip3}(x,y,z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{otherwise} \end{cases}$$

$EOTF_{PQ}(x)$	reference PQ EOTF used to convert a non-linear light PQ representation to a linear light representation
$Exp(x) = e^x$	where e is Euler's base constant 2.718 281 828....
$Floor(x)$	largest integer less than or equal to x
$iEOTF_{PQ}(x)$	inverse reference PQ EOTF used to convert a linear light representation to a non-linear light representation
$iOETF_{HLG}(x)$	inverse reference HLG OETF used to convert a non-linear light representation to a scene-referred linear light representation
$Ln(x)$	natural logarithm of x (the base-e logarithm, where e is natural logarithm base constant 2.718 281 828....)
$Log_{10}(x)$	base-10 logarithm of x
$Max(x,y) = \begin{cases} x & ; x > y \\ y & ; \text{otherwise} \end{cases}$	
$OETF_{HLG}(x)$	reference HLG OETF used to convert a scene-referred linear light representation to a non-linear light representation
$Round(x) = Sign(x) * Floor(Abs(x) + 0.5)$	
$Sign(x) = \begin{cases} 1 & ; x > 0 \\ 0 & ; x = 0 \\ -1 & ; x < 0 \end{cases}$	
$Sqrt(x) = \sqrt{x}$	

5.7 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 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 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)
"x ^y "

Table 1 (continued)

Operations (with operands x, y, and z)
"x * y", "x / y", "x ÷ y", " $\frac{x}{y}$ ", "x % y"
"x + y", "x - y" (as a two-argument operator), " $\sum_{i=x}^y f(i)$ "
"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"

6 Overview

This document is structured as follows.

- [Clause 7](#) reviews identified signalling mechanisms of HEVC, and when applicable, of AVC, relevant for HDR/WCG video coding and distribution. It also describes some common processing steps used in end-to-end processing chains such as described in [Clauses 8](#) and [9](#).
- [Clause 8](#) describes usage of HLG transfer functions and CRI or TMI SEI messages for the support of bitstream SDR backward compatibility (defined below) with a single-layer profile (e.g. HEVC Main 10).
- [Clause 9](#) describes a dual-layer HDR/WCG video coding system with bitstream SDR backward compatibility implemented with the HEVC Scalable Main 10 profile.
- [Clause 10](#) addresses the display adaptation functionality (defined below), with application examples based on the CRI SEI message. This clause includes the specific case of display SDR backward compatibility.

“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.2020-2 colour primaries). In the present document, it is defined in two modes: bitstream and display.

- 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 are 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.2100-1 reference HLG opto-electronic transfer function (OETF) and Rec. ITU-R BT.2100-1 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 static, and 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, using metadata conveyed for

instance 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.2100-1 inverse reference PQ electro-optical transfer function (EOTF) and Rec. ITU-R BT.2100-1 colour primaries). A post-decoding dynamic range adaptation (DRA) 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 adapt the decoded video signal to a target display. Techniques providing display SDR backward compatibility are 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 of displaying a certain colour volume (dynamic range and colour gamut), to a video signal suitable to a target rendering device of colour volume capabilities different from the reference display capabilities. For instance, it can be used to convert a Y’CbCr 4:2:0 10-bits Rec. ITU-R BT.2100-1 PQ signal (denoted PQ10 in the present document), originated from an HDR video master produced on a display with a given reference peak luminance, to a lower peak luminance capable display. 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, for instance using SEI messages.

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.2100-1 PQ transfer characteristics are outside of scope of this document. These aspects are specifically addressed in ITU-T H.Supp15 | ISO/IEC TR 23008-14.

7 HEVC signalling mechanisms applicable to HDR/WCG video

7.1 General

This clause provides an overview of the VUI syntax elements and SEI messages specified in HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2), applicable to HDR/WCG video and relevant to the scope of this document. The PQ, HLG transfer functions, and the IC_{TCP} colour representation are also described.

This clause is structured as follows.

- Subclause 7.2 reviews VUI signalling applicable to HDR/WCG video.
- Subclause 7.3 reviews SEI messages applicable to HDR/WCG video.
- Subclause 7.4 provides an overview of PQ and HLG transfer functions.
- Subclause 7.5 provides a description of IC_{TCP} colour representation, including conversion and coding practices related to HDR/WCG video signals represented with IC_{TCP} 4:2:0 video with Rec. ITU-R BT.2100-1 inverse reference PQ EOTF.

Conversion and coding practices related to HDR/WCG video signals represented with Y’CbCr 4:2:0 video with Rec. ITU-R BT.2100-1 HLG transfer characteristics are discussed in subclause 8.2.

7.2 VUI syntax elements

By design, metadata signalled in syntax elements of VUI is not necessary for constructing the luma or chroma samples by the decoding process, and may be ignored by the decoder. However, such syntax elements provide useful parameters or attributes of an encoded signal and can 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 Rec. ITU-T H.265 | ISO/IEC

23008-2:2017, Tables E.3, E.4 and E.5 respectively. [Table 2](#) and [Table 3](#) provide values of VUI syntax elements that indicate usage of Rec. ITU-R BT.2100-1 representation of the video signal, including matrix coefficients associated with Rec. ITU-R BT.2100-1 (same as those associated with Rec. ITU-R BT.2020-2).

Rec. ITU-R BT.2100-1 specifies HDR-TV image parameters for use in production and international programme exchange. It defines two sets of transfer functions: perceptual quantizer (PQ) and hybrid log-gamma (HLG). RGB colour primaries are defined identically as in Rec. ITU-R BT.2020-2. Rec. ITU-R BT.2100-1 describes two different luminance and colour difference signal representations: non-constant luminance (NCL) Y'CbCr and constant intensity (CI) IC_TC_P. Syntax elements of VUI in HEVC can be used to convey the metadata describing such attributes of the coded signal. The VUI transfer_characteristics syntax element either indicates the reference OETF of the source picture as a function of a source input linear optical intensity or indicates the inverse of the reference EOTF as a function of an output linear optical intensity, as described in [Table 2](#) for HEVC. RGB colour primaries are indicated using colour_primaries syntax element (set equal to 9 for Rec. ITU-R BT.2100-1/Rec. ITU-R BT.2020-2 colour primaries). Colour representation is indicated using matrix_coeffs syntax element, as described in [Table 3](#) for HEVC.

Table 2 — Values of transfer_characteristics indication in VUI in HEVC

	PQ	HLG
transfer_characteristics	16	18

Table 3 — Values of matrix_coeffs indication in VUI in HEVC

	NCL Y'CbCr	CI IC _T C _P
matrix_coeffs	9	14

NOTE 1 PQ is also defined in SMPTE ST 2084 and HLG is also defined in ARIB STD-B67.

NOTE 2 VUI syntax element values for PQ are also defined in AVC.

7.3 SEI messages applicable for HDR/WCG video

7.3.1 General

SEI messages assist in processes related to decoding, display or other purposes. They are not required for constructing the luma or chroma samples by the decoding process. HEVC and AVC specify several SEI messages applicable to HDR/WCG video. Some SEI messages convey descriptive information about the content. These SEI messages are reviewed in subclauses [7.3.2](#) to [7.3.5](#). Some other SEI messages are devoted to enabling specific post-processing of the decoded samples, e.g. signal adaptation processes. Some SEIs can be used for the conversion of decoded content from one colour volume to another colour volume. For example, this approach can apply 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 a different colour volume. This comprises conversion to an SDR version for display SDR backward compatibility use case. These various SEI messages are described in subclauses [7.3.6](#) and [7.3.7](#).

7.3.2 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 (ISO 11664-1), 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. The HEVC specification requires that all MDCV SEI messages that apply to the same CLVS have the same content.

NOTE 1 The mastering display colour volume SEI message is also defined in AVC.

NOTE 2 Mastering display colour volume metadata is also defined in SMPTE ST 2086.

7.3.3 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/m^2), among the 4:4:4 R, G, B samples of the content pictures in the coded video sequence. As for the MDCV SEI message, 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 (see Reference [12]). The CLL SEI message persists until the end of the CLVS. Rec. ITU-T H.265 | ISO/IEC 23008-2 requires that all CLL SEI messages that apply to the same CLVS have the same content.

NOTE 1 The content light level information SEI message is also defined in AVC.

NOTE 2 Corresponding metadata associated with the CLL SEI message and examples of derivation algorithms are defined in CEA-861.3.

7.3.4 Ambient viewing environment SEI message

The ambient viewing environment (AVE) SEI message characterizes 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) 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. Rec. ITU-T H.265 | ISO/IEC 23008-2 requires that all AVE SEI messages that apply to the same CLVS have the same content.

7.3.5 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 OETF (transfer_characteristics values 1, 6, 14 or 15). The ATC SEI message can be conveyed to indicate the actual OETF used to produce an SDR-compatible HDR video signal representation. This is applicable to the reference HLG OETF (transfer_characteristics value 18), further discussed in subclause 8.2. The ATC SEI message persists until the end of the CLVS. Rec. ITU-T H.265 | ISO/IEC 23008-2 requires that all ATC SEI messages that apply to the same CLVS have the same content.

7.3.6 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 implementing the tone mapping function are specified for the TMI SEI message, including 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 LUT_{TM} . 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 in [Formula \(1\)](#).

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} LUT_{TM}[R_{in}] \\ LUT_{TM}[G_{in}] \\ LUT_{TM}[B_{in}] \end{bmatrix} \quad (1)$$

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 element `tone_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 lesser dynamic range than the decoded video data.

NOTE The tone mapping information SEI message is also specified in AVC.

7.3.7 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 the 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 mapping 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 of the CRI mapping is applied in the 4:4:4 domain. When it is not activated, the mapping process can be applied in the 4:2:2 or 4:2:0 chroma sampling formats.

[Formulae \(2\)](#), [\(3\)](#) and [\(4\)](#) illustrate the application of the complete CRI model to the R, G, B values of a colour sample (although the CRI mapping can also be applied to Y’CbCr samples):

$$\begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} PreLUT_0[R_{in}] \\ PreLUT_1[G_{in}] \\ PreLUT_2[B_{in}] \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} R_2 \\ G_2 \\ B_2 \end{bmatrix} = M_{3 \times 3} \begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} PostLUT_0[R_2] \\ PostLUT_1[G_2] \\ PostLUT_2[B_2] \end{bmatrix} \quad (4)$$

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, the `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, 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 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 is analogous to the purpose of the colour description syntax of the VUI.

More than one remapping can be associated with a coded video sequence through the CRI SEI message identifier tags. This can enable simultaneous support for multiple dynamic range targets, including targets that have greater or lesser dynamic range than the decoded video data.

NOTE The colour remapping information SEI message is also specified in AVC.

7.4 Overview of PQ and HLG transfer functions

7.4.1 General

Rec. ITU-R BT.2100-1 specifies two sets of transfer functions, PQ and HLG. The specification of HLG is defined in terms of its reference OETF, while the specification of PQ is defined in terms of its reference EOTF. Thus, the HLG system is considered as a “scene-referred” system, whereas the PQ system is considered as a “display-referred” system. 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, see Rec. ITU-R BT.2100-1) which compensates for the difference in tonal perception between the environment of the camera and that of the display as further explained in Rec. ITU-R BT.2100-1:2017, Annex 1. 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 may 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 detail about the differences between scene and display-referred signals can be found in Rec. ITU-R BT.2100-1:2017, Annex A.

NOTE 1 The OOTF defined for HLG is different from the OOTF defined for PQ.

In Rec. ITU-R BT.2100-1, the PQ and HLG sets of transfer functions are defined as follows:

- The PQ OETF, converting the signal from scene linear light to non-linear representation, is defined based on the reference PQ EOTF, converting the signal from non-linear to display linear light representation, as the concatenation of two functions:

$$\text{OETF} = \text{EOTF}^{-1}(\text{OOTF})$$

- The HLG EOTF, converting the signal from non-linear to display linear light representation, is defined based on the reference HLG OETF, converting the signal from scene linear light to non-linear representation, as the concatenation of two functions:

$$\text{EOTF} = \text{OOTF}(\text{OETF}^{-1})$$

Functions “iEOTF” and “iOETF” are used in this document to denote EOTF^{-1} and OETF^{-1} , respectively.

NOTE 2 Informative methods for transcoding between PQ and HLG signals are described in ITU-R BT.2100-1:2017, Annex 2 and in Report ITU-R BT.2390-0.

7.4.2 Reference PQ EOTF

The inverse reference PQ EOTF is described in Rec. ITU-T H.265 | ISO/IEC 23008-2:2017, Table E.4 for transfer_characteristics equal to 16. The non-linear light representation V of a linear light intensity signal L_o, which takes values normalized to the range [0, 1], is computed using [Formula \(5\)](#):

$$V = iEOTF_{PQ}(L_o) = \left(\frac{c_1 + c_2 * L_o^n}{1 + c_3 * L_o^n} \right)^m \tag{5}$$

where c₁, c₂, c₃, m, and n are constants defined using [Formulae \(6\) to \(10\)](#):

$$c_1 = c_3 - c_2 + 1 = 3424 \div 4096 = 0.835\ 937\ 5 \tag{6}$$

$$c_2 = 2413 \div 128 = 18.851\ 562\ 5 \tag{7}$$

$$c_3 = 299 \div 16 = 18.687\ 5 \tag{8}$$

$$m = 2523 \div 32 = 78.843\ 75 \tag{9}$$

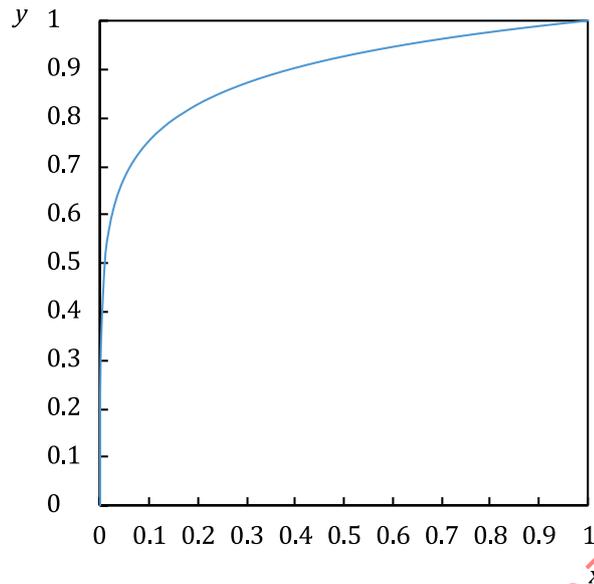
$$n = 1305 \div 8192 = 0.159\ 301\ 757\ 812\ 5 \tag{10}$$

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

The reference PQ EOTF, EOTF_{PQ}(), is described as the exact inverse of [Formula \(5\)](#). The linear light intensity signal L_o is computed from the non-linear representation V, which takes values in the range [0, 1], as per [Formula \(11\)](#):

$$L_o = EOTF_{PQ}(V) = \left(\frac{\text{Max}\left(\left(V^{1/m} - c_1\right), 0\right)}{c_2 - c_3 * V^{1/m}} \right)^{1/n} \tag{11}$$

A plot of the inverse reference PQ EOTF is shown in [Figure 1](#).

**Key**

- x normalized linear light data
 y non-linear converted signal (normalized)

Figure 1 — Graph of the inverse reference PQ EOTF (iEOTF_{PQ})

7.4.3 Reference HLG OETF

The reference HLG OETF, $OETF_{HLG}()$, as described in Rec. ITU-T H.265 | ISO/IEC 23008-2:2017, Table E.4 for transfer_characteristics equal to 18, applies as follows to linear light scene-referred R, G, B samples, normalized to the range [0, 1]. For $L_c = R, G, \text{ or } B$, [Formula \(12\)](#) applies:

$$V = OETF_{HLG}(L_c) = a * \text{Ln}(12 * L_c - b) + c \quad \text{for } 1 \geq L_c > (1 \div 12) \quad (12)$$

$$= \text{Sqrt}(3) * L_c^{0.5} \quad \text{for } (1 \div 12) \geq L_c \geq 0$$

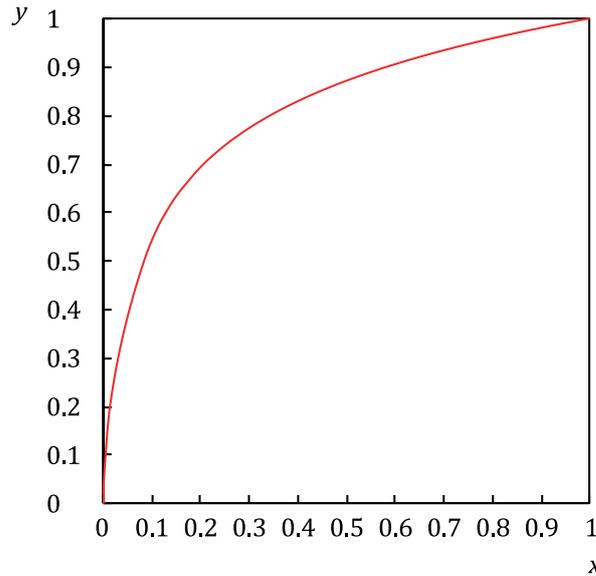
where a , b and c are constants defined as per [Formulae \(13\)](#) to [\(15\)](#):

$$a = 0.178\ 832\ 77 \quad (13)$$

$$b = 1 - 4 * a = 0.284\ 668\ 92 \quad (14)$$

$$c = 0.5 - a * \text{Ln}(4 * a) = 0.559\ 910\ 73 \quad (15)$$

A plot of the reference OETF for HLG is shown in [Figure 2](#).



Key

- x normalized linear light data
- y non-linear converted signal (normalized)

Figure 2 — Graph of the reference HLG OETF (OETF_{HLG})

The inverse of the reference HLG OETF is formulated as per [Formula \(16\)](#):

$$L_c = iOETF_{HLG}(V) = 1 \div 12 * (\text{Exp}((V - c) \div a) + b) \quad \text{for } 1 \geq V > (1 \div 2) \tag{16}$$

$$= V^2 \div 3 \quad \text{for } (1 \div 2) \geq V \geq 0$$

The inverse of the reference HLG OETF does not equate to the EOTF. For more information on the HLG EOTF, the reader can refer to Rec. ITU-R BT.2100-1, Report ITU-R BT.2390-0 and Reference [\[13\]](#).

7.5 IC_{TCP} colour representation

7.5.1 General

An overview of CI IC_{TCP} colour representation is given in Rec. ITU-R BT.2100-1 and Report ITU-R BT.2390-0. This subclause mainly provides a set of methods on processing HDR/WCG video for consumer distribution, including conversion steps for converting from a linear light RGB representation with Rec. ITU-R BT.2100-1 colour primaries to a 10-bits, narrow range, inverse reference PQ EOTF, 4:2:0, IC_{TCP} representation.

The HDR/WCG system described in this subclause follows the same workflow of 10-bits, narrow range, inverse reference PQ EOTF, 4:2:0, non-constant luminance Y'CbCr representation described in Supplement ITU-T H.Sup15 | ISO/IEC TR 23008-14. It consists of four major stages:

- a pre-encoding stage consisting of several pre-processing processes (subclause [7.5.2](#)),
- an encoding stage (subclause [7.5.3](#)),
- a decoding stage (subclause [7.5.4](#)), and
- a post-decoding stage consisting of several post-processing processes (subclause [7.5.5](#)).

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 subclause is to highlight the conversion and coding differences between Rec. ITU-R BT.2100-1 PQ IC_{TCP} and PQ Y'CbCr signals. More detail on common parts of these two systems are provided specifically in Supplement ITU-T H.Sup15 | ISO/IEC TR 23008-14.

NOTE 2 The same workflow can be used for HLG IC_{TCP} by substituting PQ with HLG.

7.5.2 Pre-encoding process

7.5.2.1 General

The pre-encoding process described in this document includes the following components, as presented in [Figure 3](#):

- a conversion component from a linear RGB data representation to a linear LMS data representation, described in subclause [7.5.2.2](#),
- a conversion component from a linear LMS data representation to a non-linear data representation using the inverse reference PQ EOTF, described in subclause [7.5.2.3](#),
- a colour format conversion component that converts data to the IC_{TCP} colour representation, described in subclause [7.5.2.4](#),
- a conversion component that converts a floating-point to a fixed-point representation (e.g. 10 bits), following the process described in subclause [7.5.2.5](#), and
- a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, following the process described in subclause [7.5.2.6](#).

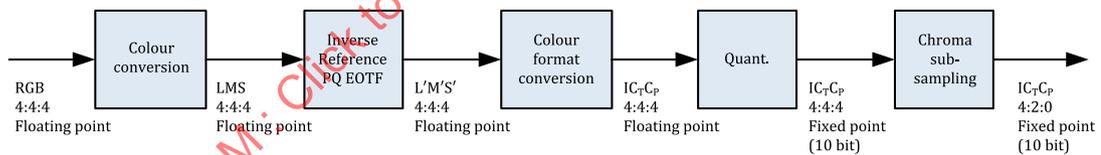


Figure 3 — Conventional PQ IC_{TCP} pre-encoding process system diagram

7.5.2.2 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 3×3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2100-1 (see [Formula 17](#))).

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \frac{1}{4096} \begin{bmatrix} 1688 & 2146 & 262 \\ 683 & 2951 & 462 \\ 99 & 309 & 3688 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (17)$$

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

The inverse reference PQ EOTF [Formula (5)] is applied to all L, M, and S linear light samples, where each component is a number between 0 (representing 0 cd/m²) and 1 (representing 10 000 cd/m²). This results in their non-linear light counterparts L', M', and S' as per Formulae (18) to (20).

$$L' = iEOTF_{PQ}(L) \tag{18}$$

$$M' = iEOTF_{PQ}(M) \tag{19}$$

$$S' = iEOTF_{PQ}(S) \tag{20}$$

7.5.2.4 Colour representation conversion: L'M'S' to IC_TC_P

Conversion from L'M'S' to IC_TC_P representation is performed using a 3 × 3 matrix conversion process (see Formula (21)).

$$\begin{bmatrix} I \\ C_T \\ C_P \end{bmatrix} = \frac{1}{4096} \begin{bmatrix} 2048 & 2048 & 0 \\ 6610 & -13613 & 7003 \\ 17933 & -17390 & -543 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{21}$$

7.5.2.5 Conversion from floating to fixed point (10-bits)

The conversion process from a floating-point representation to a fixed-point 10-bit representation is essentially a quantization step. In general, the conversion process can be expressed as per Formula (22):

$$D' = \text{Clip3}(0, (1 \ll b) - 1, \text{Round}(E' * \text{scale} + \text{offset})) \tag{22}$$

where E' is the floating-point representation of a particular component and D' is the resulting quantized value using b bits. In this document, b = 10. The scale and offset constants depend on the target range (narrow versus full range video) and the component type (luma, chroma, or colour primary components).

For the narrow range representation, the scale and offset for the luma component are set as per Formulae (23) and (24):

$$\text{scale} = 219 * (1 \ll (b - 8)) \tag{23}$$

$$\text{offset} = 1 \ll (b - 4) \tag{24}$$

and the scale and offset for the chroma components are set as per Formulae (25) and (26):

$$\text{scale} = 224 * (1 \ll (b - 8)) \tag{25}$$

$$\text{offset} = 1 \ll (b - 1) \tag{26}$$

NOTE For a 10-bit narrow range representation, DY' results in a value within the range of [64, 940]. Similarly, DCb and DCr result in values within the range of [64, 960].

For the full range representation, the scale and offset for the luma component are set as per Formulae (27) and (28):

$$\text{scale} = (1 \ll b) - 1 \tag{27}$$

$$\text{offset} = 0 \quad (28)$$

and the scale and offset for the chroma components are set as per [Formulae \(29\)](#) and [\(30\)](#):

$$\text{scale} = (1 \ll b) - 1 \quad (29)$$

$$\text{offset} = 1 \ll (b - 1) \quad (30)$$

7.5.2.6 Chroma downscaling

Converting the video data from a 4:4:4 representation to a 4:2:0 representation can follow a similar chroma down-conversion process as described in ITU-T H.Supp15 | ISO/IEC TR 23008-14. This process is explained below.

It is anticipated that consumer electronics conversion systems would typically use 2-D separable finite impulse response (FIR) linear filters for low-pass filtering the chroma data before subsampling (2:1 decimation step). Such filters would basically be of the form [Formula \(31\)](#):

$$y[n] = \sum_{i=-N}^N b_i * x[n+i] \quad (31)$$

where $x[n]$ is the input chroma signal, $y[n]$ is the filtered output chroma signal, $(2 * N)$ corresponds to the filter order or, equivalently, $(2 * N + 1)$ corresponds to the number of taps of the filter, and b_i corresponds to the coefficient of the filter at position i . For example, the two short-tap-length linear FIR filters of [Table 4](#) have been used in experiments for development of this report. Such filters can be utilized for both vertical and horizontal filtering of the chroma samples.

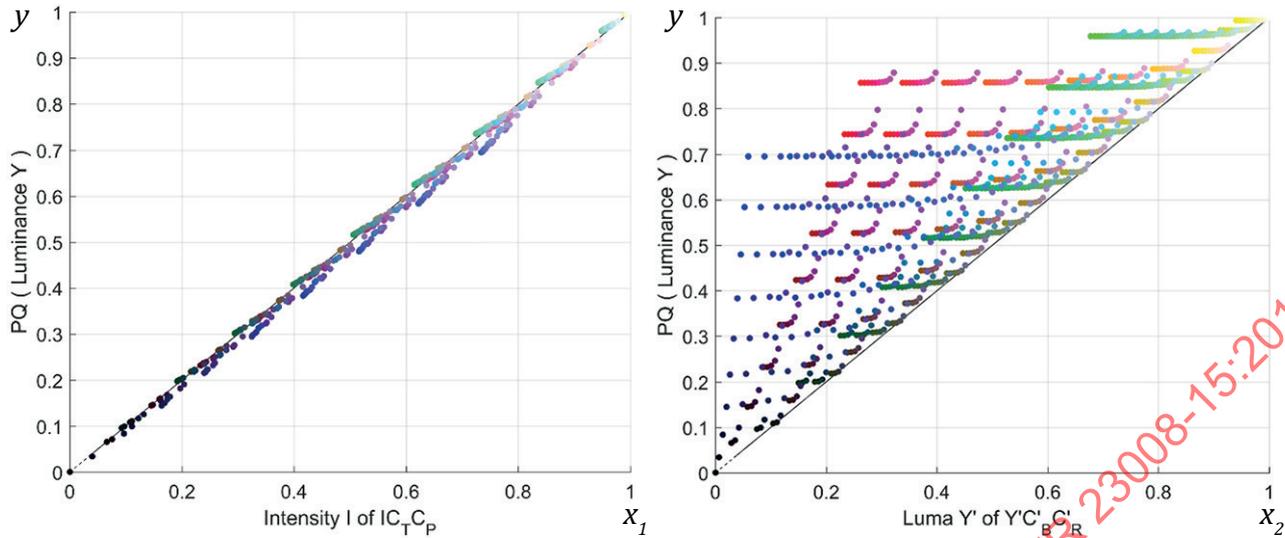
Table 4 — Filter example for chroma down-sampling

Filter	Filter coefficients		
	b_{-1}	b_0	b_1
f_0	$1 \div 8$	$6 \div 8$	$1 \div 8$
f_1	$1 \div 4$	$2 \div 4$	$1 \div 4$

7.5.2.7 Comments on IC_{TCp} Chroma down-conversion

Converting the IC_{TCp} video data from a 4:4:4 representation to a 4:2:0 representation follows the chroma down-conversion described in subclause [7.5.2.6](#).

It is observed in Reference [\[14\]](#) that the closed loop luma adjustment method described in ITU-T H.Supp15 | ISO/IEC TR 23008-14, which aims at reducing NCL Y'CbCr chroma down-sampling artefacts, does not prove to be beneficial for IC_{TCp} chroma down-conversion. This can be explained by the constant luminance property of IC_{TCp}. As shown in [Figure 4](#), the achromatic axis of Y' deviates substantially from luminance especially when colour becomes more saturated. On the other hand, the achromatic axis of IC_{TCp} (I) corresponds very closely with luminance (see also Reference [\[15\]](#)). The constant luminance property of IC_{TCp} 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.



Key
 x_1 intensity I of IC_{TcP}
 x_2 Luma Y' of $Y'CbCr$
 y PQ (Luminance Y)

Figure 4 — Luminance correlation with IC_{TcP} and $Y'CbCr$ colour representations

7.5.3 Encoding process

7.5.3.1 General

After pre-processing, the data is ready for compression. ITU-T H.Supp15 | ISO/IEC TR 23008-14 presents some guidance on how an encoder may be configured for HDR/WCG 10-bits, narrow range, PQ, 4:2:0, $Y'CbCr$ data. Two quantization methodologies, perceptual luma quantization, and chroma QP offset, are described in that document. For the HDR/WCG 10-bit, narrow range, PQ, 4:2:0, IC_{TcP} data coming out from pre-processing step, those two methodologies are still valid. Reference [16] shows that IC_{TcP} exhibits different characteristics than $Y'CbCr$ data for current test data, therefore the exact parameter settings for these two quantization methods might need to be reconfigured. One practice for IC_{TcP} encoding is to configure the encoder settings to follow the bitrate allocation behaviour of $Y'CbCr$ (see Reference [16]). An example of encoder setting is described in subclause 7.5.3.2.

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. These processes commonly consider simple distortion metrics such as mean absolute difference (MAD), or sum of squared errors (SSE), for making a variety of decisions for the encoder settings, and these may have been tuned based on SDR content characteristics. However, given the earlier observations about the differences in the characteristics between SDR and HDR/WCG content, these settings may also need to be appropriately adjusted. Furthermore, other metrics may also be more appropriate for these encoding decisions. These aspects are not explored in the context of this document.

7.5.3.2 Example of encoder setting

Two main observations are made in Reference [16]. Firstly, the Y' component in $Y'CbCr$ and the I component in IC_{TcP} have very similar compression characteristics in terms of variance. This results in similar coding bits given same luma quantizer. Reference [16] suggests that techniques designed to improve coding efficiency of Y' component (e.g., perceptual luma quantization from ITU-T H.Supp15 | ISO/IEC TR 23008-14) can be used for the I component directly. Secondly, for colour components, C_T and

C_p have a higher variance than C_b and C_r , thus modifications may be applied for technologies designed to improve chroma coding efficiency, such as the adaptive chroma QP offset model in ITU-T H.Supp15 | ISO/IEC TR 23008-14.

In the following example of chromaQP offset settings for $IC_T C_p$, 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 C_b , corresponding to C_T , and C_r , corresponding to C_p , based on the luma QP and a factor dependent on the capture and representation colour primaries. The model is expressed as per [Formulae \(32\)](#) and [\(33\)](#):

$$QP_{offsetCb} = \text{Clip3}\left(-12, 12, Cb_{offset} + \text{Clip3}\left(-12, 0, \text{Round}\left(c_{cb} * (k * QP + l)\right)\right)\right) \quad (32)$$

$$QP_{offsetCr} = \text{Clip3}\left(-12, 12, Cr_{offset} + \text{Clip3}\left(-12, 0, \text{Round}\left(c_{cr} * (k * QP + l)\right)\right)\right) \quad (33)$$

where $c_{cb} = 1$ and $Cb_{offset} = 6$ if the capture colour primaries are the same as the representation colour primaries, $c_{cb} = 0.5$ and $Cb_{offset} = 8$ if the capture colour primaries are equal to the P3D65 primaries and the representation colour primaries are equal to the Rec. ITU-R BT.2100-1 primaries, and $c_{cb} = 0.9$ and $Cb_{offset} = 6$ if the capture colour primaries are equal to the Rec. ITU-R BT.709-6 primaries and the representation primaries are equal to the Rec. ITU-R BT.2100-1 primaries.

Similarly, $c_{cr} = 1$ and $Cr_{offset} = 6$ if the capture colour primaries are the same as the representation colour primaries, $c_{cr} = 0.9$ and $Cr_{offset} = 7$ if the capture colour primaries are equal to the P3D65 primaries and the representation colour primaries are equal to the Rec. ITU-R BT.2100-1 primaries, and $c_{cr} = 1.6$ and $Cr_{offset} = 6$ if the capture colour primaries are equal to the Rec. ITU-R BT.709-6 primaries and the representation primaries are equal to the Rec. ITU-R BT.2100-1 primaries.

Finally, $k = -9.46$ and $l = 10$.

7.5.3.3 HEVC encoding

When creating an HEVC bitstream, it is suggested to set syntax elements to the values listed in [Table 5](#) in the sequence parameter set (SPS) of the bitstream. The syntax elements in [Table 5](#) are conveyed in the video usability information syntax branch of the SPS defined in Rec. ITU-T H.265 | ISO/IEC 23008-2:2017, Annex E. They may also be duplicated and carried in various application-layer headers.

Table 5 — Suggested settings for HEVC encoding of $IC_T C_p$ 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_coefs	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.2100-1, Rec. ITU-R BT.2100-1 inverse reference PQ EOTF and $IC_T C_p$ colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec.

ITU-R BT.709-6 colour primaries and Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 OETF. Chroma QP adjustment, as described in subclause 7.5.3.2 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 subclause 7.5.3.2 could be achieved by adjusting the syntax elements `cu_qp_delta_abs` and `cu_qp_delta_sign_flag`.

7.5.3.4 AVC encoding

When creating an AVC bitstream, it is suggested to set syntax elements to the values listed in Table 6 in the SPS of the bitstream. The syntax elements in Table 6 are conveyed in the video usability information syntax branch of the SPS defined in Rec. ITU-T H.264 | ISO/IEC 14496-10:2014, Annex E. They may also be duplicated and carried in various application-layer headers.

Table 6 — Suggested settings for AVC encoding of IC_{TCP} 4:2:0 PQ 10-bit signal

Syntax element	Location	Suggested value
<code>profile_idc</code>	<code>seq_parameter_set_data()</code>	110
<code>vui_parameters_present_flag</code>	<code>seq_parameter_set_data()</code>	1
<code>video_signal_type_present_flag</code>	<code>vui_parameters()</code>	1
<code>video_full_range_flag</code>	<code>vui_parameters()</code>	0
<code>colour_description_present_flag</code>	<code>vui_parameters()</code>	1
<code>colour_primaries</code>	<code>vui_parameters()</code>	9
<code>transfer_characteristics</code>	<code>vui_parameters()</code>	16
<code>matrix_coefficients</code>	<code>vui_parameters()</code>	14
<code>chroma_loc_info_present_flag</code>	<code>vui_parameters()</code>	1
<code>chroma_sample_loc_type_top_field</code>	<code>vui_parameters()</code>	2
<code>chroma_sample_loc_type_bottom_field</code>	<code>vui_parameters()</code>	2

For HDR/WCG content represented with the colour primaries of Rec. ITU-R BT.2100-1, Rec. ITU-R BT.2100-1 inverse reference PQ EOTF and IC_{TCP} colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec. ITU-R BT.709-6 colour primaries and Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 OETF. Chroma QP adjustment, as described in subclause 7.5.3.2 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 subclause 7.5.3.2 could be achieved by adjusting the syntax element `mb_qp_delta`.

7.5.4 Decoding process

When the bitstream is an HEVC bitstream, the decoding process as in Rec. ITU-T H.265 | ISO/IEC 23008-2 is performed.

When the bitstream is an AVC bitstream, the decoding process in Rec. ITU-T H.264 | ISO/IEC 14496-10 is performed.

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

7.5.5 Post-decoding process

7.5.5.1 General

The post-decoding stage includes the following components, as presented in Figure 5:

- a) a chroma up-conversion component that converts data from 4:2:0 to 4:4:4, following the process described in subclause 7.5.5.2,

- b) a conversion component that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, described in subclause 7.5.5.3,
- c) a colour format conversion component that converts data from the IC_TC_P representation back to the non-linear L'M'S' representation, described in subclause 7.5.5.4,
- d) a conversion component from the non-linear L'M'S' data representation back to a linear LMS data representation, described in subclause 7.5.5.5,
- e) a conversion component from a linear LMS data representation to a linear RGB data representation, described in subclause 7.5.5.6.

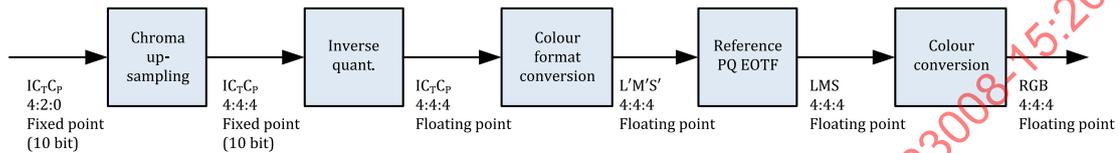


Figure 5 — Conventional PQ IC_TC_P post-decoding process system diagram

7.5.5.2 Chroma up-sampling

Chroma plane interpolation, both vertically and horizontally, is performed to convert the 4:2:0 signal to a 4:4:4 representation, using a similar chroma up-sampling process as described in ITU-T H.Supp15 | ISO/IEC TR 23008-14. It is quite likely that FIR linear filters would be used by many implementations for this process. The simple two-phase resampling filter shown in Table 7 was used for the experiments conducted for the preparation of this report. The same filter is applied both vertically and horizontally.

NOTE This is essentially a “Lanczos 2” filter; higher-precision and higher-order filters could potentially be used, especially when up sampling content of very high quality or with no compression.

Table 7 — Two-phase chroma resampling filter

Phase p	Interpolation filter coefficients			
	fc[p, -1]	fc[p, 0]	fc[p, 1]	fc[p, 2]
0	0	1	0	0
1	-1 ÷ 16	9 ÷ 16	9 ÷ 16	-1 ÷ 16

7.5.5.3 Conversion from fixed to floating point

This process can be seen as the exact inverse of the process presented in subclause 7.5.2.5. In particular, a fixed-point precision value can be converted to a floating-point precision value using Formula (34):

$$E' = \text{Clip3}(\text{minE}, \text{maxE}, (D' - \text{offset}) \div \text{scale}) \tag{34}$$

The exact same values for scale and offset as in subclause 7.5.2.5 are used according to the component type, whereas minE and maxE are equal to -0.5 and 0.5 for the chroma components respectively, and equal to 0 and 1.0 for all other colour components.

7.5.5.4 Colour representation conversion: IC_TC_P to L'M'S'

Conversion from the IC_TC_P representation back to the L'M'S' representation can be performed using a 3 × 3 matrix conversion process (see [Formula \(35\)](#)).

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} 1 & 0.008\ 609\ 037\ 037\ 933 & 0.111\ 029\ 625\ 003\ 026 \\ 1 & -0.008\ 609\ 037\ 037\ 933 & -0.111\ 029\ 625\ 003\ 026 \\ 1 & 0.560\ 031\ 335\ 710\ 679 & -0.320\ 627\ 174\ 987\ 319 \end{bmatrix} \begin{bmatrix} I \\ C_T \\ C_P \end{bmatrix} \quad (35)$$

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

The reference PQ EOTF (Formula 7-11) is applied to all L', M', and S' non-linear representations, resulting in their linear light counterparts L, M, and S as per [Formulae \(36\)](#) to [\(38\)](#):

$$L = \text{EOTF}_{PQ}(L') \quad (36)$$

$$M = \text{EOTF}_{PQ}(M') \quad (37)$$

$$S = \text{EOTF}_{PQ}(S') \quad (38)$$

7.5.5.6 Colour representation conversion: LMS to RGB

Conversion from the LMS representation back to the RGB representation can be performed using a 3 × 3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2100-1 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) (see [Formula \(39\)](#)).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.436\ 606\ 694\ 333\ 079 & -2.506\ 452\ 118\ 656\ 270 & 0.0698\ 454\ 243\ 231\ 91 \\ -0.791\ 329\ 555\ 598\ 929 & 1.983\ 600\ 451\ 792\ 291 & -0.192\ 270\ 896\ 193\ 362 \\ -0.025\ 949\ 899\ 690\ 593 & -0.098\ 913\ 714\ 711\ 726 & 1.124\ 863\ 614\ 402\ 319 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (39)$$

8 Bitstream SDR backward compatibility with single-layer coding

8.1 General

This subclause 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; Approach 2 which applies an “adaptive” HDR-to-SDR conversion process prior to the encoding, with dynamic metadata used to perform the inverse conversion after decoding. Approach 1 can be addressed with HEVC or AVC using the Rec. ITU-R BT.2100-1 HLG transfer functions. Approach 2 can be addressed with HEVC or AVC using the CRI or TMI SEI messages to convey the dynamic metadata.

8.2 Approach 1: usage of HLG for “static” bitstream SDR backward compatibility

8.2.1 General

The HLG transfer function (Rec. ITU-R BT.2100-1) was 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 Rec. ITU-R BT.2100-1 to

Rec. ITU-R BT.709-6 colour primaries) but needing only dynamic range conversion. HLG is suggested to be used with Rec. ITU-R BT.2100-1 colour primaries.

This subclause provides a set of methods on processing HDR/WCG video for consumer distribution with SDR compatibility, including conversion steps for converting from a linear light RGB representation with Rec. ITU-R BT.2100-1 colour primaries to a 10-bits, narrow range, reference HLG OETF, 4:2:0, Y'CbCr representation.

The HLG-based HDR/WCG workflow consists of four major stages:

- a pre-encoding stage described in subclause 8.2.2,
- an encoding stage described in subclause 8.2.3, including suggested signalling using VUI and ATF SEI message to enable the SDR compatibility feature,
- a decoding stage described in subclause 8.2.4, and
- a post-decoding stage described in subclause 8.2.5.

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

8.2.2 HLG pre-encoding conversion process

8.2.2.1 General

An example HLG pre-encoding conversion process is illustrated in Figure 6. It 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:

- a) a pre-conversion step, denoted “inverse system gamma” in Figure 6 and described in subclause 8.2.2.3, when the input HDR signal is display-light referred (see explanations in subclause 7.4); when the input HDR signal is scene-light referred (e.g. from camera, see explanations in subclause 7.4), this pre-conversion step is not expected to be applied,
- b) 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 reference HLG OETF as described in subclause 8.2.2.2,
- c) a colour format conversion from non-linear HLG R'G'B' 4:4:4 signal to Y'CbCr 4:4:4 as described in subclause 8.2.2.4,
- d) a conversion step that converts a floating-point to a fixed-point representation (e.g. 10 bits), narrow range, following the subclause 7.5.2.5,
- e) a chroma down-conversion that converts data from 4:4:4 to 4:2:0, for instance as described in subclause 7.5.2.6.

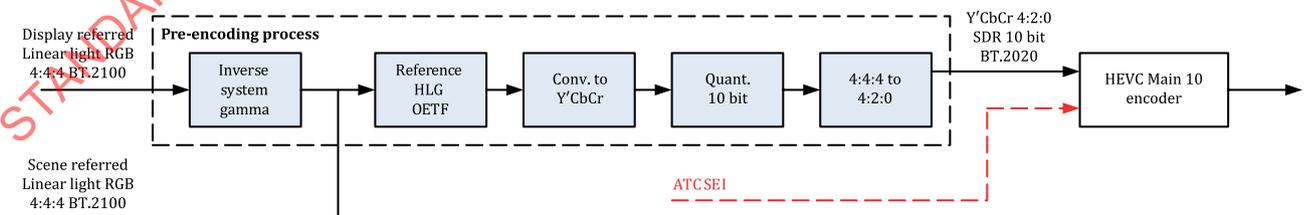


Figure 6 — Conventional HLG Y'CbCr pre-encoding process system diagram

8.2.2.2 HLG conversion from a linear RGB to a non-linear representation R'G'B'

For a scene-referred linear light input, the reference HLG OETF as defined in Formula (12) 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”) is expected to be applied to convert the display-referred video to a scene-referred signal. This step removes the inherent OOTF within the display-referred signal. Subclause 8.2.2.3 describes this preliminary conversion process.

8.2.2.3 Conversion from linear light display-referred to linear light scene-referred (inverse system gamma)

This step aims to remove the OOTF inherent in display-referred signals (see Reference [17]). 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 sample:

- Let R_{display}, G_{display}, B_{display} be the display-referred R, G, B values. R_n, G_n, B_n are derived as per [Formula \(40\)](#):

$$\begin{cases} R_n = \text{Max}(0, R_{\text{display}} \div P) \\ G_n = \text{Max}(0, G_{\text{display}} \div P) \\ B_n = \text{Max}(0, B_{\text{display}} \div P) \end{cases} \tag{40}$$

- The luminance Y_{display} is derived from R_n, G_n, B_n values as per [Formula \(41\)](#):

$$Y_{\text{display}} = a_0 * R_n + a_1 * G_n + a_2 * B_n \tag{41}$$

where a₀ = 0.262 7, b₀ = 0.678 0 and c₀ = 0.059 3 are the conventional RGB-to-Y derivation coefficients for Rec. ITU-R BT.2100-1 colour gamut primaries.

- The value of the system gamma g is estimated from the peak luminance P as per [Formula \(42\)](#):

$$g = 1.2 + 0.42 * \text{Log}_{10}(P \div 1000) \tag{42}$$

- The value g_{scale} is derived as per [Formula \(43\)](#):

$$g_{\text{scale}} = (Y_{\text{display}})^{\frac{1-g}{g}} \tag{43}$$

- The output mapped samples R_{scene}, G_{scene}, B_{scene} are finally obtained by scaling R_n, G_n, B_n by g_{scale} as per [Formula \(44\)](#):

$$\begin{cases} R_{\text{scene}} = \text{Clip3}(0, 1, g_{\text{scale}} * R_n) \\ G_{\text{scene}} = \text{Clip3}(0, 1, g_{\text{scale}} * G_n) \\ B_{\text{scene}} = \text{Clip3}(0, 1, g_{\text{scale}} * B_n) \end{cases} \tag{44}$$

8.2.2.4 Colour representation conversion: R'G'B' to Y'CbCr

Conversion from the R'G'B' to the Y'CbCr representation is commonly performed using a 3×3 matrix conversion process of the form in [Formula \(45\)](#):

$$\begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} w_{YR} & w_{YG} & w_{YB} \\ w_{CbR} & w_{CbG} & w_{CbB} \\ w_{CrR} & w_{CrG} & w_{CrB} \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \mathbf{W}^* \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (45)$$

where w_{YR} , w_{YG} , w_{YB} , w_{CbR} , w_{CbG} , w_{CbB} , w_{CrR} , w_{CrG} , and w_{CrB} are constants defined as per [Formulae \(46\)](#) to [\(54\)](#):

$$w_{YR} = 0.2627 \quad (46)$$

$$w_{YG} = 0.6780 \quad (47)$$

$$w_{YB} = 0.0593 \quad (48)$$

$$w_{CbR} = -\frac{w_{YR}}{2 * (1 - w_{YB})} = -0.139630063 \quad (49)$$

$$w_{CbG} = -\frac{w_{YG}}{2 * (1 - w_{YB})} = -0.360369937 \quad (50)$$

$$w_{CbB} = 0.5 \quad (51)$$

$$w_{CrR} = 0.5 \quad (52)$$

$$w_{CrG} = -\frac{w_{YG}}{2 * (1 - w_{YR})} = -0.459785705 \quad (53)$$

$$w_{CrB} = -\frac{w_{YB}}{2 * (1 - w_{YR})} = -0.040214295 \quad (54)$$

An alternative method to perform the same conversion process is presented in Rec. ITU-R BT.2020-2 and Rec. ITU-R BT.2100-1, where the chroma components are computed after the conversion of the luma component according to [Formulae \(55\)](#) and [\(56\)](#) as follows:

$$Cb = \frac{B' - Y'}{\alpha} \quad (55)$$

$$Cr = \frac{R' - Y'}{\beta} \quad (56)$$

with $\alpha = 2 * (1 - w_{YB})$ and $\beta = 2 * (1 - w_{YR})$.

This can be seen as equivalent to the matrix presented in [Formula \(45\)](#).

8.2.3 Encoding process

After pre-processing, the data is ready for compression. The quantization methodologies, perceptual luma quantization, and chroma QP offset derivation, presented in ITU-T H.Supp15 | ISO/IEC TR 23008-14, can be used in the case of an HLG signal. The parameter settings for these two quantization methods are anticipated to differ for the case of an HLG signal.

In the case of an HEVC bitstream, the syntax elements in Table 8 are conveyed in the VUI syntax branch of the SPS defined in Rec. ITU-T H.265 | ISO/IEC 23008-2:2017, Annex E. In the case of an AVC bitstream, the syntax elements in Table 9 are conveyed in the VUI syntax branch of the SPS defined in Rec. ITU-T H.264 | ISO/IEC 14496-10:2014, Annex E. 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 is expected to be signalled in the VUI as being an SDR signal. This is done by indicating a transfer_characteristics value corresponding to an SDR OETF, e.g. Rec. ITU-R BT.2100-1 with code value 14. For HDR HLG compliant displays, an ATC SEI message can be used, with the alternative_transfer_characteristics syntax element indicating the usage of HLG, as described in subclause 7.3.5 and illustrated in Table 10. One ATC SEI message is expected to be conveyed with each RAPAU.

Table 8 — 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 9 — 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 is expected to be set equal to 18 in both HEVC and AVC cases, which corresponds to HLG.

Table 10 — Suggested usage of ATC SEI message for Y'CbCr 4:2:0 HLG 10-bits signal with SDR compatibility

Syntax element	Location	Suggested value
preferred_transfer_characteristics	alternative_transfer_characteristics()	18

8.2.4 Decoding process

When the bitstream is an HEVC bitstream, the decoding process as in Rec. ITU-T H.265 | ISO/IEC 23008-2 is performed.

When the bitstream is an AVC bitstream, the decoding process in Rec. ITU-T H.264 | ISO/IEC 14496-10 is performed.

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

8.2.5 HLG post-decoding conversion

8.2.5.1 General

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

- a) a chroma up-conversion that converts data from Y'CbCr 4:2:0 to Y'CbCr 4:4:4, for instance as described in subclause 7.5.5.2,
- b) a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, as described in subclause 7.5.5.3,
- c) a colour representation conversion from Y'CbCr 4:4:4 to R'G'B' 4:4:4, as described in subclause 8.2.6,
- d) a conversion using the actual HLG EOTF of the input R'G'B' 4:4:4; the actual HLG EOTF is the concatenation of the inverse reference HLG OETF and of the HLG OETF; this last step aims at adapting the content to the rendering display, by applying a suitable end-to-end transfer function of the whole system (OETF); this transfer function is dependent on the viewing environment (peak brightness of the display, brightness of the surround).

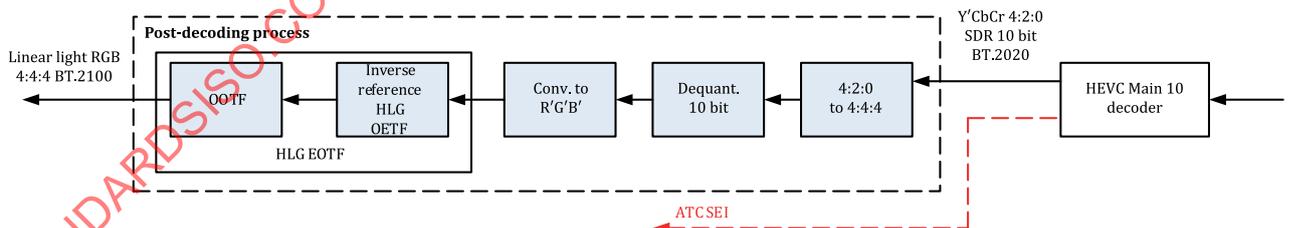


Figure 7 — Conventional HLG Y'CbCr post-decoding process system diagram

8.2.6 Colour representation conversion: Y'CbCr to R'G'B'

Conversion from the Y'CbCr representation back to the R'G'B' representation can be performed using [Formula \(57\)](#):

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} a_{RY} & a_{RCb} & a_{RCr} \\ a_{GY} & a_{GCb} & a_{GCr} \\ a_{BY} & a_{BCb} & a_{BCr} \end{bmatrix} \begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \mathbf{A} * \begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} \tag{57}$$

where **A** is the inverse matrix of **W**, provided in subclause [8.2.2.4](#), as per [Formulae \(58\)](#) to [\(63\)](#).

$$a_{RY} = a_{GY} = a_{BY} = 1 \tag{58}$$

$$a_{RCb} = a_{BCr} = 0 \tag{59}$$

$$a_{RCr} = 2 * (1 - w_{YR}) = 1.474\ 6 \tag{60}$$

$$a_{GCb} = -\frac{2 * w_{YB} * (1 - w_{YB})}{w_{YG}} = -0.164\ 553\ 126\ 843\ 660 \tag{61}$$

$$a_{GCr} = -\frac{2 * w_{YR} * (1 - w_{YR})}{w_{YG}} = -0.571\ 353\ 126\ 843\ 660 \tag{62}$$

$$a_{BCr} = 2 * (1 - w_{YB}) = 1.881\ 4 \tag{63}$$

If high precision is possible, the matrix in [Formula \(64\)](#) can be used:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.000\ 000\ 000\ 000\ 00 & 1.474\ 600\ 000\ 000\ 00 \\ 1 & -0.164\ 553\ 126\ 843\ 66 & -0.571\ 353\ 126\ 843\ 66 \\ 1 & 1.881\ 400\ 000\ 000\ 00 & 0.000\ 000\ 000\ 000\ 00 \end{bmatrix} \tag{64}$$

For systems with limited precision, a lower precision representation of the above matrix, e.g., retaining only 6 digits of precision, could be used instead.

8.3 Approach 2: usage of SEI messages for “dynamic” bitstream SDR backward compatibility

8.3.1 General

This subclause is focused on dynamic bitstream SDR backward compatibility support using a single-layer coding framework. [Figure 8](#) provides an overview of a single-layer distribution system enabling bitstream SDR backward compatibility by means of dynamic metadata. 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. An optional master input SDR signal can also be accommodated for parameter settings as described in subclause [8.3.5](#). Metadata is 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, intended to be as close as possible to the original HDR signal. The decoded SDR video can be directly rendered on an SDR display without adaptation.

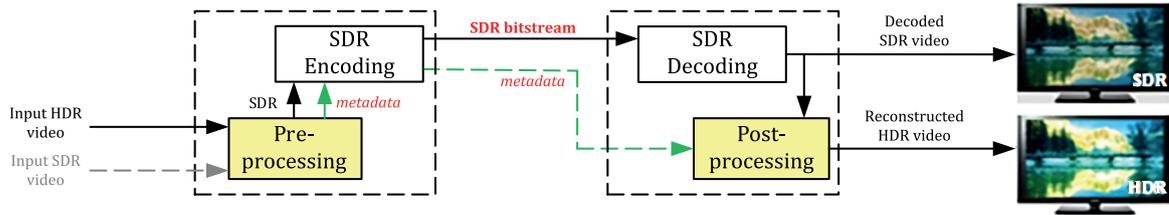


Figure 8 — Bitstream SDR backward compatibility using single layer coding system

Various approaches using HEVC SEI messages for this task are possible; three are illustrated. Two of them use the CRI SEI message, and operate in the Y'CbCr 4:2:0 and Y'CbCr 4:4:4 domains, respectively. The third one uses the TMI SEI message, and operates in the R'G'B' 4:4:4 domain. Other configurations are also possible.

Examples of HDR/WCG distribution systems that support bitstream SDR compatibility are illustrated in [Figure 9](#) (using CRI SEI message and applying in Y'CbCr 4:2:0 domain), [Figure 10](#) (using CRI SEI message and applying in Y'CbCr 4:4:4 domain), and [Figure 11](#) (using TMI SEI message and applying in R'G'B' 4:4:4 domain).

The bitstream SDR backward compatibility functionality is also addressed in Rec. ETSI TS 103 433. ETSI TS 103 433 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 (see Reference [18]). 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 in a user-data registered SEI message.

8.3.2 CRI applied in Y'CbCr 4:2:0 domain

In the first design, the pre-encoding conversion process (top of [Figure 9](#)) 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 (see Reference [19]):

- a conversion from an input linear RGB 4:4:4 representation to a non-linear representation, using the inverse reference PQ EOTF described in subclause [7.4.2](#), applied separately on the R, G and B components;
- a colour format conversion from non-linear PQ R'G'B' 4:4:4 signal to Y'CbCr 4:4:4 as described in subclause [8.2.2.4](#);
- a conversion step that converts a floating-point to a fixed-point representation (i.e. 10 bits), narrow range, as described in subclause [7.5.2.5](#);
- a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, for instance as described in subclause [7.5.2.6](#), resulting in a PQ 10-bits 4:2:0 Y'CbCr signal (PQ10);
- a DRA step that applies three different mapping 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 Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 transfer characteristics and Rec. ITU-R BT.2020-2 colour primaries, is then encoded, using an HEVC Main 10 compliant encoder. The DRA mapping functions can be implemented in the shape of 1D-LUTs that directly apply to the PQ10 Y', Cb and Cr components, in 4:2:0 format. The inverse DRA functions can be conveyed using the three Pre-LUTs of the CRI SEI message.

After HEVC Main 10 compliant decoding, the decoded 4:2:0 Y'CbCr signal has Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 transfer characteristics and Rec. ITU-R BT.2020-2 colour primaries. The post-decoding

inverse conversion processing (bottom of Figure 9) is the inverse of the pre-encoding processing. It is made of the following steps:

- a) an inverse DRA process, converting the SDR 10-bits Y'CbCr 4:2:0 signal into a PQ10 compatible signal using the inverse DRA mapping functions;
- b) a chroma up-conversion that converts data from Y'CbCr 4:2:0 to Y'CbCr 4:4:4, as for instance as described in subclause 7.5.5.2;
- c) a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, as described in subclause 7.5.5.3;
- d) a colour representation conversion from Y'CbCr 4:4:4 to R'G'B' 4:4:4, as described in subclause 8.2.6;
- e) a conversion from the input R'G'B' 4:4:4 to linear RGB 4:4:4, using the reference PQ EOTF described in subclause 7.4.2, applied on the R', G' and B' components separately.

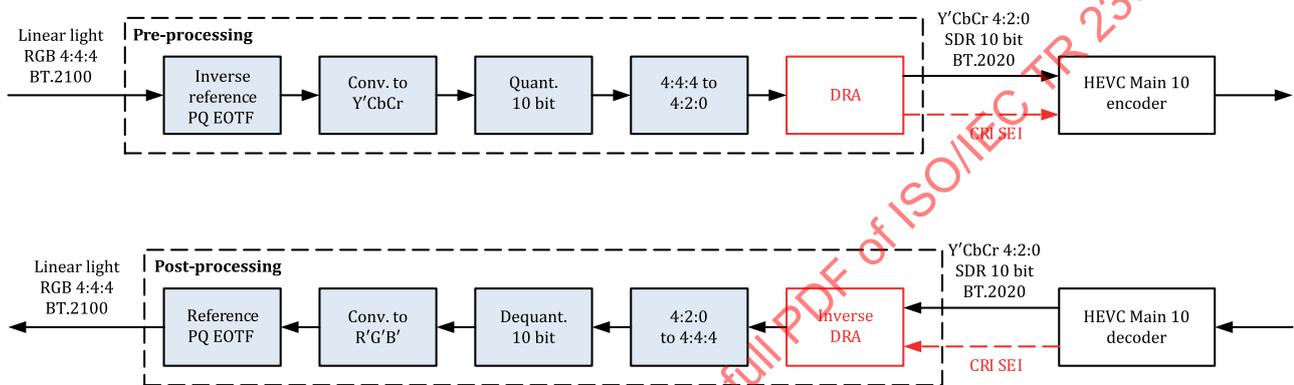


Figure 9 — Pre-processing (top) and post-processing (down) for conversion from SDR/Rec. ITU-R BT.2020-2 to HDR/Rec. ITU-R BT.2100-1 with CRI applied in Y'CbCr 4:2:0 domain

8.3.3 CRI applied in Y'CbCr 4:4:4 domain

The second design (Figure 10), 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:

- a) a conversion step from an input linear RGB 4:4:4 representation to a non-linear representation, using the inverse reference PQ EOTF, described in subclause 7.4.2, applied on the R, G and B components separately;
- b) a colour format conversion from non-linear PQ R'G'B' 4:4:4 signal to Y'CbCr 4:4:4 as described in subclause 8.2.2.4;
- c) a conversion step that converts a floating-point to a fixed-point representation (i.e. 10 bits), narrow range, as described in subclause 7.5.2.5;
- d) 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;
- e) a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, for instance as described in subclause 7.5.2.6, resulting in a 10-bits SDR 4:2:0 Y'CbCr signal.

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

- a) a chroma up-conversion component that converts data from Y'CbCr 4:2:0 to Y'CbCr 4:4:4, for instance as described in subclause 7.5.5.2;

- b) 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;
- c) a conversion step that converts a fixed-point representation, i.e. 10 bits, to a floating-point representation, as described in subclause 7.5.5.3;
- d) a colour representation conversion from $Y'CbCr$ 4:4:4 to $R'G'B'$ 4:4:4, as described in subclause 8.2.6;
- e) a conversion from the input $R'G'B'$ 4:4:4 to linear RGB 4:4:4, using the reference PQ EOTF described in subclause 7.4.2, applied on the R, G and B' components separately.

Working in $Y'CbCr$ 4:4:4 domain enables using the complete CRI model to convey the inverse DRA metadata, 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/Rec. ITU-R BT.709-6 colour primaries while the input HDR colour primaries are Rec. ITU-R BT.2100-1. In this case the three-by-three matrix can be of use for converting from Rec. ITU-R BT.2020-2 colour gamut to Rec. ITU-R BT.709-6 colour gamut.

NOTE In the case of non-backward compatible applications, the CRI SEI message can also be used to work in a colour space other than $Y'CbCr$.

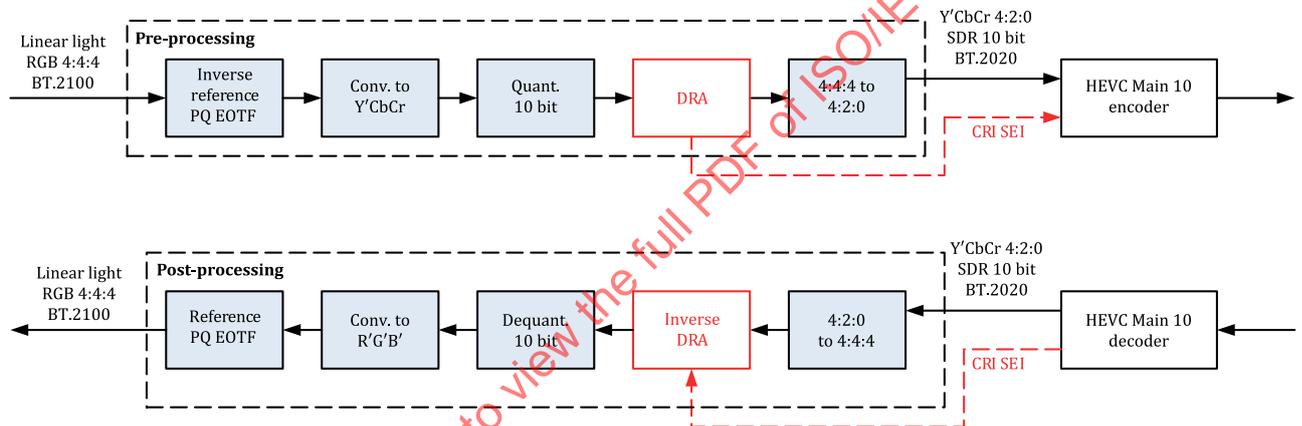


Figure 10 — Pre-processing (top) and post-processing (down) for conversion from SDR/Rec. ITU-R BT.2020-2 to HDR/Rec. ITU-R BT.2100-1 with CRI applied in $Y'CbCr$ 4:4:4 domain

8.3.4 TMI applied in $R'G'B'$ 4:4:4 domain

In the third design (Figure 11), 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:

- a) a conversion step from an input linear RGB 4:4:4 representation to a non-linear representation, using the inverse reference PQ EOTF, described in subclause 7.4.2, applied on the R, G and B components separately;
- b) a conversion step that converts a floating-point to a fixed-point M-bits full range representation (typically M being set to 14 or 16 bits), as described in subclause 7.5.2.5 with the bit-depth parameter b set to M;
- c) 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 Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 transfer characteristics and Rec. ITU-R BT.2020-2 colour primaries);
- d) a colour format conversion from non-linear $R'G'B'$ 4:4:4 signal to $Y'CbCr$ 4:4:4 as described in subclause 8.2.2.4;

- e) a chroma down-conversion that converts data from 4:4:4 to 4:2:0, following a similar process, for instance as described in subclause 7.5.2.6, resulting in an SDR 4:2:0 Y'CbCr signal;
- f) a re-quantization to 10 bits.

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

- a) a re-quantization to M-bits of the input decoded signal having Rec. ITU-R BT.709-6/Rec. ITU-R BT.2020-2 transfer characteristics and Rec. ITU-R BT.2020-2 colour primaries;
- b) a chroma up-conversion that converts data from Y'CbCr 4:2:0 to Y'CbCr 4:4:4, for instance as described in subclause 7.5.5.2;
- c) a colour representation conversion from Y'CbCr 4:4:4 to R'G'B' 4:4:4, as described in subclause 8.2.6;
- d) 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;
- e) a conversion component that converts a M-bits fixed-point representation, to a floating-point representation, as described in subclause 7.5.5.3;
- f) a conversion from the input R'G'B' 4:4:4 to linear RGB 4:4:4, using the reference PQ EOTF described in subclause 7.4.2, applied on the R', G' and B' components separately.

The inverse DRA mapping 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. One of the pre- or post-LUTs of the CRI SEI message can also be used to model this function.

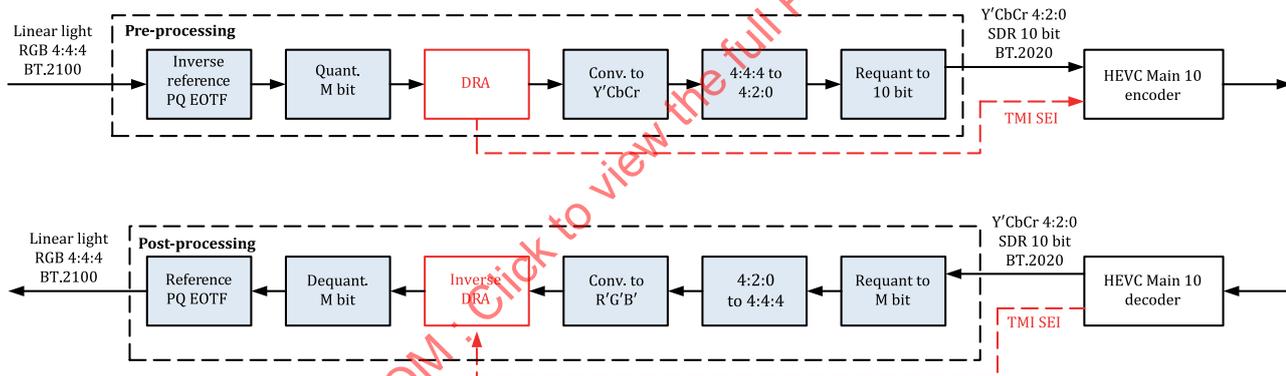


Figure 11 — Pre-processing (top) and post-processing (down) for conversion from SDR/Rec. ITU-R BT.2020-2 to HDR/Rec. ITU-R BT.2100-1 with TMI applied in RGB 4:4:4 domain

8.3.5 Derivation of DRA functions

In the three examples presented in subclauses 8.3.2, 8.3.3 and 8.3.4, 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. Examples of automatic derivation algorithms can be found in Reference [20].

It is expected to generate at least one CRI or TMI SEI message per scene. In scenes with large temporal changes, it can be preferable to guarantee smooth SDR rendering to generate several CRI or TMI SEI messages per scene (up to one per picture). At least, one CRI or TMI SEI message is expected to be conveyed with each RAPAU.

8.3.6 Settings with colour remapping information SEI message

For the adaptive bitstream SDR backward compatibility use case described in subclause 8.3, with the implementation shown in Figure 9, 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 9), 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 U_{SDR} and V_{SDR} as per Formulae (65) to (67):

$$Y'_{\text{HDR}} = \text{PreLUT}_0 [Y'_{\text{SDR}}] \quad (65)$$

$$U_{\text{HDR}} = \text{PreLUT}_1 [U_{\text{SDR}}] \quad (66)$$

$$V_{\text{HDR}} = \text{PreLUT}_2 [V_{\text{SDR}}] \quad (67)$$

When the conversion process applies to the decoded Y'CbCr 4:2:0 signal upsampled to 4:4:4 (implementation shown in Figure 10), 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} , U_{SDR} and V_{SDR} as per Formulae (68) to (70):

$$\begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} \text{PreLUT}_0 [Y'_{\text{SDR}}] \\ \text{PreLUT}_1 [U_{\text{SDR}}] \\ \text{PreLUT}_2 [V_{\text{SDR}}] \end{bmatrix} \quad (68)$$

$$\begin{bmatrix} Y_2 \\ U_2 \\ V_2 \end{bmatrix} = M_{3 \times 3} \begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} \quad (69)$$

$$\begin{bmatrix} Y'_{\text{HDR}} \\ U_{\text{HDR}} \\ V_{\text{HDR}} \end{bmatrix} = \begin{bmatrix} \text{PostLUT}_0 [Y_2] \\ \text{PostLUT}_1 [U_2] \\ \text{PostLUT}_2 [V_2] \end{bmatrix} \quad (70)$$

Table 11 indicates suggested parameters settings of HEVC Main 10 bitstreams to properly render the decoded video on displays compatible with SDR/Rec. ITU-R BT.2020-2 representation (using VUI information) and to conduct the HDR reconstruction to PQ10 compatible representation (using CRI information).

Table 11 — 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 (Rec. ITU-R BT.709-6) or 9 (Rec. ITU-R BT.2020-2) depending on SDR target
transfer_characteristics	vui_parameters()	14 (SDR OETF)