International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Series H
Supplement 18
(10/2017)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

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

ITU-T H-series Recommendations – Supplement 18

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Telepresence	H.420–H.429
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
VEHICULAR GATEWAYS AND INTELLIGENT TRANSPORTATION SYSTEMS (ITS)	
Architecture for vehicular gateways	H.550–H.559
Vehicular gateway interfaces	H.560–H.569
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619
Advanced multimedia services and applications	H.620–H.629
Ubiquitous sensor network applications and Internet of Things	H.640–H.649
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	
General aspects	H.700–H.719
IPTV terminal devices	H.720–H.729
IPTV middleware	H.730–H.739
IPTV application event handling	H.740–H.749
IPTV metadata	H.750–H.759
IPTV multimedia application frameworks	H.760–H.769
IPTV service discovery up to consumption	H.770–H.779
Digital Signage	H.780–H.789
E-HEALTH MULTIMEDIA SERVICES AND APPLICATIONS	
Personal health systems	H.810–H.819
Interoperability compliance testing of personal health systems (HRN, PAN, LAN, TAN and WAN)	H.820–H.859
Multimedia e-health data exchange services	H.860–H.869

For further details, please refer to the list of ITU-T Recommendations.

Supplement 18 to ITU-T H-series Recommendations

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

Summary

Supplement 18 to ITU-T H-series Recommendations relates to high dynamic range (HDR) and wide colour gamut (WCG) video distribution, based on the high efficiency video coding (HEVC) standard (ITU-T H.265 | ISO/IEC 23008-2), and when applicable on the advanced video coding (AVC) standard (ITU-T H.264 | ISO/IEC 14496-10). This Supplement serves several purposes. It provides a survey of the signalling mechanisms (video usability information indicators and supplemental enhancement information messages) for handling HDR/WCG video with HEVC, and when applicable, with AVC. It describes example processing and coding chains for HDR/WCG video in different ITU-R BT.2100 representations. It discusses single- and dual-layer approaches to enable backward compatibility with standard dynamic range (SDR) systems. SDR systems are defined as legacy decoding systems that do not have the ability to detect and properly display HDR/WCG video content. It describes methods for adapting HDR/WCG video content for use with display technology having different degrees and types of dynamic range and colour gamut capability.

This Supplement is technically aligned with ISO/IEC TR 23008-15 and it complements the material provided in Supplement ITU-T H Suppl. 15 | ISO/IEC TR 23008-14, entitled *Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics*.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H Suppl. 18	2017-10-27	16	11.1002/1000/13441

Keywords

AVC, HDR, HEVC, WCG.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this publication, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this publication is voluntary. However, the publication may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the publication is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the publication is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this publication may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the publication development process.

As of the date of approval of this publication, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this publication. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

Page

1 Scope ... 1

2 References ... 1

3 Definitions ... 2

4 Abbreviations and acronyms ... 3

5 Conventions ... 4
- 5.1 General ... 4
- 5.2 Arithmetic operators ... 5
- 5.3 Bit-wise operators ... 5
- 5.4 Assignment operators ... 5
- 5.5 Relational, logical, and other operators ... 6
- 5.6 Mathematical functions ... 6
- 5.7 Order of operations ... 6

6 Overview ... 7

7 HEVC signalling mechanisms applicable to HDR/WCG video ... 8
- 7.1 General ... 8
- 7.2 VUI syntax elements ... 8
- 7.3 SEI messages applicable for HDR/WCG video ... 9
- 7.4 Overview of PQ and HLG transfer functions ... 11
- 7.5 IC_TC_P colour representation ... 13

8 Bitstream SDR backward compatibility with single-layer coding ... 20
- 8.1 General ... 20
- 8.2 Approach 1: usage of HLG for "static" bitstream SDR backward compatibility ... 20
- 8.3 Approach 2: usage of SEI messages for "dynamic" bitstream SDR backward compatibility ... 24

9 Bitstream SDR backward compatibility with dual-layer SHVC coding ... 30
- 9.1 General ... 30
- 9.2 Encoding and decoding stages ... 30

10 Display adaptation ... 31
- 10.1 General ... 31
- 10.2 Case of display SDR backward compatibility ... 32

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 Supplement relates to HDR/WCG video coding and distribution, using single-layer or dual-layer coding, with the high efficiency video coding (HEVC), and when applicable, the advanced video coding (AVC) signalling. This Supplement 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_TC_P 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.

Supplement 18 to ITU-T H-series Recommendations

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

1 Scope

This Supplement reviews approaches for processing and coding of high dynamic range/ wide colour gamut (HDR/WCG) video content. The purpose of this Supplement is to provide a set of publicly referenceable methods for the operation of advanced video coding (AVC) or high efficiency video coding (HEVC) video coding systems adapted for compressing HDR/WCG video for consumer distribution applications. This Supplement 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 red, green, blue (RGB) linear light representation video signals into video signals with IC_TC_P colour representation and perceptual quantizer (PQ) transfer function, or with Y′CbCr colour representation and hybrid log-gamma (HLG) transfer function (IC_TC_P, PQ and HLG are defined in Rec. ITU-R BT.2100). Some high-level approaches for compressing these signals using either the AVC or HEVC video coding standards are provided. A description of post-decoding processing steps is also included for converting back to a linear light, 4:4:4 RGB representation. This Supplement 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 colour remapping information (CRI) and tone mapping information (TMI) SEI messages, using dual-layer approach with the Scalable Main 10 profile of HEVC and an SDR compatible base layer. Finally, this Supplement 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.

The Supplement is technically aligned with ISO/IEC TR 23008-15 and it complements the material provided in [ITU-T H Suppl.15] | [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 References

– Recommendation ITU-T H.264 (in force), *Advanced video coding for generic audiovisual services*.

– ISO/IEC 14496-10 (in force), *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*.

– Recommendation ITU-T H.265 (in force), *High Efficiency Video Coding*.

– ISO/IEC 23008-2 (in force), *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding*.

– Supplement ITU-T H.Sup15 (2017), *Conversion and Coding Practices for HDR/WCG Y'CbCr 4:2:0 Video with PQ Transfer Characteristics*.

– Recommendation ITU-R BT.709-6 (2015), *Parameter values for the HDTV standards for production and international programme exchange*.

– Recommendation ITU-R BT.2020-2 (2015), *Parameter values for ultra-high definition television systems for production and international programme exchange*.

– Recommendation ITU-R BT.2100-1 (2017), *Image parameter values for high dynamic range television for use in production and international programme exchange*.

– Report ITU-R BT.2390-0 (2016), *High dynamic range television for production and international programme exchange*.

– ARIB STD-B67 Association of Radio Industries and Businesses (2015), *Essential parameter values for the extended image dynamic range television (EIDRTV) system for programme production*.

– CEA-861.3 (2015), *HDR Static Metadata Extensions*, https://standards.cta.tech/kwspub/published_docs/CEA-861.3-Preview.pdf.

– ETSI TS 103 433 (in force), *High-Performance Single Layer Directly Standard Dynamic Range (SDR) Compatible High Dynamic Range (HDR) System for use in Consumer Electronics devices (SL-HDR1)*.

– ISO 11664-1 (2007), *Colorimetry – Part 1: CIE standard colorimetric observers*.

– ISO/IEC TR 23008-14 (2017), *Information technology -- High efficiency coding and media delivery in heterogeneous environments – Part 14: Conversion and coding practices for HDR/WCG Y'CbCr 4:2:0 video with PQ transfer characteristics*.
– SMPTE ST 2084 (2014), *High Dynamic Range Electro-Optical Transfer Function for Mastering Reference Display*.
– SMPTE ST 2086 (2014), *Mastering Display Color Volume Metadata for High Luminance and Wide Color Gamut Images*.
– SMPTE ST 2094 (2016), *Dynamic Metadata for Color Transforms of HDR and WCG Images*.
– [Baroncini2016] V. Baroncini, W. Husak, A. K. Ramasubramonian, J. Sole, G. J. Sullivan, Y.-K. Wang, Y. Ye (2016), *SHVC Supplemental Verification Test Report*, JCTVC-X1004, 24th JCT-VC meeting, Geneva, CH, May 2016.
– [Bordes2016] P. Bordes, P. Andrivon, X. Li, Y. Ye (2016), Overview of Color Gamut Scalability, *IEEE Trans. Circuits and Systems for Video Technology*, Mar. 2016.
– [Borer2015] T. Borer, A. Cotton (2015), A Display Independent High Dynamic Range Television System, *BBC Research and Development white paper*, WHP 309, Sep. 2015, http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP309.pdf.
– [Boyce2015] J.M. Boyce, Y.Ye, A.K. Ramasubramonian (2015), Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard, *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 26, Issue 1, pp 20-34, July 2015.
– [Francois2016] E. François, L. van de Kerkhof, A Single-Layer HDR Video Coding Framework With SDR Compatibility, *in Proceedings of the IBC*, Amsterdam, Sept 2016.
– [Minoo2016] K. Minoo, T. Lu, P. Yin, L. Kerovsky, D. Rusanovskyy, E. Francois, *Description of the reshaper parameters derivation process in ETM reference software*, JCTVC-W0031, 24th JCT-VC Meeting, San Diego, USA, Feb. 2016.
– [Pindoria2016] M. Pindoria and S. Thompson, Image Adaptation Requirements for High Dynamic Range Video Under Reference and Non-Reference Viewing Conditions, *in Proceedings of the IBC*, Amsterdam, Sept 2016.
– [Pu2016a] F. Pu, T. Lu, P. Yin, T. Chen, W. Husak (2016), *AHG13: on Luma Adjustment*, JCTVC-X0043, 24th JCT-VC meeting, Geneva, CH, May 2016.
– [Pu2016b] F. Pu, T. Lu, P. Yin, T. Chen, W. Husak (2016), *AHG13:* IC_TC_P *Compression Using HEVC Main 10*, JCTVC-X0050, 24th JCT-VC meeting, Geneva, CH, May 2016.
– [Pytlarz2016] J.A. Pytlarz, K.D. Thurston, D. Brooks, P. Boon, R. Atkins (2016), Real time cross-mapping of high dynamic range images, *in Proceedings of the IBC*, Amsterdam, Sept 2016.
– [Rusanovskyy2016] D. Rusanovskyy, A.K. Ramasubramonian, D. Sansli, J. Sole, M. Karczewicz (2016), *Usage of CRI for guided mapping (dynamic range adaptation)*, JCTVC-X0060, 24th JCT-VC meeting, Geneva, CH, May 2016.
– [Topiwala2016] P. Topiwala, W. Dai, M. Krishnan, *AHG14: Tone Mapping and Related SEIs for HDR Coding*, JCTVC-Y0042, 25th JCT-VC Meeting, Chengdu, CN, Oct. 2016.
– [Ye2016] Y. Ye, V. Baroncini, Y.-K. Wang (2016), *Verification Test Report for Scalable HEVC (SHVC)*, JCTVC-W1004, 23rd JCT-VC meeting, San Diego, CA, USA, Feb. 2016.
– [Zink2015] M. Zink, M. D. Smith (2015), Managing HDR content production and display device capabilities, *Proc. International Broadcasting Convention 2015 (IBC 2015)*, Amsterdam, NL, Sep. 2015, http://www.ibc.org/__media/PDF/michael_zink_managing_hdr_content_production_and_display_device_capabilities.pdf.

3 Definitions

This Supplement defines the following terms. The definitions used in the AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10) and HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2) standards also apply.

3.1 **dynamic range adaptation (DRA)**: A mapping process to convert content from one colour volume to another colour volume.

3.2 **electro-optical transfer function (EOTF)**: The electro-optical transfer function, which converts the non-linear video signal into a quantity of output linear light (e.g., light emitted by a display).

3.3 **full range**: A range in a fixed-point (integer) representation that spans the full range of values that could be expressed with that bit depth, such that, for 10-bit signals, black corresponds to code value 0 and peak white corresponds to code value 1023 for Y′, as per the full range definition from Rec. ITU-R BT.2100.

3.4 **hybrid log-gamma (HLG)**: One of the two sets of transfer functions specified in Rec. ITU-R BT.2100; Rec. ITU-R BT.2100 specifies the reference HLG OETF and its inverse.

3.5 **narrow range**: A range in a fixed-point (integer) representation that does not span the full range of values that could be expressed with that bit depth 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, as per the narrow range definition from [ITU-R BT.2100].

3.6 **opto-electronic transfer function (OETF)**: The opto-electronic transfer function, which converts source input linear optical intensity (e.g., light input to a camera) into the non-linear video signal.

3.7 **opto-optical transfer function (OOTF)**: The opto-optical transfer function, which has the role of applying the "rendering intent" on video signal; in general, an OOTF is a concatenation of an OETF, artistic adjustments and an EOTF.

3.8 **perceptual quantizer (PQ)**: One of the two sets of transfer functions specified in Rec. ITU-R BT.2100; Rec. ITU-R BT.2100 specifies the reference PQ EOTF and its inverse.

3.9 **random access point access unit (RAPAU)**: An 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)**: A specified EOTF for use under specific viewing environment, named the reference viewing environment.

3.11 **reference opto-electronic transfer function (reference OETF)**: A 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; Table 3 of Rec. ITU-R BT.2100 provides reference viewing environment parameters for HDR programme material.

4 Abbreviations and acronyms

This Supplement uses the following abbreviations and acronyms. The abbreviations and acronyms used in the AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10) and HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2) standards also apply.

ATC Alternative Transfer Characteristics

AVC Advanced Video Coding (as 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

DRA Dynamic Range Adaptation

EOTF Electro-Optical Transfer Function

FIR Finite Impulse Response

HDR High Dynamic Range

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

HLG Hybrid Log-Gamma (as specified in Rec. ITU-R BT.2100)

IC_TC_P Alternative colour space representation to Y′CbCr (as specified in Rec. ITU-R BT.2100)

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

LUT Look-up Table

MAD Mean Absolute Difference

MDCV Mastering Display Colour Volume

NCL Non-Constant Luminance

OETF Opto-Electronic Transfer Function

OOTF Opto-optical Transfer Function

PQ Perceptual Quantizer (as specified in Rec. ITU-R BT.2100)

PQ10 HDR content representation that utilizes the Rec. ITU-R BT.2100 colour primaries, the Rec. ITU-R BT.2100 reference PQ EOTF, and the Rec. ITU-R BT.2100 Y′CbCr colour space representation with 10 bits per sample in the 4:2:0 chroma sampling format

QP Quantization Parameter

RAPAU Random Access Point Access Unit

RGB Red, Green, and Blue (colour system using these 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 The CIE 1931 colour space; Y corresponds to the luminance signal

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

5 Conventions

5.1 General

The mathematical operators used in this Supplement 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

The following arithmetic operators are defined as follows:

+	Addition
−	Subtraction (as a two-argument operator) or negation (as a unary prefix operator)
*	Multiplication, including matrix multiplication
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.
÷	Used to denote division in mathematical formulae where no truncation or rounding is intended.
$\frac{x}{y}$	Used to denote division in mathematical formulae where no truncation or rounding is intended.
$\sum_{i=x}^{y} f(i)$	The summation of f(i) with i taking all integer values from x up to and including y.
x % y	Modulus. Remainder of x divided by y, defined only for integers x and y with x >= 0 and y > 0.

5.3 Bit-wise operators

The following bit-wise operators are defined as follows:

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

5.4 Assignment operators

The following assignment operators are defined as follows:

=	Assignment operator
++	Increment, i.e., x++ is equivalent to x = x + 1; when used in an array index, evaluates to the value of the variable prior to the increment operation.
− −	Decrement, i.e., x− − is equivalent to x = x − 1; when used in an array index, evaluates to the value of the variable prior to the decrement operation.
+=	Increment by amount 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

The following operators are defined as follows:

==	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

The following mathematical functions are defined as follows:

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

$\mathrm{Ceil}(x)$ the smallest integer greater than or equal to x.

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

$\mathrm{EOTF}_{PQ}(x)$ the reference PQ EOTF used to convert a non-linear light PQ representation to a linear light representation.

$\mathrm{Exp}(x) = e^x$ where e is Euler's base constant 2.718 281 828....

$\mathrm{Floor}(x)$ the largest integer less than or equal to x.

$\mathrm{iEOTF}_{PQ}(x)$ the inverse reference PQ EOTF used to convert a linear light representation to a non-linear light representation.

$\mathrm{iOETF}_{HLG}(x)$ the inverse reference HLG OETF used to convert a non-linear light representation to a scene-referred linear light representation.

$\mathrm{Ln}(x)$ the natural logarithm of x (the base-e logarithm, where e is natural logarithm base constant 2.718 281 828...).

$\mathrm{Log10}(x)$ the base-10 logarithm of x.

$$\mathrm{Max}(x, y) = \begin{cases} x & ; \quad x > y \\ y & ; \quad \text{otherwise} \end{cases}$$

$\mathrm{OETF}_{HLG}(x)$ the reference HLG OETF used to convert a scene-referred linear light representation to a non-linear light representation.

$$\mathrm{Round}(x) = \mathrm{Sign}(x) * \mathrm{Floor}(\mathrm{Abs}(x) + 0.5)$$

$$\mathrm{Sign}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x = 0 \\ -1 & ; \quad x < 0 \end{cases}$$

$$\mathrm{Sqrt}(x) = \sqrt{x}$$

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 Supplement 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"
"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

The Supplement is structured as follows.

- clause 7 reviews identified signalling mechanisms of HEVC, 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 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., ultra-high definition (UHD) SDR display with ITU-R BT.2020 colour primaries). In this Supplement, it is defined in two modes:

- In HDR/WCG distribution systems that support "bitstream" SDR backward compatibility, the decoded video signal from a standard-compliant decoder (e.g., HEVC Main 10 decoder) can be directly displayed on an SDR-capable display without adaptation. Two categories of "bitstream" SDR backward compatibility 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 reference HLG opto-electronic transfer function (OETF) and Rec. ITU-R BT.2100 colour primaries, that can be directly displayed on an HDR-capable display or an SDR-capable display, without adaptation. In this context, the HDR processing chain is a static function, not dependent on the input video data.
- In "dynamic" bitstream SDR backward compatibility, the decoded video is an SDR signal. A post-processing step can be further used to reconstruct an HDR signal, 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 ITU-R BT.2100 inverse reference PQ electro-optical transfer function (EOTF) and ITU-R BT.2100 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 to display 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 ITU-R BT.2100 PQ signal (denoted PQ10 in this Supplement), 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 non-constant luminance (NCL) Y′CbCr 4:2:0 video with ITU-R BT.2100 PQ transfer characteristics are outside of scope of this Supplement. These aspects are specifically addressed in Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 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 Supplement. The PQ, HLG transfer functions, and the IC_TC_P colour representation are also described.

This clause is structured as follows.

– clause 7.2 reviews VUI signalling applicable to HDR/WCG video.

– clause 7.3 reviews SEI messages applicable to HDR/WCG video.

– clause 7.4 provides an overview of PQ and HLG transfer functions.

– clause 7.5 provides a description of IC_TC_P colour representation, including conversion and coding practices related to HDR/WCG video signals represented with IC_TC_P 4:2:0 video with ITU-R BT.2100 inverse reference PQ EOTF.

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

7.2 VUI syntax elements

By design, metadata signalled in syntax elements of VUI in HEVC is not mandatory for constructing the luma or chroma samples by the decoding process, and may be ignored by the decoder. However, such syntax elements 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 Tables E.3, E.4 and E.5 of the HEVC specification respectively. Table 7-1 and Table 7-2 below provide values of VUI syntax elements that indicate usage of Rec. ITU-R BT.2100 representation of the video signal, including matrix coefficients associated with Rec. ITU-R BT.2100 (same as those associated with Rec. ITU-R BT.2020).

Rec. ITU-R BT.2100 specifies HDR-TV image parameters for use in production and international programme exchange. It defines two sets of transfer functions: PQ and HLG. RGB colour primaries are defined identically as in Rec. ITU-R BT.2020. Rec. ITU-R BT.2100 describes two different luminance and colour difference signal representations: 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 7-1 for HEVC. RGB colour primaries are indicated using colour_primaries syntax element (set equal to 9 for ITU-R BT.2100/ITU-R BT.2020 colour primaries). Colour representation is indicated using matrix_coeffs syntax element, as described in Table 7-2 for HEVC.

Table 7-1 – Values of transfer_characteristics indication in VUI in HEVC

	PQ	**HLG**
transfer_characteristics	16	18

Table 7-2 – Values of matrix_coeffs indication in VUI in HEVC

	NCL Y′CbCr	**CI IC_TC_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 clauses 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 clauses 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 (see ISO 11664-1), and provides the minimum and maximum linear light luminance of the monitor (expressed in candelas per square metre, denoted cd/m^2). 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 coded layer-wise video sequence (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 [Zink2015]. The CLL SEI message persists until the end of the CLVS. The HEVC specification 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. The HEVC specification 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 clause 8.2. The ATC SEI message persists until the end of the CLVS. The HEVC specification 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 $\mathrm{LUT_{TM}}$. 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 below.

$$\begin{bmatrix} \mathrm{R_{out}} \\ \mathrm{G_{out}} \\ \mathrm{B_{out}} \end{bmatrix} = \begin{bmatrix} \mathrm{LUT_{TM}}[\mathrm{R_{in}}] \\ \mathrm{LUT_{TM}}[\mathrm{G_{in}}] \\ \mathrm{LUT_{TM}}[\mathrm{B_{in}}] \end{bmatrix} \quad (7\text{-}1)$$

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 less 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 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 the CRI mapping must be 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.

The following formulae 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):

$$\begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} \text{PreLUT}_0[R_{in}] \\ \text{PreLUT}_1[G_{in}] \\ \text{PreLUT}_2[B_{in}] \end{bmatrix} \tag{7-2}$$

$$\begin{bmatrix} R_2 \\ G_2 \\ B_2 \end{bmatrix} = M_{3x3} \begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} \tag{7-3}$$

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} \text{PostLUT}_0[R_2] \\ \text{PostLUT}_1[G_2] \\ \text{PostLUT}_2[B_2] \end{bmatrix} \tag{7-4}$$

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 less 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 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), which compensates for the difference in tonal perception between the environment of the camera and that of the display as further explained in Annex 1 of Rec. ITU-R BT.2100. 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 Annex 1 of Rec. ITU-R BT.2100.

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

In Rec. ITU-R BT.2100, PQ and HLG 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:

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

– 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:

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

Functions "iEOTF" and "iOETF" are used in this Supplement to represent EOTF^{-1} and OETF^{-1}, respectively.

NOTE 2 – Informative methods for transcoding between PQ and HLG signals are described in Annex 2 of Rec. ITU-R BT.2100 and in Rec. ITU-R BT.2390.

7.4.2 Reference PQ EOTF

The inverse reference PQ EOTF is described in Table E.4 of the HEVC specification 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 as:

$$V = \mathrm{iEOTF_{PQ}}(\,L_o\,) = \left(\frac{c_1 + c_2 * L_o^n}{1 + c_3 * L_o^n}\right)^m \quad (7\text{-}5)$$

where c_1, c_2, c_3, m, and n are constants defined as follows:

$$c_1 = c_3 - c_2 + 1 = 3424 \div 4096 = 0.835\,937\,5 \quad (7\text{-}6)$$

$$c_2 = 2413 \div 128 = 18.851\,562\,5 \quad (7\text{-}7)$$

$$c_3 = 299 \div 16 = 18.687\,5 \quad (7\text{-}8)$$

$$m = 2523 \div 32 = 78.843\,75 \quad (7\text{-}9)$$

$$n = 1305 \div 8192 = 0.159\,301\,757\,812\,5 \quad (7\text{-}10)$$

The peak value of 1 for L_o is ordinarily intended to correspond to an intensity level of 10 000 cd/m^2, while the value of 0 for L_o is ordinarily intended to correspond to an intensity level of 0 cd/m^2.

The reference PQ EOTF, $\mathrm{EOTF_{PQ}}(\ )$, is described as the exact inverse of Formula 7-5. The linear light intensity signal L_o is computed from the non-linear representation V, which takes values in the range [0, 1], as follows:

$$L_o = \mathrm{EOTF_{PQ}}(\,V\,) = \left(\frac{\mathrm{Max}\left(\left(V^{1/m} - c_1\right), 0\right)}{c_2 - c_3 * V^{1/m}}\right)^{1/n} \quad (7\text{-}11)$$

A plot of the inverse reference PQ EOTF is shown in Figure 7-1.

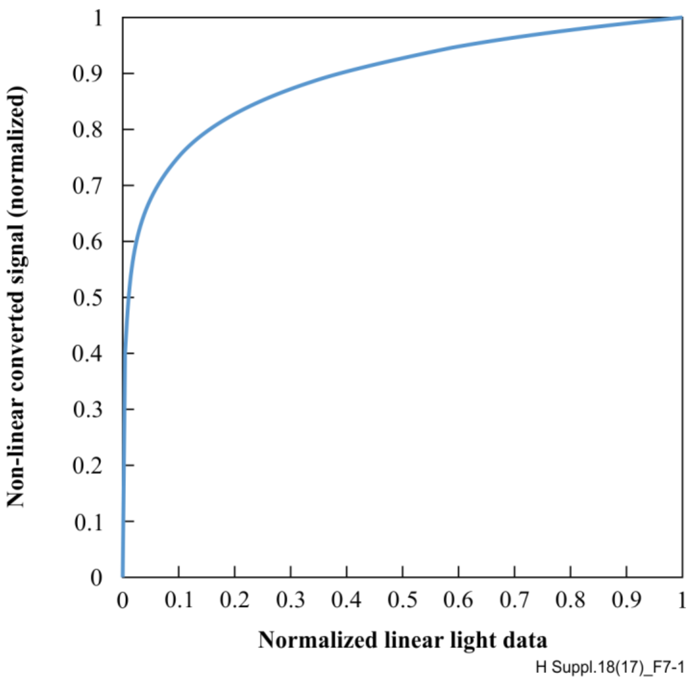

Figure 7-1 – Graph of the inverse reference PQ EOTF (iEOTFP$_{PQ}$)

7.4.3 Reference HLG OETF

The reference HLG OETF, $\mathrm{OETF_{HLG}}(\ )$, as described in Table E.4 of the HEVC specification 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, or B, the following applies:

$$V = OETF_{HLG}(L_c) = a * Ln(12 * L_c - b) + c \quad \text{for } 1 >= L_c > (1 \div 12) \tag{7-12}$$

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

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

$$a = 0.178\ 832\ 77 \tag{7-13}$$

$$b = 1 - 4 * a = 0.284\ 668\ 92 \tag{7-14}$$

$$c = 0.5 - a * Ln(4 * a) = 0.559\ 910\ 73 \tag{7-15}$$

A plot of the reference OETF for HLG is shown in Figure 7-2.

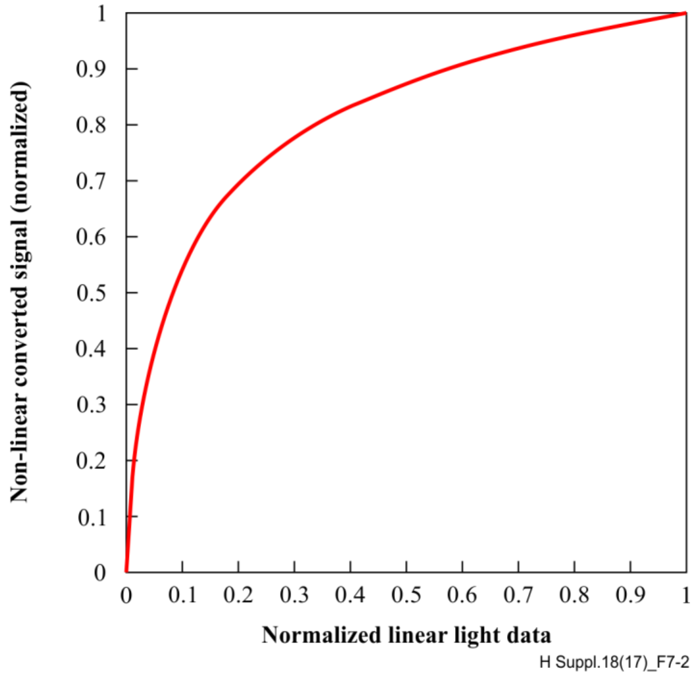

Figure 7-2 – Graph of the reference HLG OETF ($OETF_{HLG}$)

The inverse of the reference HLG OETF is formulated as follows:

$$L_c = iOETF_{HLG}(V) = 1 \div 12 * (Exp((V - c) \div a) + b) \quad \text{for } 1 >= V > (1 \div 2) \tag{7-16}$$

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

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, Report ITU-R BT.2390 and [Borer2015].

7.5 IC_TC_P colour representation

7.5.1 General

An overview of CI IC_TC_P colour representation is given in Rec. ITU-R BT.2100 and Report ITU-R BT.2390. This clause 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 ITU-R BT.2100 colour primaries to a 10-bits, narrow range, inverse reference PQ EOTF, 4:2:0, IC_TC_P representation.

The HDR/WCG system described in this clause 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 | Technical Report ISO/IEC 23008-14. It consists of four major stages:

– a pre-encoding stage consisting of several pre-processing processes (clause 7.5.2),

– an encoding stage (clause 7.5.3),

– a decoding stage (clause 7.5.4), and

– a post-decoding stage consisting of several post-processing processes (clause 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 clause is to highlight the conversion and coding differences between BT.2100 PQ IC_TC_P and BT.2100 PQ Y′CbCr signal. More detail on common parts of these two systems are provided specifically in Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14.

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

7.5.2 Pre-encoding process

7.5.2.1 General

The pre-encoding process described in this Supplement includes the following components, as presented in Figure 7-3:

a) a conversion component from a linear RGB data representation to a linear LMS data representation, described in clause 7.5.2.2,
b) a conversion component from a linear LMS data representation to a non-linear data representation using the inverse reference PQ EOTF, described in clause 7.5.2.3,
c) a colour format conversion component that converts data to the IC_TC_P colour representation, described in clause 7.5.2.4,
d) a conversion component that converts a floating-point to a fixed-point representation (e.g., 10 bits), following the process described in clause 7.5.2.5, and
e) a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, following the process described in clause 7.5.2.6.

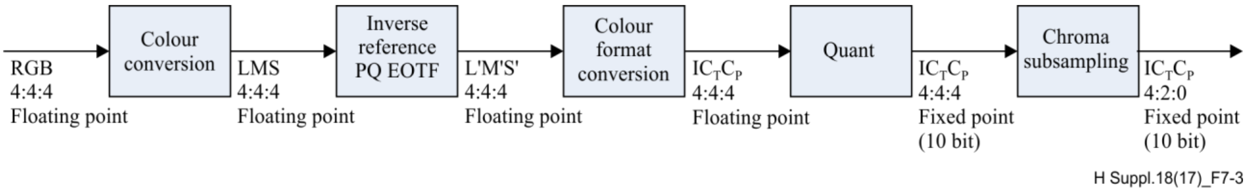

Figure 7-3 – Conventional PQ IC_TC_P pre-encoding process system diagram

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 3x3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2100.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \frac{1}{4096}\begin{bmatrix} 1688 & 2146 & 262 \\ 683 & 2951 & 462 \\ 99 & 309 & 3688 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (7\text{-}17)$$

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

The inverse reference PQ EOTF (Formula 7-5) is applied to all L, M, and S linear light samples, where each component is a number between 0 (representing 0 cd/m^2) and 1 (representing 10 000 cd/m^2). This results in their non-linear light counterparts L′ , M′, and S′ as follows.

$$L' = iEOTF_{PQ}(L) \quad (7\text{-}18)$$

$$M' = iEOTF_{PQ}(M) \quad (7\text{-}19)$$

$$S' = iEOTF_{PQ}(S) \quad (7\text{-}20)$$

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 3x3 matrix conversion process.

$$\begin{bmatrix} I \\ C_T \\ C_P \end{bmatrix} = \frac{1}{4096} \begin{bmatrix} 2048 & 2048 & 0 \\ 6610 & -13613 & 7003 \\ 17933 & -17390 & -543 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \quad (7\text{-}21)$$

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:

$$D' = \text{Clip3}(\ 0, (1 << b) - 1, \text{Round}(E' * \text{scale} + \text{offset})\ ) \quad (7\text{-}22)$$

where E′ is the floating-point representation of a particular component and D′ is the resulting quantized value using b bits. In this Supplement, 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:

$$\text{scale} = 219 * (1 << (b - 8)\ ) \quad (7\text{-}23)$$

$$\text{offset} = 1 << (b - 4) \quad (7\text{-}24)$$

and the scale and offset for the chroma components are set as:

$$\text{scale} = 224 * (1 << (b - 8)\ ) \quad (7\text{-}25)$$

$$\text{offset} = 1 << (b - 1) \quad (7\text{-}26)$$

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:

$$\text{scale} = (1 << b) - 1 \quad (7\text{-}27)$$

$$\text{offset} = 0 \quad (7\text{-}28)$$

and the scale and offset for the chroma components are set as:

$$\text{scale} = (1 << b) - 1 \quad (7\text{-}29)$$

$$\text{offset} = 1 << (b - 1) \quad (7\text{-}30)$$

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 Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14. This process is explained below.

It is anticipated that a considerable amount of consumer electronics conversion systems would 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:

$$y[n] = \sum_{i=-N}^{N} b_i * x[n + i] \quad (7\text{-}31)$$

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 7-3 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 7-3 – Filter example for chroma down-sampling

Filter	Filter coefficients		
	b_{-1}	b_0	b_1
f_0	1÷8	6÷8	1÷8
f_1	1÷4	2÷4	1÷4

7.5.2.7 Comments on IC_TC_P Chroma down-conversion

Converting the IC_TC_P video data from a 4:4:4 representation to a 4:2:0 representation follows the chroma down-conversion described in clause 7.5.2.6.

It is observed in [Pu2016a] that the closed loop luma adjustment method described in Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14, which aims at reducing NCL Y′CbCr chroma down-sampling artefacts, does not prove to be beneficial for IC_TC_P chroma down-conversion. This can be explained by the constant luminance property of IC_TC_P. As shown in Figure 7-4, the achromatic axis of Y′ deviates substantially from luminance especially when colour goes more saturated. On the other hand, the achromatic axis of IC_TC_P (I) corresponds very closely with luminance (see also [Pytlarz2016]). The constant luminance property of IC_TC_P can reduce chroma leakage errors that can be introduced when spatially sub-sampling the chroma components (such as the 4:2:0 widely used for compression) compared with NCL Y′CbCr.

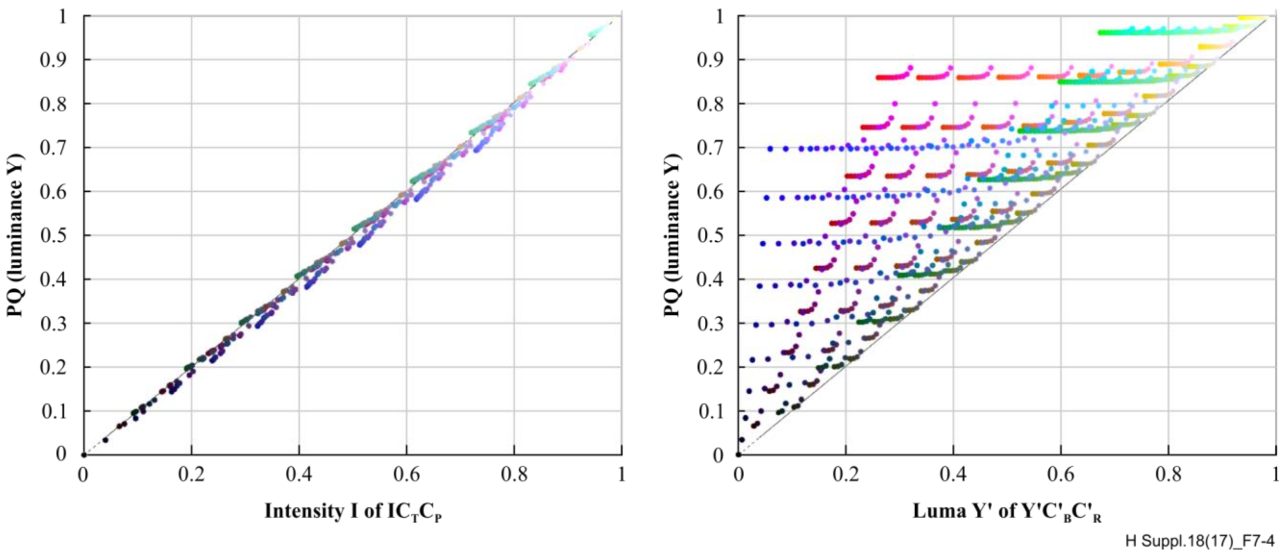

Figure 7-4 – Luminance correlation with IC_TC_P and Y′CbCr colour representations

7.5.3 Encoding process

7.5.3.1 General

After pre-processing, the data is ready for compression. Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 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 quantization parameter (QP) offset, are described. For the HDR/WCG 10-bit, narrow range, PQ, 4:2:0, IC_TC_P data coming out from pre-processing step, those two methodologies are still valid. [Pu2016b] shows that IC_TC_P 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_TC_P encoding is to configure the encoder settings to follow the bitrate allocation behavior of Y′CbCr [Pu2016b]. An example of encoder setting is described in clause 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 decision process, and may have been tuned based on SDR content characteristics. Given, however, the earlier observations about the differences in the characteristics between SDR and HDR/WCG content, these processes 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 Supplement.

7.5.3.2 Example of encoder setting

Two main observations are made in [Pu2016b]. Firstly, the Y′ component in Y′CbCr and I component in IC_TC_P have very similar compression characteristics in terms of variance. This results in similar coding bits given same luma quantizer. [Pu2016b] suggests that techniques designed to improve coding efficiency of Y′ component (e.g., perceptual luma quantization from Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14) can be used for the I component directly. Secondly, for colour components, C_T and C_P have a higher variance than Cb and Cr, thus modifications may be applied for technologies designed to improve chroma coding efficiency, such as the adaptive chroma QP offset model in Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14.

In the following example of chroma QP offset settings for IC_TC_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 Cb, corresponding to C_T, and Cr, 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:

$$\mathrm{QPoffsetCb} = \mathrm{Clip3}\left(-12, 12, \mathrm{CbOffset} + \mathrm{Clip3}\left(-12, 0, \mathrm{Round}\left(\mathrm{c_{cb}} * (\mathrm{k} * \mathrm{QP} + \mathrm{l})\right)\right)\right) \tag{7-32}$$

$$\mathrm{QPoffsetCr} = \mathrm{Clip3}\left(-12, 12, \mathrm{CrOffset} + \mathrm{Clip3}\left(-12, 0, \mathrm{Round}\left(\mathrm{c_{cr}} * (\mathrm{k} * \mathrm{QP} + \mathrm{l})\right)\right)\right) \tag{7-33}$$

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

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

Finally, $\mathrm{k} = -9.46$ and $\mathrm{l} = 10$.

7.5.3.3 HEVC encoding

When creating the HEVC bitstream it is suggested to set syntax elements to the values listed in Table 7-4 in the sequence parameter set (SPS) of the bitstream. The syntax elements in Table 7-4 are conveyed in the video usability information syntax branch of the SPS defined in Annex E of the HEVC specification. They may also be duplicated and carried in various application-layer headers.

Table 7-4 – Suggested settings for HEVC encoding of IC_TC_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_coeffs	vui_parameters( )	14
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2

For HDR/WCG content represented with the colour primaries of Rec. ITU-R BT.2100, Rec. ITU-R BT.2100 inverse reference PQ EOTF and IC_TC_P colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec. ITU-R BT.709 colour primaries and Rec. ITU-R BT.709/BT.2020 OETF. Chroma QP adjustment, as described in clause 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 clause 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 the AVC bitstream it is suggested to set syntax elements to the values listed in Table 7-5 in the SPS of the bitstream. The syntax elements in Table 7-5 are conveyed in the video usability information syntax branch of the SPS defined in Annex E of the AVC specification. They may also be duplicated and carried in various application-layer headers.

Table 7-5 – Suggested settings for AVC encoding of IC_TC_P 4:2:0 PQ 10-bit signal

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

For HDR/WCG content represented with the colour primaries of Rec. ITU-R BT.2100, Rec. ITU-R BT.2100 inverse reference PQ EOTF and IC_TC_P colour representation, the video characteristics is typically different compared to the video characteristics of SDR content represented with Rec. ITU-R BT.709 colour primaries and Rec. ITU-R BT.709/BT.2020 OETF. Chroma QP adjustment, as described in clause 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 clause 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 the HEVC specification is performed.

When the bitstream is an AVC bitstream, the decoding process in the AVC specification 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:

a) a chroma up-conversion component that converts data from 4:2:0 to 4:4:4, following the process described in clause 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 clause 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 clause 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 clause 7.5.5.5,

e) a conversion component from a linear LMS data representation to a linear RGB data representation, described in clause 7.5.5.6.

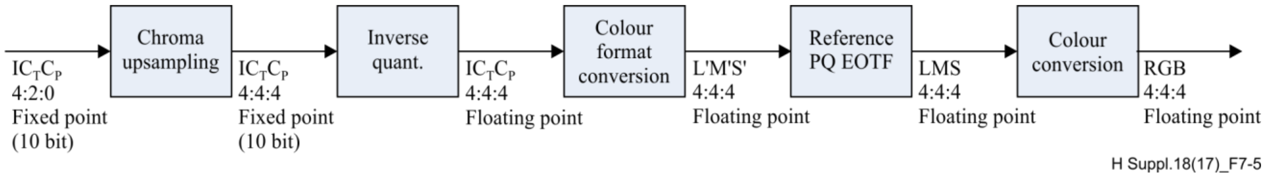

Figure 7-5 – Conventional PQ IC_TC_P post-decoding process system diagram

7.5.5.2 Chroma upscaling

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 Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 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 7-6 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 7-6 – Two-phase chroma resampling filter

Phase p	Interpolation filter coefficients			
	f_C[p, −1]	f_C[p, 0]	f_C[p, 1]	f_C[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 clause 7.5.2.5. In particular, a fixed-point precision value can be converted to a floating-point precision value using the following formula:

$$E' = \text{Clip3}(\ \text{minE}, \text{maxE}, (D' - \text{offset}) \div \text{scale}\ ) \tag{7-34}$$

The exact same values for scale and offset as in clause 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 3x3 matrix conversion process.

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} 1 & 0.008\,609\,037\,037\,933 & 0.111\,029\,625\,003\,026 \\ 1 & -0.008\,609\,037\,037\,933 & -0.111\,029\,625\,003\,026 \\ 1 & 0.560\,031\,335\,710\,679 & -0.320\,627\,174\,987\,319 \end{bmatrix} \begin{bmatrix} I \\ C_T \\ C_P \end{bmatrix} \tag{7-35}$$

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 follows:

$$L = EOTF_{PQ}(\ L'\ ) \tag{7-36}$$

$$M = EOTF_{PQ}(\ M'\ ) \tag{7-37}$$

$$S = EOTF_{PQ}(\ S'\ ) \tag{7-38}$$

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 3x3 matrix conversion process, where RGB colour primaries are in accordance with Rec. ITU-R BT.2100 and each component, L, M, S, R, G, and B, is a number between 0.0 (representing no light) and 1.0 (representing 10 000 cd/m^2).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.436\,606\,694\,333\,079 & -2.506\,452\,118\,656\,270 & 0.0698\,454\,243\,231\,91 \\ -0.791\,329\,555\,598\,929 & 1.983\,600\,451\,792\,291 & -0.192\,270\,896\,193\,362 \\ -0.025\,949\,899\,690\,593 & -0.098\,913\,714\,711\,726 & 1.124\,863\,614\,402\,319 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (7\text{-}39)$$

8 Bitstream SDR backward compatibility with single-layer coding

8.1 General

This clause addresses the use cases of bitstream SDR compatibility. Two approaches are described: Approach 1 which uses a static HDR opto-electrical transfer function (OETF) to deliver "native" bitstream SDR backward compatibility; 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 ITU-R BT.2100 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 (see Rec. ITU-R BT.2100) has been designed to provide some level of backward compatibility with the legacy SDR systems. The legacy system is taken to be the SDR systems not needing colour gamut conversion (for example, not needing conversion from ITU-R BT.2100 to ITU-R BT.709 colour primaries) but needing only dynamic range conversion. HLG is suggested to be used with ITU-R BT.2100 colour primaries.

This clause 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 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:

1) a pre-encoding stage described in clause 8.2.2,

2) an encoding stage described in clause 8.2.3, including suggested signalling using VUI and ATF SEI message to enable the SDR compatibility feature,

3) a decoding stage described in clause 8.2.4, and

4) a post-decoding stage described in clause 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

HLG pre-encoding conversion process is illustrated in Figure 8-1. The conversion chain converts an input linear RGB 4:4:4 signal to HLG 10-bits Y′CbCr 4:2:0 signal. It consists of the following successive steps:

a) a pre-conversion step, denoted "inverse system gamma" in Figure 8-1 and described in clause 8.2.2.3, when the input HDR signal is display-light referred (see explanations in clause 7.4); when the input HDR signal is scene-light referred (e.g., from camera, see explanations in clause 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 clause 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 clause 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 clause 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 clause 7.5.2.6.

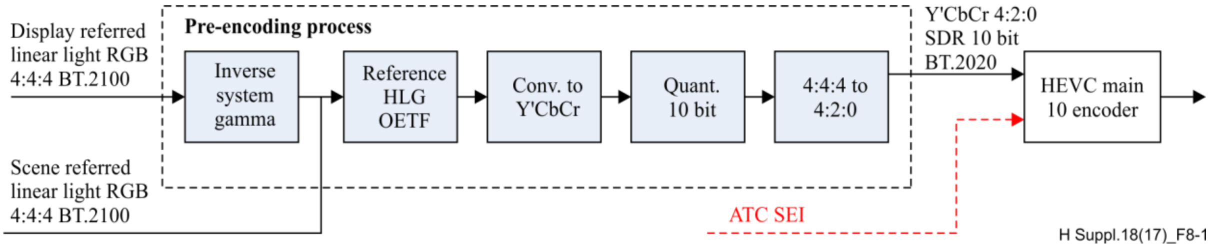

Figure 8-1 – Conventional HLG Y′CbCr pre-encoding process system diagram

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 Formulae 7-12 from clause 7.4.3 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. Clause 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 [Pindoria2016]. 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/m^2. 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 follows:

$$\begin{cases} R_n = \text{Max}(0, R_{display} \div P) \\ G_n = \text{Max}(0, G_{display} \div P) \\ B_n = \text{Max}(0, B_{display} \div P) \end{cases} \quad (8\text{-}1)$$

– The luminance $Y_{display}$ is derived from R_n, G_n, B_n values:

$$Y_{display} = a0 * R_n + a1 * G_n + a2 * B_n \quad (8\text{-}2)$$

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

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

$$g = 1.2 + 0.42 * \text{Log10}(P \div 1000) \quad (8\text{-}3)$$

– The value g_{scale} is derived as follows:

$$g_{scale} = \left(Y_{display}\right)^{\frac{1-g}{g}} \quad (8\text{-}4)$$

– 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 follows:

$$\begin{cases} R_{scene} = \text{Clip3}(0,1, g_{scale} * R_n) \\ G_{scene} = \text{Clip3}(0,1, g_{scale} * G_n) \\ B_{scene} = \text{Clip3}(0,1, g_{scale} * B_n) \end{cases} \quad (8\text{-}5)$$

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:

$$\begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} w_{YR} & w_{YG} & w_{YB} \\ w_{CbR} & w_{CbG} & w_{CbB} \\ w_{CrR} & w_{CrG} & w_{CrB} \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \mathbf{W} * \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (8\text{-}6)$$

where w_{YR}, w_{YG}, w_{YB}, w_{CbR}, w_{CbG}, w_{CbB}, w_{CbR}, w_{CbG}, and w_{CbB} are constants defined as follows:

$$w_{YR} = 0.262\,7 \quad (8\text{-}7)$$

$$w_{YG} = 0.678\,0 \quad (8\text{-}8)$$

$$w_{YB} = 0.059\ 3 \quad (8\text{-}9)$$

$$w_{CbR} = -\frac{w_{YR}}{2*(1-w_{YB})} = -0.139\ 630\ 063 \quad (8\text{-}10)$$

$$w_{CbG} = -\frac{w_{YG}}{2*(1-w_{YB})} = -0.360\ 369\ 937 \quad (8\text{-}11)$$

$$w_{CbB} = 0.5 \quad (8\text{-}12)$$

$$w_{CrR} = 0.5 \quad (8\text{-}13)$$

$$w_{CrG} = -\frac{w_{YG}}{2*(1-w_{YR})} = -0.459\ 785\ 705 \quad (8\text{-}14)$$

$$w_{CrB} = -\frac{w_{YB}}{2*(1-w_{YR})} = -0.040\ 214\ 295 \quad (8\text{-}15)$$

An alternative method to perform the same conversion process is presented in Rec. ITU-R BT.2020 and Rec. ITU-R BT.2100, where the chroma components are computed after the conversion of the luma component according to Formula 8-6 as follows:

$$\text{Cb} = \frac{\text{B}' - \text{Y}'}{alpha} \quad (8\text{-}16)$$

$$\text{Cr} = \frac{\text{R}' - \text{Y}'}{beta} \quad (8\text{-}17)$$

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

This can be seen as equivalent to the matrix presented in Formula 8-6.

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 Supplement ITU-T H.Sup15 | Technical Report ISO/IEC 23008-14, can be used in case of HLG signal. The parameter settings for these two quantization methods are anticipated to differ for the case of an HLG signal.

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., BT.2100 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 clause 7.3.5 and illustrated in **Error! Reference source not found.**. One ATC SEI message is expected to be conveyed with each random access point access unit (RAPAU).

Table 8-1 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 HLG 10-bits signal with SDR compatibility

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	2 (Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	9
transfer_characteristics	vui_parameters( )	14
matrix_coeffs	vui_parameters( )	9
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2

Table 8-2 – Suggested settings for AVC encoding of Y′CbCr 4:2:0 HLG 10-bits signal with SDR compatibility

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

NOTE – For HDR-only systems, with no SDR backward compatibility, the VUI transfer_characteristics element is expected to be set equal to 18 in both HEVC and AVC cases, which corresponds to HLG.

Table 8-3 – 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 the HEVC specification is performed.

When the bitstream is an AVC bitstream the decoding process in the AVC specification 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 8-2. 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 clause 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 clause 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 clause 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 OOTF; 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 (OOTF); this transfer function is dependent on the viewing environment (peak brightness of the display, brightness of the surround).

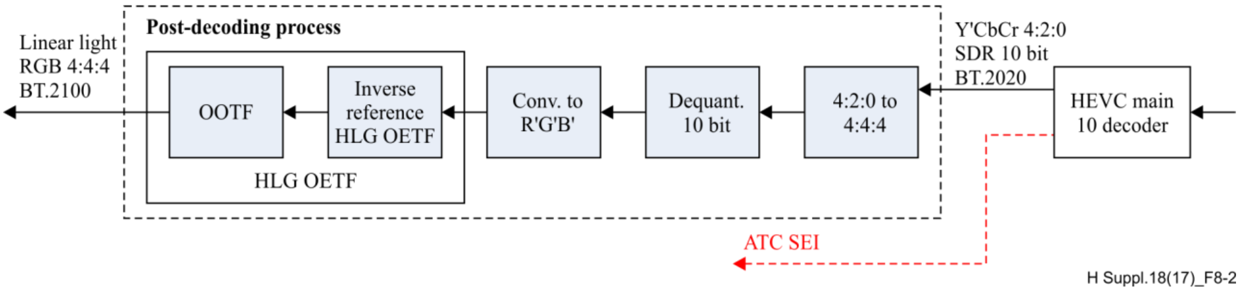

Figure 8-2 – Conventional HLG Y′CbCr post-decoding process system diagram

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 the following formula:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} a_{RY} & a_{RCb} & a_{RCr} \\ a_{GY} & a_{GCb} & a_{GCr} \\ a_{BY} & a_{BCb} & a_{BCr} \end{bmatrix} \begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \mathbf{A} * \begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} \quad (8\text{-}18)$$

where **A** is the inverse matrix of **W**, provided in clause 8.2.2.4.

$$a_{RY} = a_{GY} = a_{BY} = 1 \quad (8\text{-}19)$$

$$a_{RCb} = a_{BCr} = 0 \quad (8\text{-}20)$$

$$a_{RCr} = 2 * (1 - w_{YR}) = 1.474\,6 \quad (8\text{-}21)$$

$$a_{GCb} = -\frac{2 * w_{YB} * (1 - w_{YB})}{w_{YG}} = -0.164\,553\,126\,843\,660 \quad (8\text{-}22)$$

$$a_{GCr} = -\frac{2 * w_{YR} * (1 - w_{YR})}{w_{YG}} = -0.571\,353\,126\,843\,660 \quad (8\text{-}23)$$

$$a_{BCr} = 2 * (1 - w_{YB}) = 1.881\,4 \quad (8\text{-}24)$$

If high precision is possible, the following matrix can be used:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.000\,000\,000\,000\,00 & 1.474\,600\,000\,000\,00 \\ 1 & -0.164\,553\,126\,843\,66 & -0.571\,353\,126\,843\,66 \\ 1 & 1.881\,400\,\,000\,000\,00 & 0.000\,000\,000\,000\,00 \end{bmatrix} \quad (8\text{-}25)$$

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 clause is focused on dynamic bitstream SDR backward compatibility support using a single-layer coding framework. Figure 8-3 provides an overview of a single-layer distribution system enabling bitstream SDR backward compatibility with 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. 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.

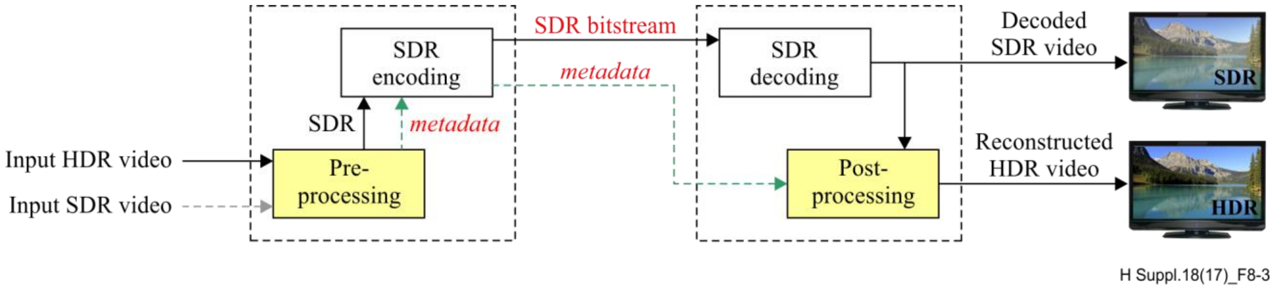

Figure 8-3 – Bitstream SDR backward compatibility using single layer coding system

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 8-4 (using CRI SEI message and applying in Y′CbCr 4:2:0 domain), Figure 8-5 (using CRI SEI message and applying in Y′CbCr 4:4:4 domain), and Figure 8-6 (using TMI SEI message and applying in R′G′B′ 4:4:4 domain).

The bitstream SDR backward compatibility functionality is also addressed in [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 [Francois2016]. 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 8-4) converts an input linear RGB 4:4:4 signal to SDR 10-bits Y′CbCr 4:2:0 signal by applying the following successive steps [Rusanovskyy2016]:

a) a conversion from an input linear RGB 4:4:4 representation to a non-linear representation, using the inverse reference PQ EOTF described in clause 7.4.2, applied separately on the R, G and B components,
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 clause 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 clause 7.5.2.5,
d) a chroma down-conversion component that converts data from 4:4:4 to 4:2:0, for instance as described in clause 7.5.2.6, resulting in a PQ 10-bits 4:2:0 Y′CbCr signal (PQ10),
e) 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 ITU-R BT.709/ITU-R BT.2020 transfer characteristics and ITU-R BT.2020 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 BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries. The post-decoding inverse conversion processing (bottom of Figure 8-4) 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 clause 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 clause 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 clause 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 clause 7.4.2, applied on the R′, G′ and B′ components separately.

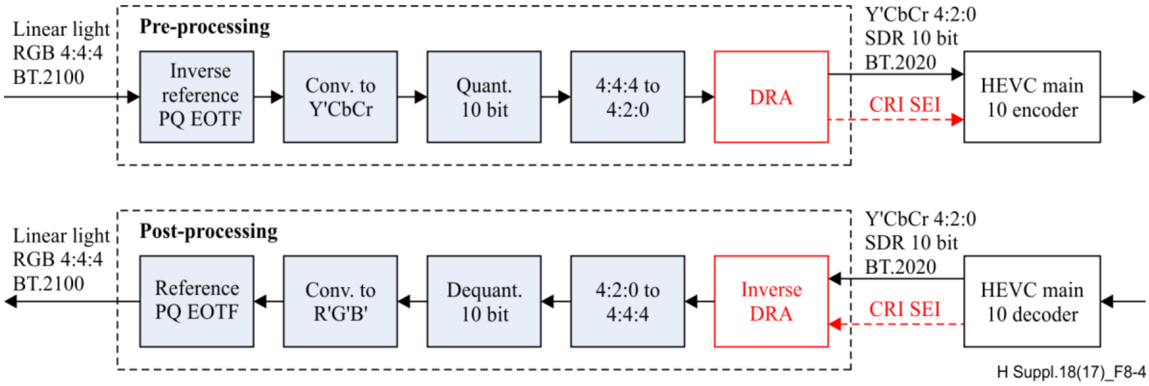

Figure 8-4 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2100 with CRI applied in Y′CbCr 4:2:0 domain

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

The second design (Figure 8-5), similarly to the first one, makes use of the CRI SEI message to perform the inverse DRA process. However, DRA and inverse DRA apply in Y′CbCr 4:4:4 domain, instead of Y′CbCr 4:2:0 domain.

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

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 clause 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 clause 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 clause 7.5.2.5,
d) a 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 clause 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 clause 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 clause 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 clause 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 clause 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 BT.709 colour primaries while the input HDR colour primaries are BT.2100. In this case the three-by-three matrix can be of use for converting from BT.2100 colour gamut to BT.709 colour gamut.

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

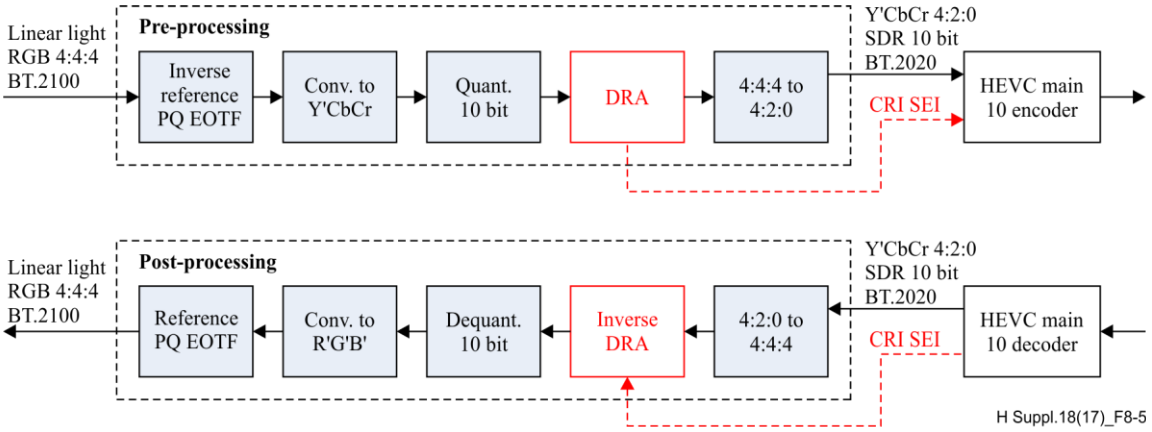

Figure 8-5 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2100 with CRI applied in Y′CbCr 4:4:4 domain

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

In the third design (Figure 8-6), the TMI SEI message is used to perform the inverse DRA and it applies to a R′G′B′ 4:4:4 signal.

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

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 clause 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 clause 7.5.2.5 with the bit-depth parameter b set to M,
c) a DRA step applying to the M-bits PQ R′G′B′ 4:4:4 signal to produce an SDR R′G′B′ 4:4:4 signal (having BT.709/BT.2020 transfer characteristics and BT.2020 colour primaries),
d) a colour format conversion from non-linear R′G′B′ 4:4:4 signal to Y′CbCr 4:4:4 as described in clause 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 clause 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 BT.709/BT.2020 transfer characteristics and BT.2020 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 clause 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 clause 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 clause 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 clause 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.

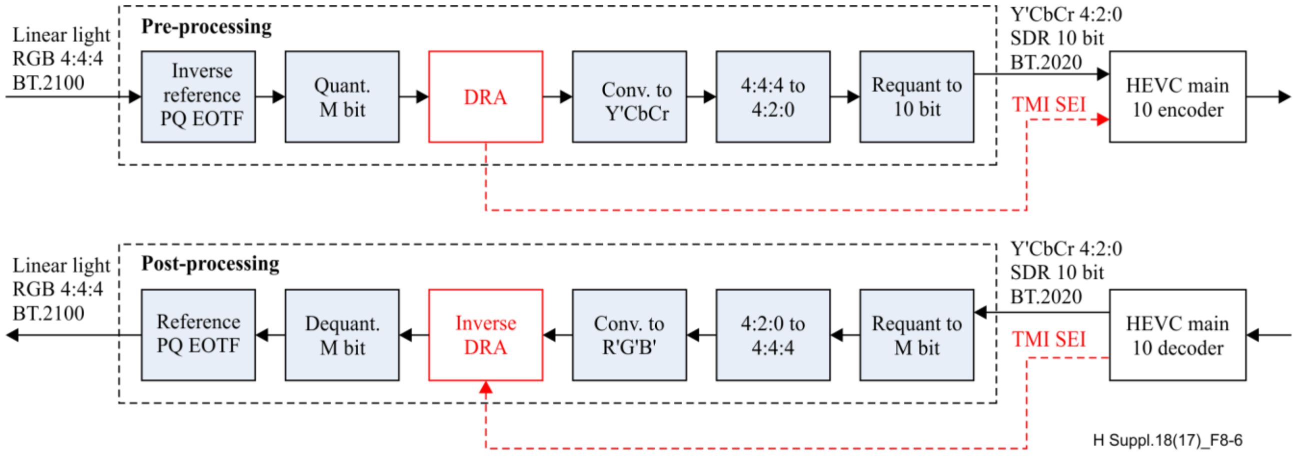

Figure 8-6 – Pre-processing (top) and post-processing (down) for conversion from SDR/BT.2020 to HDR/BT.2100 with TMI applied in RGB 4:4:4 domain

8.3.5 Derivation of DRA functions

In the three examples presented in clauses 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 [Minoo2016].

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 clause 8.3, with the implementation shown in Figure 8-4, the inverse DRA functions are coded using the three Pre-LUTs of a CRI SEI message. The conversion process directly applies to the decoded Y′CbCr 4:2:0 signal (Figure 8-4), and the three-by-three matrix and the Post-LUTs are not activated. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma sample Y'_{SDR} and each chroma samples U_{SDR} and V_{SDR} as follows:

$$Y'_{HDR} = \text{PreLUT}_0[Y'_{SDR}] \quad (8\text{-}26)$$

$$U_{HDR} = \text{PreLUT}_1[U_{SDR}] \quad (8\text{-}27)$$

$$V_{HDR} = \text{PreLUT}_2[V_{SDR}] \quad (8\text{-}28)$$

When the conversion process applies to the decoded Y′CbCr 4:2:0 signal upsampled to 4:4:4 (implementation shown in Figure 8-5), the full CRI model, made of the three Pre-LUTs, the intermediate three-by-three matrix, and the three Post-LUTs, can be activated. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma and chroma samples Y'_{SDR}, U_{SDR} and V_{SDR} as follows:

$$\begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} \text{PreLUT}_0[Y'_{SDR}] \\ \text{PreLUT}_1[U_{SDR}] \\ \text{PreLUT}_2[V_{SDR}] \end{bmatrix} \quad (8\text{-}29)$$

$$\begin{bmatrix} Y_2 \\ U_2 \\ V_2 \end{bmatrix} = M_{3x3} \begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} \quad (8\text{-}30)$$

$$\begin{bmatrix} Y'_{HDR} \\ U_{HDR} \\ V_{HDR} \end{bmatrix} = \begin{bmatrix} \text{PostLUT}_0[Y_2] \\ \text{PostLUT}_1[U_2] \\ \text{PostLUT}_2[V_2] \end{bmatrix} \quad (8\text{-}31)$$

Table 8-4 indicates suggested parameters settings of HEVC Main 10 bitstreams to properly render the decoded video on displays compatible with SDR/BT.2020 representation (using VUI information) and to conduct the HDR reconstruction to PQ10 compatible representation (using CRI information).

Table 8-4 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 10-bits signal with bitstream SDR compatibility using CRI

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	2 (Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0 (narrow range)
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	1 (BT.709) or 9 (BT.2020) depending on SDR target
transfer_characteristics	vui_parameters( )	14 (SDR OETF)
matrix_coeffs	vui_parameters( )	1 (BT.709) or 9 (BT.2020) depending on SDR target
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2
colour_remap_video_signal_info_present_flag	colour_remapping_info( )	1
colour_remap_transfer_function	colour_remapping_info( )	16 (BT.2100 PQ)
colour_remap_full_range_flag	colour_remapping_info( )	0
colour_remap_primaries	colour_remapping_info( )	9 (BT.2100/BT.2020)
colour_remap_matrix_coefficients	colour_remapping_info( )	9 (BT.2100/BT.2020)
colour_remap_output_bit_depth	colour_remapping_info( )	10

8.3.7 Settings with tone mapping information SEI message

For the use case described in clause 8.3, with the implementation example shown in Figure 8-6, the inverse DRA function is coded using one 1D-LUT of a TMI SEI message. Two implementation examples of the inverse DRA processes are provided below.

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

$$\begin{bmatrix} R_{HDR} \\ G_{HDR} \\ B_{HDR} \end{bmatrix} = \begin{bmatrix} LUT_{TM}[R_{SDR}] \\ LUT_{TM}[G_{SDR}] \\ LUT_{TM}[B_{SDR}] \end{bmatrix} \tag{8-32}$$

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

$$Y_{HDR\_PQ} = LUT_{TM}[Y_{SDR}] \tag{8-33}$$

For each sample value X = R, G or B:

$$X_{HDR} = \text{Clip3}\left(0, (\ 1 \ll M\ ) - 1, \frac{Y_{HDR\_PQ}}{Y_{SDR}} * X_{SDR}\right) \tag{8-34}$$

where M is the bit-depth of the output samples resulting from the inverse DRA application (corresponding to the TMI SEI message syntax element target_bit_depth). This second implementation can be beneficial to limit the colour hue shift resulting from the SDR-to-HDR conversion (see [Topiwala2016] for a fuller exposition).

Table 8-5 indicates suggested parameters settings of HEVC Main 10 bitstreams to properly render the decoded video on displays compatible with SDR/BT.2020 representation (using VUI information) and to conduct the HDR reconstruction to PQ10 compatible representation (using TMI information).

Table 8-5 – Suggested settings for HEVC encoding of Y′CbCr 4:2:0 10-bits signal with bitstream SDR compatibility using TMI

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	2 (Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0 (narrow range)
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	9
transfer_characteristics	vui_parameters( )	14 (SDR OETF)
matrix_coeffs	vui_parameters( )	9 (BT.2020)
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2
coded_data_bit_depth	tone_mapping_info ( )	M (typically 14 bits)
target_bit_depth	tone_mapping_info ( )	M (typically 14 bits)
tone_map_model_id	tone_mapping_info ( )	3 (piece-wise linear model)
num_pivots	tone_mapping_info ( )	32

9 Bitstream SDR backward compatibility with dual-layer SHVC coding

9.1 General

Scalable high efficiency video coding (SHVC), the scalable form of HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2, [Boyce2015]), is based on a multi-layer framework. In dual-layer case, one base layer and one enhancement layer are coded. The enhancement layer coding uses inter-layer prediction mechanisms that produce an inter-layer picture used as reference picture for temporal prediction. The HEVC Scalable Main 10 profile includes colour mapping prediction specifically designed for inter-layer prediction from a given colour volume (relating to the base layer video) to another one (relating to the enhancement layer video). In an SDR-backward compatible system, this profile can be used to convey in a dual layer-bitstream an SDR video and an HDR video. The input of the coding system is an SDR content as base layer, and its HDR version as enhancement layer. The SDR version can result from an HDR-to-SDR conversion such as the one mentioned in clause 8.3, or can be provided as an input SDR master (see Figure 9-1).

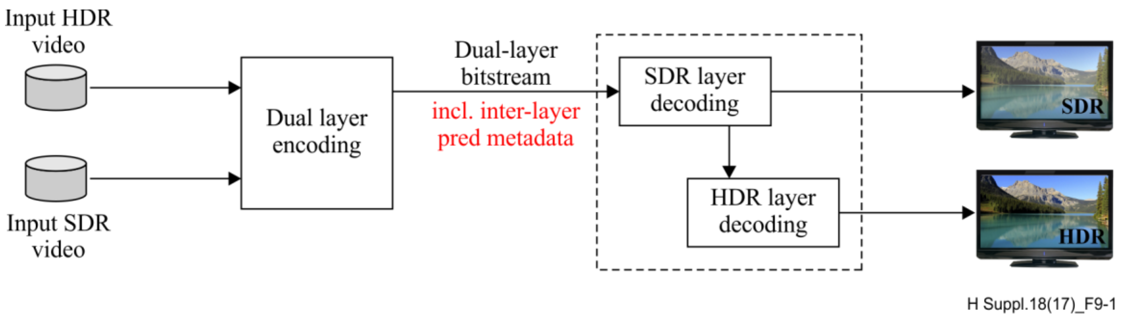

Figure 9-1 – Example of SDR backward compatible solution implemented with SHVC

9.2 Encoding and decoding stages

When creating the dual-layer SHVC bitstream it is suggested to set syntax elements in the base layer and enhancement layer sequence parameter sets (SPS) of the bitstream to the values listed in Table 9-1 (SDR base layer) and Table 9-2 (HDR enhancement layer). The syntax elements in Table 9-1 and Table 9-2 are conveyed in the Video Usability Information syntax branch of the SPS defined in Annex E of the HEVC specification. They may also be duplicated and carried in various application-layer headers.

Table 9-1 – Suggested settings for base layer (Y′CbCr 4:2:0 SDR 10-bits) in a dual-layer SHVC bitstream

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	2 (Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0 (narrow range)
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	1 (BT.709) or 9 (BT.2020) depending on SDR input
transfer_characteristics	vui_parameters( )	14 (SDR OETF)
matrix_coeffs	vui_parameters( )	1 (BT.709) or 9 (BT.2020) depending on SDR input
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2

Table 9-2 – Suggested settings for enhancement layer (Y′CbCr 4:2:0 PQ 10-bits) in a dual-layer SHVC bitstream

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	7 (Scalable Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0 (narrow range)
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	9 (BT.2100/BT.2020)
transfer_characteristics	vui_parameters( )	16 (BT.2100 PQ)
matrix_coeffs	vui_parameters( )	9 (BT.2100/BT.2020)
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2

The SDR video is coded as the base layer of SHVC encoding conforming to the Main 10 profile, and the HDR video is coded as the enhancement layer conforming to the Scalable Main 10 profile. When the colour containers in which the SDR and HDR are represented are different, SHVC colour gamut scalability (CGS) tool [Bordes2016] can apply to encode the content with high coding efficiency [Ye2016], [Baroncini2016]. Even if the colour containers in which the SDR and HDR are coded are the same, the colour volume occupied by the SDR and HDR pictures is typically different. Even in such cases, the CGS tool can encode the video with good coding efficiency.

The bitstream can be decoded by an HEVC Main 10 profile compliant decoder to generate the reconstructed SDR video directly compatible with SDR-only capable rendering devices. The dual-layer bitstream is can be decoded by an SHVC Scalable Main 10 profile compliant decoder, with the output from the base layer resulting in an SDR video representation and the enhancement layer resulting in a HDR video representation.

10 Display adaptation

10.1 General

Display adaptation aims at converting a decoded HDR video (for instance a decoded PQ10 signal) to a version adapted to a target rendering device. For instance, it may be used to convert a PQ10 video mastered for a display with a given peak luminance, to a display capable of displaying a lower peak luminance. The concept is illustrated in Figure 10-1. The conversion process can be driven by metadata produced during the mastering stage, and conveyed with the main bitstream. CRI or TMI SEI messages can be used for this purpose.

Display SDR backward compatibility, using CRI or TMI SEI messages, is a specific example of display adaptation, when the rendering device has only SDR display capabilities. This specific case is further described in clause 10.2. More generally, the adaptation of an HDR content with a given colour volume, to a version with a different target colour volume, can be achieved using CRI or TMI SEI messages. If multiple target colour volumes are considered, multiple instances of CRI or TMI SEI messages are coded. Identifier tags signalled in these SEI messages can be used to identify the different target colour volumes.

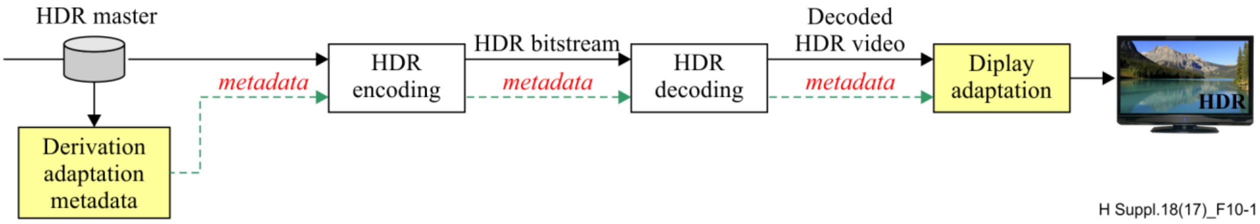

Figure 10-1 – Illustration of the concept of display adaptation

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

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

For HLG systems, Rec. ITU-R BT.2100 describes display adaptation as part of the display EOTF.

10.2 Case of display SDR backward compatibility

10.2.1 Conversion and coding process example

Dynamic mapping from HDR to SDR at the rendering side is the reverse problem of bitstream SDR backward compatibility studied in clause 8.3. In this Supplement, this use case is named display SDR backward compatibility.

The system design is deduced from Figure 8-4 and Figure 8-5, by placing a forward DRA process after the decoder and by signalling DRA parameters using SEI messages such as CRI [Rusanovskyy2016].

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

In the first example (Figure 10-2), the DRA operates in the Y′CbCr 4:2:0 domain. It applies three mapping functions to the Y′, Cb, and Cr components of the decoded PQ10 signal, for instance using the three Pre-LUTs of a CRI SEI message. The functions may be implemented in the shape of 1D-LUTs that directly apply to the PQ10 Y′, Cb and Cr components. The 3x3 matrix and Post-LUTs are not activated.

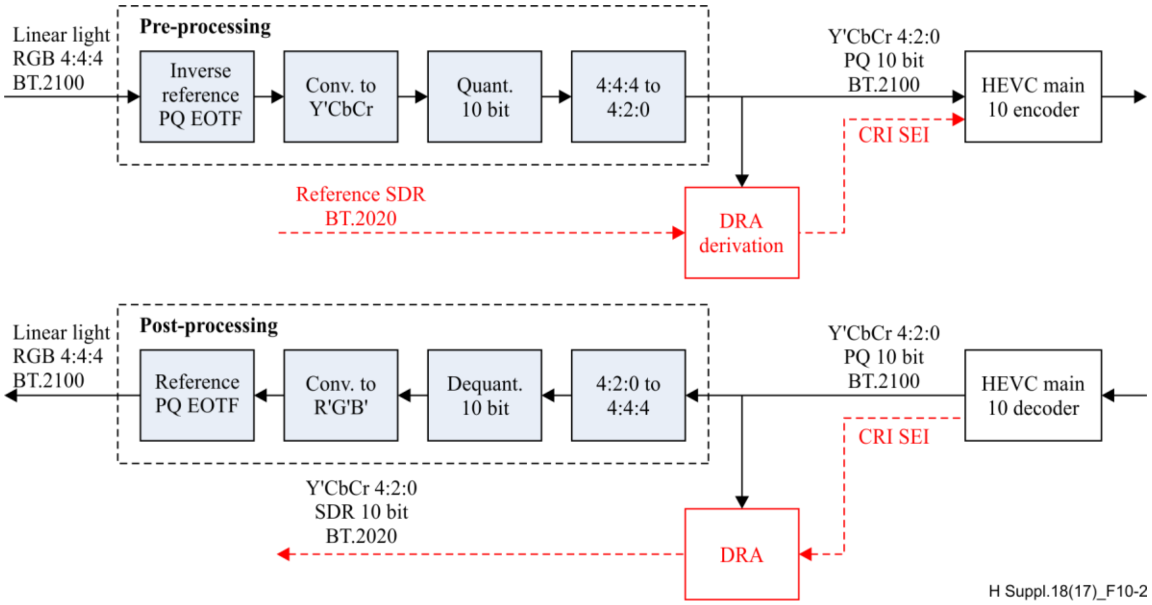

Figure 10-2 – Example of pre-processing (top) and post-processing (down) for display SDR backward compatibility with CRI applied in Y′CbCr 4:2:0 domain

In the second example (Figure 10-3), the DRA operates after the conversion of the decoded Y′CbCr 4:2:0 10-bits PQ BT.2100 video to 4:4:4. The complete CRI model may be used, which may be relevant when the target SDR colour primaries are BT.709 while the HDR content primaries are BT.2100. In that case, the three-by-three matrix can be beneficial to perform the conversion from BT.2100 to BT.709 primaries.

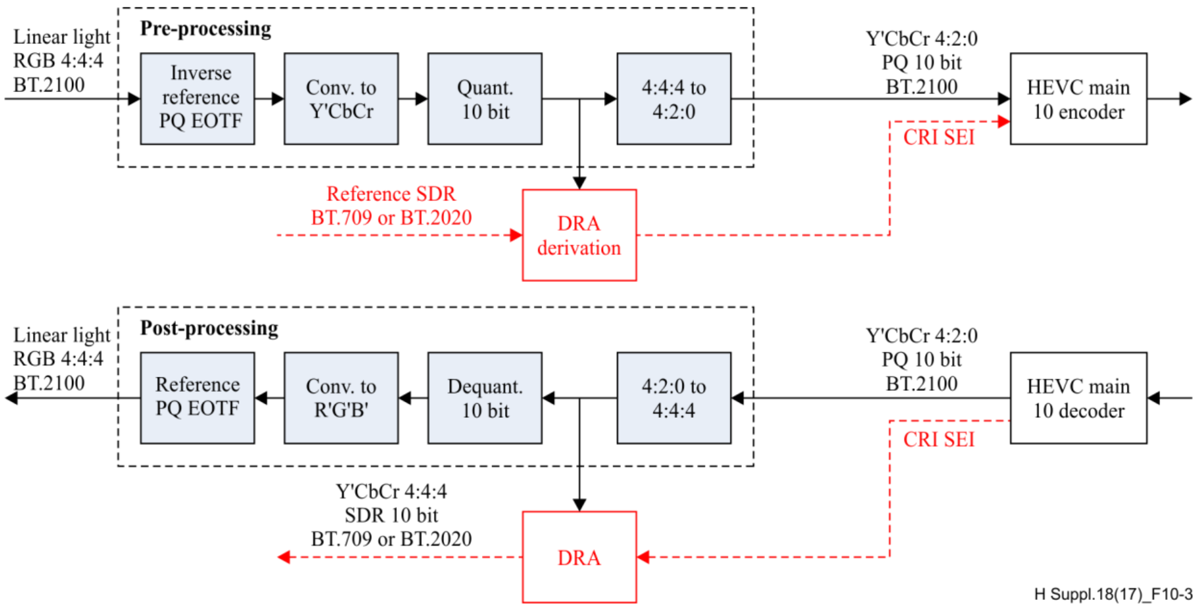

Figure 10-3 – Example of pre-processing (top) and post-processing (down) for display SDR backward compatibility with CRI applied in Y′CbCr 4:4:4 domain

10.2.2 Using colour remapping information SEI message

In the implementation illustrated in Figure 10-2, the DRA operates in Y′CbCr 4:2:0 domain. The three DRA functions can be implemented in 4:2:0 using the three Pre-LUTs of a CRI SEI message. The three-by-three matrix and the Post-LUTs are not activated. The conversion process can directly apply to the decoded 4:2:0 signal. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma sample Y'_{HDR} and each chroma sample U_{HDR} and V_{HDR} as follows:

$$Y'_{SDR} = PreLUT_0[Y'_{HDR}] \quad (10\text{-}1)$$

$$U_{SDR} = PreLUT_1[U_{HDR}] \quad (10\text{-}2)$$

$$V_{SDR} = PreLUT_2[V_{HDR}] \quad (10\text{-}3)$$

In the implementation illustrated in Figure 10-3, the DRA operates in Y′CbCr 4:4:4 domain. The full CRI model can be used, involving three Pre-LUTs, three-by-three matrix, and three Post-LUTs. This may be applicable, when the use case requires an SDR backward compatibility to SDR BT.709 colour primaries while the input HDR colour primaries are BT.2100. Activation of the matrix can help converting from one colour representation to the other one. The SDR-to-HDR conversion achieved at post-processing stage applies to each luma and chroma samples Y'_{HDR}, U_{HDR} and V_{HDR} as follows:

$$\begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} = \begin{bmatrix} PreLUT_0[Y'_{HDR}] \\ PreLUT_1[U_{HDR}] \\ PreLUT_2[V_{HDR}] \end{bmatrix} \quad (10\text{-}4)$$

$$\begin{bmatrix} Y_2 \\ U_2 \\ V_2 \end{bmatrix} = M_{3x3} \begin{bmatrix} Y_1 \\ U_1 \\ V_1 \end{bmatrix} \quad (10\text{-}5)$$

$$\begin{bmatrix} Y'_{SDR} \\ U_{SDR} \\ V_{SDR} \end{bmatrix} = \begin{bmatrix} PostLUT_0[Y_2] \\ PostLUT_1[U_2] \\ PostLUT_2[V_2] \end{bmatrix} \quad (10\text{-}6)$$

Error! Reference source not found. indicates suggested parameters settings of HEVC Main 10 bitstreams to correctly display video on PQ10 compatible devices and to conduct a guided mapping from SDR to HDR with CRI post-processing.

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

Syntax element	Location	Suggested value in HEVC
general_profile_space	profile_tier_level()	0
general_profile_idc	profile_tier_level()	2 (Main 10)
vui_parameters_present_flag	seq_parameter_set_rbsp( )	1
video_signal_type_present_flag	vui_parameters( )	1
video_full_range_flag	vui_parameters( )	0 (narrow range)
colour_description_present_flag	vui_parameters( )	1
colour_primaries	vui_parameters( )	9 (BT.2100/BT.2020)
transfer_characteristics	vui_parameters( )	16 (BT.2100 PQ)
matrix_coeffs	vui_parameters( )	9 (BT.2100/BT.2020)
chroma_loc_info_present_flag	vui_parameters( )	1
chroma_sample_loc_type_top_field	vui_parameters( )	2
chroma_sample_loc_type_bottom_field	vui_parameters( )	2
colour_remap_video_signal_info_present_flag	colour_remapping_info( )	1
colour_remap_transfer_function	colour_remapping_info( )	14 (SDR OETF)
colour_remap_full_range_flag	colour_remapping_info( )	0
colour_remap_primaries	colour_remapping_info( )	1 (BT.709) or 9 (BT.2020) depending on target SDR
colour_remap_matrix_coefficients	colour_remapping_info( )	1 (BT.709) or 9 (BT.2020) depending on target SDR
colour_remap_output_bit_depth	colour_remapping_info( )	10

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2018