TECHNICAL REPORT

ISO/IEC TR 23008-15

First edition
2018-08

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

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

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

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

Reference number
ISO/IEC TR 23008-15:2018(E)

Contents

Page

Foreword

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

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

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

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

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

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. A technically aligned twin text is published as ITU-T H.Sup18.

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

Introduction

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

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

This document serves several purposes:

— It provides a survey of identified video usability information (VUI) syntax elements and supplemental enhancement information (SEI) messages specified in HEVC and AVC applicable for HDR/WCG video.

— It covers conversion and coding chains using the IC_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.

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

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

1 Scope

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

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

NOTE The document complements the material provided in ITU-T H.Sup15 | ISO/IEC TR 23008-14, which is focused on conversion and coding practices for non-constant luminance (NCL) Y′CbCr video signals using the PQ transfer function.

2 Normative references

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

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

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

3 Terms and definitions

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

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

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
dynamic range adaptation
DRA
mapping process to convert content from one colour volume to another colour volume

3.2
electro-optical transfer function
EOTF
function which converts a non-linear video signal into a quantity of output linear light

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

3.3
full range
range in a fixed-point (integer) representation that spans the full range of values that could be expressed with that bit depth, such that, for 10-bit signals, black corresponds to code value 0 and peak white corresponds to code value 1023 for Y′

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

3.4
hybrid log-gamma
HLG
one set of transfer functions offering a degree of compatibility with legacy displays by more closely matching the previously established television transfer curves

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

3.5
narrow range
range in a fixed-point (integer) representation that does not span the full range of values that could be expressed with that bit depth such that, for 10-bit representations, the range from 64 (black) to 940 (peak white) is used for Y′ and the range from 64 to 960 is used for Cb and Cr

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

3.6
opto-electronic transfer function
OETF
function which converts a source input linear optical intensity into a non-linear video signal

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

3.7
opto-optical transfer function
OOTF
function which has the role of applying the “rendering intent” on video signal

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

3.8
perceptual quantizer
PQ
one set of transfer functions achieving a very wide range of brightness levels for a given bit depth using a non-linear transfer function that is finely tuned to match the human visual system

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

3.9
random access point access unit
RAPAU
access unit in the bitstream at which the initiation of the decoding process for some or all subsequent pictures in the bitstream is intended to be feasible

3.10
reference electro-optical transfer function
reference EOTF
specified EOTF for use under specific viewing environment, named the reference viewing environment

3.11
reference opto-electronic transfer function
reference OETF
specified OETF implemented within cameras, to ensure consistency of the image between cameras from different manufacturers

3.12
reference viewing environment
parameters to establish a reproducible viewing environment for critical viewing of material that can provide repeatable results from one facility to another when viewing the same material

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

4 Abbreviated terms

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

ATC alternative transfer characteristics

AVC advanced video coding, specified in Rec. ITU-T H.264 | ISO/IEC 14496-10

AVE ambient viewing environment

CGS colour gamut scalability

CI constant intensity

CL constant luminance

CLL content light level

CLVS coded layer-wise video sequence

CRI colour remapping information

FIR finite impulse response

HDR high dynamic range

HEVC high efficiency video coding, specified in Rec. ITU-T H.265 | ISO/IEC 23008-2

IC_TC_P alternative colour space representation to Y′CbCr, specified in Rec. ITU-R BT.2100-1

LMS long, medium, and short wavelength-based colour space, specified in Rec. ITU-R BT.2100-1

LUT look-up table

MAD mean absolute difference

MDCV mastering display colour volume

NCL non-constant luminance

PQ10 HDR content representation that utilizes the Rec. ITU-R BT.2100-1 colour primaries, the Rec. ITU-R BT.2100-1 reference PQ EOTF, and the Rec. ITU-R BT.2100-1 Y′CbCr colour space representation with 10 bits per sample in the 4:2:0 chroma sampling format

QP quantization parameter

RGB colour system using red, green, and blue components

SDR standard dynamic range

SEI supplemental enhancement information

SHVC scalable high efficiency video coding

SPS sequence parameter set

SSE sum of squared errors

TMI tone mapping information

UHD ultra-high definition

VUI video usability information

WCG wide colour gamut

XYZ CIE 1931 colour space; Y corresponds to the luminance signal

Y′CbCr colour space representation commonly used for video/image distribution as a way of encoding RGB information, also commonly expressed as YCbCr, $Y'C_BC_R$, 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 document are similar to those used in the C programming language. However, the results of integer division and arithmetic shift operations are defined more precisely, and additional operations are defined, such as exponentiation and real-valued division. Numbering and counting conventions generally begin from 0, e.g., “the first” is equivalent to the 0-th, “the second” is equivalent to the 1-th, etc.

5.2 Arithmetic operators

+	addition
−	subtraction (as a two-argument operator) or negation (as a unary prefix operator)
*	multiplication, including matrix multiplication
x^y	exponentiation (Denotes x to the power of y. In other contexts, such notation is used for superscripting not intended for interpretation as exponentiation.)
/	integer division with truncation of the result toward zero [For example, 7/4 and (−7)/(−4) are truncated to 1 and (−7)/4 and 7/(−4) are truncated to −1.]
÷	division in mathematical formulae where no truncation or rounding is intended
$\frac{x}{y}$	division in mathematical formulae where no truncation or rounding is intended
$\sum_{i=x}^{y} f(i)$	summation of *f*(*i*) with i taking all integer values from x up to and including y
x % y	modulus (Remainder of x divided by y, defined only for integers x and y with x >= 0 and y > 0.)

5.3 Bit-wise operators

&	bit-wise "and" (When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.)
|	bit-wise "or" (When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.)
^	bit-wise "exclusive or" (When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.)
x >> 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

= assignment operator

++ increment, i.e., x++ is equivalent to x = x + 1; when used in an array index, evaluates to the value of the variable prior to the increment operation

−− decrement, i.e., x−− is equivalent to x = x − 1; when used in an array index, evaluates to the value of the variable prior to the decrement operation

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

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

5.5 Relational, logical and other operators

== equality operator

!= not equal to operator

!x logical negation “not”

> larger than operator

< smaller than operator

>= larger than or equal to operator

<= smaller than or equal to operator

&& conditional/logical “and” operator
(Performs a logical “and” of its Boolean operators, but only evaluates the second operand if necessary.)

|| conditional/logical “or” operator
(Performs a logical “or” of its Boolean operators, but only evaluates the second operand if necessary.)

a ? b : c ternary conditional
(If condition a is true, then the result is equal to b; otherwise the result is equal to c.)

5.6 Mathematical functions

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

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

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

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

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

$\text{Floor}(x)$ largest integer less than or equal to x

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

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

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

$\text{Log10}(x)$ base-10 logarithm of x

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

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

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

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

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

5.7 Order of operations

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

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

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

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

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

Table 1 — Operation precedence from highest (at top of table) to lowest (at bottom of table)

Operations (with operands x, y, and z)
"x++", "x−−"
"!x", "−x" (as a unary prefix operator)
"x^y"

Table 1 *(continued)*

Operations (with operands x, y, and z)
"x * y", "x / y", "x ÷ y", "$\frac{x}{y}$", "x % y"
"x + y", "x − y" (as a two-argument operator), "$\sum_{i=x}^{y} f(i)$"
"x << y", "x >> y"
"x < y", "x <= y", "x > y", "x >= y"
"x = = y", "x != y"
"x & y"
"x | y"
"x && y"
"x | | y"
"x ? y : z"
"x..y"
"x = y", "x += y", "x −= y"

6 Overview

This document is structured as follows.

— Clause 7 reviews identified signalling mechanisms of HEVC, and when applicable, of AVC, relevant for HDR/WCG video coding and distribution. It also describes some common processing steps used in end-to-end processing chains such as described in Clauses 8 and 9.

— Clause 8 describes usage of HLG transfer functions and CRI or TMI SEI messages for the support of bitstream SDR backward compatibility (defined below) with a single-layer profile (e.g. HEVC Main 10).

— Clause 9 describes a dual-layer HDR/WCG video coding system with bitstream SDR backward compatibility implemented with the HEVC Scalable Main 10 profile.

— Clause 10 addresses the display adaptation functionality (defined below), with application examples based on the CRI SEI message. This clause includes the specific case of display SDR backward compatibility.

"SDR backward compatibility" relates to the ability of HDR/WCG video coding and distribution systems to produce a video signal suitable for SDR-only capable rendering devices (e.g. UHD SDR display with Rec. ITU-R BT.2020-2 colour primaries). In the present document, it is defined in two modes: bitstream and display.

— In HDR/WCG distribution systems that support "bitstream" SDR backward compatibility, the decoded video signal from a standard-compliant decoder (e.g. HEVC Main 10 decoder) can be directly displayed on an SDR-capable display without adaptation. Two categories of "bitstream" SDR backward compatibility are considered:

  — In "static" bitstream SDR backward compatibility, the decoded video is an HDR signal, for instance, Y′CbCr 4:2:0 10-bits with the Rec. ITU-R BT.2100-1 reference HLG opto-electronic transfer function (OETF) and Rec. ITU-R BT.2100-1 colour primaries, that can be directly displayed on an HDR-capable display or an SDR-capable display, without adaptation. In this context, the HDR processing chain is static, and not dependent on the input video data.

  — In "dynamic" bitstream SDR backward compatibility, the decoded video is an SDR signal. A post-processing step can be further used to reconstruct an HDR signal, using metadata conveyed for instance in CRI or TMI SEI messages. In this context, the HDR processing chain is dynamic, and adapts to the input video data.

— In HDR/WCG distribution systems that support "display" SDR backward compatibility, the decoded video signal from a standard-compliant decoder (e.g. HEVC Main 10 decoder) is an HDR signal (for instance, Y′CbCr 4:2:0 10-bits with Rec. ITU-R BT.2100-1 inverse reference PQ electro-optical transfer function (EOTF) and Rec. ITU-R BT.2100-1 colour primaries). A post-decoding dynamic range adaptation (DRA) process is applied to the decoded video signal to produce an SDR video signal that can be displayed on an SDR-capable display. The adaptation process can use metadata, conveyed for example in CRI SEI messages, to perform this conversion.

"Display adaptation" is a generic term covering techniques of video signal processing which adapt the decoded video signal to a target display. Techniques providing display SDR backward compatibility are considered as a subset of display adaptation. Display adaptation techniques aim at converting an HDR/WCG video signal, originally produced for a reference display capable of displaying a certain colour volume (dynamic range and colour gamut), to a video signal suitable to a target rendering device of colour volume capabilities different from the reference display capabilities. For instance, it can be used to convert a Y′CbCr 4:2:0 10-bits Rec. ITU-R BT.2100-1 PQ signal (denoted PQ10 in the present document), originated from an HDR video master produced on a display with a given reference peak luminance, to a lower peak luminance capable display. Display adaptation could also increase the colour volume, if desired. Another term used in the industry for display adaptation is regrading. Display adaptation can be driven by metadata transmitted along with the video bitstream, for instance using SEI messages.

Conversion and coding practices related to production and compression of HDR/WCG video signal represented with NCL Y′CbCr 4:2:0 video with Rec. ITU-R BT.2100-1 PQ transfer characteristics are outside of scope of this document. These aspects are specifically addressed in ITU-T H.Sup15 | ISO/IEC TR 23008-14.

7 HEVC signalling mechanisms applicable to HDR/WCG video

7.1 General

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

This clause is structured as follows.

— Subclause 7.2 reviews VUI signalling applicable to HDR/WCG video.

— Subclause 7.3 reviews SEI messages applicable to HDR/WCG video.

— Subclause 7.4 provides an overview of PQ and HLG transfer functions.

— Subclause 7.5 provides a description of IC_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 Rec. ITU-R BT.2100-1 inverse reference PQ EOTF.

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

7.2 VUI syntax elements

By design, metadata signalled in syntax elements of VUI is not necessary for constructing the luma or chroma samples by the decoding process, and may be ignored by the decoder. However, such syntax elements provide useful parameters or attributes of an encoded signal and can be utilized in the video system design. Examples of VUI parameters relevant to HDR/WCG video system design include colour primaries, transfer characteristics and matrix coefficients specified in Rec. ITU-T H.265 | ISO/IEC

