TECHNICAL REPORT

ISO/IEC TR 23008-14

First edition
2018-08

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

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

Partie 14: Conversion et pratiques de codage pour la vidéo HDR/WCG Y′CbCr 4:2:0 avec caractéristiques de transfert PQ

Reference number
ISO/IEC TR 23008-14:2018(E)

© ISO/IEC 2018



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2018

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

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

Published in Switzerland

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**

4 Abbreviated terms ... **3**

5 Conventions ... **4**
5.1 General ... 4
5.2 Arithmetic operators ... 4
5.3 Bit-wise operators ... 5
5.4 Assignment operators ... 5
5.5 Relational, logical, and other operators ... 5
5.6 Mathematical functions ... 6
5.7 Order of operations ... 7

6 Overview ... **7**

7 Pre-encoding process ... **8**
7.1 General ... 8
7.2 Pre-encoding process stages ... 9
7.2.1 Conversion from a linear to a non-linear light representation: RGB to R′G′B′ ... 9
7.2.2 Colour representation conversion: R′G′B′ to non-constant luminance Y′CbCr ... 11
7.2.3 Chroma down-conversion ... 12
7.2.4 Floating-point to fixed-point (narrow range) 10 bit conversion ... 14
7.3 Closed loop pre-encoding conversion — Luma adjustment ... 16
7.3.1 General ... 16
7.3.2 Luma adjustment — Iterative approach ... 17
7.3.3 Luma adjustment — Closed form solution ... 22

8 Encoding process ... **24**
8.1 General ... 24
8.2 Perceptual luma quantization ... 24
8.2.1 General ... 24
8.2.2 Example of luma-dependent adaptive quantization ... 25
8.3 Chroma QP offset ... 26
8.3.1 General ... 26
8.3.2 Example of chroma QP offset settings ... 27
8.4 Other encoding aspects ... 29
8.5 HEVC encoding ... 29
8.6 AVC encoding ... 29

9 Decoding process ... **30**

10 Post-decoding processes ... **30**
10.1 General ... 30
10.2 Conversion from a fixed-point to a floating-point representation ... 31
10.3 Chroma up-sampling ... 31
10.4 Colour representation conversion: non-constant luminance Y′CbCr to R′G′B′ ... 32
10.5 Conversion from a non-linear to a linear light representation: R′G′B′ to RGB ... 32

Annex A (informative) **Supplemental enhancement information (SEI) messages** ... **34**

Bibliography ... **36**

© ISO/IEC 2018 – All rights reserved

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 Joint 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.Sup15.

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

© ISO/IEC 2018 – All rights reserved

Introduction

High dynamic range (HDR) video is a type of video content in which the sample values span 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. On the other hand, wide colour gamut (WCG) video 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 provides a set of recommended guidelines on processing of consumer distribution HDR/WCG video. This includes recommendations for converting a video signal, in a linear light RGB representation with Rec. ITU-R BT.2020 colour primaries, to a 10-bit, narrow range, PQ encoded (as defined in SMPTE ST 2084 and Rec. ITU-R BT.2100), 4:2:0, non-constant luminance Y′CbCr representation. These guidelines may also apply to other representations with higher bit depth or other colour formats, such as 4:4:4 Y′CbCr 12 bit video. The scope of this document is illustrated in Figure 1.

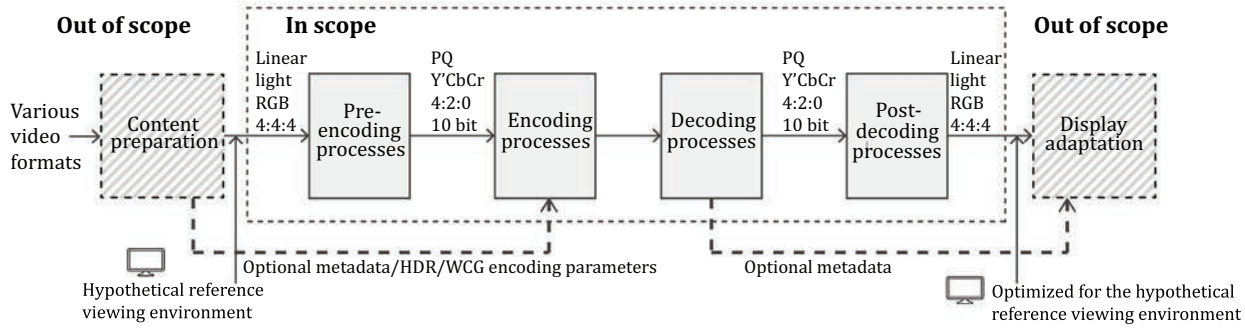


Figure 1 — Illustration of the scope of this document

The content preparation step, as well as the display adaptation step, are considered to be out of the scope of this document. However, metadata generated during the content preparation step may be passed through the encoder-decoder chain and can significantly affect the display adaptation step. The content preparation step may include filtering and image enhancement processing such as de-noising, colour correction, and sharpening, as well as other processes. Such methods are deliberately not described in this document. The processing steps described in this document are made available for reference only and the document does not contain any elements of normative nature. It is possible to replace one or more of the processing steps described in this document; for example, in order to reduce computational complexity or to improve fidelity. This document's intention is to provide some guidelines for operating an HDR/WCG video system that is constrained to code a 10-bit, PQ (as defined in SMPTE ST 2084 and Rec. ITU-R BT.2100), 4:2:0, non-constant luminance Y′CbCr signal representation. This configuration is also aligned with the HDR10 media profile defined in DECE v2.1, the interface defined in CTA 861G and the restrictions in the BD-ROM specifications. The processing steps in this document are designed for the case when the same hypothetical reference viewing environment (HRVE) is used before and after the HDR/WCG system. This document does not account for when the viewing environment used after the HDR/WCG system is different from the viewing environment used as the HRVE. In particular, display adaptation, such as the techniques described in the SMPTE ST 2094 standards, are not considered in this document. Report ITU-R BT.2390 contains additional information on viewing environments and examples of parameters that may be appropriate to apply for practical HDR/WCG systems. This document does not provide a description of any preferred HRVE, but acknowledges the fact that in many applications of HDR/WCG video, it may be desirable to have a well-defined HRVE description in order to ensure alignment between content preparation and content consumption.

© ISO/IEC 2018 – All rights reserved

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

1 Scope

This document provides guidance on the processing of high dynamic range (HDR) and wide colour gamut (WCG) video content. The purpose of this document is to provide a set of publicly referenceable recommended guidelines for the operation of AVC or HEVC video coding systems adapted for compressing HDR/WCG video for consumer distribution applications. This document includes a description of processing steps for converting from 4:4:4 RGB linear light representation video signals into non-constant luminance (NCL) Y′CbCr video signals that use the perceptual quantizer (PQ) transfer function defined in SMPTE ST 2084 and Rec. ITU-R BT.2100. Although the focus of this document is primarily on 4:2:0 Y′CbCr 10 bit representations, these guidelines are also applicable to other representations with higher bit depth or other colour formats, such as 4:4:4 Y′CbCr 12 bit video. In addition, this document provides some high-level recommendations for compressing these signals using either the AVC or HEVC video coding standards. A description of post-decoding processing steps is also included for converting these NCL Y′CbCr signals back to a linear light, 4:4:4 RGB representation.

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.

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

Rec. ITU-T H.265 | ISO/IEC 23008-2, *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 http://www.iso.org/obp

3.1
electro-optical transfer function
EOTF
function used in the post-decoding process to convert from a non-linear representation to a linear representation

© ISO/IEC 2018 – All rights reserved

3.2
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.

3.3
inverse electro-optical transfer function
inverse EOTF
function used in the pre-encoding process to convert from a linear representation to a non-linear representation, computed as the inverse of the *EOTF* (3.1)

Note 1 to entry: In this document, the pre-encoding process is assumed to operate on HDR/WCG video content that has been prepared for a hypothetical reference viewing environment as shown in Figure 1. The content preparation step may contain processing such as applying an *opto-optical transfer function (OOTF)* (3.6), in which the HDR/WCG video is converted from one linear representation (corresponding to the scene) to another linear representation (corresponding to the display). The OOTF has the role of applying a "rendering intent". In systems where no such OOTF is applied in the content preparation step, the process of converting from a linear representation (corresponding to the scene) to a non-linear representation is typically called the *opto-electrical transfer function (OETF)* (3.5).

3.4
narrow range
range in a fixed-point (integer) representation that does not span the *full range* (3.2) 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.

Note 2 to entry: Narrow range is, in some applications, called by synonyms such as: "limited range", "video range", "legal range", "SMPTE range" or "standard range".

3.5
opto-electrical transfer function
OETF
function that converts linear scene light into the video signal, typically within a camera

3.6
opto-optical transfer function
OOTF
function that maps relative scene linear light (typically the camera output signal) to display linear light (typically, the signal driving a mastering monitor)

3.7
random access point access unit
RAPAU
access unit in the bitstream containing an intra coded picture with the property that all pictures following the intra coded picture in output order can be correctly decoded without using any information preceding the random access point access unit in the bitstream

3.8
transfer function
function that can be *EOTF* (3.1), *inverse EOTF* (3.3), *OETF* (3.5), inverse OETF, *OOTF* (3.6), or inverse OOTF

© ISO/IEC 2018 – All rights reserved