TECHNICAL REPORT

ISO/IEC TR 23091-4

First edition
2019-08

Information technology — Coding-independent code points —

Part 4: Usage of video signal type code points

Reference number
ISO/IEC TR 23091-4:2019(E)

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

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.

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) or the IEC list of patent declarations received (see http://patents.iec.ch).

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

For an explanation of 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 www.iso.org/iso/foreword.html.

This document was prepared by ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information* in collaboration with ITU-T. The technically identical text is published as ITU-T Series H Supplement 19 (03/2019).

A list of all parts in the ISO/IEC 23091 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document discusses video signal property description code points and their combinations that are widely used in production and video content workflows. Video properties and values are usually expressed in "metadata" that can exist across production and distribution workflows. Knowledge of these properties and their combinations has value as content is processed in the end-to-end production-to-distribution workflow chain.

The combinations of all possible expressible video properties as code point values could hypothetically result in hundreds or thousands of permutations; but many of those combinations are rarely or never used in practice. For example, it is highly unlikely that perceptual quantization (PQ) transfer characteristics function specified in Rec. ITU-R BT.2100 would be combined with the colour primaries specified in Rec. ITU-R BT.601. Only a small subset of the possible combinations is used in practice.

This document is intended to help the producers of various content processing tools to avoid processing mistakes that can cause video quality degradation due to having incorrect assumptions made about video property combinations. There are only a few limited sets of video property combinations that are widely used in present-day video production and distribution equipment chains. This document describes these limited sets of combinations that are currently widely used and describes how the associated signal type metadata is carried to aid in the automation of content workflows across various domains of capture, production and distribution. Lastly, this document aims to help its readers, especially toolset developers, to repurpose tools to work properly across several domains (e.g., capture, production, production distribution, and service distribution) where similar video conversion functions (e.g., chroma sub-sampling or colour space conversions) may be performed.

Information technology — Coding-independent code points —

Part 4:
Usage of video signal type code points

1 Scope

This document describes common industry representation practices for the usage of video signal type code points, as these properties are conveyed across video content production and distribution carriage systems.

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 — High efficiency video coding*

Rec. ITU-T H.273 | ISO/IEC 23091-2, *Information technology — Coding-independent code points — Part 2: Video*

3 Terms and definitions

For the purposes of this document, the terms and definitions in Rec. ITU-T H.265 | ISO/IEC 23008-2, Rec. ITU-T H.264 | ISO/IEC 14496-10 and Rec. ITU-T H.273 | ISO/IEC 23091-2 and the following apply.

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

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

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

3.1
colour volume
space of all colours and intensities that a device or signal can reproduce or convey

3.2
creative intent
desired vision of the content creator who adjusts and approves the appearance of rendered content in the production process

Note 1 to entry: Examples of a content creator are a director, cinematographer, videographer, editor or colourist.

3.3
electro-optical transfer function
EOTF
function to map a non-linear video signal to display linear light

3.4
full range
range in a fixed-point (integer) representation that spans the full range of values that could be expressed with that bit depth

3.5
inverse electro-optical transfer function
inverse EOTF
function that is the inverse of an *EOTF* (3.3)

3.6
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

Note 1 to entry: Narrow range is, in some applications, referred to by synonyms such as: “limited range”, “video range”, “legal range”, “SMPTE range” or “standard range”.

3.7
opto-electrical transfer function
OETF
function to map relative scene linear light to a non-linear video signal

3.8
opto-optical transfer function
OOTF
function to map relative scene linear light to display linear light

3.9
random access point access unit
RAPAU

access unit in a video 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 it in the bitstream

3.10
transfer function
function among any of the following: *EOTF* (3.3), *inverse EOTF* (3.5), *OETF* (3.7), inverse OETF, *OOTF* (3.8), or inverse OOTF

4 Abbreviated terms

2K	informally used to refer to an HD resolution (1920 × 1080 for television or 2048 × 1080 for film)
4K	informally used to refer a UHD resolution (3840 × 2160 for television or 4096 × 2160 for film)
8K	informally used to refer to a UHD resolution (7680 × 4320 or 8192 × 4320)
AVC	advanced video coding (Rec. ITU-T H.264 | ISO/IEC 14496-10)
CICP	coding-independent code points (Rec. ITU-T H.273 | ISO/IEC 23091-2)
GBR	green, blue and red component colour system in linear light domain; same as RGB, although emphasizing that the green component is handled as the primary colour component by some technical elements of the video coding technology NOTE The colour representation does not indicate the media component order in a coded representation. For example, GBR represents the same component colour system as RGB.

G′B′R′ green, blue and red component colour system in a non-linear domain associated with a transfer function which maps the linear light domain to a more perceptually uniform domain; same as R′G′B′, although emphasizing that the green component is handled as the primary colour component by some technical elements of the video coding technology

NOTE The colour representation does not indicate the media component order in a coded representation. For example, G′B′R′ represents the same component colour system as R′G′B′.

HD high definition

HDR high dynamic range

HEVC high efficiency video coding (Rec. ITU-T H.265 | ISO/IEC 23008-2)

HLG hybrid log-gamma (as defined in Rec. ITU-R BT.2100)

HVS human visual system

LCD liquid crystal display

LED light-emitting diode

LUT look-up table

MDCV mastering display colour volume

MXF material exchange format (as defined in SMPTE ST 377-1)

N/A not applicable

N/R not required

NCG narrow colour gamut (typically as per Rec. ITU-R BT.709)

NCL non-constant luminance

OLED organic light-emitting diode

PQ perceptual quantizer (as defined in Rec. ITU-R BT.2100)

QP quantization parameter

RGB red, green and blue component colour system in linear light domain

NOTE The colour representation does not indicate the media component order in a coded representation. For example, RGB represents the same component colour system as GBR.

R′G′B′ red, green and blue component colour system in a non-linear domain associated with a transfer function which maps the linear light domain to a more perceptually uniform domain

NOTE The colour representation does not indicate the media component order in a coded representation. For example, R′G′B′ represents the same component colour system as G′B′R′.

SD standard definition

SDR standard dynamic range

SEI supplemental enhancement information

UHD ultra-high definition

UL universal label (as defined in SMPTE ST 377-1)

VUI video usability information (a sequence-level syntax structure in HEVC and AVC bit-streams)

WCG wide colour gamut (a gamut substantially wider than the gamut conveyed by Rec. ITU-R BT.709, e.g., as per Rec. ITU-R BT.2020 or Rec. ITU-R BT.2100)

XYZ CIE 1931 colour space (wherein Y corresponds to the luminance signal)

Y′CbCr luma (Y′), chroma blue (Cb) and chroma red (Cr) colour representation defined by a matrix transformation relationship to an R′G′B′ colour system

NOTE A Y′CbCr representation is commonly used for video/image distribution as a way of encoding RGB information. Such a representation is also commonly expressed as YCbCr, $Y'C_BC_R$, or $Y'C'_BC'_R$, and can also be known as YUV in some documents. The relationship between Y′CbCr and R′G′B′ considered in this document is defined by matrix coefficients specified in Rec. ITU-R BT.601, Rec. ITU-R BT.709, Rec. ITU-R BT.2020 or Rec. ITU-R BT.2100. Unlike the CIE-Y component in the linear-light XYZ representation, the non-linear, the approximately perceptual uniform Y′ in this representation might not be representing true luminance, regardless of the transfer function.

5 Overview

This document discusses video signal property description code points and their combinations that are widely used in production and video content workflows. Video properties and values are usually expressed in "metadata" that can exist across production and distribution workflows. Knowledge of these properties and their combinations has value as content is processed in the end-to-end production-to-distribution workflow chain.

The combinations of all possible expressible video properties as code point values could hypothetically result in hundreds or thousands of permutations; but many of those combinations are rarely or never used in practice. For example, it is highly unlikely that perceptual quantization (PQ) transfer characteristics function specified in Rec. ITU-R BT.2100 would be combined with the colour primaries specified in Rec. ITU-R BT.601. Only a small subset of the possible combinations is used in practice.

This document is written to provide information to help the producers of various content processing tools to avoid processing mistakes that can cause video quality degradation due to having incorrect assumptions made about video property combinations. There are only a few limited sets of video property combinations that are widely used in present-day video production and distribution equipment chains. This document describes these limited sets of combinations that are currently widely used and describes how the associated signal type metadata is carried to aid in the automation of content workflows across various domains of capture, production and distribution. Lastly, this document aims to help its readers, especially toolset developers, to repurpose tools to work properly across several domains (e.g., capture, production, production distribution, and service distribution) where similar video conversion functions (e.g., chroma subsampling or colour space conversions) may be performed.

The coding-independent code points (CICP) specification for video (Rec. ITU-T H.273 | ISO/IEC 23091-2) defines code points and fields that identify some properties of video signals. These are defined independently from how these properties are carried in a coded video-layer bitstream such as an HEVC or AVC bitstream, which could differ depending on bitstream format. The compressed representation is sometimes considered to be a temporary, compacted state for distribution or delivery of the video signal, while the reconstructed video signal output from a video decoder may be interpreted as having the same meaning as a video signal immediately prior to compression by a compression encoder.

Subclauses 7.2 and 7.3 define system identifier tags combinations of the described commonly used values of such video signal property combinations that apply across domains. In addition, these subclauses also identify how the video property values are carried in the signal processing workflow.

6 Workflow domains

Figure 1 illustrates workflow domains (capture, production, production distribution and service distribution) in which video content may exist, be edited or be converted. Typical content workflows across these domains are theatrical/scripted TV or live events. There are many similar video processing functions that can be performed in each domain and often these may be repeated in the next successive domain.

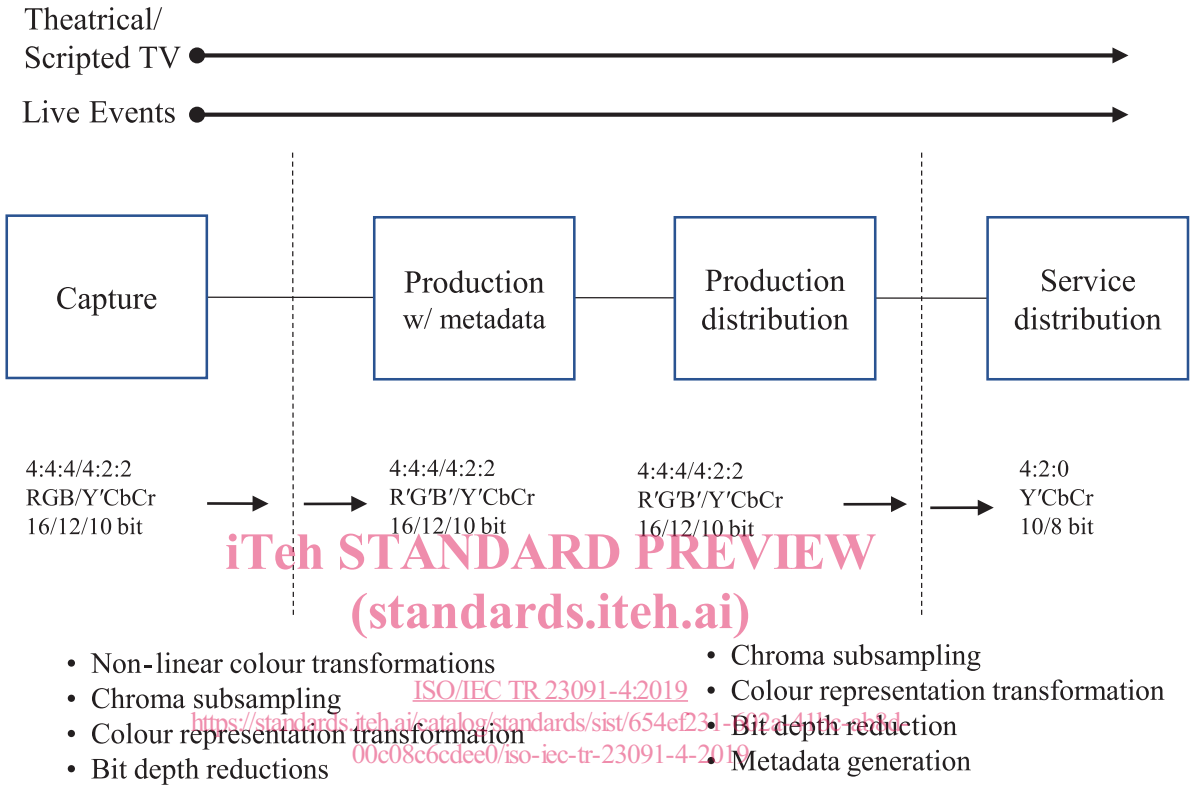

Figure 1 — Video workflows through different carriage domains

In the capture domain, content is created through sensors on cameras converting optical signals into a digital format. Content is retained at its highest informational format, although some conversions may be performed to reduce transport bandwidth demands.

In the interface to the production domain, content undergoes further processing transformations such as non-linear transformations, chroma subsampling (e.g., 4:4:4 to 4:2:2), colour representation changes (e.g., RGB to Y′CbCr NCL), and bit depth reduction (e.g., 16 bits per sample to 10 bits per sample). For theatrical/scripted TV workflows entering in the production domain, content can be added to by computer-generated imagery sources, overlaid with graphics, and colour graded using a mastering display. For live event workflows, there is always a real-time constraint, which limits content processing to real-time operations. After the colour grading, both static and dynamic metadata may be generated that are to be attached to the content workflow. However, for live events, the generation of highly customized metadata may not be practical, and metadata may need to be generated further downstream by automated content analysis approaches.

In the production distribution domain, some additional processing is done to the content to further reduce transport bandwidth demands. This may include some sample-wise processing transformations (chroma subsampling and bit depth) and compression (e.g., using HEVC or AVC) but mostly employing spatial compression techniques.

For 4:2:0 chroma subsampling operations, it is important to make known the relative location alignment of the initial subsampling location processing of the content to avoid unnecessary quality degradation upon further content processing. For purposes of this document, this property is described