|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  29th Meeting: Macao, CN, 19–25 Oct. 2017 | Document: JCTVC-AC1003 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Usage of video signal type code points (Draft 1)** | | |
| *Status:* | Output Document approved by JCT-VC | | |
| *Purpose:* | Draft Text | | |
| *Author(s) or Contact(s):* | Yasser Syed  Chad Fogg  Lars Borg  Chris Seeger  Alexis Tourapis  Gary Sullivan | Tel: Email: | +1 303-246-8413 [yasser\_syed@comcast.com](mailto:yasser_syed@comcast.com)  [chad.fogg@gmail.com](mailto:chad.fogg@gmail.com)  [borg@adobe.com](mailto:borg@adobe.com)  chris.seeger@nbcuni.com  [alexismt@apple.com](mailto:alexismt@apple.com)  [garysull@miscrosoft.com](mailto:garysull@miscrosoft.com) |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document contains a draft of a technical report on prominent video property description points combinations applied in the production industry. This output document is to be used for a discussion on a planned Technical Report (23091-4), which is a non-normative document of industry best practices on describing existing combinations of video properties, some of which are coding points and their representations in different carriage systems. The combinations of properties and the permutations of all possible values being considered can reach hundreds of choices to describe a stream and some of these combinations are not expected to be used in practice (such as using the ITU-R BT.2100 perceptual quantization (PQ) transfer characteristics with ITU-R BT.601 colour primaries). Only a small subset of these combinations (such as BT.2100 colour primaries with BT.2100 PQ and hybrid log-gamma transfer characteristics) is used in practice. This document aims to avoid mistaken assumptions on video property combinations made by vendors of various content processing tools that may be part of a production equipment chain. With the increased usage of high-dynamic range and the increased use of look-up tables in television systems, these content processing mistakes could increasingly become magnified. Lastly, this Technical Report aims to help its readers, especially toolset developers, in avoiding unnecessary, complicated tailoring of their content processing tools to specific areas of the workflow, and thus improve the repurposing of their tools to different parts of the workflow.

# Scope

This document describes non-normative industry best practices on the usage of existing combinations of video properties and their representations in different carriage systems for video content production.

# References

1. W16694 Geneva Jan 2017- 23091-2 Information technology – Coding-independent code points – Part 2: Video. [Ed. Note YS. Reference H.273, which is the ITU version; For ISO/IEC, can still reference ISO/IEC 23001-8 and transition over to reference the new ISO/IEC standard once it is published. Remove reference to WG 11 document.]
2. [SMPTE 2067-21:2016 Interoperable Master Format- Application #2E](http://ieeexplore.ieee.org/document/7560818/).
3. SMPTE ST 2086:2014:Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images.
4. SMPTE 377-1:2011 Am1:2012 Am2:2012: Material Exchange Format (MXF)- File Format Specification.
5. ISO/IEC 14496-10/ ITU-T H.264, Information technology – Coding of audio-visual objects Part 10: Advanced Video Coding.
6. ISO/IEC 23008-2/ ITU-T H.265, Information Technology – High efficiency coding and media delivery in heterogeneous environments Part 2: High efficiency video coding.
7. ITU-R BT.2100-1:2017 Image parameter values for high dynamic range television for use in production and international programme exchange.
8. SMPTE EG 432-1. Digital Source Processing – Color Processing for D-Cinema
9. SMPTE RP 431-2. D-Cinema Quality – Reference Projector and Environment.

[Ed: P3 will be more formally defined in a new SMPTE specification, expected to be completed by end of year 2018. That should supersede reference [8] and possibly [9] above.]

# Definitions

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

|  |  |
| --- | --- |
| * 1. **a** | **electro-optical transfer function (EOTF):** The function used in the post-decoding process to convert from a non-linear representation to a linear representation. |
|  | **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. |
|  | **inverse electro-optical transfer function (inverse EOTF**): A 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.  NOTE 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 **Error! Reference source not found.**. The content preparation step may contain processing such as applying an opt-optical transfer function (OOTF), 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 opt-electrical transfer function (OETF). |
|  | **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 Rec. ITU-R BT.2100.  NOTE to entry –  Narrow range is, in some applications, called by synonyms such as: “limited range”, “video range”, “legal range”, “SMPTE range” or “standard range”. |
|  | **opto-electrical transfer function (OETF):** The function that converts linear scene light into the video signal, typically applied within a camera. |
|  | **opto-optical transfer function (OOTF):** A function that maps relative scene linear light (typically the camera output signal) to display linear light (typically, the signal driving a mastering monitor). |
|  | **random access point access unit (RAPAU):** An 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. |
|  | **transfer function:** In this document, a transfer function refers to any of the following; EOTF, inverse EOTF, OETF, inverse OETF, OOTF, or inverse OOTF. |

# Abbreviations and acronyms

This document uses the following abbreviations and acronyms: [Ed. Note: Consider using a table to ensure alignment of abbreviation and its definition]

AVC Advanced Video Coding (Rec. ITU-T H.264 | ISO/IEC 14496-10)

EOTF Electro-Optical Transfer Function

HD High Definition

HDR High Dynamic Range

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

HVS Human Visual System

LUT Look-up Table

NCL Non-Constant Luminance

PQ Perceptual Quantizer (as defined in SMPTE ST 2084 and Rec. ITU-R BT.2100)

QP Quantization Parameter

RAPAU Random Access Point Access Unit

RGB Colour System using linear light Red, Green, and Blue components

R′G′B′ Colour System using non-linear (mapped to approximately perceptually uniform steps) Red, Green, and Blue components. The mapping between linear RGB and non-linear R′G′B′ is indicated by the transfer function.

SDR Standard Dynamic Range

SEI Supplemental Enhancement Information

OETF Opto-Electrical Transfer Function

OOTF Opto-Optical Transfer Function

VUI Video Usability Information. A sequence level sub-header in AVC and HEVC bitstreams.

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′CBCR, 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 linear light XYZ representation, the non-linear, approximately perceptual uniform Y′ in this representation might not be representing the same quantity, regardless of the transfer function. Y′ is commonly referred to as “luma”. Cb and Cr are commonly referred to as “chroma”. It could also be known as YUV in other domains.

# Overview

Part 2 of the Coding-Independent Code Points (CICP) standard (ISO/IEC 23091-2) defines code points and fields that identify the video signal type, independent from its compressed representation in a coded video layer bitstream, such as an AVC or HEVC bitstream. The compressed representation is often considered to be a temporary, compact distribution delivery state of the video signal, while the reconstructed video signal output from the decoder should be interpreted as having the same meaning as the video signal immediately prior to compression, indicated by the signal type code points conveyed by metadata mechanisms of the compressed stream such as the VUI header. The combinations of properties and the permutations of all possible values being considered can reach hundreds of choices to describe a signal type, and some of these combinations are not expected to be used in practice (such as using the ITU-R BT.2100 perceptual quantization (PQ) transfer characteristics with ITU-R BT.601 colour primaries). Only a small subset of these combinations (such as BT.2100 colour primaries with BT.2100 PQ /HLG transfer characteristics, Y’CbCr with narrow range signaling) are expected to be used in the industry. These properties, usually expressed in "metadata", can exist across the production and distribution workflows and knowledge of these properties and combinations has value as content gets processed in the E2E [End-to-End] production to distribution workflow chain.

This planned Technical Report is a non-normative document of industry best practices on describing existing combinations of video properties and their representations in different carriage systems. This document aims to avoid mistaken assumptions on video property combinations made by vendors of various content processing tools that may be part of a production equipment chain. With the increased usage of high-dynamic range and the increased use of look-up tables in television systems, these content processing mistakes could increasingly become magnified. Lastly, this Technical Report aims to help its readers, especially toolset developers, avoiding unnecessary, complicated tailoring of their content processing tools to specific areas of the workflow, and thus improve the repurposing of their tools to different parts of the workflow.

This Technical Report will first talk about the types of common standard colour volumes and the properties associated to define them. Then for each common standard colour volume what different parameters of the associated properties are in use in industry. Lastly depending on domain (production, distribution) how each of these associated properties are carried.

# Domains- capture, production, production distribution, service distribution

The following diagram describes the different domains that video content may exist and can be edited or converted in a workflow. Similar conversions may be done in different domains and tools need to understand what conversions happened in the preceding domain.



Figure 1 Video workflows through different carriage domains

[Ed. The full and narrow range level property is not always carried in some wrappers, but common practice should be that YCbCr uses narrow range and RGB uses full range. ATSC 3.0 allows either Full-Range or Narrow-Range signals with the PQ TF.]

# Common combinatorial descriptions

Separately defining video properties with all of their possible enumerated values can fully describe all cases of the video property in isolation. [Ed. clarify this sentence ] Describing video content requires looking collectively at the groups of video properties. Some combination of values of the video properties may collectively not make sense (e.g. xxx) or are not being used by the industry. This section defines commonly associated video properties and those combinations of values in use for the content industry today. Additionally, it defines for each Domain (Capture, Production, Production-Distribution, Service-Distribution) what methods of carriage are commonly used to convey these video properties.

## Colour properties descriptions

Common Colour Volumes describes combinations of video properties that are needed to do conversions on colour volumes. [Ed. Note (Alexis): need to define colour volume in the definitions table ] These may be conversions to other colour volumes, changes in bit depth, changes in colour sampling, non-linear optimizations and may also include transformations based on carriage and bit rate restrictions. Tools performing these conversions need to know the values of these video properties to decode the video and consistently perform its operations.

Carriage formats for colour properties in each of the domains (Capture, Production, Production Distribution, Service Distribution) contain the same payload but in different wrappers. In the capture and production domains, the colour property information can be carried in an MXF wrapper using a Generic Picture Essence descriptor as described by Annex C of ST 2067-21[[1]](#footnote-1). Colour Volume Information in the distribution domain can be carried within the video stream as syntax information in the selected video format such as MPEG-4 AVC, HEVC, MPEG-2 etc, through VUI headers according to ISO/IEC 29001-2. The full and narrow range level video property is not carried in any wrapper but can be identified through the system identifier tags which may be sent as out-of-stream information.

[ Ed. YS: (From Feedback) The full and narrow range level property is not always carried in some wrappers, but common practice should be that YCbCr uses narrow range and RGB uses full range. Also ATSC 3.0 allows either Full-Range or Narrow-Range signals with the PQ TF.]

### Defined properties

* Colour Primary [ref- coding points]
* Electro-Optical Transfer Function [ref- coding points]
* Colour Representation [ref- coding points]
* Full/Narrow Level Range Scale [ref]

### Common descriptions and carriage: Standard Dynamic Range- Narrow Colour Gamut

This colour volume describes SDR video, which includes the majority of the production and distribution workflows used in the industry today. There are several combinations of values of these video properties that are used for this colour volume. Table 1 describes these combinations. Note there are several one-way operations that can be performed for this colour Volume including bit depth reductions, colour sampling reductions, and full to narrow level range scale operations.

The following system identifier tags are described:

* BT.709 YCC
* BT.709 RGB
* BT.601

|  |  |  |  |
| --- | --- | --- | --- |
| Standard Dynamic Range & Narrow Colour Gamut | | | |
| System Identifier | BT.709 YCC | BT.709  RGB | BT.601 RGB |
| Colour properties |  |  |  |
| Colour Primaries | BT.709 | BT.709 | BT.601 |
| Electro-Optical Transfer Function (EOTF) (Logical) | BT.1886 | BT.1886 | BT.1886 |
| Colour Representation | Y′CbCr | R′G′B′ | R′G′B′ |
| Full/Narrow Range | Narrow | Full | Full |
|  |  |  |  |
| AVC/HEVC VUI parameters |  |  |  |
| colour primaries | 1 | [1] | [5 or 6] |
| transfer\_characteristics | 1 | [1] | [1] |
| matrix\_coeffs | 1 | [0] | [0] |
| VideoFullRangeFlag | 0 | [1] | [1] |
|  |  |  |  |
| SMPTE MXF parameters |  |  |  |
| IMF colour ID | Colour 3 |  |  |
| Colour Primaries | 06.0E.2B.34.04.01.01.06.04.01.01.01.03.03.00.00 | 06.0E.2B.34.04.01.01.06.04.01.01.01.03.03.00.00 |  |
| Transfer Characteristic | 06.0E.2B.34.04.01.01.0E.04.01.01.01.01.09.00.00 | 06.0E.2B.34.04.01.01.01.04.01.01.01.01.09.00.00 | 06.0E.2B.34.04.01.01.01.04.01.01.01.01.09.00.00 |
| Coding Equations | 06.0E.2B.34.04.01.01.0D.04.01.01.01.02.06.00.00 |  |  |
| ′ | 8242 | 2032 | [UNICODE char for math prime] |

Table 1 SDR common colour volume descriptions[[2]](#footnote-2)

[ Ed . Should BT.601 RGB remain in the above table or add BT.601 YCC. ?]

[ Ed. Colour primaries for BT.601 in SMPTE MXF seems to be the same as those for RGB. However, the transfer characteristics are not. YS: need to double check ]

### Common descriptions and carriage: Standard Dynamic Range- Wide Colour Gamut

This colour volume describes SDR with Wide Colour Gamut video, which is keyed off the colour primary video property. Please note that the same colour is described with two different values depending on the colour primary container used. It is important for tools to understand what colour volume it is operating in to make sure the conversion is handled consistently.

The following system identifier tags are described:

* BT.2020 Gamma YCC
* BT.2020 Gamma RGB



Table 2 SDR w/ WCG common colour volume descriptions[[3]](#footnote-3)

[Ed. Note (Alexis) : MXF values seem wrong and seem to be copying the BT.709 ones. Please doublecheck]

[Ed. GJS: Format as a Word table with formatting consistent with usual specification text documents.]

[Ed. GJS: What do those hex strings mean?]

### Common descriptions and carriage: High Dynamic Range- Wide Colour Gamut

This colour volume describes what is known as HDR video, which is used in current Ultra High Definition video technologies. Tools including static and dynamic LUTs, tone mapping, and bit rate compression need to understand these combination values to produce consistent conversions.

The following system identifier tags are described:

* BT.2100 PQ YCC
* BT.2100 HLG YCC
* BT.2100 PQ RGB
* BT.2100 HLG RGB



Table 3 HDR/WCG common colour volume descriptions

[Ed. Note (Alexis) : Both of these transfer functions are in BT.2100. Please use that reference also.]

## Common pixel format workflow descriptions

Several video properties can be associated as video pixel format information. These types of video properties can be converted as it moves through each domain (Capture, Production, Production Distribution, Service Distribution) in the workflow. For distribution domains, bit depth, colour subsampling, colour form, and compression profile can be discovered through codec profile definition and stream syntax. The remainder of video properties in this description is not carried in any wrapper but can be identified through the system identifier tags which may be sent as out of stream information.

### Defined properties

* Bit Depth [ref]
* Colour SubSampling[[4]](#footnote-4) [ref]
* Colour Form [ref]
* Compression Dimension [??]
* Compression Profile [??][[5]](#footnote-5)

The pixel format parameters and values are described in the following table:



Table 4 Pixel format description parameters & values

[Ed. Note (Alexis) : There are also intra only and still profiles. This does not seem to include those.]

## Common descriptions and carriage: pixel format workflow descriptions for live

The following system identifier tags are described:

* Y’CrCb-444-16-N-UC
* Y’CrCb-422-12-N-UC
* Y’CrCb-422-10-N-UC
* Y’CrCb-420-8-ST-AVCMH
* Y’CrCb-420-10-ST-HEVCM10
* Y’CrCb-420-8-ST-HEVCM10
* Y’CrCb-420-8-ST-HEVCM
* Y’CrCb-420-8-ST-MPEG2M

[Ed. Note: Is there also 4:2:0 Uncompressed pixel formats?]

Table 5 Pixel format workflow system identifiers- live

## Common descriptions and carriage: pixel format workflow descriptions for file

The following system identifier tags are described:

* R’G’B’-444-16-N-UC

[ED. Note (Alexis): Some compression (potentially lossless) may be used for such formats, e.g. within a TIFF file. Is this sufficient?]

* Y’CrCb-420-8-ST-AVCMH
* Y’CrCb-420-10-ST-HEVCM10
* Y’CrCb-420-8-ST-HEVCM10
* Y’CrCb-420-8-ST-HEVCM
* Y’CrCb-420-8-ST-MPEG2M

Table 6 Pixel format workflow system identifiers- file

## Mastering Display Colour Volume descriptions

The Mastering Display Information describes the colour volume of a display that was used for viewing during the authoring of the video content. It is the display where creative work is performed during the mastering process, and can be represented by a combination of video display properties. When the authored content is shown on other displays, this type of video display information can be used to more closely reproduce the creative intent originally achieved in the mastering suite. Again, only certain combinations of video display properties are used in the mastering display environment. Tools performing conversions of video content would be able to preserve more of the intended colour reproduction with knowledge of these combinations of in-use values for the video display property parameters.

### Defined Properties

* Mastering Display Primaries [ref- ST 2086]
* Mastering Display White Point Chromaticity [ref- ST 2086]
* Mastering Display Maximum Luminance [ref- ST 2086]
* Mastering Display Minimum Luminance [ref- ST 2086]

[Ed. Note (Alexis): These also exist in the HEVC and AVC specs. Ideally they should also be in CICP and we should be referencing those specs also.YS: Agreed

Also why reference each individual property of a group?YS: The identifiers differ with respect to different values of these parameters.]

### Common descriptions and carriage: MDCV descriptions

The following system identifier tags are described:

* 709,D65,1000- represents display colour volume environments dealing with mastering of SDR/HDR content on LED/LCD displays capable of 1000 cd/m2 of peak brightness but restricted to BT.709 colourspace.
* 709,D65,4000- represents display colour volume environments dealing with mastering of SDR/HDR content on LED/LCD displays capable of 4000 cd/m2 of peak brightness but restricted to BT.709 colourspaces.
* 2020,D65,1000- represents display colour volume environments dealing with mastering of HDR content on OLED displays capable of 1000 cd/m2 of peak brightness and wider BT.2020 Colourspace.
* P3,D65,1000- represents display colour volume environments dealing with mastering of HDR content on OLED displays capable of 1000 cd/m2 of peak brightness, having overall a smaller gamut than BT.2020 (except for a slightly deeper red P3 primary), often used in mastering digital cinema content as described in Table 6-10 of [8]. Various white points (D60, D62, D65) are defined in Table G-1 of [9]

[Ed: Examples; table entry for content colour volumes at 1200 and and 4000 cd/m^2 for BT.2020/BT.2100 signals?.]

[Ed. TF needs to be understood for mastering displays input. Not relevant for ST 2086 however ]

[Ed. most of today’s video is 709-Gamma D65 with 200- 300 cd/m2 ]

[Ed. YS: This is one of those areas that we probably need to discuss]

Carriage formats for mastering display colour volume in each of the domains (Capture, Production, Production Distribution, Service Distribution) contain the same payload but in different wrappers. In the capture and production domains, the MDCV information can be carried in an MXF wrapper using Generic Picture Essence descriptor as described by Annex C of ST 2067-21[[6]](#footnote-6). In the distribution domain, the mastering display colour volume information is carried the mastering display colour volume SEI message (payload 137) that needs to be repeated at least every random access point access unit (RAPAU).

Table 7 Mastering display colour volume descriptions

[Ed. Note (Alexis) : For MXF, instead of numbers, should the actual byte values be written?

[Ed. For BT.2020, the red primary has an x coordinate of 0.708 \* 50000 = 35400. Please verify all points

YS: Yes needs to be double checked. ]

[Ed. Maximum luminance seems to be scaled by 10000, not 1000 as reported on the first column, row 6.]

[Ed. There is nothing that mentions that the values in this table are scaled by 50000 ].

[Ed. In the SEI messages as well CICP spec, the order is green, blue, red. Not red, green, blue. So this seems maybe problematic. Please fix.]

[Ed. Check whether MXF and IMF values for ST 2086 carriage may be in R,G,B order. SMPTE 2067-21 ]

# [Document End]

1. In the capture, production, and distribution domains, the colour property information can also be carried as metadata in a MOV wrapper in the ‘nclc’ atom. [↑](#footnote-ref-1)
2. For Quicktime MOV Wrappers , Colour Primary-EOTF-Colour Representation parameter values are combined into 1 entry called NCLC that is the MPEG values of these parameters in a dash delinated string sequence. For example, Colour Primary = 1/ EOTF=1/Colour Representation = 1 are combined into “1-1-1” for the NCLC value. [↑](#footnote-ref-2)
3. For Quicktime MOV Wrappers , Colour Primary-EOTF-Colour Representation parameter values are combined into 1 entry called NCLC that is the MPEG values of these parameters in a dash delinated string sequence. For example, Colour Primary = 1/ EOTF=1/Colour Representation = 1 are combined into “1-1-1” for the NCLC value. [↑](#footnote-ref-3)
4. Information about chroma sample location use is defined in HEVC- 23008-2 Edition 5 (TBD). BT.2020/BT.2100 chroma location is 2, BT.709 chroma location is 1. [↑](#footnote-ref-4)
5. In some production domain compression profiles, Apple ProRes HQ for broadcast and Apple ProRes 4444 for film and graphics could also be in use. [↑](#footnote-ref-5)
6. In the capture, production, and distribution domains, display colour volume information can also be carried as metadata in an MOV wrapper in the ‘mdvc’ atom that accompanies video formats such as RAW, DPX, OpenEXR. [↑](#footnote-ref-6)