**3GPP TSG-SA4 Meeting #131S4-250033**

**Geneva, Switzerland, 17th Feb 2025 - 21st Feb 2025**

**Source: Qualcomm Incorporated, Tencent**

**Title: Pseudo-CR on [VOPS] Video Operation Points**

**Spec: 3GPP TS26.265v0.5.1**

**Agenda item: 9.5**

**Document for: Decision**

**1. Introduction**

Video Operation Points are underspecified in the current version.

**2. Reason for Change**

Video Operation Points are underspecified in the current version.

**3. Conclusions**

Video Operation Points are underspecified in the current version.

**4. Proposal**

It is proposed to agree the following changes to 3GPP TS26.265v0.5.1.

\* \* \* First Change \* \* \* \*

### 4.4.2 Video signal parameters

Video signals considered in this specification are represented by a sequence of pictures, where a *picture* can represent either an array of *luma* samples in a monochrome format or an array of luma samples and two corresponding arrays of *chroma* samples in a 4:2:0, 4:2:2, or 4:4:4 colour format. Only *progressive* signals are considered. A component refers to an array or single sample from one of the three arrays (luma and two chroma) that compose a picture. The Luma component represents a sample array or single sample representing the monochrome signal related to the primary colours (denoted with the symbol *Y*), and a chroma component represents a sample array or single sample representing one of the two colour difference signals related to the primary colours, represented by the symbols *Cb* and *Cr*.

Video signals are typically described by a set of parameters that are required for the proper rendering of the decoded signal. Table 4.4.2-1 documents typical video signal parameters and provides a definition and/or reference.

Table 4.4.2-1 Video Signal Parameters

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Definition | 3GPP restrictions | Service or Application restrictions |
| Spatial Resolution width | The number of active samples per line for the luma component.  Example values are 1280 or 1920 for HD, and 3840 for UHD.  NOTE: The width does not restrict the encoding resolution to fixed values. Cropping parameters can be indicated that prescribe decoders the need to remove spatial video samples in a partially filled coding block that are not intended for presentation. | No restrictions | Restrictions possible |
| Spatial Resolution height | The number of active lines per picture for the luma component.  Example values are 720 or 1080 for HD, and 2160 for UHD.  NOTE: The height does not restrict the encoding resolution to fixed values. Cropping parameters can be indicated that prescribe decoders the need to remove spatial video samples in a partially filled coding block that are not intended for presentation. | No restrictions | Restrictions possible |
| Scan Type | Indicates the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273.  Typical value is progressive | Progressive only |  |
| Chroma format indicator | Indicates whether the picture has only a luma component or that the picture has three colour components that consist of a luma component and two associated chroma components, such that the width and height of each chroma component are the width and height of the luma component divided by a factor defined by the chroma format as defined in Rec. ITU-T H.274, clause 7.3. | 4:2:0 |  |
| Bit depth | Indicates the bit depth for the samples of the luma component and the samples of the two associated chroma components.  Note that in general, the bit depth of the luma component and of the two associated chroma components may differ.  Typical values are 8 or 10 bits. | 8 or 10 bits  Luma and chroma components shall not differ |  |
| Colour primaries | Indicates the chromaticity coordinates of the source colour primaries as specified in clause 8.1 of Rec. ITU-T H.273.  Typical values are 1 to refer to Rec. ITU-R BT.709-6 [bt709] or 9 to refer to Rec. ITU-R BT.2020-2 and Rec. ITU-R BT.2100-2. | BT.709 or BT.2020/BT.2100 |  |
| Transfer Characteristics | Either indicates the reference opto-electronic transfer characteristic function of the source picture as a function of a source input linear optical intensity input or indicates the inverse of the reference electro-optical transfer characteristic function as a function of an output linear optical intensity as defined in clause 8.2 of Rec. ITU-T H.273.  Typical values are 1 to refer to Rec. ITU-R BT.709-6, 14 to refer to Rec. ITU-R BT.2020-2 (10 bit), 16 to refer to the Rec. ITU-R BT.2100-2 perceptual quantization (PQ) system, or 18 to refer to the Rec. ITU-R BT.2100-2 hybrid log-gamma (HLG) system | BT.709 SDR, BT.2020 SDR, BT.2100 PQ, or BT.2100 HLG |  |
| Matrix Coefficients | Describes the matrix coefficients used in deriving the luma and chroma signals from the green, blue, and red primaries. A video full range flag may be supplied with this parameter specifying the scaling and offset values applied in association with the Matrix coefficients. For detailed definition refer to clause 8.2 of Rec. ITU-T H.273.  Typical values are 1 to refer to the non constant luminance YCbCr representation in Rec. ITU-R BT.709-6 or 9 to refer to the non constant luminance YCbCr representations in Rec. ITU-R BT.2020-2 and Rec. ITU-R BT.2100-2. | YCbCr BT.709, YCbCr BT.2020, or YCbCr BT.2100 |  |
| Frame rate | Typical values, using frames per second, are: 120, 120/1.001, 100, 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 | No restrictions | services may only permit a restricted subset |
| Frame packing | Indicates a frame packing arrangement, if present, as defined in clause 8.4 of Rec. ITU-T H.273. | Typically restricted to no frame packing. | Some applications may use frame packing. |
| Projection | Indicates a projection, if present, as defined in Rec. ITU-T H.274, clause 7.3, and typically refers to packing arrangements in clause 8.6 of Rec. ITU-T H.274. | Typically restricted to no projection. | Some applications may use projections. |
| Sample aspect ratio | Indicates width-to-height aspect ratio of the luma samples of the associated pictures as defined in clause 7.3 of Rec. ITU-T H.273.  Typical value is 1 | No specific restrictions, but 1 is expected. |  |
| Chroma sample location type | Specifies the location of the chroma samples relative to the luma samples for frames as defined in Rec. ITU-T H.273, clause 8.7.  Typical values are 0 (chroma samples are horizontally co-sited with and vertically centered between the first luma sample at the top-left corner and the first two luma samples at the top-left corner, respectively) or 2 (chroma samples are co-sited with the luma sample at the top-left corner).  Note that a value of 1 is common for still images. | No specific restrictions, but 0 is expected if not present. For HDR the value is typically set to 2. |  |
| Range | Specifies how luma and chroma samples are represented in digital video as defined in Rec. ITU-T H.273, clause 8.3 using the parameter VideoFullRangeFlag.  For video applications only the value set to 0 is used, i.e. the video range or restricted range is applied where the luma values range from 16 to 235 in an 8-bit system, and chroma values range from 16 to 240. For 10-bit systems, the values are multiplied by 4.  Note that for still images full range (value set to 1) is commonly used. | No specific restrictions, but 0 is expected if not present. |  |
| Stereoscopic Video | Visual media may be stereoscopic, in which case a view is available to be presented to the left eye and another view is available to be presented simultaneously to the right eye. The presentation of both the left and right views allows for an effect known as stereopsis, which can be defined as "the perception of depth produced by the reception in the brain of visual stimuli from both eyes in combination; binocular vision."  For signal representations, [3dtv] recommends that the Left and Right eyes comply to regular image formats such as Rec. ITU-R BT.709 and any necessary 3D-specific metadata is incorporated with the data. Hence, for stereoscopic video, two synchronized video signals are available, each with identical format parameters (such as the ones defined in this table).  Additional metadata that may be added with stereoscopic video:  - “Hero eye” is the default eye in a stereo (stereoscopic) video pair, often determined by tags set by the cameras used to capture the video. If so signaled, this indicates that the other stereo eye view is derived from the specified stereo eye and may be useful when choosing which eye to use in a monoscopic viewing environment. There is no requirement that either of the two eyes (or views) is tagged as the hero eye, in which case no hero eye tagging may be present. |  |  |

\* \* \* Next Change \* \* \* \*

#### 4.4.3.2 High-Definition TV

3GPP High-Definition TV (HDTV) formats are defined based on Rec. ITU-R BT-709-6 [bt709]. 3GPP HDTV formats shall conform to Rec. ITU-R BT-709-6 [bt709] with the following restrictions:

- Only the following formats are included 24/P, 25/P, 30/P, 50/P and 60/P. Interlace and progressive segmented frame signals are excluded.

- Only the Non-Constant Luminance YCbCr signal format is considered.

An informative summary of the parameters of a 3GPP HDTV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.2-1.

Table 4.4.3.2-1 Video Signal Parameters for 3GPP HDTV format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Spatial Resolution width | the number of active samples per line is 1920.  NOTE 1: Down-sampled resolutions may be created for distribution, for example in case of adaptive streaming. |
| Spatial Resolution height | the number of active lines per picture for the luma component is 1080.  NOTE 2: Down-sampled resolutions may be created for distribution, for example in case of adaptive streaming.  NOTE 3: To accommodate the block coding structure of a given specification, quite often the encoded signal may be padded. In such cases, normative cropping is typically applied to remove spatial samples that are not intended to be presented.. |
| Scan Type | The source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive. |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted values are 8 or 10 bit. The bit depth is the same for all samples. |
| Colour primaries | Only the value 1, as defined in clause 8.2 of Rec. ITU-T H.273, is permitted. |
| Transfer Characteristics | Only the value 1, as defined in clause 8.2 of Rec. ITU-T H.27,3 is permitted. |
| Matrix Coefficients | Only the value 1, as defined in clause 8.2 of Rec. ITU-T H.273, is permitted. |
| Frame rates | The permitted values are 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | The location of the chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 0 (chroma samples are horizontally co-sited with and vertically centered between the first luma sample at the top-left corner and the first two luma samples at the top-left corner, respectively). |
| Range | The restricted video range is used. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted values are 8 or 10 bit. |
| Colour primaries | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Transfer Characteristics | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Matrix Coefficients | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Frame rates | The permitted values are 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 0 (chroma samples are horizontally co-sited with and vertically centered between the first luma sample at the top-left corner and the first two luma samples at the top-left corner, respectively). |
| Range | The restricted video range is used. |

\* \* \* Next Change \* \* \* \*

## 4.5 Common Bitstream definitions

### 4.5.1 General

This clause defines common definitions for bitstreams that are used in capability definitions in the remainder of this document.

### 4.5.2 AVC Bitstreams

The following definitions are provided for AVC/ITU-T H.264 [h264] bitstreams.

### 4.5.3 HEVC Bitstreams

The following definitions are provided for HEVC/ITU-T H.265 [h265] bitstreams.

An **HEVC-Progressive Bitstream** is definedas an HEVC/ITU-T H.265 [h265] bitstream that has set the following flags in the active Sequence Parameter Set (SPS):

- general\_progressive\_source\_flag shall be set to 1,

- general interlaced\_source\_flag shall be set to 0,

- general\_non\_packed\_constraint\_flag shall be set to 1, and

- general\_frame\_only\_constraint\_flag shall be set to 1.

An **HEVC-Format** Bitstream is defined as an HEVC/ITU-T H.265 [h265] bitstream for which the

- Video Parameter Sets (VPS) NAL units as defined in Recommendation ITU-T H.265 / ISO/IEC 23008-2 [h265] may be present, but the Bitstream shall be valid if the Receiver ignores the VPS.

- the Video Usability Information (VUI) is present in the active Sequence Parameter Set, i.e. the vui\_parameters\_present\_flag shall be set to 1.

- in the VUI,

- the aspect ratio information is present, i.e. the aspect\_ratio\_info\_present\_flag value shall be set to 1,

- the colour parameter information is present, i.e. video\_signal\_type\_present\_flag value shall be set to 1 and the colour\_description\_present\_flag value shall be set to 1.

- only video range signals are used, i.e. the video\_full\_range\_flag shall be set to 0,

- no overscan signalling is present, i.e. the overscan\_info\_present\_flag shall be set to 0,

- the chroma location shall be signalled, i.e. chroma\_loc\_info\_present\_flag shall be set to 1,

- The timing information may be present. If the timing information is present, i.e. the value of vui\_timing\_info\_present\_flag is set to 1, then the values of vui\_num\_units\_in\_tick and vui\_time\_scale shall be set according to the frame rates allowed for each operation point. The timing information present in the video Bitstream should be consistent with the timing information signalled at the system level. The frame rate shall not change between two RAPs. fixed\_frame\_rate\_flag value, if present, shall be set to 1.

NOTE: CMAF does recommend to not change the frame rate within an entire CMAF track.

\* \* \* Next Change \* \* \* \*

#### 4.4.3.3 High Dynamic Range TV

3GPP High Dynamic Range (HDR) TV formats are defined based on Rec. ITU-R BT-2100-2 [bt2100]. 3GPP HDR TV formats shall conform to ITU-R BT-2100-2 [bt2100] with the following restrictions:

- Only 4:2:0 colour subsampling is considered

- Only the Non-Constant Luminance Y'C'BC'R signal format is considered

- Only 10-bit representations are considered

An informative summary of the parameters of a 3GPP HDR TV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.3-1.

Table 4.4.3.3-1 Video Signal Parameters for 3GPP HDR TV format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Picture aspect ratio | 16:9 |
| Spatial Resolution width x height | 7680 × 4320, 3840 × 2160, 1920 × 1080  NOTE: To accommodate the block coding structure of a given specification, quite often the encoded signal may be padded. In such cases, normative cropping is typically applied to remove spatial samples that are not intended to be presented.. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted value is 10 bit. |
| Colour primaries | Only the value 9 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Transfer Characteristics | Only the values 14 (for SDR with WCG), 16 (for PQ) and 18 (for HLG) as defined in clause 8.2 of Rec. ITU-T H.273 are permitted. |
| Matrix Coefficients | Only the value 9 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Frame rates | The permitted values are 120, 120/1.001,100, 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 2 (chroma samples are co-sited with the luma samples at the top-left corner). |
| Range | The restricted video range is used. |

#### 

\* \* \* Next Change \* \* \* \*

#### 4.4.3.4 3GPP Stereoscopic Cinema Format

The stereoscopic 3D TV format uses two signals, one for the left eye and another view for the right eye as defined in Table 4.4.2-1. The components for each eye closely follow the specifications of the 3GPP HDR signals, but there are some restrictions and extensions, namely:

- Only 4:2:0 colour subsampling is considered.

- Frame rates include high frame rate for movies, namely 48 fps.

- the spatial resolution is restricted to a maximum value of 4K.

- Only the Non-Constant Luminance Y'C'BC'R signal format is considered.

An informative summary of the parameters of a 3GPP Stereoscopic 3D TV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.4-1.

Table 4.4.3.4-1 Video Signal Parameters for 3GPP Stereoscopic 3D Cinema format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Picture aspect ratio | 16:9 |
| Spatial Resolution width x height | 3840 × 2160, 1920 × 1080  NOTE: To accommodate the block coding structure of a given specification, quite often the encoded signal may be padded. In such cases, normative cropping is typically applied to remove spatial samples that are not intended to be presented. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted values are 8 or 10 bit. 8 bit is only permitted for SDR. |

|  |  |
| --- | --- |
| Colour primaries  Transfer Characteristics  Matrix Coefficients | Only the following value combinations are permitted: (1, 1, 1), (9, 14, 9), (9, 16, 9), and (9, 18, 9) for SDR HD, SDR UHD, HDR PQ, and HDR HLG, respectively. |
| Frame rates | The permitted values are 60, 60/1.001, 48, 48/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | For SDR HD, the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 0.  For SDR UHD, HDR PQ, and HLG, the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 2. |
| Range | The restricted video range is used. |
| Stereoscopic Video | A signal for the Left and for the Right Eye is provided whereby the signals have the identical parameters as above and are timely synchronized. |

## \* \* \* Next Change \* \* \* \*4.5 Common Bitstream Constraints

### 4.5.1 General

This clause defines common definitions for bitstreams that are used in capability definitions in the remainder of this document.

### 4.5.2 AVC Bitstreams

The following definitions are provided for AVC/ITU-T H.264 [h264] bitstreams.

### 4.5.3 HEVC Bitstreams

The following definitions are provided for HEVC/ITU-T H.265 [h265] bitstreams.

For an HEVC/ITU-T H.265 [h265] bitstream, *progressive constraints* are defined that the following flags in the active Sequence Parameter Set (SPS):

- general\_progressive\_source\_flag shall be set to 1,

- general interlaced\_source\_flag shall be set to 0,

- general\_non\_packed\_constraint\_flag shall be set to 1, and

- general\_frame\_only\_constraint\_flag shall be set to 1.

For an HEVC/ITU-T H.265 [h265] bitstream, *VUI constraints* are defined:

- Video Parameter Sets (VPS) NAL units as defined in Recommendation ITU-T H.265 / ISO/IEC 23008-2 [h265] may be present, but the Bitstream shall be valid if the Receiver ignores the VPS.

- The Video Usability Information (VUI) is present in the active Sequence Parameter Set, i.e. the vui\_parameters\_present\_flag shall be set to 1.

- In the VUI,

- the aspect ratio information is present, i.e. the aspect\_ratio\_info\_present\_flag value shall be set to 1,

- the colour parameter information is present, i.e. video\_signal\_type\_present\_flag value shall be set to 1 and the colour\_description\_present\_flag value shall be set to 1.

- only video range signals are used, i.e. the video\_full\_range\_flag shall be set to 0,

- no overscan signalling is present, i.e. the overscan\_info\_present\_flag shall be set to 0,

- the chroma location shall be signalled, i.e. chroma\_loc\_info\_present\_flag shall be set to 1,

- the timing information may be present. If the timing information is present, i.e. the value of vui\_timing\_info\_present\_flag is set to 1, then the values of vui\_num\_units\_in\_tick and vui\_time\_scale shall be set according to the frame rates allowed for each operation point. The timing information present in the video Bitstream should be consistent with the timing information signalled at the system level. The frame rate shall not change between two RAPs. fixed\_frame\_rate\_flag value, if present, shall be set to 1.

NOTE: CMAF does recommend to not change the frame rate within an entire CMAF track.

## 4.6 Reference API parameters

### 4.6.1 Introduction

When media is played back, the decoder and the playback pipeline need to be initialized. For this purpose, certain parameters are required. In CTA-5003 [DPC], a media playback model is described that is aligned with HTML 5.1 and the <video> element, as well as the Media Source Extensions.

### 4.6.2 Video Decoder API Parameters

Based on CTA-5003 [DPC], Table 4.6.2-1 provide relevant parameters that need to be attached to the content, in order to establish media playback properly, and serve as an API. The parameters are used for the following purposes:

- to identify the capability of the device in order to check whether the signal can be played back

- to initialize the decoding and playback platform to allocate the resources for decoding and rendering

Table 4.6.2-1 Video Decoder API Parameters

|  |  |  |
| --- | --- | --- |
| Parameter | Restrictions | Status |
| width | Specifies the width of a video player, in pixels | required |
| height | Specifies the width of a video player, in pixels. | required |
| media type | Specifies the media type of the component, in this case video | required |
| format | Specifies the format of the media, for example mp4 | required |
| profiles | Specifies the profile of the format, for example 'cmfc' | optional |
| codecs | Specifies through a well-defined string the codec used for the signal | required |
| Video format parameters | Specifies additional video format parameters as defined in Table 4.4.2.1 to describe the signal and to initialize the encoder. | optional |

Editor’s Note: The capability of such API for decoding and playback of multilayer content, e.g. for stereoscopic content needs to be documented.

### 4.6.3 Video Encoder API Parameters

Video encoder API parameters are for further study.

\* \* \* Next Change \* \* \* \*

5.3 Single-Instance Decoding Capabilities

5.3.1 AVC Decoding Capabilities

The following decoding capabilities are defined:

**- AVC-FullHD-Dec**: the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 4.0 [h264] bitstreams.

**- AVC-UHD-Dec:** the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 5.1 [h264] bitstreams with the following additional requirements:

- the maximum VCL Bit Rate is constrained to be 120 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively; and,

- the bitstream does not contain more than 10 slices per picture.

**- AVC-8K-Dec:** the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 6.1 [h264] bitstreams with the following requirements:

- the maximum VCL Bit Rate is constrained to be 120 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively; and,

- the bitstream does not contain more than 16 slices per picture.

- the bitstream shall not include horizontal motion vector component values that exceed the range from −2048 to 2047, inclusive, or that have vertical motion vector component values that exceed the range from −512 to 511, inclusive, in units of ¼ luma sample displacement. This constraint should be indicated by using values of log2\_max\_mv\_length\_horizontal less than or equal to 11 and values of log2\_max\_mv\_length\_vertical less than or equal to 9.

5.3.2 HEVC Decoding Capabilities

The following decoding capabilities are defined:

- **HEVC-HD-Dec**: the capability to decode bitstreams conforming to both, HEVC/ITU-T H.265 Main Profile, Main Tier, Level 3.1 [h265] bitstreams with *progressive* constraints as defined in clause 4.5.3.

- **HEVC-FullHD-Dec**: the capability to decode bitstreams conforming to HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 4.1 [h265] bitstreams with *progressive* constraints as defined in clause 4.5.3.

- **HEVC-UHD-Dec**: the capability to decode bitstreams conforming to HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 5.1 [h265] bitstreams with *progressive* constraints as defined in clause 4.5.3.

- **HEVC-8K-Dec**: the capability to decode bitstreams conforming to HEVC/ITU-T H.265 Main10 Profile, Main Tier, Level 6.1 [h265] bitstreams with *progressive* constraints as defined in clause 4.5.3 and further constraints:

- **MV-HEVC-UHD-Dec**: the capability to decode bitstreams with an HEVC/ITU-T H.265 Main 10 Profile base layer (layer\_id=0), and a single HEVC/ITU-T H.265 Multiview Main 10 [or Multiview Extended 10] layer (layer\_id=1) [h265]. Each layer shall conform to Main Tier, Level 5.1, while the device should be capable of supporting single layer decoding of HEVC/ITU-T H.265 Main 10 Profile bitstreams at Main Tier, Level 5.2. All layers shall follow the *progressive* constraints as defined in clause 4.5.3.

Editor’s Note: Adding operating point(s) for frame packed stereoscopic video is FFS.

\* \* \* Next Change \* \* \* \*

## 5.4 Single-Instance Encoding Capabilities

The following encoding capabilities are defined:

**- AVC-FullHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *AVC-FullHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

- up to 245,760 macroblocks per second;

- up to a frame size of 8,192 macroblocks;

- up to 240 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bit;

NOTE 1: The 3GPP HDTV format if restricted to 8 bit as defined in clause 4.4.3.2 may be encoded with an AVC-FullHD-Enc capable encoder.

- **HEVC-HD-Enc**: the capability to encode a video signal with

- up to 33,177,600 luma samples per second;

- up to a luma picture size of 983,040 samples;

- up to 120 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bit;

to a bitstream that is decodable by a decoder that is **HEVC-HD-Dec** capable as defined in clause 5.3.

NOTE 2: A restricted version of the 3GPP HDTV format as defined in clause 4.4.3.2 may be encoded with an HEVC-HD-Enc capable encoder.

**- HEVC-FullHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *HEVC-FullHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

- up to 133,693,440 luma samples per second;

- up to a luma picture size of 2,228,224 samples;

- up to 240 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being either 8 or 10 bit;

NOTE 3: The 3GPP HDTV format as defined in clause 4.4.3.2 may be encoded with an HEVC-FullHD-Enc capable encoder. A restricted version of the 3GPP HDR TV format as defined in clause 4.4.3.3 may be encoded with an HEVC-FullHD-Enc capable encoder.

**- HEVC-UHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *HEVC-UHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

- up to 534,773,760 luma samples per second;

- up to a luma picture size of 8,912,896 samples;

- up to 480 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being either 8 or 10 bit;

NOTE 4: The 3GPP HDTV format as defined in clause 4.4.3.2 may be encoded with an HEVC-FullHD-Enc capable encoder. A restricted version of the 3GPP HDR TV format as defined in clause 4.4.3.3 may be encoded with an HEVC-FullHD-Enc capable encoder.

\* \* \* Next Change \* \* \* \*

## 5.5 Multi-Instance Decoding Capabilities

The following multi-instance decoding capabilities are defined:

**- AVC-FullHD-Dec-2**: The capability of supporting up to two (*N*=2) concurrent decoder instances with the aggregate capabilities of *AVC-FullHD-Dec* as defined in clause 5.4.

**- AVC-UHD-Dec-4**: The capability of supporting up to four (*N*=4) concurrent decoder instances with the aggregate capabilities of *AVC-UHD-Dec* as defined in clause 5.4.

**- HEVC-UHD-Dec-4:** The capability of supporting up to four (*N*=4) concurrent decoder instances with the aggregate capabilities of *HEVC-UHD-Dec* as defined in clause 5.4.

**- UHD-Dec-4**: The capability supporting up to four (*N*=4) concurrent decoder instances with either:

- the aggregate capabilities of *AVC-UHD-Dec-4* as defined in this clause,

- the aggregate capabilities of *HEVC-UHD-Dec-4* as defined in this clause, or,

- the capability of decoding up to 4 bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-FullHD-Dec* or *HEVC-FullHD-Dec* as defined in clause 5.4.

**- AVC-8K-Dec-8:** The capability of supporting up to eight (*N*=8)concurrent decoder instances with the aggregate capabilities of *AVC-8K-Dec* as defined in clause 5.4.

**- HEVC-8K-Dec-8:** The capability of supporting up to eight (*N*=8)concurrent decoder instances with the aggregate capabilities of *HEVC-8K-Dec* as defined in clause 5.4.

**- 8K-Dec-8**: The capability supporting up to eight (*N*=8)concurrent decoder instances with either:

- the aggregate capabilities of *AVC-8K-Dec-8* as defined in this clause,

- the aggregate capabilities of *HEVC-8K-Dec-8* as defined in this clause, or,

- the capability of decoding up to:

- eight bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-FullHD-Dec* or *HEVC-FullHD-Dec* as defined in clause 5.4; or,

- four bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-UHD-Dec* or *HEVC-UHD-Dec* as defined in clause 5.4.

\* \* \* Next Change \* \* \* \*

# 6 Video Operation Points

Editor’s Note: A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and a video encoding format.

* See again S4-240619 for existing ones

## 6.1 Introduction

Video operation points define a restricted subset of representation signals and media capabilities. For each Video Operation Point, requirements for the Bitstream and for the Receiver are defined.

Table 6.1-1 provides an overview of defined video operation points.

Table 6.1-1 Overview of Video Operation Points

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Video Format | Decoding Capabilities | Definition |
| 3GPP-AVC-HD | 3GPP-HDTV (see clause 4.4.3.2) | AVC-FullHD-Dec (see clause 5.4) | 6.2.2 |
| 3GPP-HEVC-HD | 3GPP-HDTV (see clause 4.4.3.2) | HEVC-FullHD-Dec (see clause 5.4) | 6.3.2 |
| 3GPP-HEVC-HD-HDR | 3GPP-HDR (see clause 4.4.3.3) | HEVC-FullHD-Dec (see clause 5.4) | 6.3.3 |
| 3GPP-HEVC-UHD-HDR | 3GPP-HDR (see clause 4.4.3.3) | HEVC-UHD-Dec (see clause 5.4) | 6.3.4 |
| 3GPP-HEVC-3D | 3GPP-3DTV (see clause 4.4.3.4) | HEVC-UHD-Dec-2 (see clause 5.5) | 6.3.4 |
| 3GPP-MVHEVC-3D | 3GPP-3DTV (see clause 4.4.3.4) | MVHEVC-UHD-2 (see clause 5.5) | 6.3.6 |

## 6.2 AVC Video Operation Points

### 6.2.1 Introduction

The clause defines operation points for AVC. The video Bitstream and Receiver shall conform to Recommendation ITU-T H.264 [h264] with the restrictions described in this clause.

### 6.3.2 3GPP AVC HD Operation Point

#### 6.3.2.1 Introduction

The AVC HD Operation Point permits consistent distribution of HD-based video using AVC. The remainder of this clause 6.3.2 defines the Bitstream and Receiver requirements for the 3GPP-AVC-HD receiver.

Editor’s Note: Details need to be completed.

## 6.3 HEVC Video Operation Points

### 6.3.1 Introduction

The clause defines operation points for HEVC. The video Bitstream and Receiver shall conform to Recommendation ITU-T H.265 [h265] with the restrictions described in this clause.

### 6.3.2 3GPP HEVC HD Operation Point

#### 6.3.2.1 Introduction

The HEVC HD Operation Point permits consistent distribution of HD-based video using HEVC. The remainder of this clause 6.3.2 defines the Bitstream and Receiver requirements for the 3GPP-HEVC-HD receiver.

#### 6.3.2.2 Bitstream Requirements

A 3GPP-HEVC-HD Bitstream shall conform to the following requirements

- the Bitstream shall be an **HEVC-Progressive Bitstream** as defined in clause 4.5.3.

- the Bitstream shall be an **HEVC-Format** Bitstream as defined in clause 4.5.3.

- the Representation Format included in the Bitstream shall conform to the 3GPP-HDTV Representation format as defined in clause 4.4.3.2.

- the Bitstream shall be decodable by a decoder with **HEVC-FullHD-Dec** decoding capabilities.

Based on this, the following additional restrictions apply

- The chroma sub-sampling shall be 4:2:0 and the value of chroma\_format\_idc shall be set to 1.

- The aspect\_ratio\_idc value shall be set to 1, indicating a square pixel format.

- In the VUI, the values of colour\_primaries, transfer\_characteristics and matrix\_coeffs each shall be set to 1.

- The value of chroma\_sample\_loc\_type\_top\_field shall be set to 0.

The timing information may be present.

- If the timing information is present, i.e. the value of vui\_timing\_info\_present\_flag is set to 1, then the values of vui\_num\_units\_in\_tick and vui\_time\_scale shall be set according to the frame rates allowed for each operation point. The timing information present in the video Bitstream should be consistent with the timing information signalled at the system level.

- The frame rate shall not change between two RAPs. fixed\_frame\_rate\_flag value, if present, shall be set to 1.

#### 6.3.2.3 Receiver Requirements

Receivers conforming to the Operation Point 3GPP-HEVC-HD shall support decoding and rendering Bitstreams with the restrictions defined in clause 6.3.2.2.

NOTE 1: Rendering includes adherence to the parameters signalled in the bitstream to characterize the distributed Representation format.

Receivers should ignore the content of all Video Parameter Sets (VPS) NAL units as defined in Recommendation ITU-T H.265 / ISO/IEC 23008-2 [6].

NOTE 2: The VPS may be present to address requirements in other Operation Points, but the Bitstream also conforms to this Operation point.

There are no requirements on output timing conformance for H.265/HEVC decoding (Annex C of [6]). The Hypothetical Reference Decoder (HRD) parameters, if present, should be ignored by the Receiver.

### 6.3.3 3GPP HEVC HDR Operation Point

#### 6.3.3.1 Introduction

The HEVC HDR Operation Point permits consistent distribution of High Dynamic Range based video using HEVC. The remainder of this clause 6.3.3 defines the Bitstream and Receiver requirements for the 3GPP-HEVC-HDR receiver.

#### 6.3.3.2 Bitstream Requirements

A 3GPP-HEVC-HDR Bitstream shall conform to the following requirements

- the Bitstream shall be an **HEVC-Progressive Bitstream** as defined in clause 4.5.3.

- the Bitstream shall be an **HEVC-Format** Bitstream as defined in clause 4.5.3.

- the Representation Format included in the Bitstream shall conform to the 3GPP HDR TV Representation format as defined in clause 4.4.4.2.

- the Bitstream shall be decodable by a decoder with **HEVC-UHD-Dec** decoding capabilities.

Based on this, the following additional restrictions apply

- The chroma sub-sampling shall be 4:2:0 and the value of chroma\_format\_idc shall be set to 1.

- The aspect\_ratio\_idc value shall be set to 1, indicating a square pixel format.

- In the VUI, the values of colour\_primaries and matrix\_coeffs each shall be set to 9, and the value of transfer\_characteristics shall be set to one of the following values: 14 (for SDR with WCG), 16 (for PQ) and 18 (for HLG).

- The value of the chroma\_sample\_loc\_type\_top\_field shall be set to 2.

The timing information may be present.

- If the timing information is present, i.e. the value of vui\_timing\_info\_present\_flag is set to 1, then the values of vui\_num\_units\_in\_tick and vui\_time\_scale shall be set according to the frame rates allowed for each operation point. The timing information present in the video Bitstream should be consistent with the timing information signalled at the system level.

- The frame rate shall not change between two RAPs. fixed\_frame\_rate\_flag value, if present, shall be set to 1.

#### 6.3.3.3 Receiver Requirements

Receivers conforming to this Operation Point 3GPP-HEVC-HDR shall support decoding and rendering Bitstreams with the restrictions defined in clause 6.3.3.2.

NOTE 1: Rendering includes adherence to the parameters signalled in the bitstream to characterize the distributed Representation format.

Receivers should ignore the content of all Video Parameter Sets (VPS) NAL units as defined in Recommendation ITU-T H.265 / ISO/IEC 23008-2 [6].

NOTE 2: The VPS may be present to address requirements in other Operation Points, but the Bitstream also conforms to this Operation point.

There are no requirements on output timing conformance for H.265/HEVC decoding (Annex C of [6]). The Hypothetical Reference Decoder (HRD) parameters, if present, should be ignored by the Receiver.

### 6.3.4 3GPP HEVC UHD HDR

Editor’s Note: Details need to be completed.

### 6.3.5 3GPP HEVC 3D

Editor’s Note: Details need to be completed.

### 6.3.6 3GPP MVHEVC 3D

Editor’s Note: Details need to be completed.

\* \* \* End of Changes \* \* \* \*