**3GPP TSG-SA4 Meeting #131S4-250115**

**Geneva, 17th - 21st February 2025**

|  |
| --- |
| *CR-Form-v12.2* |
| **PSEUDO CHANGE REQUEST** |
|  |
|  | **26.265** | **CR** | **pseudo** | **rev** | **-** | **Current version:** | **0.5.0** |  |
|  |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **X** | Radio Access Network |  | Core Network | **X** |

|  |
| --- |
|  |
| ***Title:***  | [VOPS] Frame packed stereoscopic operating points |
|  |  |
| ***Source to WG:*** | Apple Inc., Samsung Electronics |
| ***Source to TSG:*** |  |
|  |  |
| ***Work item code:*** | VOPS |  | ***Date:*** | 2025-02-10 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***Release:*** | Rel-19 |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:Rel-8 (Release 8)Rel-9 (Release 9)Rel-10 (Release 10)Rel-11 (Release 11)…Rel-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19)* |
|  |  |
| ***Reason for change:*** | See discussion paper on Frame packed stereoscopic content for SA4 specifications. To summarize, there is a varying support for frame packed content in SA4 specifications:1. For generic MTSI clients in TS 26.114, for **HEVC** specifically, **frame packing is disallowed**. However, for the **special case of ITT4RT it is used normatively**.
2. MBMS specification provides support for frame packed stereoscopic 3D content for **H.264/AVC only**.
3. Various **HEVC decoding** configurations with frame packing are used in TS 26.118.
4. Misc. specs such as PSS (TS 26.234) need to be analyzed and decided upon.

Given the relatively non-uniform support for frame packing, it is proposed to provide uniform frame packing support for HEVC content in draft TS 26.265. Additionally, draft TS 26.265 text in Table 4.4.2-1 stating “Typically restricted to no frame packing, but applications may use frame packing” should be revised as there is no typical restriction. |
|  |  |
| ***Summary of change:*** | Adds frame packed stereo operating point. |
|  |  |
| ***Consequences if not approved:*** | Stereo operating points will remain missing. |
|  |  |
| ***Clauses affected:*** | 4.4.2, 4.4.3.4, 5.2.2, 5.3.2 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  | **X** |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  | **X** |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  | **X** |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** | Revision 1:* Added details on the SEI message needed for frame packing.
* Miscellaneous corrections including editorial ones.
* Added an EN that impact on clause 6 needs to be specified.
 |

\* \* \* 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** |
| 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 direct restrictions, but services may provide subsets. |
| Spatial Resolution height | The number of active lines per picture for the luma component.Typical values are 720 or 1080.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 direct restrictions, but services may provide subsets. |
| 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 |
| 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 Rec. ITU-R BT.2100-2 perceptual quantization (PQ) system, or 18 to refer to 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 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 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. | BT.709 or BT.2020/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 direct restrictions, but 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. | 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, but 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 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 VideoFullRangeFlag parameter. 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 is commonly used. | No specific restrictions, but 0 is expected if not present. |
| Stereoscopic Video | Visual media may be stereoscopic, in which 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 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 \* \* \* \*

4.4.3.4 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 YCbCr 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 × 1080NOTE: For 1080, typically the encoded signal has 1088 lines and cropping is applied to remove spatial samples that are not 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 primariesTransfer CharacteristicsMatrix 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. |
|  |  |
| 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 \* \* \* \*

5.2.2 Codec & profile & Levels

This specification defines capabilities based on the following video codec profile and levels:

- AVC/H.264 Progressive High Profile Level 3.1,

- AVC/H.264 Progressive High Profile Level 4.0,

- AVC/H.264 Progressive High Profile Level 4.2,

- AVC/H.264 Progressive High Profile Level 5.1,

- AVC/H.264 Progressive High Profile Level 6.1,

- HEVC/H.265 Main Profile Main Tier Level 3.1,

- HEVC/H.265 Main-10 Profile Main Tier Level 4.1,

- HEVC/H.265 Main-10 Profile Main Tier Level 5.0,

- HEVC/H.265 Main-10 Profile Main Tier Level 5.1,

- HEVC/H.265 Main 10 Profile Main Tier, Level 5.2,

- HEVC/H.265 Main-10 Profile Main Tier Level 6.1,

- HEVC/H.265 Multiview Main 10 Profile Main Tier Level 5.1,

- HEVC/H.265 Multiview Extended 10 Profile Main Tier Level 5.1.

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

5.3.2 HEVC Decoding Capabilities

The following decoding capabilities are defined:

- **HEVC-HD-Dec**: the capability to decode HEVC/ITU-T H.265 Main Profile, Main Tier, Level 3.1 [h265] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1 in all coded video sequences in the bitstream:.

- **HEVC-FullHD-Dec**: the capability to decode HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 4.1 [h265] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1 in all coded video sequences in the bitstream:.

- **HEVC-UHD-Dec**: the capability to decode HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 5.1 [h265] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1 in all coded video sequences in the bitstream:.

- **HEVC-8K-Dec**: the capability to decode HEVC/ITU-T H.265 Main10 Profile, Main Tier, Level 6.1 [h265] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1 with the following further limitations in all coded video sequences in the bitstream:

- **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 have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

Editor’s Note: Adding operating point(s) for 8k stereoscopic is FFS.

- **HEVC-Frame-Packed-Stereo-Dec**: the capability to decode HEVC/ITU-T H.265 Main 10 Profile bitstreams at Main Tier, Level 5.2. Such bitstreams shall have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, and general\_frame\_only\_constraint\_flag equal to 1 in all coded video sequences in the bitstream. If such bitstreams contain coded video sequences with the flag general\_non\_packed\_constraint\_flag set to 0, the frame packing arrangement SEI message can be present in such coded video sequences, with the following limitations:

If the frame packing arrangement SEI message is present in a coded video sequence, it shall have the following characteristics:

- The frame packing arrangement SEI message shall be present for the first frame in the coded video sequence and its information shall either persist for a series of frames or repeated for all frames in the coded video sequence.

- All parameters relating to a frame packing arrangement SEI message shall remain the same for the entire coded video sequence.

- The value of frame\_packing\_arrangement\_type shall be set to either the value of 3 for the side-by-side packing arrangement, or the value of 4 for the top-bottom/over-under packing arrangement.

- The value of quincunx\_sampling\_flag shall be set to 0.

- The value of content\_interpretation\_type shall be set to either 1 or 2.

- The value of spatial\_flipping\_flag shall be set to 0.

- The value of frame0\_flipped\_flag shall be set to 0.

- The value of field\_views\_flag shall be set to 0.

- The value of current\_frame\_is\_frame0\_flag shall be set to 0.

- The values of frame0\_grid\_position\_x, frame0\_grid\_position\_y, frame1\_grid\_position\_x, and frame1\_grid\_position\_y, shall remain the same throughout the coded video sequence.

- If the value of upsampled\_aspect\_ratio\_flag is set to 0, indicating the presence of full resolution frame packed video, then aspect\_ratio\_idc shall be set to 1. All parameters shall remain the same for the entire coded video sequence.

-

If the value of upsampled\_aspect\_ratio\_flag is set to 1, indicating the presence of half resolution frame packed video, then aspect\_ratio\_idc shall be set to 1.

Frame packing information could also be indicated through external means.

Bitstreams supported under this decoding capability are not required to be associated with frame packing information for all coded video sequences. It is also possible that such information, when present, may defer from one coded video sequence to another.

Editor’s Note: Impact on clause 6 (Video Operation Points) need to be specified.

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