|  |  |
| --- | --- |
| **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, 19–24 Oct. 2017 | Document: JCTVC-AC0028 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On the relation of frame packing arrangement and region-wise packing SEI messages** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Miska M: Hannuksela Nokia Technologies Tampere, Finland | Tel: Email: | +358 10 44 88 000 miska.hannuksela@nokia.com |
| *Source:* | Nokia | | |

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

# Abstract

It is proposed to include the frame packing arrangement scheme signalling within the region-wise packing SEI message. The frame packing arrangement and region-wise packing SEI messages are proposed to be mutually exclusive: when a frame packing arrangement SEI message applies to a picture, no region-wise SEI message shall apply to the same picture.

It is asserted that thanks to the proposed design the interpretation of the frame packing arrangement SEI message does not depend on the presence of the region-wise packing SEI message or vice versa.

It is noted that a similar proposal has been made for OMAF boxes in MPEG M41460.

# Problem Statement

It is asserted that the HEVC SoDAM3 has the following problems or shortcomings:

1. Frame packing arrangement SEI is misinterpreted by decoders that omit the decoding of the region-wise packing SEI message.

When region-wise packing SEI message is present in the bitstream, the arrangement indicated by the frame packing arrangement SEI message is intended to indicate the frame packing of the projected picture, i.e., the picture that results when the regions of the decoded picture have been "unpacked" according to the region-wise packing SEI message. However, the semantics of the frame packing arrangement SEI message, as specified in clause D.3.16 of HEVC, apply to the decoded picture. Decoders are not required to decode all SEI messages for conforming output. Decoders not parsing the region-wise packing SEI message would interpret the frame packing arrangement SEI message incorrectly, when both SEI messages are present in the bitstream.

1. Correct interpretation of the frame packing arrangement SEI message requires that the applicable region-wise packing SEI message, if any, is decoded first.

Both messages can persist through multiple access units, but their persistence needs not be synchronized. These messages are not required to be present within the same access unit. When present in the same access, their order is not constrained.

As a consequence, when a decoder interprets a frame packing arrangement SEI message, it has to check which region-wise packing SEI message, if any, is applicable. This checking or bookkeeping has to start from the beginning of each CLVS and has to also cover all NAL units of the coded picture that the frame packing arrangement SEI message is associated with.

1. No benefit in keeping the packing information in separate SEI messages

When omnidirectional video SEI messages are present, the frame packing arrangement SEI message shall be used in a constrained manner, i.e., only allowing indicating that projected pictures have side-by-side, top-bottom, and temporal interleaving for stereoscopic video. The frame packing arrangement information used with omnidirectional video could be indicated in a more compact manner within region-wise packing metadata than by separate frame packing arrangement SEI messages.

The frame packing arrangement SEI message also includes the following functionality, which in are not used or assumed to have a constant default value, when omnidirectional video SEI messages are present:

* Quincunx sampling indication
* Horizontal or vertical flipping
* Indication if pictures are coded fields, and mapping of coded fields of certain parity to a constituent frame 0 or 1
* Indication of constituent frames are self-contained from each other in terms of prediction dependencies
* The relative (sub-pixel-accurate) sampling grid position between constituent frames
* Indication whether the VUI sample aspect ratio information applies before or after upsampling inferred by the frame packing arrangement scheme

# Proposed specification text changes relative to JCTVC-AB1005-v1

## Region-wise packing SEI message syntax

*Carry out the changes indicated by the red font:*

|  |  |
| --- | --- |
| regionwise\_packing( payloadSize ) { | **Descriptor** |
| **rwp\_cancel\_flag** | u(1) |
| if( !rwp\_cancel\_flag ) { |  |
| **rwp\_persistence\_flag** | u(1) |
| **rwp\_stereo\_idc** | u(2) |
| if( rwp\_stereo\_idc > 0 ) |  |
| **pic0\_is\_right\_view\_flag** | u(1) |
| else |  |
| **rwp\_reserved\_zero\_bit** | u(1) |
| **rwp\_reserved\_zero\_3~~6~~bits** | u(3~~6~~) |
| if( rwp\_stereo\_idc < 3) |  |
| pic\_packing\_struct( ) |  |
| else { |  |
| pic\_packing\_struct( ) // for temporally interleaved picture 0 |  |
| pic\_packing\_struct( ) // for temporally interleaved picture 1 |  |
| } |  |
| } |  |
| } |  |

|  |  |
| --- | --- |
| pic\_packing\_struct( ) { | **Descriptor** |
| **num\_packed\_regions** | u(8) |
| **proj\_picture\_width** | u(16) |
| **proj\_picture\_height** | u(16) |
| for( i = 0; i < num\_packed\_regions; i++ ) { |  |
| **rwp\_reserved\_zero\_4bits**[ i ] | u(4) |
| **packing\_type**[ i ] | u(4) |
| if( packing\_type[ i ]  = =  0 ) { |  |
| **proj\_region\_width**[ i ] | u(16) |
| **proj\_region\_height**[ i ] | u(16) |
| **proj\_region\_top**[ i ] | u(16) |
| **proj\_region\_left**[ i ] | u(16) |
| **transform\_type**[ i ] | u(3) |
| **rwp\_reserved\_zero\_5bits**[ i ] | u(5) |
| **packed\_region\_width**[ i ] | u(16) |
| **packed\_region\_height**[ i ] | u(16) |
| **packed\_region\_top**[ i ] | u(16) |
| **packed\_region\_left**[ i ] | u(16) |
| } |  |
| } |  |
| } |  |

## Region-wise packing SEI message semantics

*Add the following:*

**rwp\_stereo\_idc** equal to 0 specifies that projected pictures are monoscopic. rwp\_stereo\_idc equal to 1 or 2 specify the frame packing arrangement for projected pictures identically to frame\_packing\_arrangement\_type equal to 3 or 4, respectively. rwp\_stereo\_idc equal to 3 specifies that all of the following apply:

* The decoded pictures in the scope of this SEI message are assigned in decoded order to be temporally interleaved picture 0 or 1 in an alternating manner.
* Each temporally interleaved picture 0 has the region-wise packing arrangement specified in the first pic\_packing\_struct( ) syntax structure included in the SEI message.
* Each temporally interleaved picture 1 has the region-wise packing arrangement specified in the second pic\_packing\_struct( ) syntax structure included in the SEI message.,
* Each pair of temporally interleaved pictures 0 and 1 in decoding order contain the content for a single projected picture having the top-bottom frame packing arrangement.

When rwp\_stereo\_idc is equal to 3, the SEI message shall persist in an even number of pictures in the same CLVS.

**pic0\_is\_right\_view\_flag** equal to 0 specifies that constituent frame 0 represents the left view, and constituent frame 1 represents the right view. pic0\_is\_right\_view\_flag equal to 1 specifies that constituent frame 0 represents the right view, and constituent frame 1 represents the left view.

## Region-wise packing SEI message semantics

*In D.3.41.5.1, carry out the changes indicated by the red font:*

– If neither a frame packing arrangement SEI message with frame\_packing\_arrangement\_cancel\_flag equal to 0 that applies to the picture nor a region-wise packing SEI message with rwp\_cancel\_flag equal to 0 that applies to the picture is ~~not~~ present, StereoFlag, TopBottomFlag, and SideBySideFlag are all set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 1.

– Otherwise, if a frame packing arrangement SEI message with frame\_packing\_arrangement\_cancel\_flag equal to 0 that applies to the picture is present, the following applies:

– StereoFlag is set equal to 1.

– If the value of frame\_packing\_arrangement\_type of the frame packing arrangement SEI message is equal to 3, TopBottomFlag is set equal to 0, SideBySideFlag is set equal to 1, HorDiv1 is set equal to 2 and VerDiv1 is set equal to 1.

– Otherwise, if the value of frame\_packing\_arrangement\_type of the frame packing arrangement SEI message is equal to 4, TopBottomFlag is set equal to 1, SideBySideFlag is set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 2.

– Otherwise, TopBottomFlag is set equal to 0, SideBySideFlag is set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 1.

– Otherwise, the following applies:

– If the value of rwp\_stereo\_idc of the region-wise packing SEI message is equal to 0, StereoFlag, TopBottomFlag, SideBySideFlag are all set equal to 0, and HorDiv1 and VerDiv1 are both set equal to 1.

– Otherwise, if the value of rwp\_stereo\_idc of the region-wise packing SEI message is equal to 1, StereoFlag is set equal to 1, TopBottomFlag is set equal to 0, SideBySideFlag is set equal to 1, HorDiv1 is set equal to 2 and VerDiv1 is set equal to 1.

– Otherwise, StereoFlag is set equal to 1, TopBottomFlag is set equal to 1, SideBySideFlag is set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 2.

# Patent rights declaration(s)

**Nokia Technologies may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**