|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  27th Meeting: Hobart, AU, 31 March – 7 April 2017 | Document: JCTVC-AA0022 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Comments on and proposed changes to the draft text for additional SEI messages** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | **Ye-Kui Wang** Qualcomm Incorporated 5775 Morehouse Drive San Diego, CA 92130, USA  **Gary J. Sullivan** Microsoft Corporation 1 Microsoft Way Redmond, WA 98052, USA | Tel: Email: | +1 858 651 8345 [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com)  +1 425 703 5308 [garysull@microsoft.com](mailto:garysull@microsoft.com) |
| *Source:* | Qualcomm Incorporated, Microsoft Corporation | | |

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

# Abstract

This document provides some comments on the draft text in JCTVC-Z1005 ("HEVC Additional Supplemental Enhancement Information (Draft 1)"), and proposes some text changes to address some (not all) of the comments.

Some of the comments are on general topics that apply to all or some of the new SEI messages, and other comments are specifically on the omnidirectional projection indication SEI message and the content colour volume SEI message. Some remarks about related aspects of the basis text are also included.

The text changes are in the attachment, with changes marked relative to JCTVC-Z1005. Some asserted obvious minor editorial changes, e.g., missing or unnecessary spaces, are also included.

Note that although some of the comments originated from the second author, the proposed solutions were primarily drafted by the first author and have not necessarily all been reviewed by the second author due to lacking of time.

# List of comments and suggestions

## General comments and suggestions

1. **Comment:** The scalability issues and CLVS concept seem not adequately addressed in the specification of the new SEI messages – especially for the regional nesting and MCTS nesting SEI messages.

**Suggestion:** In Annex F and onwards, add the new SEI message types into the lists VclAssociatedSeiList and PicUnitRepConSeiList. Make sure the CLVS concept is covered in the semantics of each SEI message.

1. **Comment:** Clause 7.4.2.4.4 contains a NOTE that discusses the (scalability) nesting of some SEI messages. This seems to be the only place in the main body text (prior to clause 10 which specifies the sub-bitstream extraction process) that discusses this concept. It seems like it would be desirable to avoid discussing this concept in that location.

**Suggestion:** The first sentence in the NOTE should be kept in 7.4.2.4.4. The last sentence, i.e., "Consequently, hypothetical reference decoder (HRD) parameters carried in non-nested buffering period, picture timing and decoding unit information SEI messages apply to access units based on such access unit boundary detection." may be moved to a NOTE somewhere in Annex C, with appropriate rephrasing as needed.

1. **Comment:** The phrases of "nested SEI message" and "not nested" are used in many places in the published HEVC specification as well as in JCTVC-Z1005. However, up to now, there are four different nesting SEI messages defined, and the use of these phrases is confusing, since the reader needs to know which type of nesting is being discussed in each context. For example, 3.85 defines "nested SEI message" as "An SEI message that is contained in a scalable nesting SEI message" (and 3.88 defines a "non-nested SEI message" in the similar manner). The current draft specification of the new draft SEI message seems to violate this definition, because it is discussing a nesting of SEI messages that is not related to the scalable nesting SEI message.

**Suggestion:** Throughout the HEVC specification, replace all single-word instances of "nested" with "scalably-nested" (e.g., there is no single-word instance of "nested" in "non-nested"), and all instances of "non-nested" with "non-scalably-nested", except in 3.158 (which can be kept unchanged) and Table F.4 in F.14.3.1 (which is suggested to be changed to not use the phrase of "nested SEI message" but rather the phrase of "SEI message contained in the bitstream partition nesting SEI message").

In the context of the regional nesting SEI message, define and consistently use the phrases "regionally-nested" and "non-regionally-nested", and in the context of the MCTS nesting SEI message, define and consistently use the phrases "MCTS-nested" and "non-MCTS-nested".

1. **Comment:** For some of the SEI messages, the stated persistence scope in Table D.1 is "The CLVS containing the SEI message". Though Table D.1 is informative, if the design intent is indeed as stated in Table D.1 for these SEI messages, there needs to be a constraint to require the content of all SEI messages of the type in a CLVS to be identical. Such examples include the temporal MCTS SEI message and the omnidirectional projection information SEI message.

**Suggestion:** Systematically check all SEI messages for which the persistence scope stated in Table D.1 is "The CLVS containing the SEI message" on whether there is a constraint to require the content of all SEI messages of the type in a CLVS to be identical, and if not, discuss what is the actual intent, and either add this constraint or change the persistence scope stated in Table D.1.

1. **Comment:** There are some different specified behaviors in treating the SEI related IDs in the current HEVC specification that are specified in a similar manner as the following:

"Values of \*\_id from 0 to 255 and from 512 to 231 − 1 may be used as determined by the application. Values of \*\_id from 256 to 511, inclusive, and from 231 to 232 − 2, inclusive, are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore ... containing a value of \*\_id in the range of 256 to 511, inclusive, or in the range of 231 to 232 − 2, inclusive, and bitstreams shall not contain such values."

A search of "256 to 511" would find the following list:

* pan\_scan\_rect\_id
* scene\_id
* second\_scene\_id
* snapshot\_id
* progressive\_refinement\_id
* tone\_map\_id
* frame\_packing\_arrangement\_id
* mcts\_id[ i ]
* knee\_function\_id
* colour\_remap\_id
* ilcts\_id[ i ]
* regional\_nesting\_id

The following ones are specified as "that may be used to identify the purpose (or usage) of ...":

* pan\_scan\_rect\_id
* tone\_map\_id
* frame\_packing\_arrangement\_id
* mcts\_id[ i ]
* knee\_function\_id
* colour\_remap\_id, ilcts\_id[ i ]
* regional\_nesting\_id

The following ones are really used as an identifier of an object (e.g., scene, region) or an operation (e.g., progressive refinement operation):

* scene\_id
* second\_scene\_id
* snapshot\_id
* progressive\_refinement\_id

Some places say that decoders encountering a reserved value of such an ID shall ignore the ID itself (pan\_scan\_rect\_id, scene\_id, second\_scene\_id, snapshot\_id, progressive\_refinement\_id, ilcts\_id[ i ], and regional\_nesting\_id), while some other places say to ignore the entire SEI message (tone\_map\_id, frame\_packing\_arrangement\_id, mcts\_id[ i ], knee\_function\_id, and colour\_remap\_id).

Note that the two categories of semantics of identifying the purpose/usage or identifying an object/operation and the two different ignoring behaviors are fully mixed: all four possible combinations exist.

The values are either reserved or to be determined by the application. Reserved values are not allowed to be used, and under many circumstances, it may be ill-advised to use application-determined values for some of the SEI messages unless some particular usage is specified by some application standard.

A minor, editorial issue is that ", inclusive" is missing in some expressions of the value range in such places.

In the semantics of ilcts\_id[ i ], the specified decoder behaviour is to ignore the ID itself when it has a reserved value, but with an additional instruction to remove the syntax element from the bitstream and discard it. This is inconsistent with the semantics provided for similar uses.

**Suggestion:** Discuss the questions and try to find answers to them and establish greater consistency in the text.

For all these IDs (pan\_scan\_rect\_id, scene\_id, second\_scene\_id, snapshot\_id, progressive\_refinement\_id, tone\_map\_id, frame\_packing\_arrangement\_id, mcts\_id[ i ], knee\_function\_id, colour\_remap\_id, and ilcts\_id[ i ]), it may be best to just specify that when it has a reserved value, the decoder is only required to ignore the value of ID within the reserved range, rather than ignoring the entire SEI message.

Systematically check all such places and add the missing ", inclusive" in the expressions of value ranges.

Additionally, provide an explicit warning (similar to what was recently done for NAL unit type codes in the range of UNSPEC48..UNSPEC63) about potential "collisions" of interpretation for the application-specific ID values.

1. **Comment:** In Table D.1, the stated persistence scopes of some SEI messages are not accurate, including the content colour volume SEI message, the motion-constrained tile set extraction information set SEI message, and the temporal MCTS SEI message.

**Suggestion:** Change the persistence scopes stated in Table D.1 to be accurate, for the content colour volume SEI message, the motion-constrained tile set extraction information set SEI message, and the temporal MCTS SEI message.

1. **Comment:** In some places, including three instances in D.3.3, two instances in D.3.33, and a few instances in JCTVC-Z1005 (for the inference rules of ccv\_primaries\_x[ c ], ccv\_primaries\_y[ c ], ccv\_min\_luminance\_value, ccv\_max\_luminance\_value, and ccv\_avg\_luminance\_value of the content volume SEI message, and for the five yaw/pitch/roll syntax elements of the omnidirectional projection indication SEI message), the specifications of inference of values of absent syntax elements use the wording "should be inferred" or "can be inferred". However, in other (hundreds of) places for inference of values in the HEVC specification, the wording "is/are inferred" is used. Note that there is no wording of "shall be inferred" anywhere in the published HEVC specification or JCTVC-Z1005.

**Suggestion:** Consistently use the wording "is/are inferred" for specifying the inference of values of absent syntax elements, including the specifically identified places (with some possible exception for ccv\_avg\_luminance\_value, for which two possible inference rules are discussed).

## Comments and suggestions for the content colour volume SEI message

1. **Comment:** In the following text:

"Otherwise (the value of transferCharacteristics is not equal to 1, 6, 7, 14, 15, or 18), the exact inverse of the transfer function specified in Table E.4 should be used to convert the non-linear signal to a linear representation."

The wording "the exact inverse of the transfer function specified in Table E.4" is unclear for some possible values of transferCharacteristics. Some of the cases do not even have a transfer function specified in Table E.4.

**Suggestion:** Clarify the intent or prohibit using the SEI message in such cases.

1. **Comment:** The texts of the three NOTEs are largely overlapping and redundant, merging of them into one would make sense.

**Suggestion:** Merge NOTEs 1 to 3 into one NOTE.

## Comments and suggestions for the omnidirectional projection SEI message

1. **Comment:** The omnidirectional projection SEI message is said to persist for the current layer in decoding order from the current picture until the end of the CLVS. The stated persistence scope in Table D.1 is "The CLVS containing the SEI message". These two places are not completely aligned.

**Suggestion:** Do either of the following:

* Add a constraint to require that the applicable omnidirectional projection shall be the same for all pictures within a CLVS.
* Since it does not make sense to have different projections across layers, it should make sense to just say that the SEI message applies to all pictures in the CVS (instead of CLVS), both in this clause and in Table D.1, and require that the nuh\_layer\_id in the SEI NAL unit containing this SEI message to be 0.

1. **Comment:** Figures D.12 and D.14 are not referenced in the text.

**Suggestion:** Add a reference to each of the two figures at an appropriate place in the text. For example, a NOTE could be added to refer to each of the two figures.

1. **Comment:** In multiple places, the omnidirectional projection indication SEI message is mistakenly written as the spherical orientation SEI message.

**Suggestion:** Change all instances of "spherical orientation" to "omnidirectional projection indication".

1. **Comment:** The specified values for geometry\_type and projection\_type start from 1 instead of 0. However, it seems that there is no reason to reserve different sets of values.

**Suggestion:** Specify geometry\_type and projection\_type starting from value 0 instead, and reserve values 1 and above (or remove the problem in another manner as discussed in the next comment about these syntax elements).

1. **Comment:** The syntax elements geometry\_type and projection\_type are basically unused and seem potentially unnecessary.

geometry\_type seems not really useful as it seems not expected to have a different geometry than spherical surface coming soon.

It might be simpler just to remove those and rename the SEI message to say it is an equirectangular projection indication SEI message. Also, the specification of constraints on their values is missing, and the specification of what a decoder should do if it encounters reserved values is missing.

**Suggestion:** Remove geometry\_type.

Discuss and decide whether to remove projection\_type, taking into account that this depends on whether it would be preferable to use different SEI messages for different projection types (e.g., if both ERP and CMP are allowed).

For geometry\_type (if to be kept) and projection\_type, specify value constraints and reserved values, and specify that decoders shall ignore omnidirectional projection indication SEI messages with reserved values of either of the two syntax elements.

1. **Comment:** The equations might be simplified, made less redundant, or otherwise editorially improved. Note that similar or related equations and expressions can be found in 7.4.3.2.1, D.3.4, D.3.29, D.3.31, and E.3.1.

**Suggestion:** Simplify the equations if feasible, both in the added clauses and in the prior basis text..

1. **Comment:** There is inconsistent usage of underscores, hyphens, and minus signs in D.3.41.

"geometry type" is missing an underscore between the words at the beginning of its semantics.

Hyphens are used instead of minus signs in a few places in the semantics.

"sub geometry" should be hyphenated or concatenated.

**Suggestion:** Correct the usage of underscores, hyphens, and minus signs.

1. **Comment:** In the following:

**"sub\_geometry\_flag** indicates the presence of sub geometry elements that further describe or limit the geometry of this SEI message."

The wording "that further describe or limit the geometry of this SEI message" is unclear.

**Suggestion:** Change the semantics of sub\_geometry\_flag to be the following:

**"sub\_geometry\_flag** indicates the presence of sub-geometry elements that describe the area on the geometry (e.g., the spherical surface when geometry\_type is equal to 1) covered by the cropped output picture."

1. **Comment:** The term "projection image" is used in the semantics of five yaw/pitch/roll syntax elements, to refer to the area on the spherical surface covered by the cropped output picture. However, firstly the term is not defined. Secondly, calling a region on the sphere a "projection image" does not seem straightforward or intuitive.

**Suggestion:** Change all instances of "the projection image" to "the area on the spherical surface covered by the cropped output picture", or define a term to refer to this and use the defined term, e.g., "represented spherical region".

1. **Comment:** The provided coordinate conversion equations and figures may not reflect the convention that has typically been established for the geometry of ERP projection, the order of operations for the conversion process, or the way the conversion has been specified in other closely-related documents.

**Suggestion:** Check the equations, and adjust as appropriate.

1. **Comment:** The video using the omnidirectional projection indication SEI message should not be interpreted as ordinary 2D rectangular video. However, there is no way to indicate to existing decoders that the video is unusual.

**Suggestion:** For the semantics of the general\_non\_packed\_constraint\_flag, treat the new SEI message in a similar manner as the segmented rectangular frame packing arrangement SEI message.

1. **Comment:** The value range of roll\_center is specified to be −18000 and 18000, inclusive. However, it should be the same as that of yaw\_center, i.e., −18000 to 17999, inclusive.

**Suggestion:** Define the range of roll\_center in a consistent manner with that for yaw\_center.

1. **Comment:** Note that each of the angular coordinates (φ, θ) may go out of the ranges of [−18000, 17999] and [−9000, 9000], respectively, per the equations.

**Suggestion:** Discuss and decide whether the values of the angular coordinates (φ, θ) should be modulo to be within their respective ranges, and if so, change the equations accordingly.

1. **Comment:** These two variables CenterLeftOffsetC and CenterTopOffsetC are only used in this subclause and therefore should be defined as local variables.

**Suggestion:** Replace CenterLeftOffsetC and CenterTopOffsetC with centerLeftOffsetC and centerTopOffsetC, respectively.

1. **Comment:** There is a constraint that disallows the presence of frame packing arrangement SEI messages and segmented rectangular frame packing arrangement SEI messages in a CLVS when there is an omnidirectional projection indication SEI message. This constraint disallows the signalling of projected stereoscopic 360 video using frame packing.

**Suggestion:** Remove the constraint, and specify that the projection mapping process independently applies to the samples of each constituent frame that are within the cropped output picture.

Ensure that the provided equations are adequate for this usage (either by using “generalized” equations or by providing specific equations for each case).

1. **Comment:** In the following text:

"When geometry\_type is equal to 1, the projection indicator is given in spherical coordinates, where φ is the azimuth (longitude) and θ is the elevation (latitude) as depicted in Figure D.13."

Shouldn't roll be mentioned here as well now that it is also depicted in figure D.13?

**Suggestion:** Add the mentioning of roll into the sentence.

# Patent rights declaration(s)

**Qualcomm Incorporated and Microsoft Corporation do not have any current or pending patent rights relating to the technology described in this contribution (to the extent of the personal awareness of the authors).**