|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  22nd Meeting: Geneva, CH, 15–21 Oct. 2015 | Document: JCTVC-V0038 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Effective Colour Volume SEI** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Alexis Michael Tourapis,  Yeping Su, David Singer  Apple, Inc. 1 Infinite Loop Cupertino, CA 95014 USA | Tel: Email: | +1-408-228-7983 atourapis@apple.com |
| *Source:* | Apple, Inc. | | |

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

# Abstract

This contribution requests a new MPEG-4 AVC and HEVC “Effective Colour Volume” SEI message, that would indicate the effective colour volume occupied by the current layer of a coded video stream (CVS). This SEI message can be seen as an extension of the “Content Light Level Information” (CLL) SEI that is already supported by these two standards, and can provide additional information to a decoder about the characteristics and limitations of the video signal. This information can then be used to appropriately process the content for display or other types of processing, e.g. to assist in processes such as colour conversion, clipping, colour gamut tailoring, and tone mapping, among others.

# Introduction

The new version [2] of MPEG HEVC [1] includes a new SEI message [3] that provides information about the maximum video intensity contained in the video stream. This includes two variables, the MaxFALL variable that describes the Maximum Frame Average Light Level, and the MaxCLL parameter that describes the Maximum Content Light Level in a scene. These parameters are fixed per sequence and could be used by video decoders to assist with the display process of the content.

Although this SEI message can be quite useful especially for High Dynamic Range (HDR) and Wide Colout Gamut (WCG) applications, it could be argued that it does not provide a comprehensive description about the characteristics of the content. In particular, apart from the content’s effective dynamic range (or maximum light level), it is quite common to see content that may be represented in a wide colour gamut representation/container, e.g. BT.2020 [4], but in reality only cover a much narrower colour volume. For example, all the material used for the recent MPEG HDR/WCG CfE activity have used the BT.2020 colour volume container and the ST 2084 transfer function [6], however effectively only covered BT.709 [5] or P3D65 [7] colour primaries. The maximum value for all samples was also restricted to 4000 cd/m2. Nevertheless, even though it could be argued that the signal representation is suboptimal for these scenarios, there are certain advantages on using such fixed, wider representations, such as achieving interoperability while future-proofing applications relating to HDR/WCG. However, knowledge of the actual limitations of the original signal, without also having to perform any direct analysis on the video data, can still be quite useful if available. For example, it was demonstrated in [8] that knowledge of the relationship between the luma and chroma parameters (i.e. matrix coefficient parameters) in a system can considerably improve colour clipping and avoid the introduction of artifacts at the boundaries of the colour container. However, it could be argued that knowledge of the actual occupied/effective colour volume can further improve the clipping process, since now colour values can be better constrained within their originally occupied space. In another example, a different system may desire to perform additional processing, e.g. denoising, frame rate conversion, or rescaling, before display. If these operations are of limited in precision, knowledge of the effective color volume could allow this system to appropriately convert the data in a color space that would maximize its performance, e.g. by considering only the effective color volume. For example, if we know that data represented in a BT.2020 container and with the ST 2084 transfer function only cover BT.709 primaries and have a maximum brightness of 1000 cd/m2, it may be more appropriate to convert the data into a BT.709 container and use a transfer function that covers the entire brightness range. Similar knowledge could better assist a display of particular capabilities to perform better or reduced complexity tone mapping or colour gamut tailoring among others.

We propose specifying a new SEI message, named the “Effective Colour Volume” SEI message, that can provide a more refined description of the colour characteristics and volume occupied by the content. This SEI is capable of regional signaling of parameters as well as signalling of not only 3-primary color coordinates and value limits but also 5-primary and 6-primary parameters. The inclusion of such models in this SEI is inspired by new display paradigms that employ these models, and can provide a more flexible and accurate description model for the effective color volume of a signal.

# Proposed Text

1. **General SEI message syntax**

|  |  |
| --- | --- |
| sei\_payload( payloadType, payloadSize ) { | **Descriptor** |
| if( nal\_unit\_type = = PREFIX\_SEI\_NUT ) |  |
| if( payloadType = = 0 ) |  |
| buffering\_period( payloadSize ) |  |
| else if( payloadType = = 1 ) |  |
| pic\_timing( payloadSize ) |  |
| else if( payloadType = = 2 ) |  |
| …… |  |
| else if( payloadType = = 183 ) |  |
| ambient\_viewing\_environment( payloadSize ) |  |
| else if( payloadType = = 184 ) |  |
| effective\_colour\_volume( payloadSize ) |  |
| else |  |
| reserved\_sei\_message( payloadSize ) |  |
| else /\* nal\_unit\_type = = SUFFIX\_SEI\_NUT \*/ |  |
| if( payloadType = = 3 ) |  |
| filler\_payload( payloadSize ) |  |
| else if( payloadType = = 4 ) |  |
| user\_data\_registered\_itu\_t\_t35( payloadSize ) |  |
| else if( payloadType = = 5 ) |  |
| user\_data\_unregistered( payloadSize ) |  |
| else if( payloadType = = 17 ) |  |
| progressive\_refinement\_segment\_end( payloadSize ) |  |
| else if( payloadType = = 22 ) |  |
| post\_filter\_hint( payloadSize ) |  |
| else if( payloadType = = 132 ) |  |
| decoded\_picture\_hash( payloadSize ) |  |
| else |  |
| reserved\_sei\_message( payloadSize ) |  |
| if( more\_data\_in\_payload( ) ) { |  |
| if( payload\_extension\_present( ) ) |  |
| **reserved\_payload\_extension\_data** | u(v) |
| **payload\_bit\_equal\_to\_one** /\* equal to 1 \*/ | f(1) |
| while( !byte\_aligned( ) ) |  |
| **payload\_bit\_equal\_to\_zero** /\* equal to 0 \*/ | f(1) |
| } |  |
| } |  |

1. Effective colour volume SEI message syntax

|  |  |
| --- | --- |
| effective\_colour\_volume( payloadSize ) { | Descriptor |
| **ecv\_cancel\_flag** | u(1) |
| if( !ecv\_cancel\_flag ) { |  |
| **ecv\_persistence\_flag** | u(1) |
| **ecv\_ctb\_partitions\_y\_minus1** | ue(v) |
| **ecv\_ctb\_partitions\_x\_minus1** | ue(v) |
| for( j = 0; j < ecv\_ctb\_partitions\_y\_minus1 + 1; j++ ) { |  |
| for( i = 0; i < ecv\_ctb\_partitions\_x\_minus1 + 1; i++ ) { |  |
| **ecv\_three\_primary\_representation\_flag**[ j ][ i ] | u(1) |
| **ecv\_five\_primary\_representation\_flag**[ j ][ i ] | u(1) |
| **ecv\_six\_primary\_representation\_flag**[ j ][ i ] | u(1) |
| **ecv\_native\_primary\_limits\_flag**[ j ][ i ] | u(1) |
| **ecv\_xyz\_primary\_limits**[ j ][ i ] | u(1) |
| if(ecv\_three\_primary\_representation\_flag == 1) { |  |
| for( c = 0; c < 3; c++ ) { |  |
| **ecv\_three\_primary\_x**[ j ][ i ] [ c ] | u(16) |
| **ecv\_three\_primary\_y**[ j ][ i ] [ c ] | u(16) |
| **ecv\_three\_primary\_min**[ j ][ i ] **[** c ] | u(32) |
| **ecv\_three\_primary\_max**[ j ][ i ] [ c ] | u(32) |
| } |  |
| } |  |
| if(ecv\_five\_primary\_representation\_flag == 1) { |  |
| for( c = 0; c < 5; c++ ) { |  |
| **ecv\_five\_primary\_x**[ j ][ i ] [ c ] | u(16) |
| **ecv\_five\_primary\_y**[ j ][ i ] [ c ] | u(16) |
| **ecv\_five\_primary\_min**[ j ][ i ] **[** c ] | u(32) |
| **ecv\_five\_primary\_max**[ j ][ i ] [ c ] | u(32) |
| } |  |
| } |  |
| if(ecv\_six\_primary\_representation\_flag == 1) { |  |
| for( c = 0; c < 6; c++ ) { |  |
| **ecv\_six\_primary\_x**[ j ][ i ] [ c ] | u(16) |
| **ecv\_six\_primary\_y**[ j ][ i ] [ c ] | u(16) |
| **ecv\_six\_primary\_min**[ j ][ i ] **[** c ] | u(32) |
| **ecv\_six\_primary\_max**[ j ][ i ] [ c ] | u(32) |
| } |  |
| } |  |
| if(ecv\_native\_primary\_limits\_flag == 1) { |  |
| for( c = 0; c < 3; c++ ) { |  |
| **ecv\_native\_primary\_min**[ j ][ i ] **[** c ] | i(32) |
| **ecv\_native\_primary\_max**[ j ][ i ] [ c ] | i(32) |
| } |  |
| } |  |
| if(ecv\_xyz\_primary\_limits\_flag == 1) { |  |
| for( c = 0; c < 3; c++ ) { |  |
| **ecv\_xyz\_primary\_min**[ j ][ i ] **[** c ] | u(32) |
| **ecv\_xyz\_primary\_max**[ j ][ i ] [ c ] | u(32) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

1. Effective Colour Volume SEI message semantics

The effective colour volume SEI message provides information about the colour volume characteristics of the associated picture.

**ecv\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous effective colour volume SEI message in output order that is associated with one or more primary picture layers to which this SEI applies. ecv\_cancel\_flagequal to 0 indicates that effective colour volume information follows.

**ecv\_persistence\_flag** specifies the persistence of the effective colour volume SEI message.

ecv\_persistence\_flag equal to 0 specifies that the effective colour volume applies to the current decoded picture only.

Let picA be the current picture. ecv\_persistence\_flag equal to 1 specifies that the effective colour volume SEI message persists for the current layer in output order until any of the following conditions are true:

– A new CLVS of the current layer begins.

– The bitstream ends.

– A picture picB in the current layer in an access unit containing an effective colour volume SEI message that is applicable to the current layer is output for which PicOrderCnt( picB ) is greater than PicOrderCnt( picA ), where PicOrderCnt( picB ) and PicOrderCnt( picA ) are the PicOrderCntVal values of picB and picA, respectively, immediately after the invocation of the decoding process for the picture order count of picB.

**ecv\_ctb\_partitions\_y\_minus1** specifies the number of vertical partitions for which colour volume information can be signalled. The height of each partition ecvPartition(j,i) at index (j,i) can be computed as follows:

numYParts = ecv\_ctb\_partitions\_y\_minus1 + 1

ecvPartHeight(j,i) **=** round((j + 1) \* PicHeightInCtbsY / numYParts)

– round(j \* PicHeightInCtbsY / numYParts)

**ecv\_ctb\_partitions\_x\_minus1** specifies the number of vertical partitions for which colour volume information can be signalled. The width of each partition ecvPartition (j,i) at index (j,i) can be computed as follows:

numXParts = ecv\_ctb\_partitions\_y\_minus1 + 1

ecvPartWidth(j,i) **=** round((j + 1) \* PicWidthInCtbsY / numXParts)

* round(j \* PicWidthInCtbsY / numXParts)

**ecv\_three\_primary\_representation\_flag**[ j ][ i ]equal to 1 indicates that a three primary effective colour volume representation for the current partition ecvPartition(j,i) is present. ecv\_three\_primary\_representation\_flag[ j ][ i ] equal to 0 indicates that this information is not present.

**ecv\_five\_primary\_representation\_flag**[ j ][ i ]equal to 1 indicates that a five primary effective colour volume representation for the current partition ecvPartition(j,i) is present. ecv\_five\_primary\_representation\_flag[ j ][ i ] equal to 0 indicates that this information is not present.

**ecv\_six\_primary\_representation\_flag**[ j ][ i ]equal to 1 indicates that a six primary effective colour volume representation for the current partition ecvPartition(j,i) is present. ecv\_five\_primary\_representation\_flag[ j ][ i ] equal to 0 indicates that this information is not present.

If ecv\_three\_primary\_representation\_flag[ j ][ i ], ecv\_three\_primary\_representation\_flag[ j ][ i ], and ecv\_three\_primary\_representation\_flag[ j ][ i ] are all equal to 0, then no explicit refinement of the colour volume of the current partition is provided.

**ecv\_native\_primary\_limits\_flag**[ j ][ i ] equal to 1 indicates that the native/current primary brightness limits are present. ecv\_native\_primary\_limits\_flag[ j ][ i ] equal to 0 indicates that this information is not present.

**ecv\_xyz\_primary\_limits**[ j ][ i ] equal to 1 indicates that the CIE-1931 XYZ primary brightness limits are present. ecv\_xyz\_primary\_limits\_flag[ j ][ i ] equal to 0 indicates that this information is not present.

**ecv\_three\_primary\_x**[ j ][ i ][ c ] and **ecv\_three\_primary\_y**[ j ][ i ][ c ] specify the normalized x and y chromaticity coordinates, respectively, of the colour primary component c for the three primary colour volume representation for partition ecvPartition(j,i) in increments of 0.00002, according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15). Index value c equal to 0 should correspond to the green primary, c equal to 1 should correspond to the blue primary, and c equal to 2 should correspond to the red colour primary. The values of ecv\_three\_primary\_x[ j ][ i ][ c ] and ecv\_three\_primary\_y[ j ][ i ][ c ] shall be in the range of 0 to 50 000, inclusive. If these syntax elements are not present in this SEI message, then no three primary colour coordinates are explicitly defined.

**ecv\_three\_primary\_min**[ j ][ i ][ c ] and **ecv\_three\_primary\_max**[ j ][ i ][ c ] specify the nominal maximum and minimum values, respectively, of the signals in partition ecvPartition(j,i) after conversion into the three colour primary colour space, in units of 0.0001 candelas per square metre. When present, ecv\_three\_primary\_min[ j ][ i ][ c ] shall be less than ecv\_three\_primary\_max[ j ][ i ][ c ]. If not present, no explicit limit for these primaries is specified.

**ecv\_five\_primary\_x**[ j ][ i ][ c ] and **ecv\_five\_primary\_y**[ j ][ i ][ c ] specify the normalized x and y chromaticity coordinates, respectively, of the colour primary component c for the five primary colour volume representation for partition ecvPartition(j,i) in increments of 0.00002, according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15). Index value c equal to 0 should correspond to the green primary, c equal to 1 should correspond to the cyan primary, c equal to 2 should correspond to the blue colour primary, c equal to 3 should correspond to the red primary, and c equal to 4 should correspond to the yellow primary. The values of ecv\_five\_primary\_x[ j ][ i ][ c ] and ecv\_five\_primary\_y[ j ][ i ][ c ] shall be in the range of 0 to 50 000, inclusive. If these syntax elements are not present in this SEI message, then no five primary colour coordinates are explicitly defined.

**ecv\_five\_primary\_min**[ j ][ i ][ c ] and **ecv\_five\_primary\_max**[ j ][ i ][ c ] specify the nominal maximum and minimum values, respectively, of the signals in partition ecvPartition(j,i) after conversion into the five colour primary colour space, in units of 0.0001 candelas per square metre. When present, ecv\_five\_primary\_min[ j ][ i ][ c ] shall be less than ecv\_five\_primary\_max[ j ][ i ][ c ]. If not present, no explicit limit for these primaries is specified.

**ecv\_six\_primary\_x**[ j ][ i ][ c ] and **ecv\_six\_primary\_y**[ j ][ i ][ c ] specify the normalized x and y chromaticity coordinates, respectively, of the colour primary component c for the six primary colour volume representation for partition ecvPartition(j,i) in increments of 0.00002, according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15). Index value c equal to 0 should correspond to the green primary, c equal to 1 should correspond to the cyan primary, c equal to 2 should correspond to the blue colour primary, c equal to 3 should correspond to the magenda primary, c equal to 4 should correspond to the red primary, and c equal to 5 should correspond to the yellow primary. The values of ecv\_six\_primary\_x[ j ][ i ][ c ] and ecv\_six\_primary\_y[ j ][ i ][ c ] shall be in the range of 0 to 50 000, inclusive. If these syntax elements are not present in this SEI message, then no six primary colour coordinates are explicitly defined.

**ecv\_six\_primary\_min**[ j ][ i ][ c ] and **ecv\_six\_primary\_max**[ j ][ i ][ c ] specify the nominal maximum and minimum values, respectively, of the signals in partition ecvPartition(j,i) after conversion into the six colour primary colour space, in units of 0.0001 candelas per square metre. When present, ecv\_five\_primary\_min[ j ][ i ][ c ] shall be less than ecv\_five\_primary\_max[ j ][ i ][ c ]. If not present, no explicit limit for these primaries is specified.

**ecv\_native\_primary\_min**[ j ][ i ][ c ] and **ecv\_native\_primary\_max**[ j ][ i ][ c ] specify the nominal maximum and minimum values, respectively, of the signals in partition ecvPartition(j,i) in their native colour primaries in units of 0.0001 candelas per square metre. When present, ecv\_native\_primary\_min[ j ][ i ][ c ] shall be less than ecv\_native\_primary\_max[ j ][ i ][ c ]. If not present, no explicit limit for these primaries is specified.

**ecv\_xyz\_primary\_min**[ j ][ i ][ c ] and **ecv\_xyz\_primary\_max**[ j ][ i ][ c ] specify the nominal maximum and minimum values, respectively, of the signals in partition ecvPartition(j,i) after conversion into the xyz colour space, in units of 0.0001 candelas per square metre. Index value c equal to 0 should correspond to the X primary, c equal to 1 should correspond to the Y primary, and c equal to 2 should correspond to the Z colour primary. When present, ecv\_xyz\_primary\_min[ j ][ i ][ c ] shall be less than ecv\_xyz\_primary\_max[ j ][ i ][ c ]. If not present, no explicit limit for these primaries is specified.

# Conclusion

This contribution proposes a new SEI message, named the “Effective Colour Volume” SEI message, that would indicate the effective colour volume occupied by the current layer of a coded video stream (CVS). This information can be used to appropriately process the content for display or other types of processing, e.g. to assist in processes such as colour conversion, clipping, colour gamut tailoring, and tone mapping, among others.

# References

1. ISO/IEC 23008:2014 Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding, Second Edition
2. R. Joshi, S. Liu, G. Sullivan, G. Tech, J. Xu, and Y. Ye, “High Efficiency Video Coding (HEVC) Screen Content Coding: Draft 4”, MPEG/JCTVC document M36870/JCTVC-U1005, Warsaw, PL, Jun. 2015
3. C. Fogg, J. Helman, M. Smith, and M. Zink, “Content light level information SEI”, MPEG/JCTVC document M35567/JCTVC-T0101, Switzerland, Feb. 2015
4. Recommendation ITU-R BT.2020, “Parameter values for ultra-high definition television systems for production and international programme exchange” (2012).
5. Recommendation ITU-R BT.709-5, “Parameter values for the HDTV standards for production and international programme exchange”
6. SMPTE ST 2084:2014, “High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays”, 2014
7. Society of Motion Picture and Television Engineers EG 432-1 (2010), Digital Source Processing – Color Processing for D-cinema
8. A.M. Tourapis, Y. Su, D. Singer, J. Sole, D. Rusanovskyy, M. Karczewicz, C. Fogg, A. Duenas, F. Bossen, “Clipping during 4:2:0 to 4:4:4 conversions”, MPEG document M36272, Warsaw, PL, Jun. 2015

# Patent rights declaration(s)

**Apple Inc 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).**