|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  33rd Meeting: Macao, CN, 6–12 October 2018 | Document: JCTVC-AG0025-v1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Motion constrained slice set sub-bitstream extraction for AVC** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Robert Skupin, Yago Sanchez,  Thomas Schierl, Thomas Wiegand | email: | robert.skupin@hhi.fraunhofer.de |
| *Source:* | Fraunhofer HHI | | |

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

# Abstract

HEVC has recently been amended with an MCTS sub-bitstream extraction process which is asserted to be useful in 360-degree video applications. This document proposes to add similar functionality to AVC for integration into the current amendment.

The proposed syntax allows extraction of conformant AVC sub-bitstreams of motion constraint slice sets spanning the picture width.

# Introduction

HEVC has recently been amended with a set of SEI messages related to HDR and 360-degree video and a similar effort is currently underway for AVC [1]. In HEVC, the MCTS extraction related SEI messages (MCTS EIS and MCTS EISN SEI messages) allow sub-bitstream extraction of rectangular regions within a video, specifically HEVC tiles.

Apart from applications such as region-of-interest decoding in broadcast applications [2], these messages have proven useful in content preparation for services using the HEVC based viewport-dependent profile (VDP) of OMAF [3]. For instance, in OMAFs HEVC VDP, a client can select among different combinations of the multi-resolution or multi-quality tiles wherein each combination represents another preferred viewing direction. In the respective content preparation, videos are encoded in multiple resolutions or qualities using motion constrained tiles which are then offered to the client as separate HEVC streams. In order to generate the separate HEVC streams, VR-IF guidelines [4] recommend usage of the MCTS sub-bitstream extraction process based on the MCTS EIS SEI messages. This procedure can be used in an open source OMAF implementations such as [5] or [6].

The following Figure 1 illustrates the workflow from content preparation to client-side tile selection with red dashes indicating the step from encoded video to individual stream containing MCTS.



Figure 1: Workflow of OMAF HEVC based viewport dependent profile [3]. Extraction marked with red dashes.

The OMAF specification also defines an AVC based viewport dependent profile for support of legacy devices. The following Figure 2 illustrates the content preparation workflow comparable to the above for an OMAF AVC VDP service. The video is adjusted in the raw domain before encoding so that each envisioned “tile” (numbers 1-8 in Figure 2) can be covered by a consecutive set of macro block in raster scan order, e.g. by stacking the “tiles” vertically. Encoding of the video entails partitioning of the video into the envisioned number of slices, wherein constraints in the motion compensated prediction are enforced at the “tile” boundaries in order to code the “tiles” independently. Once encoded, the coded video is split into separate AVC streams as indicated with red dashed in Figure 2.

However, a defined extraction procedure for slice based sub-streams is lacking in AVC. Therefore, this document proposes to add functionality similar to the MCTS EIS and MCTS EISN SEI messages to AVC to provide a conformant extraction procedure.



Figure 2: Workflow of OMAF AVC based viewport dependent profile [3]. Extraction marked with red dashes.

While the MCTS extraction procedure in HEVC relies on the definition of MCTS through the TMCTS SEI message, the profile targeted in OMAF AVC VDP (AVC Progressive High) lacks similar functionalities. as FMO usage is disallowed. Hence, all slices belong to a single slice group with slice\_group\_id equal to zero. Therefore, usage of the existing Motion-constrained slice group set SEI message cannot be used for the same purpose as the TMCTS SEI message in HEVC as a single slice\_group\_id value does not allow distinguishing the motion-constrained slices.

# Proposal

The proposal can be characterized by the following key aspects:

* Similar in spirit to respective HEVC SEI messages and processes
* Definition of Motion Constrained Slice Sets (MCSS) through addresses of first and last MB in MCSS EIS SEI message
* MCSS EIS SEI message for carriage of MCSS-nested parameter sets
* MCSS EISN SEI message for carriage of MCSS-nested SEI messages
* MCSS sub-bitstream extraction process based on the two SEI messages

The proposal assumes adaptive frame field coding is disabled.

**Design decisions**

In order to avoid changes to existing AVC SEI messages such as the Motion-constrained slice group set SEI message, the proposal provides a “stand alone” solution that does not rely on other SEI messages and that does not require FMO.

The use case addressed through the proposal is the one illustrated in Figure 2, i.e. slices targeted for extraction span the width of the coded picture and are vertically stacked on top of each other. In order to allow the same degree of flexibility of extraction as in the HEVC scenario, i.e. extraction of horizontal subsets of coded pictures, an AVC based solution in lack of FMO would have to resort to a slice structure as illustrated in Figure 3. However, the illustrated slice structure has a very disadvantageous coding efficiency and therefore, the slice addressing approach of the below proposed design (through macroblock lines) does not support extraction of horizontal subsets.



Figure 3: Slice structure allowing extraction of horizontal subsets.

# Specification

Motion-constrained slice set extraction information sets SEI message syntax

|  |  |  |
| --- | --- | --- |
| mcss\_extraction\_info\_sets( payloadSize ) { | **C** | **Descriptor** |
| **num\_info\_sets\_minus1** | 5 | ue(v) |
| for( i = 0; i  <=  num\_info\_sets\_minus1; i++ ) { |  |  |
| **num\_mcss\_minus1**[ i ] | 5 | ue(v) |
| for( j = 0; j  <=  num\_mcss\_minus1[ i ]; j++ ) { |  |  |
| **first\_mb\_in\_mcss**[ i ][ j ] | 5 | ue(v) |
| **last\_mb\_in\_mcss**[ i ][ j ] | 5 | ue(v) |
| } |  |  |
| **num\_sps\_in\_info\_set\_minus1**[ i ] | 5 | ue(v) |
| for( j = 0; j  <=  num\_sps\_in\_info\_set\_minus1[ i ]; j++ ) |  |  |
| **sps\_rbsp\_data\_length**[ i ][ j ] | 5 | ue(v) |
| **num\_pps\_in\_info\_set\_minus1**[ i ] | 5 | ue(v) |
| for( j = 0; j  <=  num\_pps\_in\_info\_set\_minus1[ i ]; j++ ) |  |  |
| **pps\_rbsp\_data\_length**[ i ][ j ] | 5 | ue(v) |
| while( !byte\_aligned( ) ) |  |  |
| **mcss\_alignment\_bit\_equal\_to\_zero** | 5 | f(1) |
| for( j = 0; j  <=  num\_sps\_in\_info\_set\_minus1[ i ]; j++ ) |  |  |
| for( k = 0; k  <  sps\_rbsp\_data\_length[ i ][ j ]; k++ ) |  |  |
| **sps\_rbsp\_data\_byte**[ i ][ j ][ k ] | 5 | u(8) |
| for( j = 0; j  <=  num\_pps\_in\_info\_set\_minus1[ i ]; j++ ) |  |  |
| for( k = 0; k  <  pps\_rbsp\_data\_length[ i ][ j ]; k++ ) |  |  |
| **pps\_rbsp\_data\_byte**[ i ][ j ][ k ] | 5 | u(8) |
| } |  |  |
| } |  |  |

Motion-constrained slice set extraction information sets SEI message semantics

The motion-constrained slice set extraction information sets SEI message provides supplemental information that can be used in the motion-constrained slice set (MCSS) sub-bitstream extraction as specified below to generate a conforming bitstream for an MCSS. The information consists of a number of extraction information sets, each defining a number of MCSS and containing RBSP bytes of the replacement NAL units with nal\_unit\_type equal to 7 and NAL units with nal\_unit\_type equal to 8 to be used during the MCSS sub-bitstream extraction process. Each extraction information set can be shared by multiple MCSS, i.e., is used for extraction of any of these MCSS.

This SEI message indicates that inter prediction over boundaries of identified MCSS is constrained as specified below. When present, the message shall only appear where it is associated, as specified in clause 7.4.1.2.3, with an IDR access unit.

The target picture set for this SEI message contains all consecutive primary coded pictures in decoding order starting with the associated primary coded IDR picture (inclusive) and ending with the following primary coded IDR picture (exclusive) or with the very last primary coded picture in the bitstream (inclusive) in decoding order when there is no following primary coded IDR picture. The MCSS is a collection of one or more slices, identified by addresses of first and last macroblocks of the included slices in the MCSS as indicated in this SEI message. When separate\_colour\_plane\_flag is equal to 1, the term "primary coded pictures" represents the parts of the corresponding primary coded pictures that correspond to the NAL units having the same colour\_plane\_id.

This SEI message indicates that, for each picture in the target picture set, the inter prediction process is constrained as follows: No sample value outside the identified MCSS, and no sample value at a fractional sample position that is derived using one or more sample values outside the identified MCSS is used for inter prediction of any sample within the identified MCSS.

The MCSS is defined by addresses of the first and last macroblock of the primary coded picture and the MCSS index is the value of the variable j within the loop of the num\_mcss\_minus1[ i ] + 1 sets of i-th extraction information set specified by this SEI message. It is a requirement of bitstream conformance that following conditions are true

mb\_adaptive\_frame\_field\_flag == 0 (X-X)

first\_mb\_in\_mcss[ i ][ j ] % PicWidthInMbs == 0. (X-X)

last\_mb\_in\_mcss[ i ][ j ] % PicWidthInMbs == PicWidthInMbs (X-X)

When more than one MCSS extraction information sets SEI message is present for the primary coded pictures of the target picture set, these MCSS extraction information sets SEI messages shall contain identical content.

**num\_info\_sets\_minus1** plus 1 specifies the number of extraction information sets contained in the MCSS extraction information sets SEI message. The value of num\_info\_sets\_minus1 shall be in the range of 0 to 2047, inclusive.

The i-th extraction information set is assigned an MCSS extraction information set identifier value equal to i.

**num\_mcss\_minus1**[ i ] plus 1 specifies the number of MCSS that share the i-th extraction information set. The value of num\_mcss\_minus1[ i ] shall be in the range of 0 to 2047, inclusive.

**first\_mb\_in\_mcss**[ i ][ j ] specifies the address of the first macroblock in the j-th MCSS in the i-th extraction information set. The value of first\_mb\_in\_mcss[ i ][ j ] shall be in the range of 0 to PicSizeInMbs − 1, inclusive.

**last\_mb\_in\_mcss**[ i ][ j ] specifies the address of the last macroblock in the j-th MCSS in the i-th extraction information set. The value of last\_mb\_in\_mcss[ i ][ j ] shall be in the range of 0 to PicSizeInMbs − 1, inclusive.

**num\_sps\_in\_info\_set\_minus1**[ i ] plus 1 specifies the number of replacement NAL units with nal\_unit\_type equal to 7 in the i-th extraction information set. The value of num\_sps\_in\_info\_set\_minus1[ i ] shall be in the range of 0 to 31, inclusive.

**sps\_rbsp\_data\_length**[ i ][ j ] specifies the number of RBSP data bytes of the j-th replacement NAL unit with nal\_unit\_type equal to 7 in the i-th extraction information set.

**num\_pps\_in\_info\_set\_minus1**[ i ] plus 1 specifies the number of replacement NAL units with nal\_unit\_type equal to 8 in the i-th extraction information set. The value of num\_pps\_in\_info\_set\_minus1[ i ] shall be in the range of 0 to 255, inclusive.

**pps\_rbsp\_data\_length**[ i ][ j ] specifies the number of RBSP data bytes of the j-th replacement NAL unit with nal\_unit\_type equal to 8 in the i-th extraction information set.

**mcss\_alignment\_bit\_equal\_to\_zero** shall be equal to 0.

**sps\_rbsp\_data\_byte**[ i ][ j ][ k ] contains the k-th byte of the RBSP data of the j-th replacement NAL unit with nal\_unit\_type equal to 7 in the i-th extraction information set.

**pps\_rbsp\_data\_byte**[ i ][ j ][ k ] contains the k-th byte of the RBSP data of the j-th replacement NAL unit with nal\_unit\_type equal to 8 in the i-th extraction information set.

The MCSS sub-bitstream extraction process is specified as follows:

* Let a bitstream inBitstream, a target MCSS index mcssIdxTarget, and a target MCSS extraction information set identifier mcssEisIdTarget be the inputs to the MCSS sub-bitstream extraction process.
* The output of the MCSS sub-bitstream extraction process is a sub-bitstream outBitstream derived as follows:

– The bitstream outBitstream is set to be identical to the bitstream inBitstream.

– The lists ausWithSps. and ausWithPps are set to consist of all access units within outBitstream containing non-VCL NAL units with nal\_unit\_type equal to 7, or 8.

– Remove all SEI NAL units that contain non-MCSS-nested SEI messages.

NOTE 1 – A "smart" bitstream extractor might include appropriate non-MCSS-nested SEI messages in the extracted MCSS sub-bitstream, provided that the SEI messages applicable to the MCSS sub-bitstream were present as MCSS-nested SEI messages in the original bitstream.

– Remove from outBitstream all of the following NAL units:

– VCL NAL units that contain slices with macroblock addresses lower than first\_mb\_in\_mcss[ mcssEisIdTarget ][ mcssIdxTarget ] and higher than last\_mb\_in\_mcss[ mcssEisIdTarget ][ mcssIdxTarget ].

– Non-VCL NAL units with nal\_unit\_type equal to 7, or 8.

– Insert into each access unit within the list ausWithSps in outBitstream num\_sps\_in\_info\_set\_minus1[ mcssEisIdTarget ] plus 1 NAL units with nal\_unit\_type equal to 7 generated from the RBSP data of the list of replacement NAL units with nal\_unit\_type equal to 7 in the mcssEisIdTarget-th MCSS extraction information set. For each NAL unit NAL unit with nal\_unit\_type equal to 7 that is generated, the nal\_ref\_idc is set equal to 1.

– Insert into each access unit within the list ausWithPps in outBitstream num\_pps\_in\_info\_set\_minus1[ mcssEisIdTarget ] plus 1 NAL units with nal\_unit\_type equal to 8 generated from the RBSP data of the list of replacement NAL units with nal\_unit\_type equal to 8 in the mcssEisIdTarget-th MCSS extraction information set. For each NAL unit NAL unit with nal\_unit\_type equal to 8 that is generated, the nal\_ref\_idc is set equal to 1.

NOTE 2 – The values of pic\_parameter\_set\_idin the replacement PPSs should be identical to the values of pic\_parameter\_set\_idof the removed PPSs to retain the value of slice\_pic\_parameter\_set\_id in slice headers of the original bitstream.

– For each remaining VCL NAL units in outBitstream, adjust the slice header as follows:

– For the first VCL NAL unit within each access unit, set the value of first\_mb\_in\_slice equal to 0.

– For each remaining VCL NAL units in outBitstream, let MbAddrRs be the value of the raster scan address of the last MB in the previous VCL NAL unit in bitstream order within a primary coded picture of outBitstream, set the value of first\_mb\_in\_slice equal to MbAddrRs + 1.

It is a requirement of bitstream conformance for the input bitstream that any output sub-bitstream that is the output of the MCSS sub-bitstream extraction process specified in this clause shall be a conforming bitstream.

Motion-constrained slice set extraction information nesting SEI message syntax

|  |  |  |
| --- | --- | --- |
| mcss\_extraction\_info\_nesting( payloadSize ) { | **C** | **Descriptor** |
| **num\_mcss\_minus1** | 5 | ue(v) |
| for( i = 0; i <=  num\_mcss\_minus1; i++ ) { |  |  |
| **first\_mb\_in\_mcss**[ i ] | 5 | ue(v) |
| **last\_mb\_in\_mcss**[ i ] | 5 | ue(v) |
| } |  |  |
| while( !byte\_aligned( ) ) |  |  |
| **sei\_nesting\_zero\_bit** /\* equal to 0 \*/ | 5 | f(1) |
| do |  |  |
| sei\_message( ) |  |  |
| while( more\_rbsp\_data( ) ) |  |  |
|  |  |  |
| } |  |  |

Motion-constrained slice set extraction information nesting SEI message semantics

The motion-constrained slice sets extraction information nesting SEI message, also referred to as the MCSS nesting SEI message, provides a mechanism to carry and associate the SEI messages with bitstream subsets corresponding to one or more MCSSs. An SEI message contained in an MCSS nesting SEI message is referred to as MCSS-nested or an MCSS-nested SEI message, and an SEI message that is not contained in an MCSS nesting SEI message is referred to as a non-MCSS-nested SEI message.

In the MCSS sub-bitstream extraction process as specified in the semantics of the MCSS extraction information sets SEI message, the MCSS-nested SEI messages applicable to the MCSSs in an MCSS set in an access unit can be included in the corresponding access unit of the extracted sub-bitstream as non-MCSS-nested SEI messages.

An MCSS-nesting SEI message shall not be present for a picture unless the picture belongs to the target picture set of a MCSS extraction information set SEI message. This MCSS extraction information set SEI message is referred to as the associated MCSS extraction information set SEI message of the MCSS-nesting SEI message.

An SEI NAL unit containing an MCSS nesting SEI message shall not contain any other SEI message that is not MCSS-nested in the MCSS nesting SEI message.

**num\_mcss\_minus1** plus 1 specifies the number of MCSS sets that the MCSS-nested SEI messages in this SEI message are associated with. The value of num\_mcss\_cminus1 shall be in the range of 0 to 2047, inclusive.

**first\_mb\_in\_mcss**[ i ] specifies the address of the first macroblock in the i-th MCSS that the MCSS-nested SEI messages in this SEI message are associated with. The value of first\_mb\_in\_mcss[ i ] shall be in the range of 0 to PicSizeInMbs − 1, inclusive.

**last\_mb\_in\_mcss**[ i ] specifies the address of the last macroblock in the i-th MCSS that the MCSS-nested SEI messages in this SEI message are associated with. The value of last\_mb\_in\_mcss[ i ] shall be in the range of 0 to PicSizeInMbs − 1, inclusive.

**mcss\_nesting\_zero\_bit** shall be equal to 0.

# References

1. C. Fogg, W. Husak, G. J. Sullivan, A. M. Tourapis, Y.-K. Wang, “Additional Supplemental Enhancement Information for AVC (Draft 3)”, JCTVC-AF1006, July 2018
2. R. Skupin, Y. Sanchez, C. Hellge, T. Schierl (HHI), T. Wrede, T. Christophory (SES), “AHG12: MCTS extraction usage scenario and showcase”, JCTVC-Z0032, Jan 2017
3. ISO/IEC 23090-2: Omnidirectional MediA Format
4. VR-Industry Forum (VR-IF) Guidelines, online: https://www.vr-if.org/guidelines/
5. A. Hourunranta, M. Lepistö, M. M. Hannuksela, E. Aksu, “[Systems Plenary][OMAF] Source code availability of Omnidirectional Application Format (OMAF) creator and player”, m43788, July 2018
6. Fraunhofer HHI, “Freely available tools for HEVC-based viewport-dependent OMAF video profile”, online: <https://www.hhi.fraunhofer.de/index.php?id=1608&L=1>

# Patent rights declaration(s)

**Fraunhofer HHI may have current or pending patent rights relating to the technology described in** t**his 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).**