|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  11th Meeting: Shanghai, CN, 10–19 Oct. 2012 | Document: JCTVC-K0142 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG9: Modification of SEIs specified in AVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Kimihiko Kazui, Junpei Koyama, Satoshi Shimada and Akira Nakagawa  4-1-1 Kamikodanaka, Nakahara-ku 211-8588 Kawasaki JAPAN | Tel: Email: | +81-44-754-2639  kazui.kimihiko@jp.fujitsu.com koyama.junpei@jp.fujitsu.com  sshimada@jp.fujitsu.com anaka@jp.fujitsu.com |
| *Source:* | FUJITSU LABORATORIES LTD. | | |

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

# Abstract

In this contribution it is proposed to modify the definitions of several SEIs in HEVC draft specification text version 8, whose syntax and semantics are specified in the AVC specification, in order to adapt those SEIs to the HEVC specification.

# Introduction

In draft specification text version 8[3], the syntax and semantics of the following SEIs are the same as specified in the AVC specification [1][2].

* Pan-scan rectangle SEI (\*)
* Filler payload SEI
* User data registered by Rec. ITU-T T.35 SEI
* User data unregistered SEI
* Scene information SEI
* Full-frame snapshot SEI
* Progressive refinement segment start SEI (\*)
* Progressive refinement segment end SEI
* Film grain characteristics SEI
* Deblocking filter display preference SEI (\*)
* Post-filter hint SEI
* Frame packing arrangement SEI (\*)

In the definition of those SEIs associated with (\*), the concepts such as "frame\_num" and "MB", which are not relevant to the HEVC specification, are used. Therefore it is not appropriate to specify those SEIs by just referring to the AVC specification.

In this contribution it is proposed to modify those SEIs, whose syntax and semantics are specified in the AVC specification, in order to adapt those SEIs to the HEVC specification.

# Proposal

In this section, proposed modification on pan-scan rectangle SEI, progressive refinement segment start SEI, deblocking filter display preference SEI and frame packing arrangement SEI is described. The changes in those SEIs from the AVC specification are recorded in this document.

## Clause 3: Definition

### Problem

The terms relating to interlaced video, such as "field" and "frame", are used in the semantic part of some SEIs. However, those terms are not defined in the clause 3 "Definitions".

### Proposed changes

*Insert in subclause 3:*

**field**: An assembly of alternate rows of a *frame*. A *frame* is composed of two *fields*, a *top field* and a *bottom field*.

NOTE – *field* is not referred in the decoding process.

**frame**: A *frame* consists of two *fields*, a *top field* and a *bottom field.*

NOTE – *frame* is not referred in the decoding process.

*In the end of the definition of "picture", add:*

NOTE – When both field\_pic\_flag and progressive\_source\_flag are equal to 0, a *picture* is a *frame* consisting of two *fields*, a *top field* and a *bottom field*. When field\_pic\_flag is equal to 0 and progressive\_source\_flag is equal to 1, a *picture* is a progressive *frame*. When field\_pic\_flag is equal to 1, a *picture* consists of one *field*, either a *top field* or a *bottom field*.

## Pan-scan rectangle SEI

### Problem

The terms "field" and "frame", which are not used in the decoding process, are used in the semantics of pan\_scan\_cnt\_minus1. It shall be linked with the syntax elements specified in field indication SEI.

The variable PicWidthInMbs (and its specific number '16') and frame\_crop\_*yyy*\_offset (where *yyy* is either "left", "right", "top" or "bottom"), used in the semantics of pan\_scan\_rect\_left\_offset, pan\_scan\_rect\_right\_offset, pan\_scan\_rect\_top\_offset and pan\_scan\_rect\_bottom\_offset, are not defined in HEVC.

### Proposed change

*Replace D.1.3 by*

|  |  |  |
| --- | --- | --- |
| pan\_scan\_rect( payloadSize ) { | C | Descriptor |
| **pan\_scan\_rect\_id** | 5 | ue(v) |
| **pan\_scan\_rect\_cancel\_flag** | 5 | u(1) |
| if( !pan\_scan\_rect\_cancel\_flag ) { |  |  |
| **pan\_scan\_cnt\_minus1** | 5 | ue(v) |
| for( i = 0; i <= pan\_scan\_cnt\_minus1; i++ ) { |  |  |
| **pan\_scan\_rect\_left\_offset[** i **]** | 5 | se(v) |
| **pan\_scan\_rect\_right\_offset[** i **]** | 5 | se(v) |
| **pan\_scan\_rect\_top\_offset[** i **]** | 5 | se(v) |
| **pan\_scan\_rect\_bottom\_offset[** i **]** | 5 | se(v) |
| } |  |  |
| **pan\_scan\_rect\_repetition\_period** | 5 | ue(v) |
| } |  |  |
| } |  |  |

*Replace D.2.3 by:*

The pan-scan rectangle SEI message syntax elements specify the coordinates of a rectangle relative to the cropping rectangle of the sequence parameter set. Each coordinate of this rectangle is specified in units of one-sixteenth sample spacing relative to the luma sampling grid.

**pan\_scan\_rect\_id** contains an identifying number that may be used to identify the purpose of the pan-scan rectangle (for example, to identify the rectangle as the area to be shown on a particular display device or as the area that contains a particular actor in the scene). The value of pan\_scan\_rect\_id shall be in the range of 0 to 232 − 2, inclusive.

Values of pan\_scan\_rect\_id from 0 to 255 and from 512 to 231 − 1 may be used as determined by the application. Values of pan\_scan\_rect\_id from 256 to 511 and from 231 to 232 − 2 are reserved for future use by ITU-T | ISO/IEC. Decoders encountering a value of pan\_scan\_rect\_id in the range of 256 to 511 or in the range of 231 to 232 − 2 shall ignore (remove from the bitstream and discard) it.

**pan\_scan\_rect\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous pan-scan rectangle SEI message in output order. pan\_scan\_rect\_cancel\_flag equal to 0 indicates that pan-scan rectangle information follows.

**pan\_scan\_cnt\_minus1** specifies the number of pan-scan rectangles that are present in the SEI message. pan\_scan\_cnt\_minus1 shall be in the range of 0 to 2, inclusive. pan\_scan\_cnt\_minus1 equal to 0 indicates that a single pan-scan rectangle is present that applies to the decoded picture. If the decoding picture consists of two fields (i.e. field\_pic\_flag is equal to 0 and progressive\_source\_flag is equal to 0), the single pan-scan rectangle applies to two fields. pan\_scan\_cnt\_minus1 shall be equal to 0 when the current picture is a field (i.e. field\_pic\_flag is equal to 1) or a progressive frame (i.e field\_pic\_flag is equal to 0 and progressive\_source\_flag is equal to 1). pan\_scan\_cnt\_minus1 equal to 1 indicates that two pan-scan rectangles are present, the first of which applies to the first field of the picture in output order and the second of which applies to the second field of the picture in output order. pan\_scan\_cnt\_minus1 equal to 2 indicates that three pan-scan rectangles are present, the first of which applies to the first field of the picture in output order, the second of which applies to the second field of the picture in output order, and the third of which applies to a repetition of the first field as a third field in output order.

**pan\_scan\_rect\_left\_offset[** i **]**, **pan\_scan\_rect\_right\_offset[** i **]**, **pan\_scan\_rect\_top\_offset[** i **]**,and **pan\_scan\_rect\_bottom\_offset[** i **]**, specify, as signed integer quantities in units of one-sixteenth sample spacing relative to the luma sampling grid, the location of the pan-scan rectangle. The values of each of these four syntax elements shall be in the range of −231 + 1 to 231 − 1, inclusive.

The pan-scan rectangle is specified, in units of one-sixteenth sample spacing relative to a luma frame sampling grid, as the region with frame horizontal coordinates from 16\*CropUnitX \* pic\_crop\_left\_offset + pan\_scan\_rect\_left\_offset[ i ] to 16 \* ( CtbSizeY \* PicWidthInCtbsY − CropUnitX \* pic\_crop\_right\_offset ) + pan\_scan\_rect\_right\_offset[ i ] − 1 and with vertical coordinates from 16 \*CropUnitY \* pic\_crop\_top\_offset + pan\_scan\_rect\_top\_offset[ i ] to 16 \* ( CtbSizeY \* PicHeightInCtbsY − CropUnitY \* pic\_crop\_bottom\_offset ) + pan\_scan\_rect\_bottom\_offset[ i ] − 1, inclusive. The value of 16 \* CropUnitX \* pic\_crop\_left\_offset + pan\_scan\_rect\_left\_offset[ i ] shall be less than or equal to 16 \* ( CtbSizeY \* PicWidthInCtbsY − CropUnitX \* pic\_crop\_right\_offset ) + pan\_scan\_rect\_right\_offset[ i ] − 1; and the value of 16 \* CropUnitY \* pic\_crop\_top\_offset + pan\_scan\_rect\_top\_offset[ i ] shall be less than or equal to 16 \* ( CtbSizeY \* PicHeightInCtbsY − CropUnitY \* pic\_crop\_bottom\_offset ) + pan\_scan\_rect\_bottom\_offset[ i ] − 1.

When the pan-scan rectangular area includes samples outside of the cropping rectangle, the region outside of the cropping rectangle may be filled with synthesized content (such as black video content or neutral grey video content) for display.

**pan\_scan\_rect\_repetition\_period** specifies the persistence of the pan-scan rectangle SEI message and may specify a picture order count interval within which another pan-scan rectangle SEI message with the same value of pan\_scan\_rect\_id or the end of the coded video sequence shall be present in the bitstream. The value of pan\_scan\_rect\_repetition\_period shall be in the range of 0 to 16 384, inclusive. When pan\_scan\_cnt\_minus1 is greater than 0, pan\_scan\_rect\_repetition\_period shall not be greater than 1.

pan\_scan\_rect\_repetition\_period equal to 0 specifies that the pan-scan rectangle information applies to the current decoded picture only.

pan\_scan\_rect\_repetition\_period equal to 1 specifies that the pan-scan rectangle information persists in output order until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing a pan-scan rectangle SEI message with the same value of pan\_scan\_rect\_id is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ).

pan\_scan\_rect\_repetition\_period equal to 0 or equal to 1 indicates that another pan-scan rectangle SEI message with the same value of pan\_scan\_rect\_id may or may not be present.

pan\_scan\_rect\_repetition\_period greater than 1 specifies that the pan-scan rectangle information persists until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing a pan-scan rectangle SEI message with the same value of pan\_scan\_rect\_id is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + pan\_scan\_rect\_repetition\_period.

pan\_scan\_rect\_repetition\_period greater than 1 indicates that another pan-scan rectangle SEI message with the same value of pan\_scan\_rect\_id shall be present for a picture in an access unit that is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + pan\_scan\_rect\_repetition\_period; unless the bitstream ends or a new coded video sequence begins without output of such a picture.

## Progressive refinement segment start SEI

### Problem

The variables relating to frame\_num, which is not used in HEVC, are used in the semantics of progressive refinement segment start SEI.

### Proposed change

*Replace D.1.10 by:*

|  |  |  |
| --- | --- | --- |
| progressive\_refinement\_segment\_start( payloadSize ) { | C | Descriptor |
| **progressive\_refinement\_id** | 5 | ue(v) |
| **num\_refinement\_steps\_minus1** | 5 | ue(v) |
| } |  |  |

*Replace D.2.10 by:*

The progressive refinement segment start SEI message specifies the beginning of a set of consecutive coded pictures that is labelled as the current picture followed by a sequence of one or more pictures of refinement of the quality of the current picture, rather than as a representation of a continually moving scene.

The tagged set of consecutive coded pictures shall continue until one of the following conditions is true. When a condition below becomes true, the next slice to be decoded does not belong to the tagged set of consecutive coded pictures:

– The next slice to be decoded belongs to an IDR picture.

– num\_refinement\_steps\_minus1 is greater than 0 and the pic\_order\_cnt\_lsb of the next slice to be decoded is (currPicOrderCntLsb + num\_refinement\_steps\_minus1 + 1) % MaxPicOrderCntLsb, where currPicOrderCntLsb is the value of pic\_order\_cnt\_lsb of the picture in the access unit containing the SEI message.

– num\_refinement\_steps\_minus1 is 0 and a progressive refinement segment end SEI message with the same progressive\_refinement\_id as the one in this SEI message is decoded.

The decoding order of pictures within the tagged set of consecutive pictures should be the same as their output order.

**progressive\_refinement\_id** specifies an identification number for the progressive refinement operation. progressive\_refinement\_id shall be in the range of 0 to 232 − 2, inclusive.

Values of progressive\_refinement\_id in the range of 0 to 255, inclusive, and in the range of 512 to 231 − 1, inclusive, may be used as determined by the application. Values of progressive\_refinement\_id in the range of 256 to 511, inclusive, and in the range of 231 to 232 − 2, inclusive, are reserved for future use by ITU-T | ISO/IEC. Decoders encountering a value of progressive\_refinement\_id in the range of 256 to 511, inclusive, or in the range of 231 to 232 − 2, inclusive, shall ignore (remove from the bitstream and discard) it.

**num\_refinement\_steps\_minus1** specifies the number of reference frames in the tagged set of consecutive coded pictures as follows:

– If num\_refinement\_steps\_minus1 is equal to 0, the number of reference frames in the tagged set of consecutive coded pictures is unknown.

– Otherwise, the number of reference frames in the tagged set of consecutive coded pictures is equal to num\_refinement\_steps\_minus1 + 1.

num\_refinement\_steps\_minus1 shall be in the range of 0 to MaxPicOrderCntLsb − 1, inclusive.

## Deblocking filter display preference SEI

### Problem

Unlike in AVC, sample adaptive offset process is performed after deblocking filter process. Those two filter process should be considered jointly for display preference.

Subclauses referred in the semantics of deblocking filter display preference SEI are not appropriate.

The name of syntax corresponding to "max\_dec\_frame\_buffering" is changed in the HEVC specification.

### Proposed change

*Replace D.1 with:*

|  |  |
| --- | --- |
| sei\_payload( payloadType, payloadSize ) { | Descriptor |
| if( payloadType = = 0 ) |  |
| buffering\_period( payloadSize ) |  |
| else if( payloadType = = 1 ) |  |
| pic\_timing( payloadSize ) |  |
| else if( payloadType = = 2 ) |  |
| pan\_scan\_rect( payloadSize ) |  |
| else 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 = = 6 ) |  |
| recovery\_point( payloadSize ) |  |
| else if( payloadType = = 9 ) |  |
| scene\_info( payloadSize ) |  |
| else if( payloadType = = 15 ) |  |
| full\_frame\_snapshot( payloadSize ) |  |
| else if( payloadType = = 16 ) |  |
| progressive\_refinement\_segment\_start( payloadSize ) |  |
| else if( payloadType = = 17 ) |  |
| progressive\_refinement\_segment\_end( payloadSize ) |  |
| else if( payloadType = = 19 ) |  |
| film\_grain\_characteristics( payloadSize ) |  |
| else if( payloadType = = 20 ) |  |
| inloop\_filter\_display\_preference( payloadSize ) |  |
| else if( payloadType = = 22 ) |  |
| post\_filter\_hint( payloadSize ) |  |
| else if( payloadType = = 23 ) |  |
| tone\_mapping\_info( payloadSize ) |  |
| else if( payloadType = = 45 ) |  |
| frame\_packing\_arrangement( payloadSize ) |  |
| else if( payloadType = = 47 ) [Ed. (GJS): Check numbering w.r.t. AVC.] |  |
| display\_orientation( payloadSize ) |  |
| else if( payloadType = = 128 ) |  |
| sop\_description( payloadSize ) |  |
| else if( payloadType = = 129 ) |  |
| field\_indication( payloadSize ) |  |
| else if( payloadType = = 130) |  |
| decoded\_picture\_hash( payloadSize ) |  |
| else if( payloadType = = 131) |  |
| active\_parameter\_sets( payloadSize ) |  |
| else if( payloadType = = 132 ) |  |
| sub\_pic\_timing( payloadSize ) |  |
| else |  |
| reserved\_sei\_message( payloadSize ) |  |
| if( !byte\_aligned( ) ) { |  |
| **bit\_equal\_to\_one** /\* equal to 1 \*/ | f(1) |
| while( !byte\_aligned( ) ) |  |
| **bit\_equal\_to\_zero** /\* equal to 0 \*/ | f(1) |
| } |  |
| } |  |

*Replace the title and content of D.1.13 by:*

**D.1.13 Inloop filter display preference SEI message syntax**

|  |  |  |
| --- | --- | --- |
| inloop\_filter\_display\_preference( payloadSize ) { | **C** | **Descriptor** |
| **inloop\_display\_preference\_cancel\_flag** | 5 | u(1) |
| if( !inloop\_display\_preference\_cancel\_flag ) { |  |  |
| **display\_prior\_to\_inloop\_preferred\_flag** | 5 | u(1) |
| **dec\_pic\_buffering\_constraint\_flag** | 5 | u(1) |
| **inloop\_display\_preference\_repetition\_period** | 5 | ue(v) |
| } |  |  |
| } |  |  |

*Replace the title and content of D.2.13 by:*

**D.2.13 Inloop filter display preference SEI message semantics**

This SEI message provides the decoder with an indication of whether the display of the cropped result of the in-loop filter process specified in subclause 8.7 or of the cropped result of the picture construction process prior to the in-loop process specified in subclause 8.6.5 is preferred by the encoder for the display of each decoded picture that is output.

NOTE 1 – The display process is not specified in this Recommendation | International Standard. The means by which an encoder determines what to indicate as its preference expressed in an inloop filter display preference SEI message is also not specified in this Recommendation | International Standard, and the expression of an expressed preference in an inloop filter display preference SEI message does not impose any requirement on the display process.

**inloop\_display\_preference\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous inloop filter display preference SEI message in output order. inloop\_display\_preference\_cancel\_flag equal to 0 indicates that a display\_prior\_to\_inloop\_preferred\_flag and inloop\_display\_preference\_repetition\_period follow.

NOTE 2 – In the absence of the inloop filter display preference SEI message, or after the receipt of an inloop filter display preference SEI message in which inloop\_display\_preference\_cancel\_flag is equal to 1, the decoder should infer that the display of the cropped result of the in-loop filter process specified in subclause 8.7 is preferred over the display of the cropped result of the picture construction process prior to the in-loop filter process specified in subclause 8.6.5 for the display of each decoded picture that is output.

**inloop\_prior\_to\_deblocking\_preferred\_flag** equal to 1 indicates that the encoder preference is for the display process (which is not specified in this Recommendation | International Standard) to display the cropped result of the picture construction process prior to the in-loop filter process specified in subclause 8.6.5 rather than the cropped result of the in-loop filter process specified in subclause 8.7 for each picture that is cropped and output as specified in Annex C. display\_prior\_to\_inloop\_preferred\_flag equal to 0 indicates that the encoder preference is for the display process (which is not specified in this Recommendation | International Standard) to display the cropped result of the in-loop filter process specified in subclause 8.7 rather than the cropped result of the picture construction process prior to the in-loop filter process specified in subclause 8.6.5 for each picture that is cropped and output as specified in Annex C.

NOTE 3 – The presence or absence of the inloop filter display preference SEI message and the value of display\_prior\_to\_inloop\_preferred\_flag does not affect the requirements of the decoding process specified in this Recommendation | International Standard. Rather, it only provides an indication of when, in addition to fulfilling the requirements of this Recommendation | International Standard for the decoding process, enhanced visual quality might be obtained by performing the display process (which is not specified in this Recommendation | International Standard) in an alternative fashion. Encoders that use the inloop filter display preference SEI message should be designed with an awareness that unless the encoder restricts its use of the DPB capacity specified in Annex A for the profile and level in use, some decoders may not have sufficient memory capacity for the storage of the result of the picture construction process prior to the inloop filter process specified in subclause 8.6.5 in addition to the storage of the result of the in-loop filter process specified in subclause 8.7 when reordering and delaying pictures for display, and such decoders would therefore not be able to benefit from the preference indication. By restricting its use of the DPB capacity, an encoder can be able to use at least half of the DPB capacity specified in Annex A while allowing the decoder to use the remaining capacity for storage of unfiltered pictures that have been indicated as preferable for display until the output time arrives for those pictures.

**dec\_pic\_buffering\_constraint\_flag** equal to 1 indicates that the use of the picture buffering capacity of the HRD decoded picture buffer (DPB) as specified by sps\_max\_dec\_pic\_buffering[ 0 ] has been constrained such that the coded video sequence will not require a decoded picture buffer with more than Max( 1, sps\_max\_dec\_pic\_buffering[ 0 ] ) picture buffers to enable the output of the decoded filtered or unfiltered pictures, as indicated by the inloop filter display preference SEI messages, at the output times specified by the dpb\_output\_delay of the picture timing SEI messages. dec\_pic\_buffering\_constraint\_flag equal to 0 indicates that the use of the picture buffering capacity in the HRD may or may not be constrained in the manner that would be indicated by dec\_pic\_buffering\_constraint\_flag equal to 1.

For purposes of determining the constraint imposed when dec\_pic\_buffering\_constraint\_flag is equal to 1, the quantity of picture buffering capacity used at any given point in time by each picture buffer of the DPB that contains a picture shall be derived as follows:

– If both of the following criteria are satisfied for the picture buffer, the picture buffer is considered to use two picture buffers of capacity for its storage.

– The picture buffer contains a picure that is marked as "used for reference", and

– The picture buffer contains a picture for which both of the following criteria are fulfilled:

– The HRD output time of the picture is greater than the given point in time.

– It has been indicated in an inloop filter display preference SEI message that the encoder preference for the picture is for the display process to display the cropped result of the picture construction process prior to the in-loop filter process specified in subclause 8.6.5 rather than the cropped result of the in-loop filter process specified in subclause 8.7.

– Otherwise, the picture buffer is considered to use one picture buffer of DPB capacity for its storage.

When dec\_pic\_buffering\_constraint\_flag is equal to 1, the picture buffering capacity used by all of the picture buffers in the DPB that contain pictures, as derived in this manner, shall not be greater than Max( 1, sps\_max\_dec\_pic\_buffering[ 0 ] ) during the operation of the HRD for the coded video sequence.

The value of dec\_pic\_buffering\_constraint\_flag shall be the same in all inloop filter display preference SEI messages of the coded video sequence.

**inloop\_display\_preference\_repetition\_period** specifies the persistence of the inloop filter display preference SEI message and may specify a picture order count interval within which another inloop filter display preference message or the end of the coded video sequence shall be present in the bitstream. The value of inloop\_display\_preference\_repetition\_period shall be in the range 0 to 16 384, inclusive.

inloop\_display\_preference\_repetition\_period equal to 0 specifies that the inloop filter display preference SEI message applies to the current decoded picture only.

inloop\_display\_preference\_repetition\_period equal to 1 specifies that the inloop filter display preference SEI message persists in output order until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing an inloop filter display preference SEI message is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ).

inloop\_display\_preference\_repetition\_period greater than 1 specifies that the inloop filter display preference SEI message persists until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing an inloop filter display preference SEI message is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + inloop\_display\_preference\_repetition\_period.

inloop\_display\_preference\_repetition\_period greater than 1 indicates that another inloop filter display preference SEI message shall be present for a picture in an access unit that is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + inloop\_display\_preference\_repetition\_period; unless the bitstream ends or a new coded video sequence begins without output of such a picture.

## Frame packing arrangement SEI

### Problem

There are AVC-specific terms for interlaced tools.

### Proposed change

*Replace D.1.16 by:*

|  |  |  |
| --- | --- | --- |
| frame\_packing\_arrangement( payloadSize ) { | **C** | **Descriptor** |
| **frame\_packing\_arrangement\_id** | 5 | ue(v) |
| **frame\_packing\_arrangement\_cancel\_flag** | 5 | u(1) |
| if( !frame\_packing\_arrangement\_cancel\_flag ) { |  |  |
| **frame\_packing\_arrangement\_type** | 5 | u(7) |
| **quincunx\_sampling\_flag** | 5 | u(1) |
| **content\_interpretation\_type** | 5 | u(6) |
| **spatial\_flipping\_flag** | 5 | u(1) |
| **frame0\_flipped\_flag** | 5 | u(1) |
| **field\_views\_flag** | 5 | u(1) |
| **current\_frame\_is\_frame0\_flag** | 5 | u(1) |
| **frame0\_self\_contained\_flag** | 5 | u(1) |
| **frame1\_self\_contained\_flag** | 5 | u(1) |
| if( !quincunx\_sampling\_flag &&  frame\_packing\_arrangement\_type != 5 ) { |  |  |
| **frame0\_grid\_position\_x** | 5 | u(4) |
| **frame0\_grid\_position\_y** | 5 | u(4) |
| **frame1\_grid\_position\_x** | 5 | u(4) |
| **frame1\_grid\_position\_y** | 5 | u(4) |
| } |  |  |
| **frame\_packing\_arrangement\_reserved\_byte** | 5 | u(8) |
| **frame\_packing\_arrangement\_repetition\_period** | 5 | ue(v) |
| } |  |  |
| **frame\_packing\_arrangement\_extension\_flag** | 5 | u(1) |
| } |  |  |

*Replace D.2.16 by:*

This SEI message informs the decoder that the output cropped decoded picture contains samples of multiple distinct spatially packed constituent frames that are packed into one frame using an indicated frame packing arrangement scheme. This information can be used by the decoder to appropriately rearrange the samples and process the samples of the constituent frames appropriately for display or other purposes (which are outside the scope of this Specification).

This SEI message may be associated with pictures that are either frames (i.e. field\_pic\_flag is equal to 0) or fields (i.e. field\_pic\_flag is equal to 1). The frame packing arrangement of the samples is specified in terms of the sampling structure of a frame in order to define a frame packing arrangement structure that is invariant with respect to whether a picture is a single field of such a packed frame or is a complete packed frame.

**frame\_packing\_arrangement\_id** contains an identifying number that may be used to identify the usage of the frame packing arrangement SEI message. The value of frame\_packing\_arrangement\_id shall be in the range of 0 to 232 − 2, inclusive.

Values of frame\_packing\_arrangement\_id from 0 to 255 and from 512 to 231 − 1 may be used as determined by the application. Values of frame\_packing\_arrangement\_id from 256 to 511 and from 231 to 232 − 2 are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all frame packing arrangement SEI messages containing a value of frame\_packing\_arrangement\_id in the range of 256 to 511 or in the range of 231 to 232 − 2, and bitstreams shall not contain such values.

**frame\_packing\_arrangement\_cancel\_flag** equal to 1 indicates that the frame packing arrangement SEI message cancels the persistence of any previous frame packing arrangement SEI message in output order. frame\_packing\_arrangement\_cancel\_flag equal to 0 indicates that frame packing arrangement information follows.

**frame\_packing\_arrangement\_type** indicates the type of packing arrangement of the frames as specified in Table D‑8.

Table D‑8 – Definition of frame\_packing\_arrangement\_type

|  |  |
| --- | --- |
| **Value** | **Interpretation** |
| 0 | Each component plane of the decoded frames contains a "checkerboard" based interleaving of corresponding planes of two constituent frames as illustrated in Figure D‑1. |
| 1 | Each component plane of the decoded frames contains a column based interleaving of corresponding planes of two constituent frames as illustrated in Figure D‑2 and Figure D‑3. |
| 2 | Each component plane of the decoded frames contains a row based interleaving of corresponding planes of two constituent frames as illustrated in Figure D‑4 and Figure D‑5. |
| 3 | Each component plane of the decoded frames contains a side-by-side packing arrangement of corresponding planes of two constituent frames as illustrated in Figure D‑6, Figure D‑7, and Figure D‑10. |
| 4 | Each component plane of the decoded frames contains a top-bottom packing arrangement of corresponding planes of two constituent frames as illustrated in Figure D‑8 and Figure D‑9. |
| 5 | The component planes of the decoded frames in output order form a temporal interleaving of alternating first and second constituent frames as illustrated in Figure D‑11. |

NOTE 1 – Figure D‑1 to Figure D‑10 provide typical examples of rearrangement and upconversion processing for various packing arrangement schemes. Actual characteristics of the constituent frames are signalled in detail by the subsequent syntax elements of the frame packing arrangement SEI message. In Figure D‑1 to Figure D‑10, an upconversion processing is performed on each constituent frame to produce frames having the same resolution as that of the decoded frame. An example of the upsampling method to be applied to a quincunx sampled frame as shown in Figure D‑1 or Figure D‑10 is to fill in missing positions with an average of the available spatially neighbouring samples (the average of the values of the available samples above, below, to the left and to the right of each sample to be generated). The actual upconversion process to be performed, if any, is outside the scope of this Specification.

NOTE 2 – The sample aspect ratio (SAR) indicated in the VUI parameters should indicate the output picture shape for the packed decoded frame output by a decoder that does not interpret the frame packing arrangement SEI message. In the examples shown in Figure D‑1 to Figure D‑10, the SAR produced in each upconverted colour plane would be the same as the SAR indicated in the VUI parameters, since the illustrated upconversion process produces the same total number of samples from each constituent frame as existed in the packed decoded frame.

NOTE 3 – When the output time of the samples of constituent frame 0 differs from the output time of the samples of constituent frame 1 (i.e., when field\_views\_flag is equal to 1 or frame\_packing\_arrangement\_type is equal to 5) and the display system in use presents two views simultaneously, the display time for constituent frame 0 should be delayed to coincide with the display time for constituent frame 1. (The display process is not specified in this Recommendation | International Standard.)

NOTE 4 – When field\_views\_flag is equal to 1 or frame\_packing\_arrangement\_type is equal to 5, the value 0 for fixed\_frame\_rate\_flag is not expected to be prevalent in industry use of this SEI message.

NOTE 5 – frame\_packing\_arrangement\_type equal to 5 describes a temporal interleaving process of different views.

All other values of frame\_packing\_arrangement\_type are reserved for future use by ITU-T | ISO/IEC. It is a requirement of bitstream conformance that the bitstreams shall not contain such other values of frame\_packing\_arrangement\_type.

**quincunx\_sampling\_flag** equal to 1 indicates that each colour component plane of each constituent frame is quincunx sampled as illustrated in Figure D‑1 or Figure D‑10, and quincunx\_sampling\_flag equal to 0 indicates that the colour component planes of each constituent frame are not quincunx sampled.

When frame\_packing\_arrangement\_type is equal to 0, it is a requirement of bitstream conformance that quincunx\_sampling\_flag shall be equal to 1. When frame\_packing\_arrangement\_type is equal to 5, it is a requirement of bitstream conformance that quincunx\_sampling\_flag shall be equal to 0.

NOTE 6 – For any chroma format (4:2:0, 4:2:2, or 4:4:4), the luma plane and each chroma plane is quincunx sampled as illustrated in Figure D‑1 when quincunx\_sampling\_flag is equal to 1.

**content\_interpretation\_type** indicates the intended interpretation of the constituent frames as specified in Table D‑1. Values of content\_interpretation\_type that do not appear in Table D‑1 are reserved for future specification by ITU-T | ISO/IEC.

For each specified frame packing arrangement scheme, there are two constituent frames that are referred to as frame 0 and frame 1.

Table D‑1 – Definition of content\_interpretation\_type

|  |  |
| --- | --- |
| **Value** | **Interpretation** |
| 0 | Unspecified relationship between the frame packed constituent frames |
| 1 | Indicates that the two constituent frames form the left and right views of a stereo view scene, with frame 0 being associated with the left view and frame 1 being associated with the right view |
| 2 | Indicates that the two constituent frames form the right and left views of a stereo view scene, with frame 0 being associated with the right view and frame 1 being associated with the left view |

NOTE 7 – The value 2 for content\_interpretation\_type is not expected to be prevalent in industry use of this SEI message. However, the value was specified herein for purposes of completeness.

**spatial\_flipping\_flag** equal to 1, when frame\_packing\_arrangement\_type is equal to 3 or 4, indicates that one of the two constituent frames is spatially flipped relative to its intended orientation for display or other such purposes.

When frame\_packing\_arrangement\_type is equal to 3 or 4 and spatial\_flipping\_flag is equal to 1, the type of spatial flipping that is indicated is as follows:

– If frame\_packing\_arrangement\_type is equal to 3, the indicated spatial flipping is horizontal flipping.

– Otherwise (frame\_packing\_arrangement\_type is equal to 4), the indicated spatial flipping is vertical flipping.

When frame\_packing\_arrangement\_type is not equal to 3 or 4, it is a requirement of bitstream conformance that spatial\_flipping\_flag shall be equal to 0. When frame\_packing\_arrangement\_type is not equal to 3 or 4, the value 1 for spatial\_flipping\_flag is reserved for future use by ITU-T | ISO/IEC. When frame\_packing\_arrangement\_type is not equal to 3 or 4, decoders shall ignore the value 1 for spatial\_flipping\_flag.

**frame0\_flipped\_flag**, when spatial\_flipping\_flag is equal to 1, indicates which one of the two constituent frames is flipped.

When spatial\_flipping\_flag is equal to 1, frame0\_flipped\_flag equal to 0 indicates that frame 0 is not spatially flipped and frame 1 is spatially flipped, and frame0\_flipped\_flag equal to 1 indicates that frame 0 is spatially flipped and frame 1 is not spatially flipped.

When spatial\_flipping\_flag is equal to 0, it is a requirement of bitstream conformance that frame0\_flipped\_flag shall be equal to 0. When spatial\_flipping\_flag is equal to 0, the value 1 for spatial\_flipping\_flag is reserved for future use by ITU-T | ISO/IEC. When spatial\_flipping\_flag is equal to 0, decoders shall ignore the value of frame0\_flipped\_flag.

**field\_views\_flag** equal to 1 indicates that all pictures in the current coded video sequence are coded as field pictures (i.e. field\_pic\_flag is equal to 1). All fields of a particular parity are considered a first constituent frame and all fields of the opposite parity are considered a second constituent frame. When frame\_packing\_arrangement\_type is not equal to 2, it is a requirement of bitstream conformance that the field\_views\_flag shall be equal to 0. When frame\_packing\_arrangement\_type is not equal to 2, the value 1 for field\_views\_flag is reserved for future use by ITU‑T | ISO/IEC. When frame\_packing\_arrangement\_type is not equal to 2, decoders shall ignore the value of field\_views\_flag.

**current\_frame\_is\_frame0\_flag** equal to 1, when frame\_packing\_arrangement is equal to 5, indicates that the current decoded frame is constituent frame 0 and the next decoded frame in output order is constituent frame 1, and the display time of the constituent frame 0 should be delayed to coincide with the display time of constituent frame 1. current\_frame\_is\_frame0\_flag equal to 0, when frame\_packing\_arrangement is equal to 5, indicates that the current decoded frame is constituent frame 1 and the previous decoded frame in output order is constituent frame 0, and the display time of the constituent frame 1 should not be delayed for purposes of stereo-view pairing.

When frame\_packing\_arrangement\_type is not equal to 5, the constituent frame associated with the upper-left sample of the decoded frame is considered to be consitutuent frame 0 and the other constituent frame is considered to be constituent frame 1. When frame\_packing\_arrangement\_type is not equal to 5, it is a requirement of bitstream conformance that current\_frame\_is\_frame0\_flag shall be equal to 0. When frame\_packing\_arrangement\_type is not equal to 5, the value 1 for current\_frame\_is\_frame0\_flag is reserved for future use by ITU-T | ISO/IEC. When frame\_packing\_arrangement\_type is not equal to 5, decoders shall ignore the value of current\_frame\_is\_frame0\_flag.

**frame0\_self\_contained\_flag** equal to 1 indicates that no inter prediction operations within the decoding process for the samples of constituent frame 0 of the coded video sequence refer to samples of any constituent frame 1. frame0\_self\_contained\_flag equal to 0 indicates that some inter prediction operations within the decoding process for the samples of constituent frame 0 of the coded video sequence may or may not refer to samples of some constituent frame 1. When frame\_packing\_arrangement\_type is equal to 0 or 1, it is a requirement of bitstream conformance that frame0\_self\_contained\_flag shall be equal to 0. When frame\_packing\_arrangement\_type is equal to 0 or 1, the value 1 for frame0\_self\_contained\_flag is reserved for future use by ITU-T | ISO/IEC. When frame\_packing\_arrangement\_type is equal to 0 or 1, decoders shall ignore the value of frame0\_self\_contained\_flag. Within a coded video sequence, the value of frame0\_self\_contained\_flag in all frame packing arrangement SEI messages shall be the same.

**frame1\_self\_contained\_flag** equal to 1 indicates that no inter prediction operations within the decoding process for the samples of constituent frame 1 of the coded video sequence refer to samples of any constituent frame 0. frame1\_self\_contained\_flag equal to 0 indicates that some inter prediction operations within the decoding process for the samples of constituent frame 1 of the coded video sequence may or may not refer to samples of some constituent frame 0. When frame\_packing\_arrangement\_type is equal to 0 or 1, it is a requirement of bitstream conformance that frame1\_self\_contained\_flag shall be equal to 0. When frame\_packing\_arrangement\_type is equal to 0 or 1, the value 1 for frame1\_self\_contained\_flag is reserved for future use by ITU-T | ISO/IEC. When frame\_packing\_arrangement\_type is equal to 0 or 1, decoders shall ignore the value of frame1\_self\_contained\_flag. Within a coded video sequence, the value of frame1\_self\_contained\_flag in all frame packing arrangement SEI messages shall be the same.

When quincunx\_sampling\_flag is equal to 0 and frame\_packing\_arrangement\_type is not equal to 5, two (x, y) coordinate pairs are specified to determine the indicated luma sampling grid alignment for constituent frame 0 and constituent frame 1, relative to the upper left corner of the rectangular area represented by the samples of the corresponding constituent frame.

NOTE 9 – The location of chroma samples relative to luma samples can be indicated by the chroma\_sample\_loc\_type\_top\_field and chroma\_sample\_loc\_type\_bottom\_field syntax elements in the VUI parameters.

**frame0\_grid\_position\_x** (when present) specifies the x component of the (x, y) coordinate pair for constituent frame 0.

**frame0\_grid\_position\_y** (when present) specifies the y component of the (x, y) coordinate pair for constituent frame 0.

**frame1\_grid\_position\_x** (when present) specifies the x component of the (x, y) coordinate pair for constituent frame 1.

**frame1\_grid\_position\_y** (when present) specifies the y component of the (x, y) coordinate pair for constituent frame 1.

When quincunx\_sampling\_flag is equal to 0 and frame\_packing\_arrangement\_type is not equal to 5 the (x, y) coordinate pair for each constituent frame is interpreted as follows:

– If the (x, y) coordinate pair for a constituent frame is equal to (0, 0), this indicates a default sampling grid alignment specified as follows:

– If frame\_packing\_arrangement\_type is equal to 1 or 3, the indicated position is the same as for the (x, y) coordinate pair value (4, 8), as illustrated in Figure D‑2 and Figure D‑6.

– Otherwise (frame\_packing\_arrangement\_type is equal to 2 or 4), the indicated position is the same as for the (x, y) coordinate pair value (8, 4), as illustrated in Figure D‑4 and Figure D‑8.

– Otherwise, if the (x, y) coordinate pair for a constituent frame is equal to (15, 15), this indicates that the sampling grid alignment is unknown or unspecified or specified by other means not specified in this Recommendation | International Standard.

– Otherwise, the x and y elements of the (x, y) coordinate pair specify the indicated horizontal and vertical sampling grid alignment positioning to the right of and below the upper left corner of the rectangular area represented by the corresponding constituent frame, respectively, in units of one sixteenth of the luma sample grid spacing between the samples of the columns and rows of the constituent frame that are present in the decoded frame (prior to any upsampling for display or other purposes).

NOTE 10 – The spatial location reference information frame0\_grid\_position\_x, frame0\_grid\_position\_y, frame1\_grid\_position\_x, and frame1\_grid\_position\_y is not provided when quincunx\_sampling\_flag is equal to 1 because the spatial alignment in this case is assumed to be such that constituent frame 0 and constituent frame 1 cover corresponding spatial areas with interleaved quincunx sampling patterns as illustrated in Figure D‑1 and Figure D‑10.

NOTE 11 – When frame\_packing\_arrangement\_type is equal to 2 and field\_views\_flag is equal to 1, it is suggested that frame0\_grid\_position\_y should be equal to frame1\_grid\_position\_y.

**frame\_packing\_arrangement\_reserved\_byte** is reserved for future use by ITU-T | ISO/IEC. It is a requirement of bitstream conformance that the value of frame\_packing\_arrangement\_reserved\_byte shall be equal to 0. All other values of frame\_packing\_arrangement\_reserved\_byte are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) the value of frame\_packing\_arrangement\_reserved\_byte.

**frame\_packing\_arrangement\_repetition\_period** specifies the persistence of the frame packing arrangement SEI message and may specify a frame order count interval within which another frame packing arrangement SEI message with the same value of frame\_packing\_arrangement\_id or the end of the coded video sequence shall be present in the bitstream. The value of frame\_packing\_arrangement\_repetition\_period shall be in the range of 0 to 16 384, inclusive.

frame\_packing\_arrangement\_repetition\_period equal to 0 specifies that the frame packing arrangement SEI message applies to the current decoded frame only.

frame\_packing\_arrangement\_repetition\_period equal to 1 specifies that the frame packing arrangement SEI message persists in output order until any of the following conditions are true:

– A new coded video sequence begins.

– A frame in an access unit containing a frame packing arrangement SEI message with the same value of frame\_packing\_arrangement\_id is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ).

frame\_packing\_arrangement\_repetition\_period equal to 0 or equal to 1 indicates that another frame packing arrangement SEI message with the same value of frame\_packing\_arrangement\_id may or may not be present.

frame\_packing\_arrangement\_repetition\_period greater than 1 specifies that the frame packing arrangement SEI message persists until any of the following conditions are true:

– A new coded video sequence begins.

– A frame in an access unit containing a frame packing arrangement SEI message with the same value of frame\_packing\_arrangement\_id is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + frame\_packing\_arrangement\_repetition\_period.

frame\_packing\_arrangement\_repetition\_period greater than 1 indicates that another frame packing arrangement SEI message with the same value of frame\_packing\_arrangement\_frames\_id shall be present for a frame in an access unit that is output having PicOrderCnt( ) greater than PicOrderCnt( CurrPic ) and less than or equal to PicOrderCnt( CurrPic ) + frame\_packing\_arrangement\_repetition\_period; unless the bitstream ends or a new coded video sequence begins without output of such a frame.

**frame\_packing\_arrangement\_extension\_flag** equal to 0 indicates that no additional data follows within the frame packing arrangement SEI message. It is a requirement of bitstream conformance that the value of frame\_packing\_arrangement\_extension\_flag shall be equal to 0. The value 1 for frame\_packing\_arrangement\_extension\_flag is reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value 1 for frame\_packing\_arrangement\_extension\_flag in a frame packing arrangement SEI message and shall ignore all data that follows within a frame packing arrangement SEI message after the value 1 for frame\_packing\_arrangement\_extension\_flag.



Figure D‑1 – Rearrangement and upconversion of checkerboard interleaving   
(frame\_packing\_arrangement\_type equal to 0)



Figure D‑2 – Rearrangement and upconversion of column interleaving   
with frame\_packing\_arrangement\_type equal to 1, quincunx\_sampling\_flag equal to 0,  
and (x, y) equal to (0, 0) or (4, 8) for both constituent frames



Figure D‑3 – Rearrangement and upconversion of column interleaving with  
frame\_packing\_arrangement\_type equal to 1, quincunx\_sampling\_flag equal to 0,   
(x, y) equal to (0, 0) or (4, 8) for constituent frame 0 and (x, y) equal to (12, 8) for constituent frame 1



Figure D‑4 – Rearrangement and upconversion of row interleaving with  
frame\_packing\_arrangement\_type equal to 2, quincunx\_sampling\_flag equal to 0,  
and (x, y) equal to (0, 0) or (8, 4) for both constituent frames



Figure D‑5 – Rearrangement and upconversion of row interleaving with  
frame\_packing\_arrangement\_type equal to 2, quincunx\_sampling\_flag equal to 0,   
(x, y) equal to (0, 0) or (8, 4) for constituent frame 0, and (x, y) equal to (8, 12) for constituent frame 1



Figure D‑6 – Rearrangement and upconversion of side-by-side packing arrangement with  
frame\_packing\_arrangement\_type equal to 3, quincunx\_sampling\_flag equal to 0,  
and (x, y) equal to (0, 0) or (4, 8) for both constituent frames



Figure D‑7 – Rearrangement and upconversion of side-by-side packing arrangement with frame\_packing\_arrangement\_type equal to 3, quincunx\_sampling\_flag equal to 0,  
(x, y) equal to (12, 8) for constituent frame 0, and (x, y) equal to (0, 0) or (4, 8) for constituent frame 1



Figure D‑8 – Rearrangement and upconversion of top-bottom packing arrangement with  
frame\_packing\_arrangement\_type equal to 4, quincunx\_sampling\_flag equal to 0,  
and (x, y) equal to (0, 0) or (8, 4) for both constituent frames



Figure D‑9 – Rearrangement and upconversion of top-bottom packing arrangement with frame\_packing\_arrangement\_type equal to 4, quincunx\_sampling\_flag equal to 0,  
(x, y) equal to (8, 12) for constituent frame 0, and (x, y) equal to (0, 0) or (8, 4) for constituent frame 1



Figure D‑10 – Rearrangement and upconversion of side-by-side packing arrangement with quincunx sampling  
(frame\_packing\_arrangement\_type equal to 3 with quincunx\_sampling\_flag equal to 1)



Figure D‑11 – Rearrangement of a temporal interleaving frame arrangement   
(frame\_packing\_arrangement\_type equal to 5)

# References

[1] Rec. ITU-T H.264 (in force), *Advanced video coding for generic audiovisual services*.

[2] ISO/IEC 14496-10: in force, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*

[3] "High Efficiency Video Coding (HEVC) text specification draft 8", JCTVC-J1003, July 2012, Stockholm Sweden

[4] Proposed Editorial Improvements for High efficiency video coding (HEVC) Text Specification Draft 8, JCTVC-K0030, October 2012, Shanghai China

# Patent rights declaration(s)

**FUJITSU LIMITED 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).**