|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  14th Meeting: Vienna, AT, 25 July – 2 Aug. 2013 | Document: JCTVC-N0212r1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On inter-layer parameter set** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Yong He, Yan Ye, Yuwen He 9710 Scranton R-D, #250  San Diego, CA 92121  USA | Tel: Email: | 1-858-210-4807 Yong.He@InterDigital.com |
| *Source:* | InterDigital Communications, Inc. | | |

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

# Abstract

This contribution document proposes a new parameter set, inter-layer parameter set (IPS), to address the redundant issue of SPS syntax elements across multiple layers. The proposed IPS aggregates groups of parameters shared by multiple layers, and each layer could refer to the IPS parameters. As a result, the number of bits used for SPS would be significantly reduced when multiple layers shares the identical parameter values. The IPS could further remove the parsing dependency when the parameters placed in different parameter sets could be included in the same IPS. In addition, the proposed new parameter set also avoids the problems caused by SPS to SPS, or VPS to SPS parameter set prediction.

# Introduction

Current high level syntax design allows multiple layers to refer to same SPS, which is typically used for multi-view and SNR scalability. While for spatial scalability, each layer may refer to different SPS due to different video resolutions. Given the fact that the majority parameters in SPS would be identical across multiple layers, it is desirable to save bitrate by removing such redundancy and sharing the same parameters among multiple layers.

Various solutions were proposed in 13th JCTVC meetings to save the bitrate [1][2][3][4]. All these proposals can be grouped into two kinds of approaches, the first one is SPS to SPS prediction which predicts the enhancement layer SPS parameters such as scaling list, reference picture set (RPS), etc, from the SPS parameters of the base layer or another dependent layer; another approach is VPS to SPS prediction which relocates the shared parameters across layers to the VPS, and then predicts these shared SPS parameters of each layer in the VPS.

Both approaches have pros and cons. The SPS to SPS prediction may introduce the parameter set dependency among different layers. The approach to use VPS to hold the shared parameters may not align with the design philosophy of VPS, which is for capability exchange and session negotiation. Another concern to use VPS as place holder is that VPS would be carried in a base layer and the base layer could be rather large after sub-bitstream extraction if all shared parameters are added into VPS. In addition, any change of any parameter values in one layer has to require a new VPS.

In 13th JCTVC meeting, AHG suggested the following criteria for the design of VPS and SPS for HEVC extensions :

1. the VPS should (per its primary design purpose) include parameters that are useful for bitstream extraction and/or capability exchange,
2. parameters that are not required for bitstream extraction and/or capability exchange should not be present in the VPS,
3. Sequence-level parameters that may change in some layers while not affecting the general layer structure should not be present in the VPS.
4. For the design of the SPS with nuh\_layer\_id great than 0, it is suggested that parameters present in the VPS should not be duplicated in the SPS (except for those parameters that are required for base layer decoding), and sequence-level parameters that are not present in the VPS should be included in the SPS.
5. It is considered desirable for DPB related parameters to be present in the VPS extension.

Given the above restriction, majority parameters that might be shared among multiple layers could not be present in the VPS and its extensions to save the overhead bit.

# Inter-layer Parameter Set

A new parameter set, inter-layer parameter set (IPS), is proposed here to aggregate these high level syntax elements that can be shared among multiple layers. Each layer could then refer to the IPS parameters to save the overhead bit. There are several benefits by introducing IPS:

1. Since IPS is not carried in the base layer, the size of IPS does not impact the base layer sub-stream.
2. To keep VPS intact for the purpose of bitstream extraction, session negotiation and capability exchange.
3. To improve the high level signaling efficiency by facilitating the prediction of shared parameters across multiple layers.
4. To remove the parsing dependency because the parameters used to be placed in different parameter sets are now included in the same IPS and the parsing of each parameter does not rely on the parsing result from another different parameter set.

## IPS Syntax structure

For the parameters currently present in SPS, the parameters serving the similar purpose are grouped into one subset such as video format subset, coding parameter subset, scaling list subset, scaled offset subset, and VUI subset. In one IPS, each subset can have a plurality of parameter values; this allows a given enhancement layer to refer to a plurality of parameter values by indexing into IPS and the subset. In such way, the minimum number of syntax elements can be signaled to cover parameters used by multiple layers. Table 1 is the syntax table showing the proposed IPS.

1. Proposed Inter-layer Parameter Set RBSP Syntax

|  |  |
| --- | --- |
| Inter\_layer\_parameter\_set\_rbsp( ) { | Descriptor |
| **ips\_inter\_layer\_ parameter\_set\_id** | u(4) |
| **num\_video\_format\_subsets** | ue(v) |
| for ( idx = 0; idx < num\_video\_format\_subsets; idx++ ) |  |
| ips\_video\_format\_subset(idx) |  |
| **num\_coding\_param\_subsets** | ue(v) |
| for ( idx = 0; idx < num\_coding\_param\_subsets; idx++) |  |
| ips\_coding\_param\_subset(idx) |  |
| **num\_pcm\_param\_subsets** | ue(v) |
| for ( i = 0; i < num\_pcm\_param\_subsets; idx ++) |  |
| ips\_pcm\_param\_subset(idx) |  |
| **num\_scaling\_list\_subsets** | ue(v) |
| for ( idx = 0; idx < num\_scaling\_list\_subsets; idx ++) |  |
| ips\_scaling\_list\_subset(idx) |  |
| **num\_short\_term\_ref\_pic\_sets** | ue(v) |
| for( i = 0; i < num\_short\_term\_ref\_pic\_sets; i++) |  |
| short\_term\_ref\_pic\_set( i ) |  |
| **num\_long\_term\_ref\_pic\_sets** | ue(v) |
| for( i = 0; i < num\_long\_term\_ref\_pic\_sets; i++) { |  |
| **num\_long\_term\_ref\_pics\_sps[i]** | ue(v) |
| for( j = 0; j < num\_long\_term\_ref\_pics\_sps[i]; j++ ) { |  |
| **lt\_ref\_pic\_poc\_lsb\_sps**[i][j] | u(v) |
| **used\_by\_curr\_pic\_lt\_sps\_flag**[i][j] | u(1) |
| } |  |
| } |  |
| **num\_scaled\_ref\_layer\_offset\_subsets** | ue(v) |
| for (idx = 0; idx < num\_scaled\_ref\_layer\_offset\_subsets; idx ++) |  |
| ips\_scaled\_ref\_layer\_offset\_subset(idx) |  |
| **num\_vui\_param\_subsets** | ue(v) |
| for (i = 0; i < num\_vui\_param\_subsets; i++) |  |
| ips\_vui\_param\_subset(idx) |  |
| **num\_vui\_extension\_param\_subsets** | ue(v) |
| for (i = 0; i < num\_vui\_extension\_param\_subsets; i++) |  |
| ips\_vui\_extension\_param\_subset(idx) |  |
| rbsp\_trailing\_bits( ) |  |
| } |  |

The semantics of the syntax elements in IPS are defined as follows:

**ips\_inter\_layer\_parameter\_set\_id** identifies the IPS for reference by other syntax elements.

**num\_video\_format\_subsets** specifies the number of video format structure (ips\_video\_format\_subset).

**num\_coding\_param\_susets** specifies the number of coding parameter structure (ips\_coding\_param\_subset).

**num\_pcm\_param\_subsets** specifies the number of PCM coding parameter structure (ips\_pcm\_param\_subset).

**num\_scaling\_list\_subsets** specifies the number of scaling list structure (ips\_scaling\_list\_subset).

**num\_scaled\_ref\_layer\_offset\_subset** specifies the number of scaled reference layer offset structures(ips\_scaled\_ref\_layer\_offset\_subset).

**num\_vui\_param\_subsets** specifies the VUI parameter structure (ips\_vui\_param\_subset).

**num\_vui\_param\_subsets** specifies the SPS extension VUI parameter structure (ips\_vui\_extension\_param\_subset).

Table 2 is the parameter subsets specified within IPS. Each subset is a group of parameters currently signaled in SPS serving the similar purpose.

1. Parameter subsets in IPS

|  |  |
| --- | --- |
| ips\_video\_format\_subset( idx) { | Descriptor |
| **chroma\_format\_idc** | ue(v) |
| if( chroma\_format\_idc = = 3 ) |  |
| **separate\_colour\_plane\_flag** | u(1) |
| **pic\_width\_in\_luma\_samples** | ue(v) |
| **pic\_height\_in\_luma\_samples** | ue(v) |
| **conformance\_window\_flag** | u(1) |
| if( conformance\_window\_flag ) { |  |
| **conf\_win\_left\_offset** | ue(v) |
| **conf\_win\_right\_offset** | ue(v) |
| **conf\_win\_top\_offset** | ue(v) |
| **conf\_win\_bottom\_offset** | ue(v) |
| } |  |
| **bit\_depth\_luma\_minus8** | ue(v) |
| **bit\_depth\_chroma\_minus8** | ue(v) |
| } |  |

|  |  |
| --- | --- |
| ips\_coding\_param\_subset( idx) { | Descriptor |
| **log2\_max\_pic\_order\_cnt\_lsb\_minus4** | ue(v) |
| **sps\_sub\_layer\_ordering\_info\_present\_flag** | u(1) |
| for( i = ( sps\_sub\_layer\_ordering\_info\_present\_flag ? 0 : sps\_max\_sub\_layers\_minus1 );  i <= sps\_max\_sub\_layers\_minus1; i++ ) { |  |
| **sps\_max\_dec\_pic\_buffering\_minus1**[i] | ue(v) |
| **sps\_max\_num\_reorder\_pics**[i] | ue(v) |
| **sps\_max\_latency\_increase\_plus1** [ i ] | ue(v) |
| } |  |
| **log2\_min\_luma\_coding\_block\_size\_minus3** | ue(v) |
| **log2\_diff\_max\_min\_luma\_coding\_block\_size** | ue(v) |
| **log2\_min\_transform\_block\_size\_minus2** | ue(v) |
| **log2\_diff\_max\_min\_transform\_block\_size** | ue(v) |
| **max\_transform\_hierarchy\_depth\_inter** | ue(v) |
| **max\_transform\_hierarchy\_depth\_intra** | ue(v) |
| } |  |

|  |  |
| --- | --- |
| ips\_pcm\_param\_subset( idx) { | Descriptor |
| **pcm\_sample\_bit\_depth\_luma\_minus1** | u(4) |
| **pcm\_sample\_bit\_depth\_chroma\_minus1** | u(4) |
| **log2\_min\_pcm\_luma\_coding\_block\_size\_minus3** | ue(v) |
| **log2\_diff\_max\_min\_pcm\_luma\_coding\_block\_size** | ue(v) |
| **pcm\_loop\_filter\_disabled\_flag** | u(1) |
| **}** |  |

|  |  |
| --- | --- |
| ips\_scaling\_list\_subset( idx) { | Descriptor |
| **scaling\_list\_enabled\_flag** | u(1) |
| if( scaling\_list\_enabled\_flag ) { |  |
| **sps\_scaling\_list\_data\_present\_flag** | u(1) |
| if( sps\_scaling\_list\_data\_present\_flag ) |  |
| scaling\_list\_data(idx ) |  |
| } |  |

|  |  |
| --- | --- |
| ips\_scaled\_ref\_layer\_offset\_subset( idx) { | Descriptor |
| **scaled\_ref\_layer\_offset\_present\_flag** |  |
| if (scaled\_ref\_layer\_offset\_present\_flag{ |  |
| **scaled\_ref\_layer\_left\_offset** | se(v) |
| **scaled\_ref\_layer\_top\_offset** | se(v) |
| **scaled\_ref\_layer\_right\_offset** | se(v) |
| **scaled\_ref\_layer\_bottom\_offset** | se(v) |
| **}** |  |
| **}** |  |

|  |  |
| --- | --- |
| ips\_vui\_param\_subset( idx) { | Descriptor |
| **vui\_parameters\_present\_flag** |  |
| if (vui\_parameters\_present\_flag) |  |
| vui\_parameters() |  |
| } |  |

|  |  |
| --- | --- |
| ips\_vui\_extension\_param\_subset( idx) { | Descriptor |
| **num\_ilp\_restricted\_ref\_layers** | | ue(v) |
| for( i = 0; i < num\_ilp\_restricted\_ref\_layers; i++ ) { | |  |
| **min\_spatial\_segment\_offset\_plus1**[ i ] | | ue(v) |
| if( min\_spatial\_segment\_offset[ i ] > 0 ) { | |  |
| **ctu\_based\_offset\_enabled\_flag**[ i ] | | u(1) |
| if( ctu\_based\_offset\_enabled\_flag[ i ] ) | |  |
| **min\_horizontal\_ctu\_offset\_plus1**[ i ] | | ue(v) |
| } | |  |
| } | |  |
| } |  |

With the proposed IPS, the SPS of each enhancement can refer the index of each parameter subset and the size of SPS would be significantly reduced. Table 3 is the proposed SPS structure.

1. Proposed Sequence Parameter Set and extension RBSP syntax

|  |  |
| --- | --- |
| seq\_parameter\_set\_rbsp( ) { | Descriptor |
| **sps\_video\_parameter\_set\_id** | u(4) |
| if (nuh\_layer\_id > 0) |  |
| **sps\_inter\_layer\_parameter\_set\_id** | u(4) |
| if( nuh\_layer\_id = = 0 ) { |  |
| **sps\_max\_sub\_layers\_minus1** | u(3) |
| **sps\_temporal\_id\_nesting\_flag** | u(1) |
| profile\_tier\_level( 1, sps\_max\_sub\_layers\_minus1 ) |  |
| } |  |
| **sps\_seq\_parameter\_set\_id** | ue(v) |
| if (nuh\_layer\_id == 0) { |  |
| **...** |  |
| **}** |  |
| else { |  |
| **ips\_video\_format\_subsets\_index** | ue(v) |
| **ips\_coding\_param\_subsets\_index** | ue(v) |
| **ips\_scaling\_list\_subsets\_index** | ue(v) |
| **amp\_enabled\_flag** | u(1) |
| **sample\_adaptive\_offset\_enabled\_flag** | u(1) |
| **pcm\_enabled\_flag** | u(1) |
| if( pcm\_enabled\_flag ) |  |
| **ips\_pcm\_param\_subsets\_index** | ue(v) |
| **sps\_temporal\_mvp\_enabled\_flag** | u(1) |
| **strong\_intra\_smoothing\_enabled\_flag** | u(1) |
| **ips\_vui\_param\_subsets\_index** | ue(v) |
| } |  |
| **sps\_extension\_flag** | u(1) |
| if( sps\_extension\_flag ) { |  |
| sps\_extension( ) |  |
| **sps\_extension2\_flag** | u(1) |
| if( sps\_extension2\_flag ) |  |
| while( more\_rbsp\_data( ) ) |  |
| **sps\_extension\_data\_flag** | u(1) |
| **}** |  |
| rbsp\_trailing\_bits( ) |  |
| } |  |

|  |  |
| --- | --- |
| sps\_extension( ) { | **Descriptor** |
| **inter\_view\_mv\_vert\_constraint\_flag** | u(1) |
| **ips\_vui\_extension\_param\_subsets\_index** | ue(v) |
| sps\_extension\_vui\_parameters( ) |  |
| **ips**\_**scaled\_ref\_layer\_offset\_subsets\_index** | ue(v) |
| **num\_scaled\_ref\_layer\_offsets** | ue(v) |
| for( i = 0; i < num\_scaled\_ref\_layer\_offsets; i++) { |  |
| **scaled\_ref\_layer\_left\_offset**[ i ] | se(v) |
| **scaled\_ref\_layer\_top\_offset**[ i ] | se(v) |
| **scaled\_ref\_layer\_right\_offset**[ i ] | se(v) |
| **scaled\_ref\_layer\_bottom\_offset**[ i ] | se(v) |
| **}** |  |
| } |  |

**sps\_inter\_layer\_parameter\_subset\_id** specifies the value of the ips\_inter\_layer\_parameter\_set\_id of the active IPS.

**ips\_video\_format\_subsets\_index** specifies the index, into the list of video format syntax structures included in the active IPS. The range of ips\_video\_format\_subsets\_index is from 0 to num\_video\_format\_subsets\_index, exclusive.

**ips\_coding\_param\_subsets\_index** specifies the index, into the list of coding parameter syntax structures included in the active IPS. The range of ips\_coding\_param\_subsets\_index is from 0 to num\_coding\_param\_subsets\_index, exclusive.

**ips\_scaling\_list\_subsets\_index** specifies the index, into the list of scaling list syntax structures included in the active IPS. The range of ips\_scaling\_list\_subsets\_index is from 0 to num\_scaling\_list\_subsets\_index, exclusive.

**ips\_pcm\_param\_subsets\_index** specifies the index, into the list of PCM parameter syntax structures included in the active IPS. The range of ips\_pcm\_param\_subsets\_index is from 0 to num\_pcm\_param\_subsets\_index, exclusive.

**ips\_vui\_param\_subsets\_index** specifies the index, into the list of VUI syntax structures included in the active IPS. The range of ips\_vui\_param\_subsets\_index is from 0 to num\_vui\_param\_subsets\_index, exclusive.

**ips\_vui\_extension\_param\_subsets\_index** specifies the index, into the list of VUI extension syntax structures included in the active IPS. The range of ips\_vui\_param\_subsets\_index is from 0 to num\_vui\_extension\_param\_subsets\_index, exclusive.

**ips**\_**scaled\_ref\_layer\_offset\_subset\_index** specifies the index, into the list of video format syntax structures included in the active IPS. The range of ips\_scaled\_ref\_layer\_offset\_subsets\_index is from 0 to num\_scaled\_ref\_layer\_offset\_subsets\_index, exclusive.

These syntax structure indexes allow a given enhancement layer to derive a plurality of parameter values by indexing into IPS and the particular subset.

Because IPS is only applicable to the one or more enhancement layers in a scalable coding system, the nuh\_layer\_id value of an IPS NAL unit shall not be 0 for a conforming bitstream. For example, a conforming bitstream may have the nuh\_layer\_id of all IPS NAL units equal to 1.

## The order of IPS RBSP and its activation

An IPS RBSP includes parameters that can be referred to by one or more SPS RBSPs. Each IPS RBSP is initially considered not active at the start of the operation of the decoding process. At most one IPS RBSP is considered active at any given moment during the operation of the decoding process, and the activation of any particular IPS RBSP results in the deactivation of the previously-activated IPS RBSP (if any).

When an IPS RBSP is not already active and it is referred to by activation of an SPS RBSP (in which sps\_inter\_layer\_parameter\_set\_id is equal to the ips\_inter\_layer\_parameter\_set\_id value), , it is activated. This IPS RBSP is called the active IPS RBSP until it is deactivated by the activation of another IPS RBSP. An IPS RBSP, with that particular value of ips\_inter\_layer\_parameter\_subset\_id, shall be available to the decoding process prior to its activation.

# Results

Table 4 shows the number of bits of each parameter subset and SPS for spatial 2x resolutions. Majority of parameter subsets, such as coding parameters and stRPS, would be identical among multiple layers. By allocating these parameters into IPS, the size of SPS can be reduced significantly.

1. Bit count of each parameter subset for CTC

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video format | Coding | PCM | Scaling list | stRPS | ltRPS | Scaled ref layer offsets | VUI | Total SPS  (bits) |
| RA (960x540) | 50 | 35 | 1 | 1 | 139 | 1 | 0 | 1 | 245 |
| RA (1920x1080) | 52 | 35 | 1 | 1 | 139 | 1 | 10 | 1 | 257 |
| LDB/LDP  (960x540) | 50 | 31 | 1 | 1 | 130 | 1 | 0 | 1 | 232 |
| LDB/LDP  (1920x1080) | 52 | 31 | 1 | 1 | 130 | 1 | 10 | 1 | 244 |

Table 5 shows the bits comparison between current SPS and the proposed IPS approach for different number of spatial scalable enhancement layers. Over 40% bits could be saved when 3 or more enhancement layers are available.

1. Bit count comparison

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Anchor | N0212 | | | Bit saving Percentage |
| # ELs |  | SPS (bit) | SPS (bit) | IPS(bit) | Total (bit) |  |
| 2 | RA | 502 | 56 | 302 | 358 | 28.7% |
| LDP/LDB | 476 | 30 | 289 | 319 | 33% |
| 3 | RA | 761 | 86 | 302 | 388 | 49% |
| LDP/LDB | 722 | 47 | 289 | 336 | 53.5% |
| 4 | RA | 1020 | 116 | 302 | 418 | 59% |
| LDP/LDB | 968 | 64 | 289 | 353 | 63.5% |

# Reference

1. JCTVC-M0211, AHG9 On Scaling Signaling for SHVC, April 2013
2. JCTVC-M0268, AHG9 VPS and SPS designs in HEVC 3DV and scalable extensions, April 2013
3. JCTVC-M0134, AHG9 Inter-layer SPS prediction for HEVC extensions, April 2013
4. JCTVC-M0140, AHG9 Interlayer RPS prediction, April 2013
5. JCTVC-M0165, On parameter set prediction, April 2013JCTVC-M0450, Report of the joint BoG on high level syntax, April 2013

# Patent rights declaration(s)

**InterDigital Communications, 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).**