|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  2nd Meeting: Shanghai, CN, 13–19 Oct. 2012 | Document: JCT3V-C0041 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Proposed VPS extension semantics and editorial cleanups to syntax** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jill Boyce  Ye-Kui Wang  Sachin Deshpande | Email: [jill@vidyo.com](mailto:jill@vidyo.com)  [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com)  [sdeshpande@sharplabs.com](mailto:sdeshpande@sharplabs.com) |  |
| *Source:* | Vidyo, Qualcomm, Sharp | | |

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

# Abstract

The output document from the Shanghai meeting JCTVC-K1007/JCT3V-B1007 contains the working design of the syntax for the NAL unit header and VPS extension for the scalability and multiview extensions, but lacks semantics for many of the syntax elements in the VPS extension. This contribution proposes semantics for those VPS extension syntax elements for which the semantics were missing in JCTVC-K1007/JCT3V-B1007. A correction for an error in the VPS syntax of JCTVC-K1007/JCT3V-B1007 is also proposed. Additionally, some purely editorial changes to the syntax table are proposed which simplify the semantics. Changes from JCTVC-K1007/JCT3V-B1007 are highlighted.

# Video parameter set extension syntax and semantics

## Video parameter set extension syntax

An editorial change is proposed to include a length 16 array of 1-bit scalability\_mask\_flag syntax elements rather than a single 16-bit field, to simplify referencing of individual bits.

The loop over the dimension\_id[ i ][ j ] syntax element is corrected to avoid referencing the no longer present num\_dimensions\_minus1 syntax element, with some additional editorial changes which simplify the semantics. The changes relatives to JCTVC-K1007/JCT3V-B1007 are highlighted in the table below.

|  |  |
| --- | --- |
| vps\_extension( ) { | Descriptor |
| while( !byte\_aligned( ) ) |  |
| **vps\_extension\_byte\_alignment\_reserved\_one\_bit** | u(1) |
| **avc\_base\_layer\_flag** | u(1) |
| for( i = 0, numScalabilityTypes = 0 ; i <16; i++ ) { |  |
| **scalability\_mask\_flag**[ i ] | u(1) |
| numScalabilityTypes += scalability\_mask\_flag[ i ] |  |
| } |  |
| for( i = 0; i <NumScalabilityTypes; i++ ) { |  |
| **dimension\_id\_len\_minus1**[ i ] | u(3) |
| } |  |
| **vps\_nuh\_layer\_id\_present\_flag** | u(1) |
| // layer specific information |  |
| for( i = 1; i <= vps\_max\_layers\_minus1; i++ ) { |  |
| if( vps\_nuh\_layer\_id\_present\_flag ) |  |
| **layer\_id\_in\_nuh**[ i ] | u(6) |
| for( j = 0; j < numScalabilityTypes; j++ ) |  |
| **dimension\_id**[ i ][ j ] | u(v) |
| } |  |
| for( i = 1; i <= vps\_max\_layers\_minus1 ; i++ ) |  |
| profile\_tier\_level( 1, vps\_max\_sub\_layers\_minus1 ) |  |
| for( i = 1; i <= vps\_max\_layers\_minus1; i++ ) { |  |
| // layer dependency |  |
| **num\_direct\_ref\_layers**[ i ] | u(6) |
| for( j = 0; j < num\_direct\_ref\_layers[ i ]; j++ ) |  |
| **ref\_layer\_id**[ i ][ j ] | u(6) |
| } |  |
| } |  |

## Video parameter set extension semantics

Additional semantics are proposed, which had been missing in JCTVC-K1007.

**vps\_extension\_byte\_alignment\_reserved\_one\_bit** shall be equal to 1.

**avc\_base\_layer\_flag** equal to 1 specifies that the base layer conforms to Rec. ITU-T H.264 | ISO/IEC 14496-10, equal to 0 specifies that it conforms to this specification.

When avc\_base\_layer\_flag equal to 1, in the Rec. ITU-T H.264 | ISO/IEC 14496-10 conforming base layer, after applying the Rec. ITU-T H.264 | ISO/IEC 14496-10 decoding process for reference picture lists construction the output reference picture lists refPicList0 and refPicList1 (when applicable) does not contain any pictures for which the TemporalId is greater than TemporalId of the coded picture. All sub-bitstreams of the Rec. ITU-T H.264 | ISO/IEC 14496-10 conforming base layer, that can be derived using the sub-bitstream extraction process as specified in Rec. ITU­T H.264 | ISO/IEC 14496-10 subclause G.8.8.1 with any value for temporal\_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Rec. ITU­T H.264 | ISO/IEC 14496-10 Annexes A, G and H.

**scalability\_mask\_flag** [ i ] equal to 1 indicates that dimension\_id syntax elements corresponding to the i-th scalability dimension in Table XX are present. scalability\_mask\_flag[ i ] equal to 0 indicates that dimension\_id syntax elements corresponding to the i-th scalability dimension are not present.

Table XX. Mapping of ScalabiltyId to scalability dimensions

|  |  |  |
| --- | --- | --- |
| **scalability\_mask\_flag**  **index** | **Scalability dimension** | **ScalabiltyId mapping** |
| 0 | spatial | DependencyId |
| 1 | quality | QualityId |
| 2 | depth | DepthFlag |
| 3 | multiview | ViewId |
| 4 - 15 | Reserved | Reserved |

**dimension\_id\_len\_minus1**[ j ] plus1 specifies the length, in bits, of the dimension\_id[ i ][ j ] syntax element.

**vps\_nuh\_layer\_id\_present\_flag** specifies whether the layer\_id\_in\_nuh[ i ] syntax is present.

**layer\_id\_in\_nuh**[ i ] specifies the value of the nuh\_layer\_id syntax element in VCL NAL units of the i-th layer. When not present, the value of layer\_id\_in\_nuh[ i ] is inferred to be equal to i.

**dimension\_id**[ i ][ j ] specifies the identifier of the j-th scalability dimension type of the i-th layer. When not present, the value of dimension\_id[ i ][ j ] is inferred to be equal to 0. The number of bits used for the representation of dimension\_id[ i ][ j ] is dimension\_id\_len\_minus1[ j ] + 1 bits.

The following applies:

for( i = 0; i <= vps\_max\_layers\_minus1; i++ )

for( smIdx= 0, j =0; smIdx< 16; smIdx ++ )

if( ( i != 0 ) && scalability\_mask\_flag[ smIdx ] )

ScalabilityId[ i ][ smIdx ] = dimension\_id[ i ][ j++ ]

else

ScalabilityId[ i ][ smIdx ] = 0

The following applies:

- For each layer indexed by i = 0..vps\_max\_layers\_minus1 + 1

* DependencyId[ layer\_id\_in\_nuh[ i ] ] = ScalabilityId[ i ][ 0 ]
* QualityId[ layer\_id\_in\_nuh[ i ] ] = ScalabilityId[ i ][ 1 ]
* DepthFlag[ layer\_id\_in\_nuh[ i ] ] = ScalabilityId[ i ][2 ]
* ViewId[ layer\_id\_in\_nuh[ i ] ] = ScalabilityId[ i ][ 3 ]

**num\_direct\_ref\_layers**[ i ] specifies the number of layers the i-th layer directly references. [Ed.(YK): Add the exact meaning of a layer directly referencing another layer.]

**ref\_layer\_id**[ i ][ j ] identifies the j-th layer the i-th layer directly references.

# Patent rights declaration(s)

**Vidyo may have IPR 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).**

**Qualcomm may have IPR 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).**

**Sharp may have IPR 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).**