|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  15th Meeting: Geneva, CH, 23 Oct. – 1 Nov. 2013 | Document: JCTVC-O0168 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On derivation of slice information and motion information for inter-layer reference picture in SHVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Xiaoyu Xiu, Yan Ye, Yong He, Yuwen He  9710 Scranton R-D, #250 San Diego, CA 92121 USA | Tel: Email: | 1-858-210-4830 [Xiaoyu.Xiu@InterDigital.com](mailto:Xiaoyu.Xiu@InterDigital.com)  [Yan.Ye@InterDigital.com](mailto:Yan.Ye@InterDigital.com)  [Yong.He@InterDigital.com](mailto:Yong.He@InterDigital.com)  [Yuwen.He@InterDigital.com](mailto:Yuwen.He@InterDigital.com) |
| *Source:* | InterDigital Communications, Inc. | | |

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

# Abstract

This contribution proposes to simplify the derivation of slice information and motion information for inter-layer reference picture in SHVC. Firstly, it is proposed to always associate an inter-layer reference picture with a single slice. Secondly, it is proposed that in the case when reference layer picture is coded with multiple slices and at least two slices have different slice information, the slice associated with the inter-layer reference picture is set to I-Slice and all motion blocks of the inter-layer reference picture are set to MODE\_INTRA.

# Slice and motion information derivation

In the SHVC draft 3 [1], if the reference layer is used for inter-layer motion prediction of the current layer (specified by the direct\_dependency\_flag and direct\_dependency\_type in VPS), the motion information of the corresponding reference layer picture is resampled to the resolution of the current layer. The motion information of the inter-layer reference picture is resampled in units of blocks, and may be used to predict the motion vectors of the current layer by using the inter-layer reference picture as the collocated picture for TMVP process.

According to the current TMVP design, the slice-level information of the collocated picture, such as slice type and reference picture lists, is needed to derive the scaled motion vector predictor [2]. Similarly, to resample the motion information from the reference layer, slice information associated with each block in the inter-layer reference picture should be also properly generated. If the reference layer picture is coded using multiple slices (each slice could have unique slice type and/or reference picture lists), both the slice partitions and the slice information need to be mapped from the reference layer picture to the inter-layer reference picture. In spatial scalability, the directly resampled slice partitions of the inter-layer reference picture may not be aligned with the CTB boundaries. Figure 1 shows one example of 1.5x spatial scalability, in which the up-sampled slice boundaries cut across the CTB boundaries in the current layer. In this case, the resampled slice partitions need to be re-arranged along those CTB boundaries. Such slice partitioning re-arrangement could incur additional complexity in decoding process.



Figure 1 Example of resampling slice partitions.

To avoid such slice partition adjustment, a solution was proposed in JCTVC-N0344 [3] to make the whole inter-layer reference picture associated to a single slice, that is, the first slice of the reference layer picture, and to directly copy the slice information of the first slice of the reference layer picture. In case of any two slices in the reference layer picture have different slice type and/or reference picture lists, the abovementioned slice information of the single slice for the inter-layer reference picture cannot be used for the TMVP derivation of the current layer. For example, it is assumed that the two slice partitions Slice\_1 and Slice\_2 in Figure 1 have different slice information and the block CTB\_A and the block CTB\_B in the inter-layer reference picture was originally located in Slice\_1 and Slice\_2 respectively. After the slice mapping process in [3], the slice information of Slice\_1 will be used for both CTB\_A and CTB\_B when they are utilized as the collocated block of the TMVP process. However, the combination of CTB\_B’s motion information and Slice\_1’s slice information will lead to an incorrect TMVP candidate for the motion vector prediction, resulting in a decreased coding efficiency of the current layer. Therefore, a bit-stream constraint was proposed in [3] that the inter-layer reference picture cannot be used for TMVP derivation of the current layer if any two slices in the reference picture have different slice information.

In this contribution, an alternative solution is proposed to generate slice information and motion information for inter-layer reference picture. Similar to [3], it is proposed to associate the whole inter-layer reference picture with a single slice. If there is one slice in the reference layer picture or if there are multiple slices in the reference layer picture and the slice information of all the slices are the same, the slice information of the inter-layer reference picture is directly copied from that of the first slice in the reference layer picture. However, different from [3] which set the slice information of the inter-layer reference picture equal to that of the first slice in the reference layer picture when any two reference layer slices have different slice information, it is proposed to directly set the slice type of the inter-layer reference picture to be I-Slice. Correspondingly, the motion information of all the blocks in the inter-layer reference pictures are set to MODE\_INTRA, which is equivalent to disabling inter-layer motion prediction for the current layer. In summary, the following is proposed:

1. The whole inter-layer reference picture is associated with one single slice.
2. If there is one single slice in the reference layer picture of if there are multiple slices in the reference layer picture and all the slices have the same slice type and reference picture lists, the slice type and the reference picture lists of the inter-layer reference picture are set equal to that of the first slice in the reference layer picture. And the motion information resampling process is conducted to generate the motion field of the inter-layer reference picture.
3. If there are multiple slices in the reference layer picture and any two slices have different slice type and/or reference picture lists, the slice type of the inter-layer reference picture is set to be I-Slice. And the motion information of all the blocks of the inter-layer reference picture are set equal to MODE\_INTRA, i.e., zero motion vectors and invalid reference picture index.

# Specification changes

The following changes are made based on the current SHVC working draft 3 [1]. All the changes are highlighted in yellow.

H.8.1.4 Resampling process for inter layer reference pictures

Input to this process is:

– a decoded reference layer picture rlPic

– a variable rLId specifies the layer id of reference layer picture .

Output of this process is the resampled reference layer picture rsPic.

The variables PicWidthInSamplesL and PicHeightInSamplesL are set equal to pic\_width\_in\_luma\_samples and pic\_height\_in\_luma\_samples, respectively.

The variables RefLayerPicWidthInSamplesL and RefLayerPicHeightInSamplesL are set equal to the width and height of the decoded reference layer picture rlPic in units of luma samples, respectively.

The variables PicWidthInSamplesC, PicHeightInSamplesC, RefLayerPicWidthInSamplesC, and RefLayerPicHeightInSamplesC are derived as follows:

PicWidthInSamplesC = PicWidthInSamplesL / subWidthC (H‑9)  
PicHeightInSamplesC = PicHeightInSamplesL / subHeightC (H‑10)  
RefLayerPicWidthInSamplesC = RefLayerPicWidthInSamplesL / subWidthC (H‑11)  
RefLayerPicHeightInSamplesC = RefLayerPicHeightInSamplesL / subHeightC (H‑12)

The variable currLayerId is set equal to nuh\_layer\_id of the current picture. The variable dRlIdx is set equal to DirectRefLayerIdx[ currLayerId ][ rLId ].

The variables ScaledRefLayerLeftOffset, ScaledRefLayerTopOffset, ScaledRefLayerRightOffset and ScaledRefLayerBottomOffset are derived as follows:

ScaledRefLayerLeftOffset = scaled\_ref\_layer\_left\_offset[ dRlIdx ] << 1 (H‑13)  
ScaledRefLayerTopOffset = scaled\_ref\_layer\_top\_offset[ dRlIdx] << 1 (H‑14)  
ScaledRefLayerRightOffset = scaled\_ref\_layer\_right\_offset[ dRlIdx ] << 1 (H‑15)  
ScaledRefLayerBottomOffset = scaled\_ref\_layer\_bottom\_offset[ dRlIdx ] << 1 (H‑16)

The variables ScaledRefLayerPicWidthInSamplesL and ScaledRefLayerPicHeightInSamplesL are derived as follows:

ScaledRefLayerPicWidthInSamplesL = PicWidthInSamplesL –   
 ScaledRefLayerLeftOffset  – ScaledRefLayerRightOffset (H‑17)  
ScaledRefLayerPicHeightInSamplesL = PicHeightInSamplesL –   
 ScaledRefLayerTopOffset – ScaledRefLayerBottomOffset (H‑18)

The variables ScaleFactorX and ScaleFactorY are derived as follows:

ScaleFactorX = ( ( RefLayerPicWidthInSamplesL << 16 ) + ( ScaledRefLayerPicWidthInSamplesL >> 1 ) ) / ScaledRefLayerPicWidthInSamplesL (H‑19)  
ScaleFactorY = ( ( RefLayerPicHeightInSamplesL << 16 ) + ( ScaledRefLayerPicHeightInSamplesL >> 1 ) ) / ScaledRefLayerPicHeightInSamplesL (H‑20)

The following steps are applied to derive the resampled inter layer reference picture rsPic.

– if PicWidthInSamplesL is equal to RefLayerPicWidthInSamplesL and PicHeightInSamplesL is equal to RefLayerPicHeightInSamplesL and the values of ScaledRefLayerLeftOffset, ScaledRefLayerTopOffset, ScaledRefLayerRightOffset and ScaledRefLayerBottomOffset are all equal to 0

* + rsPic is set equal to rlPic.

– otherwise, rsPic is derived as follows:

* + The PicOrderCntVal value of rsPic is set equal to the PicOrderCntVal value of rlPic.
  + The variable rsSlice is defined as the slice of rsPic and all the coding tree blocks in rsPic are associated with rsSlice. rsSlice is derived as follows:
  + If there is only one slice in rlPic, the slice\_type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of rsSlice are set equal to the values of slice\_type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of the slice of rlPic, respectively. When rsSlice is a P or B slice, for i in the range of 0 to num\_ref\_idx\_l0\_active\_minus1 inclusive, reference picture with index i in reference picture list L0 of rsSlice is set equal to reference picture with index i in reference picture list L0 of the slice of rlPic. When rsSlice is a B slice, for i in the range of 0 to num\_ref\_idx\_l1\_active\_minus1 inclusive, reference picture with index i in reference picture list L1 of rsSlice is set equal to reference picture with index i in reference picture listL1 of the slice of rlPic.
  + Otherwise if there are more one slices in rlPic and the values of slice type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of all the slices in rlPic are identical, and for i in the range of 0 to num\_ref\_idx\_l0\_active\_minus1 inclusive all the reference pictures with index i in reference picture list L0 of all the slices in rlPic are identical, and for i in the range of 0 to num\_ref\_idx\_l1\_active\_minus1 inclusive all the reference pictures with index i in reference picture list L1 of all the slices in rlPic are identical, the slice\_type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of rsSlice are set equal to the values of slice\_type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of the first slice of rlPic, respectively. When rsSlice is a P or B slice, for i in the range of 0 to num\_ref\_idx\_l0\_active\_minus1 inclusive, reference picture with index i in reference picture list L0 of rsSlice is set equal to reference picture with index i in reference picture list L0 of the first slice of rlPic. When rsSlice is a B slice, for i in the range of 0 to num\_ref\_idx\_l1\_active\_minus1 inclusive, reference picture with index i in reference picture list L1 of rsSlice is set equal to reference picture with index i in reference picture list L1 of the first slice of rlPic.
  + Otherwise, the slice type, num\_ref\_idx\_l0\_active\_minus1 and num\_ref\_idx\_l1\_active\_minus1 of rsSlice are set equal to I, -1 and -1 respectively.
  + When SamplePredEnabledFlag[ currLayerId ][ rLId ] is equal to 1, the picture sample resampling process as specified in subclause H.8.1.4.1 is invoked with the sample arrays of reference layer picture rlPic as input, and with the sample arrays of resampled picture rsPic as output.
  + When MotionPredEnabledFlag[ currLayerId ][ rLId ] is equal to 1, the picture motion field resampling process as specified in subclause H.8.1.4.2 is invoked with refereence layer picture rlPic and its motion field as inputs, and with the motion field of resampled picture rsPic as output.

H.8.1.4.2 Resampling process of picture motion field

Inputs to this process are:

– the decoded reference layer picture rlPic,

– the variable rlPicMotion specifying the motion field of the reference layer picture rlPic,

Output of this process is rsPicMotion specifying the motion field of the resampled picture rsPic.

The motion field of rlPic specified by rlPicMotion consists of:

– a ( RefLayerPicWidthInSamplesL ) x ( RefLayerPicHeightInSamplesL ) array predModeRL specifies the prediction modes of the reference layer picture rlPic,

– two ( RefLayerPicWidthInSamplesL ) x ( RefLayerPicHeightInSamplesL ) arrays refIdxLXRL specify the reference indices of the reference layer picture rlPic, with X = 0,1,

– two ( RefLayerPicWidthInSamplesL ) x ( RefLayerPicHeightInSamplesL ) arrays mvLXRL specify the luma motion vectors of the reference layer picture rlPic, with X = 0,1,

– two ( RefLayerPicWidthInSamplesL ) x ( RefLayerPicHeightInSamplesL ) arrays predFlagLXRL specify the prediction list utilization flags of the reference layer picture rlPic, with X = 0,1.

The resampled motion field specified by rsPicMotion consists of:

– a ( PicWidthInSamplesL ) x ( PicHeightInSamplesL ) array predMode specifies the prediction modes of the resampled picture,

– two ( PicWidthInSamplesL ) x ( PicHeightInSamplesL ) arrays refIdxLX specify the reference indexes of the resampled picture, with X = 0,1,

– two ( PicWidthInSamplesL ) x ( PicHeightInSamplesL ) arrays mvLX  specify the luma motion vectors of the resampled picture, with X = 0,1,

* two ( PicWidthInSamplesL ) x ( PicHeightInSamplesL ) arrays predFlagLX specify the prediction list utilization flags of the resampled picture, with X = 0,1.

For each luma sample location xPb = 0 ... ( ( PicWidthInSamplesL + 15 ) >> 4 ) − 1 and yPb = 0 … ( ( PicHeightInSamplesL + 15 ) >> 4) − 1,

– The variables xP and yP are set to ( xPb  << 4 ) and ( yPb  << 4 ), respectively,

* The variables predMode[xP][yP], refIdxLX[xP][yP], mvLX[xP][yP] and predFlagLX[xP][yP], with X = 0,1, of the resampled picture are derived by invoking inter layer motion parameters derivation process specified in subclause H.8.1.4.2.1 with the luma location ( xP, yP ), rlPic , predModeRL, refIdxLXRL, mvLXRL and predFlagLXRL, with X = 0,1, and rsSlice given as input.

H.8.1.4.2.1 Derivation process for inter layer motion parameters

Inputs to this process are

– a luma location ( xP, yP ) specifying the top-left sample of the current luma prediction block relative to the top-left luma sample of the resampled picture,

– the decoded reference layer picture rlPic

– the reference layer prediction mode array predModeRL,

– the reference layer reference index arrays refIdxL0RL and refIdxL1RL

– the reference layer motion vector arrays mvL0RL and mvL1RL

– the reference layer prediction list utilization flag arrays predFlagL0RL and predFlagL1RL.

* the slice of the resampled picture rsSlice.

Outputs of this process are

– a derived prediction mode predMode,

– two derived motion vectors mvL0 and mvL1

– two derived reference indices refIdxL0 and refIdxL1

– two derived prediction list utilization flags predFlagL0 and predFlagL1.

The variables predMode, mvLX, refIdxLX, refPicOrderCntLX, and predFlagLX are derived as follows.

1. The center location (xPCtr, yPCtr) of the luma prediction block is derived as follows

xPCtr = xP + 8 (H‑44)

yPCtr = yP + 8 (H‑45)

1. The derivation process for reference layer luma sample location specified in subclause H.6.1 is invoked with luma location ( xPCtr, yPCtr ) given as the inputs and ( xRef, yRef ) as output.
2. The collocated position (xRL, yRL) is derived as follows:

xRL = ( ( xRef + 4 ) >> 4 ) << 4 (H‑46)

yRL = ( ( yRef + 4 ) >> 4 ) << 4 (H‑47)

1. The prediction mode predMode[ xP ][ yP ] is derived as follows:

* If the slice type of rsSlice is I, predMode[ xP ][ yP ] is set to MODE\_INTRA.
* Otherwise i~~I~~f ( xRL < 0 ) or ( xRL >= RefLayerPicWidthInSamplesL ) or ( yRL < 0 ) or ( yRL >= RefLayerPicHeightInSamplesL ), predMode[ xP ][ yP ] is set to MODE\_INTRA.
* Otherwise, the following applies:

predMode[ xP ][ yP ] = predModeRL[ xRL ][ yRL ] (H‑48)

1. The variables mvL0[ xP ][ yP ], mvL1[ xP ][ yP ], refIdxL0[ xP ][ yP ], refIdxL1[ xP ][ yP ], predFlagL0[ xP ][ yP ] and predFlagL1[ xP ][ yP ] are derived as follows:

* If predMode[ xP ][ yP ] is equal to MODE\_INTER, the following applies
* The variable colPb specifies the luma prediction block covering the location given by ( xRL, yRL ) inside the reference layer picture specified by rlPic. The 16 x 16 coding block with top-left luma location ( xP, yP ) relative to the top-left luma sample of the resampled picture is set associated with the slice containing the prediction block colPb inside the reference layer picture specified by rlPic.
* For each X = 0, 1, the following applies:
  + - * + The variables refIdxLX[ xP ][ yP ] and predFlagLX[ xP ][ yP ] are derived as follows:

refIdxLX[ xP ][ yP ] = refIdxLXRL[ xRL ][ yRL ] (H‑49)

predFlagLX[ xP ][ yP ] = predFlagLXRL[ xRL ][ yRL ] (H‑50)

* + - * + The variable mvLX[ xP ][ yP ][ 0 ] is derived as follows:
        + If ScaledRefLayerPicWidthInSamplesL is not equal to RefLayerPicWidthInSamplesL, mvLX[ xP ][ yP ][ 0 ] is derived as follows:

scaleFactorMVX = Clip3( −4096, 4095, ( ( ScaledRefLayerPicWidthInSamplesL << 8 ) + ( RefLayerPicWidthInSamplesL >> 1 ) ) / RefLayerPicWidthInSamplesL) (H‑51)

mvLX[ xP ][ yP ][0] = Clip3( −32768, 32767, Sign(scaleFactorMVX \*   
mvLXRL[ xRL ][ yRL ][ 0 ] ) \*  ( ( Abs ( scaleFactorMVX \* mvLXRL[ xRL ][ yRL ][ 0 ] )  
 + 127 ) >> 8 ) ) (H‑52)

* + - * + Otherwise, the following applies:

mvLX[ xP ][ yP ][ 0 ] = mvLXRL[ xRL ][ yRL ][ 0 ] (H‑53)

* + - * + The variable mvLX[ xP ][ yP ][ 1 ] is derived as follows:
        + If ScaledRefLayerPicHeightInSamplesL is not equal to RefLayerPicHeightInSamplesL, mvLX[ xP ][ yP ][ 1 ] is derived as follows:

scaleFactorMVY = Clip3( −4096, 4095, ( ( ScaledRefLayerPicHeightInSamplesL << 8 ) + ( RefLayerPicHeightInSamplesL >> 1 ) ) / RefLayerPicHeightInSamplesL) (H‑54)

mvLX[ xP ][ yP ][ 1 ] = Clip3( −32768, 32767, Sign(scaleFactorMVY \*   
mvLXRL[ xRL ][ yRL ][ 1 ] ) \*  ( ( Abs  ( scaleFactorMVY \* mvLXRL[ xRL ][ yRL ][ 1 ] )  
 + 127 ) >> 8 ) ) (H‑55)

* + - * + Otherwise, the following applies:

mvLX[ xP ][ yP ][ 1 ] = mvLXRL[ xRL ][ yRL ][ 1 ] (H‑56)

* Otherwise (predMode[ xP ][ yP ] is equal to MODE\_INTRA), the following applies:
* both components of mvL0[ xP ][ yP ] and mvL1[ xP ][ yP ] are set to 0, refIdxL0[ xP ][ yP ] and refIdxL1[ xP ][ yP ] are set to –1, predFlagL0[ xP ][ yP ] and predFlagL1[ xP ][ yP ] are set to 0.

# 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).**

# References

1. J. Chen, J. Boyce, Y. Ye and M. M. Hannuksela, “SHVC Working Draft 3”, JCTVC document JCTVC-N1008, Vienna, Austria, August 2013.
2. B. Bross, W-J. Han, J-R. Ohm, G. J. Sullivan, “High Efficiency Video Coding (HEVC) Text Specification Draft 10”, JCTVC document JCTVC-L1003, Geneva, Switzerland, January, 2013.
3. J. Chen, V. Seregin, X. Li, K. Rapaka and M. Karczewicz, “Derivation of picture and slice information for resampled interlayer reference picture”, JCTVC document JCTVC-N0344, Vienna, Austria, August 2013.