|  |  |
| --- | --- |
| **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-K0238 (m26574) |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | Inter-layer intra prediction mode coding for the scalable extension of HEVC | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Zhijie Zhao, Junyong Si, Joern Ostermann Appelstr. 9A Institut für Informationsverarbeitung(TNT) 30167 Hannover Germany | Tel: Email: | +49-511-7625316 {zhao,si,ostermann}@tnt.uni-hannover.de |
| *Source:* | Leibniz Universitaet Hannover | | |

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

# Abstract

This contribution presents two approaches for the inter-layer intra coding of the scalable extension of HEVC. In the scalable extension of H.264/AVC, inter-layer intra prediction (ILIP) is used to reduce the redundancy between two spatial layers of intra pictures. Except the blocks coded using ILIP, other blocks are coded as a single layer coding. In this contribution, two approaches are proposed to utilize the base layer intra prediction mode to improve the coding efficiency of the CUs not being coded by ILIP in the context of the scalable HEVC coding. The first approach directly utilizes the intra prediction mode from the corresponding base layer PU as the intra prediction mode of the enhancement layer PU, while the second approach uses the intra prediction mode of the corresponding base layer PU as an additional MPM candidate. ILIP and the two proposed approaches are implemented based on HM 6.1. The implemented scheme achieves an average luma BD-rate reduction of 31.9% for class B sequences relative to simulcast high resolution reference used by Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding in dyadic all intra spatial scalability.

# Introduction

Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (CfP) issued jointly by ITU-T SG 16 WP 3 (VCEG) and ISO/IEC JTC1/SC29/WG11 (MPEG) [1] in July 2012. In this contribution, two approaches are proposed to respond to the Category 1 intra-only spatial scalability. This proposal is based on the HM software version 6.1 with the addition of supporting spatial scalable coding, and two tools inter-layer intra prediction mode coding (ILIPM) and base layer based MPM coding (BL-MPM) are implemented to improve the coding efficiency of inter-layer coding for intra coded block.

In HM 6.1, the PU size can be 64x64, 32x32, 16x16, 8x8 and 4x4. There are 35 Intra modes, among which 33 are directional prediction modes and the other two are Planar and DC modes. The intra prediction modes of all PU sizes are coded using 3MPM (Most Probable Mode) scheme [2] [3]. For the current PU, the prediction mode of the top neighbor PU and the prediction mode of the left neighbor PU are first selected as the two most probable modes. Another mode is added as the third most probable mode depending on the modes of the left and above neighbor PUs. If one of the left or above neighbor PU is not available, DC mode is added as a candidate for the most probable mode. If the prediction mode of the current PU is the same as one of the most probable modes, a shorter code word at maximal 3 bits is used to encode the current mode. Otherwise, the remaining mode is coded using a long code word at 5 bits.

For the enhancement layer coding, in case one CU is coded using IPIL, no intra prediction modes need to be coded and the MPM coding does not happen. Although IPIL is very efficient, there are still some CUs coded by single layer intra coding instead of IPIL. In case one CU is not coded by IPIL, RD optimization process compares the RD cost of the available intra prediction modes for a PU and the one which demonstrates the best RD performance is chosen as the prediction mode of the current PU. Then MPM process is used to code the intra prediction mode of the current PU. Since the co-located pictures of the base layer and enhancement layer share the similar structure and texture, it may be true that the prediction mode of the current PU in the enhancement layer uses the same intra prediction mode as the corresponding PU in the base layer. Therefore, we proposed two approaches utilizing the intra prediction mode of base layer to improve the coding efficiency of CUs which are not coded by IPIL in the enhancement layer.

# Up-sampling filters

In [4], DCT interpolation filter [5] is used for 2x up-sampling of the reconstructed base layer picture, using 8-tap for luma and 4-tap for chroma. The same filter proposed in [4] is used in the inter-layer intra prediction (ILIP) in our scalable implementation. The luma filter coefficients are shown in , and the chroma filter coefficients are shown in .

Table 1 8-tap DCTIF filter coefficients for luma

|  |  |
| --- | --- |
| Position | Filter coefficients |
| 1/12 | { -1, 1, -4, 63, 6, -2, 1, 0 } |
| 2/12 | { -1, 3, -7, 61, 12, -5, 2, -1 } |
| 3/12 | { -1, 4, -10, 57, 19, -7, 3, -1 } |
| 4/12 | { -1, 4, -11, 52, 26, -8, 3, -1 } |
| 5/12 | { -1, 5, -12, 47, 32, -9, 3, -1 } |
| 6/12 | { -1, 4, -11, 40, 40, -11, 4, -1 } |

Table 2 8-tap DCTIF filter coefficients for chroma

|  |  |
| --- | --- |
| Position | Filter coefficients |
| 1/12 | { -2, 62, 5, -1 } |
| 2/12 | { -4, 59, 11, -2 } |
| 3/12 | { -4, 54, 16, -2 } |
| 4/12 | { -5, 50, 22, -3 } |
| 5/12 | { -5, 43, 30, -4 } |
| 6/12 | { -4, 36, 36, -4 } |

The same idea as the Inter-layer intra prediction (ILIP) proposed in the scalable extension of H.264/AVC [6] is used as the ILIP for the scalable extension of HEVC for intra coded unit. However, DCTIF filter instead of the filter used in JSVM is used as the up-sampling filter to do 2x up-sampling for the reconstructed base layer picture.

After up-sampling process, the difference in pixel domain between the CU in the enhancement layer and the up-sampled corresponding CU in the base layer are transform and entropy coded. A new prediction mode, base layer mode, is added to the CU level to indicate whether one CU is coded by ILIP. The signaling of the enhancement layer’s CU, PU, and TU sizes in the base and enhancement layer is coded independently. In case one CU in the enhancement layer is coded by ILIP, no intra prediction mode and partition information (always 2Nx2N) need to be coded in the enhancement layer. Transform partitions are still coded in the ILIP coded PU.

# Proposed methods

## Inter-layer intra prediction mode coding

In this approach, a new mode, called inter-layer intra prediction mode (ILIPM), is added as an extra intra prediction mode for an intra coded CU in the enhancement layer. Figure 1 shows the relationship between the corresponding base layer PU and enhancement layer PU. Assume that the intra prediction mode of the current PU in the enhancement layer is, while the intra prediction mode of the corresponding PU in the base layer is . The CU depth level of is , and the CU depth level of is . The partition size of and are and , respectively. If the partition size is 2Nx2N, . If the partition size is NxN, then ILIPM is only used as a candidate mode if the following condition is true:

(1)

Figure Relationship between the current PU in the enhancement layer and the corresponding PU in the base layer

Enhancement layer

Base layer

If Eq. 1 is true, the RD optimization process adds IPILM as an extra intra prediction mode. In the IPILM mode, the current prediction unit utilizes the intra prediction mode of the corresponding prediction unit in the base layer as its intra prediction mode. At the same time, a flag, **intra\_bl\_mode\_flag**, is introduced before **prev\_intra\_luma\_pred\_flag** in the coding unit syntax. Intra\_bl\_mode\_flag is only coded in the coding unit syntax if Eq. 1 is true. If intra\_bl\_mode\_flag is true, prev\_intra\_luma\_pred\_flag and possible rem\_intra\_luma\_pred\_mode are not coded any more. If Eq. 1 is true and IPILM is not chosen as the intra prediction mode for the current prediction unit , intra\_bl\_mode\_flag is set to false and MPM coding as in the single layer coding is performed.

## Base layer based MPM coding

As mentioned in section , MPM process is used to code the intra prediction mode in HM. In single layer coding, the top neighbor CU, the left neighbor PU and another mode which depends on the availability of top and left neighbor CUs and theirs intra prediction modes is selected as MPM modes.

In the enhancement layer coding, we propose to use the intra prediction mode of the corresponding PU in the base layer as a new additional MPM candidate, called base layer based MPM coding (BL-MPM). A flag, **bl\_mode\_mpm\_flag**, is introduced to indicate whether the intra prediction mode of the corresponding PU in the base layer is used as a candidate MPM mode. The bl\_mode\_mpm\_flag is coded in the coding unit syntax only if **prev\_intra\_luma\_pred\_flag** is false. As shown in , if the encoder cannot find a matched prediction mode in the list of 3MPM candidates for the current , then prev\_intra\_luma\_pred\_flag is set to flase and the intra prediction mode of is used as a new MPM candidate. If the intra prediction mode of is equal to the intra prediction mode of , then bl\_mode\_mpm\_flag is set to true and rem\_intra\_luma\_pred\_mode needs not be coded any more. If the intra prediction mode of is not equal to the intra prediction mode of, Bl\_mode\_mpm\_flag is set to false and rem\_intra\_luma\_pred\_mode is coded as the single layer HM coding.

# Syntax and semantics description

The proposed changes to the syntax of the enhancement layer coding unit are highlighted in .

**intra\_bl\_mode\_flag**[ x0 + i ][ y0+ j ], **prev\_intra\_luma\_pred\_flag**[ x0 + i ][ y0 + j ], **mpm\_idx[** x0 + i **][** y0 + j **], bl\_mode\_mpm\_flag**[ x0 + i ][ y0+ j ] and **rem\_intra\_luma\_pred\_mode**[ x0 + i ][ y0 + j ]specify the intra prediction mode for luma samples. The array indices x0 + i, y0 + j specify the location ( x0 + i, y0 + j ) of the top-left luma sample of the considered prediction block relative to the top-left luma sample of the picture. When intra\_bl\_mode\_flag[ x0 + i ][ y0 + j ] is available and is equal to 1, the intra prediction mode is inferred from the corresponding predcition unit in the base layer. Otherweise, when prev\_intra\_luma\_pred\_flag[ x0 + i ][ y0 + j ] is equal to 1, the intra prediction mode is inferred from a neighbouring intra-predicted prediction unit. When prev\_intra\_luma\_pred\_flag[ x0 + i ][ y0 + j ] is equal to 0, bl\_mode\_mpm\_flag[ x0 + i ][ y0 + j ] is available. If bl\_mode\_mpm\_flag[ x0 + i ][ y0 + j ] is equal to 1, the intra prediction mode is inferred from the corresponding prediction unit in the base layer.

# Compression performance discussion

The coding configuration used for the base layer is identical to the anchors for the all intra case. The important configuration parameters for the enhancement layer are shown in .

Table Configuration parameters for the enhancement layer

|  |  |
| --- | --- |
|  | Value |
| ALF | 1 |
| SAO | 1 |
| NSQT | 0 |
| LMChroma | 1 |
| InternalBitDepth | 8 |
| ALFMaxNumFilter | 128 |
| SAOInterleaving | 1 |

The BD-rate performances can be found in the attached excel sheet.

# Conclusions

In this contribution, ILIP is implemented for the scalable extension of HEVC. Two additional approaches are proposed as complementary methods of ILIP to improve the coding efficiency of inter-layer intra coding. We propose to evaluate our intra prediction method in the core experiments of the scalable extension of HEVC and to consider ILIPM and BL-MPM as valuable tools in the new standard.

Table 3 Syntax

|  |
| --- |
| coding\_unit( x0, y0, log2CbSize ) { |
| CurrCbAddrTS = MinCbAddrZS[ x0 >> Log2MinCbSize ][ y0 >> Log2MinCbSize ] |
| if( transquant\_bypass\_enable\_flag ) { |
| **cu\_transquant\_bypass\_flag** |
| } |
| if( slice\_type != I ) |
| **skip\_flag[** x0 **][** y0 **]** |
| if( skip\_flag[ x0 ][ y0 ] ) |
| prediction\_unit( x0, y0 , log2CbSize ) |
| else { |
| if( slice\_type != I ) |
| **pred\_mode\_flag** |
| if( PredMode != MODE\_INTRA | | log2CbSize = = Log2MinCbSize ) |
| **part\_mode** |
| x1 = x0 + ( ( 1 << log2CbSize ) >> 1 ) |
| y1 = y0 + ( ( 1 << log2CbSize ) >> 1 ) |
| x2 = x1 − ( ( 1 << log2CbSize ) >> 2 ) |
| y2 = y1 − ( ( 1 << log2CbSize ) >> 2 ) |
| x3 = x1 + ( ( 1 << log2CbSize ) >> 2 ) |
| y3 = y1 + ( ( 1 << log2CbSize ) >> 2 ) |
| if( PredMode = = MODE\_INTRA ) { |
| if( PartMode = = PART\_2Nx2N && pcm\_enabled\_flag &&  log2CbSize >= Log2MinIPCMCUSize &&  log2CbSize <= Log2MaxIPCMCUSize ) |
| **pcm\_flag** |
| if( pcm\_flag ) { |
| **num\_subsequent\_pcm** |
| NumPCMBlock = num\_subsequent\_pcm + 1 |
| while( !byte\_aligned( ) ) |
| **pcm\_alignment\_zero\_bit** |
| for( i = 0; i < 1 << ( log2CbSize << 1 ); i++ ) |
| **pcm\_sample\_luma**[ i ] |
| for( i = 0; i < ( 1 << ( log2CbSize << 1 ) ) >> 1; i++ ) |
| **pcm\_sample\_chroma**[ i ] |
| NumPCMBlock− − |
| } else { |
| pbOffset = ( PartMode = = PART\_NxN ) ? ( ( 1 << log2CbSize ) >> 2 ) : 0 |
| for( j = 0; j <= pbOffset; j = j + pbOffset ) |
| for( i = 0; i <= pbOffset; i = i + pbOffset ) { |
| if( ( 1 << log2CbSize\_enh ) >> trafoDepth = = (( 1 << log2CbSize\_base ) >> trafoDepth ) << 1) |
| **intra\_bl\_mode\_flag**[ x0 + i ][ y0+ j ] |
| if(!intra\_bl\_mode\_flag[ x0 + i ][ y0+ j ] || log2CbSize\_enh + trafoDepth\_enh -  1 != log2CbSize\_base +trafoDepth\_base ) |
| **prev\_intra\_luma\_pred\_flag**[ x0 + i ][ y0+ j ] |
| } |
| for( j = 0; j <= pbOffset; j = j + pbOffset ) |
| for( i = 0; i <= pbOffset; i = i + pbOffset ) { |
| if(!intra\_bl\_mode\_flag[ x0 + i ][ y0+ j ] || log2CbSize\_enh + trafoDepth\_enh -  1 != log2CbSize\_base +trafoDepth\_base ){ |
| if( prev\_intra\_luma\_pred\_flag[ x0 + i ][ y0+ j ] ) |
| **mpm\_idx**[ x0 + i ][ y0+ j ] |
| Else{ |
| **bl\_mode\_mpm\_flag**[ x0 + i ][ y0+ j ] |
| If(!bl\_mode\_mpm\_flag[ x0 + i ][ y0+ j ]) |
| **rem\_intra\_luma\_pred\_mode**[ x0 + i ][ y0+ j ] |
| } |
| } |
| } |
| **intra\_chroma\_pred\_mode**[ x0 ][ y0 ] |
| } |
| } else { |
| if( PartMode = = PART\_2Nx2N ) |
| prediction\_unit( x0, y0 , log2CbSize ) |
| else if( PartMode = = PART\_2NxN ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x0, y1 , log2CbSize ) |
| } else if( PartMode = = PART\_Nx2N ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x1, y0 , log2CbSize ) |
| } else if( PartMode = = PART\_2NxnU ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x0, y2 , log2CbSize ) |
| } else if( PartMode = = PART\_2NxnD ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x0, y3 , log2CbSize ) |
| } else if( PartMode = = PART\_nLx2N ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x2, y0 , log2CbSize ) |
| } else if( PartMode = = PART\_nRx2N ) { |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x3, y0 , log2CbSize ) |
| } else { /\* PART\_NxN \*/ |
| prediction\_unit( x0, y0 , log2CbSize ) |
| prediction\_unit( x1, y0 , log2CbSize ) |
| prediction\_unit( x0, y1 , log2CbSize ) |
| prediction\_unit( x1, y1 , log2CbSize ) |
| } |
| } |
| if( !pcm\_flag ) { |
| if( PredMode != MODE\_INTRA &&   !(PartMode = = PART\_2Nx2N && merge\_flag[x0][y0]) ) |
| **no\_residual\_data\_flag** |
| if( !no\_residual\_data\_flag ) { |
| MaxTrafoDepth = ( PredMode = = MODE\_INTRA ?   max\_transform\_hierarchy\_depth\_intra + IntraSplitFlag :   max\_transform\_hierarchy\_depth\_inter ) |
| transform\_tree( x0, y0, x0, y0, x0, y0, log2CbSize, log2CbSize, log2CbSize, 0, 0 ) |
| } |
| } |
| } |
| } |

# Bibliography

|  |  |
| --- | --- |
| [1] | *ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG 16 WP 3, ISO/IEC JTC1/SC29/WG11 N12957,* Stockholm, Sweden, July 2012. |
| [2] | E. Maani and A. Tabatabai, *Intra mode coding using logical mode numbering, JCTVC-H0407,* San Jose, 2011. |
| [3] | X. Zhang, S. Liu, M. Guo, X. Guo and S. Lei, *Non-CE6: Intra mode coding with fixed length binarization, JCTVC-H0435,* San Jose, 2011. |
| [4] | J. Boyce, D. Hong, W. Jang and A. Abbas, *Information for HEVC scalability extension, JCTVC-G078,* Geneva, 2011. |
| [5] | K. McCann, W.-J. Han, I.-K. Kim and e. al., *Video coding technology proposal by Samsung (and BBC), JCTVC-A124,* Dresden, 2010. |
| [6] | *ITU-T Recommendation H.264, “Advanced Video Coding for Generic Audiovisual Services,” March, 2010..* |
|  |  |
|  |  |

# Patent rights declaration(s)

**Leibniz Universität Hannover 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).**