|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  13th Meeting: Incheon, KR, 18–26 Apr. 2013 | Document: JCTVC-M0465 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | Signaling of Phase Offset in Up-sampling Process and Chroma Sampling Location | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Liwei Guo Jianle Chen, and Marta Karczewicz  J. Dong, Y. Ye, and Y. He,  K. Ugur and J. Lainema | Tel: Email: | [liweig@qti.qualcomm.com](mailto:liweig@qti.qualcomm.com)  [jie.dong@interdigital.com](mailto:jie.dong@interdigital.com)  [kemal.ugur@nokia.com](mailto:kemal.ugur@nokia.com) |
| *Source:* | Qualcomm Incorporated, InterDigital Communications, Inc., Nokia | | |

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

# Abstract

In spatial scalability coding, reference layer pictures are down-sampled versions of current layer pictures. The down-sampling locations are not normative part in the standard and can have different phase shifts. To avoid the mismatch between the down-sampling phase and the up-sampling phase, this contribution proposes to have indication of phase shift for the up-sampling process. In addition, this contribution proposes to signal chroma sampling locations and define the up-sampling filters for all 16 phases.

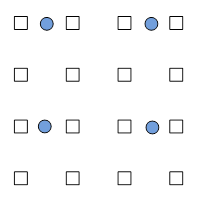
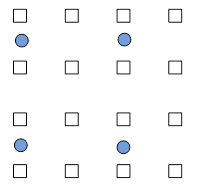
# Introduction

In scalable video coding, base layer (i.e., reference layer) pictures are down-sampled versions of enhancement layer pictures. As a result, up-sampling need to be performed to base layer pictures for inter-layer texture prediction. While the up-sampling process (including the filter coefficients and the up-sampling positions) is defined in the standard, the down-sampling process is left as non-normative. (Figure 1 shows two possible down-sampling phase shift). If the phases (location) used in the up-sampling and the down-sampling have mismatch, coding efficiency will be affected [1] [2] [3].

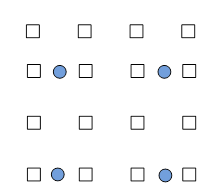
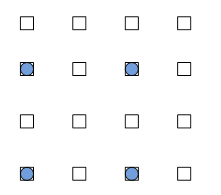
 

Figure Down-sampling location. Left: zero-phase; right: symmetric

For 420 vieo format, H.264 and HEVC allows 6 different chroma sampling locations (as shown in Figure.2). In current SHVC position mapping, the chroma sampling location (b) is assumed.



1. (b) (c)



(d) (e) (f)

Figure 2 Color sampling location

# Proposed

In this contribution, it is proposed to signal sampling grid shift information so an up-sampling filter with correct phase can be used. This requires the definition of up-sampling filter at all 16 phases. It is also proposed to signal chroma sampling locations in SPS.

|  |  |
| --- | --- |
| sps\_extension ( ) { | Descriptor |
|  |  |
| sampling\_grid\_information() |  |
| **}** |  |

|  |  |
| --- | --- |
| sampling\_grid\_information( ) { | Descriptor |
| **phase\_offset\_present\_flag** | u(1) |
| if ( sampling\_grid\_info\_present\_flag ) { |  |
| **horizontal\_phase\_offset16** | ue(v) |
| **vertical\_phase\_offset16** | ue(v) |
| **chroma\_phase\_x\_flag** | u(1) |
| **chroma\_phase\_y** | u(2) |
| } |  |

**phase\_offset\_present\_flag** equal to 1 specifies that the syntax elements horizontal\_phase\_offset16 and vertical\_phase\_offset16 are present in the bitstream.

**horizontal\_phase\_offset16** specifies the horizontal phase offset of the samples in the current layer with respect to lower layer in 1/16-th pixel units and it is used to calculate the reference layer sample locations used in reseampling. The value of horizontal\_phase\_offset16 should be in the range 0 to 7 inclusive. When horizontal\_phase\_offset16 is not present, the value of horizontal\_phase\_offset16 is inferred to be zero.

**vertical\_phase\_offset16** specifies the vertical phase offset of the samples in the current layer with respect to lower layer in 1/16-th pixel units and it is used to calculate the reference layer sample locations used in reseampling. The value of vertical\_phase\_offset16 should be in the range 0 to 7 inclusive. When vertical\_phase\_offset16 is not present, the value of vertical\_phase\_offset16 is inferred to be zero.

**chroma\_phase\_x \_flag** specifies the horizontal phase shift of the chroma components in units of half luma samples of a picture or layer picture in the CVS. When chroma\_phase\_x\_flag is not present, it shall be inferred to be equal to 0. The phase shift refers to the spatial displacement between the top left chroma sample and the top left luma sample.

**chroma\_phase\_y** specifies the vertical phase shift of the chroma components in units of half luma samples of a picture or layer picture in the CVS. When chroma\_phase\_y is not present, it shall be inferred to be equal to 0. The value of chroma\_phase\_y shall be in the range of 0 to 2, inclusive. The phase shift refers to the spatial displacement between the top left chroma sample and the top left luma sample.

The position calculation is as follows:

1. The variables phaseXC and phaseYC are derived by

phaseXC = chroma\_phase\_x\_flag   
phaseYC = chroma\_phase\_y

1. The variable xRef16 is derived as follows:

If cIdx is equal to 0, the variable xRef16 is derived as follow:

iShiftX = 16;

iAddX = 1 << ( iShiftX - 5 );

iShiftXM4 = iShiftX - 4;

iScaleX = ( ( PicWRL << iShiftX ) + ( ScaledW >> 1 ) ) / ScaledW;

xRef16 = ( (xP\*iScaleX + iAddX) >> iShiftXM4 )-horizontal\_phase\_offset16;

– Otherwise, the variable xRef16 is derived as follow:

iShiftX = 16;

iAddX = ( ( (PicWRL\* phaseXC) << (iShiftX - 2) ) +(ScaledW >> 1) ) / ScaledW+ (1 << ( iShiftX - 5 ));

iDeltaX= 4\* phaseXC;

iShiftXM4 = iShiftX - 4;

iScaleX = ( ( PicWRL << iShiftX ) + ( ScaledW >> 1 ) ) / ScaledW;

xRef16 = ( (xP\*iScaleX + iAddX) >> iShiftXM4 ) – iDeltaX - (horizontal\_phase\_offset16>>1);

1. The variable yRef16 is derived as follows:

– If cIdx is equal to 0, the variable yRef16 is derived as follow:

iShiftY = 16;

iAddY = 1 << ( iShiftY - 5 );

iShiftYM4 = iShiftY - 4;

iScaleY = ( ( PicHRL << iShiftY ) + ( ScaledH >> 1 ) ) / ScaledH;

yRef16  = ( (yP\*iScaleY + iAddY) >> iShiftYM4) - vertical\_phase\_offset16;

– Otherwise, the variable yRef16 is derived as follow:

iShiftY = 16;

iAddY = ( ( (PicHRL\* phaseYC) << (iShiftY - 2) ) +(ScaledH >> 1) ) / ScaledH+ (1 << ( iShiftY - 5 ));

iDeltaY= 4\* phaseYC;

iShiftYM4 = iShiftY - 4;

iScaleY = ( ( PicHRL << iShiftY ) + ( ScaledH >> 1 ) ) / ScaledH;

yRef16 = ((yP\*iScaleY + iAddY) >> iShiftYM4) – iDeltaY - (vertical\_phase\_offset16>>1)

Table Coefficients of Luma Up-sampling filter

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| phase p | interpolation filter coefficients | | | | | | | |
| fC[ p, 0 ] | fC[ p, 1 ] | fC[ p, 2 ] | fC[ p, 3 ] | fC[ p, 4 ] | fC[ p, 5 ] | fC[ p, 6 ] | fC[ p, 7 ] |
| 0 | 0 | 0 | 0 | 64 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | -3 | 63 | 4 | -2 | 1 | 0 |
| 2 | 0 | 2 | -6 | 61 | 9 | -3 | 1 | 0 |
| 3 | -1 | 3 | -8 | 60 | 13 | -4 | 1 | 0 |
| 4 | -1 | 4 | -10 | 58 | 17 | -5 | 1 | 0 |
| 5 | -1 | 4 | -11 | 52 | 26 | -8 | 3 | -1 |
| 6 | -1 | 4 | -11 | 50 | 29 | -9 | 3 | -1 |
| 7 | -1 | 4 | -11 | 45 | 34 | -10 | 4 | -1 |
| 8 | -1 | 4 | -11 | 40 | 40 | -11 | 4 | -1 |
| 9 | -1 | 4 | -10 | 34 | 45 | -11 | 4 | -1 |
| 10 | -1 | 3 | -9 | 29 | 50 | -11 | 4 | -1 |
| 11 | -1 | 3 | -8 | 26 | 52 | -11 | 4 | -1 |
| 12 | 0 | 1 | -5 | 17 | 58 | -10 | 4 | -1 |
| 13 | 0 | 1 | -4 | 13 | 60 | -8 | 3 | -1 |
| 14 | 0 | 1 | -3 | 9 | 61 | -6 | 2 | 0 |
| 15 | 0 | 1 | -2 | 4 | 63 | -3 | 1 | 0 |

Table Coefficients of Chroma up-sampling filter

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| phase p | interpolation filter coefficients | | | |
| fC[ p, 0 ] | fC[ p, 1 ] | fC[ p, 2 ] | fC[ p, 3 ] |
| 0 | 0 | 64 | 0 | 0 |
| 1 | -2 | 62 | 4 | 0 |
| 2 | -2 | 58 | 10 | -2 |
| 3 | -4 | 56 | 14 | -2 |
| 4 | −4 | 54 | 16 | −2 |
| 5 | −6 | 52 | 20 | −2 |
| 6 | −6 | 46 | 28 | −4 |
| 7 | -4 | 42 | 30 | -4 |
| 8 | -4 | 36 | 36 | -4 |
| 9 | -4 | 30 | 42 | -4 |
| 10 | -4 | 28 | 46 | -6 |
| 11 | -2 | 20 | 52 | -6 |
| 12 | -2 | 16 | 54 | -4 |
| 13 | -2 | 14 | 56 | -4 |
| 14 | -2 | 10 | 58 | -2 |
| 15 | 0 | 4 | 62 | -2 |

# Experimental results

Sequences with two different down-sampled base layer versions are used in the experiment: sequences with zero-phase down-sampling (i.e. sequences in SHVC CTC) and sequences with symmetric down-sampling [4]. Other simulation settings follow common test conditions [5].

In the first experiment, base layer sequences are generated with zero-phase down-sampling. The anchor is with the symmetric up-sampling. With the help of the signaling of phase offset, the proposed method can select the matched up-sampling phase and achieve 5.7%-10.4% Y BD-rate reduction.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **AI HEVC 2x** | | | **AI HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -15.2% | -16.5% | -17.2% |  |  |  |
| Class B | -8.5% | -9.2% | -9.0% | -9.6% | -11.0% | -11.1% |
| **Overall (Test vs Ref)** | -10.4% | -11.3% | -11.4% | -9.6% | -11.0% | -11.1% |
| **Overall (Test vs single layer)** | 12.4% | 13.7% | 13.4% | 10.3% | 10.3% | 9.7% |
| **EL only (Test vs Ref)** | -18.3% | -19.1% | -19.1% | -25.5% | -26.7% | -26.9% |
| Enc Time[%] | 102.4% | | | 102.9% | | |
| Dec Time[%] | 101.6% | | | 101.2% | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **RA HEVC 2x** | | | **RA HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -10.7% | -6.3% | -7.2% |  |  |  |
| Class B | -5.8% | -3.8% | -2.9% | -7.1% | -5.7% | -5.1% |
| **Overall (Test vs Ref)** | -7.2% | -4.5% | -4.1% | -7.1% | -5.7% | -5.1% |
| **Overall (Test vs single layer)** | 19.3% | 32.0% | 33.0% | 16.5% | 28.5% | 30.9% |
| **EL only (Test vs Ref)** | -13.2% | -10.5% | -10.0% | -18.7% | -17.1% | -16.6% |
| Enc Time[%] | 106.7% | | | 102.5% | | |
| Dec Time[%] | 101.5% | | | 101.9% | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **LD-P HEVC 2x** | | | **LD-P HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -8.6% | -3.1% | -3.7% |  |  |  |
| Class B | -4.6% | -2.9% | -1.9% | -6.7% | -4.5% | -4.2% |
| **Overall (Test vs Ref)** | -5.7% | -2.9% | -2.4% | -6.7% | -4.5% | -4.2% |
| **Overall (Test vs single layer)** | 26.1% | 36.5% | 38.2% | 22.5% | 33.1% | 36.0% |
| **EL only (Test vs Ref)** | -9.9% | -7.2% | -6.7% | -16.3% | -14.1% | -13.8% |
| Enc Time[%] | 105.7% | | | 103.9% | | |
| Dec Time[%] | 101.4% | | | 101.7% | | |

In the second experiment, sequences are generated with symmetric down-sampling. The anchor is with zero-phase up-sampling. With the help of signaling of phase offset, the proposed method can select the matched up-sampling phase and achieve around 5.6%-9.4% BD-rate reduction.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **AI HEVC 2x** | | | **AI HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -13.9% | -16.1% | -16.8% |  |  |  |
| Class B | -7.6% | -8.9% | -8.8% | -8.2% | -10.2% | -10.2% |
| **Overall (Test vs Ref)** | -9.4% | -11.0% | -11.1% | -8.2% | -10.2% | -10.2% |
| **Overall (Test vs single layer)** | 12.9% | 13.8% | 13.5% | 10.8% | 10.4% | 9.7% |
| **EL only (Test vs Ref)** | -17.0% | -18.5% | -18.6% | -22.5% | -24.4% | -24.6% |
| Enc Time[%] | 97.2% | | | 101.9% | | |
| Dec Time[%] | 93.7% | | | 97.8% | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **RA HEVC 2x** | | | **RA HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -10.2% | -5.8% | -7.2% |  |  |  |
| Class B | -5.4% | -3.8% | -3.0% | -6.4% | -5.2% | -4.7% |
| **Overall (Test vs Ref)** | -6.8% | -4.4% | -4.2% | -6.4% | -5.2% | -4.7% |
| **Overall (Test vs single layer)** | 19.7% | 32.2% | 33.1% | 16.8% | 28.5% | 31.1% |
| **EL only (Test vs Ref)** | -12.6% | -10.2% | -10.0% | -17.2% | -15.9% | -15.4% |
| Enc Time[%] | 97.6% | | | 105.1% | | |
| Dec Time[%] | 97.9% | | | 98.0% | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **LD-P HEVC 2x** | | | **LD-P HEVC 1.5x** | | |
|  | Y | U | V | Y | U | V |
| Class A | -8.4% | -2.9% | -4.2% |  |  |  |
| Class B | -4.4% | -3.1% | -2.3% | -6.4% | -4.7% | -4.2% |
| **Overall (Test vs Ref)** | -5.6% | -3.1% | -2.9% | -6.4% | -4.7% | -4.2% |
| **Overall (Test vs single layer)** | 26.5% | 36.7% | 38.1% | 22.7% | 33.0% | 35.9% |
| **EL only (Test vs Ref)** | -9.8% | -7.3% | -7.1% | -15.6% | -13.9% | -13.6% |
| Enc Time[%] | 98.4% | | | 101.5% | | |
| Dec Time[%] | 98.2% | | | 99.1% | | |

# Conclusions

This contribution proposes to have indication of phase shift for the up-sampling process so the up-sampling phase can match the down-sampling phase in the generation of the reference layer signals. In addition, this contribution proposes to adopt SVC reference layer sample location derivation process (which supports different chroma sampling location in reference layer and current layer) and the up-sampling filters for all 16 phases. Experimental results show the matched up-sampling can achieve 5%-10% Y BD-rate direction.

# References

1. J. Dong, Y. Ye, and Y. He, "sampling based on sampling grid information for aligned inter layer prediction," JCTVC-M0188, Apr. 2013.
2. L. Guo, J. Chen, and M. Karczewicz, “Signaling of Phase Offset in Up-sampling Process”, JCTVC-M0322, Apr., 2013.
3. K. Ugur, and J. Lainema, “AHG13: Signaling phase offset for upsampling in SHVC”, JCTVC-M0231, Apr. 2013.
4. Jie Dong, Yong He, and Yan Ye, "Downsampling filter for anchor generation for scalable extensions of HEVC," m23485, 99th MPEG meeting, San Jose, USA, Feb. 2012.

1. [X. Li](mailto:lxiang@qti.qualcomm.com), [J. Boyce](mailto:jill@vidyo.com), [P. Onno](mailto:patrice.onno@crf.canon.fr), and [Y. Ye](mailto:yan.ye@interdigital.com), “Common SHM test conditions and software reference configurations”, JCTVC-L1009, Geneva, Switzerland, Jan. 2013.

# Patent rights declaration(s)

**Qualcomm Incorporated 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).**

**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).**

**Nokia 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).**