|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  11th Meeting: Shanghai, CN, 10-19 Oct., 2012 | Document: JCTVC-K0364 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **A suggested initial software model for HEVC scalable video coding** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jianle Chen Krishna Rapaka Xiang Li Vadim Seregin Liwei Guo Marta Karczewicz Geert Van der Auwera Joel Sole Xianglin Wang Chengjie Tu Ying Chen  5775 Morehouse Dr. San Diego, CA 92121, USA | Tel: Email: | +1-858-651-8028  [cjianle@qti.qualcomm.com](mailto:cjianle@qti.qualcomm.com) [krapaka@qti.qualcomm.com](mailto:krapaka@qti.qualcomm.com) [lxiang@qti.qualcomm.com](mailto:lxiang@qti.qualcomm.com)  [vseregin@qti.qualcomm.com](mailto:vseregin@qti.qualcomm.com) [liweig@qti.qualcomm.com](mailto:liweig@qti.qualcomm.com)  [martak@qti.qualcomm.com](mailto:martak@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

This proposal provides a candidate for an initial software model for the scalability extension of HEVC.

The suggested software only includes the basic tools which are asserted to have sufficient coding efficiency and low complexity.

# Contents

[Abstract 1](#_Toc337899753)

[Contents 1](#_Toc337899754)

[1 Introduction 2](#_Toc337899755)

[2 Algorithm description 2](#_Toc337899756)

[2.1 Up-sampling filter 2](#_Toc337899757)

[2.2 Intra-BL prediction 3](#_Toc337899758)

[2.3 Inter-layer motion parameter prediction 3](#_Toc337899759)

[3 Syntax and semantics description 4](#_Toc337899760)

[3.1 Coding unit syntax 5](#_Toc337899761)

[3.2 Coding unit semantics 6](#_Toc337899762)

[4 Compression performance discussion 6](#_Toc337899763)

[4.1 Category 1 (HEVC base layer) 6](#_Toc337899764)

[4.1.1 BD-rate over simulcast based on EL+BL actual rate 6](#_Toc337899765)

[4.1.2 BD-rate over simulcast based on EL+BL target rate 6](#_Toc337899766)

[4.1.3 BD-rate savings over simulcast based on EL-only actual rate 6](#_Toc337899767)

[4.1.4 BD-rate savings over simulcast based on EL-only target rate 6](#_Toc337899768)

[5 Complexity analysis 6](#_Toc337899769)

[5.1 Encoding time and measurement methodology 6](#_Toc337899770)

[5.2 Decoding time and measurement methodology and comparison vs. anchor bitstreams decoded by HM 7](#_Toc337899771)

[6 Patent rights declaration 7](#_Toc337899772)

# Introduction

This proposal covers the following categories (tick box ⌧)

⌧ Category 1 (HEVC base layer) spatial scalability

⌧ Category 1 (HEVC base layer) intra-only spatial scalability

⌧ Category 1 (HEVC base layer) SNR scalability

🞏 Category 2 (AVC base layer) spatial scalability

⌧ The proposal obeys the constraints under section 5 of the CfP (if box is not ticked, explain cases where constraints are violated)

# Algorithm description

A framework of multi-loop decoding is employed and the following two basic coding tools are proposed on top of the HEVC tools:

* **Intra-BL prediction**: A coding tool that uses collocated BL block to predict EL samples.
* **Inter-layer MV prediction**: A coding tool that employs BL motion parameter on Merge and AMVP list derivation process.

## Up-sampling filter

The basic function of the up-sampling process is to generate a picture that has the resolution of the current layer picture from the BL picture. The sample position mapping for the up-sampling process used in this proposal is similar to that in H.264/AVC SVC. The fractional sample position, denoted as phase, is quantized in units of 1/16-pel. Therefore, an up-sampling filter set may employ up to sixteen filters. The zero phase offset is used for the top-left sample position of a picture.

The eight-tap filter set shown in Table 1 is used for the luma component up-sampling. These filters are designed in a similar manner as the HEVC luma motion compensation interpolation filters. It should be noted that the filter coefficients for phases 4 and 8 exactly match the quarter-pel and half-pel HEVC luma interpolation filters.

Table Luma filter for inter-layer intra prediction modes

|  |  |
| --- | --- |
| Phase | Filter coefficients |
| 0/16 | { 0, 0, 0, 64, 0, 0, 0, 0 } |
| 1/16 | { 0, 1, -3, 63, 4, -2, 1, 0 } |
| 2/16 | { 0, 2, -6, 61, 9, -3, 1, 0 } |
| 3/16 | { -1, 3, -8, 60, 13, -4, 1, 0 } |
| 4/16 | { -1, 4, -10, 58, 17, -5, 1, 0 } |
| 5/16 | { -1, 4, -11, 53, 25, -8, 3, -1 } |
| 6/16 | { -1, 4, -11, 50, 29, -9, 3, -1 } |
| 7/16 | { -1, 4, -11, 45, 34, -10, 4, -1 } |
| 8/16 | { -1, 4, -11, 40, 40, -11, 4, -1 } |

A four-tap filter set shown in Table 2 is used for the chroma component up-sampling process. The filter is applied to all the inter-layer prediction modes.

Table Chroma filter

|  |  |
| --- | --- |
| Phase | Filter coefficients |
| 0/16 | { 0, 64, 0, 0 } |
| 1/16 | { -2, 62, 4, 0 } |
| 2/16 | { -2, 58, 0, -2 } |
| 3/16 | { -4, 56, 14, -2 } |
| 4/16 | { -4, 54, 16, -2 } |
| 5/16 | { -6, 52, 20, -2 } |
| 6/16 | { -6, 48, 26, -4 } |
| 7/16 | { -4, 42, 30, -4 } |
| 8/16 | { -4, 36, 36, -4 } |

## Intra-BL prediction

Under the framework of multi-loop decoding, intra-BL mode indicates that the pixels in the collocated BL block are used as the prediction for the EL block.

## Inter-layer motion parameter prediction

As a part of inter-layer syntax prediction in this proposal, base layer MVs are used in enhancement layer coding.

In HEVC, motion vectors are compressed after being coded and such compressed motion vectors can be later used as TMVP for coding other frames. In this proposal, the uncompressed base layer MVs are used in enhancement layer coding. In other words, motion vector compression for a base layer is performed after it is used for inter-layer prediction.



Figure Center3 block

More specifically, the base layer MVs used in inter-layer prediction are obtained at a location co-located to the Center3 block in EL as shown in Figure 1, which is the same as that used for TMVP derivation in HEVC.

Additionally, when the spatial resolutions of BL and EL pictures are different, the base layer MVs are scaled according to the spatial resolution ratio between the two layers. For example, when the ratio is 3/2, motion vector scaling is performed according to the equation.

In this proposal, base layer MVs are used for both merge mode and AMVP mode in enhancement layer coding. The derived base layer MV is inserted as the first candidate in merge list and added after the TMVP candidate in AMVP list.

Additionally, validity check is performed on BL reference index. If its reference picture is not present in the EL reference picture list, the base layer MV is marked as “unavailable” and is not used in merge candidate list. For a bi-directional MV candidate from base layer, this reference index validity check is done for each reference picture list respectively. If base layer MV is marked as “unavailable” in one reference list, the bi-directional MV candidate is converted into uni-directional before using. And, if base layer MV is marked as “unavailable” in both reference lists, the candidate is not used in merge list.

For merge list construction at an enhancement layer, a new type of artificial merge candidates is generated by adding offset values to the first MV candidate in the merge candidate list. Such candidates are called offset MV candidates. If a merge candidate list is empty, zero MV with zero reference indexes is inserted prior to generating offset MV candidates. In this proposal, the four possible offset values of include , , and , with and representing the offset value for the horizontal and vertical components of a motion vector respectively. Such offset MV candidates are placed right after combined MV candidates in an EL merge candidate list. The maximum number of merge and AMVP motion vector candidates is kept the same as in HEVC, i.e. five for merge and two for AMVP.

# Syntax and semantics description

This proposal inserts few flags and indices in the coding unit level syntax to indicate whether additional coding tools are applied for that unit. The modified part of the HEVC syntax (based on HEVC text specification draft 6 [2]) is highlighted in the following syntax tables.

The syntax design is mainly based on the HEVC syntax/semantics, thus the syntax structure and the related semantics of HEVC are reused with additions and modifications, highlighted in yellow. In sub-sections that provide semantics of only the newly introduced syntax elements, the text is not highlighted.

## Coding unit syntax

|  |  |
| --- | --- |
| coding\_unit( x0, y0, log2CbSize ) { | Descriptor |
| CurrCbAddrTS = MinCbAddrZS[ x0 >> Log2MinCbSize ][ y0 >> Log2MinCbSize ] |  |
| if( slice\_type != I ) |  |
| **skip\_flag[** x0 **][** y0 **]** | ae(v) |
| if( skip\_flag[ x0 ][ y0 ]){ |  |
| prediction\_unit( x0, y0 , log2CbSize ) |  |
| } |  |
| else { |  |
| **intra\_bl\_flag** | ae(v) |
| if (intra\_bl\_flag) { | ae(v) |
| if( slice\_type != I ) |  |
| **pred\_mode\_flag** | ae(v) |
| if( PredMode != MODE\_INTRA | | log2CbSize = = Log2MinCbSize ) |  |
| **part\_mode** | ae(v) |
| 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( 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 ) |  |
| transform\_tree( x0, y0, x0, y0, log2CbSize, log2CbSize, log2CbSize, 0, 0 ) |  |
| } |  |
| } |  |

## Coding unit semantics

The specifications in sub-clause 3.1 apply with the following additions.

**intra\_bl\_skip\_flag**[ x0 ][  y0 ] equal to 1 specifies that the current coding unit is coded in intra-BL mode, and no more syntax elements are parsed after intra\_bl\_skip\_flag[ x0 ][ y0 ]. intra\_bl\_skip\_flag [ x0 ][ y0 ] equal to 0 specifies that more syntax elements are to be parsed. The array indices x0, y0 specify the location ( x0, y0 ) of the top-left luma sample of the considered coding block relative to the top-left luma sample of the picture.

**intra\_bl\_flag** equal to 1 specifies that the current coding unit is coded in intra-BL mode. intra\_bl\_flag equal to 0 specifies that coding unit is not coded in intra-BL mode.

When intra\_bl\_flag is not present, it is inferred to be equal to 0.

# Compression performance discussion

## Category 1 (HEVC base layer)

The proposed methods are tested according to the HEVC SVC CfP [1]. As requested in the CfP, four different types of BD-rate results compared to simulcast are provided in subsections 4.1.1 to 4.1.4.

### BD-rate over simulcast based on EL+BL actual rate

On average, xx% (All Intra) and xx% (Random Access) luma BD-rate reduction (based on EL+ BL actual rate) is obtained over simulcast. Clearly, the proposed methods are much more efficient than simulcast.

### BD-rate over simulcast based on EL+BL target rate

On average, xx% (All Intra) and xx% (Random Access) luma BD-rate reduction is obtained over simulcast when the EL+BL target rate instead of the EL+BL actual rate is used in the BD-rate calculation. Due to rate matching, the actual rate is close to the target rate so that the BD-rate reduction in this subsection is similar to that in the previous subsection.

### BD-rate savings over simulcast based on EL-only actual rate

On average, xx% (All Intra) and xx% (Random Access) luma BD-rate reduction is obtained over simulcast when the EL-only actual rate is used in the BD-rate calculation. Compared to the results based on the EL+BL actual rate, the gain by this measure is higher.

### BD-rate savings over simulcast based on EL-only target rate

On average, xx% (All Intra) and xx% (Random Access) luma BD-rate reduction is obtained over simulcast when the EL target rate is used in the BD-rate calculation. Similarly, the BD-rate reduction is close to that based on the EL actual rate due to the rate matching.

# Complexity analysis

## Encoding time and measurement methodology

Encoding time TSVC of the proposed SVC encoder is compared to anchor high resolution encoding time TAH. The geometric mean of encoding time ratio is summarized in the following table.

|  |  |
| --- | --- |
| Test conditions | Geometric mean of |
| AI 2x |  |
| AI 1.5x |  |
| RA 2x |  |
| RA 1.5x |  |
| RA SNR |  |

## Decoding time and measurement methodology and comparison vs. anchor bitstreams decoded by HM

Decoding time TSVC of the proposed SVC decoder is compared to anchor and high resolution decoding time TAH. The geometric mean of encoding time ratio is summarized in the following table.

|  |  |
| --- | --- |
| Test conditions | Geometric mean of |
| AI 2x |  |
| AI 1.5x |  |
| RA 2x |  |
| RA 1.5x |  |
| RA SNR |  |

# Patent rights declaration

**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).**