|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  16th Meeting: San José, US, 9–17 Jan. 2014 | Document: JCTVC-P0165 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Interlaced to progressive scalability in SHVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jianle Chen  Krishna Rapaka  Ye-kui Wang  Marta Karczewicz  5775 Morehouse Drive, San Diego, CA, 92121, U.S.A | Email: | [cjianle@qti.qualcomm.com](mailto:cjianle@qti.qualcomm.com)  [krapaka@qti.qualcomm.com](mailto:krapaka@qti.qualcomm.com)  [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com)  [martak@qti.qualcomm.com](mailto:martak@qti.qualcomm.com) |
| *Source:* | Qualcomm Inc. | | |

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

# Abstract

In this contribution, a lightweight solution to support interlaced to progressive scalability is proposed.

# Introduction

The current SHVC does not support scalable video coding with the base layer containing field pictures while the enhancement layer includes frame pictures. However migration from 1080i to 1080p with SHVC could be a potential user case in broadcasting, especially that the H.264/AVC 1080i bitstreams and related ecosystem are widely deployed nowadays. It would be worthy that SHVC could support the scalability from interlaced to progressive if the additional complexity burden is small.

# Interlaced to progressive scalability

SHVC support the two kinds of sample location alignment between different layer. One of them is that the top-left sample of the base layer and the enhancement layer are aligned. As shown in Figure 1 (a), in case base layer is a top field picture and enhancement layer is a frame picture, the location of top-left samples of base layer picture and enhancement picture are aligned. Therefore, for such case, the SHVC resampling process can be applied without change. While for the case that the base layer is a bottom field picture, as shown in Figure 1 (b), the SHVC resampling process needs minor change to take the advantage of inter-layer texture prediction.



**(a) BL: top field (b) BL: bottom field**

Figure 1 Luma sample location of interlaced base layer picture and progressive enhancement

This contribution proposes a lightweight solution to support interlaced to progressive scalability in SHVC by considering the field parity (a top field or a bottom field) during resampling process. When the base layer is a bottom field picture and enhancement is a frame picture, reference sample location is shifted by one sample vertically. A flag is signaled in slice header to indicate whether such shift is applied during resampling process which is invoked to generate the inter-layer reference picture of the current picture.

# Experimental results

To check the benefit of the proposed method, five interlaced 1080i sequences (Crowdrun, ParkJoy, DucksTakeOff, InToTree and OldTownCross) and the corresponding progressive sequences are employed in the experiments [1]. The interlaced 1080i sequence are coded as base layer and the corresponding 1080p sequence is coded in enhancement layer. Both HEVC base and AVC base were tested. The frame rate of the base layer sequence is 25 fps (i.e. 50 fields per second), the frame rate of the enhancement layer sequence is 50 fps. For all sequences, 250 frames in enhancement layer are tested.

The proposed method are tested SHM4.0 and the proposed method, with base layer coded by HEVC and AVC. When base layer is coded with HEVC, each field is coded as a picture. When base layer is coded with AVC, MB-AFF method is used. The output of each frame is interleaved as a top and a bottom field and feed to SHVC codec.

Table 1 shows the coding gain by using SHM4.0 as anchor with HEVC base layer. –9.1%, –0.2% and –4.1% BD rate saving, respectively, for AI, RA and RA with traditional IbbB GOP structure are achieved. The gain of RA with hierarchical B GOP structure is marginal because the bottom field in BL and its corresponding frame in EL are coded with highest temporal level so that the benefit of inter-layer prediction o

Table 1 Coding performance for HEVC base layer

|  |  |  |  |
| --- | --- | --- | --- |
| Configurations | BD rate saving | | |
| Y | U | V |
| All Intra | –9.1% | –9.6% | –9.5% |
| RA | –0.2% | 0.3% | 0.4% |
| RA with IbbB Gop | –4.1% | –2.6% | –1.9% |

Table 2 shows the coding gain with AVC base layer. –7.3%, –0.8% and –x.x% BD rate saving, respectively, for AI, RA and RA with traditional IbbB GOP structure are achieved.

Table 2 Coding performance for AVC base layer

|  |  |  |  |
| --- | --- | --- | --- |
| Configurations | BD rate saving | | |
| Y | U | V |
| All Intra | –7.3% | –8.6% | –8.6% |
| RA | –0.8% | 0.7% | 0.9% |
| RA with IbbB Gop | TBA | TBA | TBA |

# Conclusions

In this contribution, a lightweight solution to support interlaced to progressive scalability is proposed. The proposed method only introduced one line change in decoding process and it was asserted that the migration from 1080i to 1080p broadcasting could be a potential key application for SHVC. Therefore, we recommend the proposed method to be adopted in SHVC.

# References

1. <ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat/>

# Patent rights declaration(s)

**Qualcomm 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).**

# Spec text modification

Note: the modification or addition are highlighted.

## Syntax and semantics

**F.7.3.6.1 General slice segment header syntax**

|  |  |
| --- | --- |
| slice\_segment\_header( ) { | **Descriptor** |
| **…** |  |
| if( nuh\_layer\_id > 0 && !all\_ref\_layers\_active\_flag &&  NumDirectRefLayers[ nuh\_layer\_id ] > 0 ) { |  |
| **inter\_layer\_pred\_enabled\_flag** | u(1) |
| if( inter\_layer\_pred\_enabled\_flag && NumDirectRefLayers[ nuh\_layer\_id ] > 1) { |  |
| if( !max\_one\_active\_ref\_layer\_flag ) |  |
| **num\_inter\_layer\_ref\_pics\_minus1** | u(v) |
| if( NumActiveRefLayerPics != NumDirectRefLayers[ nuh\_layer\_id ] ) |  |
| for( i = 0; i < NumActiveRefLayerPics; i++ ) |  |
| **inter\_layer\_pred\_layer\_idc[**i ] | u(v) |
| } |  |
| } |  |
| if( field\_to\_frame\_resampling\_present\_flag && NumActiveRefLayerPics > 0 ) |  |
| **bottom\_field\_to\_frame\_resampling\_flag** | u(1) |
| **…** |  |

**bottom\_field\_to\_frame\_resampling\_flag** equal to 1 specifies that the input of the resampling process invoked to generate the interlayer reference pictures for the current picture is a bottom field picture and the output is a frame picture. The flag equal to 0 specifies that the above restriction does not apply. When not present, it is inferred to be equal to 0.

**F.7.3.2.3 Picture parameter set RBSP syntax**

|  |  |
| --- | --- |
| pic\_parameter\_set\_rbsp( ) { | Descriptor |
| **…** |  |
| if( nuh\_layer\_id > 0 ){ |  |
| **pps\_infer\_scaling\_list\_flag** | u(1) |
| field\_to\_frame\_resampling\_present\_flag | u(1) |
| } |  |
| if( pps\_infer\_scaling\_list\_flag ) |  |
| **pps\_scaling\_list\_ref\_layer\_id** | u(6) |
| else { |  |
| **pps\_scaling\_list\_data\_present\_flag** | u(1) |
| if( pps\_scaling\_list\_data\_present\_flag ) |  |
| scaling\_list\_data( ) |  |
| } |  |
| … |  |

**field\_to\_frame\_resampling\_present\_flag** equal to 1 specifies that the syntax bottom\_field\_to\_frame\_resampling\_flag may be present in the slice header segment header. field\_to\_frame\_resampling\_present\_flag equal to 0 specifies that the syntax bottom\_field\_to\_frame\_resampling\_flag is not present in the slice header segment header.

## Decoding process

H.6.2 Derivation process for reference layer sample location used in resampling

Inputs to this process are

– a variable cIdx specifying the color component index,

– a sample location ( xP, yP ) relative to the top-left sample of the color component of the current picture specified by cIdx.

Output of this process is a sample location ( xRef16, yRef16 ) specifying the reference layer sample location in units of 1/16-th sample relative to the top-left sample of the reference layer picture.

The variables offsetX and offsetY are derived as follows:

offsetX = ScaledRefLayerLeftOffset / ( ( cIdx = = 0)  ?  1 :  SubWidthC) (H‑3)  
offsetY = ScaledRefLayerTopOffset / ( ( cIdx = = 0)  ?  1 :  SubHeightC) (H‑4)

The variables phaseX, phaseY, addX and addY are derived as follows:

phaseX = ( cIdx = = 0 ) ? ( cross\_layer\_phase\_alignment\_flag << 1 ) : cross\_layer\_phase\_alignment\_flag (H‑5)  
phaseY = ( cIdx = = 0 ) ? ( cross\_layer\_phase\_alignment\_flag << 1 ) : cross\_layer\_phase\_alignment\_flag + 1 (H‑6)  
phaseY += 4 \* bottom\_field\_to\_frame\_resampling\_flag

addX = ( ScaleFactorX \* phaseX + 2 ) >> 2 (H‑7)   
addY = ( ScaleFactorY \* phaseY + 2 ) >> 2 (H‑8)

The variables xRef16 and yRef16 are derived as follows:

xRef16 = ( ( ( xP – offsetX ) \* ScaleFactorX  + addX + ( 1 << 11 ) ) >> 12 ) – ( phaseX << 2 ) (H‑9)  
yRef16 = ( ( ( yP – offsetY ) \* ScaleFactorY + addY + ( 1 << 11 ) ) >> 12 ) – ( phaseY << 2 ) (H‑10)