

The MEDIATEK logo is displayed in white, uppercase letters within an orange parallelogram shape.

MEDIATEK

AHG9: Chroma format and bitdepth constraints for multi-layer structures

Tzu-Der Chuang, Lulin Chen, Olena Chubach, Ching-Yeh Chen,
Chih-Wei Hsu, Yu-Wen Huang, Shaw-Min Lei

The bottom section of the slide features a solid orange background with a dense, repeating pattern of white line-art icons. These icons represent various objects and concepts, including a laptop, a coffee cup, a lightbulb, a desk, a camera, a puzzle piece, a whisk, a bottle, a basket, a pizza, a book, a pen, a lightbulb, a desk, a camera, a puzzle piece, a whisk, a bottle, a basket, a pizza, a book, a pen, and a lightbulb.

Presented by Tzu-Der (Peter) Chuang
17th Meeting: Brussels, BE,
7–17 January 2020

Overall Summary

- In VVC Draft 7, the chroma format and bitdepth of each video layer are independently signalled in SPS without any constraint
- Proposed chroma format constraint
 - The values of `chroma_format_idc` in all video layers shall be the same
 - The value of `chroma_format_idc` in a video layer shall be greater than or equal to the value in any lower reference video layer
- Proposed bitdepth constraint
 - The value of `bit_depth_minus8` in a video layer shall be greater than or equal to the value in any lower reference video layer

Problem Definition

- In VVC Draft 7, multi-layer structures in one coded video sequence can be supported
- In each coded layer video sequence, the `chroma_format_idc` and `bit_depth_minus8` are signalled independently in each SPS
- No constraint on the values of `chroma_format_idc` and `bit_depth_minus8` among different video layers
- Different video layers may have different chroma formats and bitdepths, but **the current VVC draft cannot support different chroma formats and bitdepths in inter-layer prediction process**

Bitstream Conformance Requirement 1: Single Chroma Format in Multiple Layers

- The chroma format of a video layer shall be the same as that of any lower reference layer

sps_seq_parameter_set_id provides an identifier for the SPS for reference by other syntax elements.

SPS NAL units, regardless of the `nuh_layer_id` values, share the same value space of `sps_seq_parameter_set_id`.

chroma_format_idc specifies the chroma sampling relative to the luma sampling as specified in clause 6.2.

It is a requirement of bitstream conformance that the value of `chroma_format_idc` of the current layer shall be the same as the value of `chroma_format_idc` of each reference layer of the current layer.

Bitstream Conformance Requirement 2: Multiple Chroma Formats in Multiple Layers

- A higher layer can have more chroma information
- The value of `chroma_format_idc` in a video layer shall be greater than or equal to that in any lower reference video layer

`sps_seq_parameter_set_id` provides an identifier for the SPS for reference by other syntax elements.

SPS NAL units, regardless of the `nuh_layer_id` values, share the same value space of `sps_seq_parameter_set_id`.

`chroma_format_idc` specifies the chroma sampling relative to the luma sampling as specified in clause 6.2.

It is a requirement of bitstream conformance that the value of `chroma_format_idc` of the current layer shall be greater than or equal to the value of `chroma_format_idc` of each reference layer of the current layer.

Decoding Process Modification: Chroma Sample Interpolation

$\text{addX} = \text{sps_chroma_horizontal_collocated_flag} ? 0 : 8 * (\text{scalingRatio}[0] - (1 \ll 14))$ (945)

$\text{addY} = \text{sps_chroma_vertical_collocated_flag} ? 0 : 8 * (\text{scalingRatio}[1] - (1 \ll 14))$ (946)

$\text{addXcur} = (\text{sps_chroma_horizontal_collocated_flag} ? 0 : 16) * \text{scalingRatio}[0] / \text{SubWidthC}$

$\text{addXref} = (\text{RefChromaHorCollocatedFlag} ? 0 : 16) * (1 \ll 14) / \text{RefSubWidthC}$

$\text{addX} = \text{addXcur} - \text{addXref}$

$\text{addYcur} = (\text{sps_chroma_vertical_collocated_flag} ? 0 : 16) * \text{scalingRatio}[1] / \text{SubHeightC}$

$\text{addYref} = (\text{RefChromaVerCollocatedFlag} ? 0 : 16) * (1 \ll 14) / \text{RefSubHeightC}$

$\text{addY} = \text{addYcur} - \text{addYref}$

Decoding Process Modification: Chroma Sample Interpolation

$$\text{refxSb}_C = (((xSb - \text{scaling_win_left_offset}) / \text{SubWidthC} \ll 5) + \text{refMvLX}[0]) * \text{scalingRatio}[0] + \text{addX} \quad (947)$$

$$\begin{aligned} \text{refx}_C = & ((\text{Sign}(\text{refxSb}_C) * ((\text{Abs}(\text{refxSb}_C) + 256) \gg 9) \\ & + xC * ((\text{scalingRatio}[0] + 8) \gg 4)) * (\text{SubWidthC}/\text{RefSubWidthC}) + \text{fRefLeftOffset} / \\ & \text{RefSubWidthC} + 16) \gg 5 \end{aligned} \quad (948)$$

$$\text{refySb}_C = (((ySb - \text{scaling_win_top_offset}) / \text{SubHeightC} \ll 5) + \text{refMvLX}[1]) * \text{scalingRatio}[1] + \text{addY} \quad (949)$$

$$\begin{aligned} \text{refy}_C = & ((\text{Sign}(\text{refySb}_C) * ((\text{Abs}(\text{refySb}_C) + 256) \gg 9) \\ & + yC * ((\text{scalingRatio}[1] + 8) \gg 4)) * (\text{SubHeightC}/\text{RefSubHeightC}) + \\ & \text{fRefTopOffset} / \text{RefSubHeightC} + 16) \gg 5 \end{aligned} \quad (950)$$

Bitstream Conformance Requirement on Different Bitdepths in Layers

- The bitdepth in a video layer shall be greater than or equal to the bitdepth in any lower reference video layer

bit_depth_minus8 specifies the bit depth of the samples of the luma and chroma arrays, **BitDepth**, and the value of the luma and chroma quantization parameter range offset, **QpBdOffset**, as follows:

$$\text{BitDepth} = 8 + \text{bit_depth_minus8} \quad (45)$$

$$\text{QpBdOffset} = 6 * \text{bit_depth_minus8} \quad (46)$$

bit_depth_minus8 shall be in the range of 0 to 8, inclusive.

It is a requirement of bitstream conformance that the value of **bit_depth_minus8** of the current layer shall be greater than or equal to the value of **bit_depth_minus8** of each reference layer of the current layer.

Decoding Process Modification: Sample Interpolation Filtering

- Consider the bitdepth of both reference and current layer

The variables shift1, shift2, shift3 and shift4 are derived as follows:

- Let the RefBitDepth be the BitDepth of the reference picture. The variable shift1 is set equal to $\text{Max}(0, \text{Min}(\text{RefBitDepth} + 4 - \text{BitDepth}, \text{RefBitDepth} - 8))$, the variable shift2 is set equal to 6, the variable shift3 is set equal to $\text{Max}(2 + \text{BitDepth} - \text{RefBitDepth}, 14 - \text{RefBitDepth})$ and the variable shift4 is set equal to $\text{Max}(0, \text{Max}(\text{BitDepth} - \text{RefBitDepth} - 4, 8 - \text{RefBitDepth}))$.

The predicted luma sample value predSampleLX_L is derived as follows:

- If both xFrac_L and yFrac_L are equal to 0, and both $\text{scalingRatio}[0]$ and $\text{scalingRatio}[1]$ are less than 20481, the value of predSampleLX_L is derived as follows:

$$\text{predSampleLX}_L = \text{refPicLX}_L[\text{xInt}_3][\text{yInt}_3] \ll \text{shift3} \quad (961)$$

- Otherwise, if yFrac_L is equal to 0 and $\text{scalingRatio}[1]$ is less than 20481, the value of predSampleLX_L is derived as follows:

$$\begin{aligned} \text{predSampleLX}_L = \\ \left(\sum_{i=0}^7 f_{LH}[\text{xFrac}_L][i] * (\text{refPicLX}_L[\text{xInt}_i][\text{yInt}_3] \ll \text{shift4}) \right) \gg \text{shift1} \end{aligned} \quad (962)$$

- Otherwise, if xFrac_L is equal to 0 and $\text{scalingRatio}[0]$ is less than 20481, the value of predSampleLX_L is derived as follows:

$$\begin{aligned} \text{predSampleLX}_L = \\ \left(\sum_{i=0}^7 f_{LV}[\text{yFrac}_L][i] * (\text{refPicLX}_L[\text{xInt}_3][\text{yInt}_i] \ll \text{shift4}) \right) \gg \text{shift1} \end{aligned} \quad (963)$$

- Otherwise, the value of predSampleLX_L is derived as follows:
 - The sample array $\text{temp}[n]$ with $n = 0..7$, is derived as follows:

$$\text{temp}[n] = \left(\sum_{i=0}^7 f_{LH}[\text{xFrac}_L][i] * (\text{refPicLX}_L[\text{xInt}_i][\text{yInt}_n] \ll \text{shift4}) \right) \gg \text{shift1} \quad (964)$$

- The predicted luma sample value predSampleLX_L is derived as follows:

$$\text{predSampleLX}_L = \left(\sum_{i=0}^7 f_{LV}[\text{yFrac}_L][i] * \text{temp}[i] \right) \gg \text{shift2} \quad (965)$$

Conclusions

- Propose two bitstream conformance requirements on chroma format, suggest to adopt one of them
 - The values of `chroma_format_idc` in all video layers shall be the same
 - The value of `chroma_format_idc` in a video layer shall be greater than or equal to the value in any lower reference video layer
- Propose one bitstream conformance requirement on bitdepth
 - The value of `bit_depth_minus8` in a video layer shall be greater than or equal to the value in any lower reference video layer



everyday genius