|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  15th Meeting: Geneva, CH, 23 Oct. – 1 Nov. 2013 | Document: JCTVC-O0161 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Non-SCE4/AHG14: Combined bit-depth and color gamut conversion with 3D LUT for SHVC color gamut scalability** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Yuwen He, Yan Ye, Jie Dong  9710 Scranton R-D, #250 San Diego, CA 92121 USA | Tel: Email: | +1-858-210-4819 [yuwen.he@interdigital.com](mailto:yuwen.he@interdigital.com) [yan.ye@interdigital.com](mailto:yan.ye@interdigital.com) [jie.dong@interdigital.com](mailto:jie.dong@interdigital.com) |
| *Source:* | InterDigital Communications, Inc. | | |

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

# Abstract

This proposal describes a combined bit-depth and color gamut conversion method with 3D LUT for SHVC color gamut scalability (CGS). In one of the SCE4 color gamut scalability tests, the base layer video format is 8-bit 1080p BT.709, and the enhancement layer video format is 10-bit 3840x2160 BT.2020. Therefore both bit-depth conversion and color gamut conversion need to be addressed in inter-layer processing, in addition to upsampling. The proposed method uses combined 3D LUT for color gamut conversion and bit-depth conversion in one step. The proposed method has three advantages compared to keeping color gamut conversion and bit-depth conversion separate, (1) higher coding efficiency, (2) higher precision and fewer rounding errors, (3) no change to upsampling in SHVC draft 3. Compared to the SCE4 anchors, the proposed scheme reportedly achieves average {Y, U, V} BD rate gain of {-15.3%, -15.7%, -22.9%} and {-10.0%, -8.7%, -16.6%} for AI and RA-2x, respectively. Compared to keeping bit-depth and color gamut conversion separate, the proposed scheme reportedly achieves average {Y, U, V} BD rate gain of {-2.4%, -2.9%, -5.3%}, and {-1.0%, -1.5%, -4.1%} for AI and RA-2x, respectively.

# Introduction

SCE4 is currently studying color gamut and bit depth scalability [3]. One of the SCE4 test cases, where the base layer is in 1080p, 8-bit, and BT.709 gamut and the enhancement layer is in 4Kx2K, 10-bit, and BT.2020 gamut, represents a practical application in which base layer is for HDTV service and enhancement layer for UHDTV service. Because the bit-depth, color gamut and resolution of the two layers are all different, bit-depth conversion, color gamut conversion and upsampling process are needed in inter-layer process to generate inter-layer reference pictures from the BL reconstructed pictures for efficient EL coding. Bordes et al. [4] proposed tri-linear interpolation based 3D LUT for color gamut conversion, which achieved accurate color gamut conversion and improved coding performance. Ugur et al. [6] and Alshina et al. [7] proposed combined bit-depth conversion and upsampling process to reduce complexity while keeping high upsampling precision.

In this contribution, a combined 3D LUT is proposed to perform bit-depth conversion and color gamut conversion in one step. In the combined 3D LUT method, the LUT inputs are in 8-bit, and the LUT outputs are in 10-bit. By considering bit-depth conversion distortion and color gamut conversion jointly in the 3D LUT derivation process, higher inter layer prediction accuracy and therefore higher EL coding efficiency is achieved. Figure 1 is the workflow of the inter-layer process for the proposed method that combines color gamut conversion and bit-depth conversion. The combined 3D LUT is applied in color gamut conversion, and the upsamping process is kept exactly the same as SHM-3.0.1 [1] without any changes. In comparison, Figure 2 is the workflow of the inter-layer process where color gamut conversion is kept separate from bit-depth conversion, and where bit-depth conversion is combined with upsampling as proposed in [6][7]. The 3D LUT with 8-bit inputs and 8-bit outputs is used in color gamut conversion in Figure 2.



Figure . Color gamut conversion and bit-depth conversion combined process



Figure 2. Color gamut conversion and bit-depth conversion separate process

In order to reduce the computation complexity of color gamut conversion with 3D LUT, the tri-linear interpolation is substituted by tetrahedral interpolation in 3D LUT interpolation process. As explained in [4], tri-linear interpolation uses 8 vertices of the smallest octant encompassing the point to be interpolated for linear-interpolation in three directions. Take Figure 3 as an example, P(y,u,v) is the point to be interpolated, and it is encompassed in the octant whose vertices are P0, P1, P2, P3, P4, P5, P6, P7. The tri-linear interpolation is calculated with Equation (1)(2)(3) for each component.

(1)

(2)

(3)

; ;

; ;

; ;

In tetrahedral interpolation, it only uses 4 vertices of the tetrahedron containing the point to be interpolated for calculation. The point P in Figure 3 is encompassed in the tetrahedron whose vertices are P0, P1, P5, P7.The tetrahedral interpolation is calculated in Equation (4)(5)(6) for each component.

+dy

(4)

+dy

(5)

+dy

(6)

=

Compared with tri-linear interpolation, tetrahedral interpolation is simpler. We also compared the color gamut conversion accuracy and coding performance between tetrahedral and tri-linear, the performance of tetrahedral interpolation is almost the same as tri-linear interpolation.



Figure . Tetrahedral interpolation

There are 6 choices to determine the tetrahedron containing the point to be interpolated in the octant given that P0 and P7 have to be included in the tetrahedron. Figure 4 lists 6 cases. The lookup table is designed for fast decision instead of checking the relationship of each two components: y and u, y and v, u and v.



Figure . Tetrahedron to encompass the point P to be interpolated

# 3D LUT derivation and coding

The 3D LUTs used in our simulations are trained offline. The 3D LUT derivation processes for the methods in Figure 1 and Figure 2 are depicted in Figure 5 and Figure 6, respectively. The 3D LUT derivation for the SCE4 test case with 10-bit BL and 10-bit EL is shown in Figure 7. All 3D LUTs are off-line trained based on base layer reconstructed signal and enhancement layer downsampled signal at the whole sequence level with Least Square method. The bit-depth conversion in Figure 6 is the same as that in content generation [5], where simple offset and shifting is applied.



Figure 5. 10-bit 3D LUT derivation for base 8-bit coding configuration



Figure 6. 8-bit 3D LUT derivation for base 8-bit coding configuration



Figure 7. 10-bit 3D LUT derivation for base 10-bit coding configuration

We use the same entropy coding method as that in [8] for 3D LUT coding. The input and the output bit-depth are signaled in the 3D LUT to support both 8-bit input LUT and 10-bit input LUT. In our simulations, the 3D LUT is signaled in the PPS extension, in the same way as in [4][8]. Table 1 are the parameters of color gamut conversion with 3D LUT coded in PPS.

Table 1. Color gamut conversion parameter signaling in PPS

|  |  |
| --- | --- |
| color\_gamut\_conversion\_param( ) { | **Descriptor** |
| **bit\_depth\_input\_luma\_minus8** | ue(v) |
| **bit\_depth\_input\_chroma\_delta** | ue(v) |
| **bit\_depth\_output\_luma\_delta** | ue(v) |
| **bit\_depth\_output\_chroma\_delta** | ue(v) |
| **layer\_num\_of\_3D\_LUT** | ue(v) |
| 3D\_LUT() |  |
| } |  |

**bit\_depth\_input\_luma\_minus8** : (bit\_depth\_input\_luma\_minus8+8) is the bit-depth of the input luma signal for 3D LUT.

**bit\_depth\_input\_chroma\_delta** : (bit\_depth\_input\_chroma\_delta+bit\_depth\_input\_luma) is the bit-depth of the input chroma signal for 3D LUT.

**bit\_depth\_output\_luma\_delta** : (bit\_depth\_output\_luma\_delta+bit\_depth\_input\_luma) is the bit-depth of the output luma signal for 3D LUT, which is used for clipping operations in color gamut conversion process.

**bit\_depth\_output\_chroma\_delta** : (bit\_depth\_output\_chroma\_delta+bit\_depth\_input\_luma) is the bit-depth of the output chroma signal for 3D LUT, which is used for clipping operations in color gamut conversion process.

**layer\_num\_of\_3D\_LUT**: it is number of layers of 3D LUT octant tree. It is used to calculate the size of 3D LUT. The size of 3D LUT in one dimension is equal to ((1<<( layer\_num\_of\_3D\_LUT-1))+1), e.g. if layer\_num\_of\_3D\_LUT is 4, then the total 3D LUT size is 9x9x9.

3D\_LUT(): it is same as that defined in [8] with extension to support that the input and output bit-dpeth are different.

# Simulation results

The compression performance is measured using BD rate compared with SCE4 anchors, using the SCE4 test conditions [2]. There are two coding configurations: 10-bit base and 8-bit base. For 10-bit base coding, 10-bit 3D LUT derived as previously described is always used. For 8-bit base coding, we conducted two tests: 1) we used the proposed method in Figure 1; 2) we used the alternative method in Figure 2. The 3D LUT size is 17x17x17 for all tests. Table 2 gives the detailed average BD rate reduction for combined 3D LUT (as in Figure 1) compared with SCE4 anchors. Table 3 gives the detailed average BD rate reduction for 8-bit 3D LUT (as in Figure 2) compared with SCE4 anchors. Table 4 gives the detailed average BD rate reduction for the proposed method in Figure 1, using the 8-bit 3D LUT method in Figure 2 as reference.

As shown in Table 2, compared with SCE4 anchors, the proposed combined bit-depth and color gamut conversion scheme achieves average {Y, U, V} BD rate gain of {-15.3%, -15.7.0%, -22.9%}, and {-10.0%, -8.7%, -16.6%} for AI and RA-2x, respectively. As shown in Table 4, compared with using the 8-bit 3D LUT method in Figure 2, the proposed method achieves average {Y, U, V} BD rate gain of {-2.4%, -2.9%, -5.3%}, and {-1.0%, -1.5%, -4.1%} for AI and RA-2x, respectively.

We also decoded 8-bit base anchor bitstreams and those bitstreams of 8-bit base test with 10-bit 3D LUT on the desktop to compare the decoding time. The decoding time is 94.6% for AI-2x, and 105.6% for RA-2x compared with anchor, respectively.

Readers are referred to the accompanying spreadsheets for further details. The encoding time is not reliable because the cores in grid for simulations are not homogeneous. The decoding time was collected from the same workstation.

“SHM3.0.1\_scbd\_bugfix(IDCC)\_vs\_8bit\_Tetra3DLUT\_dectime\_update.xls” is the data sheet for the results of using 8 bit 3D LUT for base layer 8 bit coding test compared with SCE4 anchors.

“SHM3.0.1\_scbd\_bugfix(IDCC)\_vs\_10bit\_Tetra3DLUT\_dectime\_update.xls” is the data sheet for the results of using 10 bit 3D LUT for base layer 8 bit coding test compared with SCE4 anchors.

“SHM3.0.1\_scbd\_bugfix(IDCC)\_8bit\_vs\_10bit\_Tetra3DLUT\_dectime\_update.xls” is the data sheet for the results of using 10 bit 3D LUT for base layer 8 bit coding test compared with using 8-bit 3D LUT.

Table 2. Average BD rate reduction for combined 3D LUT compared with SCE4 anchors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **AI HEVC 2x 10-bit base** | | | **AI HEVC 2x 8-bit base** | | |
|  | Y | U | V | Y | U | V |
|  |  |  |  |  |  |  |
| Class A+ | -15.4% | -15.5% | -22.2% | -15.3% | -15.7% | -22.9% |
| **Overall (Test vs Ref)** | -15.4% | -15.5% | -22.2% | -15.3% | -15.7% | -22.9% |
| **Overall (Test vs single layer)** | 7.1% | 9.3% | 0.5% | 9.8% | 11.3% | 1.9% |
| **Overall (Ref vs single layer)** | 26.8% | 29.4% | 28.9% | 29.8% | 32.1% | 32.1% |
| **EL only (Test vs Ref)** | -28.2% | -27.8% | -33.7% | -28.3% | -28.3% | -34.6% |
| **Overall (Test EL+BL vs single EL+BL)** | -30.0% | -28.5% | -35.2% | -28.2% | -27.3% | -34.4% |
| Enc Time[%] | 122.6% | | | 82.5% | | |
| Dec Time[%] | 95.6% | | | 94.6% | | |
|  |  |  |  |  |  |  |
|  | **RA HEVC 2x 10-bit base** | | | **RA HEVC 2x 8-bit base** | | |
|  | Y | U | V | Y | U | V |
|  |  |  |  |  |  |  |
| Class A+ | -10.5% | -9.1% | -17.2% | -10.0% | -8.7% | -16.6% |
| **Overall (Test vs Ref)** | -10.5% | -9.1% | -17.2% | -10.0% | -8.7% | -16.6% |
| **Overall (Test vs single layer)** | 18.3% | 23.3% | 9.6% | 20.3% | 24.7% | 11.2% |
| **Overall (Ref vs single layer)** | 32.3% | 35.5% | 33.0% | 33.8% | 36.4% | 33.9% |
| **EL only (Test vs Ref)** | -18.9% | -17.0% | -24.6% | -18.4% | -16.5% | -23.9% |
| **Overall (Test EL+BL vs single EL+BL)** | -21.3% | -17.0% | -27.0% | -20.1% | -16.4% | -26.2% |
| Enc Time[%] | 114.8% | | | 91.1% | | |
| Dec Time[%] | 105.6% | | | 105.6% | | |

Table 3. Average BD rate reduction for 8 bit 3D LUT compared with SCE4 anchors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **AI HEVC 2x 10-bit base** | | | **AI HEVC 2x 8-bit base** | | |
|  | Y | U | V | Y | U | V |
|  |  |  |  |  |  |  |
| Class A+ | -15.4% | -15.5% | -22.2% | -13.2% | -13.2% | -18.9% |
| **Overall (Test vs Ref)** | -15.4% | -15.5% | -22.2% | -13.2% | -13.2% | -18.9% |
| **Overall (Test vs single layer)** | 7.1% | 9.3% | 0.5% | 12.6% | 14.7% | 7.2% |
| **Overall (Ref vs single layer)** | 26.8% | 29.4% | 28.9% | 29.8% | 32.1% | 32.1% |
| **EL only (Test vs Ref)** | -28.2% | -27.8% | -33.7% | -24.6% | -24.3% | -29.4% |
| **Overall (Test EL+BL vs single EL+BL)** | -30.0% | -28.5% | -35.2% | -26.0% | -24.7% | -30.5% |
| Enc Time[%] | 122.6% | | | 77.7% | | |
| Dec Time[%] | 95.6% | | | 97.6% | | |
|  |  |  |  |  |  |  |
|  | **RA HEVC 2x 10-bit base** | | | **RA HEVC 2x 8-bit base** | | |
|  | Y | U | V | Y | U | V |
|  |  |  |  |  |  |  |
| Class A+ | -10.5% | -9.1% | -17.2% | -9.1% | -7.3% | -13.2% |
| **Overall (Test vs Ref)** | -10.5% | -9.1% | -17.2% | -9.1% | -7.3% | -13.2% |
| **Overall (Test vs single layer)** | 18.3% | 23.3% | 9.6% | 21.5% | 26.7% | 15.7% |
| **Overall (Ref vs single layer)** | 32.3% | 35.5% | 33.0% | 33.8% | 36.4% | 33.9% |
| **EL only (Test vs Ref)** | -18.9% | -17.0% | -24.6% | -16.8% | -14.4% | -19.9% |
| **Overall (Test EL+BL vs single EL+BL)** | -21.3% | -17.0% | -27.0% | -19.1% | -15.0% | -22.9% |
| Enc Time[%] | 114.8% | | | 85.5% | | |
| Dec Time[%] | 105.6% | | | 106.5% | | |

Table 4. Average BD rate reduction for 10-bit 3D LUT compared with 8-bit 3D LUT

|  |  |  |  |
| --- | --- | --- | --- |
|  | **AI HEVC 2x 8-bit base** | | |
|  | Y | U | V |
|  |  |  |  |
| Class A+ | -2.4% | -2.9% | -5.3% |
| **Overall (Test vs Ref)** | -2.4% | -2.9% | -5.3% |
| **Overall (Test vs single layer)** | 9.8% | 11.3% | 1.9% |
| **Overall (Ref vs single layer)** | 12.6% | 14.7% | 7.2% |
| **EL only (Test vs Ref)** | -5.0% | -5.4% | -8.0% |
| **Overall (Test EL+BL vs single EL+BL)** | -28.2% | -27.3% | -34.4% |
| Enc Time[%] | 106.1% | | |
| Dec Time[%] | 96.9% | | |
|  |  |  |  |
|  | **RA HEVC 2x 8-bit base** | | |
|  | Y | U | V |
|  |  |  |  |
| Class A+ | -1.0% | -1.5% | -4.1% |
| **Overall (Test vs Ref)** | -1.0% | -1.5% | -4.1% |
| **Overall (Test vs single layer)** | 20.3% | 24.7% | 11.2% |
| **Overall (Ref vs single layer)** | 21.5% | 26.7% | 15.7% |
| **EL only (Test vs Ref)** | -2.0% | -2.5% | -5.4% |
| **Overall (Test EL+BL vs single EL+BL)** | -20.1% | -16.4% | -26.2% |
| Enc Time[%] | 106.6% | | |
| Dec Time[%] | 99.1% | | |

# Conclusions

In this proposal, the bit-depth conversion and color gamut conversion are combined into one combined 3D LUT for SHVC color gamut scalability coding. The bit-depth difference between base layer and enhancement layer is taken into account during 3D LUT derivation process. Compared with keeping bit-depth conversion and color gamut conversion separate, the combined method can achieve higher coding efficiency , with average BD rate gains of {-2.4%, -2.9%, -5.3%} for AI and {-1.0%, -1.5%, -4.1%} for RA-2x.

# Patent rights declaration(s)

**InterDigital Communications, Inc. 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).**

# References

1. J. Chen, J. Boyce, Y. Ye, M. M. Hannuksela, “Scalable High Efficiency Video Coding Test Model 3 (SHM3)”, JCTVC-N1007, Jul. 2013.
2. X. Li, J. Boyce, P. Onno, Y. Ye, “Common SHM test conditions and software reference configurations”, JCTVC-N1009, Jul. 2013.
3. A. Segall, P. Bordes, C. Auyeung, X. Li, E. Alshina, A. Duenas, “Description of Core Experiment SCE4: Color Gamut and Bit-Depth Scalability”, JCTVC-N1104, Jul. 2013.
4. P. Bordes, P. Andrivon, F. Hiron, “AHG14: Color Gamut Scalable Video Coding using 3D LUT: New Results”, JCTVC-N0168, Jul. 2013.
5. P. Andrivon, P. Bordes, “AHG14: Wide Color Gamut Test Material Creation”, JCTVC-N0163, Jul. 2013.
6. K. Ugur, A. Aminlou, “AHG14: On resampling & color gamut scalability”, JCTVC-N0146, Jul. 2013.
7. E. Alshina, A. Alshin, “AHG14: On bit-depth scalability support”, JCTVC-N0218, Jul. 2013.
8. P. Bordes, P. Andrivon, F. Hiron, P. Salmon, R. Boitard, “SCE4: Results on 5.3-test1 and 5.3-test2”, JCTVC-O0159, Oct. 2013