

JCTVC-R0179 NON-SCE1: 3D LUT WITH LARGER SIZES



invention | collaboration | contribution

Yan Ye, Yuwen He
InterDigital Communications, Inc.

18th JCT-VC meeting, July 2014

Introduction

- 3D LUT based Color Gamut Scalability was adopted into SHVC draft 6 at the last meeting
- Maximum LUT size was limited to 8x2x2
 - At most $8 \times 2 \times 2 \times 4 \times 3 = 384$ coefficients need to be stored
- Syntax elements lengths allow for larger LUT sizes
 - Offline discussions between meetings

colour_mapping_table() {	Descriptor
cm_octant_depth	u(2)
cm_y_part_num_log2	u(2)
cm_input_luma_bit_depth_minus8	u(3)
cm_input_chroma_bit_depth_delta	se(v)
cm_output_luma_bit_depth_minus8	u(3)
cm_output_chroma_bit_depth_delta	se(v)
cm_res_quant_bits	u(2)
colour_mapping_octants(0, 0, 0, 0, $1 \ll \text{cm_octant_depth}$)	
}	

- Keep current syntax design for future extensibility
 - Unmodified SHM6.0 provides additional coding gain with larger LUT sizes
 - Encoder optimization can provide further coding gain
 - Max 2 partitions for U/V may not be sufficient when bit-depth increases

8x1x1 LUT

- 8x1x1 is smaller than 8x2x2 but not allowed

cm_y_part_num_log2 specifies the number of partitions of the smallest colour mapping table octant for the luma component. The variables YOctantNum and YPartNum are derived as follows. In bitstreams conforming to this version of this Specification, the value of **cm_y_part_num_log2** shall be in the range of 0 to 2, inclusive. Other value for **cm_y_part_num_log2** is reserved for future use by ITU-T | ISO/IEC.

- Suggest to allow 8x1x1 with the following text change given 8x1x1 has no additional complexity

cm_y_part_num_log2 specifies the number of partitions of the smallest colour mapping table octant for the luma component. The variables YOctantNum and YPartNum are derived as follows. In bitstreams conforming to this version of this Specification, the value of (**cm_y_part_num_log2 + cm_octant_depth**) shall be in the range of 0 to 3, inclusive. Other value for **cm_y_part_num_log2** is reserved for future use by ITU-T | ISO/IEC.

RD cost based 3D LUT size selection

- SHM6.0 uses threshold-based LUT size control
 - LUT size is decreased when the signaling cost of 3D LUT is over 3% of the previously coded picture at the same temporal level
 - LUT size is increased when its signaling cost is below 0.5% of the previously coded picture at the same temporal level
- We propose RD cost based LUT size selection using weighted distortion

$$J_{\text{Cost}}(s) = w \cdot D(s) + \lambda \cdot \text{bits}(s)$$

where “w” depends on slice type and temporal level

- Two additional conditions:
 - Disable PPS update of 3D LUT for temporal level 2 and 3
 - Disable PPS update of 3D LUT for RASL pictures

Performance evaluation, max 8x8x8 LUT, AI

SHM6.0

	AI HEVC 1x 10-bit base			AI HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-1.7%	-3.4%	-3.5%	-1.6%	-2.2%	-3.2%
Overall (Test vs Ref)	-1.7%	-3.4%	-3.5%	-1.6%	-2.2%	-3.2%
Overall (Test vs single layer)	16.7%	6.8%	-11.9%	9.1%	8.9%	0.5%
Overall (Ref vs single layer)	18.7%	10.4%	-9.0%	10.9%	11.4%	3.7%
EL only (Test vs Ref)	-10.7%	-13.1%	-13.1%	-3.4%	-4.1%	-5.2%
Overall (Test EL+BL vs single EL+BL)	-37.9%	-43.0%	-53.7%	-28.7%	-28.9%	-35.5%
Overall (Ref EL+BL vs single EL+BL)	-36.9%	-41.4%	-52.5%	-27.5%	-27.2%	-33.3%
Enc Time[%]	77.9%			84.5%		
Dec Time[%]	71.5%			72.6%		

SHM6.0 + proposed encoder change

	AI HEVC 1x 10-bit base			AI HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-2.1%	-4.3%	-4.7%	-1.8%	-2.4%	-3.5%
Overall (Test vs Ref)	-2.1%	-4.3%	-4.7%	-1.8%	-2.4%	-3.5%
Overall (Test vs single layer)	16.2%	6.0%	-13.1%	9.0%	8.7%	0.2%
Overall (Ref vs single layer)	18.7%	10.4%	-9.0%	10.9%	11.4%	3.7%
EL only (Test vs Ref)	-14.0%	-16.3%	-16.9%	-3.8%	-4.5%	-5.7%
Overall (Test EL+BL vs single EL+BL)	-38.2%	-43.5%	-54.5%	-28.8%	-29.1%	-35.7%
Overall (Ref EL+BL vs single EL+BL)	-36.9%	-41.4%	-52.5%	-27.5%	-27.2%	-33.3%
Enc Time[%]	83.6%			110.3%		
Dec Time[%]	76.1%			94.5%		

Performance evaluation, max 8x8x8 LUT, RA

SHM6.0

	RA HEVC 1x 10-bit base			RA HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-1.2%	-2.4%	-2.6%	-0.4%	-0.7%	-1.1%
Overall (Test vs Ref)	-1.2%	-2.4%	-2.6%	-0.4%	-0.7%	-1.1%
Overall (Test vs single layer)	27.6%	20.0%	-2.7%	20.7%	21.5%	11.2%
Overall (Ref vs single layer)	29.1%	22.9%	-0.5%	21.3%	22.5%	12.4%
EL only (Test vs Ref)	-5.5%	-6.7%	-6.9%	-0.8%	-1.1%	-1.5%
Overall (Test EL+BL vs single EL+BL)	-30.6%	-33.6%	-46.8%	-19.7%	-18.0%	-25.9%
Overall (Ref EL+BL vs single EL+BL)	-30.0%	-32.7%	-46.2%	-19.4%	-17.5%	-25.2%
Enc Time[%]	79.8%			81.8%		
Dec Time[%]	61.6%			71.6%		

SHM6.0 + proposed encoder change

	RA HEVC 1x 10-bit base			RA HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-1.9%	-4.5%	-5.1%	-0.9%	-1.9%	-2.6%
Overall (Test vs Ref)	-1.9%	-4.5%	-5.1%	-0.9%	-1.9%	-2.6%
Overall (Test vs single layer)	26.6%	17.7%	-5.0%	20.2%	20.2%	9.6%
Overall (Ref vs single layer)	29.1%	22.9%	-0.5%	21.3%	22.5%	12.4%
EL only (Test vs Ref)	-8.8%	-11.6%	-12.3%	-1.8%	-2.9%	-3.6%
Overall (Test EL+BL vs single EL+BL)	-31.2%	-35.2%	-48.4%	-20.1%	-19.0%	-27.2%
Overall (Ref EL+BL vs single EL+BL)	-30.0%	-32.7%	-46.2%	-19.4%	-17.5%	-25.2%
Enc Time[%]	80.7%			100.6%		
Dec Time[%]	64.5%			80.7%		

Performance evaluation, max 8x4x4 LUT, AI

SHM6.0

	AI HEVC 1x 10-bit base			AI HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-0.9%	-1.5%	-1.9%	-0.8%	-0.9%	-1.4%
Overall (Test vs Ref)	-0.9%	-1.5%	-1.9%	-0.8%	-0.9%	-1.4%
Overall (Test vs single layer)	17.7%	8.9%	-10.6%	10.0%	10.4%	2.3%
Overall (Ref vs single layer)	18.7%	10.4%	-9.0%	10.9%	11.4%	3.7%
EL only (Test vs Ref)	-6.2%	-6.9%	-7.4%	-1.8%	-1.8%	-2.4%
Overall (Test EL+BL vs single EL+BL)	-37.6%	-42.4%	-53.5%	-28.2%	-28.0%	-34.3%
Overall (Ref EL+BL vs single EL+BL)	-36.9%	-41.4%	-52.5%	-27.5%	-27.2%	-33.3%
Enc Time[%]	77.9%			84.9%		
Dec Time[%]	72.0%			74.3%		

SHM6.0 + proposed encoder change

	AI HEVC 1x 10-bit base			AI HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-0.9%	-1.5%	-1.8%	-0.8%	-0.9%	-1.3%
Overall (Test vs Ref)	-0.9%	-1.5%	-1.8%	-0.8%	-0.9%	-1.3%
Overall (Test vs single layer)	17.7%	8.9%	-10.5%	10.0%	10.4%	2.4%
Overall (Ref vs single layer)	18.7%	10.4%	-9.0%	10.9%	11.4%	3.7%
EL only (Test vs Ref)	-6.1%	-6.8%	-7.2%	-1.8%	-1.8%	-2.3%
Overall (Test EL+BL vs single EL+BL)	-37.6%	-42.3%	-53.5%	-28.2%	-28.0%	-34.3%
Overall (Ref EL+BL vs single EL+BL)	-36.9%	-41.4%	-52.5%	-27.5%	-27.2%	-33.3%
Enc Time[%]	101.8%			84.9%		
Dec Time[%]	98.8%			73.5%		

Performance evaluation, max 8x4x4 LUT, RA

SHM6.0

	RA HEVC 1x 10-bit base			RA HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-0.3%	-0.3%	-0.5%	-0.1%	0.1%	-0.2%
Overall (Test vs Ref)	-0.3%	-0.3%	-0.5%	-0.1%	0.1%	-0.2%
Overall (Test vs single layer)	28.7%	22.6%	-0.9%	21.1%	22.5%	12.2%
Overall (Ref vs single layer)	29.1%	22.9%	-0.5%	21.3%	22.5%	12.4%
EL only (Test vs Ref)	-1.9%	-1.8%	-2.0%	-0.2%	0.0%	-0.3%
Overall (Test EL+BL vs single EL+BL)	-30.2%	-32.9%	-46.3%	-19.5%	-17.5%	-25.4%
Overall (Ref EL+BL vs single EL+BL)	-30.0%	-32.7%	-46.2%	-19.4%	-17.5%	-25.2%
Enc Time[%]	78.5%			81.4%		
Dec Time[%]	60.7%			71.7%		

SHM6.0 + proposed encoder change

	RA HEVC 1x 10-bit base			RA HEVC 2x 8-bit base		
	Y	U	V	Y	U	V
Class A+	-0.8%	-1.4%	-1.9%	-0.4%	-0.5%	-0.9%
Overall (Test vs Ref)	-0.8%	-1.4%	-1.9%	-0.4%	-0.5%	-0.9%
Overall (Test vs single layer)	28.0%	21.3%	-2.1%	20.8%	21.9%	11.4%
Overall (Ref vs single layer)	29.1%	22.9%	-0.5%	21.3%	22.5%	12.4%
EL only (Test vs Ref)	-4.1%	-4.8%	-5.3%	-0.8%	-0.9%	-1.4%
Overall (Test EL+BL vs single EL+BL)	-30.6%	-33.7%	-47.1%	-19.8%	-18.0%	-26.0%
Overall (Ref EL+BL vs single EL+BL)	-30.0%	-32.7%	-46.2%	-19.4%	-17.5%	-25.2%
Enc Time[%]	100.1%			80.9%		
Dec Time[%]	81.6%			71.3%		

Conclusion

- Suggest to keep current syntax lengths of `cm_octant_depth` and `cm_y_part_num_log2` for future extensibility
 - 3D LUT with larger sizes provide additional coding gain
- Allow 8x1x1 LUT size in this version of SHVC
- Suggest to adopt the proposed RD cost based LUT size selection
 - For max 8x8x8 LUT size, compared to SHM6.0 encoder, {-0.7%, -2.1%, -2.5%} reduction for RA 1x and {-0.5%, -1.2%, -1.5%} reduction for RA 2x

Thanks Technicolor for cross-checking!
(JCTVC-R0242)