

JCTVC-R0147:
SCCE5 Test 3.2.1: In-loop color-space transform

Li Zhang, Jianle Chen, Joel Sole, Marta Karczewicz (Qualcomm)
Xiaoyu Xiu, Yuwen He, Yan Ye (InterDigital)

Summary of the proposed method

- In-loop color transform (ICT)

- Lossy

- ✦ YCoCg

- Lossless

- ✦ YCoCg-R

- ICT is applied to residual signals

- CU-level switch on/off

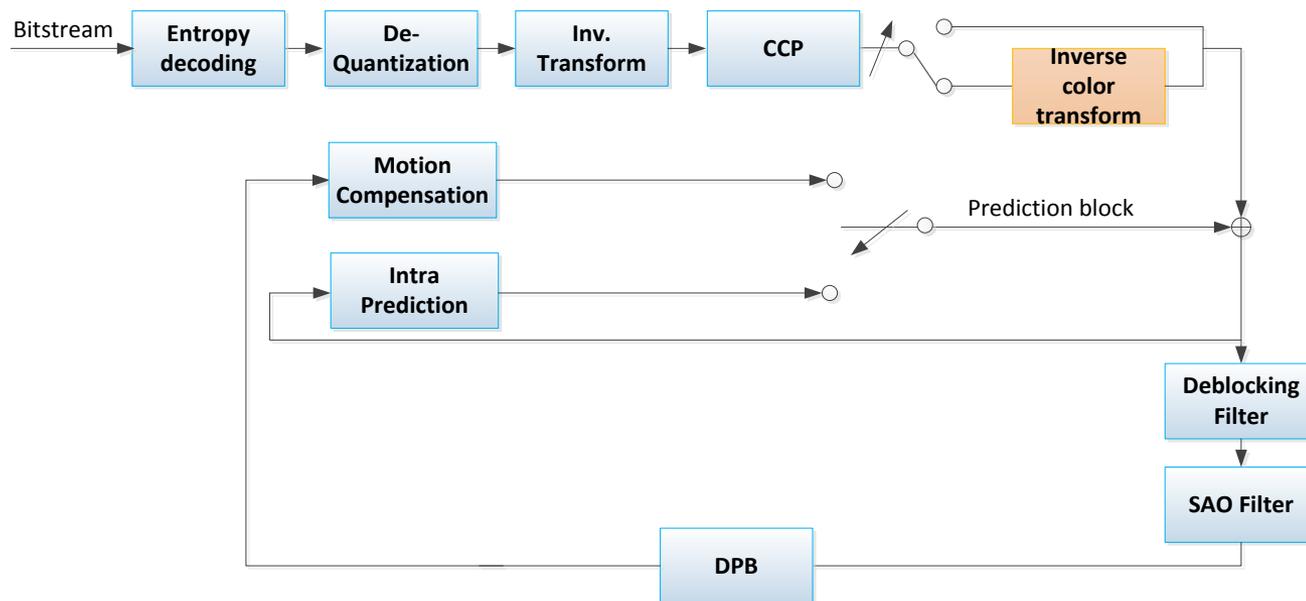
- Non-normative coding tools to further improve the coding efficiency for RGB inputs

- Encoder speedup

In-loop Transform Framework

At the decoder side,

- One module 'inverse color transform' is added after conventional inverse transform and cross-component prediction
- Inverse color transform is applied when the CU-level color transform flag is equal to 1



Transform matrix and quantization parameters

$$\begin{bmatrix} Y \\ Co \\ Cg \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} / 4$$

Forward transform

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} Y \\ Co \\ Cg \end{bmatrix}$$

Inverse transform

- **Non-normalized transform**

- Norm (Y,Cg) = $\sqrt{6}/4 \approx 0.61$
- Norm (Co) = $\sqrt{8}/4 \approx 0.71$

- **Different QP is used in YCoCg space to normalize the transform**

- (QP - 5, QP-3, QP-5) are used for Y, Co, and Cg component, respectively

Non-normative coding tools

- **Modified block vector estimation process for IntraBC**

- Block vectors are estimated in the YCoCg color space for CUs coded with IntraBC and with adaptive color transform on

- **Weighted distortion for chroma components**

- Initialized lambda for chroma components is further adjusted

$$\lambda_{chroma} = \frac{\lambda_{luma}}{W_{QP}} = \lambda_{luma} * 2^{-\text{delta}(QP)/3}$$

wherein delta(QP) is dependent on the input QP, as shown in the following table.

QP	[0, 14]	[15, 29]	[30, 36]	[37, 38]	[39, 40]	[41, 42]	[43, 52]
delta(QP)	0	-1	-2	-3	-4	-5	-6

Encoder speed-up

• Skipping the second pass

- For RGB sequences, the rate-distortion selection is first checked in transformed (YCoCg) space. While for YUV sequences, the first pass is done in original (YUV) domain.
- When one of the following two conditions is true, the second pass is skipped
 - ✦ There is no non-zero coefficient in the first pass coding
 - ✦ The 2nd-pass RD cost of its parent CU with smaller depth is larger than the 1st-pass RD cost

• Intra mode RD selection

- When color transform is enabled, only the RD cost of DM mode is checked

• For intraBC and inter modes

- Motion estimation and motion compensation only performed once

Experimental results

- Whole proposal, full-frame based IntraBC, lossy coding

	All Intra			Random Access			Low delay B		
	G/Y	B/U	R/V	G/Y	B/U	R/V	G/Y	B/U	R/V
RGB, text & graphics with motion, 1080p	-10.7%	-4.6%	-4.8%	-10.1%	-3.4%	-3.6%	-8.7%	-1.9%	-1.7%
RGB, text & graphics with motion,720p	-17.2%	-4.7%	-7.9%	-19.9%	-5.3%	-10.2%	-20.2%	-3.7%	-8.6%
RGB, mixed content, 1440p	-20.9%	-7.7%	-7.7%	-27.6%	-11.0%	-10.4%	-28.2%	-8.4%	-7.3%
RGB, mixed content, 1080p	-15.9%	-3.6%	-4.8%	-19.9%	-5.4%	-6.2%	-20.1%	-1.2%	-1.9%
RGB, Animation, 720p	-25.2%	-17.7%	-14.9%	-25.6%	-16.0%	-11.6%	-24.3%	-11.3%	-5.0%
RGB, camera captured, 1080p	-25.3%	-5.1%	-9.9%	-27.9%	-5.4%	-14.0%	-26.2%	-1.8%	-12.1%
YUV, text & graphics with motion, 1080p	-2.0%	-1.4%	-1.8%	-1.3%	-1.1%	-1.5%	-1.7%	-1.8%	-1.8%
YUV, text & graphics with motion,720p	-1.3%	-1.4%	-2.1%	-1.9%	-1.7%	-2.7%	-2.1%	-2.1%	-3.5%
YUV, mixed content, 1440p	-0.1%	-1.1%	-1.2%	0.2%	-1.1%	-0.9%	0.2%	-1.0%	-1.0%
YUV, mixed content, 1080p	-0.2%	-0.5%	-0.6%	0.1%	-0.2%	-0.7%	0.1%	-1.6%	-1.0%
YUV, Animation, 720p	0.2%	-6.7%	-3.8%	-0.1%	-7.1%	-3.3%	-0.1%	-6.4%	-4.2%
YUV, camera captured, 1080p	0.1%	-0.1%	0.0%	0.4%	-0.1%	0.2%	0.3%	-0.1%	-0.1%
Average RGB	-18.0%	-6.1%	-7.8%	-20.5%	-6.6%	-9.1%	-20.0%	-4.1%	-6.6%
Average YUV	-0.9%	-1.5%	-1.6%	-0.8%	-1.5%	-1.6%	-1.0%	-1.8%	-2.1%
Enc Time[%]		141%			110%			109%	
Dec Time[%]		100%			96%			97%	

Experimental results

- Whole proposal, 2-CTU based IntraBC, lossy coding

	All Intra			Random Access			Low delay B		
	G/Y	B/U	R/V	G/Y	B/U	R/V	G/Y	B/U	R/V
RGB, text & graphics with motion, 1080p	-10.5%	-4.3%	-4.3%	-10.2%	-3.2%	-3.2%	-8.9%	-1.6%	-1.4%
RGB, text & graphics with motion,720p	-17.2%	-4.5%	-7.8%	-19.6%	-5.2%	-9.9%	-20.5%	-3.8%	-8.9%
RGB, mixed content, 1440p	-19.9%	-7.0%	-7.0%	-26.8%	-10.6%	-10.2%	-28.3%	-8.4%	-7.3%
RGB, mixed content, 1080p	-15.6%	-3.4%	-4.7%	-19.3%	-5.2%	-6.1%	-20.4%	-1.5%	-2.2%
RGB, Animation, 720p	-25.2%	-17.7%	-14.9%	-25.6%	-16.0%	-11.6%	-24.3%	-11.4%	-5.0%
RGB, camera captured, 1080p	-25.3%	-5.1%	-9.9%	-27.9%	-5.3%	-14.0%	-26.2%	-1.8%	-12.1%
YUV, text & graphics with motion, 1080p	-1.6%	-1.2%	-1.6%	-1.4%	-1.3%	-1.6%	-1.8%	-1.7%	-1.8%
YUV, text & graphics with motion,720p	-0.9%	-1.1%	-1.7%	-1.8%	-1.7%	-3.0%	-2.3%	-2.2%	-4.0%
YUV, mixed content, 1440p	0.0%	-1.4%	-1.4%	0.1%	-0.9%	-1.0%	0.2%	-1.2%	-0.9%
YUV, mixed content, 1080p	-0.1%	-0.6%	-0.6%	0.0%	-0.6%	-0.8%	-0.2%	-0.7%	-1.4%
YUV, Animation, 720p	0.2%	-6.7%	-3.8%	0.0%	-6.8%	-3.3%	-0.1%	-6.0%	-3.8%
YUV, camera captured, 1080p	0.1%	-0.1%	-0.1%	0.5%	0.1%	0.2%	0.3%	-0.2%	-0.2%
Average RGB	-17.8%	-5.9%	-7.5%	-20.3%	-6.4%	-8.9%	-20.2%	-4.1%	-6.6%
Average YUV	-0.6%	-1.4%	-1.4%	-0.8%	-1.5%	-1.7%	-1.1%	-1.8%	-2.2%
Enc Time[%]	142%			111%			112%		
Dec Time[%]	99%			97%			98%		

Experimental results

- Whole proposal, lossless coding

full-frame based IntraBC

2-CTU based IntraBC

	All Intra		Random Access		Low Delay B		All Intra		Random Access		Low Delay B	
	Bit-rate saving (Total)	Bit-rate saving (Average)										
RGB, text & graphics with motion, 1080p	1.3%	1.4%	0.9%	0.9%	0.9%	0.7%	1.3%	1.4%	0.9%	0.9%	0.9%	0.7%
RGB, text & graphics with motion,720p	1.0%	1.0%	0.8%	1.1%	0.7%	1.1%	1.0%	1.0%	0.8%	1.1%	0.7%	1.1%
RGB, mixed content, 1440p	1.7%	1.7%	1.0%	1.0%	1.1%	1.1%	1.7%	1.7%	1.0%	1.0%	1.1%	1.1%
RGB, mixed content, 1080p	1.5%	1.5%	0.4%	0.4%	0.1%	0.1%	1.5%	1.5%	0.4%	0.4%	0.1%	0.1%
RGB, Animation, 720p	4.3%	4.3%	1.5%	1.5%	1.5%	1.5%	4.3%	4.3%	1.5%	1.5%	1.5%	1.5%
RGB, camera captured, 1080p	0.4%	0.2%	-0.2%	-0.2%	-0.1%	-0.2%	0.4%	0.2%	-0.2%	-0.2%	-0.1%	-0.2%
YUV, text & graphics with motion, 1080p	0.4%	0.5%	0.4%	0.3%	0.4%	0.3%	0.4%	0.5%	0.4%	0.3%	0.4%	0.3%
YUV, text & graphics with motion,720p	0.2%	0.1%	0.5%	0.2%	0.6%	0.2%	0.2%	0.1%	0.5%	0.2%	0.6%	0.2%
YUV, mixed content, 1440p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
YUV, mixed content, 1080p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
YUV, Animation, 720p	1.6%	1.6%	0.3%	0.3%	0.4%	0.4%	1.6%	1.6%	0.3%	0.3%	0.4%	0.4%
YUV, camera captured, 1080p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Enc Time[%]	148%		105%		107%		148%		105%		107%	
Dec Time[%]	102%		94%		99%		92%		94%		99%	

Additional experimental results

- Proposal wo/ non-normative tools, full-frame based IntraBC, lossy coding

	All Intra			Random Access			Low delay B		
	G/Y	B/U	R/V	G/Y	B/U	R/V	G/Y	B/U	R/V
RGB, text & graphics with motion, 1080p	-7.4%	-3.4%	-3.4%	-6.3%	-2.8%	-2.9%	-5.5%	-2.2%	-2.1%
RGB, text & graphics with motion,720p	-14.4%	-5.9%	-8.2%	-15.8%	-5.7%	-9.6%	-16.5%	-5.7%	-9.5%
RGB, mixed content, 1440p	-18.5%	-9.2%	-9.0%	-23.8%	-11.3%	-10.8%	-24.6%	-10.7%	-9.5%
RGB, mixed content, 1080p	-13.5%	-5.3%	-5.9%	-16.1%	-6.2%	-6.3%	-16.4%	-4.2%	-4.6%
RGB, Animation, 720p	-22.5%	-20.3%	-16.9%	-21.6%	-18.4%	-13.4%	-20.1%	-16.4%	-9.1%
RGB, camera captured, 1080p	-21.2%	-6.5%	-11.4%	-25.9%	-7.8%	-17.0%	-23.5%	-4.9%	-16.5%
YUV, text & graphics with motion, 1080p	-2.0%	-1.4%	-1.8%	-1.3%	-1.1%	-1.5%	-1.7%	-1.8%	-1.8%
YUV, text & graphics with motion,720p	-1.3%	-1.4%	-2.1%	-1.9%	-1.7%	-2.7%	-2.1%	-2.1%	-3.5%
YUV, mixed content, 1440p	-0.1%	-1.1%	-1.2%	0.2%	-1.1%	-0.9%	0.2%	-1.0%	-1.0%
YUV, mixed content, 1080p	-0.2%	-0.5%	-0.6%	0.1%	-0.2%	-0.7%	0.1%	-1.6%	-1.0%
YUV, Animation, 720p	0.2%	-6.7%	-3.8%	-0.1%	-7.1%	-3.3%	-0.1%	-6.4%	-4.2%
YUV, camera captured, 1080p	0.1%	-0.1%	0.0%	0.4%	-0.1%	0.2%	0.3%	-0.1%	-0.1%
Average RGB	-15.0%	-7.0%	-8.2%	-16.8%	-7.2%	-9.4%	-16.6%	-6.3%	-8.5%
Average YUV	-0.9%	-1.5%	-1.6%	-0.8%	-1.5%	-1.6%	-1.0%	-1.8%	-2.1%
Enc Time[%]		146%			112%			110%	
Dec Time[%]		101%			98%			97%	

Additional experimental results

- Proposal wo/ non-normative tools, 2-CTU based IntraBC, lossy coding

	All Intra			Random Access			Low delay B		
	G/Y	B/U	R/V	G/Y	B/U	R/V	G/Y	B/U	R/V
RGB, text & graphics with motion, 1080p	-7.2%	-3.4%	-3.3%	-6.3%	-2.8%	-2.9%	-5.8%	-2.3%	-2.2%
RGB, text & graphics with motion,720p	-14.3%	-6.0%	-8.5%	-15.5%	-5.8%	-9.5%	-16.7%	-6.0%	-9.9%
RGB, mixed content, 1440p	-17.4%	-8.4%	-8.3%	-23.1%	-11.1%	-10.6%	-24.7%	-10.9%	-9.8%
RGB, mixed content, 1080p	-12.9%	-5.1%	-5.7%	-15.5%	-6.0%	-6.3%	-16.6%	-5.0%	-4.9%
RGB, Animation, 720p	-22.5%	-20.4%	-16.9%	-21.7%	-18.4%	-13.5%	-20.1%	-16.5%	-9.2%
RGB, camera captured, 1080p	-21.2%	-6.5%	-11.4%	-25.8%	-7.8%	-17.0%	-23.5%	-4.9%	-16.5%
YUV, text & graphics with motion, 1080p	-1.6%	-1.2%	-1.6%	-1.4%	-1.3%	-1.6%	-1.8%	-1.7%	-1.8%
YUV, text & graphics with motion,720p	-0.9%	-1.1%	-1.7%	-1.8%	-1.7%	-3.0%	-2.3%	-2.2%	-4.0%
YUV, mixed content, 1440p	0.0%	-1.4%	-1.4%	0.1%	-0.9%	-1.0%	0.2%	-1.2%	-0.9%
YUV, mixed content, 1080p	-0.1%	-0.6%	-0.6%	0.0%	-0.6%	-0.8%	-0.2%	-0.7%	-1.4%
YUV, Animation, 720p	0.2%	-6.7%	-3.8%	0.0%	-6.8%	-3.3%	-0.1%	-6.0%	-3.8%
YUV, camera captured, 1080p	0.1%	-0.1%	-0.1%	0.5%	0.1%	0.2%	0.3%	-0.2%	-0.2%
Average RGB	-14.7%	-6.9%	-8.1%	-16.6%	-7.2%	-9.3%	-16.7%	-6.5%	-8.7%
Average YUV	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Enc Time[%]		145%			112%			112%	
Dec Time[%]		98%			99%			98%	

Additional experimental results

- Proposal wo/ non-normative tools, lossless coding

full-frame based IntraBC

2-CTU based IntraBC

	All Intra		Random Access		Low Delay B			All Intra		Random Access		Low Delay B	
	Bit-rate saving (Total)	Bit-rate saving (Average)	Bit-rate saving (Total)	Bit-rate saving (Average)	Bit-rate saving (Total)	Bit-rate saving (Average)		Bit-rate saving (Total)	Bit-rate saving (Average)	Bit-rate saving (Total)	Bit-rate saving (Average)	Bit-rate saving (Total)	Bit-rate saving (Average)
RGB, text & graphics with motion, 1080p	0.6%	0.5%	0.8%	0.3%	0.7%	0.2%		0.5%	0.3%	0.6%	0.1%	0.6%	0.1%
RGB, text & graphics with motion, 720p	0.8%	0.9%	0.7%	1.0%	0.7%	1.1%		0.9%	0.9%	0.8%	1.0%	0.7%	1.1%
RGB, mixed content, 1440p	2.0%	2.0%	1.1%	1.1%	1.1%	1.1%		1.7%	1.6%	1.0%	1.0%	1.1%	1.1%
RGB, mixed content, 1080p	1.6%	1.6%	0.4%	0.4%	0.1%	0.1%		1.4%	1.4%	0.4%	0.4%	0.1%	0.1%
RGB, Animation, 720p	4.3%	4.3%	1.5%	1.5%	1.5%	1.5%		4.3%	4.3%	1.5%	1.5%	1.5%	1.5%
RGB, camera captured, 1080p	0.4%	0.2%	-0.2%	-0.2%	-0.1%	-0.2%		0.4%	0.2%	-0.2%	-0.2%	-0.1%	-0.2%
YUV, text & graphics with motion, 1080p	0.4%	0.4%	0.4%	0.3%	0.5%	0.3%		0.4%	0.5%	0.4%	0.3%	0.4%	0.3%
YUV, text & graphics with motion, 720p	0.2%	0.1%	0.5%	0.2%	0.6%	0.2%		0.2%	0.1%	0.5%	0.2%	0.6%	0.2%
YUV, mixed content, 1440p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
YUV, mixed content, 1080p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
YUV, Animation, 720p	1.6%	1.6%	0.3%	0.3%	0.3%	0.3%		1.6%	1.6%	0.3%	0.3%	0.4%	0.4%
YUV, camera captured, 1080p	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Enc Time[%]	148%		119%		109%			146%		107%		107%	
Dec Time[%]	101%		100%		101%			101%		95%		96%	

- Thanks Microsoft for the cross-check (JCTVC-R0111)

Conclusions

- Residual-domain adaptive color transform is proposed
 - Only one inverse color transform module is added at the decoder side
 - additional complexity of inverse color transform is minor since only addition and shift operations are required
- Significant coding gains for RGB inputs are observed, e.g, under full-frame based IntraBC,
 - With non-normative coding tools, the gains of {G, B, R} are:
 - AI: {-18.0%, -6.1%, -7.8%}
 - RA: {-20.5%, -6.6%, -9.1%}
 - LD: {-20.0%, -4.1%, -6.6%}
 - Without non-normative coding tools, the gains of {G, B, R} are:
 - AI: {-15.0%, -7.0%, -8.2%}
 - RA: {-16.8%, -7.2%, -9.4%}
 - LD: {-16.6%, -6.3%, -8.5%}