Results of CE 10.2 on JPEG XS by USTC

wg1m105026

Yao Li, Zhuoyuan Li, Li Li and Dong Liu
Intelligent Visual Data Coding (iVC) Laboratory,
University of Science and Technology of China (USTC)

1. Main goals

This experiment aims to perform to assess the quality and complexity of the IPC screen coding tools as provided in document wg1m104003.

JPEG XS High444.12 Profile and High Throughput JPEG 2000 (HTJ2K) are used as anchor for quality and complexity evaluation and comparison.

2. Test sequences

The following test material compiled from the JVET and JPEG XS test image set are used for the objective quality evaluation. Sequences blacked out are selected from JPEG XS test image set.

Table 1. Test sequences

Class	Name	Frames selected (POC)
TGM	FlyingGraphics_1920x1080	0, 60, 120, 180, 240
	Desktop_1920x1080	0, 120, 240, 360, 480
	Console_1920x1080	0, 120, 240, 360, 480
	ChineseEditing_1920x1080	0, 120, 300, 480, 540
	IPX_XsPowerPoint3_1920x1080	42, 62, 88, 168, 393
	FHG_ScreenStudent_1920x1080	2
	WebBrowsing_1280x720	210
	Map_1280x720	0, 240, 420
	Programming_1280x720	2, 5, 8
	SlideShow_1280x720	1, 4, 5, 6
M	**RICHTER_ScreenContent_4096x2160**	1, 68
	BLENDER_TearsOfSteel_4096x1714	1290, 1332
	APPLE_BasketballScreen_2560x1440	0
	MissionControlClip2_2560x1440	180, 240, 500
	MissionControlClip3_1920x1080	240, 320, 360
A	**BLENDER_Sintel2_4096x1744**	4606, 4685
	ArenaOfValor_1920x1080	0, 360, 420, 540
	Robot_1280x720	0, 360
CC	**ARRI_AlexaHelicopterView_3840x2160**	48
	EBU_PendulusWide_3840x2160	1

3. Objective quality results

RGB PSNR between original sequence and reconstructed sequences are used to evaluate the objective quality of different codec. Average delta PSNR over target rates on different sequences, and average delta PSNR over class TGM and class M on different target rates, between the test codec and JPEG XS

Dept. EEIS, No. 443 Huangshan Road, Hefei 230027, Anhui Province, China

High444.12 Profile, are reported to demonstrate the objective quality ($\Delta PSNR \coloneqq PSNR_{test} - PSNR_{XSHigh}$). Test target rates are 0.75, 1.0, 1.5, 2.0, 3.0 and 4.0.

Encoding parameters and commands for different codecs are listed as follows.

- JPEG XS High444.12 profile, JPEG XS High444.12 profile with IPC
  - DWT 5,2, uniform quantization, separate sign packing, in PSNR-optimized configuration;
  - Lookahead for rate control is one precinct, and buffer model contains 16 lines, as defined in 21122-2

```
jxs_encoder.exe -D -c "p=High444.12;rate=${bpp};nlx=5;nly=2;
qpih=1;fs=1;gains=psnr;budget_lines=16" -w "${img_wdt}" -h
"${img_hgt}" -d "${img_bd}" "${file}" "${strm_path}"
```

- HTJ2K (Kakadu's implementation, version 8.4.1)
  - DWT 5,5 with 9-7 wavelet in a low-latency configuration, with PSNR optimized bounds

```
kdu_compress -i "${file}" -o "${strm_path}" -rate "${bpp}"
Corder=PCRL Clevels=5 Cprecincts=\{64,8192\},\{32,8192\},
\{16,8192\},\{8,8192\},\{4,8192\},\{2,8192\} Cblk=\{8,256\}
Cdecomp=B\(-:-:-\) Qstep=0.0001 -no_weights Scbr=\{1,66\}
Cmodes=HT Cplex=\{6,EST,0.25,-1\} -flush_period 64
```

It should also be noted that the primary goal of JPEG XS is to provide a low-latency lightweight image coding system[1]. Under a 5h2v asymmetric DWT decomposition configuration, the end-to-end latency of a JPEG XS codec can be controlled under 32 lines which was selected as requirement by WG1[1]. This low latency property distinguishes JPEG XS with other "fast" standards or codec implementations.

To evaluate the objective quality of the test codecs in a fair latency setting, we also configure the HTJ2K to a 5h2v 5/3 DWT decomposition. Commands for low latency HTJ2K encoding are as follows, where *the fundamental end-to-end latency for this compression setup is just 24 lines* [2].

```
kdu_compress -i "${file}" -o "${strm_path}" -rate "${bpp}"
Corder=PCRL Clevels=5 Cprecincts=\{8,8192\},\{4,8192\},
\{2,8192\} Cblk=\{4,1024\} Cdecomp=B\(-:-:-\),B\(-:-:-\),
H\(-\) Qstep=0.0001 Catk=2 Kkernels:I2=I5X3 -no_weights
Scbr=\{1,10\} Cmodes=HT Cplex=\{6,EST,0.25,-1\} -flush_period
```

The result of a simplified version of IPC, namely IPC-S, with search range reduced to one precinct and chroma search skipped, is also presented.

Table 2. Average delta PSNR over test rates

Class	Seq	delta PSNR (dB)			
		IPC	IPC-S	HTJ2K low latency	HTJ2K
TCM	FlyingGraphics_1920x1080	0.57	0.44	0.69	1.83
	Desktop_1920x1080	1.43	1.32	0.44	2.29
	Console_1920x1080	1.70	1.57	0.66	0.77
	ChineseEditing_1920x1080	0.28	0.21	0.38	1.85

	IPX_XsPowerPoint3_1920x1080	0.61	0.52	0.66	5.63
	FHG_ScreenStudent_1920x1080	0.70	0.47	-0.24	2.78
	WebBrowsing_1280x720	0.26	0.21	0.92	5.01
	Map_1280x720	0.29	0.21	0.59	2.93
	Programming_1280x720	0.64	0.52	0.56	2.33
	SlideShow_1280x720	0.95	0.82	0.93	4.04
Avg. TCM		0.74	0.63	0.56	2.95
M	RICHTER_ScreenContent_4096x2160	0.87	0.80	0.11	1.49
	BLENDER_TearsOfSteel_4096x1714	0.23	0.19	-0.01	2.09
	BasketballScreen_2560x1440	0.54	0.44	0.07	1.68
	MissionControlClip2_2560x1440	0.31	0.22	0.33	2.52
	MissionControlClip3_1920x1080	0.39	0.28	0.41	2.02
Avg. M		0.47	0.39	0.18	1.96
A	BLENDER_Sintel2_4096x1744	0.05	0.01	-0.18	1.36
	ArenaOfValor_1920x1080	0.08	0.02	0.04	1.25
	Robot_1280x720	0.07	0.03	0.30	1.16
Avg. A		0.07	0.02	0.06	1.26
CC	ARRI_AlexaHelicopterView_3840x2160	0.08	0.03	0.05	0.88
	EBU_PendulusWide_3840x2160	0.01	-0.01	-0.02	1.04
Avg. CC		0.04	0.01	0.01	0.96
Avg.		**0.50**	**0.41**	**0.34**	**2.25**
Avg. TGM & M		**0.65**	**0.55**	**0.43**	**2.62**

Compared with HTJ2K low latency setting, IPC and IPC-S show more significant performance improvements on sequences such as Console_1920x1080, Desktop_1920x1080, primarily because these sequences contain more vertical textures, which can facilitate the prediction in IPC. In contrast, for sequences WebBrowsing_1280x720 and Map_1280x720, the performance of IPC is inferior to HTJ2K low latency settting, as the vertical textures in these sequences are not significant.

Table 3. Average delta PSNR over class TGM

bpp	delta PSNR (dB) over TGM			
	IPC	IPC-S	HTJ2K low latency	HTJ2K
0.75	0.66	0.56	1.49	3.91
1.00	0.73	0.62	1.16	3.64
1.50	0.74	0.63	0.74	3.34
2.00	0.76	0.65	0.54	3.09
3.00	0.77	0.65	-0.01	2.61
4.00	0.79	0.67	-0.57	1.09

Table 4. Average delta PSNR over class M

bpp	delta PSNR (dB) over M			
	IPC	IPC-S	HTJ2K low latency	HTJ2K

0.75	0.62	0.50	0.95	2.70
1.00	0.55	0.47	0.56	2.31
1.50	0.47	0.40	0.17	1.92
2.00	0.43	0.35	-0.03	1.80
3.00	0.38	0.32	-0.23	1.63
4.00	0.36	0.28	-0.32	1.42

For class TGM, the performance of IPC and IPC-S shows a slight increase as the bitrate rises, while class M exhibits a decrease trend. For HTJ2K and its low latency setting, the trend of decreasing performance in higher target rates is particularly evident.

Compared to HTJ2K, the IPC scheme exhibits consistent performance improvement at all target rates.

4. Complexity analysis results

The average number of elementary operations (additions and shifts) per wavelet coefficient, including wavelet transformation, RCT color transformation, quantization, packing (block coder) and (if used) IPC intra-copy search, is utilized for encoder complexity assessment.

- DWT
  - For LeGall 5/3 kernel used in JPEG XS, it has the following form:

$$\begin{cases} c(2n+1) = x(2n+1) - \left\lfloor \frac{x(2n)+x(2n+2)}{2} \right\rfloor (high\ pass) \\ d(2n) = x(2n) + \left\lfloor \frac{c(2n-1)+c(2n+1)+2}{4} \right\rfloor (low\ pass) \end{cases},$$

    The average number of elementary operations per coefficient for one transform is 3 additions and 1 shift, which can be denoted as 3A1S. We use the same notion below for brevity.

    For H1H1, H1L1, L1H1 band, average number of operations for each wavelet coefficient is 6A2S, as one 3A1S for horizontal decomposition and one 3A1S for vertical decomposition; for H2H2, H2L2, L2H2 band, the number is 12A4S. Considering a collection of wavelet coefficients of a 5h2v decomposition, it can be inferred that the average number of operations for each wavelet coefficient is 8A3S:

    Assuming the area of the L5L2 band is 1, then the H4L2 band has an area of 2, H3L2 has an area of 4, and so on, until an area 32 for H1H1/H1L1/L1H1 band and 128 for the coefficients collection. The average number then is:

$$\frac{32\times3\times(6A2S)+8\times3\times(12A4S)+4\times(15A5S)+2\times(18A6S)+1\times2\times(21A7S)}{128} = 8A3S$$

  - For 9/7 wavelet used in HTJ2K, under the assumption that a float point operation can be approximated as one shift operation, it can be concluded that one 9/7 transform consumes 4A2S per coefficient, from figure 1. In a similar way to the derivation of 5/3 DWT, the average number per wavelet coefficient is 11A6S:

$$\frac{256\times3\times(8A4S)+64\times3\times(16A8S)+16\times3\times(24A12S)+4\times3\times(32A16S)+1\times4\times(40A20S)}{1024} = 11A6S$$

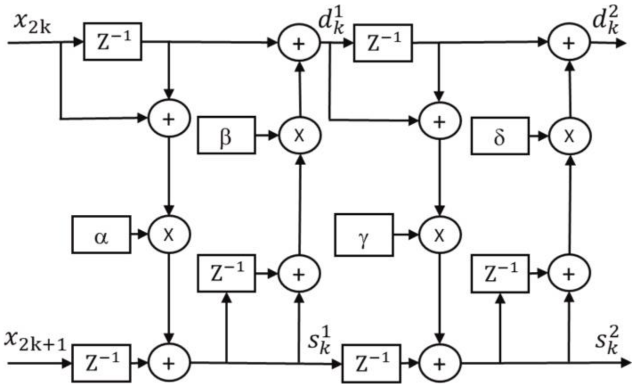

Figure 1. Lifting Scheme implementation of 9/7 DWT[3]

- RCT
  - For RGB to YCgCo conversion, the total number of operations for the total 3 components is 4A2S per pixel:

Table5. Forwards reversible multiple component transformation[4]

Syntax	Notes
`forward_rct() {`	
`for(y=0;y<Hf;y=y+1) {`	Loop over rows of the image
`for(x=0;x<Wf;x=x+1) {`	Loop over the columns of the image
$i_0 = \Omega[0,x,y]$	
$i_1 = \Omega[1,x,y]$	
$i_2 = \Omega[2,x,y]$	Retrieve input components
$o_0 = (i_0+2\times i_1+i_2)>>2$	Compute the luma component
$o_1 = i_2 - i_1$	Compute the Cb chroma component
$o_2 = i_0 - i_1$	Compute the Cr chroma component
$O[0,x,y] = o_0$	Assign the luma output
$O[1,x,y] = o_1$	Assign Cb
$O[2,x,y] = o_2$	Assign Cr
}	End of loop over columns
}	End of loop over rows
}	

  - And thus, the average number of operations per wavelet sample is 2A1S
- Quant
  - For codec without prediction, the average number is 1S; for codec with prediction (TDC, IPC), the average number is 2S, one for quantization and one for reverse quantization to generate the reconstruction.
- Search

- For JPEG XS TDC444.12 profile, the number for prediction decision per wavelet sample is 3A, as one for prediction residue calculation, one for residue cost accumulation and one for original coefficients cost accumulation
- For JPEG XS High444.12 profile with IPC, the number is 65A:
  - Available reference precincts are 2 precincts for each precinct in a slice, on average. With 16 candidates per reference precinct, there are total 32 available search candidates each IPC unit;
  - For each candidate, a sample in an IPC unit will have 2 addition operations, as one for prediction residue calculating and one for residue cost accumulating. That makes 2×32 addition operations each sample;
  - Plus one additional addition for original coefficients cost accumulating, the total number is 65A;
  - IPC-S: This number is defined on one component and when chroma skipped search is applied, this number could be reduced to 22A, based on the average across the three channels. With the reference precinct reduced to one, the number is further reduced to 11A.

- Packing
  - We added code to the OpenJPH source code and JPEG XS 2nd edition software to count addition and shift operations in the block coder module. The average value over the entire frame *APPLE_BasketBallScreen_2560x1440p_60_8b_sRGB_444_001* is reported to reflect the complexity of the block coder.
  - As this number may vary with the target coding rate, we set the target rate to 4.0 bpp to reflect the complexity in worst case.

Table 6. The number of operations per wavelet coefficient. "A" denotes addition and "S" denotes shift

Modules	JPEG XS High444.12	JPEG XS TDC444.12	JPEG XS High444.12 IPC	JPEG XS High444.12 IPC-S	HTJ2K	HTJ2K low latency
RCT	2A1S	2A1S	2A1S	2A1S	2A1S	2A1S
DWT	8A3S	8A3S	8A3S	8A3S	11A6S	8A3S
Quant	1S	2S	2S	2S	1S	1S
Prediction decision	-	3A	65A	11A	-	-
Block coder	3A3S	3A3S	3A3S	3A3S	5A5S	3A3S
Total	13A8S	16A9S	78A9S	24A9S	18A13S	13A8S

Dept. EEIS, No. 443 Huangshan Road, Hefei 230027, Anhui Province, China

References

[1] ISO/IEC JTC 1/SC 29/WG 1, “Use cases and requirements for JPEG XS v3.2,” output document wg1n100633, 2023.

[2] Kakadu Software, “Usage Examples for Demonstration Applications Supplied with Kakadu V8.4.1.”

[3] Naseer, Raja Arslan, et al. "VLSI architecture design and implementation of 5/3 and 9/7 lifting Discrete Wavelet Transform." Integration 87 (2022): 253-259.

[4] Th. Richter, T. Bruylands: “Information technology — JPEG XS low-latency lightweight image coding system, Part 1: Core coding system”, published as ISO/IEC 21122-1, (2022).

Dept. EEIS, No. 443 Huangshan Road, Hefei 230027, Anhui Province, China