Intra Pattern Copy for Screen Content Image Coding in JPEG XS

Yao Li, Zhuoyuan Li, Li Li and Dong Liu

Intelligent Visual Data Coding (iVC) Laboratory,

University of Science and Technology of China (USTC)

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

1. Abstract

This document describes an intra-pattern copy (IPC) tool proposal for JPEG XS, to improve the compression efficiency on screen contents coding (SCC). Wavelet coefficients in a precinct are predicted by copying from reconstructed coefficients of the encoded precincts that have a similar pattern. Prediction residuals instead of original coefficients are sent into subsequent rate allocation, quantization and packing modules. The prediction is performed at a spatial position-band group combined unit level, and for each unit, an optimal offset vector is searched and signaled to indicate the prediction. This proposal achieves an average of 0.66dB delta-PSNR promotion on SCC test images with an approximate 200% encoder time complexity and 100% decoder time complexity.

2. Introduction

Temporal Differential Coding (TDC) profile is introduced in JPEG XS 3rd edition to improve the coding efficiency for screen content image sequences efficiently. The TDC profile tool gradually refines the quality of the static background regions and after several image transmissions the reconstructed images can reach a very high quality. As a trade-off for the performance improvement, a framebuffer and prediction loop need to be introduced in the codec. An insight on this update on the codec architecture is that we can reuse this buffer and prediction loop to perform intra-frame prediction to further improve the coding efficiency for screen content coding (SCC) images and SCC image sequences, and thus we propose an intra pattern copy (IPC) coding tool. As the screen contents typically contain a lot of repeated patterns, especially for low-frequency decomposition bands of the discrete wavelet transform (DWT) in JPEG XS, the IPC can efficiently remove the redundancy and save coding bitrate.

Considering the low-latency and low-complexity features of the JPEG XS standard, the same as the TDC profile, the IPC prediction should also be performed in the wavelet domain. In this way, no wavelet inverse transform is required at the encoder side, and hence, no additional complexity, latency, and computational errors are introduced.

3. Method

3.1 Global description

The proposed intra pattern copy (IPC) tool is implemented on JPEG XS 2nd edition. A JPEG XS encoder and decoder with IPC extension is illustrated in Figure 1 and Figure 2, respectively.

The input image lines undergo revertible color transform (RCT) and DWT to generate wavelet coefficients which are then divided into precincts. In IPC, the neighboring wavelet coefficients in a precinct are further grouped into IPC processing units, in a cross-decomposition band approach. For each IPC unit, a prediction is retrieved by copying from reconstructed values with a similar pattern and residuals are then calculated by subtracting predictions from the original values. The codec can code either the original values or the IPC prediction residuals, depending on their coding costs, which are estimated using the bit plane count sums in this proposal. The encoded wavelet coefficients are reconstructed and buffered to provide references. An offset vector, or namely a block vector (BV), is signaled to indicate the prediction retrieval and a multi-band synchronized pattern search is applied to

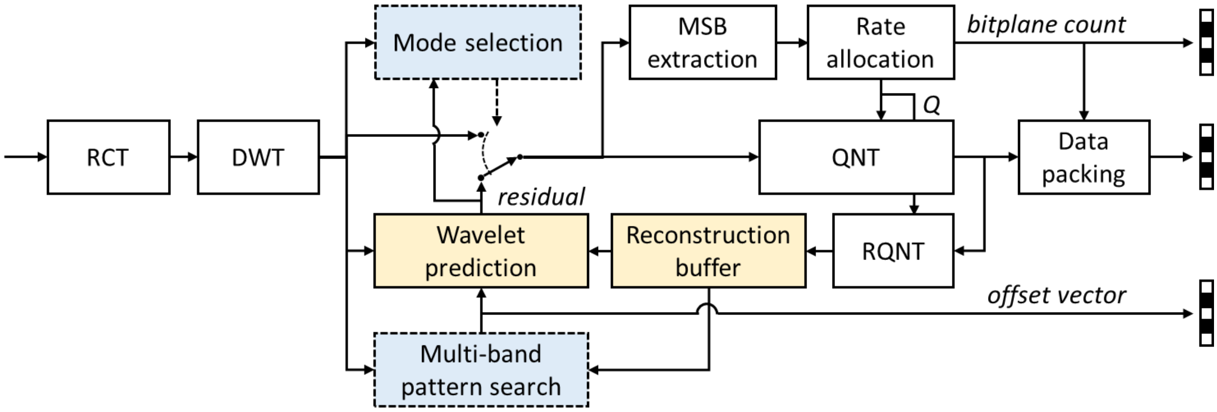

Figure 1 Illustration of a JPEG XS encoder with intra pattern copy mechanism (blocks marked with yellow and blue are newly introduced).

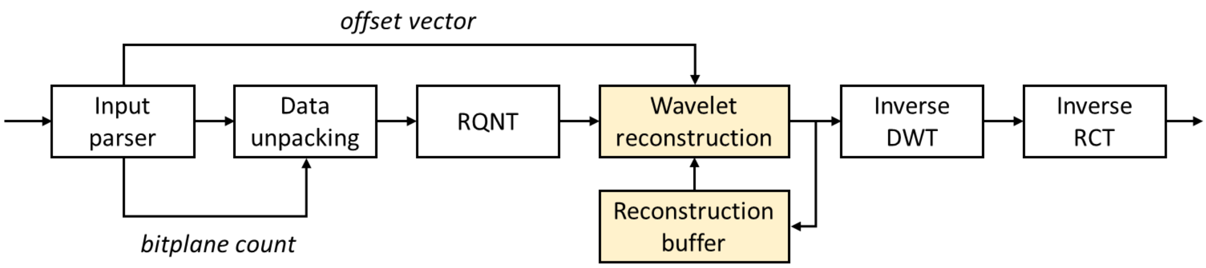

Figure 2 Illustration of a JPEG XS decoder with intra pattern copy mechanism (blocks marked with yellow are newly introduced).

find the optimal offset vectors.

Considering the constraints of independently decodable regions, the search and prediction are limited within a slice. After the encoding of a slice, the reference buffer is refreshed, and thus only a relatively small reference buffer is required, and the search range is also limited.

3.2 IPC unit and IPC group

As different regions of a screen content images normally have different patterns and textures, e.g., text area with rich, complex patterns and background area with simple and uniform patterns, the effectiveness of the proposed IPC method will vary in different image spatial regions. A spatial wavelet coefficient grouping and navigating unit, named *IPC unit*, is introduced to enable content adaptive IPC prediction selection. As illustrated in Figure 3, a hypothetical IPC unit in the pixel domain will correspond to a collection of wavelet coefficient sets of different bands, and each wavelet coefficient set of a band can be seen as a wavelet IPC unit. Different wavelet IPC units in different bands could have different sizes and positions according to their decomposition depths.

Another observation is that different decomposition bands will also have different characteristics, and the utilization of IPC should make corresponding adaptions. Based on the following two insights we further propose a decomposition band grouping and navigating unit, named *IPC group*, to achieve band characteristics adaptive IPC prediction. Firstly, low-frequency bands are low-pass filtered and normally contain smooth, simple patterns and, hence, are easier to predict. Secondly, horizontal decomposition bands mainly contain vertical textures and can be better predicted using the above reconstructed

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

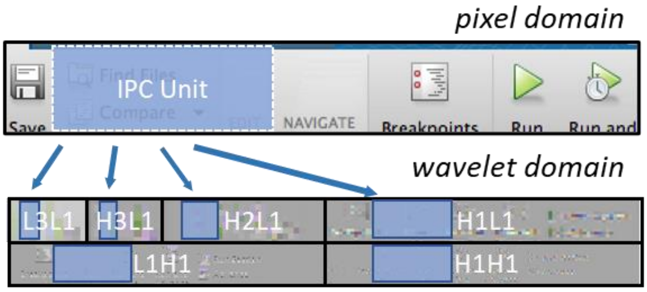

Figure 3 Illustration of the spatial grouping unit, IPC unit (3 horizontal 1 vertical decomposition)

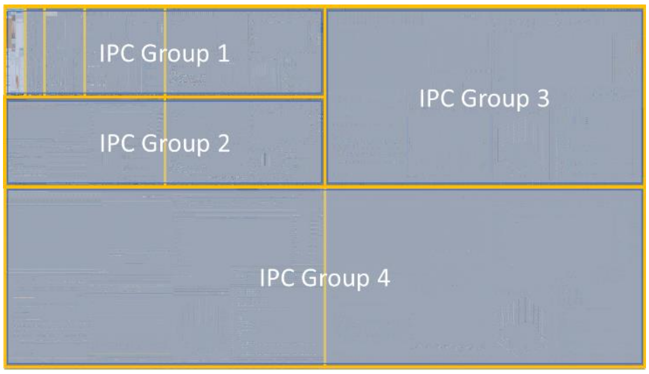

Figure 4 Illustration of the decomposition band grouping unit, IPC group (blocks with yellow boundaries denote different decomposition bands, and blocks marked with blue denote different IPC groups, 5 horizontal 2 vertical decomposition)

values than the vertical decomposition bands. Bands in the same IPC group share the same IPC selections and offset vectors.

The selection and prediction of IPC are performed in a way that jointly considers the image characteristics spatial variance and decomposition band variance, i.e., coefficients at different IPC units or different IPC groups will have different IPC selections and predictions. During the traversal process of the wavelet coefficients to be coded, we adopt an IPC unit first then IPC group order, which means we regard the IPC group index as a fast variable and the IPC unit as a slow variable.

In this proposal, the size of the IPC unit is set to be 128x4 in image pixel resolution. For a 5-horizontal 2-vertical decomposition setting, the IPC unit size is 4x1 for L5L2 band, correspondingly. In Figure 4, a possible IPC group design for 5-horizontal, 2-vertical decomposition is illustrated. Bands of different color components at the same decomposition level are in the same IPC group.

3.3 Wavelet prediction and reconstruction

The wavelet prediction and reconstruction process, performed in the units of IPC unit and IPC group, can be demonstrated by pseudo codes in Table 1 and Table 2, respectively. The IPC units indicating spatial positions and IPC groups indicating band characteristics are traversed. For each IPC unit, the offset vector of an IPC group is scaled to adapt to the different sizes of different bands, i.e., the prediction of different bands in an IPC group is synchronized.

An IPC unit in a specific band can only take references from the same band of the reference precincts. the IPC prediction and residual calculation are conducted in an appoint-wise approach by subtracting prediction values from original values.

Table 1 Pseudo code of wavelet prediction

```
reference are Ω;
for ipc_unit x in all ipc_units {
   for ipc_group φ in all ipc_groups {
      offset vector ov;
      for band b in all bands of ipc_group φ {
         ov_b = ov.scale(φ, b);
         prediction(x, b) = retrieve_prediction(x, ov_b, Ω, b)
         residual(x, b) = coefficient(x, b) - prediction(x, b);
      }
   }
}
```

Table 2 Pseudo code of wavelet reconstruction

```
reference area Ω;
for ipc_unit x in all ipc_units {
   for ipc_group φ in all ipc_groups {
      offset vector ov;
      for band b in all bands of ipc_group φ {
         ov_b = ov.scale(φ, b);
         prediction(x, b) = retrieve_prediction(x, ov_b, Ω, b)
         coefficient(x, b) = residual(x, b) + prediction(x, b);
      }
   }
}
```

3.4 Multi-band synchronized pattern search and prediction

While the encoder optimization schemes are normally not within the scope of standards, we give a brief description of the mode selection and pattern search methods used in this proposal. Bit plane count (GCLI) sum is adopted as the metric for IPC mode selection and IPC offset vector searching. The offset vector that yields the minimum prediction residual GCLI sum is searched in a certain search region. The minimum prediction residual GCLI sum is compared to the wavelet coefficient original value GCLI sum to decide whether the IPC mode is selected. As an IPC group may contain coefficients from multiple bands, for each IPC unit in an IPC group, the pattern search and GCLI sum calculation is conducted in a multi-band synchronized approach. This can be formulated as follows.

$$R_{IPC} = \min_{\boldsymbol{OV} \in \Omega} \left(\sum_{b \in \mathcal{B}} \mathrm{GCLI}\left(\boldsymbol{x}[b] - \hat{\boldsymbol{x}}_{\boldsymbol{OV}}[b] \right) + R_{\boldsymbol{OV}} \right) \tag{1}$$

$$R_{ORIG} = \sum_{b \in \mathcal{B}} \mathrm{GCLI}\left(\boldsymbol{x}[b] \right) \tag{2}$$

$$Flag_{IPC} = \begin{cases} 1, & R_{IPC} < R_{ORIG} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Where $\boldsymbol{OV}$ and Ω stands for offset vector and intra pattern search range, respectively. b denotes a band index in IPC group $\mathcal{B}$. $\boldsymbol{x}[b]$ and $\hat{\boldsymbol{x}}_{\boldsymbol{OV}}[b]$ represent the original coefficient values and reconstructed reference values of band b for current IPC unit, respectively. The coding of offset vectors takes a fixed-length approach and the corresponding coding rate $R_{\boldsymbol{OV}}$ can be precisely calculated.

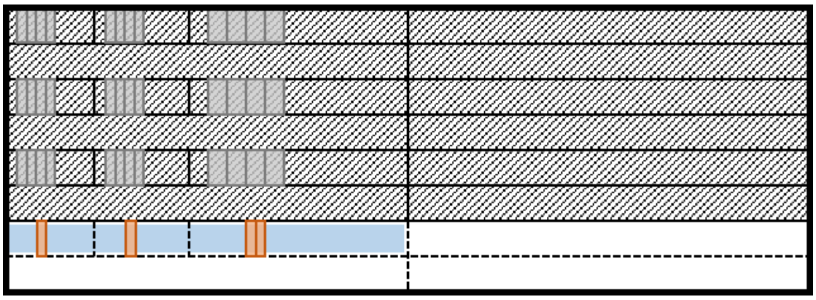

Figure 5 Illustration of the IPC search range (precincts in black shadows are buffered reconstructed references, coefficients in orange are IPC units to be coded in an IPC group, as marked in blue color, and coefficients in grey are feasible search area)

The encoding of offset vectors takes a fixed-length approach, and hence, the search range is decided by the offset vector signaling bits and the slice height. For an IPC unit, in the vertical direction, up to 3 precincts above can be searched, under the constraint of the slice boundary. In the horizontal direction, up to 2 IPC unit width positions to the left and up to 2 IPC unit width positions to the right can be searched. This contributes to a total of 12×IPC unit width candidates. It should be noted that here the search range and IPC unit width are defined on the 0^{th} band, e.g., 5L2L for 5-horizontal 2-vertical decomposition, and the search stride 1 in the 0^{th} band could be scaled for other bands correspondingly, when calculating the GCLI sum in formula(2), e.g., the stride is 4 for 3H2L. The search range is illustrated in Figure 5.

3.5 Codestream syntax

In the precinct header, an IPC flag is added to indicate whether to code the IPC prediction residual or the original coefficients for each IPC unit in an IPC group. This will additionally consume the header bitrate 0.00781 bpp for the IPC unit and IPC group setting in this proposal.

The encoding of offset vectors takes a fixed-length approach. An additional 2 bits for vertical direction and 4 bits for horizontal direction are signaled in the data sub-packet for each IPC that uses the IPC prediction mode.

4. Performance

Test sequences are selected from the JPEG XS and JVET SCC test sequences, with class TGM (text and graphics with motion), class M (mixed content), class A (animation) and class CC (camera-captured content). Sequences of class TGM and class M from JPEG XS are blacked out.

The proposed IPC tool is integrated into JPEG XS reference software 2^{nd} edition. The encoder is configured as DWT 5-horizontal, 2-vertical decomposition, separate sign packing, deadzone quantization and the slice height is 16 lines. Tested target bpps are 0.75, 1.0, 1.5, 2.0, 3.0 and 4.0.

Δ-PSNR compared to JPEG XS 2^{nd} edition High profile is used to evaluate the coding efficiency. As for the complexity evaluation, the software's encoding decoding time is provided as a reference. The results for each sequence and each bitrate are shown in Table 3, Table 4 and Table 5, respectively. The results show that the proposed IPC tool can achieve certain gains for all sequences and all bitrates.

Table 3 Average Δ PSNR and complexity of proposed IPC tool on each sequence

Summary / Sequence				
Class	Seq	Avg. Δ PSNR	EncT Ratio	DecT Ratio
TGM	FlyingGraphics_1920x1080	0.62	205%	101%
	Desktop_1920x1080	1.43	203%	99%
	Console_1920x1080	1.69	199%	101%
	ChineseEditing_1920x1080	0.30	203%	101%
	IPX_XsPowerPoint3_1920x1080	0.54	200%	103%
	FHG_ScreenStudent_1920x1080	0.72	210%	102%
	WebBrowsing_1280x720	0.30	187%	99%
	Map_1280x720	0.30	186%	99%
	Programming_1280x720	0.68	186%	98%
	SlideShow_1280x720	0.97	193%	101%
Avg. TGM		0.75	197%	100%
M	**RICHTER_ScreenContent_4096x2160**	0.88	226%	114%
	BLENDER_TearsOfSteel_4096x1714	0.22	234%	105%
	APPLE_BasketballScreen_2560x1440	0.57	201%	102%
	MissionControlClip2_2560x1440	0.34	212%	101%
	MissionControlClip3_1920x1080	0.41	207%	99%
Avg. M		0.48	216%	104%
A	BLENDER_Sintel2_4096x1744	0.04	223%	105%
	ArenaOfValor_1920x1080	0.09	218%	107%
	Robot_1280x720	0.07	201%	98%
Avg. A		0.07	214%	103%
CC	ARRI_AlexaHelicopterView_3840x2160	0.08	231%	98%
	EBU_PendulusWide_3840x2160	0.01	234%	101%
Avg. CC		0.04	232%	100%
Avg.		**0.51**	**208%**	**102%**
Avg. TGM M		**0.66**	**203%**	**101%**

Table 5 Δ PSNR and complexity of proposed IPC tool on each bitrate (Class TGM)

Summary / bpp		TGM	
bpp	Δ PSNR	EncT Ratio	DecT Ratio
0.75	0.65	180%	99%
1.00	0.76	189%	101%
1.50	0.75	196%	98%
2.00	0.78	200%	101%
3.00	0.77	206%	102%
4.00	0.82	212%	100%

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

Table 6 Δ PSNR and complexity of proposed IPC tool on each bitrate (Class M)

Summary / bpp		M	
bpp	Δ PSNR	EncT Ratio	DecT Ratio
0.75	0.63	199%	105%
1.00	0.56	207%	106%
1.50	0.49	210%	102%
2.00	0.46	218%	99%
3.00	0.41	231%	106%
4.00	0.37	232%	106%

5. Complexity Analysis

The additional latency and memory cost introduced by this proposal mainly come from the added modules as the blocks marked in yellow and blue in Fig. 1 and Fig. 2.

- Reconstruction buffer

Once a precinct is reconstructed, it could be directly written to the reconstruction buffer and be ready for the encoding/decoding of the next precinct, and hence, no significant latency is imposed for both the encoder and decoder side.

For High profile with 2 vertical decompositions and a slice height of 16 lines, the required reference buffer size is 12 lines of coefficients.

- Wavelet prediction/reconstruction

The encoder wavelet prediction and the decoder wavelet reconstruction are placed in the coding pipeline between the DWT and rate control/inverse quantization. This will impose a certain latency. At maximum (4×image_width/128) intra-prediction units need to be processed, depending on the mode selections, with each unit having a maximum size of 64x2 coefficients. This process can be parallelized.

It should be noted that wavelet reconstruction is also performed at the encoder side, but it is located outside the encoding data path and could be parallelized with the color transform and DWT of the next precinct.

The wavelet prediction and reconstruction mainly involve addition operation and no significant memory is required.

- Multi-band pattern search

For each intra-prediction unit, a maximum of 48 offset vector candidates need to be checked. Wavelet prediction and the cost (bit-plane count of the prediction residue) calculating for each offset vector candidate are performed. In total (4×W/128, W denotes the image width) intra-prediction units need to be searched for each precinct. This will impose a certain latency but can be parallelized.

A relatively small memory buffer is needed to store the temporary prediction residue and the cost.

- Mode selection

The cost (bit-plane count) of the raw wavelet coefficients in an intra-prediction unit needs to be calculated and compared with the recorded best intra-prediction cost. In total (4×image_width/128) intra-prediction units need to be processed for each precinct. This will impose a certain latency but can be parallelized.

A relatively small memory buffer is needed for the cost comparison.

Additional glue resources are required to integrate these proposed modules into JPEG XS codec.

Table 7 overviews the additional complexity and sources of the proposed IPC tool.

Table7 Additional complexity and sources of the proposed IPC tool

Module	Operation	Latency	Memory
Reconstruction buffer	Memory copy	~0	~12 lines
Wavelet prediction/reconstruction	Adder	Addition of a line of wavelet coefficients	~0
Multi-band pattern search	Adder, bit-plane count extractor, comparator	Check 48 candidates for (4xW/128) units, (one wavelet prediction and cost calculating per check)	~64x2 coefficients
Mode selection	Adder, bit-plane count extractor, comparator	Compare cost for (4xW/128) units	~0

6. Proposals

This document proposes a new screen content coding tool, namely IPC. Specifically, it involves the following adaptions to JPEG XS 2nd High profile from the perspective of the decoder view.

- The reconstructed wavelet coefficients of all precincts within a slice, except for the last precinct, are stored in a buffer to provide references for subsequent precincts. The reference buffer is refreshed when decoding a new slice.
- For a precinct to be decoded, the reconstructed wavelet coefficients can take two forms. They can either be the dequantized values of the coefficients decoded from the data subpacket alone, or the sum of the dequantized values from the data subpacket and the intra-prediction values obtained from the reference buffer.
- The selection of the wavelet reconstruction method is performed at the level of a pre-defined spatial-band unit, depending on the IPC mode flag parsed for each spatial-band unit from the precinct header stream.

- For each spatial-band unit that chooses to use intra-prediction, the prediction values are retrieved from the reference buffer, under the guidance of an offset vector parsed from the data packet stream.

The encoder should properly select the coding mode and the offset vector for each spatial-band unit, and internally perform reconstruction to build an exactly synchronized reference buffer with the decoder.

This proposal is well-compatible with the JPEG XS framework and can take advantage of the design in the TDC profile in several additional modules.

- Color transform and wavelet decomposition remain the same, as in the JPEG XS High profile.
- Rate allocation is not impacted, and only the input of the rate allocation module is changed.
- Entropy coding is unchanged, and data packing basically follows the JPEG XS 2nd edition with increased precinct header flag and offset vector syntaxes.
- The design of the reference buffer and potential reference buffer compression methods can be reused from the JPEG XS 3rd edition TDC profile but with a smaller buffer size.
- Algorithms for mode decision as well as the hardware design, can also be referenced from the TDC profile.

We propose to adopt the IPC tool and develop the JPEG XS 4th edition on top of it.

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