

Spatially-Aligned Chroma from Luma Prediction for Lossless JPEG XS Raw Image Compression

Taizo Suzuki, *Senior Member, IEEE*, Soma Yokota, and Masaki Onuki

Abstract—This study proposes a Chroma from Luma (CfL)-enhanced Star-Tetrix transform (STT), referred to as CfL-STT, for improving raw image compression in JPEG XS. The proposed CfL-STT integrates CfL prediction into the STT to predict chroma components from the luma component in CFA-sampled raw images. Unlike conventional CfL prediction designed for full-color images, the proposed method employs spatially-aligned luma samples obtained via linear interpolation along the horizontal and vertical directions to match the chroma sampling grid. This spatial alignment suppresses high-frequency sensor noise while preserving cross-channel correlation, resulting in a more decorrelated $Y\Delta DuDv$ color space. The proposed method was implemented in the JPEG XS reference software and evaluated on raw image datasets. Experimental results demonstrate that a direct application of CfL prediction yields image-dependent performance and may degrade coding efficiency due to the lack of spatial alignment, whereas the proposed CfL-STT consistently improves coding efficiency in lossless raw image compression while preserving exact reversibility.

Index Terms—Raw image compression, lossless compression, JPEG XS, Star-Tetrix transform, Chroma from Luma prediction.

I. INTRODUCTION

FULL-COLOR (RGB) images commonly used today are obtained by applying a simple yet inherently *lossy* camera processing pipeline, including black level correction, white balance, demosaicing, and gamma correction, to raw images sampled by a color filter array (CFA), which provides only one red, green, or blue component per pixel. However, with the growing demand for flexible post-editing, there is increasing interest in both lossless and lossy compression of CFA-sampled raw images (hereinafter referred to as raw images). Consequently, compression-first approaches are gaining importance alongside demosaicing-first approaches such as JPEG [1], JPEG 2000 [2], and JPEG AI [3]. This trend has been recognized in JPEG XS 2nd Edition, which introduced compression-first raw image coding as an optional operating mode [4]. In this study, we focus on the Bayer CFA, the most widely adopted RRGB layout (Fig. 1).

Understanding spectral-spatial transforms is essential for effective raw image compression [5–12]. Such transforms map CFA-sampled RRGB data into a decorrelated $Y\Delta UV$ representation, typically expressed as $Y\Delta CbCr$ or $Y\Delta CoCg$, consisting of luma, difference luma (green), and two chroma components. Among them, the Star-Tetrix transform (STT) [10], built on three lifting-based 5/3-tap wavelet transforms, has

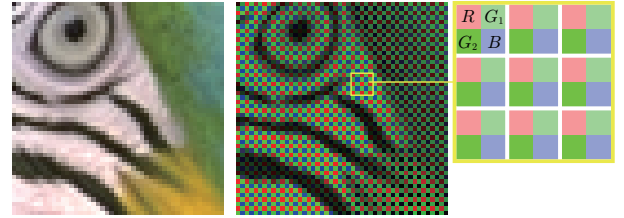


Fig. 1. Bayer pattern of a region from *Parrot* in the Kodak images dataset [19] (each 2×2 pixel block corresponds to a macropixel): (left) RGB full-color image and (right) simulated raw image with corresponding diagram.

been adopted in JPEG XS 2nd Edition for raw image compression due to its strong decorrelation capability and low computational cost. In parallel, the AOMedia Video 1 (AV1) [13] and the JPEG XL [14] introduced Chroma from Luma (CfL) prediction [15], which linearly predicts chroma components from luma to further reduce spectral redundancy in full-color images. Related cross-component prediction tools have also been studied in the context of H.266/VVC, including subset-based cross-component linear model prediction [16] and adaptive chroma prediction based on luma differences [17]. However, raw images exhibit structured sampling patterns and nonuniform sensor noise, which violate the assumptions of conventional CfL prediction and degrade its prediction accuracy when directly applied in the raw domain. Moreover, although recent learned image codecs and large language model-based predictors [18] achieve impressive rate-distortion performance at low bitrates, they primarily target semantic compression, whereas raw image coding requires faithful signal reconstruction with low complexity and strict reversibility.

This study proposes a CfL-enhanced STT (CfL-STT), in which CfL prediction is integrated into the STT to predict chroma components from the luma component. Unlike conventional CfL prediction, which assumes co-located color components, the proposed method employs spatially-aligned luma samples obtained via linear interpolation along the horizontal and vertical directions to match the chroma sampling grid. This spatial alignment not only improves prediction robustness in CFA-sampled raw images but also introduces an implicit low-pass filtering effect that suppresses high-frequency sensor noise while preserving cross-channel correlation. As a result, the $Y\Delta UV$ representation produced by STT is transformed into a more decorrelated $Y\Delta DuDv$ color space, where Du and Dv denote chroma residuals after spatially-aligned CfL prediction. Since the proposed method relies only on interpolation and simple linear regression, it incurs limited additional computational complexity compared with the baseline STT. Experimental results using the JPEG XS reference software

T. Suzuki and S. Yokota are with the Institute of Systems and Information Engineering, University of Tsukuba, Ibaraki 305-8573, Japan (e-mail: {taizo.yokota}@cs.tsukuba.ac.jp).

M. Onuki is with SonicAI Inc., Tokyo 108-0073, Japan (e-mail: monuki@sonicai.jp).

demonstrate that the proposed CfL-STT consistently improves compression performance in lossless raw image coding.

II. REVIEW AND PRELIMINARIES

A. Star-Tetrix Transform (STT)

The Star-Tetrix transform (STT) [10] is a fully reversible multi-component lifting transform designed to decorrelate both spectral and spatial redundancies inherent in Bayer CFA data. It can be regarded as a CFA-domain extension of the JPEG 2000 reversible color transform (RCT) [12], where bilinear interpolation and spatial prediction are incorporated into a unified lifting structure.

Figure 2 illustrates the implementation of STT. Let R , G , and B denote CFA samples at red, green, and blue sensor positions, and let the superscripts l , r , t , and b indicate the neighboring samples to the left, right, top, and bottom, respectively. At red and blue positions, the green signal is bilinearly interpolated from its four neighboring samples and subtracted from the measured value to form chroma components¹:

$$U = B - \left\lfloor \frac{G_2^l + G_2^r + G_1^t + G_1^b}{4} \right\rfloor, \quad (1)$$

$$V = R - \left\lfloor \frac{G_1^l + G_1^r + G_2^t + G_2^b}{4} \right\rfloor, \quad (2)$$

where $\lfloor \cdot \rfloor$ denotes a floor function. This step corresponds to the first lifting stage of a YUV-like transform that is adapted to the CFA sampling structure. By forming differences between the measured red and blue samples and the bilinearly interpolated green signal, the strong inter-channel correlation between color components is effectively removed.

Each green sample is then updated using a weighted average of the surrounding chroma components. Let w_r and w_b be non-negative integer white-balance exponents that represent powers of two derived from camera metadata. Two luma components are generated as

$$Y_1 = G_1 + \left\lfloor \frac{2^{w_r}(V^l + V^r) + 2^{w_b}(U^t + U^b)}{8} \right\rfloor, \quad (3)$$

$$Y_2 = G_2 + \left\lfloor \frac{2^{w_r}(V^t + V^b) + 2^{w_b}(U^l + U^r)}{8} \right\rfloor. \quad (4)$$

These equations perform color decorrelation and white-balance adaptation simultaneously. The powers-of-two weights 2^{w_r} and 2^{w_b} shift the effective luma axis toward the camera white point, which improves decorrelation between luma and chroma components. This study sets $w_r = w_b = 0$ for simplicity.

The next lifting stage introduces spatial decorrelation between the two luma fields. A local prediction of Y_1 is formed from the surrounding Y_2 samples:

$$\Delta = Y_1 - \left\lfloor \frac{Y_2^{l,t} + Y_2^{r,t} + Y_2^{l,b} + Y_2^{r,b}}{4} \right\rfloor. \quad (5)$$

This term represents a high-frequency spatial luma difference that is analogous to a diagonal difference signal. It captures

the mismatch between the two interleaved luma grids and is therefore well suited for efficient entropy coding.

Finally, the remaining luma field is updated using the neighboring Δ values as

$$Y = Y_2 + \left\lfloor \frac{\Delta^{l,t} + \Delta^{r,t} + \Delta^{l,b} + \Delta^{r,b}}{8} \right\rfloor. \quad (6)$$

After these four lifting steps, the CFA image is mapped into four decorrelated components Y , Δ , U , and V . Y represents low-frequency luma component, U and V represent chroma components, and Δ represents a spatial luma difference that already exhibits high-pass characteristics. Importantly, all steps are defined in an integer lifting framework, ensuring exact reversibility, which is essential for lossless raw image compression.

B. Chroma from Luma (CfL) Prediction

Chroma from Luma (CfL) prediction [15] is a linear regression model that predicts chroma components from the luma component in full-color image compression, thereby reducing residual spectral correlation. In this framework, a chroma component C , which can be either U or V , is treated as the dependent variable and is linearly predicted from the luma component Y , which serves as the independent variable. The linear prediction model is given by

$$\tilde{C} = \alpha_C Y + \beta_C, \quad (7)$$

where $\tilde{C}$ denotes the predicted chroma component and α_C and β_C are the regression parameters. The parameters are estimated by minimizing the chroma prediction error in the least-squares sense. The corresponding cost function is defined as

$$f(\alpha_C, \beta_C) = \sum_n |C(n) - \tilde{C}(n)|^2, \quad (8)$$

where n is the sample index. The optimal parameters are uniquely obtained as

$$\alpha_C = \frac{N \sum_n (Y(n)C(n)) - \sum_n Y(n) \sum_n C(n)}{N \sum_n Y(n)^2 - (\sum_n Y(n))^2}, \quad (9)$$

$$\beta_C = \frac{\sum_n C(n) - \alpha_C \sum_n Y(n)}{N}, \quad (10)$$

where N is the total number of samples. After parameter estimation, the chroma residuals

$$D_C = C - \lfloor \tilde{C} \rfloor \quad (11)$$

are encoded instead of the original chroma component C . Note that to ensure integer reversibility, the predicted chroma component $\tilde{C}$ is floored prior to subtraction. The regression parameters α_C and β_C are transmitted as side information and allow exact reconstruction of the chroma signal from the luma and the decoded residual.

This formulation assumes co-located luma and chroma samples and spatially uniform noise characteristics, assumptions that do not hold in CFA-sampled raw images. In particular, the spatial misalignment between luma and chroma samples and the presence of signal-dependent sensor noise can significantly degrade the accuracy and robustness of the regression model. Therefore, directly applying CfL prediction to CFA-domain signals may lead to suboptimal decorrelation performance.

¹In [10], the two chroma components are denoted by C_b and C_r , which correspond to U and V in this study.

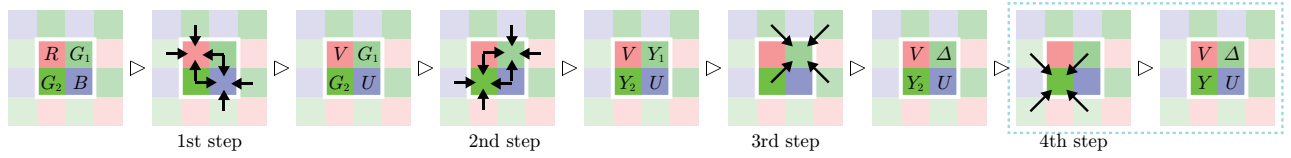


Fig. 2. Implementation of STT (arrows indicate lifting steps in the STT process).

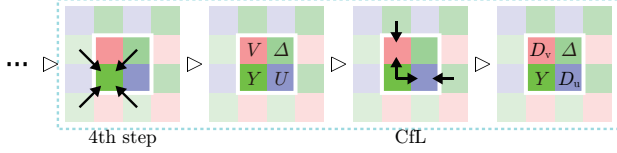


Fig. 3. Implementation of CfL-STT (arrows indicate the spatially-aligned CfL operation).

III. CfL-ENHANCED STT (CfL-STT)

To improve raw image compression performance, CfL prediction is integrated into the STT to predict chroma components U and V from the luma component Y , resulting in a $Y\Delta DuDv$ color space representation (see Fig. 3). However, as shown in Section IV, a direct application of conventional CfL prediction yields image-dependent coding behavior for raw images. This behavior arises from the lack of spatial alignment between luma and chroma samples and from the sensitivity of least-squares regression to sensor noise, as discussed in Section II-B.

We introduce a spatially-aligned CfL prediction model defined as

$$\tilde{C}^* = \alpha_C^* Y_C + \beta_C^*, \quad (12)$$

where Y_C denotes the luma signal spatially aligned with the chroma sampling grid corresponding to component C . The standard CfL prediction formulas in (9) and (10) are modified by replacing Y with Y_C , yielding

$$\alpha_C^* = \frac{N \sum_n Y_C(n) C(n) - \sum_n Y_C(n) \sum_n C(n)}{N \sum_n Y_C(n)^2 - (\sum_n Y_C(n))^2}, \quad (13)$$

$$\beta_C^* = \frac{\sum_n C(n) - \alpha_C^* \sum_n Y_C(n)}{N}. \quad (14)$$

When linear interpolation is adopted, the spatially-aligned luma signals are defined as

$$Y_U = \frac{Y^l + Y^r}{2} \quad \text{and} \quad Y_V = \frac{Y^t + Y^b}{2}. \quad (15)$$

These definitions align the luma samples with the chroma sampling grid along the horizontal and vertical directions, respectively. The interpolation in (15) is not intended to be optimal, but provides a low-complexity and robust spatially aligned luma predictor because the chroma components after STT are located between neighboring luma samples. It requires only one addition and one division-by-two operation per sample and introduces mild noise suppression without increasing the spatial support. More elaborate alignment strategies may further improve prediction accuracy, but they would increase computational cost, memory access, and line-buffer requirements.

After parameter estimation, the chroma residuals are defined by

$$D_C^* = C - \lfloor \tilde{C}^* \rfloor. \quad (16)$$

The regression parameters α_C^* and β_C^* are treated as side information for reconstruction at the decoder. This formulation preserves integer reversibility for lossless raw image compression.

In the main experiments, N in (13) and (14) denotes all samples in a frame, and the four regression parameters are included as frame-level side information. This setting evaluates the decorrelation capability of the proposed prediction, but it is not a strict line-based low-latency operation. Since (13) and (14) can be evaluated over an arbitrary set of samples, the summations can be restricted to samples within a precinct or line group, while the sample prediction in (15) only requires adjacent luma samples. For clarity, in the experimental evaluation, we refer to the proposed spatially aligned prediction as CfL-STT (Aligned) to distinguish it from a direct application of conventional CfL prediction without spatial alignment.

IV. EXPERIMENTAL RESULTS

To evaluate the proposed CfL-STT (Aligned) in lossless raw image compression, we integrated the CfL-STT module into the JPEG XS reference software [20] and compared its performance with that of the baseline STT defined in JPEG XS 2nd Edition [4]. Since the original JPEG XS reference software does not support STT-based lossless compression, we extended it to enable a fully reversible STT-based coding pipeline. In this pipeline, the non-linear transform (NLT) was disabled to ensure reversibility. The proposed CfL-STT (Aligned) replaces the standard STT by integrating CfL-based chroma prediction into the transform, resulting in a $Y\Delta DuDv$ representation instead of the conventional $Y\Delta UV$. In addition, a reference method, denoted as CfL-STT (Direct), was evaluated, corresponding to a direct application of conventional CfL prediction without spatial alignment. In this configuration, CfL prediction is applied to the STT output using luma samples available within each macropixel, which is equivalent to nearest-neighbor interpolation, i.e., $Y_U = Y^l$ and $Y_V = Y^b$. The test set consisted of ten DNG images from the MIT-Adobe FiveK dataset [21] and ten DNG images captured using an iPhone SE3 [12], with bit depths of 12 or 14 bits per sample. Coding efficiency was evaluated in terms of bitrate [bits per pixel (bpp)] relative to the original CFA-sampled raw image.

A. Decorrelation of Chroma Components

Figure 4 visualizes the decorrelated components produced by STT and the two CfL-STT variants for image *m0005*.

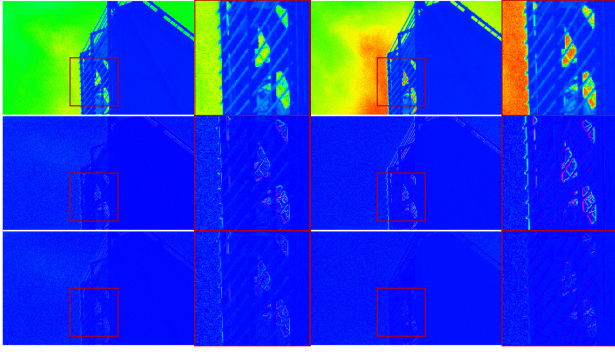


Fig. 4. Pseudo-color visualization of a particular area of the decorrelated components for image *m0005* (the same pseudo-color scale and display range are used for all panels, and colors closer to blue indicate values closer to zero): (top to bottom) (U, V), (D_U, D_V), and (D_U^*, D_V^*).

While STT produces chroma components U and V , CfL-STT (Direct) yields chroma residuals D_U and D_V , and CfL-STT (Aligned) yields chroma residuals D_U^* and D_V^* .

CfL-STT (Direct) exhibited image-dependent coding behavior. This means that the direct method reduces the bitrate for some images but increases it for others compared with the baseline STT, as quantitatively shown in Section IV-B. This image-dependent behavior is caused by the spatial mismatch between luma and chroma samples. When the nearest luma sample within a macropixel is directly used for CfL prediction, local edges and high-frequency sensor noise are not properly aligned with the chroma sampling positions. Consequently, the regression model may produce less accurate chroma predictions, broaden the residual distributions, and increase the residual entropy in some cases.

In contrast, CfL-STT (Aligned) uses interpolated luma samples that are spatially aligned with the chroma sampling grids. This alignment reduces the mismatch between the predictor and the target chroma component while mildly suppressing high-frequency sensor noise. As a result, CfL-STT (Aligned) produced more concentrated residual distributions, indicating improved decorrelation between luma and chroma components.

B. Lossless Compression

Table I reports the lossless compression results. The overhead of the four frame-level regression parameters is included in the reported bitrates, but it is not shown separately because its effect is negligible at the reported precision. CfL-STT (Direct) exhibited image-dependent coding behavior and degraded compression efficiency for some images. In contrast, CfL-STT (Aligned) consistently outperformed the baseline STT for all test images, reducing the average bitrate from 7.00 to 6.94 bpp and from 7.29 to 7.16 bpp for the two datasets, respectively. These results indicate that spatial alignment is essential for effectively applying CfL prediction in CFA-domain spectral-spatial transforms. Although the average bitrate reductions of 0.06 and 0.13 bpp are modest, the improvements were observed for all tested images.

We also conducted an encoder-side line-group simulation in which the regression parameters were estimated every

TABLE I
BITRATES [bpp] FOR LOSSLESS COMPRESSION (BOLD FONTS INDICATE THE BEST RESULTS AND BLUE/RED ONES INDICATE IMPROVED/DEGRADED RESULTS COMPARED TO STT).

Test Images	STT [10]	CfL-STT		Test Images	STT [10]	CfL-STT	
		Direct	Aligned			Direct	Aligned
#0500	6.70	6.73	6.67	<i>m0001</i>	7.67	7.75	7.52
#1000	6.70	6.72	6.67	<i>m0002</i>	7.57	7.68	7.43
#1500	8.49	8.46	8.38	<i>m0003</i>	6.89	6.88	6.88
#2000	5.51	5.56	5.45	<i>m0004</i>	7.21	7.20	7.08
#2500	6.49	6.50	6.45	<i>m0005</i>	6.55	6.56	6.42
#3000	6.11	6.13	6.04	<i>m0006</i>	7.32	7.30	7.29
#3500	7.68	7.65	7.57	<i>m0007</i>	6.62	6.65	6.44
#4000	8.88	8.88	8.87	<i>m0008</i>	8.31	8.39	8.08
#4500	7.03	7.09	6.98	<i>m0009</i>	6.92	6.88	6.81
#5000	6.36	6.40	6.27	<i>m0010</i>	7.80	7.88	7.61
Avg.	7.00	7.01	6.94	Avg.	7.29	7.32	7.16

16 component lines, corresponding to 32 input raw lines. Assuming four 32-bit floating-point coefficients per line group, the side-information overhead was about 10^{-3} bpp. Including this overhead, the average bitrates were 6.93 and 7.13 bpp for the two datasets, respectively. This suggests that the proposed formulation can retain its coding gain under local coefficient estimation without requiring full-frame statistics.

Regarding computational complexity, the additional processing mainly consists of two-tap interpolation, scalar linear prediction, and encoder-side accumulation of simple sums for coefficient estimation. Since the regression parameters are transmitted as side information, no parameter estimation is required at the decoder. The additional local memory access is limited to adjacent luma samples, although its interaction with line buffering and the JPEG XS rate allocator is left for future work. The proposed method therefore provides a practical trade-off between limited additional complexity and bitrate reduction.

Finally, a detailed analysis of backward-adaptive coefficient estimation, lossy coding including quantization effects and bitrate-dependent behavior, and their interaction with the JPEG XS rate allocator is left for future work.

V. CONCLUSION

This study proposed CfL-STT for lossless JPEG XS raw image compression. The proposed method predicts chroma components from spatially aligned luma samples, producing a more decorrelated $Y\Delta DuDv$ representation while preserving exact reversibility. Experimental results showed that direct CfL prediction can yield image-dependent performance and may degrade coding efficiency, whereas the proposed aligned prediction consistently improves lossless coding efficiency. The encoder-side line-group simulation further suggests that the coding gain can be retained under local coefficient estimation. Future work includes complete bitstream-level signaling, interaction with the JPEG XS rate allocator, and more adaptive spatial alignment strategies.

ACKNOWLEDGMENT

The authors would like to express their gratitude to C. Huang and K. Oka of our laboratory for their invaluable assistance in the initial implementation of the coding framework used in this study.

REFERENCES

- [1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electr.*, vol. 38, no. 1, pp. xviii–xxxiv, Feb. 1992.
- [2] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG2000 still image compression standard," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 36–58, Sep. 2001.
- [3] J. Ascenso, E. Alshina, and T. Ebrahimi, "The JPEG AI standard: Providing efficient human and machine visual data consumption," *IEEE MultiMedia*, vol. 30, no. 1, pp. 100–111, May 2023.
- [4] A. Descampe, T. Richter, T. Ebrahimi, S. Föbel, J. Keinert, T. Bruylants, P. Pellegrin, C. Buyschaert, and G. Rouvroy, "JPEG XS—A new standard for visually lossless low-latency lightweight image coding," *Proc. IEEE*, vol. 109, no. 9, pp. 1559–1577, May 2021.
- [5] N. Zhang and X. Wu, "Lossless compression of color mosaic images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1379–1388, Jun. 2006.
- [6] H. S. Malvar and G. J. Sullivan, "Progressive-to-lossless compression of color-filter-array images using macropixel spectral-spatial transformation," in *Proc. DCC*, Snowbird, UT, USA, Apr. 2012, pp. 3–12.
- [7] M. Hernández-Cabrero, M. W. Marcellin, I. Blanes, and J. Serra-Sagristà, "Lossless compression of color filter array mosaic images with visualization via JPEG 2000," *IEEE Trans. Multimedia*, vol. 20, no. 2, pp. 257–270, Feb. 2018.
- [8] Y. Lee, K. Hirakawa, and T. Q. Nguyen, "Camera-aware multi-resolution analysis for raw image sensor data compression," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2806–2817, Jun. 2018.
- [9] Y. Lee and K. Hirakawa, "Shift-and-decorrelate lifting: CAMRA for lossless intra frame CFA video compression," *IEEE Signal Process. Lett.*, vol. 27, pp. 461–465, Jan. 2020.
- [10] T. Richter, S. Föbel, A. Descampe, and G. Rouvroy, "Bayer pattern compression with JPEG XS," *IEEE Trans. Image Process.*, vol. 30, pp. 6557–6569, Jul. 2021.
- [11] T. Suzuki, "Wavelet-based spectral-spatial transforms for CFA-sampled raw camera image compression," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 433–444, Dec. 2020.
- [12] T. Suzuki and L. Huang, "Edge-aware extended Star-Tetrix transforms for CFA-sampled raw image compression," *IEEE Trans. Image Process.*, vol. 31, pp. 6072–6082, Sep. 2022.
- [13] J. Han, B. Li, D. Mukherjee, C.-H. Chiang, A. Grange, C. Chen, H. Su, S. Parker, S. Deng, U. Joshi, Y. Chen, Y. Wang, P. Wilkins, Y. Xu, and J. Bankoski, "A technical overview of AV1," *Proc. IEEE*, vol. 109, no. 9, pp. 1435–1462, Feb. 2021.
- [14] J. Sneyers, J. Alakuijala, L. Versari, Z. Szabadka, S. Boukourt, A. Cohen-Tidhar, M. Firsching, E. Kliuchnikov, T. Lev-Ami, E. Portis, T. Richter, and O. Watanabe, "The JPEG XL image coding system: History, features, coding tools, design rationale, and future," in *arXiv*, Jun. 2025, pp. 1–73.
- [15] L. Trudeau, N. Egge, and D. Barr, "Predicting chroma from luma in AV1," in *Proc. DCC*, Snowbird, UT, USA, Mar. 2018, pp. 374–382.
- [16] J. Huo, H. Du, X. Li, S. Wan, H. Yuan, Y. Ma, and F. Yang, "Unified cross-component linear model in VVC based on a subset of neighboring samples," *IEEE Trans. Industr. Inform.*, vol. 18, no. 12, pp. 8654–8663, Dec. 2022.
- [17] J. Huo, D. Wang, H. Yuan, S. Wan, and F. Yang, "Adaptive chroma prediction based on luma difference for H.266/VVC," *IEEE Trans. Image Process.*, vol. 32, pp. 6318–6331, Nov. 2023.
- [18] G. Delétang, A. Defossez, M. Zoran, J. Menick, N. Kalchbrenner, A. Kolesnikov, A. Arnab, Y. Aytar, J. Carreira, and A. Zisserman, "Language modeling is compression," in *Proc. ICLR*, Vienna, Austria, May 2024, pp. 1–17.
- [19] "Kodak Lossless True Color Image Suite," *Kodak [Online]*, Available: <http://r0k.us/graphics/kodak/>.
- [20] Joint Photographic Experts Group (JPEG), "JPEG XS Reference Software," <https://jpeg.org/jpegxs/software.html>, Accessed: Apr. 1, 2026.
- [21] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image pairs," in *Proc. CVPR*, Colorado Springs, CO, USA, Jun. 2011, pp. 97–104.