Inter-Component Transform for Color Video Coding

Christian Rudat[1], Christian R. Helmrich[1], Jani Lainema[2], Tung Nguyen[1], Heiko Schwarz[1,3], Detlev Marpe[1], and Thomas Wiegand[1,4]

[1]Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut (HHI), 10587 Berlin, Germany,
[2]Nokia Corporation, Tampere 33100, Finland, [3]Institute for Computer Science, Free University of Berlin, Germany,
[4]Image Communication Chair, Technical University of Berlin, Germany

***Abstract*—In natural digital images and videos, correlations between color components can be observed. These correlations can be exploited to achieve additional coding gain in modern block-based hybrid video coding. To this end, we propose the use of a block-wise, rotational inter-component transform (ICT) applied to the two residual chroma signals that result from conventional intra or inter-picture prediction. Different ICT parameterizations in terms of number and quantization of the rotational angles as well as resulting components signaled in the coded bitstream are investigated. An implementation into the currently developed Versatile Video Coding (VVC) reference software provides average bitrate savings of up to 0.7% (All Intra configuration) with negligible increases in implementation complexity and runtime. Our proposal has been adopted into the VVC draft specification text.**

***Index Terms*—color images, KLT, rotation, video coding.**

I. Introduction

Digital images and image sequences are usually composed of multiple planes, each plane representing a color component (e. g., RGB, YC$_\mathrm{b}$C$_\mathrm{r}$). In natural image content acquired via digital camera sensors, a signal correlation between these planes can be observed. To increase the coding efficiency of a picture/video compression solution such as High Efficiency Video Coding (HEVC) [1] or the currently developed Versatile Video Coding (VVC) [2], those inter-component redundancies may be exploited by reusing information from already coded components to compress another component. For example, in case of content featuring a YC$_\mathrm{b}$C$_\mathrm{r}$ color scheme (one luma plane and two chroma planes), a video coding tool named “Cross-Component Linear Model” (CCLM) [4] increases coding efficiency by allowing for intra-frame chroma prediction to be derived from the already reconstructed luma signal using a linear model.

Apart from reusing one color plane to code another, we explored techniques to further increase the coding efficiency by jointly processing signals from two or more color components in a manner that is separated from the quantization stage present in codecs like HEVC or VVC. Specifically, a switchable inter-component transform (ICT) method was developed which can be applied to multi-component residual signals in addition to a conventional (intra-component) spatial transform. In other words, separable 3D transformation is applied to multi-component residual block signals in order to achieve not only spatial but also inter-component energy compaction.

Apart from presenting the basic concept of ICT along with its underlying principles, we present three specific implementation variants and illustrate how these affect the coding efficiency in state-of-the-art video coding.

This paper is structured as follows. Section II presents popular conventional methods to exploit inter-component redundancies. After that, our ICT approach is motivated in Section III and introduced in Section IV. Furthermore, Section V gives an overview of the implementation of ICT into a state-of-the-art video codec (VVC draft 4 [2]) and of experimental results regarding the impact on coding efficiency. The paper is concluded in Section VI.

II. Related Work

A. Residual Color Transform (RCT)

In order to benefit from color space conversions (to YC$_\mathrm{b}$C$_\mathrm{r}$, for example) to improve coding efficiency while coding RGB material without color-converting the whole video sequence, block based “Residual Color Transform” (RCT) was proposed [3]. The residual signals r_R, r_G and r_B that result from intra or inter-frame predictions in RGB color space are transformed into signals r_{C1}, r_{C2} and r_{C3} with

$$\begin{pmatrix} r_{C1} \\ r_{C2} \\ r_{C3} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 1 & 0 & -1 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} r_R \\ r_G \\ r_B \end{pmatrix}.$$

At the encoder, the transform is applied before quantization while at the decoder, the inverse conversion is applied after dequantization. This process improves coding efficiency significantly due to signal energy compaction across the color components.

B. Cross-Component Linear Model (CCLM)

A common approach for exploiting inter-component correlations in $\mathrm{YC_bC_r}$ images or videos is the chroma prediction technique “Cross-Component Linear Model” (CCLM) [4]. Here, a (subsampled) version of the already reconstructed luma block signal o'_Y is adapted with use of a linear model and used to predict a chroma signal. Specifically, the prediction signal p_{Ci} for a chroma component $\mathrm{C_i}$ is generated with $p_{Ci} = \alpha \cdot o'_Y + \beta$. The model parameters α and β are derived, at both the encoder and decoder side, using linear regression applied to the already reconstructed, neighboring samples of the color components involved. Hence, the model parameters do not need to be signaled explicitly in the bitstream.

C. Cross-Component Prediction (CCP)

Another example for inter-component residual coding is so-called “ Cross-Component Prediction” (CCP) [5]. In case of $\mathrm{YC_bC_r}$, the (subsampled) dequantized luma residual signal r'_L is used to modify a dequantized chroma residual r'_C to form the final chroma residual signal r_C with $r_C = r'_C + 2^{-(p-1)} \cdot r'_L$ while the parameter p is signaled in the bitstream.

As an extension to CCP [6] the residual of the first (main) chroma component r_{CM} is used to modify the residual signal of the second (remaining) chroma component r_{CR} by means of weighted subtraction $r'_{CR} = r_{CR} - \alpha \cdot r_{CM}$ with the weighting parameter α, which is chosen by the encoder and signaled in the data bitstream. This way the modified residual signal can be coded more efficiently.

III. Motivation

The basic concept of modern, lossy video coding is to predict an image block and represent the prediction residual signal in the most efficient way. Specifically, the goal is to generate residual signals with the best possible tradeoff between bitrate and distortion. To this end, different measures to reduce signal correlation which allow for more efficient residual coding may be taken.

Digital, natural image signals typically contain a significant amount of correlation, including inter-component correlation. For example, strong block based correlations between the unquantized residual signals r_{Cb} and r_{Cr} can be observed in Fig. 1 which illustrates the distribution of the transform unit block (TU) correlation coefficient c_{CbCr} across TUs given with

$$c_{CbCr} := \frac{\sum_n (r_{Cb,n} - \overline{r}_{Cb})(r_{Cr,n} - \overline{r}_{Cr})}{\sqrt{\sum_n (r_{Cb,n} - \overline{r}_{Cb})^2 \cdot \sum_n (r_{Cr,n} - \overline{r}_{Cr})^2}}.$$

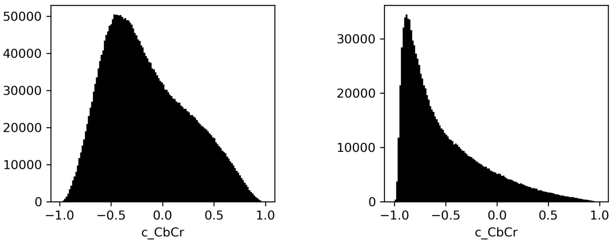

Fig. 1. Histograms of correlation coefficient c_{CbCr} for unquantized residual signals. The coefficient is measured for every nonzero C_b/C_r TU the VVC encoder (VTM 4.0.1, quantization parameter 32) chose for the first frame. Video sequence *ParkRunning3* (left) and *Campfire*.

Where $r_{Cb,n}$ and $r_{Cr,n}$ denote the unquantized residual samples within one TU at sample position n while their corresponding arithmetic mean values are given with $\overline{r}_{Cb}$ and $\overline{r}_{Cr}$.

Patterns of this kind may be exploited to increase coding efficiency by applying a rotational transform (illustrated in Fig. 2) with an angle α so that a compaction of signal energy between both components is achieved.

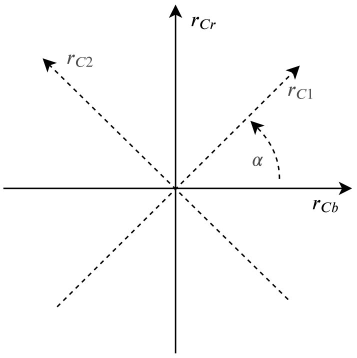

Fig. 2. Illustration of rotational transform with angle α.

IV. Proposed Approach

For the purpose of this paper, the proposed techniques are applied to the two chroma components of images and image sequences in $\mathrm{YC_bC_r}$ color space. However, the concept of ICT can easily be applied to any number of components, for example to all three components contained in RGB.

A. ICT in the coding process

The process of block based hybrid video coding (e. g., HEVC) includes several processing stages to code a block. Typically, picture blocks are reconstructed by combining quantized residual signals with prediction signals. Often it is favorable for the coding efficiency

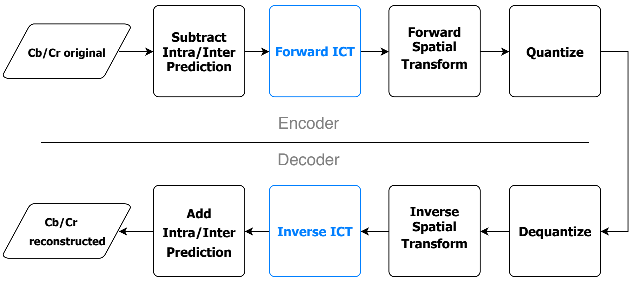
Cb/Cr original
Subtract Intra/Inter Prediction
Forward ICT
Forward Spatial Transform
Quantize
Encoder
Decoder
Cb/Cr reconstructed
Add Intra/Inter Prediction
Inverse ICT
Inverse Spatial Transform
Dequantize

Fig. 3. Location of ICT in predictive image or video coding systems.

that a spatial transform (e. g., 2D-DCT) is applied to a residual block as this process may result in signal energy compaction. This holds advantages with regards to quantization and entropy-coding of the residual signal, typically resulting in further increase of coding efficiency. However, spatial transforms can redistribute signal energy only within a single color component.

In comparison, with the introduction of ICT, correlations can be exploited further by adding another one-dimensional transformation stage applied along a third axis, the color components. In essence, multi-component spatial transform stages followed (or preceded) by an inter-component transform (ICT) stage can be understood as conjunctively forming a separable 3D-transformation, as illustrated in Figs. 3 and 4.

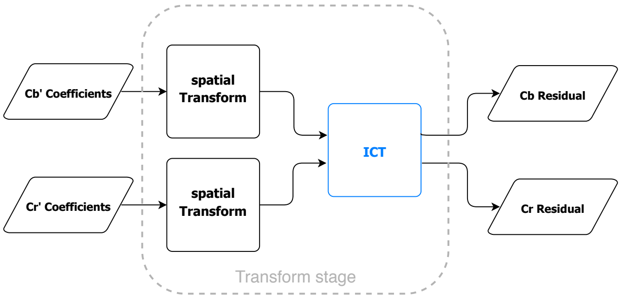

Fig. 4. Illustration of proposed complete multi-dimensional inverse transformation process along spatial and inter-component axis.

It is to be noted, that ICT is an additional and optional processing step during coding of a TU which is signaled in the bitstream. ICT can be inserted into an existing otherwise unmodified coding scheme, as long as it allows for joint processing of the corresponding residual signals.

B. Forward and inverse rotational transform

In order to exploit correlations between a pair of 2D signals $\vec{r}(x,y) = (r_{Cb}(x,y), r_{Cr}(x,y))$ at sample position (x,y) (i. e., the $\mathrm{C_b}$ and $\mathrm{C_r}$ components of a picture block) a rotational transformation yielding $\vec{r}_\alpha(x,y) = T_\alpha \cdot \vec{r}(x,y)$ with the rotation angle α may be applied in order to rotate more signal energy towards a single axis. This approach is well known from principal component analysis (PCA) [7] and discrete Karhunen-Loève transformation (KLT) used, e. g., in two-channel audio coding [8]. A possible energy-invariant transformation T_α could be given by the rotation

$$T_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Hence, applying ICT to two chroma residual signals r_{Cb} and r_{Cr} results in the transformed signals r_{C1}, r_{C2} which are then processed further (e. g., by conventional single-component spatial transformation) and transmitted

$$\begin{pmatrix} r_{C1} \\ r_{C2} \end{pmatrix} = T_\alpha \cdot \begin{pmatrix} r_{Cb} \\ r_{Cr} \end{pmatrix}.$$

During reconstruction, the (likely quantized) residual signals r'_{Cb} and r'_{Cr} can be recovered by applying an inverse inter-component transform T_α^{-1} (in case of energy-invariant transform: $T_\alpha^{-1} = T_{-\alpha}$, rotation with inverse angle) with

$$\begin{pmatrix} r'_{Cb} \\ r'_{Cr} \end{pmatrix} = T_\alpha^{-1} \cdot \begin{pmatrix} r_{C1} \\ r_{C2} \end{pmatrix}.$$

It is to be noted that in case of an angle $\alpha = \frac{\pi}{4}$, the T_α matrix corresponds to a basic Hadamard transform

$$T_{\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

C. Coding of residuals and ICT parameters

State-of-the-art video codecs like HEVC or VVC feature a residual coding stage where (transformed and quantized) residual signals are efficiently coded by using Context-Based Adaptive Binary Arithmetic Coding [10]. As the output of the ICT is of the same shape as its input (two 2D coefficient blocks of the same size), the residual coding engine is per se invariant to the application of ICT. In fact, in our approach neither the binarization nor the selection of context models is affected by whether the residual signals are processed by the ICT. Therefore, apart from signaling the ICT configuration, as described below, the bitstream syntax is not altered.

The encoder determines, on a coding-block basis, the “best” ICT parameter (angle α for least-squares optimal rotation minimizing r_{C2}) to apply for joint-chroma coding from a set of available rotation angles. In order to (implicitly) signal these selections of α to the decoder, we employ the component-wise *coded block flag* (CBF) available in HEVC and VVC which, for the given color component, indicates whether a non-zero residual block signal is coded in the bitstream (CBF $=1$). If the CBFs

TABLE I
ALLOWED VALUES OF α AND ASSOCIATED PCA ROTATION MATRIX WEIGHTS FOR INVERSE ICT IN FIXED-POINT IMPLEMENTATIONS.

Angular mode	–3	–2	–1	1	2	3
Value of α	$-\frac{\pi}{2.838}$	$-\frac{\pi}{4}$	$-\frac{\pi}{6.776}$	$\frac{\pi}{6.776}$	$\frac{\pi}{4}$	$\frac{\pi}{2.838}$
Matrix weights	$\frac{1}{2}$ 1	1 1	1 $\frac{1}{2}$	1 $-\frac{1}{2}$	1 –1	$\frac{1}{2}$ –1
(scaled T_α^{-1})	–1 $\frac{1}{2}$	–1 1	$-\frac{1}{2}$ 1	$\frac{1}{2}$ 1	1 1	1 $\frac{1}{2}$

for both components (here, C$_\text{b}$ and C$_\text{r}$) are equal to zero, the inverse ICT is disabled since it is not required at the decoder side. The remaining three possible combinations of CBF$_\text{Cb}$ and CBF$_\text{Cr}$ each indicate one particular value of α for the given block and trigger the signaling of an additional one-bit, arithmetically coded flag f per block specifying whether the ICT is actually enabled ($f = 1$). Furthermore, we investigated implementations of the ICT that omit the second resulting component r_{C2} of a given block (which has been decreased in variance by the ICT) by setting it to zero (CBF = 0).

D. ICT variants with different angle and channel count

With the *explicit* signaling of the ICT activation via f and the *implicit* signaling of the rotation angles α using the two chroma CBFs, we constructed three specific ICT variants for implementation and evaluation.

1) ICT method 1: represents, in terms of parametrization, the most constrained (but also algorithmically least complex) ICT variant investigated. It allows the selection of only a single, fixed rotation angle $\alpha = -\frac{\pi}{4}$ along with signaling of flag f when the CBFs for both chroma components C$_\text{b}$ and C$_\text{r}$ equal 1 in the given block. Also, $r_{C2} = 0$ is enforced to realize an *intensity* coding mode as in [8], which results in only one of two possible chroma channels being coded when ICT is enabled.

2) ICT method 2: extends method 1 by allowing the ICT encoder to choose, for each block, from two further angular magnitudes as well as an arbitrary angular sign, for a total of 6 angles as tabulated in Table I. For each block, the encoder selects the optimal (in terms of signal decorrelation) α and, if ICT coding with this α provides a lower cost than coding without ICT, sets $f = 1$.

3) ICT method 3: extends method 2 by removing the $r_{C2} = 0$ restriction when the CBFs for both C$_\text{b}$ and C$_\text{r}$ are equal to 1. This results in a maximally unconstrained (but also somewhat more complex) ICT variant featuring 6 angles and up to 2 signals r_{C1} and r_{C2}.

V. Experimental Results

Fixed-point realizations of the ICT methods described in Section IV were implemented into VTM 4.0.1, the (as of this writing) latest version of the reference encoding and decoding software developed as part of the currently standardized Versatile Video Coding (VVC) specification [2], [9]. To limit the implementation complexity particularly in hardware, only integer additions, subtractions, and multiplications were used, and divisions were replaced by right-shifts (equaling power-of-two divisions as suggested in Table I). To this end, T_α was multiplied by, and T_α^{-1} divided by, $\max(|\cos(\alpha)|, |\sin(\alpha)|)$, and the chroma quantization step-size and Lagrange multiplier were adjusted accordingly. For ICT methods 2 and 3, the magnitude of each block rotation angle was transmitted on a transform unit (TU) basis, while the overall sign of the rotation angles was conveyed only once per picture.

The effect of the ICT design on the coding efficiency of the VVC software was evaluated on medium and high-resolution 4:2:0 coded natural videos using Bjøntegaard delta-rate (BD-rate) calculations [11] relative to the VTM anchor configuration without activated ICT. The video selection, frame count, and basic encoder setup followed the JVET common test conditions for standard dynamic range (SDR) input [12]. To ease comparisons, the relative luma-chroma bit-allocation for each ICT extended VTM version under test was adjusted, by varying the encoder's respective Lagrange parameter, to produce similar mean chroma BD-rate results. This allows for any difference in coding efficiency between the three methods to be easily observable in the luma BD-rate results.

Tables II, III, and IV show the video-class-wise BD-rate values for the respective ICT method when activated in the context of VTM 4.0.1. It can be seen that, overall, a coding gain is achieved by all ICT variants for both All-Intra (AI, no inter-picture prediction, GOP size 1) and Random-Access (RA, intra and inter-picture prediction, GOP size 16) configurations [12]. The coding gains for AI are roughly 50% higher than those for RA since, due to the lack of efficient inter-picture prediction, fewer residual block signals are quantized to zero in the former. Using only one fixed rotation angle $\alpha = -\frac{\pi}{4}$, no overall sign transmission, and downmixing into only one channel, ICT method 1 already achieves about 0.4% BD-rate reduction with only 1–3% encoder runtime increase. ICT method 2, with its 6 angles $-\frac{\pi}{2} < |\alpha| < \frac{\pi}{2}$ to choose from and picture-wise sign transmission, yields another 0.2–0.25% in coding efficiency at the same runtime and decoder complexity. ICT method 3, allowing up to two coded chroma channels, lowers the BD-rate further by roughly 0.05% at the cost of slightly higher encoding runtimes. Overall, coding gains of about 0.7% in AI and 0.5% in RA are possible in VTM4. More detailed results, including data for low-delay configurations, are provided in [13] (method 1) and [14] (methods 2 and 3).

TABLE II
BD-RATE RESULTS FOR ICT METHOD 1 (ONE ANGLE, ONE CHANNEL). TOP: ALL INTRA (AI), BOTTOM: RANDOM ACCESS (RA).

SDR, AI	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.38%	–0.04%	0.98%	103%	101%
Class A2	–0.79%	–1.17%	0.77%	105%	101%
Class B	–0.25%	–0.21%	–2.17%	102%	101%
Class C	–0.28%	–1.26%	–3.15%	103%	99%
Class E	–0.28%	0.26%	–3.08%	102%	99%
Overall	**–0.38%**	**–0.49%**	**–1.52%**	**103%**	**100%**

SDR, RA	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.44%	–0.85%	1.55%	101%	99%
Class A2	–0.38%	–2.51%	0.31%	101%	101%
Class B	–0.16%	–0.30%	–2.03%	101%	101%
Class C	–0.17%	–0.13%	–2.31%	101%	99%
Overall	**–0.26%**	**–0.81%**	**–0.92%**	**101%**	**100%**

Further studies revealed that ICT rotations are enabled in roughly 50% of all coded chroma blocks and that BD-rate gains of up to 3.7% are achieved (AI, *ParkRunning3* sequence, class A2). These observations confirm the benefit of jointly coding image and video color components. Method 2 has been adopted into the VVC draft standard.

VI. Conclusion

We proposed the joint coding of residual chroma signals in modern hybrid image and video codecs, using rotational inter-component transformation, to achieve further signal compaction and, thereby, address remaining color-channel correlation often found in natural picture content. The approach was shown to provide up to 0.7% BD-rate reduction on average in the context of the new VVC standard [2]. Extensions to 4:4:4 color sampling and RGB coding are topics considered for future study.

TABLE III
BD-RATE RESULTS FOR ICT METHOD 2 (6 ANGLES, SINGLE CHANNEL). TOP: ALL INTRA (AI), BOTTOM: RANDOM ACCESS (RA).

SDR, AI	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.73%	0.14%	2.34%	103%	100%
Class A2	–1.37%	–1.79%	1.48%	106%	100%
Class B	–0.40%	0.21%	–1.44%	102%	100%
Class C	–0.41%	–1.73%	–2.63%	104%	100%
Class E	–0.37%	–1.60%	–3.83%	101%	101%
Overall	**–0.62%**	**–0.87%**	**–0.98%**	**103%**	**100%**

SDR, RA	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.62%	–2.73%	1.58%	101%	101%
Class A2	–0.85%	–4.01%	1.17%	101%	100%
Class B	–0.21%	–0.46%	–2.16%	100%	99%
Class C	–0.27%	–1.12%	–2.64%	100%	100%
Overall	**–0.43%**	**–1.80%**	**–0.87%**	**100%**	**100%**

TABLE IV
BD-RATE RESULTS FOR ICT METHOD 3 (6 ANGLES, 1 OR 2 CHANNELS). TOP: ALL INTRA (AI), BOTTOM: RANDOM ACCESS (RA).

SDR, AI	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.71%	–0.40%	2.21%	106%	101%
Class A2	–1.71%	–1.88%	1.22%	110%	98%
Class B	–0.41%	0.43%	–1.39%	105%	98%
Class C	–0.44%	–1.22%	–2.55%	107%	99%
Class E	–0.39%	–1.12%	–3.66%	103%	97%
Overall	**–0.68%**	**–0.72%**	**–0.99%**	**106%**	**98%**

SDR, RA	Y (Luma)	U (Cb)	V (Cr)	Enc.T.	Dec.T.
Class A1	–0.61%	–3.17%	1.84%	103%	99%
Class A2	–0.99%	–4.10%	0.84%	103%	100%
Class B	–0.22%	–0.12%	–2.18%	101%	99%
Class C	–0.24%	–0.26%	–2.41%	102%	102%
Overall	**–0.46%**	**–1.56%**	**–0.83%**	**102%**	**100%**

References

[1] ITU-T, Recommendation H.265 and ISO/IEC, Int. Standard 23008-2, “High Efficiency Video Coding,” Geneva, Feb. 2018, online: http://www.itu.int/rec/T-REC-H.265

[2] B. Bross, J. Chen, and S. Liu (editors), “Versatile Video Coding (Draft 4),” document *JVET-M1001*, Marrakech, Mar. 2019.

[3] W.-S. Kim, D.-S. Cho, D. Birinov, H. M. Kim, S.-H. Lee, and Y.-S. Seo, “RGB Video Coding Using Residual Color Transform,” *SAMSUNG Journal of Innovative Technology*, vol. 1, no. 1, pp. 21–31, Aug. 2005.

[4] K. Zhang, J. Chen, L. Zhang, X. Li, and M. Karczewicz, “Enhanced Cross-Component Linear Model for Chroma Intra-Prediction in Video Coding,” *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 3983–3997, Aug. 2018.

[5] T. Nguyen, A. Khairat, L. and D. Marpe “Adaptive Inter-Plane Prediction for RGB Content,” document *JCTVC-M0230*, Incheon, Apr. 2013.

[6] M. Siekmann, A. Khairat, T. Nguyen, D. Marpe, and T. Wiegand, “Extended Cross-Component Prediction in HEVC,” *APSIPA Trans. Signal Inf. Process.*, vol. 6, no. 3, pp 1–8, Apr. 2017.

[7] K. Pearson, “On Lines and Planes of Closest Fit to Systems of Points in Space,” *Philosophical Magazine*, vol. 2, no. 11, pp. 559–572, 1901.

[8] R. G. van der Waal and R. N. J. Veldhuis, “Subband Coding of Stereophonic Digital Audio Signals,” in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Process.*, Toronto, pp. 3601–3604, Apr. 1991.

[9] JVET, “VVCSoftware_VTM – Tags – VTM-4.0.1”, Feb. 2019, online: https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM

[10] D. Marpe, H. Schwarz, and T. Wiegand, “Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard,” *IEEE Trans. Circuits Systems Video Technol.*, vol. 13, no. 7, pp. 620–636, July 2003.

[11] G. Bjøntegaard, “Calculation of Average PSNR Differences Between RD-Curves,” document *VCEG-M33*, Austin, Apr. 2001.

[12] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Sühring, “JVET Common Test Conditions and Software Reference Configurations for SDR Video,” document *JVET-M1010*, Marrakech, Jan. 2019.

[13] J. Lainema, “CE7: Joint Coding of Chrominance Residuals (CE7-1),” document *JVET-N0054*, Geneva, Mar. 2019.

[14] C. Helmrich, C. Rudat, T. Nguyen, H. Schwarz, D. Marpe, and T. Wiegand, “CE7-Related: Joint Chroma Residual Coding with Multiple Modes,” document *JVET-N0282*, Geneva, Mar. 2019.