Cross-component prediction in HEVC

Article *in* IEEE Transactions on Circuits and Systems for Video Technology · January 2015

DOI: 10.1109/TCSVT.2015.2496821

9 authors, including:

Woo-Shik Kim
Qualcomm

Wei pu
Chinese Academy of Sciences

Jianle Chen
Qualcomm

Marta Karczewicz
Qualcomm

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology

Cross-Component Prediction in HEVC

Woo-Shik Kim, *Senior Member, IEEE,* Wei Pu, *Member, IEEE,* Ali Khairat, Mischa Siekmann, Joel Sole, *Senior Member, IEEE,* Jianle Chen, Marta Karczewicz, Tung Nguyen, Detlev Marpe, *Fellow, IEEE*

***Abstract*—Video coding in the YCbCr color space has been widely used, since it is efficient for compression, but it can result in color distortion due to conversion error. Meanwhile, coding in RGB color space maintains high color fidelity, having the drawback of a substantial bit-rate increase with respect to YCbCr coding. Cross-component prediction (CCP) efficiently compresses video content by decorrelating color components while keeping high color fidelity. In this scheme, the chroma residual signal is predicted from the luma residual signal inside the coding loop.**

This paper gives a description of the CCP scheme from several point-of-view, from theoretical background to practical implementation. The proposed CCP scheme has been evaluated in standardization communities and adopted into H.265/HEVC Range Extensions. Experimental results show significant coding performance improvements both for natural and screen content video, while the quality of all color components is maintained. The average coding gains for natural video are 17% and 5% bit-rate reduction in case of intra coding, and 11% and 4% in case of inter coding for RGB and YCbCr coding, respectively, while the average increment of encoding and decoding times in the HEVC reference software implementation are 10% and 4%, respectively.

***Index Terms*—Video compression, color decorrelation, HEVC, video coding standards.**

I. Introduction

There has been demand for high fidelity on moving picture representation, especially in professional video applications like broadcasting distribution, digital cinema, and post production [1]. A typical property in such applications is that the pictures are processed in 4:4:4 chroma format, i.e., without subsampling chroma components. Nowadays, the high fidelity demand has been extended to consumer applications thanks to the advances in multimedia related technologies including display, computing system, and content generation. In addition, the consumer application space has been expanded to screen contents coding applications such as remote desktop, screen sharing, video conference with document sharing, and gaming video, among others [2]. In these applications, 4:4:4 coding is necessary because the subsampling of the chroma components causes color distortion, thus degrading the visual quality.

Recently, following the international video coding standard H.265/MPEG-H Part 2 Version 1, also known as High Efficiency Video Coding (HEVC), the HEVC Range Extensions (HEVC RExt) [3] has been developed to meet the need in video compression to support higher bit-depth and extended chroma formats such as 4:2:2 and 4:4:4. HEVC RExt provides a new way to efficiently compress 4:4:4 contents with high color fidelity by use of prediction across color components, so-called the cross-component prediction (CCP). During HEVC RExt development, CCP has been thoroughly evaluated to show its effectiveness to achieve remarkable coding efficiency improvement both for RGB and YCbCr contents and high color fidelity at the same time, recognized as a beneficial tool in the standard [4], and adopted into HEVC RExt 4:4:4 profiles.

In this paper, the CCP scheme is presented in which the residual signals of the chroma components are predicted from the luma residual signal. To maximally decrease local correlation among components, the scale factor is calculated by linear regression for each transform unit (TU), and it is efficiently coded into bitstream so that there is no calculation burden at the decoder side. It is also shown how to implement CCP for real-time processing systems. Various experiments are performed to evaluate the CCP performance. Since CCP adapts to the input signal characteristics, it provides significant coding efficiency improvement both for RGB and YCbCr sequences. The average coding gains are 17% and 5% bit-rate reduction in case of intra coding, and 11% and 4% in case of inter coding for RGB and YCbCr, respectively. Meanwhile, the average increment of encoding and decoding times computed using the reference HM RExt software [5] are about 10% and 4%, respectively.

The rest of the paper is organized as follows. In Section II, we briefly review the impact of different color spaces on coding performance and review previous works. The overall structure of the CCP scheme and the mathematical justification are given in Section III. Practical implementation details of CCP in HEVC RExt are described in Section IV. Experiments are performed using various test sequences including camera captured natural video and screen content video in both RGB and YCbCr color space. The results are discussed in Section V. Conclusions are given in Section VI.

II. Color Representation in Video Compression

Color video plays an essential role in multimedia systems, where various color spaces are used to efficiently represent color. A color space specifies color with numerical values using multiple components [6], [7]. The RGB color space is a popular color space, where color is represented as a combination of three primary color component values (i.e., red, green, and blue). For color video representation, the YCbCr color space has been widely used. YCbCr can be

W.-S. Kim, W. Pu, J. Sole, J. Chen, and M. Karczewicz are with Qualcomm, Inc., San Diego, USA (e-mail: wskim@ieee.org).

M. Siekmann, T. Nguyen, and D. Marpe are with the Image and Video Coding Group, Video Coding & Analytics Department, Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Berlin 10587, Germany (e-mail: tung.nguyen@hhi.fraunhofer.de; detlev.marpe@hhi.fraunhofer.de).

A. Khairat was with the Image and Video Coding Group, Video Coding & Analytics Department, Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Berlin 10587, Germany (e-mail: ali.khairat@gmail.com).

easily converted from the RGB color space via a linear transformation [8]. One advantage of YCbCr is the backward compatibility with black and white TV as Y signal conveys the luminance information. In addition, chrominance bandwidth can be reduced by subsampling the Cb and Cr components with significantly less subjective impact than subsampling in RGB. Because of these advantages, YCbCr has been the major color space in video compression. However, color transform and subsampling can cause color distortion, which may not be appropriate to keep high color fidelity.

Sullivan [9] analyzed that a color conversion like YCbCr can cause color distortion due to the rounding error during conversion process. In addition, the YCbCr color space is often converted back to the RGB color space for display purposes. The rounding error during this backward conversion impacts with larger distortion than the forward conversion. Previous work in [10] shows that coding directly in the RGB space (without any color conversion) can be better in terms of coding efficiency, especially when the operating bit-rate is high. It is also discussed that the distortion in different color components are evenly distributed so that the color fidelity is better maintained in the RGB domain coding than in YCbCr.

However, usually RGB coding results in very high bit-rates. In contrast to YCbCr where the correlations among the color components are decreased by representing the chroma signals as relative differences to the luma signal, there are often significant amount of redundancies among the color components in RGB, which makes it difficult to achieve high compression efficiency.

There have been efforts to resolve the redundancy problem to achieve better color fidelity with better compression efficiency at the same time. One of the approaches is to improve the color conversion process to reduce the rounding error. The YCgCo transform was proposed in [11], where the conversion error is reduced by using the lifting scheme. Similar efforts have been made in [12] and [13] to improve the performance of the forward/backward color conversion. Even though these color conversion approaches reduce the uneven distribution of distortion among color components by reducing the rounding error, there still exists color distortion due to the rounding process during color conversion.

To further improve both color fidelity and coding efficiency, residual domain processing approaches have been proposed. First, the inter-plane prediction scheme [14], [15] was proposed for RGB coding where a simple prediction is performed between luma and chroma components. This was later extended to the residual color transform [16], where YCgCo transform is applied to the residual signal. These were the first approaches to apply color transform inside the video coding loop. However, the prediction or transform is not locally adaptive, and only adaptive at the sequence level. To further improve coding efficiency the residue sampling was proposed in [17], which was still applied globally at the sequence level. Later, a block adaptive approaches were introduced to better adapt to local characteristics. In [18] it is shown that there is a cross-over point between RGB and YCgCo rate-distortion (R-D) curves, which confirms that generally YCgCo performs better at low bit-rate and RGB coding is better in high bit-rate. By selecting the best color space for each block and applying it to residual signal, the optimal performance is achieved.

In [19], instead of a fixed color conversion, a scale factor and offset are used for each block to predict the chroma signal from the luma. The original signal is used for prediction. In contrast, the reconstructed neighboring samples are used to estimate the prediction parameters in [20]. Also, instead of predicting chroma components from luma component, an adaptive color transform is derived using the singular value decomposition based on the reconstructed neighboring samples in [21]. These approaches place burden on the decoder side, as the parameters for prediction or the transform have to be calculated for each block at the decoder side.

On the other hand, different color conversions are applied at the sequence level in [22]. In this case, it is difficult to find the optimal color space unless they are all tried, which would require significant encoding time and multi-pass coding. Another approach is to treat each color component as an independent frame and perform prediction between them using the existing coding framework like multi-view coding [10]. However, coding efficiency might be sacrificed, as the common information among color components (like partition, prediction mode, or motion vector) has to be redundantly coded.

Compared to these previous works, the advantage of the proposed CCP scheme is that it provides better coding efficiency than YCbCr coding while keeping high color fidelity as in RGB coding. This becomes possible by applying prediction between color components inside coding loop, where the prediction is applied in the residual domain after intra/inter prediction for better decorrelation with local adaptivity by signaling a scale factor. The details are given in the following section.

III. CROSS-COMPONENT PREDICTION

In this section, the proposed CCP method is explained in detail. The overall structure of the coding scheme is explained followed by the details of the prediction scheme using a linear model.

A. Encoder and Decoder Structure with CCP

In CCP, luma component is selected as the predictor component and the two chroma components are predicted from it. For convenience of notation in this paper, the luma component refers to Y or G in case of YCbCr or RGB color space, and the two chroma components are referred to as Cb and Cr or B and R in case of YCbCr or RGB color space. Note that this can be easily applied to other color spaces by selecting a component containing the most energy.

At the encoder side, first the regular intra or inter prediction in video coding is performed to generate the residual signal of each component. Then, the chroma residual signals are predicted by using the reconstructed luma residual signal to further reduce the correlation between luma and chroma components. After that, the difference between original chroma residue and its CCP predictor is coded and the luma residue

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology

is coded without any change caused by CCP. Fig. 1 (a) shows the encoder block diagram with CCP.

At the decoder side, the predictor is generated in the exactly same way as at the encoder side, which is added to the predicted chroma signal to reconstruct chroma residual signal. This is illustrated in Fig. 1 (b)

B. Prediction using a linear model

Now, it is described how the prediction is performed between components. As shown in [14], there is strong correlation among components after intra or inter prediction, and a linear model can be used to predict one from the other. Let us denote the original predictor signal (luma component) as x and the signal to be predicted (chroma component) as y. After intra or inter prediction, the residual signals are formed, which are denoted as Δx and Δy. Then, Δy can be estimated using the linear regression as

$$\widehat{\Delta y} = \alpha \cdot \Delta x + \beta. \tag{1}$$

In (1), α is calculated as

$$\alpha = \frac{E[(\Delta x - \mu_{\Delta x})(\Delta y - \mu_{\Delta y})]}{E[(\Delta x - \mu_{\Delta x})^2]}, \tag{2}$$

where $\mu_{\Delta x}$ and $\mu_{\Delta y}$ are the average of Δx and Δy, respectively, and β can be calculated as

$$\beta = \mu_{\Delta y} - \alpha \cdot \mu_{\Delta x}. \tag{3}$$

Based on the stationary property, we assume that the mean of the reference signal for intra or inter prediction and the mean of the current signal to be the same, which results in zero mean of the residual signal. This makes α in (2) as

$$\alpha = \frac{E[\Delta x \Delta y]}{E[\Delta x^2]}, \tag{4}$$

and β in (3) zero. Even though non-zero β can provide more accuracy, this will incur bitrate increment to code β in addition to α. We verified that this additional accuracy does not help to improve the coding efficiency due to bitrate increment.

At the encoder side, CCP is applied to the chroma residual signal Δy to generate the predicted signal, denoted as $\Delta^2 y$, which is represented as

$$\Delta^2 y = \Delta y - \alpha \cdot \Delta x'. \tag{5}$$

where $\Delta x'$ is the reconstructed luma residual signal. Then, $\Delta^2 y$ is compressed using transform, quantization, and entropy coding, and signaled in the bitstream. At the decoder side, the CCP predicted chroma residual signal is reconstructed as $\Delta^2 y'$. Finally, the chroma residual signal is reconstructed as

$$\Delta y' = \Delta^2 y' + \alpha \cdot \Delta x'. \tag{6}$$

Note that at the encoder side it is possible to use the original luma residual signal for prediction in (5), i.e., Δx instead of $\Delta x'$. This will be explained in detail in Section IV-B.

C. Analysis of correlation among color components

Given the details of the proposed CCP scheme in the above, this section provides the background of the CCP design, especially by explaining the reason for residual domain processing and local adaptivity. Color decorrelation is usually achieved by a color transform (e.g., from RGB to YCbCr) applied outside a coding loop. However, in CCP, as explained above, the residual signal is used for the prediction instead of the original signal, which brings a color decorrelation inside a coding loop. There are several benefits of operating in residual signal domain instead of original signal domain.

First, there is higher correlation among components in residual signal which leads to more coding efficiency improvement. The CCP in HEVC RExt is applied only when the residual signals of luma and chroma components are generated using the same prediction mode. In other words, CCP is not applied when different intra prediction mode is used for different components as less correlation among components is expected. In case of inter prediction, CCP can be always applied as the same prediction mode is applied to all the components in HEVC RExt.

Secondly, compared to the approaches that use the original signal, CCP requires less delay, which facilitates parallel processing. For example, in [20] the correlation in the reconstructed signal is used both at the encoder and decoder sides. This requires full reconstruction of luma signal to encode chroma signal at the encoder side and to reconstruct chroma signal at the decoder side. This disturbs parallel processing of luma and chroma components, since chroma coding cannot be performed until reconstruction of luma component is finished, which could cause significant delay in case of real time encoding processing as well as significant increment of decoder implementation cost. On the other hand, in CCP parallel processing of luma and chroma components is facilitated, as it does not require full reconstruction of the predictor component.

Thirdly, residual domain process causes less rounding error compared to the outside color transform, which provides ability to keep higher color fidelity. This is because the amount of information is significantly reduced after intra/inter prediction, which reduces the rounding error accordingly. For example, the residual signal tends to be sparser than in the original signal. If the residual signal is zero, the rounding error becomes zero.

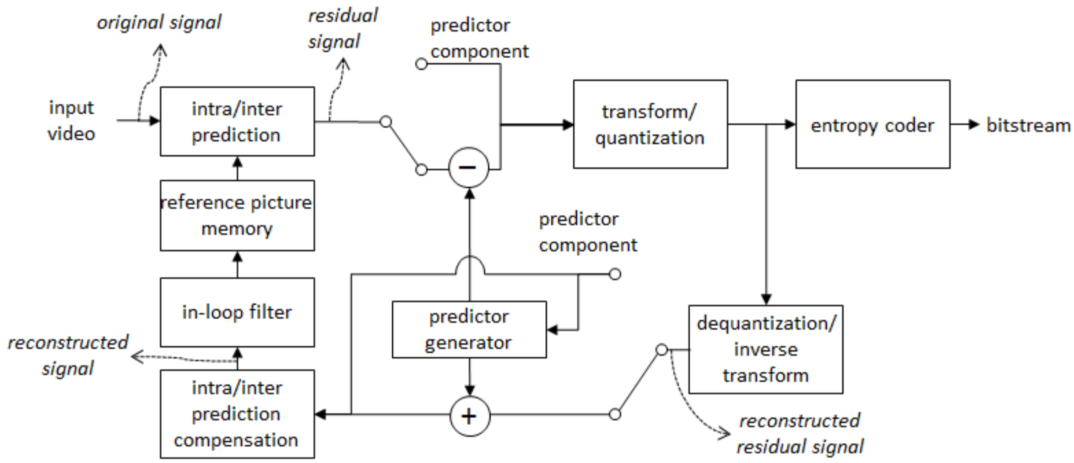
Besides these, in-loop color decorrelation like CCP enables local adaptivity by adjusting the scale factor, α. When local correlation is very low, it is better to code each component separately. In this case, α is set to 0 not to apply any prediction across components. On the other hand, when local correlation is high, α is increased for higher decorrelation.

IV. CCP IN HEVC REXT

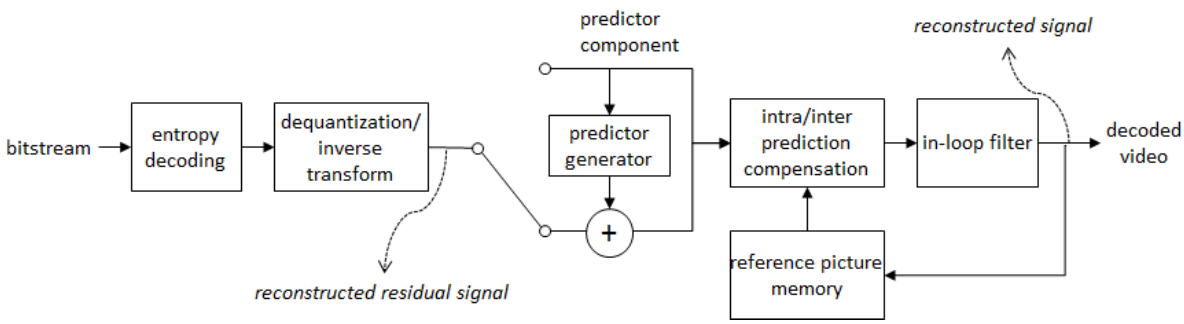
A. Description of CCP process

This section describes how CCP is specifically defined in HEVC RExt in detail. First, luma component is set as the predictor component, and two chroma components are predicted separately from the luma component. Therefore,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology



(a) Encoder.



(b) Decoder.

Fig. 1. Block diagram of encoder and decoder with CCP.

there are two α values, one for the Cb or B component, and the other one for the Cr or R component. These values are coded into bitstream, so there is no need to calculate these values at the decoder side. The signaling of α occurs at TU level in order to maximally decrease the local correlation.

As α is signaled, it is necessary to limit its range and precision by clipping and quantization for efficient compression. The correlation among components can vary depend on the input color space and sequence characteristics. For example, the distribution of α calculated using (4) is quite different for YCbCr and RGB sequences: the distributions is narrower and mostly centered around zero for YCbCr, while wider and centered around one for RGB. To capture these characteristics, α is allowed to have both negative and positive value within the range of [-1, 1] and nonlinearly quantized within the set $\{0, \pm0.125, \pm0.25, \pm0.5, \pm1\}$. To avoid floating point operation in practice, the α value is multiplied by eight $\bar{\alpha} = \{0, \pm1, \pm2, \pm4, \pm8\}$, and a right shift by 3 is applied to compensate for it as follows,

$$\text{Encoder}: \Delta^2 y = \Delta y - (\bar{\alpha} \cdot \Delta x) >> 3, \tag{7}$$

$$\text{Decoder}: \Delta y' = \Delta^2 y' + (\bar{\alpha} \cdot \Delta x') >> 3, \tag{8}$$

where $\bar{\alpha}$ is the scaled α, and $>>$ is the right-shift operator.

The HEVC RExt CCP process at the decoder side is specified as below. It incorporates flexibility to handle different bit-depths for the luma and chroma samples,

$$\text{Decoder}: \Delta y' = \Delta^2 y' + (\bar{\alpha} \cdot ((\Delta x' << b_{\text{chroma}}) >> b_{\text{luma}})) >> 3, \tag{9}$$

where b_{luma} and b_{chroma} are the bit-depth of luma and chroma samples, respectively.

For efficient coding of $\bar{\alpha}$ using context-based adaptive binary arithmetic coding (CABAC), bins are generated as follows. The first bin indicates whether $\bar{\alpha}$ is zero or not. If not, $log_2|\bar{\alpha}| + 1$ is coded with a truncated unary binarization, followed by one bin to indicate the sign of $\bar{\alpha}$. If the luma residual signal of the current TU is all zero, $\bar{\alpha}$ is set to zero and CCP is not applied. Furthermore, in case of intra, CCP is only applied when luma and chroma share the same prediction mode. Otherwise, $\bar{\alpha}$ signaling does not occur.

B. Encoder implementation

This section studies efficient ways to implement CCP to facilitate real-time processing and to improve coding efficiency.

In Fig. 1 (a), the predictor for CCP is generated using the reconstructed residual signal of the luma component, and it is subtracted from the original chroma residual signal. This implies that the scale factor, α is calculated based on the correlation between the reconstructed luma residue and the chroma residue, and the prediction is performed using the reconstructed luma residue. As the reconstructed luma residual signal is used to compensate the prediction during decoding process, this provides a precise estimation of the chroma residual signal. However, this layout imposes delay of the chroma processing, as the derivation of α cannot be performed until the luma residual signal is reconstructed. Fig. 2 (a) illustrates the delay caused by CCP in this case.

Alternatively, α can be calculated using the original residual signal. In this case, once intra/inter prediction is performed for all components, α can be calculated based on the original luma and chroma residual signals. And then the CCP prediction can be performed once the luma residual signal is reconstructed. This scheme facilitates parallel encoding processing of luma and chroma components. Fig. 2 (b) shows the processing steps for this approach.

It is even possible to calculate α before CCP process block in Fig. 2 to further reduce complexity. For example, if α is calculated using the original signals of luma and chroma before intra/inter prediction, the delay caused by CCP is further reduced, as the calculation can be performed in parallel along with intra/inter prediction. Below is the description of α derivation based on the correlation in the original luma and chroma signals.

First, let $\hat{y}$ represent the estimation of the original chroma signal using the original luma signal x by linear regression as

$$\hat{y} = \alpha' \cdot x + \beta'. \tag{10}$$

Under the stationarity assumption, the reference signal for intra/inter prediction can be represented using the same α' and β' as

$$\widehat{y_{\mathrm{ref}}} = \alpha' \cdot x_{\mathrm{ref}} + \beta'. \tag{11}$$

By subtracting (11) from (10), the following is obtained

$$\hat{y} - \widehat{y_{\mathrm{ref}}} = \alpha' \cdot x - \alpha' \cdot x_{\mathrm{ref}} = \alpha' \cdot (x - x_{\mathrm{ref}}) = \alpha' \cdot \Delta x. \tag{12}$$

By comparing (12) to (1) with β set to zero, it is noticed that α can be approximated to α' as

$$\alpha \approx \alpha', \tag{13}$$

where α' can be calculated using the original signals as

$$\alpha' = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - (\sum x)^2}, \tag{14}$$

where N is the number of pixels in a block.

For optimal parameter selection an encoder implementation of CCP can perform R-D optimization checking the R-D cost of every possible α values, which would require significant amount of complexity. Alternatively, R-D optimization can be performed by comparing only two cases, CCP on and off. If the R-D cost is smaller when CCP is not applied, $\bar{\alpha}$ is set to zero to turn CCP off. Otherwise, CCP is applied with a non-zero $\bar{\alpha}$. When the derived $\bar{\alpha}$ is zero, there is no need to perform R-D selection between CCP on and off.

Note that other techniques can be used to calculate α to further reduce complexity or improve performance, for example, by use of neighboring blocks, with potentially some performance penalty.

V. EXPERIMENTAL RESULTS

In this section, experiments are performed to evaluate the performance of CCP. The proposed CCP is implemented on top of HEVC RExt [3] reference software HM-12.0 RExt 4.1. The simulation is performed using the common test condition used during the standardization of HEVC RExt [5], which includes various RGB and YCbCr 4:4:4 test sequences. Table I shows the BD bit-rate [23] results for different test configurations of all intra, random access, and low delay B, where a positive number implies bit-rate reduction ratio compared to the anchor. For the YCbCr results, both anchor and test are coded in YCbCr, and PSNR is computed in the YCbCr color space. Full results including detailed results of each sequence are available in [24] and [25].

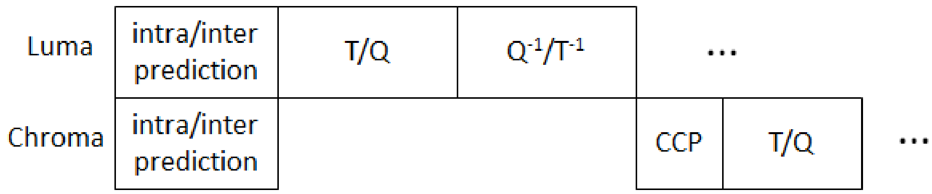
Table I shows significant coding gains are provided by CCP. Gains for RGB sequences are larger than those for YCbCr sequences. The reason is that there is generally higher correlation among components in the RGB color space than in the YCbCr space. The coding gains for YCbCr sequences are still notable even though there is less correlation in this space. This reveals CCP efficiently removes correlation among color components by applying prediction in the residual domain with local adaptivity.

It is also observed in the YCbCr case that the B-D rate gains for chroma components are larger than luma. Usually, in 4:2:0 chroma format, the coding gains in chroma are less important than luma gains, because the bit-rate occupancy of chroma components is much less than that of luma component. However, the coding gains of chroma components in 4:4:4 becomes more meaningful because of the increased bit-rate occupancy of chroma components due to not only four times larger chroma signal resolution but also more details kept without downsampling.

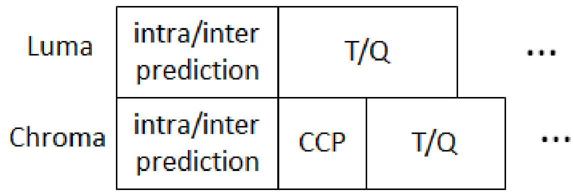
From Table I it can be also noticed that the coding gain in intra coding is larger than inter coding. Usually, intra coding requires significantly higher bit-rate than inter coding, which makes it difficult to control bit-rate and manage bitstream buffer. Also the quality of intra-coded frames is very important as they are used as a reference frame, thus affect the performance of the following inter-coded frames. As the intra prediction performance is less efficient than inter prediction performance, there tends to be more information in residual signal with high correlation among components. CCP improves the performance of intra coding significantly by compensating less efficient prediction performance. In case of random access configuration, the coding gain is larger than low delay B as more intra-coded frames are used. The coding gain in low delay B shows that CCP also provides good coding gain in inter coding as well. To verify the performance of CCP in inter coding, additional experiment is performed where CCP is applied only to inter-coded coding units. In this case, the coding gains under low delay B configuration are 10.9/7.3/8.1 % (R/G/B) for RGB 4:4:4 and 0.1/2.4/3.7 % (Y/Cb/Cr) for YCbCr 4:4:4, respectively, which shows that CCP provides good coding gain for inter coding.

Table I shows the average encoding and decoding time increment compared to the anchor. The encoding time in-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology



(a) CCP using the reconstructed luma residual signal.



(b) CCP using the original luma residual signal.

Fig. 2. Delay analysis in CCP.

crement is mainly due to R-D optimization, where CCP on and off cases are compared. The decoding time increment by CCP is 4-5% measured using the reference software. Reduced complexity could be achieved by an implementation that further exploits the parallelization capabilities of the CCP algorithm.

TABLE I
BD BIT-RATE (%) RESULTS OF THE PROPOSED METHOD.

		Y or G	Cb or B	Cr or R
All intra	RGB 4:4:4	18.1	16.6	17.3
	YCbCr 4:4:4	1.4	6.6	6.9
	Enc Time	116 %		
	Dec Time	104 %		
Random access	RGB 4:4:4	14.1	11.7	13.5
	YCbCr 4:4:4	0.5	8.2	7.2
	Enc Time	111 %		
	Dec Time	104 %		
Low delay B	RGB 4:4:4	12.6	8.7	10.1
	YCbCr 4:4:4	0.2	5.8	5.4
	Enc Time	110 %		
	Dec Time	105 %		

Additional experiments are performed to evaluate CCP for screen contents coding. Various test sequences are used according to the common test condition for screen contents coding standardization [26]. Table II shows the BD bit-rate results for screen contents coding. Gains are larger than for natural content. A reason for that is that screen contents have less noise than natural contents, so there is more correlation among components.

In Section IV-B, two alternative encoder implementation methods are presented, where the scale factor, α, is calculated using the original residue or the original pixel. Table III shows the performance of these methods on YCbCr 4:4:4 compared to the results in Table II. This shows that when the original residue is used, the coding loss is negligible, while the delay can be significantly reduced. When the original pixel is used for further complexity reduction, there is a little more loss.

Fig. 3 illustrates the R-D curves of one of the coded sequences. The horizontal axis is the total bit-rate and the vertical axis is the PSNR, for each of the three components and its average. In case of YCbCr, PSNR is measured after converting the decoded YCbCr to RGB. From Fig. 3, it can be noticed that there is an upper bound in the PSNR that YCbCr coding can achieve. This is due to rounding error during the color conversion process. It can be also noticed that at the same bit-rate, PSNR is different component by component in YCbCr coding. The PSNR is more consistent in RGB coding, leading to better color fidelity. However, at low

TABLE II
BD BIT-RATE (%) RESULTS OF THE PROPOSED METHOD FOR SCREEN CONTENTS.

		Y or G	Cb or B	Cr or R
All intra	SC RGB 4:4:4	26.0	25.7	26.0
	Animation RGB 4:4:4	25.9	25.7	24.5
	SC YCbCr 4:4:4	4.5	8.0	7.9
	Animation YCbCr 4:4:4	2.6	11.5	7.2
	RangeExt	1.7	5.6	11.1
	SC RGB Optional	24.2	24.5	24.8
	SC YCbCr Optional	7.2	8.3	7.3
	Enc Time	113 %		
	Dec Time	100 %		
Random access	SC RGB 4:4:4	23.2	23.1	23.5
	Animation RGB 4:4:4	20.0	19.9	18.4
	SC YCbCr 4:4:4	3.5	7.8	7.8
	Animation YCbCr 4:4:4	1.0	11.1	6.3
	RangeExt	0.8	8.5	11.2
	SC RGB Optional	23.4	23.5	23.7
	SC YCbCr Optional	6.0	7.9	6.4
	Enc Time	109 %		
	Dec Time	102 %		
Low delay B	SC RGB 4:4:4	20.9	20.3	20.8
	Animation RGB 4:4:4	18.9	16.9	15.6
	SC YCbCr 4:4:4	2.7	6.8	7.1
	Animation YCbCr 4:4:4	0.4	9.3	4.1
	RangeExt	0.4	5.2	7.5
	SC RGB Optional	21.1	21.1	21.5
	SC YCbCr Optional	8.6	10.2	9.1
	Enc Time	107 %		
	Dec Time	102 %		

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology

TABLE III
BD BIT-RATE (%) RESULTS OF ALTERNATIVE ENCODER IMPLEMENTATION METHODS, WHERE THE SCALE FACTOR, α, IS CALCULATED USING THE ORIGINAL RESIDUE OR THE ORIGINAL PIXEL, COMPARED TO YCBCR 4:4:4 RESULTS IN TABLE II.

		Y	Cb	Cr
Original residue	All intra	0.0	0.5	0.5
	Random access	0.1	0.5	0.5
	Low delay B	0.0	0.3	0.3
Original pixel	All intra	0.5	0.9	0.8
	Random access	0.4	1.5	1.1
	Low delay B	0.4	1.5	1.5

bit-rate, YCbCr coding shows better compression efficiency than RGB coding. As bit-rate increases, a cross-over happens between the R-D curves. By using CCP, the coding efficiency is greatly improved compared to RGB coding making the cross-over happen at significantly lower bit-rate. It can be noticed that CCP maintains the benefit of RGB coding by preserving color fidelity. At very low bit-rates, YCbCr can still perform better than RGB+CCP case. However, CCP can also be applied directly to YCbCr coding to improve coding efficiency.

Fig. 4 shows more R-D curves of various screen contents, where the vertical axis is the average PSNR. These plots also show that the YCbCr+CCP case outperforms YCbCr case, while there is still the cross-over between RGB and YCbCr coding. Sometimes, the R-D curves for RGB coding do not show a smooth arc shape. The reason is that the current HEVC RExt reference software is mainly optimized for YCbCr coding at low to middle bit-rate. Future work could include the optimization of the CCP encoder for the RGB coding performance.

Table IV shows the bit-rate results for lossless coding, where the simulation is performed according to the common test condition in [27]. CCP shows good coding gain in case of lossless coding as well.

TABLE IV
TOTAL BIT-RATE SAVINGS (%) BY CCP IN LOSSLESS CODING.

	All intra	Random access	Low delay B
SC RGB 4:4:4	16.3	9.4	7.3
Animation RGB 4:4:4	12.7	2.6	1.6
SC YCbCr 4:4:4	1.9	1.0	0.6
Animation YCbCr 4:4:4	1.8	0.7	0.6
RangeExt	1.2	0.1	0.0
SC RGB Optional	24.4	35.7	37.1
SC YCbCr Optional	2.6	0.5	0.0

Table V shows the BD bit-rate results for YCbCr 4:2:0 coding. The common test condition used for HEVC Version 1 is followed [28]. Note that CCP is adopted in HEVC RExt only for 4:4:4 chroma subsampling format. However, CCP can be applied to other chroma formats by subsampling the predictor component. Results in Table V reveal that there still can be noticeable coding gain in YCbCr 4:2:0 coding, especially in chroma components.

TABLE V
BD-BIT-RATE (%) RESULTS OF THE PROPOSED METHOD FOR YCBCR 4:2:0 CODING.

		Y	Cb	Cr
All intra	Class A	1.1	22.9	18.3
	Class B	0.2	6.4	2.6
	Class C	0.0	2.7	3.1
	Class D	-0.1	2.2	2.1
	Class E	-0.1	0.8	1.5
	Class F	0.4	2.6	3.2
	Overall	0.3	6.5	5.2
Random access	Class A	-0.1	29.9	24.9
	Class B	-0.1	9.3	3.1
	Class C	-0.2	3.6	3.7
	Class D	-0.3	3.5	2.7
	Class F	0.0	3.9	4.1
	Overall	-0.1	10.0	7.5
Low delay B	Class B	-0.2	4.2	1.2
	Class C	-0.3	1.7	1.3
	Class D	-0.4	1.4	1.6
	Class E	-0.5	0.2	1.8
	Class F	0.2	2.7	2.8
	Overall	-0.2	2.2	1.7

VI. CONCLUSION

Coding in the RGB color space provides better color fidelity compared to the YCbCr coding, while YCbCr coding can achieve better coding efficiency. In this paper, CCP is proposed where the correlation among color components are efficiently removed by applying prediction between luma and chroma residual signals using the linear regression. Experimental results show that the proposed method achieves great coding efficiency improvements for RGB coding and maintains color fidelity. CCP also provides good coding gain when applied to YCbCr coding, as the proposed scheme can adapt to the given color space through the signaling of the scale factor. The computational and implementation complexities are also analyzed to show that the proposed CCP can be implemented and used in real-time processing systems. In the future, more study can be made on how to facilitate software and hardware implementation. We also leave it as a further study how to efficiently apply CCP to various applications including high dynamic range video coding.

REFERENCES

[1] G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding standard: Overview and introduction to the Fidelity Range Extensions," in *Proc. of SPIE Applications of Digital Image Processing XXVII*, Aug. 2004.

[2] C. Lan, G. Shi, and F. Wu, "Compress compound images in H.264/MPEG-4 AVC by exploiting spatial correlation," *IEEE Trans. Image Proc.*, vol. 19, no. 4, pp. 946–957, Apr. 2010.

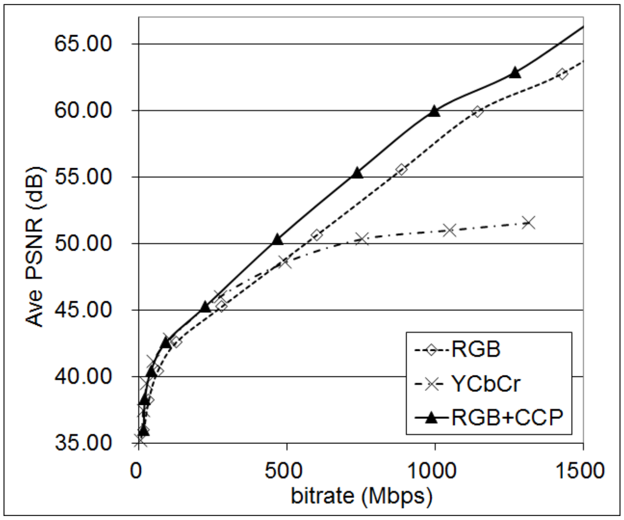
[3] *High Efficiency Video Coding (HEVC) Range Extensions text specification: Draft 7*, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Document JCTVC-Q1005, Mar. 2014.

[4] S. Lee, C. Park, E. Alshina, C. Kim, and K. McCann, "AHG5: Recommended profiling of range extension coding tools," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-Q0051, Mar. 2014.

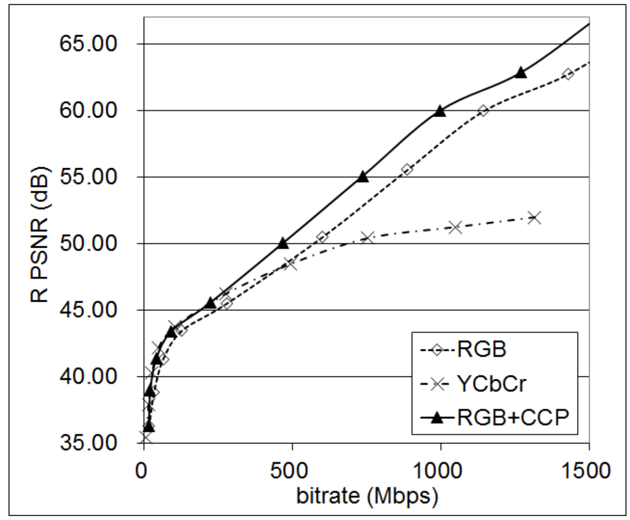
[5] D. Flynn, K. Sharman, and C. Rosewarne, "Common test conditions and software reference configurations for HEVC range extensions," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-N1006, Aug. 2013.

[6] A. Ford and A. Roberts, "Colour space conversions," University of Westminster, London, Tech. Rep., Aug. 1998.

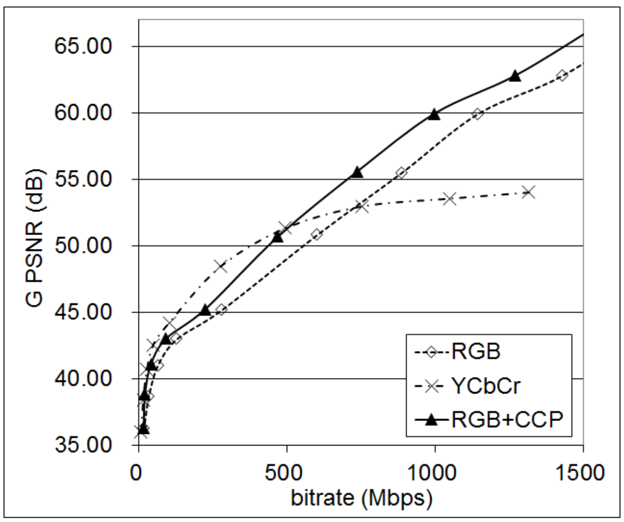
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology



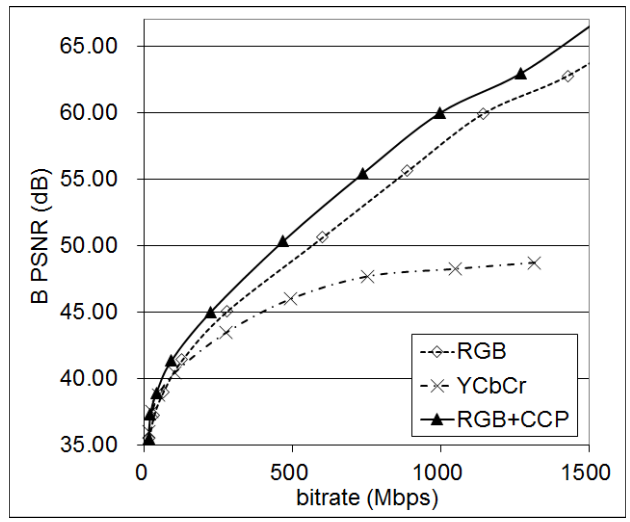
(a) Average PSNR



(b) R PSNR



(c) G PSNR



(d) B PSNR

Fig. 3. Comparison of the R-D curves of EBULupoCandleLight among YCbCr, RGB, and RGB with CCP coding. x-axis: total bit-rate; y-axis: PSNR.

[7] C. Poynton, "A guided tour of color space," in *New Foundations for Video Technology: Proceedings*. San Francisco, CA, USA: Society of Motion Picture & Television Engineers, Feb. 1995.

[8] *Parameter values for the HDTV standards for production and international programme exchange*, ITU-R Recommendation BT.709-5, Apr. 2002.

[9] G. Sullivan, "Approximate theoretical analysis of RGB to YCbCr to RGB conversion error," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-I017, Sep. 2003.

[10] A. Minezawa, S. Sekiguchi, and T. Murakami, "AHG5/AHG8: On RGB to YCbCr conversion for screen contents," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-N0115, Aug. 2013.

[11] H. Malvar and G. Sullivan, "YCoCg-R: A color space with RGB reversibility and low dynamic range," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-I014, Sep. 2003.

[12] P. Topiwala and C. Tu, "New invertible integer color transforms based on lifting steps and coding of 4:4:4 video," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-I015, Sep. 2003.

[13] H. M. Kim, W.-S. Kim, and D.-S. Cho, "A new color transform for RGB coding," in *Proc. of IEEE Int. Conf. Image Proc., ICIP 2004*, Singapore, Oct. 2004.

[14] W.-S. Kim, D.-S. Cho, and H. M. Kim, "Inter-plane prediction for RGB video coding," in *Proc. of IEEE Int. Conf. Image Proc., ICIP 2004*, Singapore, Oct. 2004.

[15] ——, "Inter-plane prediction for RGB coding," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-I023, Sep. 2003.

[16] W.-S. Kim and H. M. Kim, "Residue transform," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-J038, Dec. 2003.

[17] ——, "Residue sampling for image and video compression," in *Proc. of SPIE Visual Communications and Image Processing 2005*, 2005, pp. 12–18.

[18] D. Marpe, H. Kirchhoffer, V. George, C. Fehn, P. Kauff, and T. Wiegand, "MB-adaptive residual colour transform for 4:4:4 coding," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document JVT-R071, Jan. 2006.

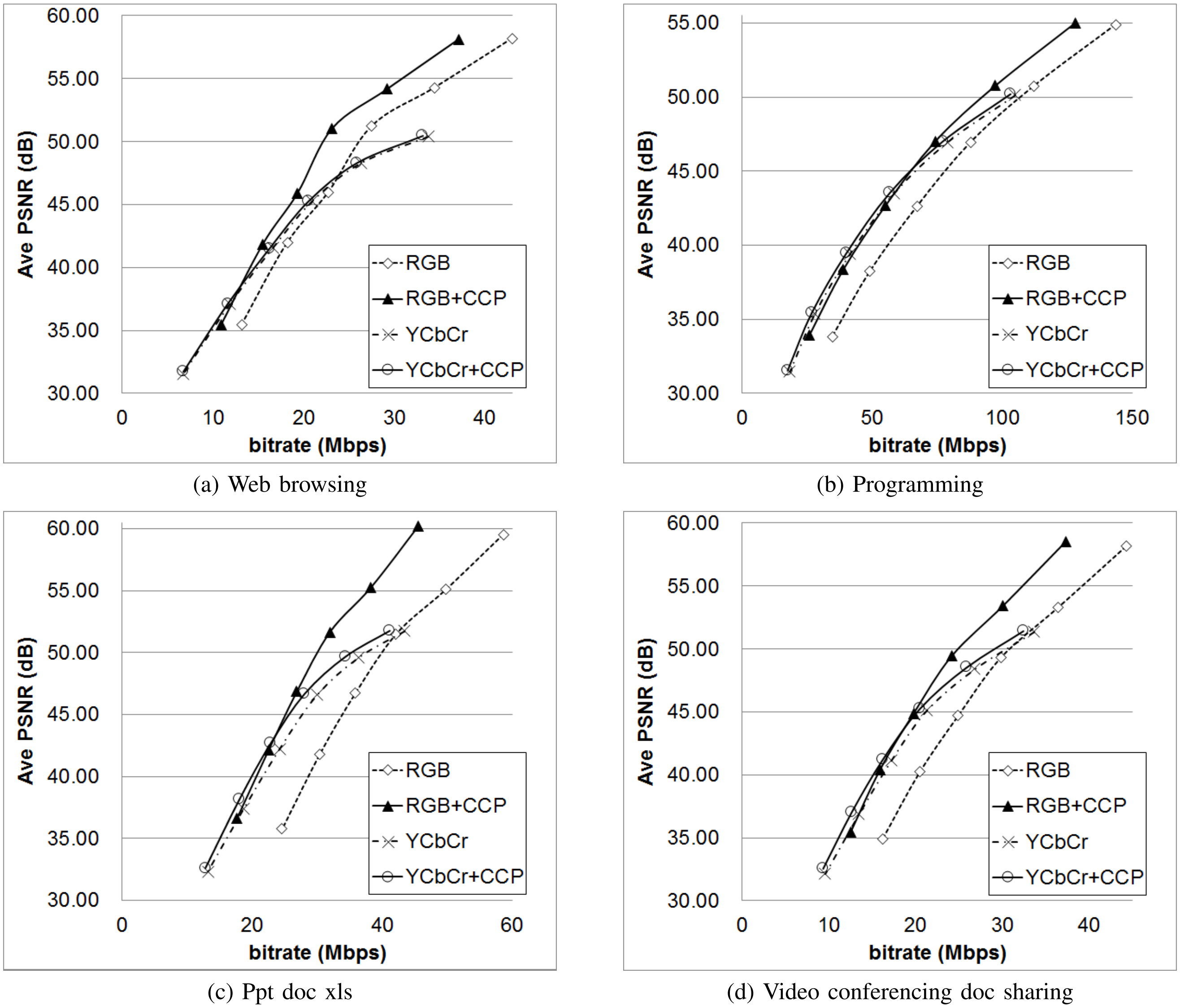
[19] B. C. Song, Y. G. Lee, and N. H. Kim, "Block adaptive inter-color compensation algorithm for RGB 4:4:4 video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 10, pp. 1447–1451, Oct. 2008.

[20] J. Kim and B. Jeon, "RCE1: EXP 1/4 The performance of extended chroma mode for non 4:2:0 format," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-N0227, Aug. 2013.

[21] K. Kawamura, T. Yoshino, and S. Naito, "AHG7: In-loop color-space transformation of residual signals for range extensions," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-L0371, Jan. 2013.

[22] W. Dai, M. Krishnan, and P. Topiwala, "RCE1: Adaptive color transforms for Range Extensions," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-M0048, Apr. 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology



(a) Web browsing

(b) Programming

(c) Ppt doc xls

(d) Video conferencing doc sharing

Fig. 4. Comparison of the rate-distortion curves of EBULupoCandleLight among YCbCr, RGB, and RGB with CCP coding. x-axis: total bit-rate; y-axis: PSNR.

[23] G. Bjøntegaard, “Calculation of average PSNR differences between RD-curves,” ITU-T SG.16 Q.6, Document VCEG-M33, Apr. 2001.

[24] W.-S. Kim and T. Nguyen, “RCE1: Summary report of HEVC Range Extensions core experiment 1 on inter-component decorrelation methods,” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-O0035, Oct. 2013.

[25] W. Pu, W.-S. Kim, J. Chen, J. Sole, and M. Karczewicz, “RCE1: Descriptions and results for experiments 1, 2, 3, and 4,” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-O0202, Oct. 2013.

[26] H. Yu, M. Budagavi, R. Cohen, A. Duenas, T. Lin, and J. Xu, “JCT-VC AHG report: Screen content coding (AHG8),” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-N0008, Aug. 2013.

[27] A. Saxena, D.-K. Kwon, M. Naccari, and C. Pang, “HEVC Range Extensions core experiment 3 (RCE3): Intra prediction techniques,” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-N1123, Aug. 2013.

[28] F. Bossen, “Common test conditions and software reference configurations,” ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document JCTVC-L1100, Jan. 2013.

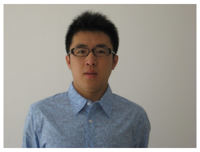
Woo-Shik Kim (SM’12) received the B.Eng. and M.Eng. degrees from the Department of Electronics Engineering, Sogang University, Korea, in 1999 and 2001, respectively, and the Ph.D. degree from the Department of Electrical Engineering, University of Southern California, CA, USA, in 2011.

He was a Senior Researcher with the Samsung Advanced Institute of Technology, from 2001 to 2006, where he contributed to the development of the international video coding standard H.264/MPEG-4 AVC. He worked as a Research Intern with the Mitsubishi Electronics Research Laboratory, Dolby Laboratories, and the Technicolor Laboratory, developing technologies for high-dynamic-range video and 3D video coding. He was a Member of Technical Staff with the Systems and Applications R&D Center, Texas Instruments Inc., TX, USA, working on the High Efficiency Video Coding standardization. He joined Qualcomm Technologies, Inc., in 2013, as a Senior Staff Engineer, and has worked on developing video processing algorithms.

He holds about 60 US patents. He received the Top 10% Paper Award at the 2011 IEEE International Workshop on Multimedia Signal Processing, and the Best Paper Award at the Samsung Technical Conference in 2004.

His research interests are image and video signal processing and coding, and an efficient multimedia system design for real-time processing under low power constraint.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology



Wei Pu (M'08-SM'15) received the B.S. and Ph.D. degrees in electrical and computer engineering from University of Science and Technology of China, Hefei, Anhui, P. R. China, in 2004 and 2009, respectively. From 2010 to 2012, he was a research scientist in department of computer science and engineering, State University of New York at Buffalo. Since 2012, he has been working on inventing and implementing new algorithms on video compression for HEVC and its extensions in Qualcomm. His current research interests include data compression, streaming, and large scale distributed system. He has over 100 publications, 4 granted patents, and over 60 pending patent applications.

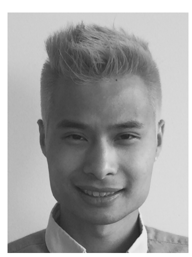
Ali Khairat earned his B.Sc. degree in Digital Media Engineering and Technology from the German University in Cairo, Egypt in 2008. He received his M.Sc. degree in Systems of Information and Multimedia Technology from the Friedrich-Alexander University of Erlangen-Nuremberg, Germany, in 2011. From 2012-2015, he has worked with the Image and Video Coding Group, in the Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Berlin. His research interests include video coding for range extensions for HEVC.

Mischa Siekmann received the degree 'Dipl.-Ing. in Electrical Engineering' from Technische Universität Braunschweig, Germany, in 2008. His studies included an exchange with the University of Southampton, U.K., and an internship with British Sky Broadcasting, London, U.K. After university he joined the Image and Video Coding Group at Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Berlin, Germany. His major research interests include image and video compression, human visual perception for video coding, and signal processing.

Joel Sole (M'01) received the M.Sc. degree in telecommunications from the Technical University of Catalonia (UPC), Barcelona, Spain, and Telecom ParisTech, Paris, France, and the Ph.D. degree from UPC in 2006. From 2006 to 2010, he was with Thomson Corporate Research, Princeton, NJ, initially as a Post-Doctoral Fellow and later as a Staff Member and Senior Scientist. Since 2010, he is with Qualcomm Incorporated, San Diego, as Senior Staff Engineer and Manager. He has participated in several standardization activities in MPEG and JCT-VC, including H.265/HEVC, HEVC RExt, Screen Content Coding and High Dynamic Range video coding.

Jianle Chen received the B.S. and Ph.D. degrees in information and communication system from Zhejiang University, Hangzhou, China, in 2001 and 2006, respectively. Since 2011, he has been a Senior Staff Engineer with Qualcomm, San Diego, CA. From 2006 to 2011, he was a senior research engineer with the Multimedia Platform Laboratory, Samsung Electronics, Suwon, Korea. Since 2006, he has been involved in the development of various video coding standards, including the HEVC standard and its scalable and screen content coding extensions. He was an editor of the SHVC Specification and Test Model. His research interests include video coding, processing and transmission. He is an IEEE Senior Member.

Marta Karczewicz received her M.Sc. and Ph.D. (with honors) degrees from the Tampere University of Technology, Finland, in 1994 and 1997, respectively. In 2006 she joined Qualcomm, where she is a Vice President of Technology in the Multimedia R&D and Standards group. From 1996 to 2006, she was a Senior Research Manager in Nokia Research Center. She was actively participating in work on MPEG-4 AVC and SVC standards within Joint Video Team (JVT) and HEVC standard within Joint Collaborative Team on Video Coding (JCT-VC). She holds 85 U.S. patents on video compression and has authored or co-authored over 30 publications, receiving EURASIPs 1997 Award for the Best Paper of the Year. Her research interests include image and video compression, rate-distortion optimization, scalar quantization, and entropy coding.



Tung Nguyen received his diploma degree in computer science (Dipl.-Inf.) from the Technical University Berlin in 2008. He was working for the Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute (HHI) as a student assistant during his study. He joined the Image Communication group as a research associate after finishing his study.

From 2010 to 2014, he was involving in the standardization activity of the Joint Collaborative Team on Video Coding (JCT-VC). It has been jointly founded by ITU-T SG16/Q6 (VCEG) and ISO/IEC JCT1/SC29 WG11 (MPEG). The JCT-VC's mandate is the development of the High Efficiency Video Coding (HEVC) video compression standard. During this time, he successfully contributed relevant coding tools to both versions of the HEVC standard.

Currently, he is a member of the Image and Video Coding group that belongs to the Video Coding & Analytics department at the HHI. His research interests are image and video processing and their efficient representation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2015.2496821, IEEE Transactions on Circuits and Systems for Video Technology

Detlev Marpe (M'00-SM'08-F'15) received the Dipl.-Math. degree (Highest Hons.) from the Technical University of Berlin, Germany, and the Dr.-Ing. degree from the University of Rostock, Germany.

In 1999, he joined the Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute (HHI), Berlin, where he is currently Head of the Video Coding & Analytics Department as well as Head of the Image and Video Coding Group. Since 1998, he has been an active contributor to the standardization activities of the ITU-T Visual Coding Experts Group, the ISO/IEC Joint Photographic Experts Group, and the ISO/IEC Moving Picture Experts Group for still image and video coding. In the development of the H.264/AVC standard, he was a Chief Architect of the CABAC entropy coding scheme, as well as one of the main technical and editorial contributors to the so-called fidelity range extensions with the addition of the High Profile in H.264/AVC. He was one of the key people in designing the basic architecture of Scalable Video Coding and Multiview Video Coding as algorithmic and syntactical extensions of H.264/AVC. During the recent development of H.265/HEVC, Dr. Marpe also successfully contributed to the first model of the corresponding standardization project and further refinements. In addition, he also made successful proposals to the standardization of its Range Extensions and 3D Extensions. The academic work of Dr. Marpe includes more than 200 publications in the areas of image and video coding. He also holds more than 250 internationally issued patents as well as numerous patent applications in this field. His current research interests include still image and video coding, signal processing for communications as well as computer vision and information theory.

Dr. Marpe was a co-recipient of two Technical Emmy Awards in recognition of his role as a key contributor and a co-editor of the H.264/AVC standard in 2008 and 2009, respectively. He received the IEEE Best Paper Award at the 2013 IEEE International Conference on Consumer Electronics – Berlin in 2013 and the SMPTE Journal Certificate of Merit in 2014. He was nominated for the German Future Prize 2012 and he was a recipient of the 2011 Karl Heinz Beckurts Award, the 2009 Best Paper Award of the IEEE Circuits and Systems Society, the Joseph-von-Fraunhofer Prize in 2004, and the Best Paper Award of the German Society for Information Technology in 2004. Since 2014, he serves as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.