



HAL
open science

Inter-Layer Prediction of Color in High Dynamic Range Image Scalable Compression

Mikaël Le Pendu, Christine Guillemot, Dominique Thoreau

► **To cite this version:**

Mikaël Le Pendu, Christine Guillemot, Dominique Thoreau. Inter-Layer Prediction of Color in High Dynamic Range Image Scalable Compression. *IEEE Transactions on Image Processing*, 2016, 25, pp.3585 - 3596. 10.1109/TIP.2016.2571559 . hal-01388961

HAL Id: hal-01388961

<https://inria.hal.science/hal-01388961>

Submitted on 27 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inter-Layer Prediction of Color in High Dynamic Range Image Scalable Compression

Mikaël Le Pendu, Christine Guillemot, and Dominique Thoreau

Abstract—This paper presents a color inter layer prediction (ILP) method for scalable coding of High Dynamic Range (HDR) video content with a Low Dynamic Range (LDR) base layer. Relying on the assumption of hue preservation between the colors of an HDR image and its LDR tone mapped version, we derived equations for predicting the chromatic components of the HDR layer given the decoded LDR layer. Two color representations are studied. In a first encoding scheme, the HDR image is represented in the classical Y’CbCr format. In addition, a second scheme is proposed using a colorspace based on the CIE u’v’ uniform chromaticity scale diagram. In each case, different prediction equations are derived based on a color model ensuring the hue preservation. Our experiments highlight several advantages of using a CIE u’v’ based colorspace for the compression of HDR content, especially in a scalable context. In addition, our inter-layer prediction scheme using this color representation improves on the state of the art ILP method which directly predicts the HDR layer u’v’ components by computing the LDR layers u’v’ values of each pixel.

Index Terms—High Dynamic Range (HDR), Tone Mapping, Color Correction, Scalability, HEVC, Inter-Layer Prediction (ILP)

I. INTRODUCTION

The emergence of High Dynamic Range technology involves new standardization efforts, both in the numerical representation of HDR pixels and in the field of image and video compression. In terms of content distribution, the question of backward compatibility is critical for enabling optimal rendering on both Low Dynamic Range and HDR displays. In particular, this article deals with the scalable compression of HDR images or videos from a LDR base layer. The focus is set on the numerical representation and the prediction of color information of the HDR enhancement layer, given the decoded LDR layer.

As regards the representation of color, the use of Y’CbCr colorspaces prevails in the compression of digital images and videos. It has several advantages such as the reduction of the redundancy between RGB color components. The separation of an image data into a luma and two chroma channels also enables the down-sampling of the chroma signal for which the human visual system is less sensitive. Finally, the straightforward conversion from Y’CbCr to displayable RGB components makes this representation essential in video coding. However, for HDR images, a Y’CbCr encoding including chroma down-sampling may introduce several types of distortions identified by Poynton et al. in a recent work [1]. The main reason is that the chroma CbCr components are not well decorrelated from the luminance. Although the same issues already occur with LDR images, the artifacts are

much less visible. Therefore, in the context of the distribution of HDR content, other color encoding schemes are emerging, based on the CIE 1976 Uniform Chromaticity Scale (i.e. u’v’ color coordinates). This representation is called uniform because of its increased perceptual uniformity in comparison to the CIE xy chromaticity scales. Furthermore, similarly to the xy coordinates, the CIE u’v’ components have no correlation with luminance, which enables a better separation of the chromatic and achromatic signals than a Y’CbCr colorspace. Therefore, downsampling the u’v’ components has no effect on the luminance. The downside, however, is that perceptual uniformity is not fully satisfied because of the loss of color sensitivity of the human eye at low luminance levels. The LogLuv TIFF image format [2] was the first attempt at using this color representation for encoding images. Thereafter, more advanced compression schemes based on the MPEG standard also used the u’v’ color representation [3]–[6]. In [1], Poynton et al. proposed a modification of the CIE u’v’ to take into account the lower accuracy of the human perception of color in dark areas. In their modified version, below a luminance threshold, the chromaticity signal is attenuated towards gray proportionally to the luma. This method avoids encoding with too much precision the color noise that may appear in dark areas.

Inter-layer prediction in the context of HDR scalability amounts to inverting the tone mapping operator (TMO) that was used to generate the LDR image of the base layer. Methods such as [7], [8] automatically generate a LDR layer with a given tone mapping operator. The TMO being known, the inverse equations may be applied for the inter-layer prediction. However, in a more general context, the TMO is unknown. This is the case, for example, when the LDR version is generated by a manual color grading process. Several approaches exist to perform the prediction without prior knowledge on the TMO used. The authors in [3] automatically determine an inverse global tone curve that is encoded in the bitstream and applied to each individual LDR pixel to recover the HDR value. This method is very efficient in the case of global tone mapping, but inaccurate when a local TMO is used. Local approaches in [6], [9]–[11] tackle this problem by applying independent tone curves for each block of the image. These methods, either global or local, are particularly suitable for the prediction of an HDR luma channel from the LDR luma. However, less attention has been given to the inter-layer prediction of chromatic components. In [3], the authors observed that many TMOs have little impact on the CIE u’v’ color coordinates of the pixels in an image. The LDR base layer u’v’ components are then used for predicting

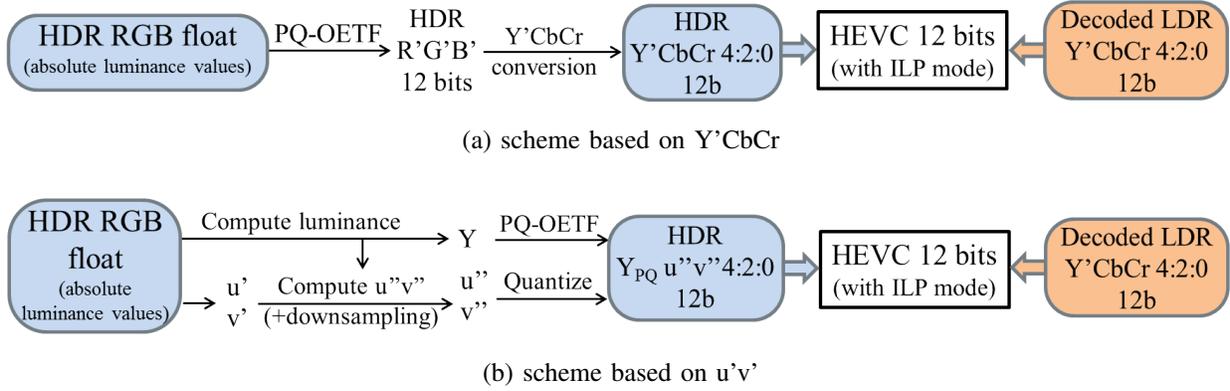


Fig. 1. Overview of the $Y'CbCr$ and $u'v'$ scalable compression schemes.

the color of the HDR enhancement layer. We show in this article under which circumstances this assumption is valid and how to generalize it to a broader range of tone mapping operators. For that purpose, we exploit general knowledge in the field of tone mapping and more precisely in the way the color information is handled. A very well-known method for generalizing any tone mapping operator to color images was developed by Schlick [12]. In [13], Tumblin and Turk then improved this method by adding a parameter for a better control of the saturation of colors in the tone mapped image. Later, several other popular TMOs in [14]–[16] used the same color correction method.

In this article, a model for predicting the color of the HDR image from the decoded LDR image and HDR luma channel is introduced. The model is derived from the color correction equations of Tumblin and Turk [13] which ensures the hue preservation between the colors of the HDR and LDR versions of the content. Since this color correction method requires a saturation parameter that might be unknown to the encoder, we developed a pre-analysis method that automatically determines the most suitable parameter value given the original HDR and LDR pair of images. This parameter is then transmitted as meta-data and used for performing predictions.

We developed two color inter-layer prediction methods using either the $Y'CbCr$ or the $u'v'$ representations for the HDR layer. We assessed these methods in a scalable coding set-up using HEVC to code the base LDR and HDR layers. For a fair comparison, we use the modified $u'v'$ coordinates proposed in [1] which are more perceptually uniform than the original $u'v'$ representation. In order to keep complete backward compatibility, the LDR layer is encoded in $Y'CbCr$ in both encoding schemes.

The remainder of the paper is organized as follows. The two encoding schemes are presented in detail in section II. Then, the color model based on the color correction of Tumblin and Turk is explained in section III. From this model, we derive in section IV the prediction equations of the chromatic components for both encoding schemes. The pre-analysis step which automatically determines the model's saturation parameter is also developed in subsection IV-D. Finally, our experimental results are presented in section V.

II. OVERVIEW OF THE SCALABLE HDR COMPRESSION SCHEME

This section presents the two considered compression schemes where the base and enhancement layers are respectively a LDR image and its HDR version. The original HDR image is calibrated and in the linear domain. The human perception of luminance being non-linear, an Opto-Electrical Transfer function (OETF) and a quantization to integers must be applied first to generate a perceptually uniform HDR signal suitable for compression. In this paper, we used the PQ-OETF function from [17], [18]. It can take input luminance values of up to 10000 cd/m^2 and outputs 12 bit integers. In their experiments, Boitard et al. [19] have reported that the PQ-OETF achieved better perceptual uniformity compared to several other EOTF functions for the encoding of HDR luminance values. Additionally, concerning the color encoding, they have also experimentally verified that the luminance and the $u'v'$ are less correlated than the luma and chroma components in the $Y'CbCr$ colorspace. The consequence is that less bits are required for a perceptually lossless encoding of colors using a $u'v'$ based colorspace compared with $Y'CbCr$. In order to confirm the potential of this representation for video compression, we have developed two inter-layer prediction methods based on either the $Y'CbCr$ or the CIE $u'v'$ representations and assessed them in a scalable set-up.

A. $Y'CbCr$ compression scheme

In the $Y'CbCr$ scheme, illustrated in Figure 1(a), the OETF is applied to the R , G , and B components independently and the resulting $R'G'B'$ components are converted to $Y'CbCr$ colorspace using the standard conversion matrix from the ITU-R BT-709 recommendations [20]. This is very similar to the colorspace generally used for the compression of LDR images, the only difference being that the usual gamma correction is replaced by the PQ-OETF which better models human perception, particularly for high luminance values.

Then, the chroma channels Cb and Cr are downsampled and the image is sent to a modified version of HEVC including our inter-layer prediction mode.

B. CIE $u'v'$ based compression scheme

In the second scheme, shown in Figure 1(b), the true luminance Y is computed and the PQ-OETF is applied only

to this achromatic component to form the luma channel Y_{PQ} . Then, the CIE $u'v'$ color coordinates are computed from the linear RGB values. The modification proposed in [1] is applied in our scheme. The modified $u''v''$ components are noted $u''v''$ and are computed with the following formula :

$$\begin{aligned} u'' &= (u' - u'_r) \cdot \frac{Y_{PQ}}{\max(Y_{PQ}, Y_{th})} + u'_r \\ v'' &= (v' - v'_r) \cdot \frac{Y_{PQ}}{\max(Y_{PQ}, Y_{th})} + v'_r \end{aligned} \quad (1)$$

where u'_r and v'_r are the $u'v'$ coordinates of the standard D65 illuminant [21] : $u'_r = 0,1978$ and $v'_r = 0.4683$. And Y_{th} is a threshold on the luma Y_{PQ} that we set to 1000 which corresponds to an absolute luminance value of 4.75 cd/m^2 . This modification allows a coarser quantization of the color in dark regions that may contain invisible color noise. In the decoding process, the $u'v'$ coordinates are retrieved by performing the inverse operations.

The two color channels are formed by quantizing the $u''v''$ pixel values. Poynton et al. [1] determined that quantizing those values to only 9 bits integers did not produce any visible artifact. However, they did not consider the HEVC based compression of both chromatic and achromatic signals. In practice, the quantization step of u'' and v'' should be chosen depending on the luma bitdepth in order to have a good bitrate allocation between luma and chromaticity. From our experiments, we have found that quantizing the chromaticity signal to 1 bit less than the luma bitdepth gave a reasonable tradeoff. Thus, 11 bits integers are used for the chromaticity. Knowing that the values of u'' and v'' are between 0 and 0.62, we apply a factor of 3302 to obtain quantized values u''_Q and v''_Q in the range $[0, 2^{11} - 1]$, as

$$u''_Q = [3302 \cdot u''], \quad v''_Q = [3302 \cdot v''], \quad (2)$$

where $[.]$ represents the rounding to the nearest integer.

Similarly to the $Y'CbCr$ scheme, the chromatic components u''_Q and v''_Q are downsampled. In order to keep compatibility with typical LDR encoding schemes, the LDR layer is encoded in the $Y'CbCr$ 4:2:0 format (i.e. $Y'CbCr$ with both horizontal and vertical chroma down-sampling).

C. Modified HEVC for Scalability

The proposed inter-layer prediction modes have been used in a scalable set-up where the base LDR and HDR layers are encoded using HEVC. The two versions of the scalable scheme using either the $Y'CbCr$ or the $u''v''$ color representation have been implemented. In the HDR enhancement layer, the encoder chooses between the existing intra and inter coding modes and the added inter-layer prediction mode for the chromatic components of the corresponding scheme. The mode decision is made at the Coding Unit (CU) level. As regards the inter-layer prediction of the luma channel, the ILP method presented in [9] is used for both schemes. This method locally determines inverse tone mapping curves on a per-block basis for predicting the HDR data from the decoded LDR version. As a result, our ILP method is not limited to the case of a LDR layer generated with a global TMO.

For both encoding schemes, the inter-layer prediction equations of the chromatic components have been derived by assuming that the base layer was generated with a TMO which applies the color correction of Tumblin and Turk [13]. More details on this color correction method are given in the next section and our prediction equations are presented in section IV.

III. TONE MAPPING COLOR CORRECTION

The color correction method used by Tumblin and Turk for generalizing any TMO to color images is illustrated in Figure 2. In this method, the TMO f , that can be either global or local, is first applied to the luminance Y . The tone mapped RGB components are then obtained based on a saturation parameter s , the tone mapped luminance $f(Y)$, and the ratio between the HDR RGB components and the original luminance Y . Since the tone mapping is performed on linear RGB values, a further gamma correction is required. The final gamma corrected LDR RGB components are then expressed by :

$$C_{LDR} = \left(\frac{C}{Y}\right)^{\frac{s}{\gamma}} \cdot f(Y)^{\frac{1}{\gamma}} \quad (3)$$

with $C = R, G, B$.

In our article, this color correction formula is considered as a model describing the relationship between the chromatic information in a HDR image and its corresponding LDR version. This choice is explained by the fact that Tumblin and Turk's color correction preserves the hues of the original HDR image in the tone mapped image. This property is very likely to be satisfied by most practical content, even when the LDR version was not generated explicitly with Equation 3 (e.g. manual color grading, etc.). Furthermore, concerning the color saturation, the parameter s gives some flexibility to the model since it can be adjusted to the content of the HDR and LDR pair of images to be encoded.

The next section describes how to derive inter-layer prediction equations from this model for the HDR chroma in both the $u''v''$ and the $Y'CbCr$ schemes. We also present an automatic procedure to determine the saturation parameter s for the content.

IV. COLOR INTER-LAYER PREDICTION

A. prediction of CIE $u'v'$ values

In the original definition of the CIE standard, the $u'v'$ color coordinates can be computed from the CIE XYZ values by :

$$\begin{aligned} u' &= \frac{4 \cdot X}{X + 15 \cdot Y + 3 \cdot Z} \\ v' &= \frac{9 \cdot Y}{X + 15 \cdot Y + 3 \cdot Z} \end{aligned} \quad (4)$$

Since the linear RGB components can be expressed as a linear combination of X, Y, and Z, we can write :

$$\begin{aligned} u' &= \frac{a_0 \cdot R + a_1 \cdot G + a_2 \cdot B}{b_0 \cdot R + b_1 \cdot G + b_2 \cdot B} \\ v' &= \frac{c_0 \cdot R + c_1 \cdot G + c_2 \cdot B}{b_0 \cdot R + b_1 \cdot G + b_2 \cdot B} \end{aligned} \quad (5)$$

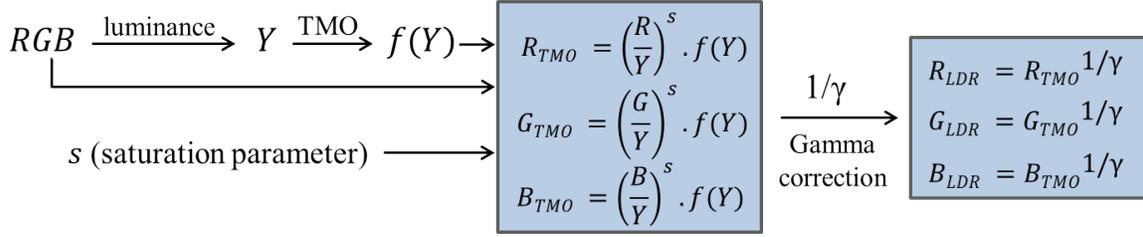


Fig. 2. Color correction formula of Tumblin and Turk [13].

where the coefficients a_0 to c_2 are fixed values depending on the chromaticities of the HDR RGB colorspace. In the case of BT-709 RGB [20], the values of the coefficients are :

$$\begin{array}{lll} a_0 = 1.650 & a_1 = 1.430 & a_2 = 0.722 \\ b_0 = 3.661 & b_1 = 11.442 & b_2 = 4.114 \\ c_0 = 1.914 & c_1 = 6.436 & c_2 = 0.650 \end{array}$$

From the model described in Equation 3 we can directly determine :

$$f(Y)^{\frac{1}{s}} = R_{LDR}^{\frac{\gamma}{s}} \cdot \frac{Y}{R} = G_{LDR}^{\frac{\gamma}{s}} \cdot \frac{Y}{G} = B_{LDR}^{\frac{\gamma}{s}} \cdot \frac{Y}{B} \quad (6)$$

Thus,

$$\begin{cases} R_{LDR}^{\frac{\gamma}{s}} = G_{LDR}^{\frac{\gamma}{s}} \cdot \frac{R}{G} \\ B_{LDR}^{\frac{\gamma}{s}} = G_{LDR}^{\frac{\gamma}{s}} \cdot \frac{B}{G} \end{cases} \quad (7)$$

Now, let us rewrite Equation 5 as :

$$u' = \frac{(a_0 \cdot \frac{R}{G} + a_1 + a_2 \cdot \frac{B}{G})}{(b_0 \cdot \frac{R}{G} + b_1 + b_2 \cdot \frac{B}{G})} \quad (8)$$

A similar equation can be found for v' . By multiplying both the numerator and the denominator by $G_{LDR}^{\frac{\gamma}{s}}$ in Equation 8 and by using Equation 7, we obtain a prediction value u'_{pred} for u' based only on the LDR RGB and the model parameters γ and s . The expression of v'_{pred} is obtained the same way :

$$\begin{aligned} u'_{pred} &= \frac{a_0 \cdot R_{LDR}^{\frac{\gamma}{s}} + a_1 \cdot G_{LDR}^{\frac{\gamma}{s}} + a_2 \cdot B_{LDR}^{\frac{\gamma}{s}}}{b_0 \cdot R_{LDR}^{\frac{\gamma}{s}} + b_1 \cdot G_{LDR}^{\frac{\gamma}{s}} + b_2 \cdot B_{LDR}^{\frac{\gamma}{s}}} \\ v'_{pred} &= \frac{c_0 \cdot R_{LDR}^{\frac{\gamma}{s}} + c_1 \cdot G_{LDR}^{\frac{\gamma}{s}} + c_2 \cdot B_{LDR}^{\frac{\gamma}{s}}}{b_0 \cdot R_{LDR}^{\frac{\gamma}{s}} + b_1 \cdot G_{LDR}^{\frac{\gamma}{s}} + b_2 \cdot B_{LDR}^{\frac{\gamma}{s}}} \end{aligned} \quad (9)$$

Hence, given the ratio between the parameters γ and s , and the decoded LDR data, we can directly predict the HDR u' and v' color components by applying the standard $u'v'$ conversion Equation 5 to the decoded LDR RGB values raised to the power $\frac{\gamma}{s}$.

This is a generalized version of Mantiuk et al's color predictions in [3] that considered $u' = u'_{LDR}$ and $v' = v'_{LDR}$ where u'_{LDR} and v'_{LDR} are computed from the linearized LDR RGB values (i.e. R_{LDR}^{γ} , G_{LDR}^{γ} and B_{LDR}^{γ}). Our prediction is equivalent in the particular case when $s = 1$. Note that in [6], Garbas and Thoma also predict the HDR layer $u'v'$ from the LDR layer $u'v'$ but the gamma correction is not taken into account in the computation of u'_{LDR} and v'_{LDR} . In this case, the prediction is thus equivalent to taking $\frac{\gamma}{s} = 1$, which is far from optimal in general since typical γ values are 2.2

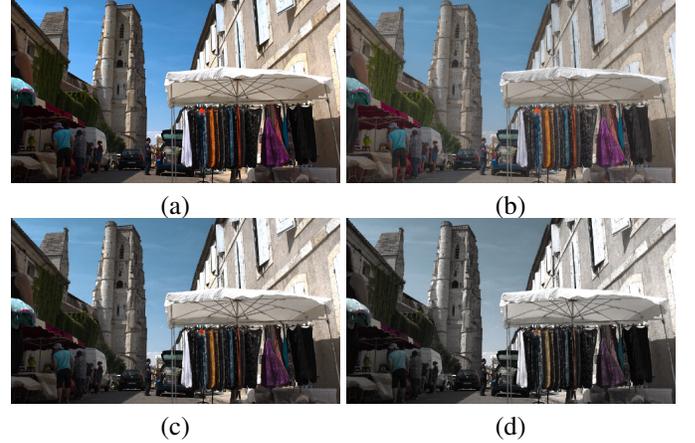


Fig. 3. $u'v'$ prediction results on a frame of the *Market3* sequence. (a) Original HDR image. (b) Tone mapped image with the TMO in [14] using $s = 0.6$ and $\gamma = 2.2$. (c) HDR color prediction from [3] (i.e. assuming $s = 1$ and $\gamma = 2.2$). (d) HDR color prediction from [6] (i.e. assuming $\frac{\gamma}{s} = 1$). For the sake of illustration, HDR images in (a), (c), and (d) are rendered with a simple gamma correction.

or 2.4 while s usually does not exceed 1. Figure 3 shows an example of color predictions produced by [3] and [6]. The base layer in Figure 3(b) was tone mapped from the original HDR image in 3(a) by the TMO [14]. This TMO explicitly uses the color correction in Equation 3 and the parameters $s = 0.6$ and $\gamma = 2.2$ were chosen. Garbas and Thoma's color predictions [6] result in too low saturation as shown in Figure 3(d). Better results are obtained in Figure 3(c) by Mantiuk et al's predictions [3] which take the gamma correction into account. However, the colors are still less saturated than in the original image because the parameter s used in the TMO was less than 1. In our method, the saturations of the original HDR image can be recovered by using the actual values of the parameters γ and s in Equation 9.

Since our compression scheme is based on the modified version $u''v''$ of the CIE $u'v'$ coordinates, the predictions u''_{pred} and v''_{pred} are formed with Equation 1 using u'_{pred} , v'_{pred} and the decoded HDR luma. Finally, u''_{pred} and v''_{pred} are multiplied by 3302 and rounded to the nearest integer, as in Equation 2, to predict the quantized values u''_Q and v''_Q .

B. prediction in $Y'CbCr$

In the case where the HDR layer is encoded in $Y'CbCr$ colorspace, a different prediction scheme is necessary. Unlike the $u'v'$ coordinates, the Cb and Cr chroma components cannot be predicted directly. First, we must predict the HDR RGB

values in the linear domain. Then, the PQ-OETF curve [17] must be applied to the predicted RGB values before computing the chroma prediction.

For the derivation of prediction equations of the RGB components, let us first define X_r and X_b as the ratios between color components :

$$X_r = \frac{R}{G}, \quad X_b = \frac{B}{G} \quad (10)$$

From the model given by Equation 3, we have :

$$\begin{aligned} R_{LDR} &= \left(\frac{R}{\bar{Y}}\right)^{\frac{s}{\gamma}} \cdot f(Y)^{\frac{1}{\gamma}} = X_r^{\frac{s}{\gamma}} \cdot \left(\frac{G}{\bar{Y}}\right)^{\frac{s}{\gamma}} \cdot f(Y)^{\frac{1}{\gamma}} \\ &= X_r^{\frac{s}{\gamma}} \cdot G_{LDR} \end{aligned} \quad (11)$$

The ratios X_r and X_b can thus be found using only the LDR RGB components :

$$X_r = \left(\frac{R_{LDR}}{G_{LDR}}\right)^{\frac{\gamma}{s}}, \quad X_b = \left(\frac{B_{LDR}}{G_{LDR}}\right)^{\frac{\gamma}{s}} \quad (12)$$

Using the ratios X_r and X_b the luminance component can be expressed as :

$$\begin{aligned} Y &= \alpha_0 \cdot R + \alpha_1 \cdot G + \alpha_2 \cdot B \\ &= (\alpha_0 \cdot X_r + \alpha_1 + \alpha_2 \cdot X_b) \cdot G \end{aligned} \quad (13)$$

Thus,

$$\begin{aligned} G &= \frac{Y}{\alpha_0 \cdot X_r + \alpha_1 + \alpha_2 \cdot X_b} \\ R &= X_r \cdot G \\ B &= X_b \cdot G \end{aligned} \quad (14)$$

where the coefficients α_0 , α_1 , and α_2 depend on the RGB colorspace used. For the BT-709 colorspace, $\alpha_0 = 0.2126$, $\alpha_1 = 0.7152$, and $\alpha_2 = 0.0722$.

However, the true luminance Y is not known in the Y'CbCr scheme. Only an approximation \tilde{Y} is obtained when the inverse PQ-OETF curve, which we denote PQ^{-1} , is applied to the luma channel Y' . The predicted RGB values can then be obtained by applying Equation 14 and by replacing Y by $\tilde{Y} = PQ^{-1}(Y')$. This can be inaccurate particularly in very saturated regions where one of the components is close to zero. It has been experimentally observed that better results are obtained by approximating the PQ-OETF function by a power function in the expression of \tilde{Y} :

$$\begin{aligned} \tilde{Y} &\approx \left(\alpha_0 \cdot R^{\frac{1}{p}} + \alpha_1 \cdot G^{\frac{1}{p}} + \alpha_2 \cdot B^{\frac{1}{p}}\right)^p \\ &\approx \left(\alpha_0 \cdot X_r^{\frac{1}{p}} + \alpha_1 + \alpha_2 \cdot X_b^{\frac{1}{p}}\right)^p \cdot G \end{aligned} \quad (15)$$

Finally the approximation for the green component G is given by :

$$G \approx \frac{\tilde{Y}}{\left(\alpha_0 \cdot X_r^{\frac{1}{p}} + \alpha_1 + \alpha_2 \cdot X_b^{\frac{1}{p}}\right)^p} \quad (16)$$

Note that for $p = 1$, this is equivalent to the previous approximation (i.e. $\tilde{Y} \approx Y$). Examples of prediction results are shown in Figure 4 with varying values of p . In our

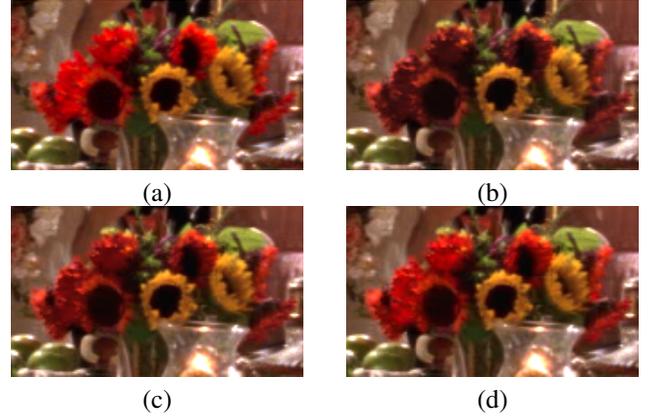


Fig. 4. Y'CbCr prediction results on a detail of a frame in the *StEM* sequence. (a) Original HDR image. (b), (c), and (d) : prediction images with respectively $p = 1$, $p = 2$, and $p = 4$. For the sake of illustration, HDR images are rendered with a simple gamma correction.

experiments, we have found that using $p = 4$ gives high-quality results in most situations.

In order to improve the predictions in dark areas, we used in our implementation a slightly modified version of the ratios X_r and X_b :

$$\tilde{X}_r = \left(\frac{R_{LDR} + \epsilon}{G_{LDR} + \epsilon}\right)^{\frac{\gamma}{s}}, \quad \tilde{X}_b = \left(\frac{B_{LDR} + \epsilon}{G_{LDR} + \epsilon}\right)^{\frac{\gamma}{s}} \quad (17)$$

where ϵ is a small value fixed to 1% of the maximum LDR value (i.e. $\epsilon = 2.55$ for a 8 bit LDR layer). Compared to the theoretical result in Equation 12, this prediction of X_r and X_b reduces the noise in dark regions where the ratios $\frac{R_{LDR}}{G_{LDR}}$ and $\frac{B_{LDR}}{G_{LDR}}$ may be too sensitive to small color errors caused by lossy compression of the LDR layer. Equation 17 also avoids singularities. The actual HDR RGB prediction is then computed from the decoded LDR RGB components and the decoded HDR luma Y' using the following equation :

$$\begin{aligned} G_{pred} &= \frac{PQ^{-1}(Y')}{\left(\alpha_0 \cdot \tilde{X}_r^{\frac{1}{p}} + \alpha_1 + \alpha_2 \cdot \tilde{X}_b^{\frac{1}{p}}\right)^p} \\ R_{pred} &= \tilde{X}_r \cdot G_{pred} \\ B_{pred} &= \tilde{X}_b \cdot G_{pred} \end{aligned} \quad (18)$$

The Cb and Cr components are finally predicted by applying back the PQ-OETF to R_{pred} , G_{pred} , and B_{pred} and by computing the chroma components.

C. Implementation details

In both the Y'CbCr and the u''v'' encoding schemes, the prediction of the chromatic components is based on the decoded LDR RGB components and the HDR luma. In our implementation, for a given block in the image, the luma block is always encoded and decoded before the chromatic components. As a result, the decoded luma block is known while encoding or decoding the u'' and v'' blocks. However, since the color components are downsampled horizontally and vertically, the same down-sampling must be performed to

the decoded luma channel. We used a simple down-sampling scheme consisting in taking the mean of the four luma pixels collocated with a given chroma pixel. In the Y'CbCr encoding scheme, the inverse PQ-OETF is applied after the luma down-sampling for the computation of \tilde{Y} .

Similarly, the LDR RGB components must be given in low resolution for performing the prediction. Since the LDR layer is originally encoded in the Y'CbCr 4:2:0 format, only the LDR luma needs to be downsampled. The low resolution LDR luma and chroma are then converted to RGB.

D. Pre-Analysis

In general, we cannot assume that the parameters s and γ used in the prediction model are known in advance. A first step thus consists in determining the parameters that best fit the HDR and LDR image pair. This can be done in a pre-processing stage before encoding. Therefore, these parameters can be estimated using the original LDR and HDR images without compression. From the prediction equations in section IV, we note that only the ratio $s' = \frac{s}{\gamma}$ must be determined.

From the color model in Equation 3, we directly obtain :

$$f(Y)^{\frac{1}{\gamma}} = R_{LDR} \cdot \left(\frac{Y}{R}\right)^{s'} = G_{LDR} \cdot \left(\frac{Y}{G}\right)^{s'} = B_{LDR} \cdot \left(\frac{Y}{B}\right)^{s'} \quad (19)$$

Thus, we want to find the value of s' that minimizes the mean square error (MSE) on all the pixels. The MSE was chosen here in order to keep the problem convex and fast to solve. For simplicity, only the red and green components are used in our minimization problem. For natural content, no difference has been observed when the blue component was taken into account. Given a pixel i , let us define the function F^i as

$$F^i(s') = \left(R_{LDR}^i \cdot \left(\frac{Y^i}{R^i}\right)^{s'} - G_{LDR}^i \cdot \left(\frac{Y^i}{G^i}\right)^{s'} \right)^2, \quad (20)$$

where R_{LDR}^i , G_{LDR}^i , R^i , and G^i are respectively the values of R_{LDR} , G_{LDR} , R , and G at pixel position i .

The estimation of the parameter s' is then expressed as

$$\hat{s}' = \underset{s'}{\operatorname{argmin}} \sum_{i=1}^n F^i(s'), \quad (21)$$

where n is the number of pixels. The problem in Equation 21 can be solved by finding the value of s' for which $\sum_{i=1}^n F^{i'}(s') = 0$, where $F^{i'}$ denotes the first derivative of F^i . Newton's iterative numerical method was used for that purpose. Given an initialization value $s'_0 = 0.4$, the value s'_k at iteration k is given by :

$$s'_k = s'_{k-1} - \frac{\sum_{i=1}^n F^{i'}(s'_{k-1})}{\sum_{i=1}^n F^{i''}(s'_{k-1})} \quad (22)$$

where the two first derivatives $F^{i'}$ and $F^{i''}$ can be determined analytically as :

$$F^{i'}(s') = A_{11}^i \cdot \left(\frac{Y^i}{R^i}\right)^{2s'} + A_{12}^i \cdot \left(\frac{Y^i}{G^i}\right)^{2s'} + A_{13}^i \cdot \left(\frac{(Y^i)^2}{R^i \cdot G^i}\right)^{s'} \quad (23)$$

$$F^{i''}(s') = A_{21}^i \cdot \left(\frac{Y^i}{R^i}\right)^{2s'} + A_{22}^i \cdot \left(\frac{Y^i}{G^i}\right)^{2s'} + A_{23}^i \cdot \left(\frac{(Y^i)^2}{R^i \cdot G^i}\right)^{s'} \quad (24)$$

with

$$\begin{aligned} A_{11}^i &= 2 \cdot \ln\left(\frac{Y^i}{R^i}\right) \cdot (R_{LDR}^i)^2 \\ A_{12}^i &= 2 \cdot \ln\left(\frac{Y^i}{G^i}\right) \cdot (G_{LDR}^i)^2 \\ A_{13}^i &= -2 \cdot \ln\left(\frac{(Y^i)^2}{R^i \cdot G^i}\right) \cdot R_{LDR}^i \cdot G_{LDR}^i \\ A_{21}^i &= A_{11}^i \cdot 2 \cdot \ln\left(\frac{Y^i}{R^i}\right) \\ A_{22}^i &= A_{12}^i \cdot 2 \cdot \ln\left(\frac{Y^i}{G^i}\right) \\ A_{23}^i &= A_{13}^i \cdot \ln\left(\frac{(Y^i)^2}{R^i \cdot G^i}\right) \end{aligned}$$

The iterative process in Equation 22 is stopped when the difference between the value of s' at two successive iterations is less than 10^{-4} . In our experiments, we observed a fast convergence and three iterations are usually sufficient to reach the precision of 10^{-4} .

In order to increase the robustness of the method, some pixels are removed from the sums in Equation 22. First, the pixels for which at least one of the HDR RGB components is less than a threshold of 0.02 are removed. Those pixels are not reliable because of the color noise that may appear in very dark regions. Moreover, too small R^i , G^i or B^i values can cause inaccurate computations of $F^{i'}$ and $F^{i''}$ in Equations 23 and 24. A second type of problem may appear for too bright pixels. In practice, after tone mapping, some pixel RGB values may exceed the maximum LDR value for one or several RGB components. A simple clipping operation is generally applied in this case in order to keep all the values in the LDR range. However, since this operation is performed on the RGB channels independently, it modifies the hue of the clipped pixels. Therefore, the assumption of hue preservation in the model is no longer satisfied. For that reason, we exclude from the computation all the pixels that exceed 99% of the maximum LDR value in at least one of the components R_{LDR} , G_{LDR} or B_{LDR} .

V. EXPERIMENTAL RESULTS

For our experiment, we have used ten HDR test sequences presented in table I. Their spatial resolution is 1920x1080 pixels. The sequences StEM WarmNight, Market3 and Tibul2, are parts of the MPEG standard sequences for HDR scalability [24]. Note that StEM WarmNight is originally one sequence

Sequence	Frames	Frame Rate	Intra period	Tone mapping	color correction	$s' (= s/\gamma)$
<i>StEM WarmNight 1</i>	3527-3551	24	24	manual (MPEG)	-	0.4099
<i>StEM WarmNight 2</i>	3729-3753	24	24	manual (MPEG)	-	0.4013
<i>Market3</i>	0-48	50	48	manual (MPEG)	-	0.3818
<i>Tibul2</i>	0-32	30	32	manual (MPEG)	-	0.3217
<i>Balloon</i>	0-24	25	24	Mantiuk et al. [14]	Tumblin and Turk [13]	0.3636
<i>Seine</i>	0-24	25	24	Mantiuk et al. [14]	Tumblin and Turk [13]	0.3636
<i>Fishing Longshot</i>	1000-1024	25	24	Fattal et al. [15]	Tumblin and Turk [13]	0.7727
<i>Bistro 1</i>	295-319	25	24	Fattal et al. [15]	Pouli et al. [22]	0.4315
<i>Carousel Fireworks 4</i>	115-139	25	24	Photographic TMO [23]	-	0.4545
<i>Showgirl 2</i>	348-372	25	24	Photographic TMO [23]	Tumblin and Turk [13]	0.3000

TABLE I
DETAIL OF THE HDR SEQUENCES AND TONE MAPPING OPERATORS USED FOR OUR EXPERIMENTS. IN THE LAST COLUMN, s' IS THE VALUE DETERMINED IN OUR PRE-ANALYSIS STEP.



Fig. 5. First frames (LDR versions) of each sequence used in the experiment.

containing two shots. In our experiments it was separated into two sequences. The sequences Balloon and Seine were produced by Binocle and Technicolor within the framework of the french collaborative project NEVEx. Finally, the sequences Fishing Longshot, Bistro 1, Carousel Fireworks 4, and Showgirl 2 are presented in [25] and are available for download. For the experiment, only the first second of each sequence was considered in order to keep reasonable computation times while showing results for a wide variety of content. Note that although the frame numbers in table I do not always start at zero, it actually corresponds to the beginning of the sequence.

Figure 5 shows the LDR versions of the first frame of each sequence used as a base layer. For the sake of simplicity, the RGB colorspace of both the LDR and HDR versions are defined with the standard BT.709 color primaries.

For the base layer of the sequences Market3, Tibul2, and StEM WarmNight, LDR versions produced by a manual color grading process were already provided in the MPEG set of sequences. Therefore, we did not apply further color correction not to interfere with the artistic intent of the producer. For the other sequences, the tone mapping operators, and for some of them, the color correction methods used for generating the

LDR layer are detailed in table I. In particular, the global version of the Photographic TMO [23] was used for the sequences Carousel fireworks 4 and Showgirl 2, while for the sequences Balloon, Seine, Fishing Longshot and Bistro 1, the local TMOs of Mantiuk et al. [14] and Fattal et al. [15] have been applied for generating the LDR images using the publicly available implementation of the pfstmo library [26]. These local TMOs were both designed to be applied on the luminance channel. They subsequently derive the LDR color components using Tumblin and Turk’s formula. For the sequence Bistro 1, we observed that more natural colors were obtained by further processing the tone mapped image using the color correction of Pouli et al. [22]. The latter method gives similar results as those obtained with the correction of Tumblin and Turk concerning the hue, but the authors have shown by visual experiments that their method better preserves the saturation of the HDR image in the tone mapped image.

All the tone mapped images were further gamma corrected with a typical 2.2 gamma value. It should be noted that for the sequences tone mapped using Tumblin and Turk’s color correction, the saturation parameter s was adjusted manually. Therefore, the value of the ratio $s' = \frac{s}{\gamma}$, which is required

by the encoder, is known in advance. In these cases, the pre-analysis step defined in subsection IV-D was able to recover the value of s' with the required precision of 10^{-4} . For all the sequences, the s' values determined in the pre-analysis step are listed in table I.

In our experiments, we have compared the $Y'CbCr$ and the $u''v''$ schemes. A first remark can be made concerning the down-sampling of the chromatic components introduced by the YUV 4:2:0 format conversion prior to the HEVC encoding. An example of down-sampling in each colorspace is shown in Figure 6. It can be seen in Figure 6(b) that the chroma down-sampling in $Y'CbCr$ may cause disturbing artifacts in areas containing saturated colors. This is due to the highly non-linear OETF applied independently to the RGB components before the conversion to $Y'CbCr$. Because of this non-linearity, a part of the luminance information is contained in the Cb and Cr component, and conversely, the luma channel Y' is influenced by the chromaticity. As a result, the chroma down-sampling causes errors in the luminance of the reconstructed image which are visually more significant than errors in colors. This problem does not occur in the $u''v''$ based colorspace since the luminance and the color components are decorrelated.

Additionally, we assessed our proposed algorithms comparatively to :

- Simulcast encoding (i.e. independent encoding of the LDR and HDR layers) for both $Y'CbCr$ and $u''v''$ color representations.
- Template-based local ILP method presented in [9] for both the luma and chroma channels.
- $u''v''$ components prediction used in Mantiuk et al's [3].

For a fair comparison, the luma channel is predicted with the template-based local ILP [9] in all the inter-layer prediction methods compared. Note that our implementation of Mantiuk et al's $u''v''$ prediction method is very close to our $u''v''$ scheme. The main difference is that the value of s in Equation 9 was fixed to 1 in order to obtain $u'_{pred} = u'_{LDR}$ and $v'_{pred} = v'_{LDR}$. The value of γ was set to 2.2 which corresponds to typical gamma correction. Furthermore, [3] directly uses the CIE $u''v''$ as color components. We have thus disabled our modification

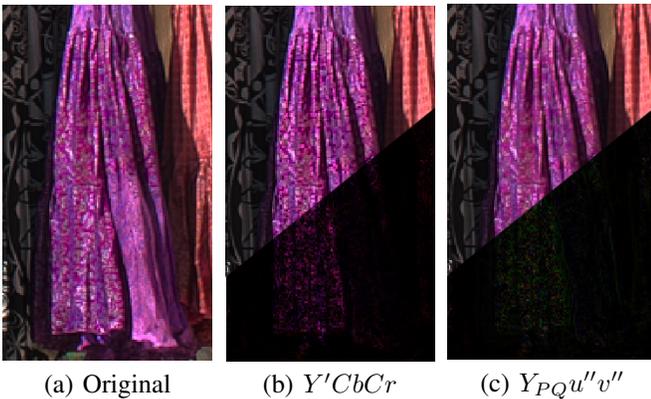


Fig. 6. Detail of a frame in the sequence Market3. (a) Original HDR image. (b) and (c) Images obtained by a down-sampling of the chromatic components using respectively $Y'CbCr$ and $Y_{PQ}u''v''$ colorspace. The bottom part shows the absolute error. For the sake of illustration, HDR images are rendered with a simple gamma correction.

of the $u''v''$ coordinates by setting to 0 the threshold value Y_{th} defined in Equation 1.

For the simulations, the encoding with our modified version of HEVC was performed with random access configuration using groups of pictures (GOPs) of 8 pictures. The period of intra frames for each sequence is given in table 1. It was chosen depending on the frame rate to correspond to approximately 1 second for each sequence.

A. Quality assessment

For assessing the quality of the decoded HDR images, we have chosen to use separate indices for the quality of the luminance signal which is achromatic and that of the chromaticity signal. The reason of this choice is that most of the existing quality metrics do not accurately account for color vision.

For instance, a common method for assessing the quality of compressed images consists in computing the peak signal-to-noise ratio (PSNR) of each of the $Y'CbCr$ components, and combining the results by a weighted sum. Alternatively, the PSNR can be computed from the perceptually quantized $R'G'B'$ components. However, the colorspace formed by the $R'G'B'$ or by the $Y'CbCr$ components only give a rough approximation of perceptual uniformity. It is particularly inaccurate in highly saturated colors, especially in the case of HDR images. Although a PSNR could be computed based on the CIE ΔE_{2000} color difference formula [27] which estimates well the perceived difference between two colors, this formula is only accurate with LDR data for which it was designed. Furthermore, new metrics have been developed specifically for HDR quality assessment, the most well-known being the HDR Visual Difference Predictor (HDR-VDP) [28]. However, they only predict luminance differences and do not consider color. Note that other metrics, TMQI [29] and TMQI-II [30], have been proposed recently for assessing the quality of the tone mapping step.

In our experiment, the quality of the HDR luminance component was assessed using the Q index of the HDR-VDP 2.2, giving a score between 0 and 100, where 100 is reached when there is no visible difference with the original luminance. This quality index is referred to as HDR-VDP(Y) in the rest of the article. The quality of the chromatic signal was assessed based on the CIE 1976 $L^*a^*b^*$ colorspace. A PSNR value is computed with Equation 25 using only the chromatic components a^* and b^* . Note that we could have used alternatively the CIE 1976 $L^*u^*v^*$ colorspace which is roughly equivalent to the CIE $L^*a^*b^*$ in terms of perceptual uniformity. However, the use of CIE $L^*a^*b^*$ is prevalent compared to the CIE $L^*u^*v^*$ in the color imaging community. Note also that the L^* component is a non-linear function of the luminance. This non-linearity was determined with the aim of perceptual uniformity by experiments based on stimuli of relatively low luminance. It is therefore only perceptually uniform for LDR data and it can be very inaccurate for modeling human perception with HDR images. For this reason, the L^* component was excluded from our index in Equation 25, and only the chromatic information contained in the a^* and b^* components was taken into account.

For the same reason, we did not use the CIE ΔE_{2000} formula which includes the differences in L^* in its expression.

$$PSNR_{a^*b^*} = 10 \cdot \log_{10} \left(\frac{1000^2}{(MSE_{a^*b^*})} \right) \quad (25)$$

where $MSE_{a^*b^*}$ is the mean square error for the a^* and b^* components (i.e. mean of the squared euclidean distance in the a^*b^* plane).

B. Rate-Distortion results

For each sequence and tested method, two Rate-Distortion curves have been determined experimentally, using either the distortion in luminance (i.e. achromatic), or the chromatic distortion index defined in Equation 25. The curves have been generated by encoding each sequence with different QP parameter values of 22, 27, 32, 37. Both the LDR and HDR layers have been encoded with the same QP value so that both layers are of comparable quality. The resulting RD curves are shown in Figure 8 for the sequences StEM WarmNight 1, Market3, Tibul2, Fishing Longshot, Bistro 1 and Showgirl 2. We have selected those 6 scenes presenting different characteristics to show the behavior of the different coding methods in various conditions. We can first note from the curves that the best compression performance, considering both the chromatic and the achromatic quality indices, is obtained with our $u''v''$ compression scheme for all the sequences.

In order to quantify our gains in comparison to other methods, we have computed the Bjontegaard Delta Rate metric [31] from the luminance distortion index and the total bitrate of all the components of both the HDR and LDR layers. Since our study focuses on the coding of the chromatic components, evaluating the rate gains only from a luminance based quality index is not enough. Therefore, we also computed the Bjontegaard Delta PSNR using the PSNR a^*b^* and the total bitrate.

The gains of our $u''v''$ scheme were computed with respect to Mantiuk et al's $u'v'$ predictions and are reported in table II. Note that fairly low rate gains are observed for most sequences since we considered only the quality of the luminance for the computation of the Delta Rate metric. This is explained by the fact that the luminance was encoded the same way

in our implementation of both methods in order to compare only the inter-layer prediction of the chromatic components. However, in the case of the sequences StEM WarmNight 1 and 2, and Showgirl 2, our method shows significant rate gains of respectively 19.7%, 13.4%, and 41.4% in comparison to Mantiuk et al's version. This is due to the color noise contained in the dark regions of those sequences which is better quantized by using the modified $u''v''$ than with the original CIE $u'v'$ color components, resulting in lower overall bitrate. Additionally, our method is more reliable for the coding of colors because of the parameter s' optimized for the image's content. In particular, for the sequences Balloon and Fishing Longshot, a gain of respectively 0.85 dB and 1.06 dB in $PSNR_{a^*b^*}$ is observed between our color ILP and that of Mantiuk et al. An example of compression and prediction results for the sequence Fishing Longshot is shown in Figure 7. Mantiuk et al's color prediction in Figure 7(b) results in too strong color saturation. The encoding of the residual in HEVC partly corrects the prediction error in the decoded image of Figure 7(c), but color artifacts remain visible. In our method, the automatic determination of the parameter s' ensures the accuracy of the color saturation in the prediction.

Similarly, we have compared our $Y'CbCr$ scheme with the method in [9] which also uses a $Y'CbCr$ encoding, but where the Cb and Cr components are predicted with the same local ILP method than the luma component Y' . The Bjontegaard gains are presented in table III. It can be seen that on average, there is little difference in terms of Rate-Distortion results when evaluating on the luminance difference. However, substantial gains are observed for our method when evaluating the chromatic distortions. In particular, for the sequences StEM WarmNight 1, StEM WarmNight 2, and Market3, whose LDR versions were produced with a manual color grading process, gains of 0.55 dB, 0.74 dB, and 0.57 dB respectively are observed in PSNR a^*b^* . It shows that the color model used in our scheme estimates well the relationship between the colors of the LDR and the HDR images even though the LDR versions were not explicitly generated with the color correction of Tumblin and Turk. The exception, however, is the sequence Tibul2 which has an overall very saturated red color. In this case, a significant loss is observed both in luminance and

Method Tested :	Ours ($u''v''$)	
Metric :	Δ Rate	Δ PSNR a^*b^*
StEM WarmNight 1	-19.7 %	0.23 dB
StEM WarmNight 2	-13.4 %	0.21 dB
Market3	-1.6 %	0.23 dB
Tibul2	-4.0 %	0.40 dB
Balloon	-1.8 %	0.85 dB
Seine	-4.0 %	0.04 dB
Fishing Longshot	-6.8 %	1.06 dB
Bistro 1	-8.9 %	0.12 dB
Carousel Fireworks 4	-3.8 %	0.16 dB
Showgirl 2	-41.4 %	0.75 dB
Average	-10.5 %	0.41 dB

TABLE II

BJONTEGAARD GAINS OF OUR $u''v''$ SCHEME WITH RESPECT TO MANTIUK ET AL. COLOR PREDICTION [3]. THE RATE GAINS ARE COMPUTED FROM THE HDR-VDP(Y) QUALITY INDEX. THE TOTAL BITRATE OF ALL THE COMPONENTS OF BOTH LAYERS IS CONSIDERED.

Method Tested :	Ours ($Y'CbCr$)	
Metric :	Δ Rate	Δ PSNR a^*b^*
StEM WarmNight 1	-0.7 %	0.55 dB
StEM WarmNight 2	5.6 %	0.74 dB
Market3	-1.2 %	0.57 dB
Tibul2	11.4 %	-0.64 dB
Balloon	7.6 %	0.19 dB
Seine	-0.3 %	0.36 dB
Fishing Longshot	-2.2 %	0.06 dB
Bistro 1	-1.1 %	0.26 dB
Carousel Fireworks 4	-2.5 %	0.69 dB
Showgirl 2	-30.0 %	0.26 dB
Average	-1.34 %	0.25 dB

TABLE III

BJONTEGAARD GAINS OF OUR $Y'CbCr$ SCHEME WITH RESPECT TO THE LOCAL ILP IN [9]. THE RATE GAINS ARE COMPUTED FROM THE HDR-VDP(Y) QUALITY INDEX. THE TOTAL BITRATE OF ALL THE COMPONENTS OF BOTH LAYERS IS CONSIDERED.

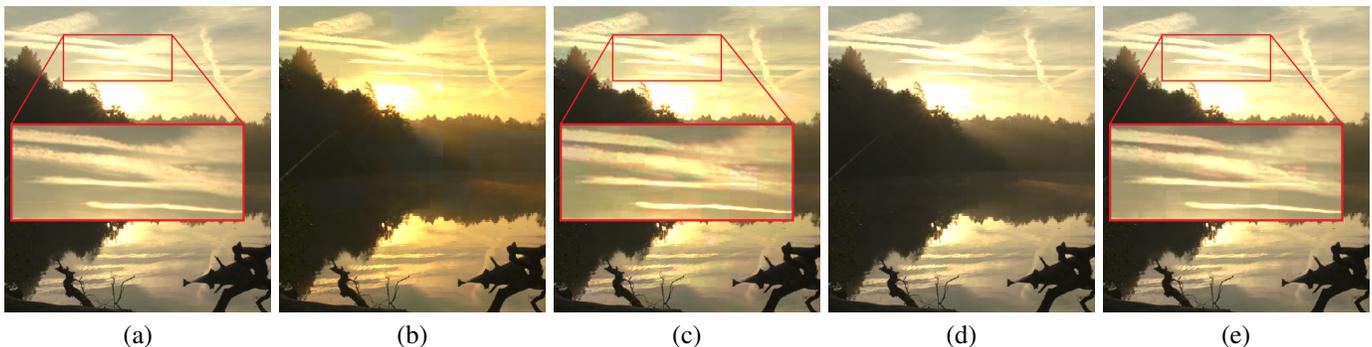


Fig. 7. Part of a frame in the sequence Fishing Longshot. (a) Original HDR image. (b) and (c) : respectively predicted and decoded image using Mantiuk et al's method [3] for the chromatic components (i.e. assuming $s' = \frac{s}{\gamma} = \frac{1}{2.2} \approx 0.45$). (d) and (e) : respectively predicted and decoded image with our method (i.e. with $s' = 0.7727$ determined in pre-analysis). The images are encoded with QP=27 for the LDR layer and QP=37 for the HDR layer. The HDR layer bitrate is 0.0145 bits per pixel in (c) and 0.0054 bits per pixel in (f). For the sake of illustration, the images are rendered with a simple gamma correction.

Method Tested :	Ours (Y'CbCr)		Ours (u"v")	
	Metric :	Δ Rate	Δ PSNR a*b*	Δ Rate
<i>STEM WarmNight 1</i>	-57.0 %	1.15 dB	-57.3 %	1.17 dB
<i>STEM WarmNight 2</i>	-46.2 %	1.14 dB	-56.2 %	1.21 dB
<i>Market3</i>	-48.1 %	1.97 dB	-49.4 %	2.22 dB
<i>Tibul2</i>	-30.4 %	0.66 dB	-48.8 %	1.84 dB
<i>Balloon</i>	-47.4 %	1.09 dB	-49.0 %	1.09 dB
<i>Seine</i>	-56.4 %	1.22 dB	-56.4 %	1.33 dB
<i>Fishing Longshot</i>	-50.1 %	1.06 dB	-51.2 %	1.07 dB
<i>Bistro 1</i>	-59.5 %	0.94 dB	-61.7 %	1.10 dB
<i>Carousel Fireworks 4</i>	-46.3 %	1.46 dB	-56.8 %	2.28 dB
<i>Showgirl 2</i>	-76.7 %	0.55 dB	-90.2 %	0.72 dB
Average	-51.8 %	1.12 dB	-57.7 %	1.40 dB

TABLE IV

BJONTEGAARD GAINS WITH RESPECT TO Y'CbCr SIMULCAST. THE RATE GAINS ARE COMPUTED FROM THE HDR-VDP(Y) QUALITY INDEX. THE TOTAL BITRATE OF ALL THE COMPONENTS OF BOTH LAYERS IS CONSIDERED.

in chrominance. This can be explained by the approximation made in Equation 15 in order to derive the prediction Equation 18. This approximation may cause artifacts in highly saturated colors as shown in the example of Figure 4. Despite the parameter p , introduced in order to reduce those artifacts, our method remains less efficient than the local ILP in [9] when the whole sequence contains very saturated colors, as in Tibul2.

Finally, table IV shows the Bjontegaard gains with respect to the Simulcast method with Y'CbCr encoding. On average, 51.8% of the bitrate is saved at equal luminance quality by using our Y'CbCr inter-layer prediction scheme. Better compression performance is obtained with our u"v" scheme which reaches an average 57.7% gain. Higher gains are also observed for the u"v" version by considering the chromatic quality index, PSNR a*b*.

VI. CONCLUSION

In the context of the scalable compression of HDR content with a LDR base layer, we have developed a new inter-layer prediction method specifically for the chromatic components. Our method is based on a model linking the colors in the HDR layer to those in the LDR layer. In particular, it follows the general assumption that the hues of the colors in an HDR image were preserved in the LDR version. In addition, the model uses a single parameter to adjust the saturations of the HDR colors in the prediction. A method is described to determine the optimal value for this parameter given an HDR image and its associated LDR version. From the model,

we derived prediction equations for two encoding schemes using different color representations of the images. In the first scheme, the classical Y'CbCr encoding is addressed while the second version uses a colorspace built from the luminance and the CIE u"v' color coordinates.

Our results show the advantages of the CIE u"v' based colorspace, which completely decorrelates the luminance and chrominance signals. This property enables a better down-sampling of the chromatic components than the usual chroma down-sampling in a Y'CbCr colorspace. Moreover, the u"v' components can be predicted more accurately from the color model than the CbCr components. We have also demonstrated that, thanks to the saturation parameter in the model, our u"v' inter-layer prediction generalizes the existing color ILP methods in the literature that uses the same u"v' representation. The experiments have confirmed that the use of the optimized saturation parameter improved the coding performance.

Regarding the coding in Y'CbCr colorspace, our ILP scheme based on the color model also shows better coding performances in most cases in comparison to other methods which directly predict the HDR layer's chroma components from those of the LDR layer.

REFERENCES

- [1] C. Poynton, J. Stessen, and R. Nijland, "Deploying wide color gamut and high dynamic range in HD and UHD," *SMPTE Mot. Imag. J.*, vol. 124, no. 3, pp. 37–49, Apr. 2015.
- [2] G. W. Larson, "Logluv encoding for full-gamut, high-dynamic range images," *J. Graph. Tools*, vol. 3, no. 1, pp. 15–31, Mar. 1998.

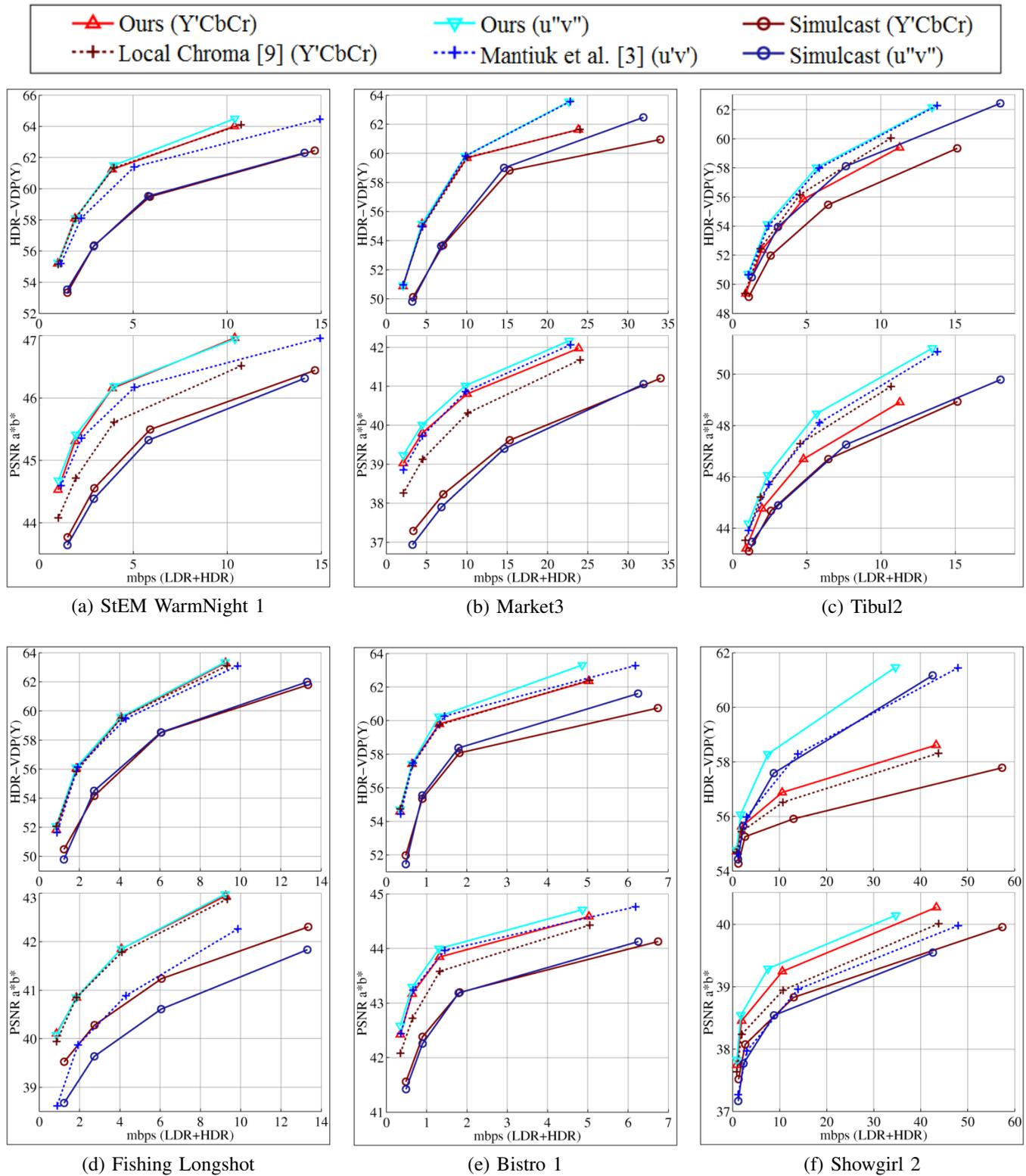


Fig. 8. Rate-Distortion curves. For each sequence, the luminance distortion is represented in the upper graph while the chromatic distortion is shown in the lower graph. The x-axis represents the total bitrate for all the components of both the HDR and LDR layers.

- [3] R. Mantiuk, A. Efremov, K. Myszkowski, and H.-P. Seidel, "Backward compatible high dynamic range MPEG video compression," *ACM Trans. Graph.*, vol. 25, no. 3, Jul. 2006.
- [4] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Lossy compression of high dynamic range images and video," *Human Vision and Electronic Imaging XI, SPIE*, vol. 6057, Feb. 2006.
- [5] A. Motra and H. Thoma, "An adaptive logluc transform for high dynamic range video compression," *17th IEEE International Conference on Image Processing (ICIP)*, pp. 2061–2064, Sep. 2010.
- [6] J.-U. Garbas and H. Thoma, "Inter-layer prediction for backwards compatible high dynamic range video coding with SVC," *Picture Coding Symposium (PCS)*, pp. 285–288, May 2012.
- [7] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. K. Ward, and W. Heidrich, "Optimizing a tone curve for backward-compatible high dynamic range image and video compression," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1558–1571, 2011.
- [8] Z. Mai, H. Mansour, P. Nasiopoulos, and R. K. Ward, "Visually favorable tone-mapping with high compression performance in bit-depth scalable video coding," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1503–1518, 2013.
- [9] M. Le Pendu, C. Guillemot, and D. Thoreau, "Local inverse tone curve learning for high dynamic range image scalable compression," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5753–5763, Dec. 2015.
- [10] S. Liu, W.-S. Kim, and A. Vetro, "Bit-depth scalable coding for high dynamic range video," *SPIE Conference on Visual Communications and Image Processing*, Jan. 2008.
- [11] A. Segall, "Scalable coding of high dynamic range video," *14th IEEE International Conference on Image Processing (ICIP)*, Oct. 2007.
- [12] C. Schlick, "Quantization techniques for visualization of high dynamic range pictures," *5th Eurographics Workshop on Rendering*, pp. 7–20, 1994.
- [13] J. Tumblin and G. Turk, "Lcis: A boundary hierarchy for detail-preserving contrast reduction," pp. 83–90, 1999.
- [14] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," *ACM Trans. Appl. Percept.*, vol. 3, no. 3, pp. 286–308, Jul. 2006.
- [15] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, Jul. 2002.
- [16] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, Jul. 2002.
- [17] S. Miller, M. Nezamabadi, and S. Daly, "Perceptual signal coding for more efficient usage of bit codes," *SMPTE Motion Imaging Journal*, Oct. 2012.
- [18] SMPTE ST 2084:2014, "High dynamic range electro-optical transfer function of mastering reference displays," Aug. 2014.
- [19] R. Boitard, R. K. Mantiuk, and T. Pouli, "Evaluation of Color Encodings for High Dynamic Range Pixels," in *Proc. SPIE 9394, Human Vision and Electronic Imaging XX*, San Francisco, 2015.
- [20] ITU-R rec. BT.709, *basic parameter values for the HDTV standard for the studio and for international programme exchange*, Geneva, 1990.
- [21] ISO 11664-2:2007(E)/CIE S 014-2/E:2006, "CIE colorimetry - part 2: Standard illuminants for colorimetry."
- [22] T. Pouli, A. Artusi, F. Banterle, A. O. Akyuz, H.-P. Seidel, and E. Reinhard, "Color correction for tone reproduction," *CIC21: Twenty-first Color and Imaging Conference*, pp. 215–220, Nov. 2013.
- [23] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, Jul. 2002.
- [24] A. Luthra, E. François, and W. Husak, "Call for evidence (CfE) for HDR and WCG video coding," *ISO/IEC JTC1/SC29/WG11 N15083*, Feb. 2015.
- [25] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, and H. Brendel, "Creating cinematic wide gamut hdr-video for the evaluation of tone mapping operators and hdr-displays," vol. 9023, 2014, pp. 90 230X–90 230X–10, sequences available for download at <https://hdr-2014.hdm-stuttgart.de/>.
- [26] G. Krawczyk and R. Mantiuk, "pfstmo tone mapping library," <http://pfstools.sourceforge.net/pfstmo.html>.
- [27] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color research and application*, vol. 30, no. 1, Feb. 2005.
- [28] M. Narwaria, R. K. Mantiuk, M. P. Da Silva, and P. Le Callet, "HDR-VDP-2.2: a calibrated method for objective quality prediction of high-

- dynamic range and standard images," *Journal of Electronic Imaging*, vol. 24, no. 1, p. 010501, 2015, code available at : <http://hdrvdp.sf.net/>.
- [29] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb 2013.
- [30] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image compression by optimizing tone mapped image quality index," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3086–3097, 2015.
- [31] G. Bjontegaard, "Calculation of average PSNR differences between RD curves," *document VCEG-M33, ITU-T VCEG Meeting*, 2001.



Mikaël Le Pendu received the Engineering degree from Ecole Nationale Supérieure des Mines (ENSM) de Nantes, France in 2012. He is currently pursuing his Ph.D. studies in Computer Science in INRIA (Institut National de Recherche en Informatique et en Automatique) and Technicolor in Rennes, France. His current research interests include signal processing, image and video compression, and High Dynamic Range imaging.



Christine Guillemot is currently Director of Research at INRIA (Institut National de Recherche en Informatique et Automatique) in France. She holds a PhD degree from ENST (Ecole Nationale Supérieure des Telecommunications) Paris (1992). From 1985 to 1997, she has been with France Télécom in the areas of image and video compression for multimedia and digital television. From 1990 to mid 1991, she has worked as "visiting scientist" at Bellcore Bell Communication research) in the USA. Her research interests are signal and image processing, and in

particular 2D and 3D image and video coding, joint source and channel coding for video transmission over the Internet and over wireless networks, and distributed source coding.

She has served as Associate Editor for *IEEE Trans. on Image Processing* (from 2000 to 2003), for *IEEE Trans. on Circuits and Systems for Video Technology* (from 2004 to 2006), and for *IEEE Trans. on Signal Processing* (2007-2009). She is currently associate editor of the *Eurasip journal on image communication* (since 2010), for the *IEEE Trans. on Image Processing* (2014-2016), and for the *IEEE journal on selected topics in signal processing* (since 2013). She has been a member of the *IEEE IMDSP* (2002-2007) and *IEEE MMSP* (2005-2008) technical committees. She is currently a member of the *IEEE IVMSP - Image Video Multimedia Signal Processing - technical committee* (since 2013). She is the co-inventor of 24 patents, she has co-authored 9 book chapters, 62 international journal publications and around 150 articles in peer-reviewed international conferences. She is IEEE fellow since January 2013.



Dominique Thoreau received his PhD degree in image processing and coding from the University of Marseille Saint-Jérôme in 1982. From 1982 to 1984 he worked for GERDSM Labs on underwater acoustic signal and image processing of passive sonar. He joined, in 1984, the Rennes Electronic Labs of Thomson CSF and worked successively on sonar image processing, on detection and tracking in visible and IR videos, and on various projects related to video coding. Currently working in Technicolor, he is involved in exploratory video compression

algorithmics dedicated to the next generation video coding schemes.