Effective color space representation for wavelet based compression of HDR images

Masahiro Okuda *
The University of Kitakyushu
Department of Information and Media Sciences

Nicola Adami †
University of Brescia
Department of Electronics for Automation

Abstract

Effective color image and video coding is usually exploited coupling a compression method with the most suitable color space representation. This work extends a previously proposed High Dynamic Range (HDR) image coding method, which combine a logarithmic color adaptation module with a JPEG-2000 codec. Here we propose to change the original preprocessing stage with a more suitable one, based on the $LogLuv$ *color space representation, in order to take fully advantage of wavelet based coding. The experimental comparison between confirms that the proposed modification improves the compression performances and simplify the overall coding scheme.*

1 Introducion

In general consumer digital images are stored by using formats such as, TIFF, JPEG etc., which have been developed considering the characteristics of conventional rendering devices such as display or printers. Independently from the compression method used to reduce the data size, these formats, named ”output referred”, usually employ 8 bit per channel (bpc) or equivalently 24 o 32 bit per pixel (bpp) for a true color image. The limited number of colors that can be reproduced by using traditional formats derives from the limitations of old rendering devices in term of reproducible dynamic. Nowadays more powerful acquisition and rendering devices, that can support dynamic range wider than 8 bpc, are becoming available also to the consumer market and consequently, there will be soon a demand for standard and efficient tools for handle and store this type of images. For example CCD or CMOS sensors in many of today’s camera models can capture wider range of brightness and output raw images, as well as in a traditional format, which have up to 16 bits for each RGB color channel. The main difference between conventional and High Dynamic Range image formats is that the later represent the radiance of scene captured by a device or generated by an artificial rendering system. For this reason HDR images are also named as ”scene referred” and usually take floating point values in the XYZ color space. This also implies that HDR images are not suitable to be displayed on conventional CRT and LCD monitor without any prior dynamic and color gamut adaptation.

Originally, High Dynamic Range images (HDR) have been used mainly in high-end application such as professional photography and digital movie production. For this reason most of the tools related to HDR manipulation and storing are mainly devoted to lossless or near lossless application [3, 14], which are required in professional application. Nevertheless, several solution for efficiently represent and store HDR images have been proposed.

Similarly to the case of ”output referred”, formats for representing HDR images can be divided into two classes. The first refers to near lossless o perceptually lossless representation while the second concerns the lossy coding. At today, in our knowledge the first class is represented almost by storage formats such RGBE [14], [3] which allow a near loss less representation. The well-known RGBE format spends one byte each for mantissa of RGB colors and one common exponent (E) component, that is $E = \lceil \log_2(\max(R_W, G_W, B_W)) + 128\rceil$. LogLuv format[8, 7] converts RGB colors to the logarithm of luminance and the u, v channels, and then uses 16 bits for the luminance, 8 bit each for the u and v channels. In these two formats, each channel is losslessly compressed by the run length coding. OpenEXR[3] expresses each of the RGB channels by the sign bit, mantissa, and exponent with 16 bits, similar to IEEE floating point. This achieves ”almost lossless” representation and have options for some compression schemes, such as run-length coding, wavelet-based lossless coding. These three formats intend to preserve as much information as possible the original information and do not exploit cor-

*This work was partly supported by a Grant-in-Aid for Young Sciences (#14750305) of Japan Society for the Promotion of Science, fund from MEXT via Kitakyushu innovative cluster project.

†Thanks to the support of Invitation Fellowship Program of Japan Society for the Promotion of Science (JSPS) ID. No.: S-06090

relation.

The HDR images often contain a large amount of information and hence time-consuming to retrieve from a storage devise and to download through the network. Thus the demand for lossy compression that achieves higher rate-distortion performance increases. The conventional HDR lossy compression schemes are categorized in two types: one layer and two layer compressions. The JPEG-HDR[15] proposed by the G. Ward is a first attempt for the two layer compression, which has a compatibility with the conventional 24 bit image formats such as JPEG. That is, a tone mapped 24 bit image is first encoded by JPEG and then the ratio between the image and the original HDR is encoded and stored in a user-available buffer of the JPEG file. The decoder can not only see the image with conventional JPEG decoder, but also the HDR image recovered by the side information. Since the operations to create the LDR from the HDR images consists of the tone mapping, gamut clipping, quantization, and these operations are often complicated and may not be open for the public, the compatibility is preferable in many applications. A similar idea has been previously proposed to encode raw images of digital cameras by Spaulding [6] and [5]. They extend the dynamic range and color gamut by applying an inverse nonlinearity to the LDR images, and then encode the difference between the extended images and the original HDR images. In the conventional two layer algorithms, they pay little attention to tone mapping operation that transforms the HDR image to the 24 bit image. To our knowledge, only [9] tackles this problem. They compute a reconstruction function implemented with Look Up Table (LUT) that represents the inverse tone mapping and then compensate the difference of the two images. Although the LUT-based method approximates the tone mapping curve well, the approximation is not smooth. The lack of smoothness may yield extra energy in high frequency, which is not desirable in a sense of compression efficiency. Although these two layer coding has an advantageous property of the compatibility, the one layer coding is superior in a sense of compression efficiency. Mantiuk et al propose a HDR video compression scheme based on the MPEG. They derive an optimal quantization strategy using the threshold versus intensity functions of the human visual systems, and quantize the full dynamic ranges to 8 bit before MPEG encoding.

Xu et al. [10] apply JPEG2000 to the HDR images. They transform an image to logarithm domain, quantize to 16-bit integers, and then input it to the JPEG2000 encoder. Although their method does not provide the compatibility with the 24 bit images, its compression efficiency outpefomrs others. The present work aims to extend and improve the aforementioned approach by changing the color transformation and the dynamic adaptation operations performed prior to input the image to the JPEG2000 encoder. The document is organized as follow. In Section 2 the original compression method, hereafter considered, is described. In Section 3 the proposed modifications are presented and motivated. Section 4 briefly describes the coding procedure while in Section 5 the performances of the original and the modified approaches are compared.

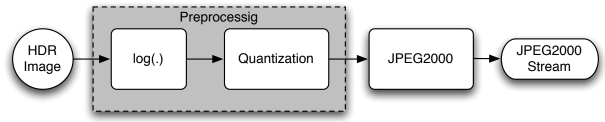


Figure 1. Original HDR image encoding scheme, Xu et al. [10]

2 Lossy compression of HDR images

The original coding scheme considered in this work, proposed by Xu et al. [10], is shown in Figure 1. Hereafter only the encoding part is described while decoding operations can be easily derived due to the symmetry of the process.

The JPEG-2000 standard do not intrinsically support the encoding of HDR images because it has been designed for madium-low dynamic range image coding. This implies that a preprocessing is required in order to adapt the original dynamic of the HDR image to the one supported by the JPEG2000 codec trying to minimize the information looses. This operation is accomplished by the first two blocks of the coding scheme (see Fig. 1). Assuming that the image values are in the RGB space, the preprocessing is described by the following relations:

$$\left[\bar{r},\bar{g},\bar{b}\right] = f\left(\left[r',g',b'\right] : n\right)$$

where the first block correspond to the logarithm transform of each color component $\left[r',g',b'\right] = \log\left([r,g,b]\right)$ and the second correspond to a dynamic scaling according to the number of bits that can be used to represent each color component and a quantization which maps a floating point value into an integer $f\left(x:n\right) = \left[\frac{(x-x_{min})}{(x_{max}-x_{min})}(2^n-1)\right]$. Although this coding scheme is characterized by compression performances which outperforms those of two layer methods some issues have been identified.

To apply the log transform of the RGB components it is not the perfect solution to adopt because JPEG2000 usually convert the input image to the YCbCr color space, in order to take advantage of its de-correlation properties. This means that because of the logarithm behaviors, artificial energy could be generated. A similar effects is introduced by the dynamic re-scaling, operated prior the float to integer conversion. In this case the dynamic of each component is rescaled to cover all the available input range with the primary intention to minimize the quantization step. As undesirable effect this operation could increase the distance between colors represented by slightly different values causing a loss of performance in the entropy encoder.

3 The proposed solution

As previously mentioned, wavelet based codecs have been designed for encoding conventional 24 bit images and then they can not effectively handle HDR images. This means that it is possible to encode a "scene referred" image but in general it will not be possible to achieve good compression performances. While the human visual system can easily accommodate the presence of very dark and very bright areas in the viewed scene because of local adaptation, classical coding schemes which uses a global approach generally tend to not encode relatively dark areas. This means that for equally consider bright and dark areas in the coding process it is required to reduce the dynamic of bright areas and at the same time to expand the dynamic of dark ones. Additionally, if the encoder process each color channel separately, the original color space representation should eventually be changed in order to have the maximum de-corelation.

The solution here proposed to tackle the necessary adaptation of HDR content is to use the $LogLuv$ color space representation. As will it will be shown in the following this approach allows for the necessary adaptation of the input images in order to exploit effective wavelet based coding without the drawback of the method described in the previous section.

The $LogLuv$ has been proposed by G. Ward [8] as a near losses HDR encoding for the TIFF image format. According to this method an image represented in the XYZ color space can be represented with a limited number of bit. Two alternative options are available: the first uses 24 bit to represent each color pixel value of the original image while the other allows a 32 bit encoding. According to Ward considerations the last seams to more flexible and provide a more accurate reproduction. Considering the encoding process of an HDR image by using the 32 bit LogLuv format the following transformation are to be applied. First of all, the original data are transformed from XYZ to $CIE(x, y)$

$$x = \frac{X}{X+Y+Z}$$
$$y = \frac{Y}{X+Y+Z}. \tag{1}$$

USC institute for creative technologies This intermediate representation is used to determine the LogLuv representation which correspond to a quantized version of the chromaticity space $u^{'}, v^{'}$ where the logarithm of the luminance is considered. This is obtained by applying the following transformations:

$$L_e = \lfloor 256(\log_2 Y + 64) \rfloor$$
$$u_e = \lfloor 410u^{'} \rfloor$$
$$v_e = \lfloor 410v^{'} \rfloor \tag{2}$$

where:

$$u^{'} = \frac{4x}{-2x+12y+3}$$
$$v^{'} = \frac{9y}{-2x+12y+3}. \tag{3}$$

By assigning to this log encoding 16 bit to L_e, where the first bit encode the sign, and 8 bit to u_e and v_e respectively it is possible to represent 38 order of magnitude which can capture the full range of perceivable light intensity [7]. In addition, the LogLuv provides a color representation perceptually similar to as the gamma-compressed YCbCr coordinates used in lossy compression method such as JPEG or JPEG2000.

Since the $LogLuv$ is used in this work only as a "color space" representation useful and not to directly store an image its 48 bit version has been adopted. This means that Equations 4 are changed as:

$$L_e = \lfloor 256(\log_2 Y + 64) \rfloor$$
$$u_e = \lfloor 32768u^{'} \rfloor$$
$$v_e = \lfloor 32768v^{'} \rfloor \tag{4}$$

and u_e and v_e are represented by using 16 bits for each component.

4 The coding Chain

The coding process, described in Section 2, has been modified according to the proposed modification. As shown in Figure 1 the preprocessing block has been replaced by the $LogLuv$

transformation. Operatively, the input HDR image values are initially mapped in the CIE XYZ color space. It has to be mentioned that this is not always required since HDR images are often represented by using this color space [3]. The XYZ representation is then transformed in the $LogLuv$ one by subsequently applying Equations 3 and 4 respectively. The image is then passed to the module that implements the JPEG-2000 compliant encoding. Contrarily

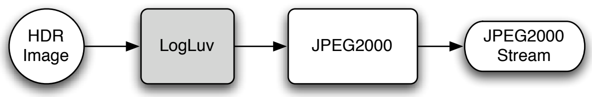


Figure 2. The proposed encoding scheme.

to what happens in the original scheme processing chain, the conversion of the input image to the YCbCr has now to be disabled.

5 Performances Evaluation

To prove the effectiveness of the method proposed in this work, its compression performances have been compared with those of the original one [10]. Tests have been conducted using HDR images obtained from several sources [13, 1, 4]. More specifically, each given image in the test set has been encoded and decoded at several different rates by using:

- the reference codec ”HDR-JP2”, freely available at the authors website [16]
- the proposed codec ”LogLuv-JP2”, which as been implemented by using the Kakadu front end available at [12].

All decoded images have been compared to the corresponding original by means of two different criteria:

- the Visual Difference Predictor (VDP) [11], which indicates the percentage of pixels in the tested image that will be perceived differently from the original with a given probability (75% or 95%). This measure considers only the luminance channel and is derived from Human Visual System models;
- The mean square error in the CIE Lab color space [2].

The joint uses of these two evaluation parameters allow for a quite complete comparison in term of both colors and luminance fidelity.

The obtained results have been plotted obtaining the corresponding Rate-Distortion curves. In Figures 3 and 4 such curves for ”Desk” and ”Apartment” images are shown. As it can be noticed the proposed method outperform the original one both in term of VDP and DE confirming that the $LogLuv$ representation improve coding efficiency. More in detail, it can be noted that the ”LogLuv-JP2” provides the lowest DE values at all bitrates. Beside, for the highest rates the VDP error values converges to 0% for both methods. Anyway the ”LogLuv-JP2” always reach this limit at a rate lower than its competitor.

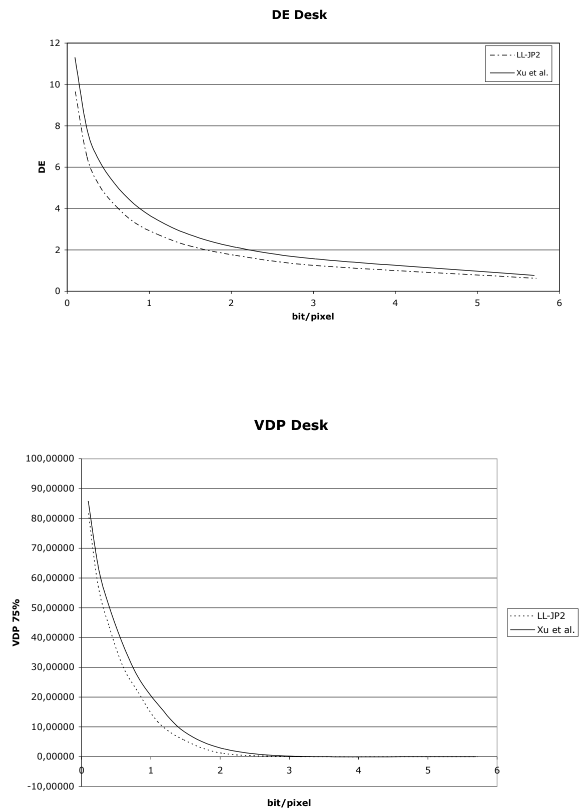
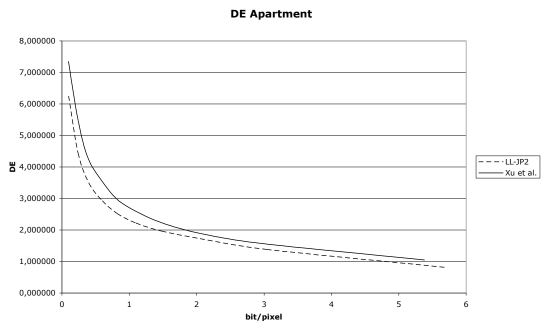


Figure 3. DE error (top) and VDP 75% for the image Desk

An example of the difference in performances is shown in Figure 5 where it is evident the effectiveness of the proposed method.

6 Conclusions

In this work we presented a wavelet based method for HDR image coding. This has been achieved coupling the $LogLuv$ color space representation with a JPEG-2000 compliant encoder wavelet based coding technology. The proposed scheme has show superior compression performance when compared with a previous similar scheme. The obtained gain is explained by the enhanced dynamic adaptation and color components de-correlation provided by the LogLuv color transformation with allows a more effective coding. Since, this has been obtained by using a standard JPEG-2000 implementation further improvement can be expected by the design of more appropriate quantization step and rate distortion measures.



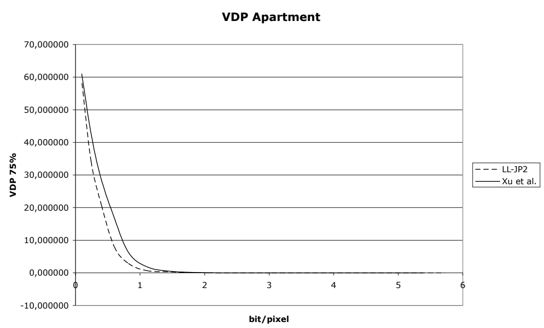


Figure 4. DE error (top) and VDP 75% for the image Apartment

Figure 5. Comparison of decoded picture Desk at 1.1 bit/pxel. Top picture refers to the proposed method (LL-JP2) while bottom pictures refers to Xu el al. method. Image courtesy ILM

References

[1] P. Debevec. Radiance maps. http://www.debevec.org/Research/HDR/.

[2] C. I. E. International commission on illumination. http://www.cie.co.at/cie/.

[3] R. B. F. Kainz and D. Hess. The openexr image file format. Technical sketch, SIGGRAPH, 2003.

[4] U. I. for Creative Technologies. High-resolution light probe image gallery. http://gl.ict.usc.edu/Data/HighResProbes/.

[5] G. J. W. K.E. Spaulding and R. Joshi. Using a residual image to extend the color gamut and dynamic range of an srgb image. In *The PICS Conference*, pages 307–314, 2003.

[6] G. W. K.E. Spaulding and E. Giorgianni. Reference input/output medium metric rgb color encodings (rimm/romm rgb). In *PICS 2000 Conference*, March 2000.

[7] G. Larson. Logluv encoding for full-gamut, high-dynamic range images. *J. Graphics Tools*, 3(1):15–31, 1998.

[8] G. W. Larson. Overcoming gamut and dynamic range limitations in digital images. In *sgi*, 1997.

[9] K. M. R. Mantiuk, A. Efremov and H. Seidel. Backward compatible high dynamic range mpeg video compression. In *SIGGRAPH*, 2006.

[10] S. P. R. Xu and C. E. Hughes. High-dynamic-range still-image encoding in jpeg 2000. *IEEE Computer Graphics and Applications*, 25(6):57–64, 2005.

[11] K. M. Rafal Mantiuk and H.-P. Seidel. Visible difference predicator for high dynamic range images. In *International Conference on Systems, Man and Cybernetics*, pages 2763–2769, 2004.

[12] D. S. Taubman. A comprehensive framework for jpeg2000. http://www.kakadusoftware.com/.

[13] G. Ward. Greg ward - anyhere software. http://www.anyhere.com/gward/hdrenc/pages/originals.html.

[14] G. Ward. Real pixels. In *Graphics Gems II*, pages 80–83. Edited by James Arvo, Academic Press, 1991.

[15] G. Ward and M. Simmons. Subband encoding of high dynamic range imagery. In *APGV ’04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 83–90, New York, NY, USA, 2004. ACM Press.

[16] R. Xu. High dynamic range image encoding. http://graphics.cs.ucf.edu/hdri/.