

PERCEPTUALLY LOSSLESS HIGH DYNAMIC RANGE IMAGE COMPRESSION WITH JPEG 2000

Yang Zhang*, David R Bull

Dept. of Electrical and Electronic Engineering
University of Bristol

Erik Reinhard

Max Planck Institute for Informatics
Saarbrücken, Germany

ABSTRACT

High Dynamic Range (HDR) technology offers high levels of immersion with a dynamic range meeting and exceeding that of the Human Visual System (HVS). A primary drawback of HDR images and video is that memory and bandwidth requirements are significantly higher than for conventional images and video. Many bits can be wasted coding redundant imperceptible information. The challenge is therefore to develop means for efficiently compressing HDR imagery to a manageable bitrate without compromising perceptual quality. In this paper, an HDR image compression method, based on an HVS optimized wavelet subband weighting method is proposed. The method has been fully integrated into a JPEG 2000 codec. Experimental results indicate that the proposed method outperforms previous approaches and operates in accordance with characteristics of the HVS, tested objectively using a HDR Visible Difference Predictor (VDP).

Index Terms— High Dynamic Range Imaging, Human Visual System, Visually Lossless Coding, JPEG 2000

1. INTRODUCTION

The human visual system can adapt from scotopic (10^{-5} – 10 cd/m^2) to photopic (10 – 10^6 cd/m^2) conditions [1] [2]. HDR (prototype) displays can now achieve a contrast ratio of 1,000,000:1 with a peak luminance of 4000 cd/m^2 . State of the art HDR imaging methods can cover a dynamic range from extremely dark (10^{-6} cd/m^2) to bright sunshine (10^8 cd/m^2) by using higher bit-depths for the luminance channel. HDR imaging allows a better dynamic range of exposures than conventional digital imaging techniques. However, an uncompressed HDR image demands significant storage space and occupies substantially more transmission bandwidth than Standard Dynamic Range (SDR) images.

To achieve HDR image compression, the Radiance RGBE/XYZE file format uses 8 bits for the red, green, and blue color channels and another 8 bits for the exponent, using run-length encoded 32 bit-per-pixel data [3]. The logluv color space encodes HDR pixels using a 16-bit logarithmic channel

representation for luminance and two 8-bit CIE chrominance (u',v') channels for chrominance [4] and is implemented as part of the public domain TIFF library. Industrial Light and Magic introduced the Extended Range format (known as OpenEXR), which encodes pixels by using 16-bit floating point values for R, G and B channels respectively [5]. These existing HDR image formats mostly use lossless compression strategies, so they are not compatible with reduced bit rate transmission and storage. A JPEG 2000 based HDR still-image encoding was proposed [6], which provides effective lossy compression, but performs poorly compared to other HDR image formats for lossless compression. Ward in 2005 [7] proposed a backward compatible HDR extension to JPEG.

In this paper, we propose a method of applying perceptually lossless HDR image coding within a JPEG 2000 framework. In LogLuv color space [4], HDR images require 16-bit for the luminance channel, which is compatible with JPEG 2000. A Discrete Wavelet Packet Transform (DWPT) based on an HVS-model is employed to reduce the imperceptible information caused by the imaging system. A Contrast Sensitivity Function (CSF) weighting is applied in the wavelet subband domain. This leads to significant improvements in terms of rate-distortion performance and visual quality, as measured with the HDR-VDP-2 visible difference predictor [8].

2. PERCEPTION-BASED WAVELET PACKET TRANSFORMATION

2.1. HVS-model for imaging applications

Several HVS models were developed over recent decades. One of the most important issues with HVS models concerns the inverse relationship between contrast sensitivity and spatial frequency. This phenomenon is described by the contrast sensitivity function (CSF), which is quantified by psychophysical experiments [9, 10]. The perceptual CSF model can be used for reduction of imperceptible information, and describes the capacity of the HVS to recognize differences in luminance and chrominance as a function of contrast and spatial frequency. Essentially, the HVS is more sensitive to lower spatial frequencies and less sensitive to high spatial frequencies. Mannos and Sakrison [11] originally presented a

*Contact author's e-mail: Yang.Zhang@bristol.ac.uk

CSF model for grayscale images as a non-linear transformation followed by a Modulation Transfer Function (MTF), as defined in the following. The spatial sampling frequency f_s (pixels per degree) is given by:

$$f_s = \frac{2v \cdot \tan(0.5^\circ)\gamma}{0.0254} \quad (1)$$

where v is the viewing distance (set to one meter in our experiments) and γ is the resolution of the display (pixels per inch). The maximum frequency represented in the CSF is defined as f_{\max} according to the Nyquist theorem, i.e. $f_{\max} = 0.5f_s$. The CSF in the luminance and chrominance channels are CSF_L , CSF_u and CSF_v :

$$\text{CSF}_L(f) = 2.6 \cdot (0.0192 + 0.114f \cdot \exp(-0.114f^{1.1})) \quad (2)$$

$$\text{CSF}_{u,v}(f) = a_1 \cdot \exp(-b_1 \cdot f^{c_1}) + a_2 \cdot \exp(-b_2 f^{c_2}) \quad (3)$$

where f is the spatial frequency in cycles per degree (cpd, measured in radians). The parameters $(a_1, b_1, c_1, a_2, b_2, c_2)$ for CSF_u are set to (5.623, 0.00001, 3.4066, 41.9363, 0.083, 1.3684). For the CSF_v , the parameters are (91.228, 0.0003, 2.803, 74.907, 0.0038, 2.601). Figure 1 shows the luminance and chrominance CSFs, where the black curve is the luminance CSF_L , the blue and red curves are the chromatic CSF_u and CSF_v respectively.

2.2. CSF Weighting in DWPT Subbands

To obtain the best trade-off between visual quality and compression ratio, the perceptual CSF model is employed in conjunction with the Discrete Wavelet Transform (DWT), which is one of the most efficient techniques for image compression and denoising and was shown to be useful for HVS modeling [12, 13, 14]. However, many applications require a finer frequency analysis (such as the nonlinear CSF weighting). One approach to achieve an accurate weighting method is to use the Continuous Wavelet Transform (CWT) to obtain better frequency resolution. The CWT provides almost unlimited flexibility but comes at the price of increased complexity. An alternative method which achieves better frequency localization whilst retaining the structure of a discrete decomposition is the Wavelet Packet Transform (WPT). Wavelet packets are functions that are defined continuously but can be discretely implemented similarly to the wavelet and scaling functions within a DWT. A signal can be adaptively decomposed according to a selection condition, based on for instance, a threshold condition (see Equation 4).

We apply the Discrete Wavelet Packet Transform (DWPT) to decompose HDR images into several spatial frequency channels with a limited range of orientations. Figure 1 shows the relation between the CSF and the wavelet subbands, where L and H represent low-pass and high-pass filtered image data. Parameters λ and Φ represent the levels and

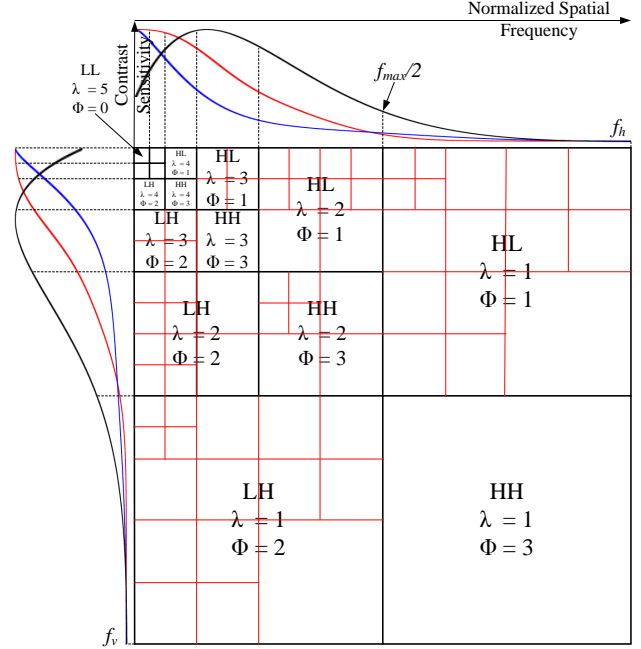


Fig. 1: Relation between luminance and chrominance CSFs and a 5-level 2D wavelet packet decomposition (see text).

orientations of the DWPT, and (f_h, f_v) are the horizontal and vertical frequency ranges which correspond to the DWPT subbands. To adapt a wavelet packet decomposition to CSF weighting, each subband is further decomposed only if a given threshold condition is met:

$$\text{DWPT}_{\lambda, \Phi}^{\theta, \eta} = \begin{cases} \text{DWPT}_{\lambda, \Phi}^{\theta, \eta} & \text{if } \overline{\text{var}}(f_h, f_v) < T_h \\ \text{DWPT}_{\lambda, \Phi}^{\theta+1, \eta} & \text{otherwise,} \end{cases} \quad (4)$$

where θ is the wavelet packet transformation level and η is the index of the wavelet packet subband. $\overline{\text{var}}(f_h, f_v)$ denotes the average variance value of the CSF in the spatial frequency range corresponding to the given DWPT subband, which is computed as follows:

$$\overline{\text{var}}(f_h, f_v) = \frac{1}{2} \left(\frac{1}{\text{card}(f_h)} \sum_{m=1}^{\text{card}(f_h)} \left(\text{CSF}(f_h(m)) - \mu_{f_h} \right)^2 + \frac{1}{\text{card}(f_v)} \sum_{n=1}^{\text{card}(f_v)} \left(\text{CSF}(f_v(n)) - \mu_{f_v} \right)^2 \right) \quad (5)$$

The preset threshold condition T_h can be computed with:

$$T_h = \frac{1}{\text{card}(f_{h_{LL}})} \sum_{i=1}^{\text{card}(f_{h_{LL}})} \left(\text{CSF}(f_{h_{LL}}(i)) - \mu_{f_{h_{LL}}} \right)^2 \quad (6)$$

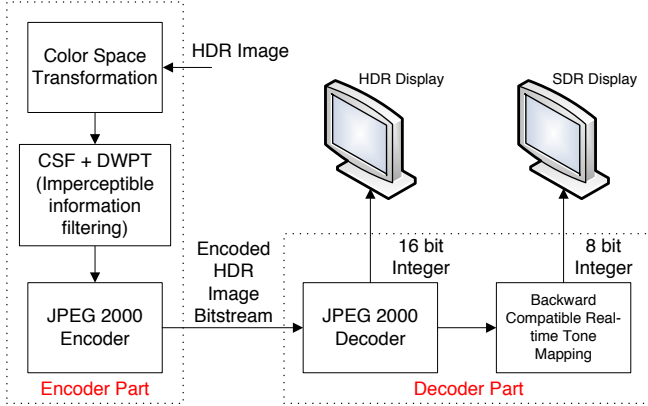


Fig. 2: Flowchart of experimental system

This type of conditional test adapts the decomposition tree to match the desired properties of the transform. In this equation, $f_{h_{LL}}$ denotes the horizontal spatial frequency range of the CSF which corresponding to the LL subband (when $\lambda = 4$ and $\Phi = 0$). Parameter μ represents the mean value of the CSF in the spatial frequency range corresponding to the DWPT subband:

$$\begin{cases} \mu_{f_h} &= \frac{1}{\text{card}(f_h)} \sum_{i=1}^{\text{card}(f_h)} \text{CSF}(f_h(i)) \\ \mu_{f_v} &= \frac{1}{\text{card}(f_v)} \sum_{i=1}^{\text{card}(f_v)} \text{CSF}(f_v(i)) \\ \mu_{f_{h_{LL}}} &= \frac{1}{\text{card}(f_{h_{LL}})} \sum_{i=1}^{\text{card}(f_{h_{LL}})} \text{CSF}(f_{h_{LL}}(i)) \end{cases} \quad (7)$$

We define $\bar{W}_{\text{CSF}}(f_h, f_v)$ as a band-average DWPT coefficient weighting function from the CSF curve in the spatial frequency domain:

$$\bar{W}_{\text{CSF}}(f_h, f_v) = \frac{\sqrt{\frac{\sum_{f_h} \text{CSF}^2(f_h)}{\text{card}(f_h)} + \frac{\sum_{f_v} \text{CSF}^2(f_v)}{\text{card}(f_v)}}}{2} \quad (8)$$

We implemented a five-level discrete 2-D wavelet packet transform followed by the above processing using the Cohen-Daubechies-Feauveau (CDF) 5/3 wavelet filter pair for lossless HDR image compression and a CDF 9/7 wavelet filter pair for lossy compression. The luminance and chrominance CSFs are employed in the DWPT domain. We weigh each subband coefficient by a normalized constant weighting value ($\bar{W}_{\text{CSF}}(f_h, f_v)$) which implements the CSF within the wavelet filtering algorithm. For the 5 level DWPT decomposition, the HVS is most sensitive in level 3 and sensitivity decreases with spatial frequency in levels 1, 2, 4 and 5 for the luminance channel. The sensitivity of the chrominance channel decreases significantly with increased spatial frequency. The

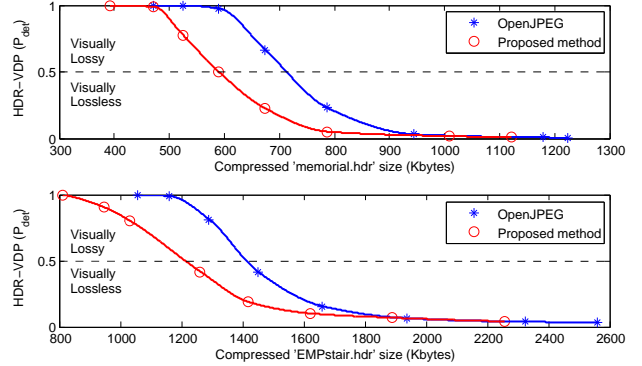


Fig. 3: Performance of the proposed method compared to the original OpenJPEG (JPEG 2000) compression system on two HDR images: memorial.hdr and EMPstair.hdr

resulting DWPT decomposition is shown in Figure 1. Finally, the filtered wavelet coefficients are inverse transformed and, if necessary, tone mapped for SDR display.

3. RESULTS AND DISCUSSION

We tested our proposed method on ten HDR images of natural scenes, as listed in Table 1. Figure 2 shows an overview of the proposed system, implemented as an extension to OpenJPEG¹ (Version 1.4). We use 16 bits per channel, the maximum bit-depth supported by JPEG 2000. To evaluate the performance of the proposed algorithm, the High Dynamic Range Visible Difference Predictor (HDR-VDP-2) [8] is used for objective testing. This algorithm is the most recent and accurate metric for assessing visible differences between HDR images, and therefore appropriate for determining the quality of our results. In this model $P_{det} \in [0, 1]$ is the probability of detecting a difference for the entire image assuming that each part of the image is equally attended. Under conservative conditions $P_{det} < 0.5$ indicates that visible differences are unlikely to be detectable.

Figure 3 shows the performance of the proposed method compared to the original OpenJPEG (JPEG 2000) compression system on two HDR test images. We use the compression ratio parameter of OpenJPEG to reduce the compressed data to a desired bit rate and HDR-VDP value. The output of our system is the compressed HDR data where the HDR-VDP value is just below the perceptible threshold to produce an optimized bit rate.

Table 1 shows the improvements offered by our approach for ten HDR test images compared to the JPEG 2000 (OpenJPEG) and numerically lossless HDR formats. Our proposed visually lossless compression method retains the same perceptual image quality and yields a significant improvement in rate distortion performance: 16.1% saving compared to JPEG

¹Available from: <http://www.openjpeg.org/index.php?menu=download>

Table 1: Storage requirements of ten test HDR image with proposed method and different HDR image formats.

HDR Image (KB)	Size (Pixels)	Proposed method	JPEG 2000 (OpenJPEG)	OpenEXR (.exr)	LogLuv (.tiff) (32 bits)	RGBE (.hdr)
memorial	512×768	590	787	1246	1190	1312
EMPstair	852×1136	1261	1451	2076	2667	3040
AtriumNight	760×1016	1158	1323	1983	2412	2547
BristolBridge	2048×1536	7549	8459	10132	7696	9895
mpi_atrium_1	1024×676	1384	1661	1916	2000	2375
BoyScoutTFalls	1000×1504	3007	3610	4497	4794	5142
BoyScoutFalls5	998×1496	2986	3583	4464	5048	5280
BoyScoutTree	998×1489	3566	4458	4224	4784	5169
sfnoma1	852×1136	1652	1936	1962	2640	2996
WardFlowers	1504×1000	3008	3609	3620	4292	4676



Fig. 4: Left: Reference image; Right: Compression result of proposed method.

2000 (OpenJPEG) on average, 38.9% saving compared to the RGBE format, 30.9% compared to the LogLuv TIFF format and 28.3% compared to the OpenEXR format. Finally, Figure 4 shows the compression results for the “memorial” HDR image. For the image on the left, our HDR image encoder compresses the original HDR image (1,312 KB) to that on the right hand side to (590 KB) offering a 55.03% reduction in storage or transformation bandwidth.

4. CONCLUSION

In this paper, an efficient visually lossless HDR image compression method is proposed that is compatible with JPEG 2000. The method employs luminance and chrominance CSF-based weighting in the DWPT domain to reduce the imperceptible information. Ten HDR images were evaluated in this study, and the experimental results indicate that the proposed method offers visually lossless quality at a significantly reduced bit rate.

5. REFERENCES

- [1] K. Myszkowski, R. Mantiuk, and G. Krawczyk, *High Dynamic Range Video*, Synthesis lectures in computer graphics and animation. Morgan & Claypool Publishers, 2008.
- [2] E. Reinhard, G. Ward, P. Debevec, S. Pattanaik, W. Heidrich, and K. Myszkowski, *High dynamic range imaging: acquisition, display and image-based lighting*, Elsevier, 2nd edition, 2010.
- [3] G.J. Ward, “The radiance lighting simulation and rendering system,” in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. ACM, 1994, pp. 459–472.
- [4] G.W. Larson, “LogLuv encoding for full-gamut, high-dynamic range images,” *Journal of Graphics Tools*, vol. 3, pp. 15–32, 1998.
- [5] R. Bogart, F. Kainz, and D. Hess, “Openexr image file format,” *ACM SIGGRAPH 2003, Sketches & Applications*, 2003.
- [6] R. Xu, S.N. Pattanaik, and C.E. Hughes, “High-dynamic-range still-image encoding in jpeg 2000,” *IEEE Computer Graphics and Applications*, vol. 25, no. 6, pp. 57–64, 2005.
- [7] G. Ward and M. Simmons, “Jpeg-hdr: A backwards-compatible, high dynamic range extension to jpeg,” in *In Proceeding of the 13th Color Imaging Conference*, 2005, pp. 283–290.
- [8] R. Mantiuk, K.J. Kim, A.G. Rempel, and W. Heidrich, “Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions,” *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, pp. 40, 2011.
- [9] P.G.J. Barten, *Contrast sensitivity of the human eye and its effects on image quality*, SPIE-International Society for Optical Engineering, 1999.
- [10] K.T. Mullen, “The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings,” *The Journal of Physiology*, vol. 359, no. 1, pp. 381–400, 1985.
- [11] J. Mannos and D. Sakrison, “The effects of a visual fidelity criterion of the encoding of images,” *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 525–536, 1974.
- [12] Y. Zhang, E. Reinhard, and D.R. Bull, “Perception-based high dynamic range video compression with optimal bit-depth transformation,” in *IEEE International Conference on Image Processing*. IEEE, 2011, pp. 1345–1348.
- [13] R. Mantiuk, A. Efremov, K. Myszkowski, H.P. Seidel, et al., “Backward compatible high dynamic range MPEG video compression,” *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 713–723, 2006.
- [14] A.P. Bradley, “A wavelet visible difference predictor,” *IEEE Transactions on Image Processing*, vol. 8, no. 5, pp. 717–730, 1999.