HDR IMAGE COMPRESSION USING OPTIMIZED TONE MAPPING MODEL

Nagisa Sugiyama, Hironori Kaida, Xinwei Xue, Takao Jinno, Nicola Adami, Masahiro Okuda†*

The University of Kitakyushu, Dpt. of Information and Media Sciences
* University of Brescia, Dpt. of Electronic for the Automation

ABSTRACT

In this paper, we propose a coding algorithm for High Dynamic Range Images (HDRI). Our encoder applies a tone mapping model based on scaled μ-Law encoding, followed by a conventional Low Dynamic Range Image (LDRI) encoder. The tone mapping model is designed to minimize the difference between the tone mapped HDRI and its LDR version. By virtue of the nature of the model, not only the quality of the HDRI but also the one of LDRI are improved, compared with a state of the art in conventional HDRI compression. Furthermore the error caused by our tone mapping model encoding is theoretically analyzed.

Index Terms— HDRI, coding, tone mapping model

1. INTRODUCTION

By adapting lights in any viewing condition, the human visual system can perceive a wider range of radiance (about 14 $\log_{10}$ units) than the one that can be captured by conventional camera sensors. In the last decade, to capture the high dynamic range of natural scene brightness, techniques have been proposed based on the multi-exposure image principle [1] - [3]. A work done by Mann et al. [2] has proposed a method for merging multiple photographs shot with different exposures, which realizes very high dynamic ranges. Devebec et al. [3] has also proposed a method to create High Dynamic Range Images (HDRI) (e.g. contrast ratio of 10^{10} : 1) and applied to high quality image based lighting. The HDRI image is often called 'scene-referred' image, since its pixel values are proportional to the intensity of scene luminance.

On the other hand, the conventional 24 bit image format is called 'output-referred'. The HDR imaging inspires many applications such as high quality CG rendering, in-vehicle sensors, camera surveillance, digital negative developments, etc, but its dynamic range is far beyond the one of conventional output devices. In many applications, the HDRI is often tone mapped to the displayable range. Many researches have proposed tone mapping operations[4]-[9]. These operations aim at reducing the high dynamic range without loss of detail. Thus although these discard a significant amount of information, one may estimate the loss by utilizing the obtained low dynamic range images.

Since the sizes of the HDR images are often huge, development for functional compression is one of the research topics. The image coding standard, JPEG 2000, provides seamless compression from 1 to 16 bits per color channel [10]. Xu et al.'s scheme [11] verifies the validity of JPEG2000 for the high dynamic range images. An HDR video compression scheme which uses MPEG2 has also been proposed [12]. Spaulding [13] proposes two layer encoding for gamut extended images. In the first layer, an image with clipped gamut (output-referred image) is encoded. In the second layer, the residual information that represents the difference between a gamut extended image and the decoded image in the first layer, is then encoded. The main advantage of this approach is that the format is compatible to existing file formats, and no extra efforts are needed to extract the output-referred image. In the field of Computer Graphics, similar concepts are adopted for the high dynamic image compression [14], [15], [16]. These two layer methods generally perform poorer than the one layer method [11] in a sense of coding efficiency.

In this paper, we introduce a coding algorithm for the HDRI which tries to combine advatages carried by the two class of methods described above. Our encoder applies a tone mapping model based on scaled μ-Law encoding. After applying the tone mapping model, we quantize the mapped image and input it to the conventional JPEG2000 encoder. The tone mapping model is designed to minimize the difference between the tone mapped HDRI and its reference LDR version (usually obtained by applying sophisticated non linear TM operators to the original HDRI). By virtue of the nature of the used μ-Law model, not only the quality of the decoded HDRI but also the one of LDRI are improved, compared with a state of the art in HDRI compression. Furthermore the error caused by out tone mapping model encoding is theoretically analyzed.

*Thanks to the support of Invitation Fellowship Programs of Japan Society for the Promotion of Science.

†This work was partly supported by a Grant-in-Aid for Young Sciences (#14750305) of Japan Society for the Promotion of Science.

2. PROPOSED ALGORITHM

The outline of our encoder is depicted in Fig.1(a), where 'TM' is a tone mapping operator and $f(\cdot)$ is an approximated model for the TM, which is discussed in detail later, and Q is quantization. Given a HDRI, we first convert it to a LDRI by a user-selected tone mapping operator. Any operator can be used unless it intends to yield unnatural effects. Using the two images, the HDRI and its tone mapped LDRI, the parameters of our tone mapping model are found by nonlinear least squares optimization. Then the HDRI is transformed by the tone mapping model with the optimized parameters, followed by the quantization. Finally the transformed image is input to JPEG2000. The decoder, illustrated in Fig.1(b), applies its inverse tone mapping model $f^{-1}(\cdot)$ after the JPEG2000 decoding. The parameters of the tone mapping model are sent to the decoder as side information.

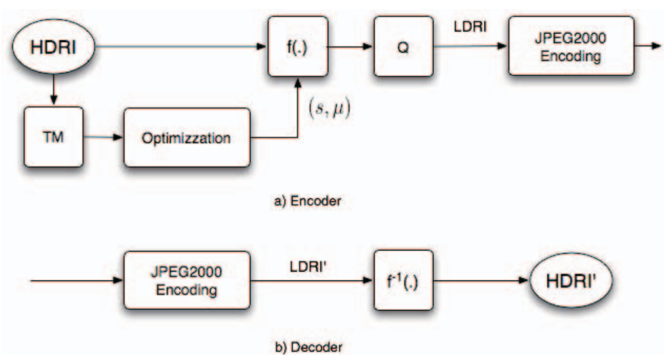


Fig. 1. Encoder and decoder

2.1. Tone mapping model

For evaluating compression performance, the criteria for image difference should be defined. The mean suqared error or equivalently SNR[1] can be a measure to evaluate the HDRIs. However these criteria do not fit visual quality especially for HDRIs due to its high dynamic range. The most effective way to evaluate the visual quality is based on the Human Visual System (HVS). Especially a nonlinear relationship to luminance is important for the HDRI compression. It is widely agreed that there is a nonlinear relationship between an amount of sensation and intensity of lights, or in other words brightness human perceives and actual luminance are not linearly related. There are many experimental results for the approximation of the nonlinearity. The well known Weber-Fechner's Law [17] indicates that the relationship between them is modeled by

$$y = k \log(x) + C, \tag{1}$$

where y and x are brightness that human retina perceives and input luminance, respectively.

[1]Since HDRIs does not have peak values explicitly, PSNR is not defined

In most applications of HDRIs, however, the images are used after tone mapping and output to current (low dynamic range) display devices. Thus it is reasonable to derive the nonlinearity model to minimize the quality of the tone mapped LDRI. Most of the tone mapping operators take the nonlinearity of the human perception models into account . The logarithmic function can be a good model and is used for the HDRI compression in [11]. This compression method does not intend to minimize the error of the input HDRI, but maximize the quality of its tone mapped version. They use the logarithm as a general tone mapping operator and do not consider any particular tone mapping. In contrast, our method optimizes a tone mapping operator to minimize the error between the HDRI and the LDRI converted by a particular tone mapping operator. Note that we show later that our scheme realizes lower distortion of the HDRI than [11] as well, even though our model is optimized for a particular operator.

We introduce the following tone mapping model, which is similar to μ-Law encoding [18]

$$f(x) = s \frac{\ln\{1 + (\mu/s)x\}}{\ln(1+\mu)}, \tag{2}$$

where s is a scaling parameter and μ is a parameter to control the "depth" of logarithm. Its inverse function is given by

$$f^{-1}(x) = \frac{s}{\mu}(e^{(x/s)\cdot\ln(1+\mu)} - 1). \tag{3}$$

We choose this function due to some desirable properties: (1) the function is controlled only by two parameters, (2) the function is monotonic and invertible, (3) the error caused by the model followed by the quantization is approximately given in closed form, which will be shown later.

To design the model, we first select a tone mapping operator that will be actually used after decoding. Then the LDRI is created by the operator. We find the parameters s and μ by minimizing a cost function:

$$\min_{s,\mu} E = \sum_i \{f(\mathcal{H}_i) - \mathcal{L}_i\}^2, \tag{4}$$

where, $\mathcal{H}$ and $\mathcal{L}$ are the HDRI and its tone mapped LDRI, respectively, and the suffix i is a pixel index.

In our method, the intensity of an input color image is calculated, and then the optimization is done for the intensity. Each of RGB channels is transformed by the same optimized model. After the image is transformed, it is uniformly quantized to integer in order to input the JPEG2000 encoder, that is,

$$y = Q\{f(\mathcal{H})\}, \tag{5}$$

where y is an input to the JPEG2000 encoder and Q is the quantization operator.

2.2. Error Analysis

Since in our framework the quantization is done in the tone mapping model encoding (5) before the JPEG2000, thus there are two sources where the quantization error occurs. Assume that the errors in (5) and in the JPEG2000 are independent, then the variance of the total error is the sum of the variance of errors caused by the two sources

$$\sigma_e^2 = \sigma_{\mu law}^2 + \sigma_{J2K}^2, \tag{6}$$

where σ_e^2, $\sigma_{\mu law}^2$, and σ_{J2K}^2 are the variances of the total error, error caused by (5) and JPEG2000, respectively. We are interested in $\sigma_{\mu law}^2$. In our case, the HDR image is transformed by (5) and then it is uniformly quantized. Letting the error of the HDRI around a pixel value x_i be in the range $[-\Delta q(x_i)/2,\ \Delta q(x_i)/2]$, then

$$\int_{-\Delta q(x_i)/2}^{\Delta q(x_i)/2} x_i^2 dx_i = \frac{1}{12}\Delta q^2(x_i) \tag{7}$$

holds. Thus the variance of the error of the HDRI is

$$\sigma_{\mu law}^2 = \frac{1}{12}\int_0^{x_m} \Delta q^2(x_i) p_x(x) dx, \tag{8}$$

where $p_x(x)$ and x_m are the pdf and the maximum value of the HDRI. Letting the stepsize of $Q(\cdot)$ in (5) be ΔQ, then

$$\frac{\Delta Q}{\Delta q(x_i)} = f'(x) \tag{9}$$

holds. From (2), the derivative of f is

$$f'(x) = \frac{1}{\ln(1+\mu)} \cdot \frac{\mu/s}{1+\mu x/s}. \tag{10}$$

We obtain the following equation from (8), (9), and (10).

$$\sigma_{\mu law}^2 = \frac{\Delta Q^2}{12}\{\ln(1+\mu)\}^2 \left(\frac{s}{\mu}\right)^2 \int_0^{x_m} \left(1+\frac{\mu}{s}x\right)^2 p_x(x)dx \tag{11}$$

Since the value μ is in the range $[100,\ \ 20000]$ in most cases, $(1+\mu x/s)^2 \approx (\mu x/s)^2$ holds. We consider the case that the maximum value of y in (5) is normalized by 1, and the quantization bits is N. Then we have $\Delta Q = 1/2^N$. Finally we have the closed form approximation of $\sigma_{\mu law}^2$.

$$\begin{aligned} \sigma_{\mu law}^2 &\approx \frac{\Delta Q^2}{12}\{\ln(1+\mu)\}^2 \int_0^{x_m} x^2 p_x(x)dx \\ &= \frac{\Delta Q^2}{12}\{\ln(1+\mu)\}^2 \sigma_H^2 \\ &= \frac{1}{12}\frac{1}{2^{2N}}\{\ln(1+\mu)\}^2 \cdot \sigma_H, \end{aligned} \tag{12}$$

where σ_H is the variance of the input HDRI.

2.3. Experimental Results

We have tested dozens of HDRIs, some of which are frequently used as sample images (Fig.2). We evaluate coding performance by two metrics. One is the mean squared error of the HDRIs. The other is the PSNR of the tone mapped LDRIs, which is a reasonable metric since in many applications the HDRIs are used after tone mapping and they are mostly designed to simulate the HVS. To minimize (4) , we apply a nonlinear least squares approximation implemented in Matlab's optimization toolbox.

Fig.3 shows the comparison with [11] for several test images, where the left and right columns indicate the MSE of the HDRI and the PSNR of the LDRI, respectively. The Reinhard's local tone mapping operator[7] is used. We use a code provided by the authors to implement [11]. We have tested several tone mapping operators [4]-[9] for the image "Dyrham Church" and "Belgium". Fig.4 shows the results of VDP-HDR [19], which is an image assessment tool that models some properties of the HVS, such as nonlinearity, frequency selectivity, direction selectivity, and masking. The VDP outputs a probability map that predicts a probability of error visibility for each pixel. Thus higher values represent that errors are more perceivable. In the figure, (+) and ($\times$) show our results and (*) is the results of HDR-JP2 [11]. The plot ($\times$) illusrates the average of the six μ-law functions that are optimized for the tone mapping operators [4]-[9]. From those examples it can be seen that our method not only improves the coding performance in the LDRI domain but also reduces the error of the HDRIs.

Note that since we contains the optimization in the encoding process, its computational complexity is higher than the conventional methods. In our experiments, the algorithm takes a few seconds for 768x512 images.

Fig. 2. Test Images: (left to right) Memorial, Atrium Night, and Dyrham Church

3. CONCLUSION

Since the HDRIs are used after tone mapping in many applications, improving the quality of LDRI is important. Our technique use the tone mapping model that is optimized to minimize the error between the HDRI and its tone mapped version for pre-processing before JPEG2000 compression. Our method outperforms the conventional method in senses both of the HDRI and LDRI.

4. REFERENCES

[1] E. Reinhard, S. Pattanaik, G. Ward and P. Debevec, “High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (Morgan Kaufmann Series in Computer Graphics and Geometric Modeling), Morgan Kaufmann Publisher 2005.

[2] S. Mann and R. Picard, ”On being ’undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures”, In Proceedings of IS&T 46th annual conference (May 1995), pp. 422–428. 1995

[3] P.E. Debevec and J. Malik. “Recovering High Dynamic Range Radiance Maps from Photographs, Proceedings of SIGGRAPH 97, Computer Graphics Proceedings, pp. 369-378 .

[4] S. N. Pattanaik, Mark D. Fairchild, James A. Ferwerda and Donald P. Greenberg, ”Multiscale model of Adaptation, Spatial Vision and Color Appearance”, Proceedings of IS&T/SID’s 6th Color Conference, Arizona, November 1998

[5] S. N. Pattanaik, Jack E. Tumblin, Hector Yee, Donald P. Greenberg,”Time-Dependent Visual Adaptation for Realistic Real-Time Image Display”, Proceedings of SIGGRAPH 2000, pp. 47-54, New Orleans, 23-28 July, 2000.

[6] Michael Ashikhmin, ”A tone mapping algorithm for high contrast images”, Proceedings of the 13th Eurographics workshop on Rendering, Pages: 145 - 156, 2002

[7] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, ”Photographic Tone Reproduction for Digital Images”, ACM Trans on Graphics. 21, 3, 267–276, 2002

[8] G.M. Johnson and M.D. Fairchild, “Rendering HDR images,” IS&T/SID 11th Color Imaging Conference, Scottsdale, 36-41, 2003

[9] ’tonemap()’ in Matlab Image processing toolbox

[10] David S. Taubman and Michael W. Marcellin, JPEG 2000: Image Compression Fundamentals, Standards and Practice, Kluwer International Series in Engineering and Computer Science,

[11] Ruifeng Xu, Sumanta N. Pattanaik, Charles E. Hughes: High-Dynamic-Range Still-Image Encoding in JPEG 2000. IEEE Computer Graphics and Applications 25(6): 57-64 (2005)

[12] R. Mantiuk, G. Krawczyk, K. Myszkowski and H. P. Siedel, ”Perception-Motivated High-Dynamic Range Video Encoding, ACM Trans. on Graphics, col23, no.3 2004, pp.773-741.

[13] Spaulding, Kevin E.; Joshi, Rajan L.; Woolfe, Geoffrey J., ”Using a residual image formed from a clipped limited color gamut digital image to represent an extended color gamut digital image”, United States Patent 6301393

[14] Ward, Greg, and Maryann Simmons, ”JPEG-HDR: A Backwards-Compatible, High Dynamic Range Extension to JPEG,” Proceedings of the Thirteenth Color Imaging Conference, November 2005.

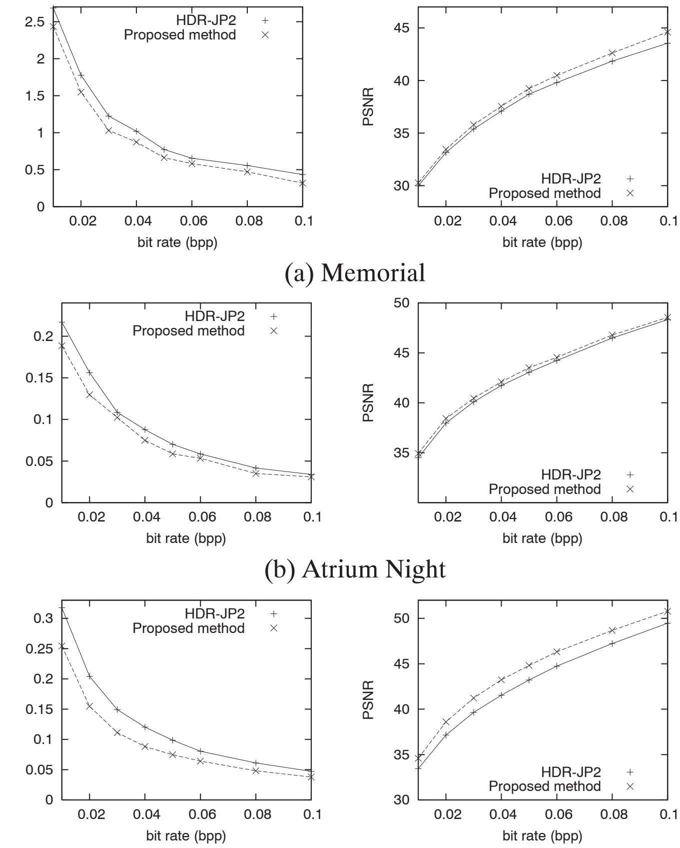
[15] R. Mantiuk, A. Efremov, K. Myszkowski, H.P. Seidel. ”Backward Compatible High Dynamic Range MPEG Video Compression”. To appear in: Proc. of SIGGRAPH ’06

[16] Masahiro Okuda, Nicola Adami, ”Two-Layer Coding Algorithm For High Dynamic Range Images based on Luminance Compensation,” Journal of Visual Communication and Image Representation, Elsevier, Vol. 18, Issue 5, pp.377-386, Oct. 2007.

[17] Spillmann, Lothar; Wener, John S, ”Visual Perception, the Neurphysiological Foundations,” Academic Press, San Diego, 1990.

[18] B.Smith, ”Instantaneous companding of quantized signals”, Bell System Technical Journal, Vol. 36, pp.653-709, May 1957.

[19] Rafal Mantiuk, Karol Myszkowski, Hans-Peter Seidel, R.l Mantiuk, K. Myszkowski, H-P Seidel, ”Visible Difference Predicator for High Dynamic Range Images, ” In: Proc. of IEEE International Conference on Systems, Man and Cybernetics, 2004. pp. 2763-2769

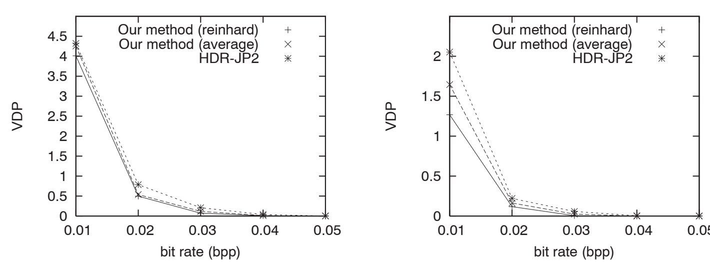


(a) Memorial

(b) Atrium Night

(c) Belgium

Fig. 3. Comparison with [11]: (Left column) MSE of HDRI, (Right column) PSNR of LDRI



(a) Dyrham Church (b) Belgium

Fig. 4. Results of VDP evaluation [19]