IEICE TRANS. FUNDAMENTALS, VOL.E94–A, NO.3 MARCH 2011

PAPER

μ-Law Based HDR Coding and Its Error Analysis

Takao JINNO[†a)], ***Member***, **Hironori KAIDA**[†], **Xinwei XUE**[†], **Nicola ADAMI**[††], ***Nonmembers***, ***and*** **Masahiro OKUDA**[†], ***Member***

SUMMARY 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 μ-Law model, not only the quality of the HDRI but also the one of the LDRI is improved, compared with a state of the art in conventional HDRI coding methods. Furthermore the error limit caused by our encoding is theoretically analyzed.
key words: *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 280 dB) than conventional camera sensors. In the last decade, to capture the high dynamic range of natural scene radiance, some techniques have been proposed based on the multi-exposure image principle [1]–[3]. Mann et al. [2] proposed a method for merging multiple photographs taken with different exposures. Devebec et al. [3] have also proposed a method to create the High Dynamic Range Images (HDRI) (e.g. contrast ratio of 10^{10} : 1) and applied it to high quality image based lighting.

The HDR imaging inspires many applications such as high quality CG rendering, in-vehicle sensors, camera surveillance, digital negative developments, and etc. Its dynamic range is, however, far beyond the one of conventional output devices, and thus the range needs to be compressed to a displayable range in order to properly express the HDRI by using a conventional 24 bit image format. This compression procedure is called “tone mapping”, Many researchers have proposed tone mapping operations [4]–[9]. These operations aim at reducing the high dynamic range without the loss of local contrast. In general the tone mapping discards a significant amount of information that the Human Visual System (HVS) is insensitive to, and keeps its detail unchanged.

Since the sizes of the HDR images are often huge, development for functional compression is one of significant 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 that uses MPEG2 has also been proposed in [12]. Spaulding [13] proposes a two layer encoding method for gamut extended images. In the first layer in [13], an image with clipped gamut (output-referred image) is encoded. In the second layer, the residual information, which represents the difference between its 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 such as JPEG, 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]–[16]. These two layer methods generally perform poorer than the one layer method [11] in the sense of coding efficiency.

In this paper, we introduce a coding algorithm for the HDRI that tries to combine advantages carried by the two classes of the 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 JPEG 2000 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 μ-Law model, not only the quality of the decoded HDRI but also the one of LDRI is improved, compared with a state of the art in HDRI compression. Furthermore the error caused by out tone mapping model encoding is theoretically analyzed.

2. Previous Research

The one layer HDR coding methods such as [11] and [12] use the tone mapping as a pre-processing function. The tone mapping plays a role of squeezing the dynamic range where the HVS is insensitive. The conventional JPEG and MPEG encoders basically does not discriminate in coding precision based on pixel intensities, that is, error on each pixel follows a uniform distribution and does not depend on its intensity. Therefore in order to control error depending on intensities, the pre-processing may be a simple and effective approach.

The tone mapping operators can be categorized into

Manuscript received July 13, 2010.
Manuscript revised October 20, 2010.
[†]The authors are with the University of Kitakyushu, Kitakyushu-shi, 808-0135 Japan.
[††]The author is with Brescia University, Italy.
a) E-mail: d09f0101@hibikino.ne.jp
DOI: 10.1587/transfun.E94.A.972

three types: global operators, local operators, and operators in the frequency domain [1]. The last two types generally outperform the global operators in the sense of contrast preservation. However unfortunately since most methods of these two types are not invertible, they cannot be employed in our framework.

The well known method by Xu et al. uses a logarithm function in [11], which is one of the simplest global operators, as the pre-processing. This choice is reasonable since it matches the luminance versus brightness nonlinearity in the HVS as the Weber-Fechner's Law indicates. However the range compression performance of the log function depends on the dynamic range and the scale of an input image. For example, it brings different effects to a HDR in cd/m^2 unit and its scaled version. Drago et al. have proposed [17] a tone mapping that can control the logarithm effect by varying the base of the log function. Although it realizes better quality than the simple log function, unfortunately it is not invertible. Another drawback of the conventional pre-processing approaches is flexibility. It generates the best quality of tone-mapped LDRI at the decoder side when the same tone mapping is applied to the decoded HDRI. Thus, it can result in undesirable range compression if a user uses a tone mapping operator that is different from the pre-processing function. To address these two issues, we introduce a more flexible tone mapping operator.

3. 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 a quantization operator. 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 generate 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 compressed by JPEG 2000. Strictly speaking, the quantization step is not mandatory, since many of the standards in image compression, including JPEG 2000, does not specify anything in encoding and the encoder that handles the images with floating point values can be implemented. However most existing implementations of the encoder support only integers. Thus we intorduce the quantization before the image compression to fully utilize the existing encoders. The decoder, illustrated in Fig. 1(b), applies its inverse tone mapping model $f^{-1}(\cdot)$ after the JPEG 2000 decoding. The parameters of the tone mapping model are sent to the decoder as side information.

3.1 μ-Law Based Tone Mapping Model

For evaluating compression performance, a criteria for image difference should be defined. The mean squared error (or equivalently SNR) is used to evaluate image quality. However these criteria do not fit perceptual visual quality especially for HDRIs due to its high dynamic range. For the HDRI evaluation, a nonlinear response of HVS should be taken into account. 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 [18] indicates that the relationship between them is modeled by

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

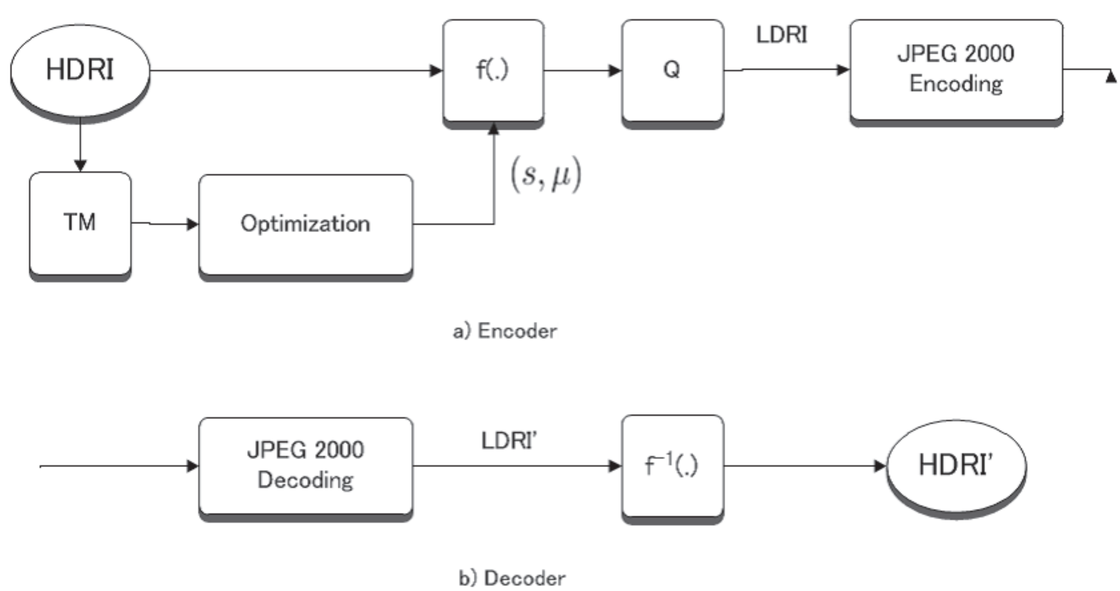


Fig. 1 Encoder and decoder.

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

In most applications of HDRIs, the images are tone-mapped for visualizing on current (low dynamic range) display devices. Thus it is reasonable to derive the nonlinearity model to maximize 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 choice and is used for the HDRI compression in [11]. This compression method does not intend to minimize the error of the input HDRI, but it maximizes 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 also realizes lower distortion of the HDRIs than [11], even though our model is designed to minimize the errors of the LDRIs.

We introduce the following tone mapping model, which is similar to μ-Law encoding [19]

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

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

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

The μ-Law curves with several parameters are shown in Fig. 2.

We choose this function because of 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 a closed form as shown in Sect. 3.2.

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 the 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 performed 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 values to the JPEG 2000 encoder, that is,

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

where y is an input to the JPEG 2000 encoder and Q is the quantization operator.

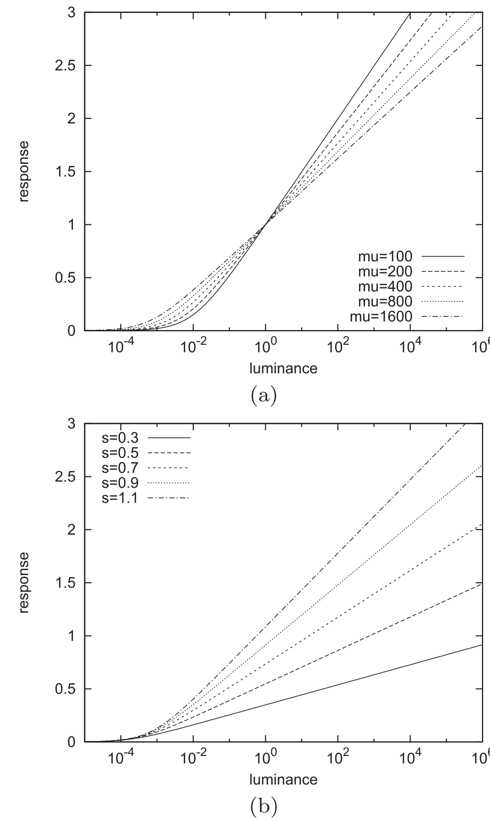


Fig. 2 μ-Law curves with several parameter settings. (a) μ = 100 to 1600, s=1 (b) s=0.3 to 1.1, μ = 1600.

3.2 Error Analysis

Most implementations of the conventional codecs such as H.264/AVC, JPEG, JPEG 2000 allow only input images with integer values. In our framework the quantization (5) is performed before the JPEG 2000 is applied. Thus there are two sources where the quantization error occurs. Assume that the errors in (5) and in the JPEG 2000 are independent, then the variance of the total error is the sum of errors caused by the two sources

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

where σ_e^2 is the total error variance. $\sigma_{\mu law}^2$, and σ_{J2K}^2 are the variances of error caused by the quantization (5) and JPEG 2000, respectively.

The error variance $\sigma_{\mu law}^2$ is not uniquely determined. It depends on the dynamic range of an input image. To fully utilize the conventional codec by exploiting rate-distortion performance, it is helpful to theoretically analyze the error.

In our case, the HDR image is transformed by the μ-Law function 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

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

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

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

where $p_x(x)$ and x_m are the probability density function 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).

$$\begin{aligned}\sigma^2_{\mu law} &= \frac{\Delta Q^2}{12}\{\ln(1+\mu)\}^2\left(\frac{s}{\mu}\right)^2 \\ &\quad \cdot \int_0^{x_m}\left(1+\frac{\mu}{s}x\right)^2 p_x(x)dx\end{aligned} \tag{11}$$

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 $\sigma^2_{\mu law}$.

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

where σ_H is the variance of the input HDRI.

4. Experimental Results

We have tested dozens of HDRIs, some of which are frequently used as sample images. Some of the sample images are listed in Fig. 3 and Table 1.

Our implementation of the encoder/decoder is same as the one of [11] except for the tone mapping. We implemented with JasPer [23]. We use the lossy mode with 9/7-tap Daubechies'wavelet, and the scalar quantization. We does not use the chrominance subsampling, which is also same as [11]. To obtain the parameters s and μ of our tone mapping model, we use the interior trust region approach for the reflective Newton method [22]

We evaluate coding performance by several 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 the tone mapping is applied and they are mostly designed to simulate the sensitivity of the HVS.

Figure 4 is the example for the tone mapping curves of the log function [11] (dotted line) and our method (solid line). Figure 5 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 for designing the μ-Law function. We use a code provided by the authors to implement [11]. Figure 6 shows the results of the VDP-HDR [20] that is an image assessment tool that models some properties of the HVS, such as

Fig. 3 Test images: (top left to bottom right) Memorial, Nave, Atrium Night, Dyrham Church, Belgium, and Adobe Fountain.

Table 1 Test images.

Image	Dynamic Range in $\log_{10}$ unit	Size	Source
Memorial Church	5.53	768×512	Image Courtesy of Paul Debevec
Nave	8.3442	480×720	Image Courtesy of Paul Debevec
Atrium Night	8.4886	1016×760	Image Courtesy of Karol Myszkowski
Belgium	5.7906	768×1024	Image Courtesy of Dani Lischinskii
Dyrham Church	4.0617	1536×2048	Image Courtesy of Greg Ward
Adobe Fountain	3.2071	2272×1704	[1]

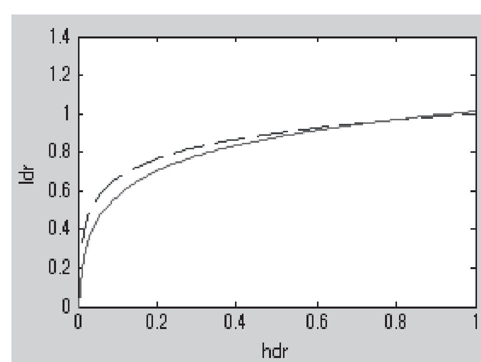
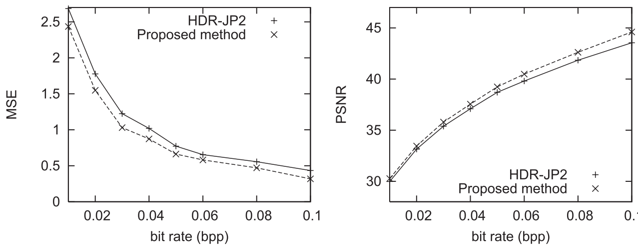
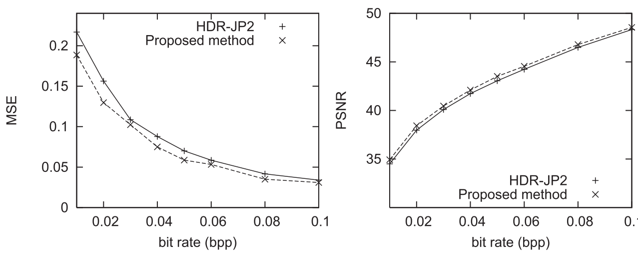


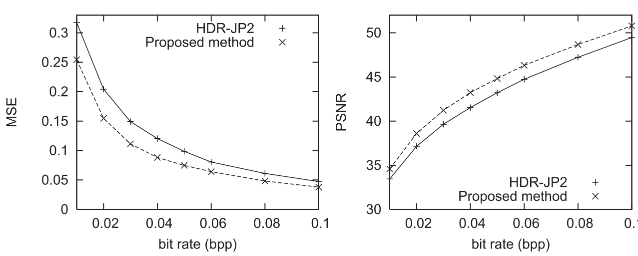
Fig. 4 Two tone mapping curves (both curves are scaled to [0, 1]): (dotted line) The conventional method [11], (solid line) our method (s = 0.6157, μ = 108.9).



(a) Memorial ($s = 0.6157, \mu = 108.9$)

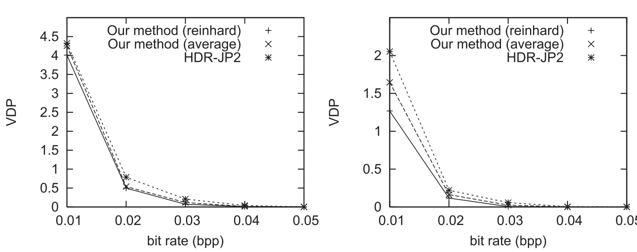


(b) Atrium Night ($s = 0.7709, \mu = 312.8$)



(c) Belgium ($s = 0.6823, \mu = 185.7$)

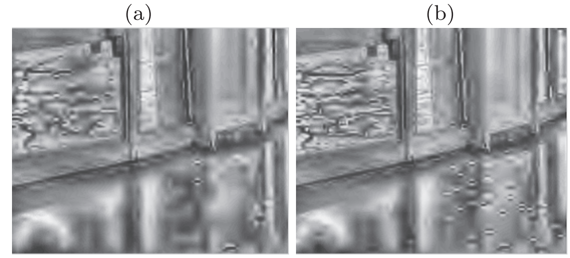
Fig. 5 Comparison with [11]: (Left column) MSE of HDRI, (Right column) PSNR of LDRI.



(a) Dyrham Church (b) Belgium

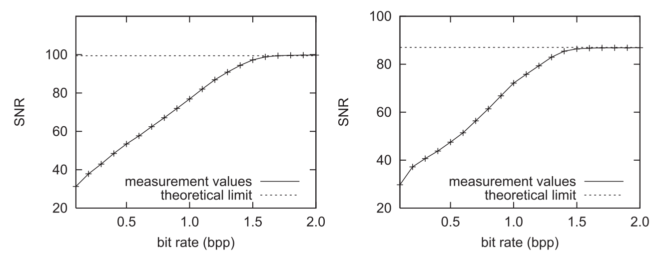
Fig. 6 Results of VDP evaluation [20].

(a) (b)



(c) (d)

Fig. 7 Results of compression: (a) [11], (b) our method, (c) Part of (a), (d) Part of (b).



Memorial

Fig. 8 Theoretical error limit.

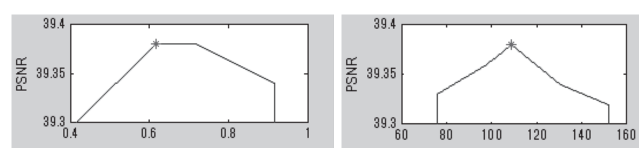


Fig. 9 Coding performance with respect to parameter variations, * indicates the optimal values: (left) plots of several s with fixed μ, (right) plots of several μ with fixed s.

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 mean that errors are more perceivable. In the figure, (+) and (×) show our results and (*) is the results of HDR-JP2 [11]. The plot (×) illustrates 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.

The images obtained by using [11] and our method are depicted in Fig. 7(a, c) and (b, d), respectively. These samples are obtained with the same setting and low bit rate (0.005 bpp). The contrast of the images are enhanced by the gradient based method [21] to improve its visibility. One can see that our method preserves contrast more in detail even at the low bit rate.

Finally we analyze the optimality of the method. Our tone mapping model is obtained by minimizing the squared error of the HDRI and LDRI. This approach does not guarantee its optimality in the sense of coding efficiency. However in practice, the coding with our model yields almost highest performance in most cases. Figure 9 shows the PSNR of LDRIs obtained with different parameters. The image "memorial" and Reinhard's local tone mapping are used for this example. The plot in Fig. 9(left) and (right) shows the results with various values of s and μ with the other parameter fixed, respectively, where '*' indicates the results of the optimized parameter. One can see the validity of our optimization.

Next we examine the validity of the error limit specified by (12) in Fig. 8, in which one can see that the rate-distortion curve approaches the limit calculated by (12), that is the dotted line, as the bit rate is elevated.

Note that since the algorithm contains the optimization procedure in the encoding process, its computational complexity is higher than the conventional methods. In our experiments, the algorithm takes a few seconds for 768×512 images.

5. Conclusion

Since the HDRIs are used after tone mapping in many applications, improving the quality of LDRI is important. Our technique uses the tone mapping model that is optimized to minimize the error between the HDRI and its tone-mapped version for pre-processing before JPEG 2000 compression. Our method outperforms the conventional method in senses both of the HDRI and LDRI.

Acknowledgement

We are grateful for the support of a Grant-in-Aid for Young Sciences (#14750305) and "Invitation Fellowship Programs for Research in Japan" of Japan Society for the Promotion of Science, and Heiwa Nakajima Foundation.

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," Proc. IS&T 46th Annual Conference, pp.422–428, May 1995.
[3] P.E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," Proc. SIGGRAPH 97, Computer Graphics Proceedings, pp.369–378, 1997.
[4] S.N. Pattanaik, M.D. Fairchild, J.A. Ferwerda, and D.P. Greenberg, "Multiscale model of adaptation, spatial vision and color appearance," Proc. IS&T/SID's 6th Color Conference, Arizona, Nov. 1998.
[5] S.N. Pattanaik, J.E. Tumblin, H. Yee, and D.P. Greenberg, "Time-dependent visual adaptation for realistic real-time image display," Proc. SIGGRAPH 2000, pp.47–54, New Orleans, July 2000.
[6] M. Ashikhmin, "A tone mapping algorithm for high contrast images," Proc. 13th Eurographics workshop on Rendering, pp.145–156, 2002.
[7] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," ACM Trans. Graphics, vol.21, no.3, pp.267–276, 2002.
[8] G.M. Johnson and M.D. Fairchild, "Rendering HDR images," IS&T/SID 11th Color Imaging Conference, pp.36–41, Scottsdale, 2003.
[9] 'tonemap()' in Matlab Image processing toolbox
[10] D.S. Taubman and M.W. Marcellin, JPEG 2000: Image Compression Fundamentals, Standards and Practice, Kluwer International Series in Engineering and Computer Science, 2000.
[11] R. Xu, S.N. Pattanaik, and C.E. Hughes, "High-dynamic-range still-image encoding in JPEG 2000," IEEE Comput. Graph. Appl., vol.25, no.6, pp.57–64, 2005.
[12] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.P. Siedel, "Perception-motivated high-dynamic range video encoding," ACM Trans. Graphics, vol.23, no.3, pp.773–741, 2004.
[13] K.E. Spaulding, R.L. Joshi, and G.J. Woolfe, "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] W. Greg and M. Simmons, "JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG," Proc. Thirteenth Color Imaging Conference, Nov. 2005.
[15] R. Mantiuk, A. Efremov, K. Myszkowski, and H.P. Seidel, "Backward compatible high dynamic range MPEG video compression," Proc. SIGGRAPH'06, 2006.
[16] M. Okuda and N. Adami, "Two-layer coding algorithm for high dynamic range images based on luminance compensation," J. Vis. Commun. Image Represent., vol.18, no.5, pp.377–386, Elsevier, Oct. 2007.
[17] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," Proc. Eurographics, pp.419–426, 2003.
[18] L. Spillmann and J.S. Wener, Visual perception, the neurophysiological foundations, Academic Press, San Diego, 1990.
[19] B. Smith, "Instantaneous companding of quantized signals," Bell Syst. Tech. J., vol.36, pp.653–709, May 1957.
[20] R. Mantiuk, K. Myszkowski, H.-P. Seidel, R.L. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Visible difference predicator for high dynamic range images," Proc. IEEE International Conference on Systems, Man and Cybernetics, pp.2763–2769, 2004.
[21] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," SIGGRAPH'02: Proc. 29th Annual Conference on Computer Graphics and Interactive Techniques, pp.249–256, July 2002.
[22] T.F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," SIAM J. Optim., vol.6, no.2, pp.418–445, May 1996.
[23] M.D. Adams and F. Kosentini, "JasPer: A software-based JPEG-2000 codec implementation," Proc. IEEE Int'l Conf. Image Processing, vol.2, no.10–13, pp.55–56, 2000.

Takao Jinno received B.E. in 2007 from the University of Kitakyushu. He has been with the graduate school of Environmental Engineering, The University of Kitakyushu. He is engaged in high dynamic range image processing. He received 2008 Excellent Student Award of the IEEE Fukuoka Section in 2009, and Hibikino Award in 2009.

Hironori Kaida received B.E. and M.E. in 2007 and 2009 from the University of Kitakyushu, respectively. He has been with Canon Inc. He is engaged in high dynamic range image processing.

Xinwei Xue received B.E. and M.E. in 2008 and 2010 from the University of Kitakyushu, respectively. She is now in Doctral Degree of Waseda University.

Nicola Adami received Ph.D. in 2002 from Brescia University. Since 2004, he has been Assistant Professor at the Engineering Faculty - University of Brescia. His research interests include multimedia signal processing, adiovisual (AV) indexing, audio visual content description processing, scaling, filtering and integration.

Masahiro Okuda received the B.E., M.E., and Dr.Eng. in 1993, 1995, 1998, respectively from Keio University, Yokohama, Japan. He was with University of California, Santa Barbara and Carnegie Mellon University as a visiting scholar in 1998 and 1999, respectively. He has been on the faculty of The University of Kitakyushu as an associate professor of Environmental Engineering since 2001. His research interests include filter design, vision/geometry coding, and multirate signal processing.