TWO LAYER SCHEME FOR ENCODING OF HIGH DYNAMIC RANGE IMAGES

Ishtiaq Rasool Khan

Department of Brain Science and Engineering
Kyushu Institute of Technology, Kitakyushu, Japan 808-0196

ABSTRACT

With advent of high dynamic range (HDR) imaging techniques, it has been possible to capture natural scenes in larger details. HDR images are tone-mapped to lower dynamic range (LDR) versions for displaying on paper or a screen. Details lost during tone-mapping are important for certain existing and future applications, and need to be preserved. However, the size of HDR images is very large and that gives rise to need of effective encoding techniques. In this paper, we present an encoding scheme for HDR images, which significantly reduces their storage requirements, with a negligible loss of information. We model an HDR image as a piecewise linear function of its tone-mapped version. The tone-mapped image and the error in modeling are stored as LDR images, and these two along with the created model, approximate the HDR image. Comparison with the existing state of the art technique is given to show the effectiveness of our proposed scheme.

***Index Terms*—** Image processing, Image coding, High dynamic range imaging, HDR, Raw images

1. INTRODUCTION

Real world scenes have much larger dynamic range (ratio between the luminance of bright and dark regions) as compared to that can be shown on photographic papers or computer screens. Modern digital cameras come with significantly larger dynamic range, and moreover, a series of pictures taken at different exposures by ordinary cameras can also be combined to create high dynamic range (HDR) maps close to the real world scenes. Only a part of information contained in these HDR images can be displayed on paper or a screen, which can display only a smaller range of contrast [1]. Several tone-mapping operators (TMO) have been presented to generate low dynamic range (LDR) versions of these HDR images for purpose of displaying. These operators typically reduce a range of over 5 orders of magnitude to 2 orders of magnitude, and use different criteria to preserve the appearance of the originals scene. A comparison of different TMOs can be seen in [2].

Some useful information contained in HDR images might be lost during tone-mapping. Although tone-mapped LDR versions are suitable for printing on paper and visualizing on most of the current displays, the additional information contained in HDR images is useful in certain applications like high quality CG rendering, sensing and surveillance etc. Moreover, HDR displays are likely to be in common use in near future, and the images stored with higher details would give a better visualization of the actual scene when displayed on them. Due to these reasons, much research is going on finding suitable formats for encoding the HDR data.

G. Ward presented a RGBE format with 4 bytes for each pixel; 3 bytes for RGB values and one byte for a shared exponent [3]. Ward presented another encoding known as LogLuv format, which separates luminance and chrominance channels and applies log encoding to luminance taking into account the human visual system [4]. Industrial Light and Magic (ILM) has released a 48 bits/pixel format (a sign bit, five bits of exponent, and a ten-bit mantissa for each channel), which covers all visible gamut and about 10.7 orders of magnitude, and is supported by many applications [5]. A detailed discussion of different HDR image formats can be seen in [6].

Since the size of raw HDR images presented in existing formats is often huge, development of HDR compression techniques is one of today's attractive research topics in image processing community. The image coding standard JPEG 2000 provides seamless compression from 1 to 16 bits per color channel [7]. Ruifeng et al.'s scheme [8] verifies the validity of JPEG2000 for the high dynamic range images. The HDR video compression scheme using MPEG2 has also been proposed [9].

Spaulding [10-11] proposed a two-layer encoding for gamut extended images. In the first layer, an image with clipped gamut is encoded, and then 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 encoded. G. Ward used a similar approach and presented his JPEG-HDR [12] extension to the existing JPEG standard. In this format, a tone-mapped LDR image is encoded as a standard JPEG image, and the ratio between its luminance and the actual HDR luminance, also encoded as a

JPEG image, is stored in one or more wrappers provided for storing additional information under JFIF/JPEG coding standards. A major advantage of this approach is its backward compatibility to the existing JPEG based applications. A non-HDR enabled JPEG application simply ignores the information contained in the wrappers and looks only at the tone-mapped LDR image stored in the first layer. Similar approach has been used by Mantiuk et al. for HDR video compression [13].

In this paper, we follow an approach similar to [10-13], and encode the dynamic range of an HDR image by splitting it into two JPEG images. The first image is a tone-mapped LDR version of the original HDR image, and the second image stores the residual information lost during tone-mapping. We reduce the dynamic range (and therefore the size) of the residual image by modeling the HDR image as piecewise linear approximation of the front layer LDR image. Unlike existing techniques that preserve the residual data only for the luminance channel and take additional measures for the information lost in other channels, we preserve the residual data for all channels. However, overall size of our residual image is still smaller than that of existing single channel techniques, due to reduction in the dynamic range as result of piecewise linear modeling. Experimental results show that our scheme better preserves the HDR details as compared to the existing techniques.

2. THE PROPOSED METHOD

Outline of our proposed encoder is given in Fig. 1, while the major steps are explained here in this section.

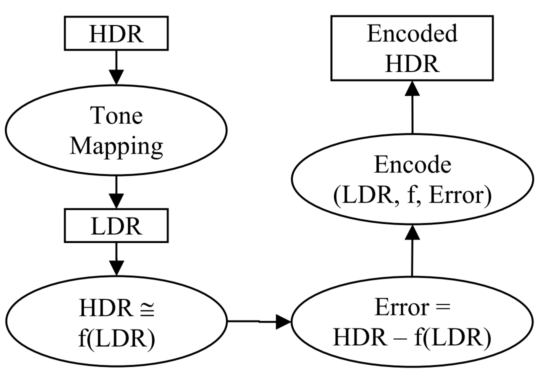

Fig. 1: The encoding scheme

2.1. Tone Mapping

Several TMOs have been presented to convert an HDR image to an LDR one, and any of those can be used with our encoding scheme presented in this paper. However, for best coding quality, a TMO that preserves the relative order of

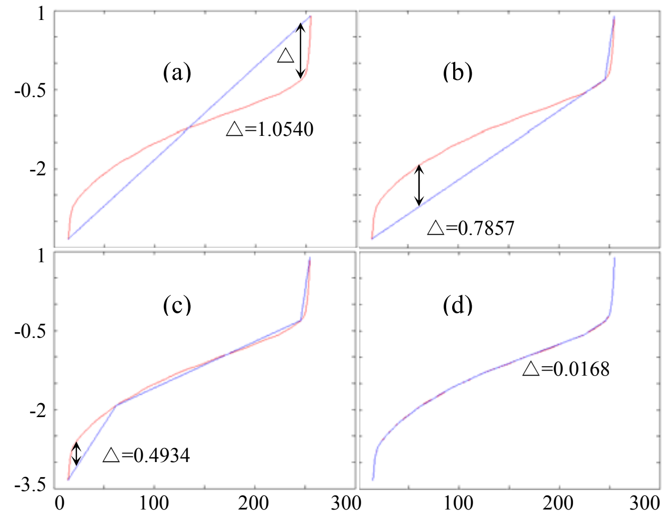

Fig. 2: Red curves show the skeleton of HDR vs. LDR plot for R channel of the Memorial church. Approximation of curve with 2, 3, 4, and 19 key points is shown in (a)-(d). △ represents the maximum error in approximation.

pixel values is required. Preserving this order means that if HDR(i) > HDR(j), then LDR(i) ≥ LDR(j), i.e., if value of i^{th} pixel is larger than value of j^{th} pixel in HDR image, then value of j^{th} pixel in LDR image should not get larger than value of i^{th} pixel after tone-mapping. We use photographic TMO operator [14] in this research, which does not necessarily satisfy this criterion; however, we force it do so by re-arranging certain pixel values, which don't obey this criteria.

2.2. Modeling of HDR as function of LDR

If we plot pixel values of an HDR image as function of its tone-mapped LDR version, following the criteria mentioned above, we would get a monotonically increasing curve, which can be represented by piecewise linear segments. However, it should be noted that tone-mapping from HDR to LDR image is not a one-to-one mapping, i.e., multiple closer HDR values might get mapped to a single LDR value, due to reduction in the dynamic range. This gives a vertical spread in HDR vs. LDR curve. We find a skeleton of the curve by averaging the HDR values corresponding to each distinct LDR value. The resultant curve is then smoothed over the neighborhood and approximation by linear segments. Such a curve for R channel of the HDR image of the Memorial Church is shown in Fig. 2(a). To increase the effective bandwidth of the dark and bright regions, $\log_{10}$ is applied to HDR values in the shown curve.

To approximate the curve by a piecewise linear function, we start with the first and the last point on the curve as the key points, and approximate it by a linear segment joining them. The point on the curve that shows largest deviation from the linear segment(s) joining the key points is selected as the next key point. The process continues iteratively until the maximum error gets below a predefined threshold. This is demonstrated in Fig. 2, where a total of 19 key points can represent the curve very accurately. For this research, we define the maximum error threshold as:

$$\Delta_{\max} = \log_{10}\left(\frac{HDR_{\max}}{HDR_{\min}}\right) \Big/ N,$$

where N is the number of distinct LDR values. For the example shown in Fig. 2, this value turns out to be 0.0175.

2.3. Error Image

As mentioned above, tone mapping is not a one-to-one function, and the key points selected above approximate the skeleton of the mapping curve. The approximating function can reproduce just as many distinct HDR values as there are distinct LDR values. Therefore, the reproduced HDR values have deviations from the actual HDR values, and these residual details are important to reproduce a closer approximation to the actual HDR image. The dynamic range of the residual data is much smaller as compared to the actual HDR image and therefore it can be stored in an 8-bit image with a reasonable accuracy.

Same procedure is repeated for G and B channels of HDR image, and the resultant 8-bit residual images are arranged in a 3-channel 24-bit image.

2.4. Encoding and decoding

The above procedure represents the HDR image by a tone-mapped LDR image, three sets of key points, and an LDR error image. We encode both LDR images in JPEG format. The key points are few in number and therefore no special coding is done for them. For the Memorial Church there are 19, 16 and 18 key points respectively for R, G, and B channels, respectively. The error image and the key points can be stored in one or more wrappers of the tone mapped LDR image. A total of 12 wrappers are provided in JFIF/JPEG coding standards for storing additional information. The applications that don't require or handle HDR data will just see the front layer of the data, i.e., the tone mapped LDR version, and therefore this encoding scheme is backward compatible to existing non-HDR enabled applications.

Decoding of data to get HDR image is quite simple. Piecewise linear approximating functions defined by the key points are used to transform each channel of the tone mapped LDR image, and the error image is added to the result to reproduce the approximation of the HDR image. The resultant values are raised to power 10 to counter the effect of $\log_{10}$ introduced during approximation process.

3. RESULTS

Our method shares the basic idea of two-layer encoding with the existing techniques presented in [10-13]; however, it differs with them in several ways as given below:

1. The residual image in existing methods is the ratio of HDR to LDR image, while in our method it is the ratio of HDR to another HDR image, generated by using our piecewise linear model (note that difference in log domain represents ratio). Therefore, dynamic range of our residual image is relatively smaller, and it can be encoded with a higher accuracy.
2. The existing methods create a residual image for the luminance channel, and take other measures to compensate the residual information of chrominance channels. However, we create residual image for all channels, and therefore better preserve the details.
3. Since we create residual image for all channels, we can work in RGB space. Therefore, unlike existing methods, transformations from RGB to XYZ, and then again back to RGB space are not required. This improves accuracy by eradicating a potential source of rounding error.

We test our proposed scheme using the HDR image of the Memorial Church and four other images, all given by the authors of [15] in the accompanying DVD, in RGBE format [3]. In Table I, we give a comparison of the size of the actual images, our encoded images, and the image encoded in JPEG-HDR [12] format. The values calculated for JPEG-HDR are obtained by using the software provided by the authors. In Table II, we compare our encoding scheme to JPEG-HDR based on mean square difference in CIE LAB color space. In Table III, we use Visual Difference Predictor (VDP) metric for HDR images [16] for comparison. This metric mimics the human perception system and estimates the fraction of pixels that would be perceived different for two images. We calculate VDP-HDR metric for probability settings of 75% and 95%, where 95% gives more conservative estimate, that is, the smaller error (see [16] for details).

All the comparisons are given at JPEG quality settings of 100% and 70% for LDR and difference images in our scheme, and the same settings used in the software provided by the authors of JPEG-HDR. It can be noted that our proposed format significantly reduces the size, as compared to JPEG-HDR at 100% quality setting, and also performs better in terms of CIE delta and VDP score. At quality setting of 70%, the size of our proposed format is further reduced significantly, but the performance is still better, even if compared to the JPEG-HDR at the maximum quality setting of 100%.

4. CONCLUSIONS

An encoding scheme for HDR images is presented. We approximate the HDR image as a piecewise linear model of a tone-mapped LDR image. The error in approximation, the approximating model, and the tone-mapped image can reproduce the actual HDR image very accurately. Experimental results obtained from a variety of HDR images show that our proposed encoding scheme performs better than state of the art existing technique.

5. REFERENCES

[1] P. Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography," *SIGGRAPH*, July 1998.

[2] A. Yoshida, V. Blanz, K. Myszkowski, H. P. Seidel, "Perceptual Evaluation of Tone Mapping Operators with Real-World Scenes," *Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging*, 2005.

[3] G. W. Larson, "Real Pixels" in "Graphics Gems II," ed. James Arvo, *Academic Press* 1991.

[4] G. W. Larson, "LogLuv encoding for full-gamut, high-dynamic range images," *Journal of Graphics Tools,* Vol. 3, No. 1, pp. 15-31, 1998.

[5] R. Bogart, F. Kainz, D. Hess, "The OpenEXR File Format," *SIGGRAPH* 2003.

[6] http://www.anyhere.com/gward/hdrenc/Encodings.pdf

[7] David S. Taubman and Michael W. Marcellin, "JPEG 2000: Image Compression Fundamentals, Standards and Practice," *Kluwer International Series in Engineering and Computer Science*, 2001.

[8] Ruifeng Xu, S. N. Pattanaik, 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.

[9] R. Mantiuk, G. Krawczyk, K. Myszkowski and H. P. Siedel, "Perception-Motivated High-Dynamic Range Video Encoding," *ACM Trans. on Graphics*, Vol. 23, No. 3, pp.773-741, 2004.

[10] K. E. Spaulding, R. L. Joshi, 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*.

[11] K. E. Spaulding, G. J. Woolfe, R. L. Joshi, "Using a Residual Image to Extend the Color Gamut and Dynamic Range of an sRGB Image," *The PICS Conference*, pp. 307-314, 2003.

[12] G. Ward, and M. Simmons, "JPEG-HDR: A Backwards-Compatible, High Dynamic Range Extension to JPEG," *Proceedings of the Thirteenth Color Imaging Conference*, Nov. 2005.

[13] R. Mantiuk, A. Efremov, K. Myszkowski, H. P. Seidel, "Backward Compatible High Dynamic Range MPEG Video Compression," *SIGGRAPH 2006.*

[14] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic Tone Reproduction for Digital Images," *ACM Tran. Graphics*, Vol. 21, No. 3, pp. 267-276, 2002.

Table I: Comparison in terms of the size of the encoded image

HDR Image (Size in KB)	Our Encoding (KB)		JPEG-HDR [12] (KB)	
	Quality 100	Quality 70	Quality 100	Quality 70
Memorial (1313)	230.5	82.0	468.9	70.1
BigTree (2322)	504.2	156.99	830.8	169.4
SpanishCanal2 (3045)	793.8	146.34	679.3	106.1
EMPStair (3040)	657.4	109.3	614.9	84.5
Sfmoma1 (2996)	635.2	88.34	565.5	70.0

Table II: Deviation from the original HDR based on CIE Delta

HDR Image	Our Encoding		JPEG-HDR [12]	
	Quality 100	Quality 70	Quality 100	Quality 70
Memorial	0.5295	0.6497	1.6872	3.0461
BigTree	0.3540	0.4484	2.3967	4.0735
SpanishCanal2	0.3176	0.4077	1.4295	2.1215
EMPStair	0.1748	0.2123	1.0558	1.4452
Sfmoma1	0.2531	0.3167	1.1143	1.4865

Table III: Deviation from the original HDR image, based on VDP measure applied to the luminance channel. The score in two rows corresponds to 75% and 95% settings of probability.

HDR Image	Our Encoding		JPEG-HDR [12]	
	Quality 100	Quality 70	Quality 100	Quality 70
Memorial	0 0	0.0089 0.0041	0.0226 0.0145	0.1317 0.0450
BigTree	0 0	0 0	0 0	0 0
SpanishCanal2	0 0	0 0	0 0	0 0
EMPStair	0 0	0 0	0 0	0.1479 0.0874
Sfmoma1	0 0	0 0	0 0	0.0029 0.0009

[15] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, "High Dynamic Range Imaging' Acquisition, Display, and Image-based Lighting," *Morgan Kaufmann Publishers*, 2005.

[16] R. Mantiuk, K. Myszkowski, H.P. Seidel, "Visual Difference Pridictor for High Dynamic Range Images," *Proc. IEEE Int. Conf. Systems, Man and Cybernetics*, 2004, pp. 2763-2769.