RGBE vs Modified TIFF for Encoding High Dynamic Range Images

Jakkarin Singnoo, Graham D. Finlayson, School of Computing Sciences, University of East Anglia, Norwich, UK

Abstract

High Dynamic Range (HDR) imaging has become more widespread in consumer imaging in the past few years, due to the emergence of methods for the recovering HDR radiance maps from multiple photographs [10]. In the domain of HDR encoding, the RGBE radiance format (.hdr) is one of the most widely used. However, conventional image editing applications do not always support this encoding and those that do take considerable time to read or write HDR images (compared with more conventional formats) and this hinders workflow productivity. In this paper we propose a simple, fast, and practical framework to extend the conventional 12 and 16-bit/channel integer TIFF gamma-encoded image format for storing such a wide dynamic range. We consider the potential of our framework for the tone-mapping application both by measuring the ΔE S-CIELAB color difference between original and encoded image, and by conducting a psychophysical experiment to evaluate the perceptual image quality of the proposed framework and compare it with an RGBE radiance encoding. The preliminary results show that our encoding frameworks work well for all images of a 65 image dataset, and give equivalent results compared to RGBE radiance formats, while both consuming much less computational cost and removing the need for a separate image coding format. The results suggest that our method, used in the normal tone mapping workflow, is a good candidate for HDR encoding and could easily be integrated with the existing TIFF image library.

Introduction

The conventional 8-bit/channel RGB gamma-encoding format cannot encode HDR data, and although standard IEEE 32-bit float format [1] is an ideal representation, it requires too much storage space (96-bit/pixel). The RGBE radiance format is the most commonly used method to encode this range of data. This format appends an 8-bits exponent channel to represent a common exponent for three 8-bit RGB mantissas at each pixel. This results in a 32-bits/pixel encoding [3]. RGBE covers a dynamic range of about 76 orders of magnitude, the relative error of this format is 1%, making this format suitable for most images. However, the format necessitates a large amount of de/encoding time, because it applies the exponent to the mantissas of each individual pixel. Further, many image editing applications do not support this encoding.

Here we investigate the possibility of using a conventional gamma-encoding format for storing HDR images. With this in mind, we first look at the existing gamma-encoded formats. We begin with scRGB (an open standard first developed by Microsoft and Hewlett-Packard). This encoding is broken into two parts. The first part is a 16-bit/channel linear RGB encoding (scRGB). The second part is a 12-bit/channel either RGB or YCC non-linear encoding (scRGB-nl) using a standard 2.2 gamma with a linear subsection near zero. This format can be thought of as the extension to the existing sRGB color space, since it uses the same color primaries and white/black points as the sRGB color space [12] but allows negative values and values above 1.0, offers a larger gamut and larger dynamic range. However, in practice the encoding is more like a Mid Dynamic Range (MDR) format, since it can cover only about 3.5 and 3.2 orders of magnitude for scRGB and scRGB-nl respectively [5]. A description of this encoding can be found in [6]. The most dynamic range available for gamma encoding, as far as we know, is of a 16-bit/channel RGB linear encoding in conventional TIFF images, which holds about 5 orders of magnitude. This simple tiff image format is widely used in photographic workflows even though this encoding requires more storage: 48-bit/pixel.

Since, HDR RGB data can span up to ten orders of magnitude, none of the schemes mentioned above can cover the necessary range. Thus, directly encoding HDR images using a conventional format does not work and will end up with losing detail outside the affordable dynamic range (both in highlight and shadow). In addition, when dealing with encoding, the quantization process also needs to be considered. Inevitably, this process will introduce errors. In term of encoding, this error can be thought as a relative step-size of the neighboring values. Thus, the *relative error* at a specified value x_i is

$$\mu(x_i) = \frac{|x_i - x_{i-1}|}{|x_i|}, \quad where \quad x_i \neq 0. \tag{1}$$

What we need is an encoding with a constant, or almost constant relative error. Unfortunately, the relative quantization error of gamma encoding continues to increase as the value gets smaller. This is shown in figure 1, there the error increases dramatically after about a magnitude of two. This is not a problem when dealing with conventional images, since the images have only two orders of magnitude. In contrast, when dealing with HDR images, we have to make sure to keep these errors as low as possible. In other words, below the threshold that humans can detect. If the error exceeds this threshold, quantization artefacts will become noticeable in the reproduction. And since, the input and/or output ranges of HDR images are arbitrary, if we directly apply a gamma to those data that have value close to zero, we will end up by introducing visible banding artefacts to our color representation in the dark regions.

In summary, the two main obstacles to using gamma-encoding for HDR imaging are the affordable dynamic range, and the relative quantization error that are introduced.

In this paper we propose a gamma-based encoding scheme that tackles both of these problems. In the next section, we describe this method and provide evidence for dynamic range increases and reduced quantization error. We then evaluate the proposed method using both computational image quality metrics and psychophysical experiments. In the last section, we conclude the work and outline future directions.

The Proposed Encoding Framework

Our proposed encoding framework is based on conventional gamma-encoded 12 and 16-bit RGB integer TIFF format which

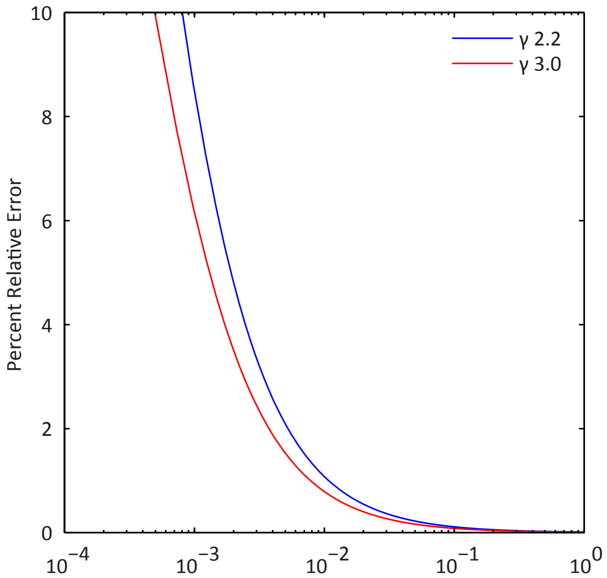

Figure 1. *Percent relative error of different gamma parameters plotted against log10 quantization steps. X-axis can be thought as the dynamic range of the image.*

are already supported by various image editing applications. Our format consists of three R, G, and B color components, 12 or 16-bits are required for each component (36 and 48-bits/pixel respectively).

To encode, the format first scales the original R, G, and B values to fit the range of [0 1] by dividing with the maximum value of all the three components and stores this value as a scaling factor. Once the input is normalized, a gamma (γ) is applied to these normalized values as follows.

$$R' = R^{1/\gamma}, G' = G^{1/\gamma}, B' = B^{1/\gamma} \quad (2)$$

Finally, in order to store the data into an integer fashion, we re-quantize these normalized values into 4096 (2^{12}) and 65536 (2^{16}) steps for the 12 and 16-bit frameworks respectively. In this way, the number of quantization steps is re-adjusted to fit the available quantization steps.

To decode, the format first apply the same γ to R', G', and B'.

$$R'' = R'^{\gamma}, G'' = G'^{\gamma}, B'' = B'^{\gamma} \quad (3)$$

Where, R'', G'', and B'' are then multiplied by the scaling factor in order to recover the original R, G, and B values.

By introducing scaling factor and re-quantization the data, we can eliminate this uncertainty of the input, minimizing the quantization error, conquer the limited dynamic range of data that the format can cope with, and allows backwards-compatibility with various existing image editing applications that support standard 12 and 16-bit TIFF format.

The proposed solution is lossy, in that it does not encode all the image data, but the goal is to produce a visually pleasing image. Thus we can take advantage of the filtering properties of human vision to mask the loss of data (as with JPEG and other lossy encoding schemes for LDR images).

For γ parameter, we base it on the standard γ of 2.2. However, in the extreme cases, like the image shown in figure 2, our 12-bit framework (top right image) introduces a visible banding artifact in the dark flat achromatic region (black paint of the ship) of the tone-mapped image. We found that if we want to encode this image without introducing any banding artifacts, yet with this γ, we have to use minimum of 14 bits (see bottom left image of figure 2) to encode this image.

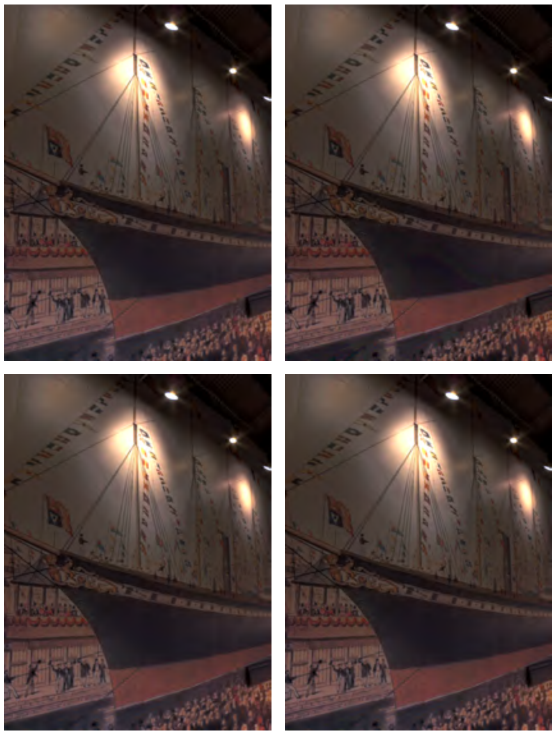

Figure 2. *Effect of bit-depth and γ on the tone-mapped image. Top left, the iCAM06 tone-mapped version of the original (32-bit float point) SS Great Britain image. Top right, the same image passed through our 12-bit Gamma encoding framework (γ of 2.2), which has introduced a banding artifact in the dark achromatic flat area. Bottom left, 14-bit framework (γ of 2.2). Bottom right, 12-bit framework with the γ of 3.0, as can be seen this image does not have any banding artifacts compare to the same bit-depth with the γ of 2.2 (Top left image).*

However, since we can vary the γ, the bit-depth becomes less important. We found that, as we increase the γ, the visible artifacts will gradually decrease, allowing us to use the same bit depth of 12 without introducing any visible artifacts. This phenomenon can be described by the relative quantization error as described in the previous section. Using larger γ reduce this error. As shown in the figure 1, the γ of 3.0 reduces the error. The greater value of γ that we use the lower the relative error we obtain. Although this work in principle, in practice, we might end up losing more data in the bright areas, this could be a problem related to other applications such as exposure compensation (which is outside the scope of this paper).

To examine these issues empirically, we analyzed the outputs of our method using a color difference metric, as will be discussed in the next section, and found that a γ of 3.0 is the optimum value, which, while eliminating any visible artifacts, also preserves most the information in the image. Coincidently, this value of the γ is roughly equal to the non-linear response of the human visual system to intensity [2]. This can be demonstrated by plotting the perceived brightness against the quantization steps as shown in figure 3. The most desirable curve is

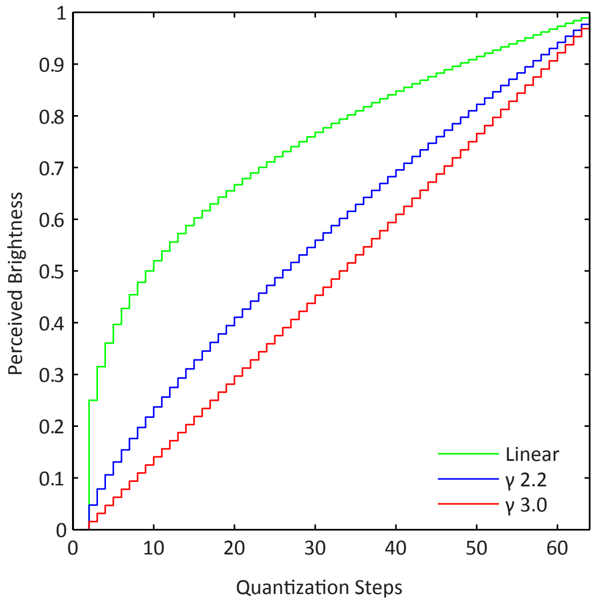

Figure 3. The perceived brightness of the quantization steps of different gamma (γ of 1, 2.2, and 3). In order to emphasize the visible steps, the number of steps in this quantization is 64 (2^6).

supposed to be linear. As shown in the figure, although γ of 2.2 gives a more even distribution of quantization steps than a linear quantization, the behavior in dark regions is still not ideal (we still see larger steps in darker regions than we do in the brighter regions). As one might guess, this can lead to visible banding artifact in the dark regions. When using a γ of 3.0, we see a more even distribution of quantization steps through the whole steps.

Testing the encoding frameworks

To evaluate the image quality of the proposed encoding frameworks, we compare the tone-mapped images resulting from our encodings with the RGBE radiance encoding to the original floating-point TIFF format. Two experiments both objective and subjective approaches have been performed.

Objective Image Quality Measurement

For the objective part, our work here is related to the work of Ward [5]. In that work, he uses the CIELAB 1994 ΔE color difference metric in order to generate encoding quality curves for different encodings of the original HDR data. However, here we are interested in the appearance of the tone-mapped versions of the HDR images. Thus, we instead investigate the color difference of the tone-mapped image after different encoding schemes have been applied. Since, they are conventional 8-bit images, we can measure the color difference directly. The color difference metric that we chose is S-CIELAB [14]. Since the ΔE S-CIELAB measures perceptual color differences. In other words, it measures how similar the reproduction to the original when viewed by an observer, and takes into account simple spatial characteristics of the human visual system.

The tone-mapped images were generated by applying two widely available tone-mapping operators-iCAM06 [7], and Meylan [8]. Figure 4 shows the 65 test-image dataset. Figure 5 shows an average percentile of pixels above a particular ΔE S-CIELAB value of different encodings for all of the test images obtained by weighing each of the test images equally. To interpret this plot, a steeper slope curve that reaches a small percentile at the smaller

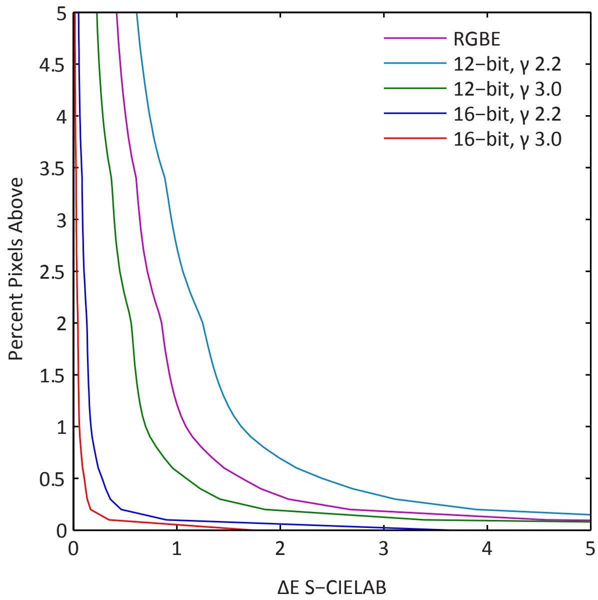

Figure 5. Average percent pixel at a particular ΔE S-CIELAB of different encodings for the entire test images.

ΔE would be the more ideal encoding.

From the figure, it is clear that, in our 16-bit framework both a γ of 2.2 and 3.0 have been out performed by the RGBE radiance format. This is not surprising at all, since the format uses 48-bit/pixel. In comparison, for our 12-bit framework, γ of 3 gives a slightly better result than RGBE, while γ of 2.2 gives the poorest result.

Table 1 summarizes the results, we report two percentile quantities taken from the plots in figure 5 for the percentage of pixels above ΔE values of both 2 and 5. These values correspond to the levels indicating a noticeable color difference to observers in ideal viewing conditions and noticeable in side-by-side images, respectively. However, in practice, one might notice a difference between a pair of images if there are more than 2% of the pixels that have a ΔE greater than 5. Since, there are only 0.61% of pixels of the worst encoding (12-bit, $\gamma = 2.2$) that have ΔE more than 5, which implies that on average both RGBE and our encoding frameworks generate perceptually identical tone-mapped images to the original floating point format.

We note that, all images obtained from our framework appear identical to the original as RGBE format does. The only exception is a SS Great Britain image resulting from the 12-bit γ of 2.2 framework, which was already shown in figure 2.

Table 1. The summary of percent of pixels above ΔE S-CIELAB of 2 and 5 for each encoding.

	RGBE	12-bit (γ 2.2)	12-bit (γ 3)	16-bit (γ 2.2)	16-bit (γ 3)
$\Delta E > 2$	0.85%	1.25%	0.56%	0.13%	0.04%
$\Delta E > 5$	0.42%	0.61%	0.23%	0.05%	0.01%

Subjective Image Quality Measurement

For the subjective image quality measurement part, a psychophysical experiment has been conducted to test the preference of our encoding frameworks in the context of tone-mapping

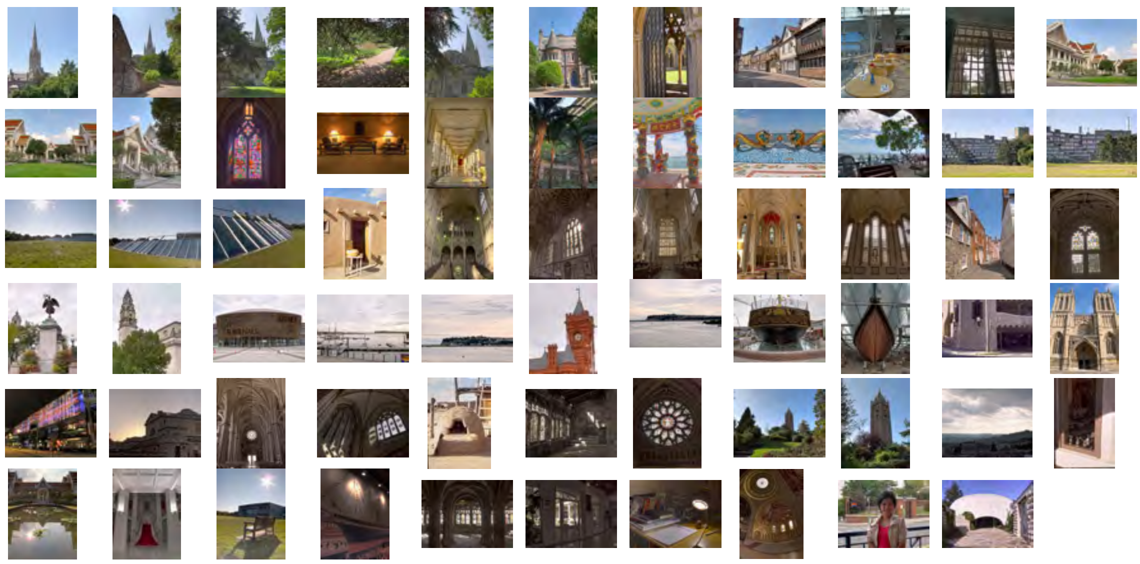

Figure 4. *The image dataset; 65 images in total, both from our own images and the well-known images by [4, 9, 10]. The last row is the ten images used in the psychophysical experiment.*

application. The goal is to generate preference scales in order to evaluate the perceived image quality of the proposed encodings compare with RGBE radiance format.

Experiment Overview

In order to evaluate the perceptual image quality of the encoding, a psychophysical experiment based on a two-alternative forced choice (2AFC) procedure for image pairs (paired comparison) [11] was conducted. Six tone-mapped images were generated by applying the same two tone-mapping operators as used in the previous section to three different HDR encodings, our 12-bit and 16-bit gamma of 2.2, and RGBE encodings. These images were evaluated by twelve observers (six males and six females) with normal color vision, naïve for the goal of the experiment under the same experimental conditions.

Experimental Design

In the experiment, the participants were asked to choose whether an image shown on the left or right was more similar to the reference (rendered 32-bit float format) image in the middle. A color calibrated monitor (HP DreamColor LP2480zx) displayed the three comparison images, each of which had a resolution 640 x 480, at 60.0 Hz. Ten real-world images with a diversity of dynamic ranges and spatial configurations (shown in the bottom row in figure 4) from the image dataset were incorporated in the experiment. With this in mind, there were a total of 120 comparisons (3 encodings, 2 tone-mapping operators, and 10 images with 2 repetitions of each image) in the experiment. The whole procedure for one participant took approximately 15 to 20 minutes. Preference scores are generated using Thurstone's Law of Comparative Judgment Case V [13].

Experimental Results

Figure 6 shows average preference scores for 10 scenes. Figure 7 shows the scores obtained for individual test images. Figure 8 sumamrises the results for each encoding; the confidence interval (error bar) for 15 observers is 0.3578. The higher the preference score for an encoding, the more often it was chosen by the observer. The results show that our 12-bit framework gives the lowest scores while our 16-bit framework has slightly higher score than the radiance RGBE encoding in total. The results also show that both iCAM and Meylan algorithms give almost the same score; this indicates the consistency of our encoding across tone-mapping algorithms. Only the 12-bit framework of SS Great Britain contains artifacts, and has the lowest scores among all images.

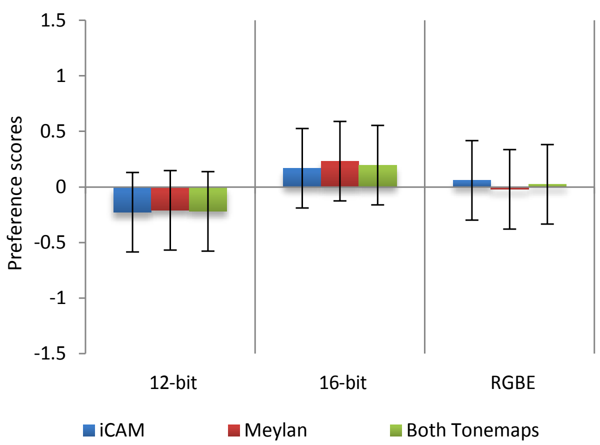

Figure 6. *Overall preference scores for encodings over 10 images (The encodings are labeled as our 12-bit Gamma TIFF encoding (12-bit), our 16-bit Gamma TIFF framework (16-bit) and RGBE Radiance format.*

Storage Size

Here, we compare our encodings with RGBE encoding with/without ZIP compression on our image dataset. The results are shown in figure 9. Compared with the original 32-bit float point format, RGBE format takes 28.75% of the original file size. Our 12 and 16-bit encodings take 37.5% and 49.92% respectively. For ZIP compression, our encodings get the benefit of integer value when dealing with compression, as can be seen both of our encodings dramatically reduce the file size compared to RGBE (RGBE is floating point). From the results, it is clear that our 12-bit encoding requires less storage space than RGBE format (22.77% compare to 24.11%), while our 16-bit encoding still requires the most storage space (30.32%). This indicates that

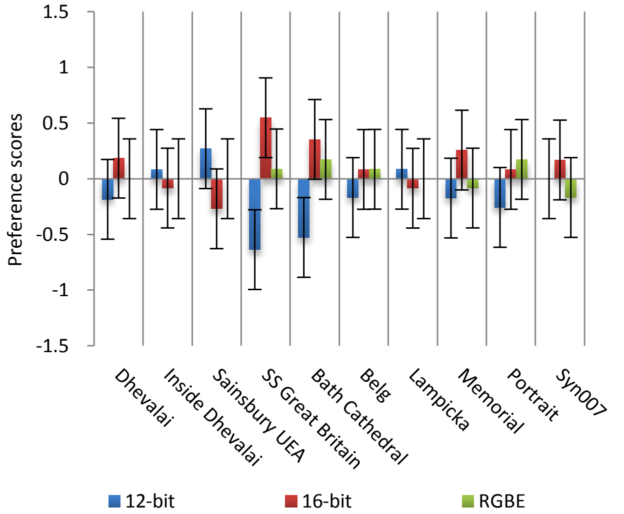

Figure 7. *Preference scores for 10 test images by image.*

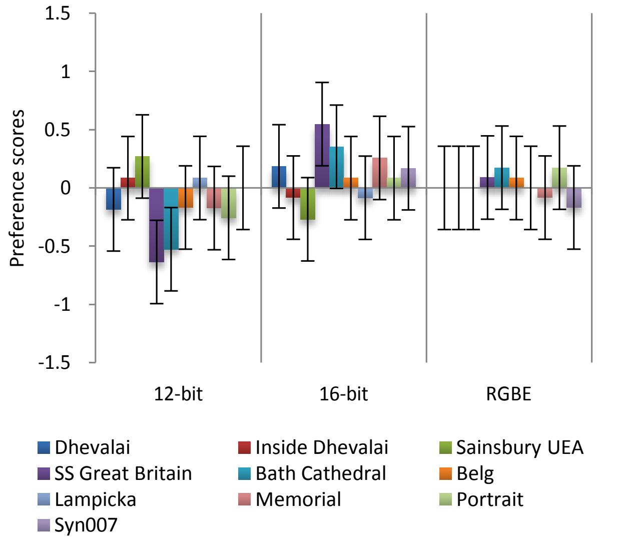

Figure 8. *Preference scores for 10 test images by encoding.*

our 12-bit framework in practice, consumes less storage space than RGBE on average, irrespective of the actual bit-depth required by the formats.

Conclusions

Our encoding frameworks can bridge the gap between conventional and HDR image editing pipelines, it's simple, fast, practical, and can be backward-compatible with any image editing applications that support 12 or 16-bit tiff encodings.

Thorough evaluations, both color difference measurement and image preference through psychophysical experiment for a set of HDR encodings have been conducted in this preliminary study.

The results show that our encoding frameworks both 12 and 16-bit, perform well in the context of a tone-mapping application, although for the 12-bit framework there is a visible banding artifact in a dark achromatic region of one image from the 65 image dataset. However, this artifact can be eliminated by the varying the γ parameter. We found that the γ of 3.0 completely eliminates this unwanted artifact on our test image.

We realize that there is an application to shifting the exposure of the image, for example, when we want to convert a camera RAW format to an 8-bit JPEG format. Especially, when we want to boost up the dark details, this might amplify the relative error in the dark regions, thus causing more visible artifacts. Indeed, for extreme manipulations of HDR images (of the kind not offered in tone mappers) then the RGBE radiance image format is likely to still be a preferred encoding standard. But, all experiments indicate that for the practical HDR workflow used by the majority of HDR enthusiasts our simple TIFF encoding would suffice (offering advantages both in simplicity, storage and computational processing).

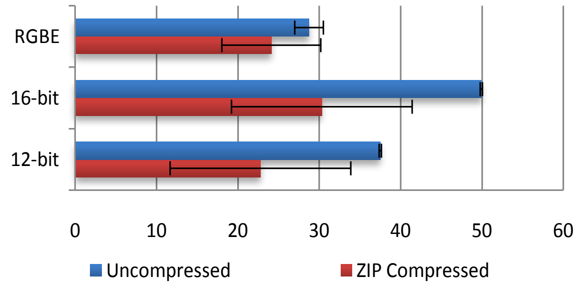

Figure 9. *The average file size (size efficiency) for each test encodings in percentage compared with the original 32-bit TIFF float format.*

Acknowledgments

The author would like to thank Dr.David Connah for his valuable comments and helpful advice on earlier drafts of this paper. Additionally, the author wishes to acknowledge all of the experiment's participants.

References

[1] ANSI/IEEE Standard 754-1985: IEEE Standard for Binary Floating-Point Arithmetic, IEEE Service Center, Piscataway, N.J., 1985.

[2] Bodmann H-W, Haubner P, Marsden AM: A unified relationship between brightness and luminance, CIE Proc. 19th Session (Kyoto) 99-102, 1979, CIE 50-1979.

[3] Erik Reinhard, Greg Ward, Sumanta Pattanaik, Paul Debevec, High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics), Morgan Kaufmann Publishers Inc., San Francisco, CA, 2005

[4] Fattal, R., Lischinski, D., and Werman, M. 2002. Gradient domain high dynamic range compression. In Proceedings of the 29th Annual Conference on Computer Graphics and interactive Techniques (San Antonio, Texas, July 23 - 26, 2002). SIGGRAPH '02. ACM, New York, NY, 249-256. DOI= http://doi.acm.org/10.1145/566570.566573

[5] Greg Ward, High Dynamic Range Image Encodings. Anyhere Software. http://www.anyhere.com/gward/hdrenc/Encodings.pdf.

[6] IEC. 2003. 61966-2-2. Extended RGB colour space - scRGB, Multimedia systems and equipment - Colour measurement and management - Part 2-2: Colour management.

[7] Jiangtao Kuang , Garrett M. Johnson , Mark D. Fairchild, iCAM06: A refined image appearance model for HDR image rendering, Journal of Visual Communication and Image Representation, v.18 n.5, p.406-414, October, 2007

[8] L. Meylan and S. Süsstrunk, High dynamic range image rendering using a Retinex-based Adaptive filter, IEEE Transactions on Image Processing, Vol. 15, Nr. 9, pp. 2820-2830, 2006.

[9] Martin Čadík, Michael Wimmer, Laszlo Neumann, Alessandro Artusi, Technical Section: Evaluation of HDR tone mapping methods using essential perceptual attributes, Computers and Graphics, v.32 n.3, p.330-349, June, 2008.

http://www.cgg.cvut.cz/members/cadikm/tmo/
[10] Paul E. Debevec and Jitendra Malik. Recovering High Dynamic Range Radiance Maps from Photographs. In SIGGRAPH 97, August 1997. http://www.debevec.org/Research/HDR/
[11] P. G. Engeldrum, Psychometric Scaling, A Toolkit for Imaging Systems Development. Imcotek Press, Winchester MA, USA, 2000.
[12] Stokes, M., Anderson, M., Chandrasekar, S., and Motta, R. A 1996. Standard Default Color Space for the Internet. www.w3.org/Graphics/Color/sRGB.
[13] THURSTONE, L. L. 1927. A law of comparative judgment. Psychological Review 34, 273.286.
[14] X. M. Zhang and B. A. Wandell, A spatial extension to CIELAB for digital color image reproduction, in: Proceedings of the SID Symposiums, vol. 27, 1996, pg. 731-734.

Author Biography

Jakkarin Singnoo received his BSc in Imaging and Printing Technology from Chulalongkorn University, Thailand (2004) and his MSc in Advanced Computer Graphics and Applications from University of East Anglia, UK (2006). He is now a PhD student in Colour Group, School of Computing Sciences at University of East Anglia. His work has focused on HDR imaging including Tone Mapping Operators and Psychophysics.

