JPEG Compatible Coding of High Dynamic Range Imagery using Tone Mapping Operators

Min Chen[1], Guoping Qiu[1], Zhibo Chen[2] and Charles Wang[2]
[1]School of Computer Science, The University of Nottingham, UK
[2]Thomson Broadband R&D (Beijing) Co., Ltd, China

Abstract. In this paper, we introduce a new method for compressing high dynamic range (HDR) imageries. Our method exploits the fact that tone mapping is a necessary operation for displaying HDR images on low dynamic range (LDR) reproduction media and seamlessly integrates tone mapping with a well-established image compression standard to produce a HDR image compression technique that is backward compatible with established standards. We present a method to handle color information so that HDR image can be gracefully reconstructed from the tone mapped LDR image and extra information stored with it. We present experimental results to demonstrate that the proposed technique works well and we also discuss future directions of this new research area.

Index Terms—High dynamic range imaging, tone mapping, compression, coding, JPEG

1. INTRODUCTION

Today's digital photography technology is very advanced. However, to achieve a level that is comparable to the amazing capability of the human visual system (HVS), technology still has a long way to go. For example, the real world scenes we experience in our daily life often have a very wide range of radiance values and HVS is capable of simultaneously perceiving scenes with dynamic ranges of over five orders of magnitude and can gradually adapt to scenes with dynamic ranges of over nine orders of magnitude. Today's predominant imaging devices all have limited dynamic range. For example, digital cameras store digital color photos in 24 bit/pixel thus only capturing a useful dynamic range of 2 orders of magnitude. Many image reproduction media such as video monitor and printing paper also have a dynamic range not exceeding 2 orders of magnitude.

High dynamic range (HDR) imaging, where the image files record the true radiance dynamic range and color gamut of the scene holds tremendous potential in advancing digital photography technology. In the past 15 years or so, researchers, mainly in the computer graphics community, have been developing HDR imaging technologies, ranging from high dynamic range radiance map capturing [1, 2] and efficient storage of 96 bit/pixel radiance maps [14, 15] to tone mapping of HDR images for display in low dynamic range reproduction media [3 - 12]. Despite this effort, many technological challenges in HDR imaging remain. For example, to display HDR images in conventional CRT or print HDR image on paper, the dynamic range of the image has to be compressed or the HDR image has to be tone mapped. Even though there has been several tone mapping methods in the literature [3 - 12], non so far can universally produce satisfactorily results. Another huge challenge is efficient storage of HDR image. Compared with conventional 24-bit/pixel images, HDR image is 96 bits/pixel and data is stored as 32-bit floating-point numbers instead of 8-bit integer numbers. The data rate using lossless coding, e.g, in OpenEXR [15], will be too high especially when it comes to HDR video. Existing lossy compressing and coding standards such as JPEG and MPEG cannot be directly applied to HDR images/videos. What is needed is lossy compression methods that is suitable for HDR images. Because of the huge file size of HDR image, effectively compressing HDR image/video has become a very important factor in pushing HDR imaging into wider applications such as digital cinema.

Even though high dynamic range display devices are being developed, e.g., [13], traditional low dynamic range reproduction devices and software will be with us for a long time yet. The authors in [14] made an interesting analogy that the transition from low dynamic range imaging to high dynamic range imaging can be compared with the transition from black and white television to color television. It would therefore be most useful to develop technologies that are backward compatible for the storage and reproduction of high dynamic range images. Therefore, HDR images should be encoded in such a way that, when using conventional reproduction media and software, we reproduce the high dynamic range images as ordinary low-dynamic range images, and when using HDR-enabled devices and software, the original HDR image will be recovered and reproduced.

In this paper, we present a novel method that combines HDR image tone mapping operators [8, 10] with established JPEG image compression standard to efficiently code HDR images. Because tone mapping is a necessary step for displaying HDR images in LDR devices, our new method intelligently exploits this operation and gracefully integrates it with a tried and tested technology. To an ordinary (low dynamic range) JPEG-enabled image viewer, a tone-mapped low-dynamic range version of the HDR image will be displayed directly, and to an HDR-enabled system such as [13], high dynamic range data can be recovered from the coded bit stream and display as high dynamic range image.

2. METHOD

Digital photography is in a transitional period in the following sense. On the one hand, high dynamic range imaging devices and software has started to appear and it has been predicted that the future of digital photography is high dynamic range imaging [16]. On the other hand, today's imaging devices, software and standards are all designed for low dynamic range images and these legacies will be with us for a long time yet. Therefore, backward compatible software that can handle high dynamic range imagery and at the same time is compatible with current software and standards will be most useful in this transitional time (which may last anything between 10 to 20 years).

Motivated by the above rationale, we propose a JPEG backward compatible coding method to efficiently store high dynamic range images. Unlike previous methods, e.g., [14], our technique intelligently exploits the fact that to display HDR images in LDR devices, tone mapping is a necessary operation and directly uses the tone mapped image as part of the compressed HDR data. As a result, the compressed bit stream contains a directly displayable tone mapped version of the HDR image. Alongside the tone-mapped image, we store a version of the image containing the detail information necessary for recovering the HDR data and wrap it in the JPEG's application markers [17]. It has to be pointed out that not all tone-mapping operators will be suitable for this application. The basic requirement is the tone mapping function must be invertible which maybe local [8] or global [10] operators.

Since HDR tone mapping is a multi-to-one mapping, tone mapped LDR image has not enough information to reconstruct its HDR radiance map. In our method, a detail layer of HDR image is stored as metadata to provide more information, especially the information of local details, to reconstruct the HDR radiance map. Fig.1 shows an overview of our HDR encoder. The tone-mapped (24-bit) image is compressed by the JPEG software in the standard way. The detail layer image is coded to 8 bit (0-255) integers by linear quantization and then again compressed by the JPEG software to be stored in the JPEG's application marker field. The encoded HDR file is a composite of two parts. One is the tone-mapped LDR image, which can be displayed by conventional image displayer. The other is quantized detail layer, which contains enough information to reproduce a close facsimile of the original HDR image. Fig. 2 shows the process of HDR image recovery from encoded file. Both tone-mapped image and quantized detail layer can be decoded by the standard JPEG decompression routines. The tone-mapped image and detail layer are converted back to the original format represented by 32-bit float-point numbers and the HDR radiance map can be reconstructed by adding these two recovered parts.

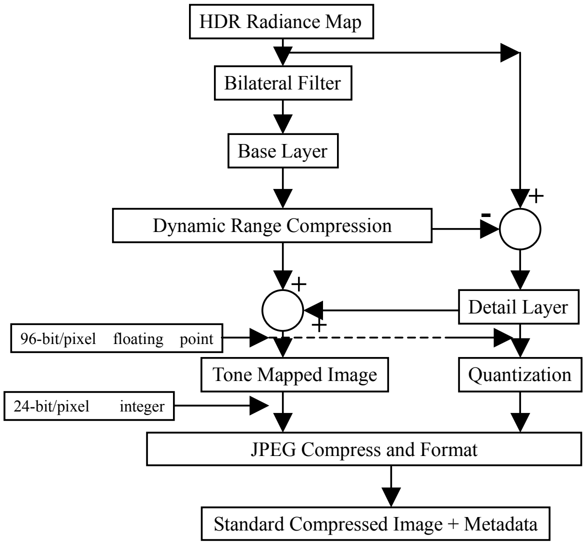

Fig. 1. HDR image encoding pipeline

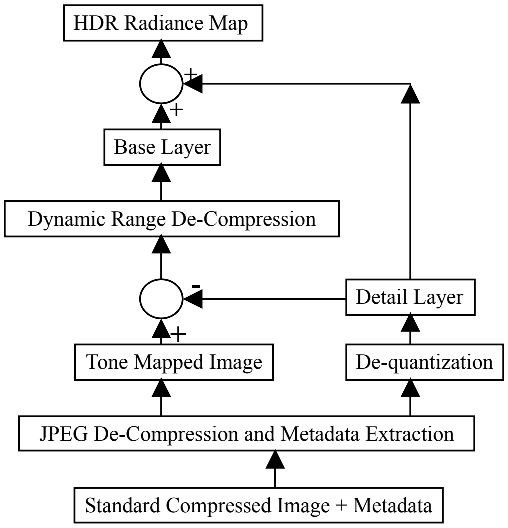

Fig. 2. HDR image decoding pipeline

2.1 Detail layer Quantization and compression

Since the tone-mapped image is encoded by standard JPEG compressor, it is most convenient for us to compress quantized detail layer by standard JPEG as well (although other data compression methods such as vector quantization may also be applied, in this paper we report results that are based on JPEG compression of detail layer).

To compress the detail layer by JPEG, we quantize the detail layer into 8 bit integer numbers (0-255). Since the magnitude of the detail layer is always very small

~(-0.5 – 0.5), we use simple linear scaling to do the quantization, which is very fast and easy to implement.

There are 16 application markers in the JFIF specification (see http://www.ijg.org/) and a single marker holds 64 Kbytes of data regardless of the image dimensions. Therefore, we have to always keep our metadata under 64 Kbytes. In practice, we keep our metadata under 60 Kbytes, and the spared 4 Kbytes are used to store other information such as the high dynamic range compression factor, the tone mapping curve etc. that are necessary for recovering the HDR data. For different sized images, we have to choose different compression quality factors to control the size of the compressed detail layer data such that they can be stored in the application markers. This is a drawback of this approach because the quality of the detail layer is dictated by the header sizes. Using other non-standard coding strategies such as VQ for storing the detail layer will not be restricted by such limitation.

2.2 Handling Color and HDR Reconstruction

In HDR image tone mapping, only the dynamic range of the luminance channel is compressed, whilst color is untouched [8, 10]. Let R_h G_h B_h be the high dynamic range pixel, R_l G_l B_l be the tone mapped low dynamic range pixel, the luminance of the high dynamic range image can be computed as

$$Y_h = 0.299R_h + 0.587G_h + 0.114B_h \qquad (1)$$

Y_h is mapped to low dynamic range luminance $Y_l = \mathrm{TM}(Y_h)$; where the tone mapping function TM is a monotonic mapping function in the case of tone mapping curve based operator [10, 11], or a simple linear scaling function in the case of bilateral filtering based tone mapping operator [8]. To produce a low dynamic range display image, we can use following formula

$$R_l = \left(\frac{R_h}{Y_h}\right)^{\gamma} Y_l \qquad G_l = \left(\frac{G_h}{Y_h}\right)^{\gamma} Y_l \qquad B_l = \left(\frac{B_h}{Y_h}\right)^{\gamma} Y_l \qquad (2)$$

However, we cannot store R_l G_l B_l directly because we cannot recover Y_l from these 3 numbers and therefore cannot recover $Y_h = \mathrm{TM}^{-1}(Y_l)$ and R_h G_h B_h. In order to be able to recover the high dynamic range pixels, we apply YC_bC_r color space to preserve colors during encoding. Compared with sRGB space used by many HDR image rendering techniques, YC_bC_r color space has two advantages, which are very import for us to reconstruct the HDR radiance map. Firstly, in sRGB space, it is very difficult to recover the rendered low dynamic range luminance Y_l since R_h G_h B_h have a non-linear relationship with Y_l caused by tone mapping operation in equation (2). Secondly, in sRGB space, some saturated pixels (larger than 255) are clipped, but in YC_bC_r space, this problem is avoided. We represent the tone-mapped low dynamic range image in YC_bC_r as follows

$$\begin{aligned} Y &= Y_l \\ C_b &= -0.169R_l - 0.331G_l + 0.5B_l \\ C_b &= 0.5R_l - 0.418G_l + 0082B_l \end{aligned} \qquad (3)$$

Note that Y is the tone mapped luminance rather than computed from R_l G_l B_l, which guarantees that the tone-mapped low dynamic range display image has a luminance signal that is identical to the tone mapped luminance Y_l which can be directly used to recover $Y_h = \mathrm{TM}^{-1}(Y_l)$.

To recover the high dynamic range pixel R'_h G'_h B'_h, we first decode the low dynamic range display image as $Y'C'_bC'_r$ and then solve the following system of nonlinear equations

$$\begin{aligned} & 0.299R'_h + 0.587G'_h + 0.114B'_h - Y'_h = 0 \\ & -0.169\left(R'_h / Y'_h\right)^{\gamma} - 0.331\left(G'_h / Y'_h\right)^{\gamma} + 0.5\left(B'_h / Y'_h\right)^{\gamma} - C'_b = 0 \\ & 0.5\left(R'_h / Y'_h\right)^{\gamma} - 0.418\left(R'_h / Y'_h\right)^{\gamma} + 0082\left(R'_h / Y'_h\right)^{\gamma} - C'_r = 0 \end{aligned} \qquad (4)$$

where Y'_h is the recovered high dynamic range luminance (see Fig. 2). For γ = 0.5 (γ controls the color of the tone mapped image and γ = 0.5 works well for most images, we used this value in all our results), we can solve the problem analytically, for any other values of γ the nonlinear system of equations can be solved using many efficient numerical methods.

3. RESULTS

We have tested our encoding method on many HDR images. We use relative mean absolute error (RMAE) and signal to noise ratio (SNR) to objectively measure the compression performances.

$$RMAE = \frac{1}{3N}\sum_{i=1}^{N}\left(\frac{\left|R_h(i) - R'_h(i)\right|}{R_h^{\max} - R_h^{\min}} + \frac{\left|G_h(i) - G'_h(i)\right|}{B_h^{\max} - B_h^{\min}} + \frac{\left|B_h(i) - B'_h(i)\right|}{B_h^{\max} - B_h^{\min}}\right)$$

$$SNR = 10\log\left(\frac{\sum_{i=1}^{N}\left(R_h^2(i) + G_h^2(i) + B_h^2(i)\right)}{\sum_{i=1}^{N}\left(\left(R_h(i) - R'_h(i)\right)^2 + \left(G_h(i) - G'_h(i)\right)^2 + \left(B_h(i) - B'_h(i)\right)^2\right)}\right)$$

We also measure the difference of tone mapped low dynamic range images before and after compression as

$$MSE_{TM} = \frac{1}{3N}\sum_{i=1}^{N}\left(\left(R_l(i) - R'_l(i)\right)^2 + \left(G_l(i) - G'_l(i)\right)^2 + \left(B_l(i) - B'_l(i)\right)^2\right)$$

Table 1 shows the compression performances of our method on 15 HDR images containing a variety of content types. It is seen that the compression ratio and reconstruction accuracy is a trade off as expected. Higher compression ratios will result in higher reconstruction error. Some authors such as [14] use some subjective image quality measures such as visible difference predictor (VDP) to measure the quality. However, objective image quality measurement is not straightforward and quantitative quality metrics are still very difficult to specify. In the absence of high dynamic range reproduction media we shall show some rendered LDR images. Fig. 3 and Fig. 4 show examples of low dynamic range images rendered from the compressed HDR data. Although these performance measures may not be ideal for evaluation, given the current state of the art of the technology, it seems that compressing HDR data using our method does not introducing much visible distortion. However, the full effect can only be assessed when high dynamic range reproduction media are available. These rendered LDR images suggest that the compression method of this paper works well.

4. Discussion and Future Work

In this paper, we have proposed a new method for compressing high dynamic range image data. The new technique takes advantage of the fact that tone mapping is a necessary operation for displaying high dynamic range images in low dynamic range media and directly use the tone mapped low dynamic range image as part of the compressed data and uses a well established image compression standard to compress the low dynamic range displayable image and the detail information necessary for the recovery of high dynamic range data. We have tested one tone-mapping operator [8] for compression and shown that our technique works quite well based on some objective distortion measures and subjective assessment of rendered low dynamic range displays.

The technique is in its early stage of development and there are many issues need further research. Firstly, we believe the general approach, i.e., the exploitation of the tone mapping operation of HDR image for developing compression technology is technically sound and that our preliminary results have demonstrated that this approach does work. The issues need further research include how to handle chromaticity information. In our current implementation, we use the same chromatic signals for both high dynamic and low dynamic range, although experimental results seem to indicate that this approach works reasonably well but we have not evaluated how much error this will introduce and formal analysis of this operation is necessary. Another long-standing question is how to measure the compression performance quantitatively. This problem existed in ordinary image compression and there are new issues in HDR image compression, especially when HDR image can only be displayed as LDR image for evaluation. These and many other questions remain open and deserve further research.

References

[1] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs", In Proceedings of SIGGRAPH'97, pp. 369 – 378

[2] T. Mitsunaga and S. K. Nayar, “High dynamic range imaging: Spatially varying pixel exposures”, Proc. CVPR’2000, vol. 1, pp. 472-479, 2000

[3] S. N. Pattanaik, James A. Ferwerda, Mark D. Fairchild, and Donald P. Greenberg, "A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display", Proceedings of SIGGRAPH'98, pp. 287-298, Orlando, July 1998

[4] J. Tumblin, J. K. Hodgins and B. K Guenter, "Two methods for display of high contrast images", ACM Trans on Graphics, vol. 18, pp. 56 - 94, 1999

[5] J. DiCarlo and B. Wandell, “Rendering high dynamic range images”, Proc. SPIE, vol.3965, pp. 392 – 401, 2001

[6] R. Fattal, D. Lischinski and M. Werman, "Gradient domain high dynamic range compression", In Proceedings of SIGGRAPH 2002, pages 249–256, 2002.

[7] E. Reinhard, M. Stark. P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images", In Proceedings of SIGGRAPH 2002, pages 267–276, 2002.

[8] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high dynamic range image", In Proceedings of SIGGRAPH 2002, pages 257–265, 2002

[9] Michael Ashikhmin. “A tone mapping algorithm for high contrast images”, in Rendering Techniques ’02 (Proceedings of the 13th Eurographics Workshop on Rendering), pages 145–156, June 26–28, 2002.

[10] J Duan, G. Qiu and M. Chen, "Comprehensive Fast Tone Mapping for High Dynamic Range Image Visualization", Pacific Graphics 2005

[11] J. Duan, G. Qiu and G. D. Finlayson, "Learning to display high dynamic range images", CGIV'2004, IS&T's Second European Conference on Color in Graphics, Imaging and Vision, Aachen, Germany, April 5-8, 2004

[12] Y. Li, L. Sharan, E. H. Adelson, “Compressing and companding high dynamic range images with subband architectures”, ACM Transactions on Graphics (TOG), 24(3), Proceedings of SIGGRAPH 2005.

[13] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, A. Vorozcovs, “High dynamic range display systems”, ACM Transactions on Graphics (Siggraph 2004), 23(3): 760-768

[14] G. Ward and M. Simmons, “Subband encoding of high dynamic range imagery”, Proceedings of the 1st Symposium on Applied perception in graphics and visualization, Pages: 83 – 90, 2004

[15] http://www.openexr.com/

[16] Jon Meyer, “The Future of Digital Imaging - High Dynamic Range Photography”, http://www.cybergrain.com/tech/hdr/

[17] http://www.jpeg.org/

file name	TM JPEG Quality	Detail layer quality	LDR MSE$_{TM}$		HDR		Compression Ratio (against RGBE .hdr file)
Size			Rendering Operators		RMAE	SNR	
Dynamic Range			Bilateral [8]	Curve [10]			
memorial	100	100	0.61	3.34	1.18E-04	13.71	4.46
512*768	90		3.15	7.87	8.94E-04	5.92	9.20
4.8	60		31.80	18.47	2.45E-03	1.64	14.44
Belgium	100	80	0.51	0.39	1.19E-03	26.92	5.77
1025*769	90		11.75	3.77	4.81E-03	11.82	13.95
4.1	60		20.53	8.64	8.26E-03	7.34	23.62
Rosette	100	100	8.45	6.03	7.17E-04	16.40	5.30
720*480	90		15.61	4.57	4.22E-03	4.75	9.80
4.4	60		27.81	8.49	6.79E-03	0.32	14.03
AtriumNight	100	80	9.75	3.09	3.69E-04	21.56	10.51
760*1016	90		24.71	17.04	2.05E-03	9.28	13.98
4.1	60		27.23	32.66	3.96E-03	4.42	23.96
Desk2	100	90	24.36	2.85	1.82E-03	22.97	7.61
644*874	90		42.07	14.05	1.70E-02	8.34	10.42
5.2	60		84.53	24.43	2.57E-02	5.93	17.47
Bathroom	100	100	91.04	0.25	5.11E-04	29.68	6.27
346*512	90		110.54	1.22	2.56E-03	18.34	7.38
4.5	60		104.04	3.37	3.96E-03	14.40	9.03
Cathedral	100	80	0.52	1.27	3.07E-03	26.02	8.82
767*1023	90		3.78	13.77	2.59E-02	12.73	12.07
4.1	60		20.35	20.97	3.86E-02	9.53	22.47
MtTamWest	100	80	117.05	0.53	2.74E-03	31.68	10.45
1214*732	90		156.06	6.47	1.10E-02	22.72	13.31
3.4	60		258.77	11.98	1.31E-02	16.55	26.16
Rend01	100	60	0.21	4.57	5.20E-04	23.60	12.34
1024*1024	90		3.05	4.44	4.08E-04	19.64	16.26
3.0	60		24.43	12.00	6.63E-04	15.54	27.43
Rend06	100	60	0.73	0.20	1.73E-04	10.32	12.34
1024*1024	90		1.23	11.77	5.29E-04	9.51	18.16
3.6	60		1.67	9.70	6.73E-04	8.80	31.99
BigFogMap	100	80	1.13	0.05	4.53E-04	2.69	17.68
751*1130	90		4.12	0.96	8.66E-04	0.70	19.74
4.1	60		11.94	3.77	1.13E-03	0.34	27.69
Tree	100	70	8.82	4.97	9.42E-03	24.54	10.25
928*906	90		47.21	32.94	4.62E-02	10.15	13.90
4.1	60		71.53	44.23	6.48E-02	7.25	25.59
Sunrendering	100	80	4.98	0.09	3.78E-05	32.93	15.83
944*400	90		20.66	1.83	1.39E-04	22.50	16.78
4.4	60		44.06	0.90	3.58E-04	17.30	20.57
DyrhamChurch	100	40	0.16	1.00	2.55E-03	29.56	13.86
2048*1536	90		2.35	4.43	1.26E-02	13.46	20.33
2.4	60		7.40	9.80	2.04E-02	9.04	43.54
NapaValley	100	40	0.58	2.54	2.55E-03	25.4	10.81
3025*2129	90		2.86	6.18	1.26E-02	12.7	32.16
3.2	60		19.65	15.87	2.04E-02	6.88	67.79

Table 1: Compression performances of the new JPEG-compatible HDR image compression technique. Note that the compression TM operator was based on bilateral method [8]. The display LDR images are rendered using both local [8] and global [10] operators.

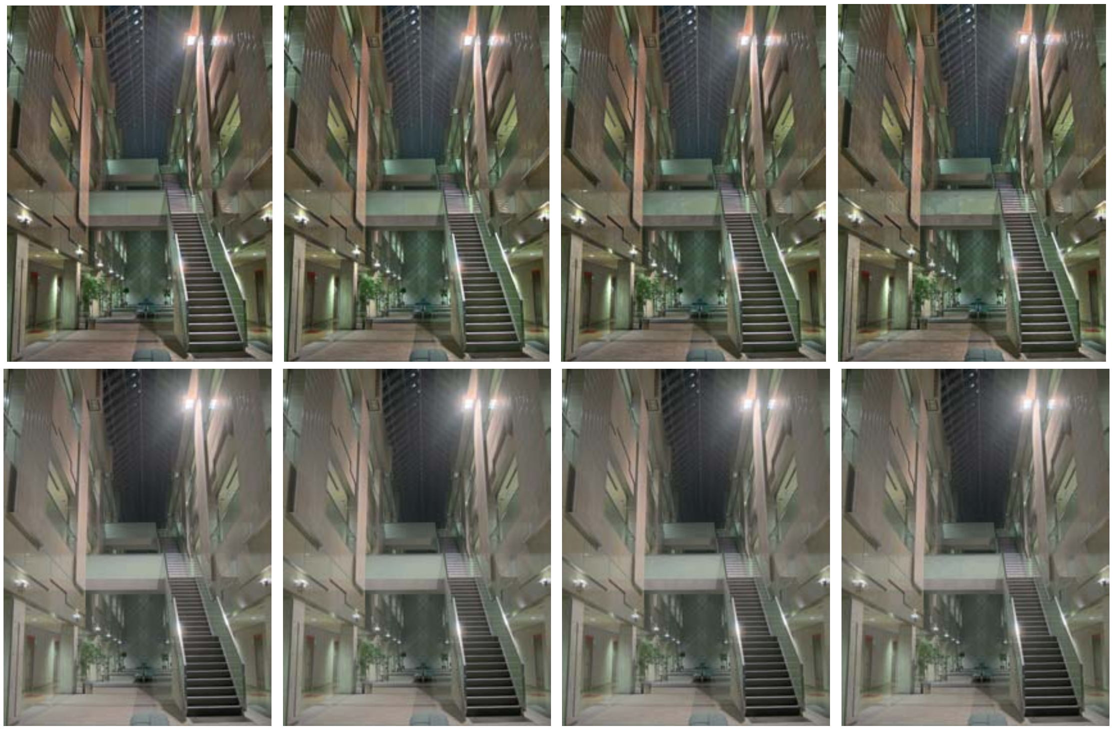

Fig.3: Top row bilateral filtering tone mapping operator [8] used for LDR rendering and compression. From left to right, rendered from non-compressed, 10.5: 1, 14:1 and 24:1 compressed HDR image. Bottom row tone reproduction curve based tone mapping operator [10] used for LDR rendering. From left to right, rendered from non-compressed, 10.5: 1, 14:1 and 24:1 compressed HDR data. Visible differences for the images rendered with the same TM operator are very small. Note different TM operators give different images.

Fig.4: Top row bilateral filtering tone mapping operator [8] used for LDR rendering and compression. From left to right, rendered from non-compressed, 6: 1, 6:1 and 9:1 compressed HDR image. Bottom row tone reproduction curve based tone mapping operator [10] used for LDR rendering. From left to right, rendered from non-compressed, 6: 1, 7:1 and 9:1 compressed HDR data. Again, visible differences for the images rendered with the same TM operator are very small. Note also different TM operators give different images.