Tone Mapping for HDR Image using Optimization – A New Closed Form Solution

Guoping Qiu, Jian Guan, Jian Duan and Min Chen
School of Computer Science and Information Technology, The University of Nottingham

Abstract

This work studies an optimization approach for designing tone reproduction curve (TRC) based tone mapping operators for the display of high dynamic range (HDR) images in low dynamic range (LDR) reproduction media. Previous work has shown that the tone mapping problem can formulated as that of optimizing a two-term cost function where adjusting the relative weightings of the two terms allows users to interactively control the appearance of the output image. However, only heuristic solutions to the tone mapping objective function have been found in past research. The main contribution of this paper is the re-formulation of the objective function to allow the introduction of a closed-form solution to the two-term tone mapping objective function. The new solution has simplified previous heuristic solutions and made this approach mathematically more elegant, computationally faster and practically easier to implement.

1. Introduction

State of the art cameras have enough spatial resolution. Reading through the technical brochures of most digital cameras these days you will see words such as "*x* mega pixels", where *x* could be anything between 4 to 16 or maybe even higher. What is lacking, however, which is also likely to be the next "big thing" in digital photography, is bit-depth, or dynamic range. It is well known that modern cameras use 8bit/pixel format to record the *R*, *G*, and *B* values. This data format allows the camera to record a useful dynamic range (the ratio between the darkest and the brightest pixel values) of about two orders of magnitude. However, the real world scenes we experience every day have far high dynamic range, which can easy exceed 4 - 5 orders of magnitude. Furthermore, human visual system has remarkable ability and can perceive real world scenes with brightness dynamic ranges of over 5 orders of magnitude and can distinguish even higher contrast through adaptation.

High dynamic range (HDR) imaging technologies are designed to produce images that faithfully depict the full visual dynamics of real world scenes. Recent technologies have made it relatively easy to create radiance maps that capture the full dynamic range of real world high contrast scenes [1]. Using the method of [1], a high dynamic range (HDR) radiance map of a scene can be generated by using a sequence of low dynamic range (LDR) images of the same scene taken under different exposure intervals. Apart from the method of [1], other techniques for capturing HDR images have also been developed by various researchers e.g., [2] and [3]. More recently, technology has also been developed to capture high dynamic range videos [7, 8].

The HDR radiance map records the full dynamic range of the scene in numerical format. However, most reproduction devices, such as computer monitors and printers, have a dynamic range of about two orders of magnitude, which is significantly lower than the dynamic range of the radiance map data. In order to reproduce HDR maps in LDR devices, mapping or tone reproduction techniques are used to map HDR values to LDR values. Tone reproduction (mapping) is a crucial step in high dynamic range digital imaging workflow, because it doesn't matter how accurate the HDR maps may be, they have to be accurately and faithfully reproduced in the LDR devices. If the reproduction techniques fail, the whole workflow fails.

In the literature, a number of techniques have been developed for tone reproduction for visualizing high dynamic range images. There are two broad categories of technology, i.e., tone reproduction curve (TRC) based and tone reproduction operator (TRO) based [10].

TRC refers to techniques that manipulate the pixel distributions. A tone reproduction method that attempted to match display brightness with real world sensations was introduced in [11]. Reference [12] presented a tone mapping method that modeled some aspects of human visual system. More recently, [9] presented an adaptive logarithmic curve method for compressing high dynamic range images. Perhaps the most comprehensive technique in this category is still that of [14], which introduced a quite sophisticated tone reproduction curve technique that incorporated models of human contrast sensitivity, glare, spatial acuity and color sensitivity.

TRO techniques involve the spatial manipulation of local neighboring pixel values, often at multiple scales. The scientific principle of this type of technique is based on the image formation model: $I(x, y) = L(x, y)\ R(x, y)$, which states that image intensity function $I(x, y)$ is the product of the illuminant function $L(x, y)$ and the scene reflectance function $R(x, y)$. Because real world reflectance $R(x, y)$ has low dynamic range (normally not exceeding 100:1), reducing the dynamic range of $I(x, y)$ can be achieved by reducing the dynamic range of $L(x, y)$ if one could separate $L(x, y)$ from $R(x, y)$. Methods based on this principle include [15], [16] and [17]. They mainly differ in the way in which they attempted to separate the illuminant component from the reflectance component. All TRO based methods can be regarded as related to the Retinex theory [21]. A direct use of the Retinex theory for high dynamic range compression has been presented in [24].

Recent development has also attempted to incorporate traditional photographic technology in the digital domain for the reproduction of high dynamic range images [18]. An impressive technology for high dynamic range compression is that of [19]. Based on the observation that human visual system is only sensitive to relative local contrast, the authors developed a multiresolution gradient domain technique. This is also a TRO type technique and the authors reported very good results that were free from halo effects.

TRO based methods involve multiresolution spatial processing and are therefore computationally very expensive. Because TRO methods could reverse local contrast, they sometimes can cause "halo" effects in the reproductions. Another difficulty of these techniques is that there are too many parameters the users have to set, which makes them quite difficult to use. TRC based methods do not involve spatial processing, they are therefore computationally very simple. This is useful in real time applications such as high dynamic range video [7, 8]. TRC techniques also preserve the intensity orders of the original scenes and avoid artifacts such as halo that is often associated with TRO based methods. One of the weaknesses of TRC approaches as compared with TRO methods is that it may cause noticeable lost of details in some images. Because of their respective merits and drawbacks of both technologies in terms of computational complexity, simple implementation and easy to use, both types of technologies are likely to co-exist for tone mapping in HDR imaging for the foreseeable future.

In this paper, we present a new solution to the optimization based TRC operators [25 – 27]. The solutions to the optimization problem in these previous works are all based on iterative heuristic either by training a neural network or based on some sort of intuitions. We will re-formulate the tone mapping objective function and introduce a closed form solution which will make this type method mathematically more elegant, computationally faster and practically easier to implement.

2. Adaptively Setting Scene Brightness

In mapping high dynamic range images to low dynamic range ones for reproduction purpose, there are two basic requirements. The first is that the overall brightness should be correct and the other is that there should be enough detail in the reproduced images. Both requirements are subjective and scene dependant. As a first step, we use following function to map the luminance of the high dynamic range image ***I*** to display luminance ***D***:

$$D(I) = (D_{max} - D_{min}) * \frac{log(I+\tau) - log(I_{min}+\tau)}{log(I_{max}+\tau) - log(I_{min}+\tau)} + D_{min} \quad (1)$$

where I_{min} and I_{max} are the minimum and maximum luminance of the scene, and D_{max} and D_{min} are the minimum and maximum luminance of the visualization devices, $\tau = \alpha(I_{max} - I_{min})$, $\alpha \geq 0$. Equation (1) ensures that the maximum and minimum luminance values of the scene are respectively mapped to the maximum and minimum luminance of the visualization device. Adjusting α will appropriately tune the overall brightness of the reproduced image. Figure 1 shows plots of the mapping curve of (1) for several different α's. It is seen that the curves distribute the world luminance to display luminance according the values of α. For example, setting α = 0 %, the lower 20% of the world luminance will be placed at over 80% of the lower end of the display luminance. The end effect is of course, that the display will be bright. By increasing the value of α, we can allocate the display luminance to world luminance differently, thus adjusting the overall brightness of the reproduction. When $\alpha \rightarrow \infty$, D(I) = I.

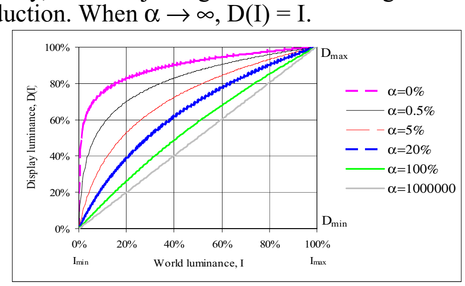

Figure 1: The global luminance-mapping curve of (1) for different values of α.

3. Optimization Approach to TRC Tone Mapping Operator Design

After an HDR image is mapped by equation (1), ignoring numerical errors, the output image still contains the same amount of information as the original input because the mapping function is monotonic and using D(I) can completely recover the original I. The reason that rendering D(I) for display will result in the lack of detail is caused by linear scaling (quantization) as illustrated in Figure 2 (a). In this case, the range of D(I) is divided into equal intervals, and pixels falling into the same interval are compressed to have the same display value. Quantization is done purely on the basis of the pixel dynamic range without taking into account the image's pixel distribution characteristics. As a consequence, in densely populated luminance intervals, compression is too aggressive (too many pixels are squeezed into one display value) resulting in a lost of detail, whilst in sparsely populated intervals, too few pixels occupy a valuable display level thus resulting the under utilization of display luminance levels. A traditional technique that takes into account pixel distribution is histogram equalization as shown in Figure 2 (b). In this case, the method divides the range of D(I) into N intervals based on the pixel distribution only. Within each interval, there are equal numbers of pixels falling onto it. The division of these intervals is purely based on the pixel population distributions and the actual pixel luminance values spanned in the intervals are not taken into account. Again pixels falling into the same interval are mapped (quantized) into the same display luminance. Although, the display luminance levels are fully utilized in this case, densely populated intervals can result in the exaggeration of contrast (the mapping curve is too steep), while in sparsely populated luminance intervals, compression is too aggressive (the mapping curve is too flat). In order to achieve the desirable results, a TRC tone-mapping algorithm should strike a balance between both techniques and combine them together (Figure 2 (c)).

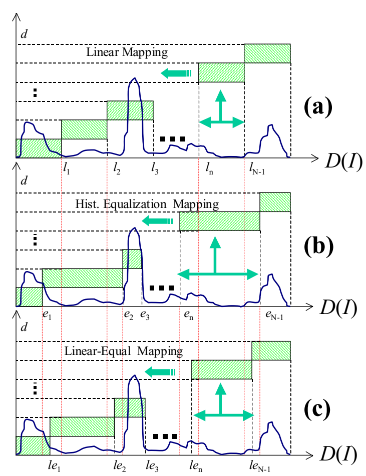

Figure 2. Mapping the output of *D*(*I*) for display. (a) Linear mapping divides the luminance range into N equal length intervals and maps pixels fall into the same interval to the same display value. (b) Histogram equalization mapping divides the luminance range into N intervals such that the numbers of pixels falling into each interval are the same. All pixels falling into the same interval are again mapped to the same display value. (c) A good tone mapping operation should divide the luminance range into N intervals in such a way that the cuts should fall in between those of the linear scaling and histogram equalization.

To implement the tone mapping strategy of Figure 2(c), i.e., to map *D*(*I*) to the display image *d* using a curve that is between linear scaling and histogram equalization, one way to achieve this is to map the image by optimizing following objective function [26]

$$E=\sum_{k=1}^{N-1}\left(le_k-\frac{k\left(D_{\max}-D_{\min}\right)}{N}\right)^2+\lambda\sum_{k=1}^{N-1}\left(\int_0^{le_i}h(x)dx-\frac{k}{N}\int_{D_{\min}}^{D_{\max}}h(x)dx\right)^2 \quad (2)$$

where λ is the Lagrange multiplier. Setting $\lambda = \infty$, optimizing E becomes histogram equalization mapping, and $\lambda = 0$, optimizing E becomes linear scaling mapping. By choosing an appropriate λ, we can strike a balance between the two extreme forms of mapping to suit individual images. An optimal solution to (2) can be found by solving following linear equations.

$$\frac{\partial E}{\partial le_k}=0 \qquad k=1,2,\cdots,N-1 \quad (3)$$

However, equation (3) is seriously under constrained, and a straightforward numerical solution to optimize E in (2) may be difficult to obtain. In [25], an iterative learning procedure was developed, and in [26, 27] a heuristic solution was used to solve the mapping problem.

4. A Closed Form Solution

In order to develop closed form solution, we would like to re-formulate equation (2). The idea is like this: instead of trying to find the cutting points le_k directly, which will be impossible to find a closed form solution, we first try to find the pixel populations $S_{le}(k)$ that falling between the cutting points le_{k-1} and le_k. Once $S_{le}(k)$ are found, then we can easily find the cutting points le_k by counting the pixels starting from the first bin, then the second, etc…

Let $S_l(k)$ be the pixel population falling between the cutting points l_{k-1} and l_k. i.e, pixel populations falling onto each linearly mapped (equal interval) bin (see Figure 2 (a)). Obviously, for a given image $S_l(k)$ is known (and fixed). Assuming the pixel population to be unity, then the number of pixels falling into the bins of histogram equalization mapping will be $S_e(k)=1/N$, where N is the total number of bins, or equivalently, the number of distinctive display levels of the output device. With these assumptions, we can now write equation (2) in the following equivalent form in terms $S_l(k)$ (known), N (given), and $S_{le}(k)$ (unknown)

$$E=\sum_{k=1}^{N}\left(S_{le}(k)-S_l(k)\right)^2+\lambda\sum_{k=1}^{N}\left(S_{le}(k)-\frac{1}{N}\right)^2 \quad (4)$$

Then, setting the partial derivatives of E with respect to $S_{le}(k)$ to zero

$$\frac{\partial E}{\partial S_{le}(k)}=\left(S_{le}(k)-S_l(k)\right)+\lambda\left(S_{le}(k)-\frac{1}{N}\right)=0$$

we have

$$S_{le}(k)=\frac{NS_l(k)+\lambda}{N(1+\lambda)} \qquad \text{for } k = 1, 2, \ldots N, \quad (5)$$

With (5), we can compute the numbers of pixels falling into each bin directly instead of having to resort to heuristic or iterative processes by previous researchers [25 – 27]. From (5), it is not difficult to see that

$$\lambda\to 0: S_{le}(k)\to\frac{NS_l(k)}{N}=S_l(k)\to \text{linear mapping}$$

$$\lambda\to\infty: S_{le}(k)\to\frac{\lambda}{N\lambda}=\frac{1}{N}\to \text{Hist. Equal. mapping}$$

Therefore, it is seen that our new solution (4) and (5) is equivalent to (2) in the sense that by adjusting the value of λ we can achieve from linear to histogram equalization mapping. Once $S_{le}(k)$ is found, then we can find the cutting points le_k for mapping the pixels very easily. All is needed is to count enough pixels for each bin.

5. Experimental Results

We have tested our solution using experimental procedures followed those in previous work [25, 25, 27] and have obtained similar results. Some examples are shown in Figures 3, 4 and 5.

Our results indicate that the optimization approach offers flexibility and simple solution to tone mapping for high dynamic range image display in low dynamic range devices. In general results are satisfactory or acceptable. However, as have been commented by previous authors, TRC based operators have the common weakness of destroying spatial details. Amongst TRC based methods, the optimization based approach which offer the flexibility of adjusting the mapping between linear and histogram equalization seems to be one of the most flexible TRC operators with clear parameter-appearance relations (by setting the value of λ). Because of the subjective nature of tone mapping, this flexibility is very important and very useful. Another advantage of TRC based optimization approach is that it can be implemented very efficiently and can be computed very fast. Our work in this paper has provided a more elegant solution to this approach which will enable even faster and simpler implementation.

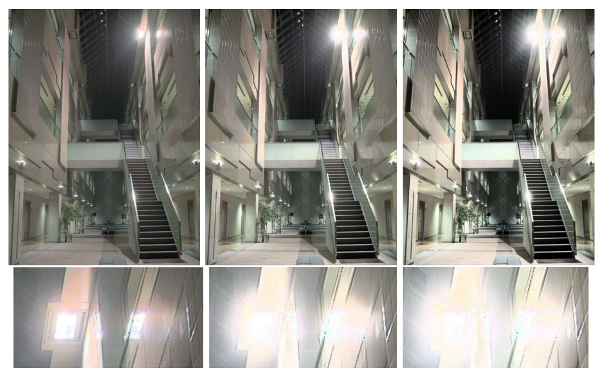

Figure 3. Top row: left to right, result of the optimization method with $\lambda = 0.5$, the method of Ward et al [14] and histogram equalization. Bottom row: amplified regions of the images above them. It is seen that the optimization method provides more flexibility. Radiance map dynamic range 3,900:1. Radiance map courtesy of Greg Ward

Figure 4. Top row: left to right, images mapped by the optimization method with $\lambda = 0.5$, method of Ward et al [14] and histogram equalization. bottom row: amplified regions of the images above them. Radiance map dynamic range 340,016:1. Radiance map courtesy of Paul Debevec.

Figure 5. More results by optimization TRC tone mapping. Radiance maps courtesy of Paul Debevec, Greg Ward and Jack Tumblin.

6. Concluding Remarks

In this paper, we have presented a closed form solution to an optimization approach to tone mapping for high dynamic range image display in low dynamic range devices. Our results indicate that the optimization-based approach can achieve good results which also has the advantage of being flexible, simple to implement and fast to compute. Our closed form solution has added mathematical elegant-ness and computational simplicity to this approach.

References

1. P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs”, Proc. ACM SIGGRAPH’97, pp. 369 – 378, 1997
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. Mann and R. W. Picard, “On being ‘undigital’ with digital cameras: extending dynamic range by combining differently exposed pictures”, IS&T’s 48th Annual Conference, Society for Imaging Science and Yechnology, Washington D. C., pp. 422 – 428, 1995
4. T.G. Stockham. Image Processing in the Context of a Visual Model, Proceedings of the IEEE, 60:828–842
5. J. Blinn. Dirty Pixels. IEEE Computer Graphics & Applications, 9(4):100–105, 1989.
6. G. Ward. Real Pixels. Graphics Gems II. Academic Press, 80–83, 1991.
7. S. B. Kang, M. Uyttendale, S. Winder and R. Szeliski, “High dynamic range video”, ACM Transactions on Graphics, vol.22, no. 3, Pages: 319 – 325, July 2003
8. R. Mantiuk, G. Krawczyk, K. Myszkowski, and H-P Seidel, “Perception-motivated High Dynamic Range Video Encoding”, Proc. of SIGGRAPH’2004, pp. 733-741, 2004
9. F. Drago, K. Myszkowski, T. Annen and N. Chiba, “Adaptive Logarithmic Mapping For Displaying High Contrast Scenes”, The Journal of Computer Graphics Forum, Vol.22, No, 3, pp. 419-426, 2003.9
10. J. DiCarlo and B. Wandell, “Rendering high dynamic range images”, Proc. SPIE, vol.3965, pp. 392 – 401, 2001
11. J. Tumblin and H. Rushmeier, “Tone reproduction for realistic images”, IEEE Computer Graphics and Applications, vol. 13, pp. 42 – 48, 1993
12. M. Ashikhmin, “A tone mapping algorithm for high contrast images”, Proc. Eurographics Workshop on Rendering, pp. 145 – 156, 2002
13. Greg Ward. “A contrast-based scalefactor for luminance display”. In Graphics Gems IV, pages 415–421. Academic Press, 1994.
14. G. W. Larson, H. Rushmeier and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes”, IEEE Trans on Visualization and Computer Graphics, vol. 3, pp. 291 – 306, 1997
15. K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang and K. Zimmerman, “Spatially nonuniform scaling functions for high contrast images”, Proc. graphics Interface’93, pp. 245 – 253, 1993
16. J. Tumblin and G. Turk, “LCIS: A boundary hierarchy for detail preserving contrast reduction”, In Proc. of ACM SIGGRAPH’99, pp. 83-90.
17. F. Durand and J. Dorsey, Fast bilateral filtering for the display of high-dynamic-range images. ACM Trans. Graph. (special issue SIGGRAPH 2002) 21, 3, 257-266, 2002
18. E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, “Photographic tone reproduction for digital images”. ACM Trans. Graph. (special issue SIGGRAPH 2002) 21, 3, 267-276, 2002.
19. R. Fattal, D. Lischinski and M. Werman, “Gradient domain high dynamic range compression”. ACM Trans. Graph. (special issue SIGGRAPH 2002) 21, 3, 249-256,2002
20. Erik Reinhard, 'Parameter Estimation for Photographic Tone Reproduction', Journal of Graphics Tools, Volume 7, Issue 1, pp 45-52, 2002.
21. E. H. Land and J. J. McCann, “Lightness and retinex theory”, Journal of the Optical society of America, vol. 61, pp. 1-11, 1971
22. 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
23. J. A. Ferwerda, S. N. Pattanaik, P. Shirley and D. P. Greenberg. " A model of visual adaptation for realistic image synthesis ". Proceedings of SIGGRAPH’96, pp. 249-258.
24. D. J. Jobson, Z. Rahman and G. A. Woodell, “A multiscale Retinex for dridging the gap between color images and the human observation of scenes”, IEEE Transactions on Image processing, vol. 6, pp. 965-976, 1997
25. 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, Volume 2, pp. 542-547 2004
26 .G. Qiu and J. Duan, “An Optimal Tone Reproduction Curve Operator for the Display High Dynamic Range Images”, Proc. IEEE **ISCAS’2005**, Kobe, Japan, May 2005
27 J. Duan and G. Qiu, "Fast Tone Mapping for High Dynamic Range Images", ICPR2004, 17th International Conference on Pattern Recognition, Volume 2, pp. 847-850, 2004