HDRi2013 - First International Conference and SME Workshop on HDR imaging (2013), paper no. 2.
Bessa M. and Mantiuk R.

Influence of Context and Content on Tone-mapping Operators

P. Korshunov and T. Ebrahimi

MMSPG, EPFL, Switzerland

Abstract

High Dynamic Range (HDR) imaging is one of the technologies that are changing photo, TV, and film industries. The popularity and full public adoption of HDR content is however hindered by the lack of standards in evaluation of quality, file formats, and compression, as well as large legacy base of Low Dynamic Range (LDR) displays that are unable to render HDR. Many tone-mapping operators, aiming to generating viewable LDR content from HDR, were developed to resolve the legacy hardware and software problem. However, there is no consensus on which tone-mapping to use and under which conditions. This paper, via a series of subjective evaluations, demonstrates the dependency of the perceptual quality of the tone-mapped LDR images on the context, such as environmental factors and display parameters, and image content. The results of subjective tests indicate a significant influence of context and content on performance of tone-mapping, and, therefore, both should be considered in applications generating LDR from HDR.

Categories and Subject Descriptors (according to ACM CCS): I.2.10 [ARTIFICIAL INTELLIGENCE]: Vision and Scene Understanding—Perceptual reasoning

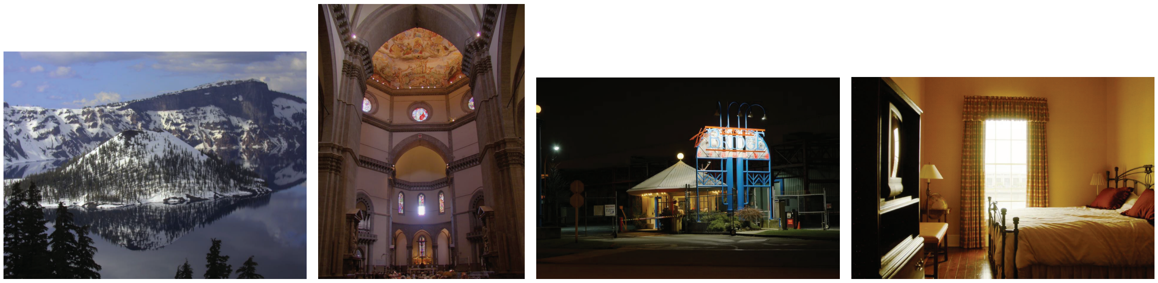
1. Introduction

Many different subjective evaluations have been previously performed to compare different tone-mapping operators for HDR images and video. Main focus of these studies was either on determining a more superior approach to tone-mapping or establishing an evaluation methodology for subjective evaluation of HDR content. As different evaluations result in different sets of best tone-mapping algorithms, it demonstrates that other factors may also affect perceptual quality of the resulted LDR images. To achieve the best possible viewing experience, it is, therefore, necessary to accustom for the factors affecting perception of resulted tone-mapped images.

In order to take human perception into account, we analyze the impact of contextual and environmental parameters on perception of tone-mapped images, such as display type, size, contrast, and brightness characteristics, as well as the type of content, in different surrounding lighting conditions. To understand how different HDR tone-mapping operators influence the perception under different conditions, we conducted a comprehensive subjective evaluation, with twenty human subjects participating in the study. Five commonly used and cited in research literature tone-mapping operators were selected for the subjective study. We used the test-bed and infrastructure consisting of displays with different sizes and characteristics, such as mobile phones, tablets, and large monitors. Environmental conditions and contextual information included the amount of environmental lighting, the way subjects viewed the images, and the backlit light of displays, as well as their size and contrast. One main novelty of this approach is evident when considering that the majority of current evaluation work in HDR ignores the context and environmental factors. By varying different environmental parameters, one could see how these factors affect the perceptual quality of the content, and use it when designing a backward compatible HDR compression.

2. Related work

Several subjective evaluation studies have been conducted in literature to compare different tone-mapping methods for HDR images. One of the first subjective evaluations of HDR images was performed by Ledda *et al.* [LCTS05]. The authors used paired comparison to evaluate the perceptual quality of six different tone-mapping algorithms. An HDR



(a) (b) (c) (d)

Figure 1: Four test HDR images (from left to right): "CraterLake1", "Duomo1", "BridgeStudios2", and "Room".

display was used as reference display for 48 subjects. The focus of this work was on the evaluation methodology for the subjective comparison of HDR images in a controlled environment. The evaluations provided the performance ranking of different tone-mapping algorithms leading to different perceptual qualities in color and gray images.

Yoshida *et al.* [YBMS05] evaluated seven different tone-mapping algorithms via subjective tests to rate the resulted images with regards to their naturalness, contrast, brightness and details of the reproduction in bright and dark regions. The overall goal was to see if different tone-mapped images are perceived differently. The results show that brightness was the largest differentiator and local-based tone-mapping is better for details in brighter regions. The focus in the work by Park and Montag [PM07] was on the scientific images (astronomic, medical, infrared, radar), where 9 tone-mapping algorithms were evaluated. Subjective paired comparison shoed no correlations between perceptual preference and scientific usefulness.

Kuang *et al.* [KYL*07] studied the overall preference of tone-mapping algorithms (9 in total) and their accuracy when compared to actual world scenes, from which they were captured. Paired comparison, rating-scale, and real-world scenes method were performed with fixed environmental parameters (lighting, luminance, screen sizes, etc.) with 19 to 23 subjects. Čadík *et al.* [vWNA08] studied how tone-mapping algorithms (14 in total) affect brightness, contrast, color reproduction, reproduction of details, and artifacts of the images. By performing subjective evaluations with real world reference and without a reference, the authors did not find any significant differences between the two testing methodologies in terms of results.

Annighöfer *et al.* [ATG10] evaluated 8 tone-mapping operators against linear tone-mapping with 51 subjects. Three tone-mapping algorithms were found performing well, consistent with previous studies. However, performance of each operator was found to be content dependent.

Compared to all these subjective tests, the goal of our study is not to find the best tone-mapping algorithm but to demonstrate the importance of other factors, such as display size, type of content, and environment on the quality of resulted tone-mapped image, etc. And the findings of these tests can be useful for building applications that would support HDR-capable and LDR legacy systems, as well as, for development and improvement of HDR compression standards and formats that are backward compatible with the existing LDR technology such as JPEG (examples of such formats can be found in [WS06, MEMS06, CQCW06]).

3. Subjective Tests

This section analyzes the suitability of the most common image and video quality evaluation methods for the subjective evaluation of HDR content and to adapt/extend these methods to take into the account contextual and environmental information. We study the effect of the environmental conditions, display characteristics, and content types on the perceptual quality of HDR images. We have designed a comprehensive methodology for subjective evaluation of quality and conducted supporting set of subjective tests to build a model of the perception of quality of HDR content by human subjects in various contexts and environments.

3.1. Evaluation data and tone-mapping algorithms

As the aim of the subjective study was not to find the best tone-mapping algorithm but to understand how they perform in different conditions, we selected the following five tone-mapping operators (shortened accordingly): "dg" by Drago *et al.* [DMAC03], "mt" by Mantiuk *et al.* [MDK08], "rh" by Reinhard and Devlin [RD05], iCAM ("ic") by Fairchild and Johnson [FJ04], and "lg" as a simple logarithm-based operator. The implementation provided in library "pfstools" was used for the first three operators and code provided in the book by Reinhard *et al.* [RWP*10] was used for the last two. These algorithms were selected to have representation of different approaches to tone-mapping, such as global operators ("lg" and "dg"), based on local information ("rh" and "ic"), and operators utilizing properties of human visual system ("mt").

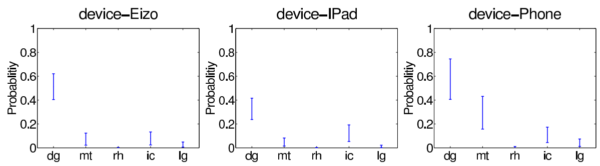


Figure 2: Subjective scores for “CraterLake1” image.

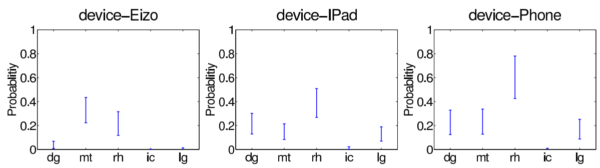


Figure 3: Subjective scores for “Duomo1” image.

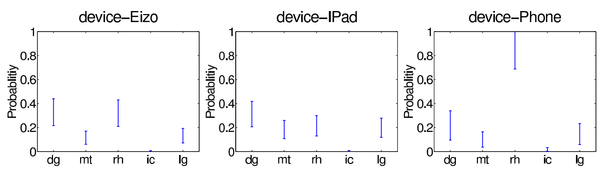


Figure 4: Subjective scores for “BridgeStudios2” image.

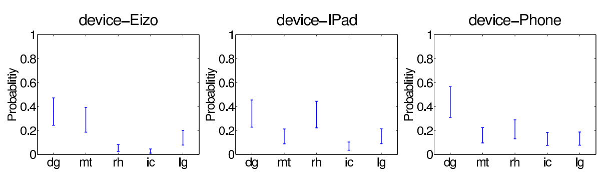


Figure 5: Subjective scores for “Room” image.

To test the selected tone-mapping operators, we used four images (see Fig. 1) from the collection provided in the book by Reinhard *et al.* [RWP*10], as the images typically used for evaluation of tone-mapping operators have very low resolution for today’s monitors. The smallest resolute of the chosen images is 1840×1224. Images are stored in Radiance file format [War91] with 32 bits per pixel. Therefore, tone-mapping operators map floating point values of HDR content in RGBE representation into 24 bits RGB. The images were also selected in such a way to have different representation in terms of content type (indoor or outdoor) and luminance range (night and day shots).

3.2. Test environment and methodology

We have conducted subjective tests as pairwise comparison of different tone-mapping operators. Pairwise comparison of five operators with four test images comprises 40 comparison pairs in total.

The goal of the subjective evaluations of different tone-mapping operators was to find how different is the resulted perceptual visual quality under different environmental conditions, on different devices, and with different type of content. The tests were performed in a laboratory for professional subjective tests when using a monitor and in a typical office environment when tablets and mobile phones were used.

The test room was equipped with a 30” LCD Eizo monitor with resolution 2560×1600. The ambient lighting was obtained with neon lamps of 6500 K color temperature and the walls color were painted mid gray 128, as recommended in Ref. ITU12. The luminance of the Eizo screen was set with EyeOne Display 2 calibration tool to $120 cd/m^2$. We used first generation of iPad with resolution 1024×768 in experiments with tablets, and Samsung Galaxy S with resolution 800×480 for tests with mobile phones. The office for testing with tablets and mobile phones had a typical lighting of about 500 lux, and the brightness was turned to maximum in both types of devices. In practical usage scenario, the environmental ambient light can be determined by either frontal camera of the device (in case of mobile phone and tablet) or a separate web camera (in case of monitor).

Subjective tests with the monitor were performed in a *passive* mode. Each subject was sitting in front of the monitor at a distance 2 to 3 times the height of the stimuli (presented images). A pair of test stimuli in the same comparison set were played one-after-another. All possible pairs in each comparison set were used for comparison in order to obtain complete winning frequency matrices. Since the monitor used had a native resolution of 2560×1600, images could fit in the horizontal space of the display.

Each subject was asked to choose which stimulus had better quality between the two presented stimuli and to mark the answer between “first”, “same”, and “second” on the score sheet. Each stimuli of the pair was displayed for 7 seconds and 5 seconds was given to vote after the pair. Each subject had a training session followed by two separate test sessions, each of which contained 20 pairs of stimuli. Twenty subjects (12 males and 8 females) having normal or corrected-to-normal vision participated in the tests having median age 25 and ranging from 20 to 61 years old.

The experiments with mobile phones and tablets were performed in a similar way as with the monitors, except users were allowed to scroll through the pairs of images by themselves, enabling a more realistic *active* mode of subjective evaluations in such contexts.

The test images (see Fig. 1) in original ‘Radiance’ format were resized to fit different resolutions of monitors, tablets, and mobile phones. Then, all the selected five tone-mapping operators were run on each image with default settings to produce LDR versions in JPEG format, which were used in the pairwise comparisons.

4. Evaluation Results

To illustrate which tone-mapping outperforms in which conditions, we have computed judging probabilities from the subjective tests using BTL (Bradley-Terry-Luce) model [BT52], which is a commonly applied model for comparison of pairwise data. The judging probabilities per tone-mapping operator are presented in Fig. 2-5. In the figures, tone-mapping algorithms are displayed on horizontal axis with subjective judging probabilities on vertical axis. Judging probability values are obtain as average across all subjects that participated in tests for each device. Each figure presents results for each image with three subfigures corresponding to three devices (in order of appearance): Eizo monitor, iPad tablet, and Samsung mobile phone, on which the evaluations were performed. From the figures, one can identify rather easily the most and the least favorite tone-mapping algorithms for each device and for each image.

It can be noted that the tone-mapping operator by Drago *et al.* performs the best in most scenarios, while in some cases, such as for "Duomo1" image (central and left graphs in Fig. 3 for Eizo monitor and mobile phone) and for "BridgeStudio2" image (left graph in Fig. 4 for mobile phone), it shows poor performance.

5. Conclusion

This paper demonstrates, by means of rigorous subjective assessments of various tone-mapping algorithms applied to typical HDR images and rendered in various controlled and uncontrolled environments and devices, that there is no universal tone-mapping operator that always stands out when compared to others.

The choice of the best tone-mapping operator depends on the content, but also on the device used, and other environmental parameters such as back lit lighting, display type and size, environment illumination, etc. These parameters need to be explicitly taken into account when building a support for HDR images in existing LDR-based applications and display systems.

The results presented in this paper can be extended by performing a performance evaluation with a larger set of typical HDR images and by using other more sophisticated tone-mapping algorithms.

Acknowledgements

This work was conducted in the framework of the EC funded Network of Excellence VideoSense, with partial support from COST Action IC1005 – HDRi, and COST Action IC1003 – Qualinet.

Help from Shuting Cai in generating some of the HDR and tone-mapped images used in subjective evaluations performed in this paper is also acknowledged.

References

[ATG10] Annighöfer B., Tajbakhsh T., Grigat R.-R.: Prediction of results from subjective evaluation of real-time-capable tone-mapping operators applied to limited high-dynamic-range images. *J. Electron. Imaging 19* (2010). 2

[BT52] Bradley R. A., Terry M. E.: Rank analysis of incomplete block designs I: The method of paired comparisons. *Biometrika 39* (1952), 324–45. 4

[CQCW06] Chen M., Qiu G., Chen Z., Wang C.: JPEG compatible coding of high dynamic range imagery using tone mapping operators. In *Proc. of the Picture Coding Symposium 2006* (Beijing, China, Apr 2006). 2

[DMAC03] Drago F., Myszkowski K., Annen T., Chiba N.: Adaptive logarithmic mapping for displaying high contrast scenes. In *Proc. of EUROGRAPHICS 2003* (Granada, Spain, 2003), Brunet P., Fellner D. W., (Eds.), vol. 22 of *Computer Graphics Forum*, Blackwell, pp. 419–426. 2

[FJ04] Fairchild M. D., Johnson G. M.: The iCAM framework for image appearance, image differences, and image quality. *Journal of Electronic Imaging 13* (2004), 126–138. 2

[ITU12] *Methodology for the Subjective Assessment of the Quality of Television Pictures*. Recommendation ITU-R BT.500-11., Geneva, Switzerland, 2012. 3

[KYL*07] Kuang J., Yamaguchi H., Liu C., Johnson G. M., Fairchild M. D.: Evaluating HDR rendering algorithms. *ACM Trans. Appl. Percept. 4*, 2 (2007). 2

[LCTS05] Ledda P., Chalmers A., Troscianko T., Seetzen H.: Evaluation of tone mapping operators using a high dynamic range display. *ACM Trans. Graph. 24*, 3 (July 2005), 640–648. 1

[MDK08] Mantiuk R., Daly S., Kerofsky L.: Display adaptive tone mapping. *ACM Trans. Graph. 27*, 3 (Aug. 2008), 68:1–68:10. 2

[MEMS06] Mantiuk R., Efremov A., Myszkowski K., Seidel H.-P.: Backward compatible high dynamic range MPEG video compression. vol. 25, SIGGRAPH '06, pp. 713–723. 2

[PM07] Park S. H., Montag E. D.: Evaluating tone mapping algorithms for rendering non-pictorial (scientific) high-dynamic-range images. *J. Vis. Comun. Image Represent. 18*, 5 (2007), 415–428. 2

[RD05] Reinhard E., Devlin K.: Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics 11*, 1 (Jan. 2005), 13–24. 2

[RWP*10] Reinhard E., Ward G., Pattanaik S., Debevec P., Heidrich W., Myszkowski K.: *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*, 2nd ed. The Morgan Kaufmann series in Computer Graphics. Elsevier (Morgan Kaufmann), Burlington, MA, 2010. 2, 3

[vWNA08] Čadík M., Wimmer M., Neumann L., Artusi A.: Evaluation of HDR tone mapping methods using essential perceptual attributes. *Comput. Graph. 32*, 3 (2008), 330–349. 2

[War91] Ward G.: *Graphics Gems II*. Academic Press, 1991, ch. Real Pixels, pp. 80–83. 3

[WS06] Ward G., Simmons M.: JPEG-HDR: a backwards-compatible, high dynamic range extension to JPEG. In *ACM SIGGRAPH 2006 Courses* (New York, NY, USA, 2006), SIGGRAPH '06, ACM. 2

[YBMS05] Yoshida A., Blanz V., Myszkowski K., Seidel H.-P.: Perceptual evaluation of tone mapping operators with real-world scenes. vol. 5666, Human Vision Electronic Imaging X, SPIE'05, pp. 192–203. 2