Evaluation of Preferred Brightness and Detail Levels in 3D and 2D Images Based on HDR Tone Mapping

Zicong Mai, Colin Doutre, Panos Nasiopoulos, and Rabab Ward
Department of Electrical and Computer Engineering
The University of British Columbia, Canada
Vancouver, Canada
{zicongm, colind, panosn, rababw}@ece.ubc.ca

***Abstract*—High dynamic range (HDR) images provide superior picture quality by allowing a larger range of brightness levels to be captured and reproduced. Even with existing 8-bit displays, picture quality can be significantly improved if content is first captured in HDR format, and then tone-mapping is applied to convert from HDR to 8-bit, low dynamic range format. Tone mapping methods have been extensively studied for 2D images. By varying the tone mapping process, the brightness and level of details (i.e., sharpness) of the output image can be altered. In this paper, we present a study that compares: i) the preferred level of brightness, and ii) the preferred amount of details, of tone mapped images when they are displayed in 2D and 3D formats. Images at a large range of different brightness and detail levels were first generated by HDR capturing followed by tone mapping with different parameters. We performed an extensive subjective experiment that allowed participants to vary the brightness and amount of details in the output tone mapped images and select the level they thought gave the best visual quality. The results showed that there is no statistically significant difference between the preferred brightness level for 2D and 3D content. On the other hand, the subjects consistently chose a higher level of details (i.e. a sharper image) for 3D images compared to the level of details they thought was optimal in 2D mode. This result indicates that when processing 3D content, the image should be sharper than the same content viewed in 2D mode for optimal appearance.**

Keywords—3D; high dynamic range; HDR; tone-mapping; quality of experience

I. INTRODUCTION

High dynamic range (HDR) images/videos have been gaining increasing attention in the past decade because of the superior picture quality they are able to deliver. Existing low dynamic range (LDR) content allows only a limited range of contrast which is far below the capability of human eyes. HDR media supports a very large luminance range that is comparable to what the human vision systems is able to perceive. HDR signals need to be encoded with at least 10 bits per color component [1], as opposed to LDR signals which are represented by only eight bits. Although the majority of today's displays can support only LDR content, they can all provide much better picture quality if the content is first captured in HDR and then converted to LDR format. Such production pipeline, i.e., shooting in HDR and then rendering to LDR, has been gaining increasingly interest in movie/television production and high-end photography.

In order to show HDR content on existing/8-bit displays, a process that converts HDR to LDR signals needs to be conducted. This process is called tone-mapping. A number of tone-mapping operators (TMOs) have been developed under different principles [2]-[8]. Combining HDR capturing and tone mapping is beneficial even when an LDR display is used because they together produce higher quality images with much less over and under saturated areas compared to the traditional LDR capturing process. In addition, tone-mapping allows higher degrees of freedom for artists who during postproduction can decide the final effect/style of the resultant LDR image.

In a similar manner, the 3D display technology also aims at providing viewers with a more realistic visual experience by providing a sense of depth. An increasing number of theatres and households have been equipped with 3D display systems. More 3D content is being produced for satisfying the demand of the rapidly growing 3D consumer market. Ideally, content could be captured in a 3D HDR representation and viewed on a 3D HDR display, to achieve a more lifelike picture quality. However, existing 3D displays can support only 8-bit LDR content. In order for HDR content to be displayed on existing imaging systems, LDR signals need to be generated for each view of the 3D HDR pair. That is, tone-mapping needs to be applied to each view. The fact that tone-mapped images produce less under- and over-exposed areas will help the fuse of the 3D depth, and the superior picture quality of HDR tone-mapped content will also add value to the 3D representation.

By varying the tone mapping method and the related parameters, LDR images with different visual effect can be achieved. In particular, the tone mapping process can adjust the brightness and levels of 'details' in the output images. Many tone mapping methods have been developed trying to preserve local contrast, which gives images with a greater levels of details, i.e., images that look sharper and have stronger texture. Brightness and sharpness/texture have also been noted by artists to affect the visual comfort and quality of 3D content [10]. In summary, the optimal tone mapping parameters may be different for 3D images than those for 2D images.

 ISBN: 978-1-61208-202-8

Fig. 1. Scenes used in the subjective test. Since HDR content cannot be shown on the paper or most monitors, all the images displayed above are tone-mapped using the photographic TMO.

In this paper, we present a study on whether the preferred levels of brightness or details are different for tone-mapped 2D and 3D content. 3D content may need to be prepared in a different way from 2D content, in order to provide the best possible picture quality. Here, we address the problem of identifying the preferred brightness and detail levels for 2D and 3D images by performing a set of subjective experiments.

The rest of the paper is organized as follows: Section II describes the experimental setup. Results and discussion can be found in Section III. Section IV concludes the paper.

II. Experimental Setup

Our study focuses on how the preferred level of brightness and the preferred amount of details differ between 3D and 2D images. "Preferred" means the best depth impression for 3D images and the best overall picture quality for 2D images.

A. Image Preparation

Eight scenes, four indoors and four outdoors, were captured in 3D with multiple exposures, and then stereoscopic HDR images were generated by blending these exposures [10]. These scenes were selected to represent different scenarios, such as scenes containing light sources and scenes having only reflectance. Fig. 1 shows an LDR version of the eight scenes. In order to investigate the effect of brightness and the amount of details, 3D images at different levels of brightness and with different amount of details are generated from the eight HDR images respectively.

To vary the brightness levels in the LDR version with consistent effect, we chose to apply the popular photographic TMO since it provides a user parameter (key value α) for changing brightness. This tone-mapping operator simulates the dodging and burning techniques in photographic tone reproduction. For each scene (i.e., each HDR image), we altered the value of α such that i) 41 LDR images with different brightness levels were produced, from very dark to very bright, and ii) the difference in the overall image brightness between two consecutive levels is as constant as possible. Fig. 2(a) illustrates tone-mapped images at different brightness levels generated using the above approach.

To produce different amounts of details in the LDR version, we choose a popular TMO based on bilateral filtering [5]. This TMO filters an HDR image into a base (low-pass) layer as well as a detail layer and then combines the modified versions of them to yield its tone-mapped image. The original *bilateral filtering TMO* uses a fixed ratio between the base and the detail layers when binding them. In our modification, this ratio λ may be changed for adjusting the contribution of the detail layer against the base layer as can be seen in (1).

$$I_{LDR} = L_b + \lambda \cdot L_d \tag{1}$$

where I_{LDR} denotes the resulting tone-mapped image. L_b and L_d represent the base layer and the detail layer, respectively. The effect of changing this ratio λ on the amount of details is demonstrated in Fig. 2(b).

For each of the scenes, four sets of tone mapped versions of the screen were created: i) 3D representation with various brightness levels, ii) 3D representation at various detail levels, iii) 2D representation at various brightness levels, and iv) 2D representation at various detail levels. For each version, we prepared 41 different levels. In total, there are 8 (scenes) x 4 (versions) x 41 (levels) = 1312 images in our system that have to be evaluated.

(a)

(b)

Fig. 2. Demonstration of images at different brightness and detail levels: (a) - the scene “Library” at different brightness levels; (b) the scene “ICICS” at different detail levels.

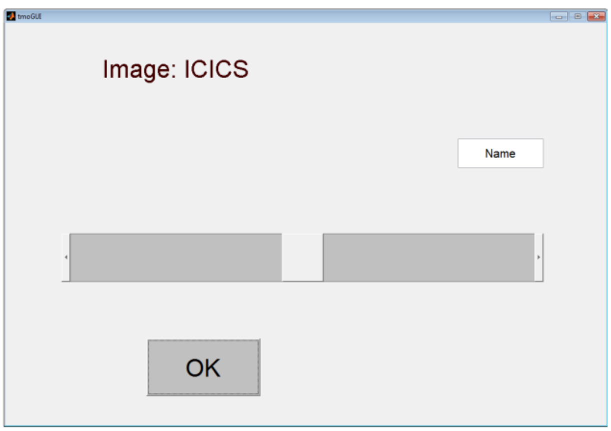


Fig. 3. Illustration of the graphic user interface for subjects to fully navigate the psychophysical experiment on their own pace and to select the images at the preferred brightness or detail levels.

B. *Testing enviorment and procedures*

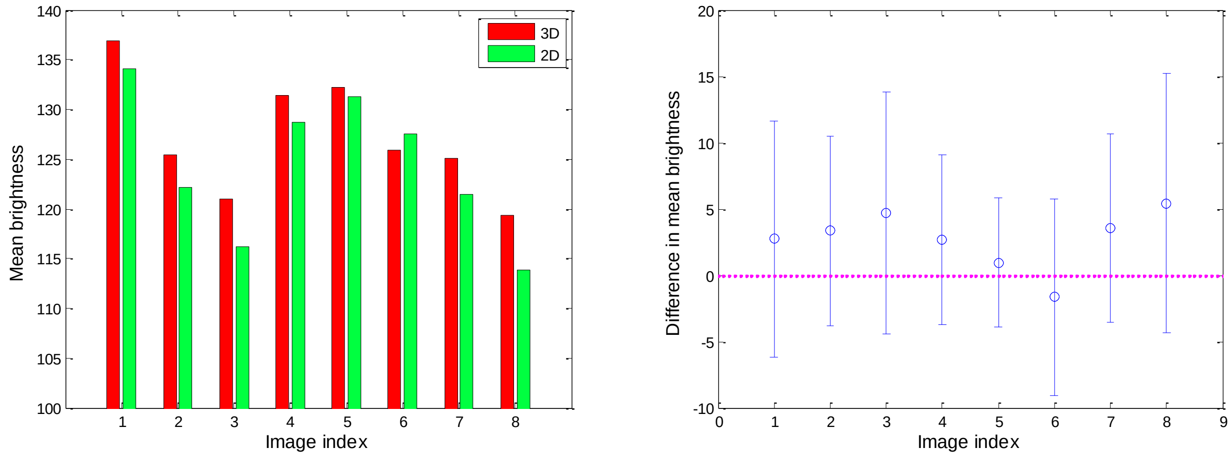
Eighteen subjects (18 – 38 years old) participated in our subjective experiment. All of them had normal or corrected vision with no/marginal experience in 3D technology. The display device used in our test was a 65” Full HD 3D display (©Panasonic, Plasma, TC-P65VT25), which uses active shutter glasses. The viewing conditions of our tests were set based on the ITU-R Recommendation BT.500-11 [13]. Viewers keep having their 3D glasses on when watching 2D videos. This guarantees that the brightness reduction is the same when watching 3D and 2D videos.

A secondary display, a 19” 2D screen, was placed closed to the viewer, where a graphic user-interface was created for subjects to select their preferred image (level) for each of the versions. Fig. 3 demonstrates the user interface. A slider bar can be found near the center of the interface. The slider consists of 41 values and each value corresponds to each brightness/detail level. The position of the slider is reset to the middle of the bar when a new scene is loaded. Once a particular position/value is chosen, an image at the selected level (brightness or amount of detail) is updated and shown immediately. As the slider is moved from left to right, an image will become brighter in the case of evaluating brightness versions and will have more details in the case of evaluating detail versions. An “OK” button sits below the slider and is used by subjects to confirm their selection. The selected value on the slider will then be recorded.

Before the test started, participants were provided with a training section which ensures they were comfortable in using the scoring interface and had their eyes adapted to the viewing conditions. In the test, they were asked to move the slider for selecting the preferred images for each version of each scene. As stated at the beginning of this section, “preferred” means the best depth impression for 3D images and the best overall picture quality for 2D images.

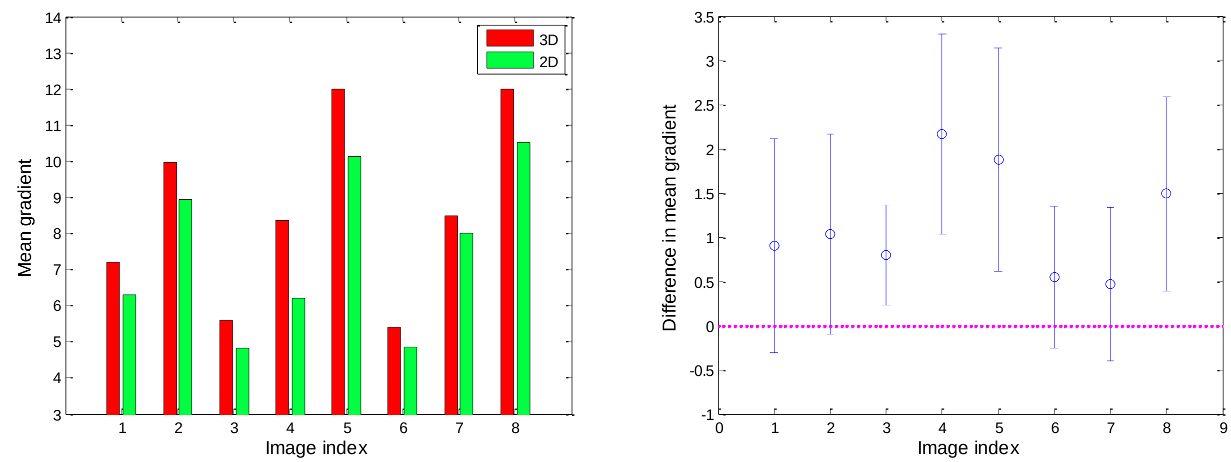
During the test, subjects know in advance the kind of image (3D or 2D) to be viewed. The procedure for each subject would proceed as follows. First, a 3D version of one scene would be shown to the subject, and he/she would adjust the slider on the GUI to select the level of brightness they thought gave the best 3D quality. Then the subject would press the OK bottom on the GUI, after which 5 seconds of grey is shown to allow the subject to rest his/her eyes. Then, the next scene would be displayed and the subject would repeat the process. After the user has done this for all 8 scenes, the slider is changed to control the amount of detail in the images, while still displaying 3D images. Again the user would adjust the slider to select what is thought to be the best amount of details for each scene. After the subject has selected the preferred amount of details for all eight scenes, the entire procedure is repeated in 2D viewing mode. That is, the user first selects what he/she considers the optimal amount of brightness for each scene, and then selects the optimal amount of details for each scene viewed in 2D.

The time of each test was completely controlled by the participants and lasted until they reached their final decision.



(a) (b)

Fig. 4. Comparison between 2D and 3D images in terms of the preferred brightness level: (a) – mean image average, (b) subtraction of the mean image average of 2D from that of 3D images. The horizontal axis denotes the image index, and 1 – 8 correspond to 'ICICS', 'MeetingTable', 'LabWindow', 'Bulletin', 'Stairs', 'SauderBuilding', 'LibraryTree', 'ChemEngEntrance', respectively. The vertical axis denotes pixel value in (a) and the difference of pixel values in (b).



(a) (b)

Fig. 5. Comparison between 2D and 3D images in terms of the preferred detail level: (a) – mean image detail, (b) subtraction of the mean image detail of 2D from that of 3D images. The horizontal axis denotes the image index, and the image order is the same as Fig. 4. The vertical axis denotes gradient value in (a) and the difference of gradient values in (b).

Although the time varies depending on the subject, no subject took more than 25 minutes to complete a test.

III. Results and Discussion

After recording the results from the psychophysical tests, we performed statistical analysis on the collected data. First, we tested for outliers based on ITU-R Recommendation BT.500-11 [13] and in our case no outlier was identified, so the data for all subjects were used.

In order to quantify the brightness and the detail levels for the different tone mapped images, we use mean image brightness B_{img} and mean image gradient D_{img}. These are defined as:

$$B_{img} = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} I(i,j)$$

$$D_{img} = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} \left| I(i,j) - I(i+1,j) \right| + \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} \left| I(i,j) - I(i,j+1) \right| \qquad (2)$$

where $I(i, j)$ denotes the pixel value at the location of (i, j) in an image I, and m and n are the dimensions of the image I.

The plot in Fig. 4(a) shows the mean image brightness (B_{img}) of the preferred 3D and the preferred 2D representations for each of the eight scenes (horizontal axis). The height of each bar is obtained by first calculating the mean image brightness of the preferred level from each of the 18 subjects and then averaging these 18 values. A general observation is that the mean brightness of the preferred 3D images is slightly higher than that of the preferred 2D counterparts for all the scenes except scene six ("Sauder"). All the mean brightness values fall in the interval between 115 and 135.

In order to gain a better understanding of the differences between 3D and 2D viewing, for each subject we calculated the difference between the average brightness of their preferred images in 3D and 2D modes as:

$$\Delta_B = B_{pref,3D} - B_{pref,2D} \quad (3)$$

where $B_{pref,3D}$ is the average brightness of the image the user selected with the slider in 3D mode (their 'preferred' image), and $B_{pre,f2D}$ is the corresponding average brightness of the preferred image they selected in 2D mode. A positive value for Δ_B indicates the subject prefers a brighter image when viewing in 3D compared to viewing in 2D. Fig. 4(b) shows the difference averaged over the eighteen participants, and also the 95% confidence interval for the differences. It is seen that although the average values of seven out of the eight scenes are above zero, all of the 95% confidence bars cross the zero axis. Therefore, there is no statistically significant difference in the preferred brightness level between a 3D image and its 2D counterpart.

The results of the preferred mean image gradient of the 3D and the 2D representations are shown in Fig. 5(a). Each point is the average of the preferred gradient selected by each of the subjects for each of the scenes. For all scenes people selected a higher level of details in 3D viewing mode compared to 2D viewing. Similar to the brightness case, we compute the difference in mean gradient between the preferred 3D and the preferred 2D images for every single subject and every scene. Fig. 5(b) shows such difference averaged over 18 subjects for each of the scenes, and the 95% confidence intervals are also provided. Positive values on the vertical axis mean that people favor more details in 3D images than their 2D counterparts. It is observed from the plot that all the average difference is above zero. Moreover, for many of the scenes, the 95% confidence interval is either completely (scenes 3, 4, 5 and 6) or majorly (scenes 1, 2, 7 and 8) above zero. A reliable conclusion can thus be drawn that people prefer a higher level of details (i.e., a sharper image with more texture) when viewing in 3D than when viewing in 2D. Since our test images cover a wide range of content (four indoor and four outdoor scenes), this conclusion will hold regardless of the nature of the content. This could be explained by the conclusion in Cormack et al. [12], that the 3D effect can be improved when the contrast is near the visibility threshold. By adjusting the tone-mapping process to produce more detail, the 3D effect in image regions where the contrast is near the threshold may be improved.

IV. CONCLUSIONS

In this paper, we studied how different are 3D and 2D HDR tone-mapped images in terms of i) the preferred brightness levels and ii) the preferred detail levels. Images at a large range of different brightness and detail levels were generated from high dynamic range capturing followed by the tone-mapping process. With such a great variety of images, we conducted an intensive subjective experiment that allows participants to select 3D and 2D images with their preference on brightness and details, respectively. Our results show that while people selected slightly brighter images in 3D viewing compared to 2D, the difference is not statistically significant. However, compared to 2D images, the subjects consistently preferred having a greater amount of details when viewing in 3D.

REFERENCES

[1] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Perception-motivated high dynamic range video encoding," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 730–738, 2004.

[2] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 21, no. 3, pp. 267–276, 2002.

[3] R. Mantiuk, S. Daly, and L. Kerofsky, "Display Adaptive Tone Mapping," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 27, no. 3, pp. 68–68, 2008.

[4] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive Logarithmic Mapping for Displaying High Contrast Scenes," *Computer Graphics Forum (Proc. of Eurographics)*, vol. 22, no. 3, pp. 419–426, 2003.

[5] F. Durand, and J. Dorsey, "Fast Bilateral Filtering for the Display of High-Dynamic-Range Images," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 21, no. 3, pp. 257–266, 2002.

[6] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, W. Heidrich, " Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image/Video Compression ", *IEEE Transactions on Image Processing (TIP)*, vol. 20, no. 6, pp. 1558-1571, 2011

[7] R. Fattal, D. Lischinski, and M. Werman, "Gradient Domain High Dynamic Range Compression," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 21, no. 3, pp. 249–256, 2002.

[8] E. Reinhard, and K. Devlin, "Dynamic Range Reduction Inspired by Photoreceptor Physiology," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 1, pp. 13-24, 2005.

[9] B. Mendiburu, "3D Movie Making – Stereoscopic Digital Cinema from Script to Screen." Elsevier, 2008.

[10] P. E. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," Proceedings of the 24th annual conference on Computer graphics and interactive techniques (*Proc.* SIGGRAPH '97), pp. 369-378, 1997.

[11] ITU-R, "Methodology for the subjective assessment of the quality of television pictures," ITU-R, Tech. Rep. BT.500-11, 2002.

[12] L.K. Cormack, S.B. Stevenson, and C.M. Schor, "Interocular correlation, luminance contrast and cyclopean processing," *Vision Research*, vol.31, no.12, pp. 2195–2207, 1991.