Dynamic Tone Mapping with Head-Mounted Displays

Matt Yu
Department of Electrical Engineering, Stanford University

Abstract

The real world consists of many scenes which contain a high dynamic range. While modern cameras are capable of capturing the dynamic range of these scenes, displays still only show a low dynamic range. Many tone map operators exist but very few consider the use of head-mounted displays. We create a dynamic tone map operator for use on panorama high dynamic range images by considering a user's head position and subsequent viewport. The tone map operator normalizes the image shown to the user by the log average luminance of the viewport. Furthermore, we use a simple model of eye adaptation to mimic the effects of light and dark adaptation. A simple A/B test shows our dynamic tone operator is preferred over a standard global tone map operator.

Figure 1: HDR content tone mapped using the Reinhard global TMO [12]. Since panorama content is spherical in nature, the content must be mapped to a plane for traditional display. In this case, the equirectangular projection is used.

1. Introduction

Recently, there has been a surge in virtual and augment reality technologies. At the forefront of these technologies are head-mounted displays which include Oculus Rift, Microsoft HoloLens, and the HTC Vive. While the main driving force behind these devices has been to create compelling games, many other applications can take advantage of the immersive experience provided by head-mounted displays. One such application is using one or more cameras to capture pictures or video of the world around a single point. When these pictures or videos are viewed on a head-mounted display, user's can view the world as if they're standing at the captured location.

This immersive content can consist of very high dynamic range (HDR) scenes. Consider, for example, a typical outdoor scene. While a traditional photographer can choose to shoot away from the sun, a photographer trying to capture all the views around a fixed point will inevitably take a picture where the sun is present. Thus, multiple exposures could be used to capture such an HDR scene. However, modern displays can still only show a limited dynamic range. The problem of HDR tone mapping is the process of reducing the dynamic range of HDR content such that the content can be displayed on a regular, limited dynamic range display. While there has been a lot of work on HDR tone mapping for traditional images, there has been relatively little work on HDR tone mapping for panorama HDR content.

The two main contributions of our work are the following:

- We propose a new HDR tone mapping operator which takes into account the fact that a user only looks at a portion of an HDR panorama.
- We introduce a simple method to mimic light and dark adaptation in human vision.

Fig. 1 shows an example of the HDR content[1] used in this project.

2. Related Work

HDR tone mapping for traditional planar images is a well studied field. In this section, we offer a very brief and incomplete review. However, for a relatively thorough review of tone mapping, [13] may be consulted.

[1]HDR panoramas used in this project can be found at http://www.hdrlabs.com/sibl/archive.html and http://www.hdri-hub.com/hdrishop/freesamples/freehdri.

Generally, HDR tone mapping operators can be broken down into global operators and local operators. Global operators apply the same mapping to all pixels and are generally fast and computationally efficient. Local operators, on the other hand, vary spatially by considering a small neighborhood around each pixel. While more computationally demanding, local operators may preserve local contrast better than global operators.

Some examples of global tone mapping operators include scaling the dynamic range by the scene's key value [12] and adaptive logarithmic mapping [2]. Moreover, both global and local tone mapping operators may consider the perceptual response of the human visual system [7, 9, 11, 5, 10, 8] to generate more realistic images. Some examples of local tone mapping operators include gradient domain HDR compression [4] and bilateral filtering [3]. Recently, there has even been work on temporally coherent tone map operators for use in such applications as HDR video [6, 1].

However, while there is plenty of work on HDR for images and videos presented on standard displays, there has been relatively few work on HDR tone mapping for use with head-mounted displays. Perhaps the closest work is [14] which performs tone-mapping with a head-mounted display but only in the context of low-vision aid and not for the generation of accurate or pleasing images.

3. Method

Due to the lack of work regarding HDR tone mapping for head-mounted displays, this work begins by considering how to extend a simple global operator for use in the situation when a user only looks at a portion of the image. Then, a simple model for human eye adaptation is introduced in the second half of the section.

3.1. Viewport Luminance Adjustment

Scaling the dynamic range of an image by the scene's key value can be seen as setting the exposure on a camera. The key value can be approximated by the log-average luminance [12, 13] of the image pixels:

$$\bar{L}_w = \frac{1}{N}\exp\Big(\sum_{x,y}\log(\delta + L_w(x,y))\Big) \tag{1}$$

where $L_w(x, y)$ is the world luminance of the pixels at location x, y. Then, the displayed luminance can be calculated as:

$$L_d(x,y) = \frac{a}{\bar{L}_w}L_w(x,y) \tag{2}$$

where a is a user parameter specifying the value which the key value of the scene is mapped. Thus, we can see that, if the scene is bright, i.e., the value of $\bar{L}_w$ is large, then the dynamic range will be mapped such that details around bright objects will be perceptible. On the other hand, detail around dark objects will be more perceptible if the key value is very dark.

This mapping, unfortunately, takes into account all values in the HDR panorama in order to compute the key value. Since the user only looks at a portion of a panorama at a time (as shown in Fig. 2a), a more accurate mapping should consider only the portion which the user can see with a head-mounted display.

Thus, we introduce a separate key value for each user viewport. Specifically, we can now introduce an additional temporal component to the key value calculation:

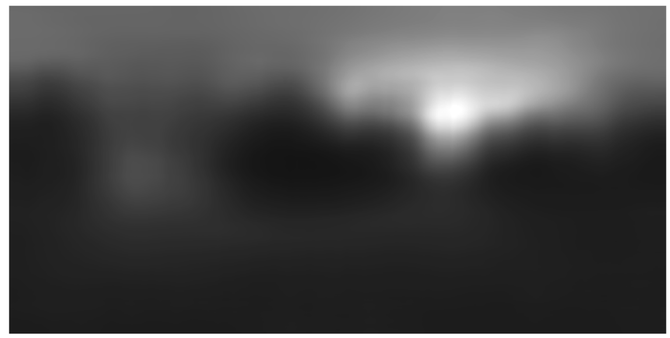
$$\bar{L}_w(V(t)) = \frac{1}{N}\exp\Big(\sum_{x,y\in V(t)}(\log(\delta + L_w(x,y)))\Big) \tag{3}$$

so that the key value is calculated only over the pixels in the user's viewport at a given time. The displayed luminance can be modified accordingly:

$$L_d(x,y,t) = \frac{a}{\bar{L}_w(V(t))}L_w(x,y) \tag{4}$$

Note that the calculation of the viewport which a user views is complicated by the fact that the viewport is a pro-

(a)



(b)

Figure 2: (a) The user's field of view (approximated by the red box) can be significantly smaller than the entire panorama. (b) The log average luminance of the viewport surrounding each pixel.

jection of the panorama onto a rectangular plane. One potential solution is to compute the log average luminance at run time, thus ensuring the log average is computed over the correct values. However, this introduces noticeable and unacceptable delay into a head-mounted display system which requires very low latency. To mitigate this problem, the log average luminance was calculated offline and stored as a lookup table at run time (see Fig. 2b). Furthermore, the viewport was approximated by a large window in the equirectangular panorama domain. This approximation works well at the regions corresponding to the equator but contains large distortions near the poles.

3.2. Simple Adaptation Model

While there has been much prior work modeling the adaptation of the human visual system, this work aims only to simulate a small factor. In particular, while light adaptation (going from a dark background to a bright background) occurs quickly, dark adaptation occurs relatively slowly. We rewrite our displayed luminance as:

$$L_d(x,y,t) = \frac{a}{y(t)} L_w(x,y) \tag{5}$$

Without considering adaptation, we have (as in our previous equation):

$$y(t) = \bar{L}_w(V(t)) \tag{6}$$

To consider adaptation, we introduce the following update rule:

$$y(t) = \alpha \bar{L}_w(V(t)) + (1-\alpha) y(t-1) \tag{7}$$

This results in a rapid approach to the target value where the rate of approach decays as the target value is reached. See Fig. 3 for an illustration of the behavior of the update rule. While the behavior is similar for both dark and light adaptation in the linear luminance domain, the perceptible effect is different. As described by Weber's law, changes in luminance are more perceptible at low background intensities than at high background intensities. Thus, modeling dark adaptation as an exponential decay to the target value will be perceived as a linear drop in key value. Modeling light adaptation as an exponential rise to the target value will be perceived as a much faster rise to the key value. In other words, while our update rule is the same for both dark and light adaptation, the user will feel as if dark adaptation occurs relatively slower than light adaptation.

4. Results

This system was deployed using an OpenGL panorama viewer in combination with an Oculus Rift DK2. HDR content along with offline computations (e.g,. viewport luminance averages) were loaded and shaders were used to dynamically tone map HDR content to a resulting texture. These textures were mapped to spheres so that the user would view a different portion of the panorama depending on their viewing direction. Due to the simplicity of the tone mapping operation and the use of offline computations, the system ran at greater than 60fps resulting in smooth operation with the head-mounted display. Fig. 4

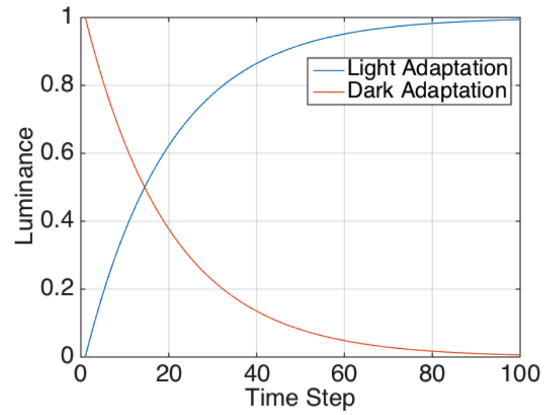


Figure 3: Simple response curves for light and dark adaptation. Note that the update occurs in the linear luminance domain.

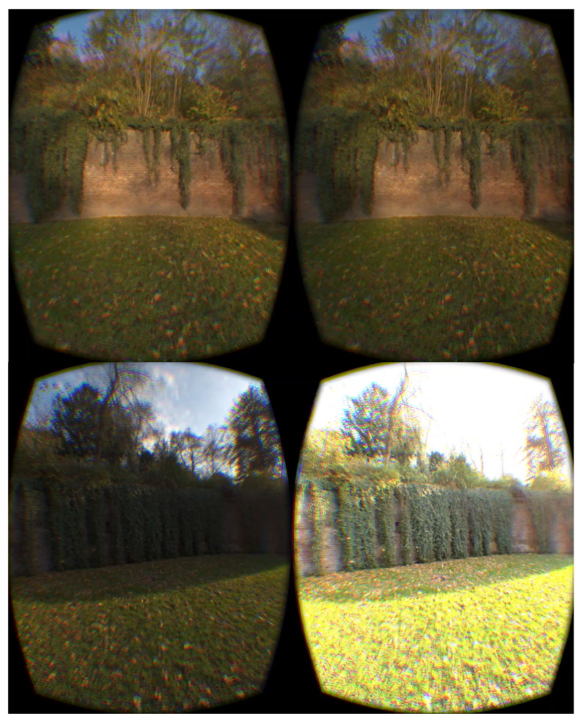


Figure 4: Global (left eye) vs. viewport (right eye) tone mapping operators. The top view and bottom view shows how the viewport method changes its tone mapping method based on the viewable pixels.

Figure 5: Dark adaptation simulation. The user has just viewed a bright scene and starts viewing a dark region at $t = 0$. As time progresses, the viewport gets brighter to simulate the effect of the user adapting from a light to dark region.

shows the difference between using a global tone mapping operator and the viewport tone mapping operator used in this report. The global operator uses the same function as the viewport operator except that the key value approximation is computed over the entire image rather than just the viewport. Fig. 5 shows the effect of dark adaptation with the global tone mapping operator used again for comparison. While the adaptation model is simple, it produces a temporally smooth and pleasing result.

To verify these results, a small subjective test was performed. A standard A/B comparison was used to compare the global and viewport based tone mapping operators. 8 adults ranging between 20-35 were shown results from both tone mapping operators and asked which they preferred. The results are shown in Tab. 1. While the test was small, there is a clear preference towards the viewport tone mapping operator.

Global	Viewport
1.5	6.5

Table 1: Results from an A/B comparison between using a global vs. a viewport tone mapping operator. Number represents the count of people who preferred that method. One person noted the differences between the methods but could not choose which he preferred (hence the 0.5).

5. Discussion

We introduced a dynamic tone mapping operator which takes into account that user wears a head-mounted display to view an HDR panorama. This allows us consider only the pixels displayed to the user at any given time rather than all pixels. This simple tone mapping operator resulted in real-time processing, suitable for use with a head-mounted display. Furthermore, we introduced a simple adaptation model which accounted for the fact that dark adaptation takes a relatively longer amount of time than light adaptation. The performance of our new tone mapping operator was verified with a small subjective study.

6. Future Work

There are a least three major avenues still left for exploration. First, the adaptation model used in this report was very simple. There has been much work in accurately modeling the adaptation of the human visual system and applying the concepts learned in this area could lead to a more realistic result. Second, humans perceive objects in their foveal vision different than in their peripheral vision. In particular, detail can only be perceived in the foveal region. This suggests that a tone mapping operator for head-mounted displays should treat these regions differently. Third, eye-tracking would allow the tone mapper to know exactly what a user is looking at. The limitations of the head-mounted display are that a user does not always look directly at the center pixels. These possibilities, along with the rapid development of new head-mounted displays and even HDR panorama video, make the study of HDR with head-mounted displays an interesting topic to study further.

References

[1] T. O. Aydin, N. Stefanoski, S. Croci, M. H. Gross, and A. Smolic. Temporally coherent local tone mapping of HDR video. *ACM Trans. Graph. ()*, 33(6):196–13, 2014.

[2] F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive Logarithmic Mapping For Displaying High Contrast Scenes. *Comput. Graph. Forum ()*, 22(3):419–426, 2003.

[3] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *SIGGRAPH*, 21(3):257–266, 2002.

[4] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *SIGGRAPH*, 21(3):249–256, 2002.

[5] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A Model of Visual Adaptation for Real-

istic Image Synthesis. *SIGGRAPH*, pages 249–258, 1996.

[6] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High dynamic range video. *ACM Transactions on Graphics*, 22(3):319–325, July 2003.

[7] P. Ledda, L. P. Santos, and A. Chalmers. A local model of eye adaptation for high dynamic range images. *Afrigraph*, pages 151–160, 2004.

[8] R. Mantiuk, S. J. Daly, and L. Kerofsky. Display adaptive tone mapping. *ACM Trans. Graph. (TOG) 27(3)*, 27(3):1, 2008.

[9] R. Mantiuk, K. Myszkowski, and H.-P. Seidel. A perceptual framework for contrast processing of high dynamic range images. *TAP*, 3(3):286–308, 2006.

[10] S. N. Pattanaik, J. Tumblin, Y. H. Yee, and D. P. Greenberg. Time-dependent visual adaptation for fast realistic image display. *SIGGRAPH*, pages 47–54, 2000.

[11] E. Reinhard and K. Devlin. Dynamic Range Reduction Inspired by Photoreceptor Physiology. *IEEE Trans. Vis. Comput. Graph. ()*, 11(1):13–24, 2005.

[12] E. Reinhard, M. M. Stark, P. Shirley, and J. A. Ferwerda. Photographic tone reproduction for digital images. *SIGGRAPH*, 21(3):267–276, 2002.

[13] E. Reinhard, G. Ward, S. N. Pattanaik, P. E. Debevec, and W. Heidrich. *High Dynamic Range Imaging - Acquisition, Display, and Image-Based Lighting (2. ed.)*. Academic Press, 2010.

[14] R. Urea, P. Martnez-Caada, J. Gmez-Lpez, C. Morillas, and F. Pelayo. Real-time tone mapping on gpu and fpga. *EURASIP Journal on Image and Video Processing*, 2012(1), 2012.