CONSISTENT TONE REPRODUCTION

Min H. Kim 
Jan Kautz

Department of Computer Science
University College London
London, UK
email: {m.kim, j.kautz}@cs.ucl.ac.uk

ABSTRACT
In order to display images of high dynamic range (HDR), tone reproduction operators are usually applied that reduce the dynamic range to that of the display device. Generally, parameters need to be adjusted for each new image to achieve good results. Consistent tone reproduction across different images is therefore difficult to achieve, which is especially true for global operators and to some lesser extent also for local operators. We propose an efficient global tone reproduction method that achieves robust results across a large variety of HDR images without the need to adjust parameters. Consistency and efficiency make our method highly suitable for automated dynamic range compression, which for instance is necessary when a large number of HDR images need to be converted.

KEY WORDS
Computer Graphics, High Dynamic Range Imaging, Tone Mapping

1 Introduction
The dynamic range (ratio of maximum to minimum luminance; usually denoted in a base-10 logarithm) of imaging devices, real-world, or artificial luminance shows large variation. High dynamic range (HDR) imaging has been introduced to record real-world radiance values, which can be much higher range than that of ordinary imaging devices. Radiance maps are usually generated by capturing several exposures [1–5] and can have a dynamic range of about nine to ten orders of magnitude. Photographic HDR images or artificial radiance maps cannot be displayed properly on low dynamic range (LDR) output devices (with about two orders of magnitude) due to the huge difference in dynamic range. Consequently, the dynamic range of the HDR scene needs to be mapped into the range of an output device, which is called tone reproduction or tone mapping.

Over the years, many different tone reproduction operators have been developed. The majority of research has focused on improving local operators, pursuing less artifacts and more efficient computation times [6–18]. Global operators have received less attention [19–22] since high contrast appearance is difficult to achieve, but on the plus side they do not suffer from halo-artifacts like many local operators and are much more efficient.

Relatively little work has been done to achieve consistent tone mapping results across a variety of image without the need to tweak parameters. Consistent tone-mapping is necessary for applications that cannot afford manual intervention and require many images to be tone-mapped, e.g., thumbnail creation for HDR images. The only work in this area [23] estimates parameters for Reinhard et al.’s tone mapper [14] and then applies the tone-mapper with those parameters. The method achieves good results, but cannot take user-preferences into account, such as a preference for higher contrast at the cost of detail.

We propose a new global tone mapping operator, which is efficient and achieves consistent results across different input HDR images. The user has two parameters at their disposal to indicate overall preferences (contrast vs. detail and light vs. dark). These parameters can be set once by the user (or left at their default values), and then achieve consistent results across different HDR inputs. Our luminance compression is also based on a sigmoidal function, like other methods before [11, 14, 22], yet we use it in such a way that inter-image consistency can be achieved.

1.1 Related Work
Tumblin and Rushmeier [19, 24] were the first to address the research question of how to render computer-generated HDR images. Their approach is to manipulate the tone-reproduction curves of HDR images. Ward et al. [20] suggested a global adaptation approach, which is based on histogram equalization. Drago et al. [21] introduced a global tone-mapping model which is based on logarithmic compression following the hypothesis by Fechner [25]. They manipulate the base of logarithm to adjust the contrast of images. Reinhard and Devlin [22] introduced a global method to mimic the physiological response of photoreceptors.

based on a bilateral decomposition of an image into a base layer and a detail layer. Only the base layer is compressed and the detail layer is added back in. Reinhard et al. [14] presented a mixed approach of global and local adaptation, for which they employed low-pass filtering through FFT. The resulting quality appears more robust compared to other approaches.

Finally, there are perceptual models that simulate human vision system (HVS) and its local adaptation in order to solve the tone-mapping problem. Ferwarda et al. [8] developed a visual adaptation model as an extension of CIECAM97s [27], which applies a threshold vs. intensity (TVI) function for tone-mapping. Pattanaik et al. [9, 11] introduced a more rigorous computational model of adaptation and spatial vision for realistic tone-reproduction. Johnson and Fairchild [16] presented a new color appearance model that modifies CIECAM02 with a spatial vision appearance model, and which is applicable to HDR tone-mapping. Since Land [28] proposed the retinex theory to simulate the HVS, many researchers [7, 12, 17] have worked on using it for HDR tone-mapping. Even though it provides better results than the other perceptual models, the computational cost is rather considerable.

Recent solutions for tone-mapping have focused on local adaptation or sub-band filtering in the frequency or gradient domain [10, 13, 14, 18]. Even though they can provide strong compression of the dynamic range, results can appear unnatural due to halo artifacts. Furthermore, computational efficiency is often lacking. Therefore, we revisit the original idea of global adaptation [19], aiming to design a general method to control the dynamic range of images, which is efficient and consistent across different HDR images. The rest of this paper will present our proposed method in detail.
2 Our Approach

Our method is based on two popular assumptions that human vision sensitivity is concentrated in the averaged log luminance which is broadly used by many tone reproduction methods [11, 14, 19, 29] and that human sensitivity statistically follows a Gaussian distribution [30–33]. It consists of two main elements: (a) dynamic range compression and (b) inverse display characterization.

2.1 Dynamic Range Compression

We will first describe our dynamic range compression method before we detail the inverse display characterization.

2.1.1 Characteristic Curve Control

Our method to compress the dynamic range of radiance map is inspired by the characteristic curve of photographic material in the log-log domain. This can be described in Hurter and Driffield’s DlogE plot [34], which plots density (logarithm of reflective luminance) against logarithm of the luminance incident on the material. The DlogE plots are used in the sense of the Jones quadrant diagram [35]. For instance in photography, the diagram consists of four quadrants of DlogE plots: original density, film density, paper density, and system transfer function. We utilize the diagram to explain our method (see Figure 1), of which axes are mirrored and show different variables in order to represent circulation of tone reproduction at each stage. Figure 2 shows the performance of the proposed method at intermediate steps corresponding to the quadrant diagram (Figure 1).

Considering that we need to adjust the dynamic range of an HDR radiance map into that of a target display, the adjustment procedure can be explained easily with the help of DlogE plots. Suppose that we have an HDR radiance map, using the church image for illustration purposes (Figure 2), which has a dynamic range (luminance) of 5.5 (1:343,512) and that we need to observe the radiance map through a display which has a dynamic range of 2.4 (1:256, 8-bits [36]). The upper part from the 1st to the 2nd quadrant in Figure 1 presents a simple transform of dynamic range from real-world luminance to output display luminance (log-log plot, counter-clockwise order). By linearly scaling the HDR radiance map to the range of display luminance in the DlogE domain (scaled by appr. 0.43), we can adjust the dynamic range of the HDR radiance map into that of display luminance. The dynamic range of these two is then identical. The scaling factor $k_1$ is computed as follows:

$$k_1 = \frac{\log L_{d_{\text{max}}} - \log L_{d_{\text{min}}}}{\log L_{s_{\text{max}}} - \log L_{s_{\text{min}}}},$$  \hspace{1cm} (1)

where $\log L_{d_{\text{max}}}$ and $\log L_{d_{\text{min}}}$ are the maximum and minimum luminances of the display signals and $\log L_{s_{\text{max}}}$ and $\log L_{s_{\text{min}}}$ are the maximum and minimum luminances of the HDR radiance map.

The dynamic-range compressed image can be computed as:

$$L_1(x, y) = \exp(k_1 \cdot \log L_0(x, y)), \hspace{1cm} (2)$$

where $L_1$ is the compressed luminance at pixel address $(x, y)$ and $L_0$ is the luminance of the HDR image at each pixel.

2.1.2 Non-Linear Dynamic Scale Factor

As shown on the upper left of Figure 2, the result of the above simple linear scaling method appears not only dark but also appears to lose contrast appearance.

When we apply the linear scale factor, the tone reproduction line can be understood to be rotated at the end of the line in the DlogE domain. We move this rotating
point to the averaged log-luminance $\mu$ — based on the initial assumption — by subtracting the mean $\mu$ before scaling, and then adding it back in the $\text{DlogE}$ domain. We then replace the linear scaling factor with a non-linear function. Many psychophysical experiments [32, 33, 37] tell us that human vision has a sigmoidal response to given stimuli (luminance) in the logarithmic domain. This means that human vision has a Gaussian-shaped sensitivity, which is the first derivative of its sigmoidal response (cumulative sensitivity) to logarithmic luminance. This also corresponds to the characteristic curve of standard photographic materials. Therefore, we do a Gaussian-weighting of the scale factor $k_1$ such that it has a peak at the averaged log-luminance $\mu$ and that it has its minimum at $k_1$ (see Figure 3). This new Gaussian-weighted scale factor $k_2(L)$ depends on the log-luminance $L = \log L_0(x, y)$ and has the range of $k_1 \leq k_2(L) \leq 1.0$. This non-linear scale factor is computed as:

$$
k_2(L) = (1 - k_1) w(L) + k_1,
$$

(3)

$$
w(x) = \exp \left( -\frac{1}{2} \frac{(x-\mu)}{\sigma^2} \right), \quad \sigma = \frac{d_0}{c_1},
$$

(4)

where $\sigma$ is the ratio of the dynamic range $d_0$ of the log-luminances of the HDR image to the user-parameter $c_1$. This adjusts the shape of Gaussian fall-off within the width of its characteristic curve. The parameter $c_1$ influences the resulting brightness and local details of the tone-mapped image. We have found that $c_1 \approx 3.0$ is the maximum level that can compress contrast without losing detail in the bright area of images (see Section 3.2).

The final non-linear mapping function is as follows (including the rotation around $\mu$):

$$
L_1(x, y) = \exp \left[ c_2 k_2 \left( \log L_0(x, y) - \mu \right) + \mu \right]
$$

(5)

We also introduce a parameter $c_2$, which will be referred to as the efficiency factor, which scales the intensity of non-linear weighting. Even though the display signal depth has the dynamic range of 2.4 (1:256), the actual dynamic range of display luminance is often lower than that of signals (e.g., Apple Cinema HD Display has a measured dynamic range of only 2.01). Therefore, the dynamic range of an HDR radiance map should be compressed more than that of the display signal depth. To combat this problem, we add the parameter $c_2$, which is 0.84 ($=2.01/2.4$) for our specific display. However, based on testing other displays, we realized that other CRT or LCD displays may produce a lower dynamic range than the Apple one such that the $c_2$ parameter should be set to lower than the above for general purpose. We have found that $c_2 \approx 0.5$ works for a wide variety of images and displays that we tested (see results section 3.2).

In order to apply our tone reproduction method to color HDR images, we first transform HDR RGB primaries into CIE XYZ coordinates by using the standard transformation matrix [38] (without gamma correction as we start with linear values), based on the assumption that the HDR RGB primaries have the same characteristics as sRGB primaries. The Y coordinate of CIE XYZ is used as the luminance input value for the proposed tone reproduction operator. After obtaining the mapped luminance layer, we scale the X and Z channels by the ratio of mapped luminance to original luminance in a similar way to Schlick [6] in order to preserve the ratio of the three primaries:

$$
C_1(x, y) = \frac{L_1(x, y)}{L_0(x, y)} \times C_0(x, y).
$$

(6)

where $L_1(x, y)$ is the mapped luminance; $L_0(x, y)$ is the original luminance; $C_1(x, y)$ are the X and Z coordinates of the mapped image; and $C_0(x, y)$ are the X and Z coordinates of source image. The Y channel is merely replaced by the mapped luminance to save computational cost and zero luminances of $L_0(x, y)$ are excluded from the calculation.

### 2.2 Inverse Display Characterization

After obtaining the tone-mapped radiance map, we use the international specification for the sRGB color space [38] to map the LDR radiance map into the display color space (CIE XYZ values are transformed into sRGB signals through the inverse transform matrix and gamma correction, corresponding to $\gamma = 2.2$ including linear ramp for dark value [38]).

In order to optimize the dynamic range of the display, we compute the histogram of the tone-mapped image and stretch the pixel levels between 1% and 99% to the range of display signals (effectively clamping values below 1% and above 99% and renormalizing to the 0%-100% range).

### 3 Results

In following section, we will present results of our proposed method in terms of efficiency, response to variation of the parameters, and robustness of its performance through comparison with others.

#### 3.1 Efficiency

We tested a large number of different photographic and computer-generated HDR images, all of which were tone-mapped into 8-bits sRGB. The results can be seen throughout this section. Images look natural, colors are vivid, and
contrast appearance is preserved well. Note that all images were computed with the same parameter set, i.e., \( c_1 = 3.0 \) and \( c_2 = 0.5 \). The computation time of the church image (Figure 2) was about 0.56 seconds for a 768 \( \times \) 512 image on a 2.0GHz Pentium 4 with unoptimized C++ code.

The doll image of 922 \( \times \) 901 in Figure 8 (top-right) took approximately 1.00 second on the same machine. See a comparison with other algorithms in Table 1.

<table>
<thead>
<tr>
<th>Methods</th>
<th>FBF</th>
<th>GC</th>
<th>SBC</th>
<th>LPR</th>
<th>ALM</th>
<th>OPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>74.2</td>
<td>17.6</td>
<td>30.7</td>
<td>5.3</td>
<td>3.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 1. Comparison of computation times (in seconds), which were produced with fast bilateral filtering (FBF) [15], gradient control (GC) [13], subband compressing (SBC) [18], local photographic reproduction (LPR) [14], adaptive logarithmic mapping (ALM) [21], and our proposed (OPR). We used the authors’, Mantuik and Krawczyk’s [39], and Reinhard et al.’s [40] codes.

### 3.2 Parameters

Figure 4 shows the results of varying parameter \( c_1 \) from Equation 4. This parameter was tested in intervals of 2.0 as (a) 1.0, (b) 3.0, (c) 5.0, and (d) 7.0. We have found that setting the parameter \( c_1 \approx 3.0 \) works well. Of course, a user can change the parameter to their liking. The important thing to note, however, is that the same parameter setting will achieve consistent results across different images.

![Figure 4](image1.png)

Figure 4. Variation of parameter \( c_1 \). We increased the parameter \( c_1 \) for the standard deviation of the Gaussian in intervals of 2.0 as (a) 1.0, (b) 3.0, (c) 5.0, and (d) 7.0. \( c_1 \approx 3.0 \) is a good compromise between contrast compression and loss of detail in bright areas. Of course, a user may have a slightly different individual preference. **Inset:** plot of weighting \( w(x) \).

The effect of varying the efficiency factor \( c_2 \) in Equation 5 is shown in Figure 5. The parameter was varied in intervals of 0.3 as (a) 0.2, (b) 0.5, (c) 0.8, and (d) 1.1. As mentioned in the previous section, \( c_2 \) should be varied depending on the actual dynamic range of the employed display device. \( c_2 = 0.2 \) is suitable for an extremely low dynamic range device, whereas \( c_2 = 1.1 \) is a better choice for higher dynamic range devices. We use \( c_2 \approx 0.5 \) for all our results (see Section 2.1 for technical details). As noted before, the user’s personal taste may require slightly higher or lower values.

### 3.3 Robustness

As stated before, it is beneficial for tone reproduction operators not to require any per-image parameter tweaking. To demonstrate how default parameters influence tone reproduction, we have tested two global tone reproduction operators (Drago et al. [21] and Reinhard and Devlin [22]) as well as ours on a set of images. The algorithms all require similar computation time but may reproduce rather different images with their default parameter sets. The results

![Figure 5](image2.png)

Figure 5. Variation of parameter \( c_2 \). We increased the parameter \( c_2 \) for the efficiency factor in 0.3 intervals as (a) 0.2, (b) 0.5, (c) 0.8, and (d) 1.1. \( c_2 = 0.5 \) works well, but this parameter can be varied depending on the display device and user preference. **Inset:** histogram of each image.

![Figure 6](image3.png)

Figure 6. Comparison of contrast. Left and middle: the results of Drago et al. [21] with different bias settings. Right: the result of our proposed method with default parameters. **Inset:** histogram of each image.
The proposed Adaptive logarithmic mapping Photoreceptor method

Figure 7. Consistency test of three global tone reproduction operators with its default parameters (top to bottom): our method, adaptive logarithmic mapping [21], and the photoreceptor method [22]. Inset: histogram of each image. Images courtesy of Industrial Light & Magic and Greg Ward.

can be found in Figure 7. As can be seen, adaptive logarithmic mapping [21] and photoreceptor [22] methods produce inconsistent results for the same default parameter set — some images are overly bright or dark while others look fine. In certain cases, see Figure 6, it may not even be possible for some previous tone-mappers to achieve good results even after parameter tweaking. Our proposed method shows consistent results across the set of images (photographic and computer-generated) without any need for parameter tweaking. More results of our method can be seen in Figure 8.

4 Conclusions

We have proposed a new global tone reproduction method that efficiently compresses high dynamic range images. Our experiments show that we achieve consistent results with a fixed set of parameters for a large variety of images. This is especially beneficial for applications that cannot afford parameter tweaking. In the future, we would like to try our algorithm on HDR video and test its temporal coherence. Furthermore, a graphics processing unit (GPU) implementation would be an interesting direction of research.

References


Figure 8. Consistency test with default parameters of our proposed method. Our method produces plausible and consistent images without any user intervention, i.e., without per-image parameter tweaking (as described in Section 2). Images courtesy of Martin Cadik, Cornell University, Paul Debevec, Yuanzhen Li, Dani Lischinski, Industrial Light & Magic, Jack Tumblin, and Greg Ward (in alphabet order of surname).