Imaging Technology

High Dynamic Range Image Acquisition and Display

Algot Sandahl

David Robín Karlsson

Linnéa Svensson

October 28, 2021

Linköping University

Abstract

This report describes the theory and implementation of HDR-image acquisition and tone mapping. Four different tone mapping methods were implemented, four global and one local. Everything was implemented in MATLAB. Results from applying HDR and tone mapping to three different exposure bracketing datasets are presented.

Contents

1 Introduction **1**

2 Method **2**

2.1 High Dynamic Range Image Acquisition 2

2.1.1 Exposure Bracketing 2

2.1.2 Estimating the Camera Response Function 2

2.1.3 Producing the Radiance Map 4

2.2 Tone Mapping 4

2.2.1 Tone Mapping using Gamma 4

2.2.2 Reinhard's Global Tone Mapping 4

2.2.3 Drago's Global Tone Mapping 5

2.2.4 Durand's Local Tone Mapping 6

3 Results **8**

3.1 High Dynamic Range Image Acquisition 8

3.2 Tone Mapping 11

4 Discussion and Conclusions **14**

Bibliography **16**

1 Introduction

A common shortcoming of digital photography is that pictures do not correctly portray what our eyes see. A frequent source of this issue is the disconnect between our eyes' and digital cameras' capability of capturing contrast ranges in a scene. It can be solved by high dynamic range (HDR) imaging, a method for capturing vastly different levels of intensities in a scene. However, this is not enough as most display mediums possess a much lower contrast range than what our eyes perceive.

In order to display the HDR image on a medium with lower contrast range, the image needs to be tone mapped. It is a method of compressing a high range of intensities into a lower range, suitable for the medium. This can be accomplished either globally or locally. Global tone mapping uses the same transformation across the entire image, whilst local tone mapping uses different transformations adapted to the local content in different areas of the image.

This report details the theory and implementation of an HDR-workflow using different tone mapping methods.

2 Method

This chapter is divided into two primary parts: the high dynamic range image acquisition part, which explains how an HDR image is obtained; and the tone mapping part, which explains how the obtained HDR image can be displayed using ordinary computer monitors.

2.1 High Dynamic Range Image Acquisition

This section describes the different parts of high dynamic range image acquisition process.

2.1.1 Exposure Bracketing

To overcome the limited dynamic range of an image capture, multiple images are captured with mostly identical settings, but with different exposure times. This results in many images that together contains HDR information of the different parts of the image, but preferably all this information should be present in one single image. This is achieved by estimating the camera response function, and then using it together with the low dynamic range images to create a high dynamic range image, or *radiance map* of the scene.

2.1.2 Estimating the Camera Response Function

The basis of recovering the camera response function from photos taken with different exposure times was presented by Debevec and Malik in 1997 [1] and is based on a property known as reciprocity that states that the exposure is the product of the irradiance E, and the exposure time Δt. This means that for example halving E can be compensated for by doubling Δt. Furthermore, the pixel intensities Z_{ij} in an image is a non-linear function of the exposure:

$$Z_{ij} = f(E_i \cdot \Delta t_j) \tag{2.1}$$

where i is a one-dimensional index over the pixels, and j is an index over the different exposure times. Since we expect a higher exposure to results in a brighter pixel, we can assume that f is monotonic and thus invertible, giving

$$f^{-1}(Z_{ij}) = E_i \cdot \Delta t_j. \tag{2.2}$$

Using the natural logarithm and defining $g = \ln f^{-1}$, this can be reformulated as

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \tag{2.3}$$

where Δt_j and Z_{ij} are known, but E_i and g has to be recovered. The pixel values Z_{ij} can only take a limited number of values (usually 256). Consequently, the function $g(z)$ only needs to be defined in these points. The problem can then be solved as a least squares problem with a regularization term that ensures the smoothness of the function; additionally, a weighting function w is used to emphasize the errors in the middle of the curve. The problem is then to minimize

$$\mathcal{O} = \sum_{i=1}^{N}\sum_{j=1}^{P}\big(w(Z_{ij})\big(g(Z_{ij}) - \ln E_i - \ln \Delta t_j\big)\big)^2 + \lambda \sum_{z=Z_{\min}+1}^{Z_{\max}-1}\big(w(z)\, g''(z)\big)^2 \tag{2.4}$$

where λ determines how much the response curve is smoothed, $Z_{\min}$ and $Z_{\max}$ are the minimum and maximum pixel values, N is the number of pixels, and P is the number of different exposure times. The second second order derivative $g''(z)$ is calculated using a simple finite difference approximation:

$$g''(z) = g(z-1) - 2g(z) + g(z+1). \tag{2.5}$$

The weighting function is a simple hat function:

$$w(z) = \begin{cases} z - Z_{\min} & \text{for } z \leq \frac{1}{2}(Z_{\min} + Z_{\max}) \\ Z_{\max} - z & \text{for } z > \frac{1}{2}(Z_{\min} + Z_{\max}) \end{cases}. \tag{2.6}$$

The catch with this method is that E_i and g can only be determined up to a factor, so to solve it the following somewhat arbitrary constraint is enforced

$$g\left(\frac{Z_{\min} + Z_{\max}}{2}\right) = 0 \tag{2.7}$$

The authors kindly provide a MATLAB function in their paper that solves the least squares problem (2.4) using singular value decomposition. This function was used in our implementation.

An issue when solving the least squared problem is that the number of pixels N in most photos is massive, and the complexity of the problem is on the order of $N \times P + Z_{\min} - Z_{\max}$. Therefore it is solved using a subset of the pixels of the image. For optimal results, this subset of pixels should be sufficiently large, and their pixel values should be uniformly distributed between $Z_{\min}$ and $Z_{\max}$. Our implementation simply vectorizes the image and sample it uniformly.

The description so far has not taken color into account, but the adaptation is quite simple: the procedure is simply performed separately for each channel. There is, however, one drawback: there are now three unknown scaling factors that affect the balance between the different channels.

2.1.3 Producing the Radiance Map

When the camera response function has been estimated, the radiance map of the scene can be constructed. From (2.3) we get

$$\ln E_i = g(Z_{ij}) - \ln \Delta t_j \tag{2.8}$$

for any j, with the important exception that if Z_{ij} is saturated or zero its value cannot be trusted. Additionally Z_{ij} contains noise from the imaging process. To remedy these problems, Debevec [1] proposes a weighted average of the estimates from the different exposures is used to calculate the radiance map:

$$\ln E_i = \frac{\sum_{j=1}^{P} w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^{P} w(Z_{ij})}. \tag{2.9}$$

Note that the weighting function w as described in (2.6) ensures that saturated pixels are not taken into account, and that the most reliable values—those in the middle of the range—are weighted the highest.

2.2 Tone Mapping

This section details the theory and implementation of four different tone mapping methods.

2.2.1 Tone Mapping using Gamma

A simple form of tone mapping is to apply gamma to the normalized input luminance L_w:

$$L_d(x,y) = AL_w^{\gamma}(x,y) \tag{2.10}$$

where γ affects the shape of the curve, and A is a scaling factor that can be used to saturate high values if desirable. The gamma can be applied to the color channels separately or to the luminance. In the latter case the final RGB image can then be calculated by scaling the input channels, for example like this

$$C_{\text{out}} = \left(\frac{C_{\text{in}}}{L_w}\right)^{s} L_d \tag{2.11}$$

where the color saturation can be controlled with s [2].

2.2.2 Reinhard's Global Tone Mapping

A well-known tone mapping method was presented by Erik Reinhard et al. in 2002 [3]. The framework behind their reasoning is the Zone System—a systematic approach for choosing film exposure. The paper also presents a local tone mapping approach based on dodging-and-burning, but we restrict our implementation to the global method.

The first step is to calculate the log-average

$$\bar{L}_w = \exp\left(\frac{1}{N}\sum_{x,y}\log\left(\delta + L_w(x,y)\right)\right) \tag{2.12}$$

where δ is a small constant that prevents the evaluation from hitting the singularity that occurs when $L_w = 0$. This average is used together with a parameter a to scale L_w:

$$L(x,y) = \frac{a}{\bar{L}_w} L_w(x,y). \tag{2.13}$$

Finally, the scaled luminance L is used to calculate the tone mapped value according to

$$L_d(x,y) = \frac{L(x,y)\left(1 + \frac{L(x,y)}{L_{\text{white}}^2}\right)}{1 + L(x,y)} \tag{2.14}$$

where L_{white} is the white point, and is defined as the maximum value of L_d if no overexposure is desired, but can also be set to other values. This step is where the the range compression happens—low values are scaled almost linearly while higher values are increasingly attenuated, this is demonstrated in figure 2.1, together with the effect of the white point. Colors can be handled like described in (2.11).

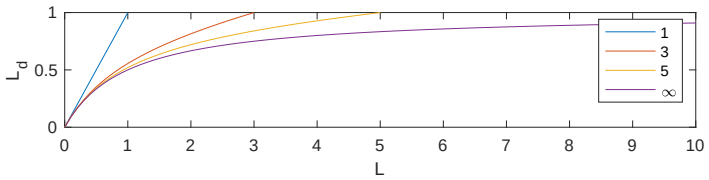


Figure 2.1: *Equation (2.14) shown for different L_{white}*

2.2.3 Drago's Global Tone Mapping

A tone mapping method presented by Drago et al. in 2003 [4] based on imitating the eye's response to light through logarithmic compression of luminance values. A defining characteristic of this method is adaptive logarithm base adjustment based on each pixel's radiance which will ensure better detail and contrast preservation in dark and medium-dark areas. Conversely, high luminance values are maximally compressed. The luminance values L_w in the scene are interpolated between $\log_2(L_w)$ and $\log_{10}(L_w)$.

To ensure smooth interpolation between the logarithm bases a bias function is used as introduced by Perlin et al. in 1989 [5]:

$$\text{bias}_b(t) = t^{\frac{\ln(b)}{\ln(0.5)}} \tag{2.15}$$

By combining (2.15) with the basic arbitrary logarithm base law $\log_{\text{base}}(x) = \frac{\ln(x)}{\ln(\text{base})}$ and intro-

ducing scaling, the displayed value L_d can be calculated as

$$L_d = \frac{L_{dmax} \cdot 0.01}{\log_{10}(L_{wmax} + 1)} \cdot \frac{\ln(L_w + 1)}{\ln\left(2 + \text{bias}_b\left(\frac{L_w}{L_{wmax}}\right) \cdot 8\right)} \tag{2.16}$$

where L_{wmax} is the highest luminance scaled by the average luminance in the scene and L_{dmax} is the highest luminance displaying capability of the displaying medium. The resulting RGB-image is then gamma corrected to adapt to non-linear displaying devices. The gamma correction is defined by the following

$$\begin{cases} L_d \cdot 4.5((\gamma - 2) \cdot 7.5) & L_d < s \\ 1.099 L_d^{\frac{0.9}{\gamma}} - 0.099 & L_d > s \end{cases} \tag{2.17}$$

where $s = \frac{0.018}{((\gamma-2)\cdot 7.5)}$. $\gamma = 2.7$ was used in calculations and was heuristically found to produce the most pleasing results.

2.2.4 Durand's Local Tone Mapping

This tone mapping method follows a basic local tone mapping workflow extended using bilateral low-pass filtering as described by Durand et al. in [6]. Shortcomings when using gaussian filtering, e.g halo-effects, are prevented with the bilateral filter's edge preserving property. The method, using bilateral low-pass filtering, is implemented as:

- Extract luminance map $L(x, y) = 0.3279 \cdot R(x, y) + 0.6557 \cdot G(x, y) + 0.01640 \cdot B(x, y)$
- Extract chrominance channels (elementwise division)
  $R_c(x, y) = R(x, y)/L(x, y)$
  $G_c(x, y) = G(x, y)/L(x, y)$
  $B_c(x, y) = B(x, y)/L(x, y)$
- Apply logarithm to luminance map: $H(x, y) = \log_{10}(L(x, y))$.
- Create base layer by applying bilateral low-pass filter $F_{LP}(x, y)$ on log-luminance channel: $H_L(x, y) = F_{LP}(x, y) * H(x, y)$
- Retrieve high-pass filtered detail layer using difference image: $H_H(x, y) = H(x, y) - H_L(x, y)$
- Calculate gamma for scaling: $\gamma = \frac{\log_{10}(C)}{H_{H,max} - H_{H,min}}$ where C is the contrast range of the displaying medium.
- Contrast reduce base layer: $H'_L(x, y) = \gamma \cdot H_L(x, y)$
- Add base- and detail layer to produce new log-luminance image: $I(x, y) = H'_H(x, y) + H_L(x, y)$
- Attenuate to retrieve new luminance map: $L'(x, y) = 10^{I(x,y)}$ (elementwise attenuation)

- Re-apply colors: (elementwise multiplication)
  $R'(x,y) = R_c(x,y) \cdot L'(x,y)$
  $G'(x,y) = G_c(x,y) \cdot L'(x,y)$
  $B'(x,y) = B_c(x,y) \cdot L'(x,y)$
- Construct HDR-image: $\mathrm{RGB}_{\mathrm{HDR}}(x,y) = R'G'B'(x,y)$

The tone mapping is then defined as

$$\mathrm{RGB}_{\mathrm{LDR}}(x,y) = \min\left(1.0, \max\left(0.0, \left(\frac{\mathrm{RGB}_{\mathrm{HDR}}(x,y)}{10^{\gamma H_{H,max}}}\right)^{0.45}\right)\right) \tag{2.18}$$

3 Results

In this chapter the results are presented.

3.1 High Dynamic Range Image Acquisition

The creation of radiance maps was tried on a few different exposure bracketing data sets (shown in figures 3.1–3.3). The camera response curves recovered from the metal ball data set (figure 3.1) is shown in figure 3.4. These plots also show the estimated irradiance values E_i plotted with the curve. Figure 3.5 shows the recovered radiance maps as false color images, and a limited range of the RGB radiance map.

Figure 3.1: *The metal ball data set (course material).*

Figure 3.2: *The kitchen data set, extracted from the EBSNet data set [7]. Available at https://github.com/wzhouxiff/EBSNetMEFNet*

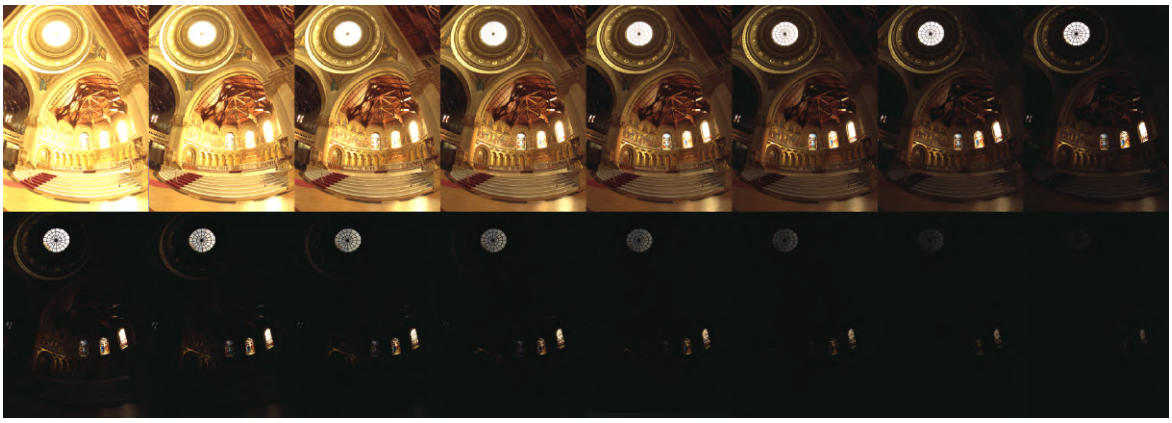


Figure 3.3: *The Memorial Church data set from [1]. Available at https://www.pauldebevec.com/Research/HDR/.*

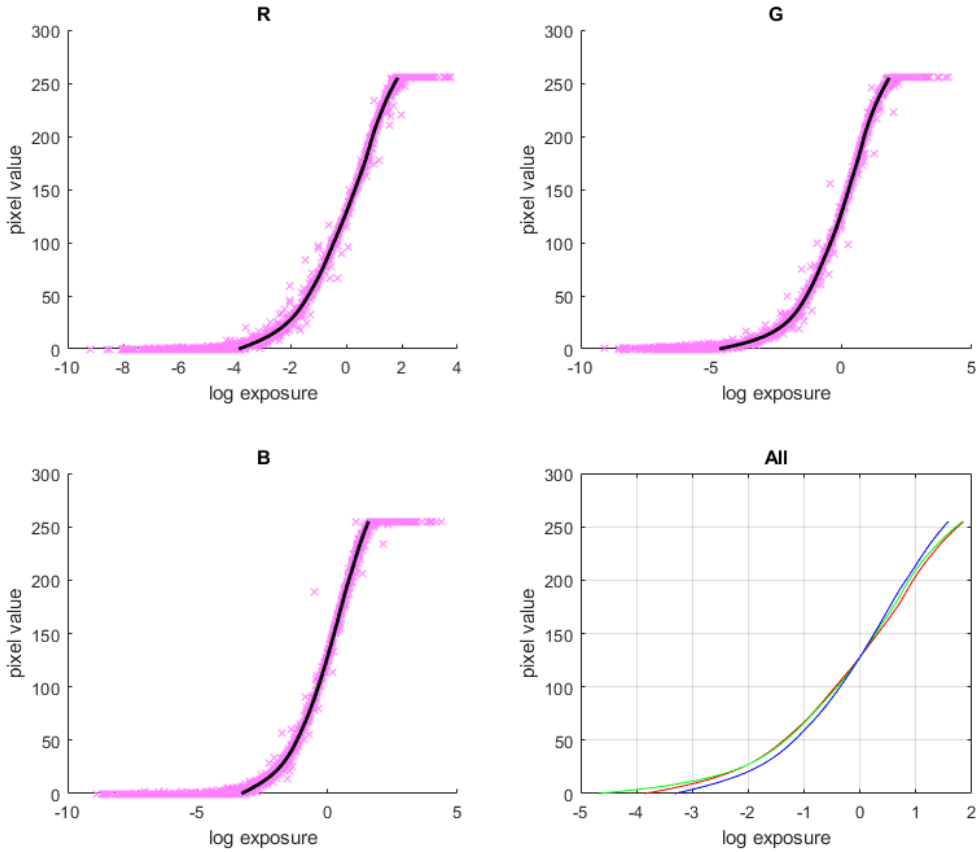


Figure 3.4: *The recovered response curves of the different color channels shown together with the estimated irradiance values from the pixel positions used in the calculations. The data set used is shown in figure 3.1.*

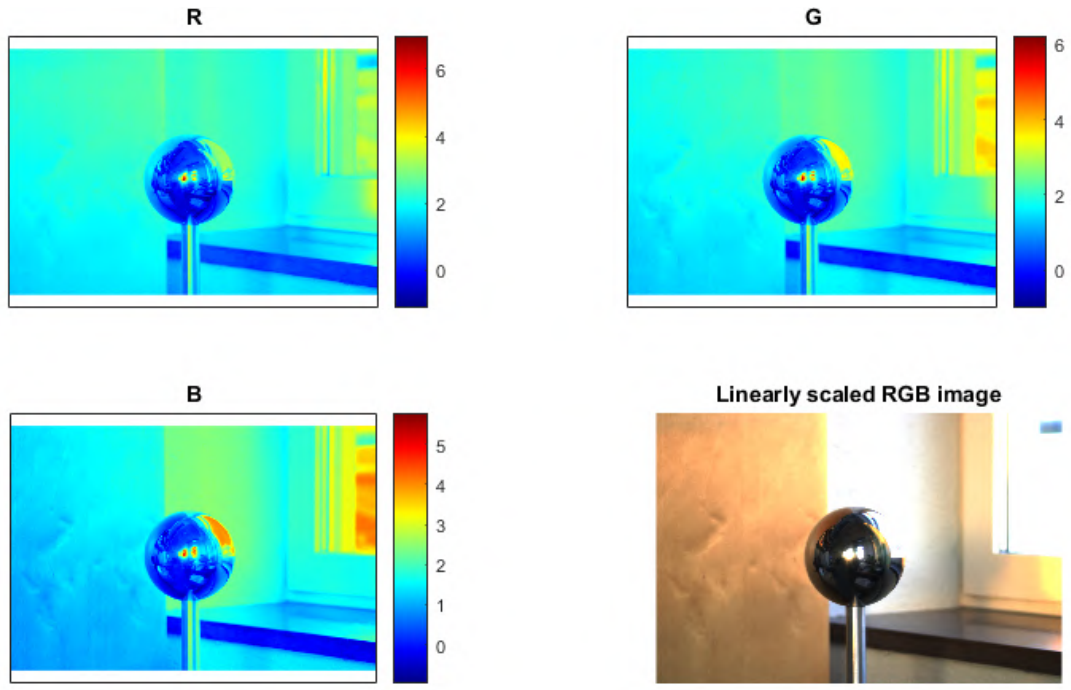


Figure 3.5: *Natural logarithm radiance maps shown for the different color channels and a linear scaling of a limited range of the radiance maps shown as an RGB image.*

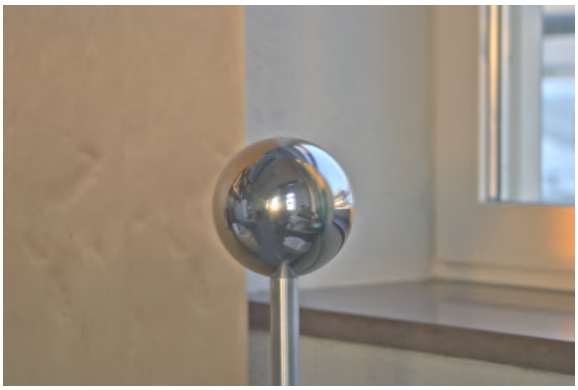
3.2 Tone Mapping

The tone mapped HDR images are presented here. Results from the Stanford Cathedral dataset can be seen in figure 3.6, the metal ball in figure 3.7 and the kitchen in figure 3.8. For all data sets the Reinhard white point is set to $\max L(x, y)$, and the parameter A for the gamma tone mapping is larger than one, resulting in high values being clipped—this is done since the image would otherwise look washed out.

(a) *Durand* **(b)** *Drago* **(c)** *MATLAB*

(d) *Gamma applied to luminance* **(e)** *Reinhard*

Figure 3.6: *The tone mapped images of the Stanford cathedral.*



(a) *Durand*

(b) *Drago*

(c) *MATLAB*

(d) *Gamma applied to luminance*

(e) *Reinhard*

Figure 3.7: *The tone mapped images of the metal ball.*

(a) *Durand*

(b) *Drago*

(c) *MATLAB*

(d) *Gamma applied to luminance*

(e) *Reinhard*

Figure 3.8: *The tone mapped images of the kitchen.*

4 Discussion and Conclusions

The exposure bracketing data sets (figure 3.1–3.3) demonstrates how no single photo captures the full dynamic range of the scenes. Nonetheless, the complete set of images contain all the information needed, and as seen in figure 3.4, the recovered camera curve fits the data well. There are for sure a few outliers, which could possibly be explained by small camera movements between the captures, but the majority of the irradiance values closely follows the curve except at the extremities. Furthermore, the logarithmic radiance maps shown in figure 3.5 exemplifies the high dynamic range that is present in many scenes. Note that as a consequence of how the process works, the recovered values are not absolute. For image display this typically does not matter, but it does introduce one problem in color images: the relation between the color channels are unknown. To accurately recover the colors, a calibration card could be used.

As seen in figures 3.6, 3.7 and 3.8, MATLAB's tone mapping method stands out somewhat from the rest. It manages to reproduce contrasts very well but the colors are not as vibrant as Durand, Drago or Reinhard. However, the contrast in the scene is questionable in some parts of some images. In figure 3.7c, the *halo* around the metal ball, present in the source images, has been amplified and is very visible in comparison to other methods. MATLAB also produces strange shadows in the kitchen dataset which can be seen on some parts of the wall and counter in figure 3.8c.

Drago's, Durand's and Reinhard's methods are all better in reproducing vibrant colors, evident from all images. Between these four, Reinhard's method seems to produce the overall darkest images. However, neither of them managed to produce the contrast MATLAB was able to in the cathedral (fig 3.6). As seen in the metal ball (fig 3.7), Durand's method produced sharper details compared to Drago which is most likely as result of the edge-preserving bilateral filter. Reinhard's method is most similar to Durand's, albeit darker. In the kitchen image (fig 3.8), Durand's method was able to reproduce the nature outside the windows much clearer than the other methods. This is also most likely related to the bilateral filter since the outdoors scene contain a lot of detail, albeit in a relatively smaller space. This time, Reinhard's method is more similar to Drago's in terms of details but is once again generally darker.

There is an argument to be made about potentially outdated technology. The newest source used as reference in this project was published in 2003 and the oldest one in 1989. Digital photography has evolved significantly since 2003 and there is no doubt a more modern approach to HDR and tone mapping is available today. However, since the digital photography industry has some well-

funded competitors these days, there is a risk that the best HDR-imaging method is proprietary and a secret to the world.

Ultimately, which tone mapping operator is the best becomes a question of what qualities are desirable. For example the “dramatic” look and good performance of the Reinhard method might be the best choice for one application, whereas the very effective local contrast compression of the Durand method might improve the visibility of all areas of the image in a way that is worth the performance penalty in another application. For machines with more resources available (like a modern PC), more methods might be viable than for one with lesser (like a cell phone). It all depends on the context and the viewer.

Bibliography

[1] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '97, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 369–378, ISBN: 0897918967. DOI: 10.1145/258734.258884. [Online]. Available: https://doi.org/10.1145/258734.258884.

[2] R. Mantiuk, R. Mantiuk, A. Tomaszewska, and W. Heidrich, "Color correction for tone mapping," *Computer Graphics Forum*, vol. 28, pp. 193–202, Apr. 2009. DOI: 10.1111/j.1467-8659.2009.01358.x.

[3] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, vol. 21, May 2002. DOI: 10.1145/566654.566575.

[4] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum*, vol. 22, no. 3, pp. 419–426, 2003. DOI: https://doi.org/10.1111/1467-8659.00689.

[5] K. Perlin and E. M. Hoffert, "Hypertexture," in *Proceedings of the 16th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '89, New York, NY, USA: Association for Computing Machinery, 1989, pp. 253–262, ISBN: 0897913124. DOI: 10.1145/74333.74359. [Online]. Available: https://doi.org/10.1145/74333.74359.

[6] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," vol. 21, Jul. 2002, pp. 257–266. DOI: 10.1145/566570.566574.

[7] Z. Wang, J. Zhang, M. Lin, J. Wang, P. Luo, and J. Ren, "Learning a reinforced agent for flexible exposure bracketing selection," in *CVPR*, 2020.