The Naturalness of Reproduced High Dynamic Range Images

Martin Čadík, Pavel Slavík
Department of Computer Science and Engineering
Faculty of Electrical Engineering, Czech Technical University in Prague
Karlovo nám. 13, 121 35 Prague, Czech Republic
cadikm@sgi.felk.cvut.cz, slavik@fel.cvut.cz

Abstract

The problem of visualizing high dynamic range images on the devices with restricted dynamic range has recently gained a lot of interest in the computer graphics community. Various so-called tone mapping operators have been proposed to face this issue. The field of tone mapping assumes thorough knowledge of both the objective and subjective attributes of an image. However, there no published analysis of such attributes exists so far. In this paper, we present an overview of image attributes which are used extensively in different tone mapping methods. Furthermore, we propose a scheme of relationships between these attributes, leading to the definition of an overall quality measure which we call naturalness. We present results of the subjective psychophysical testing that we have performed to prove the proposed relationship scheme. Our effort sets the stage for well-founded quality comparisons between tone mapping operators. By providing good definitions of the different attributes, comparisons, be they user-driven or fully automatic, are made possible at all.

1. Introduction

The dynamic range available of visual stimuli in the real world is vast. Accordingly, the human visual system (HVS) is able to adapt over a huge range of luminous intensities through a process known as visual adaptation. However, even though several different computer technologies can produce a high dynamic range luminance[1] maps (images) of synthetic graphics or real scenes, the media used to present these images can only display a few orders of magnitude in luminous intensity. This problem, i.e., the reproduction of high dynamic range images on conventional output devices, is the domain of *tone mapping* methods. A number of different tone mapping methods (or operators) have been proposed in history. Since the advantages and disadvantages of these methods are not immanently clear, a thorough comparison is very desirable. In order to conduct a comparison of tone mapping methods, it is necessary to settle upon a set of *image attributes* by which the images produced by the methods should be judged. Furthermore, the relationships and dependencies between those attributes will influence any comparison significantly and should therefore be analyzed.

In this paper, we attempt to give a list of all important attributes involved in the evaluation of a tone mapping operator, and we show which relationships exist between them. The evaluation of the attributes and their relationships leads to the definition of an overall image quality metric which we call *naturalness*. In short, the naturalness gives an idea how well the tone mapping operator is able to produce realistic images. We encourage our proposition by means of subjective perceptual study. Virtually all well-known tone mapping operators are considered in the study. The results of the study show strong correlations with our proposal.

2. Background

The input for a tone mapping method is a high dynamic range (HDR) image, which is a map of pixels with arbitrary luminance values. The output of a tone mapping method is low dynamic range (LDR) image, which is a map of pixels with values displayable by a particular output device. We distinguish *global* tone mapping methods and *local* tone mapping operators. Global tone mapping methods apply the same transformation to every pixel, while local methods are spatially varying and apply a different scale to different parts of an image. See a review by Devlin [5] for more information on a particular method.

The dynamic range of light encountered in the real world or in predictive global illumination simulations is enormous. The human visual system can face this range by

---

[1] Luminance is the amount of visible light leaving a differential point on a surface in a given differential direction. The standard unit of luminance is candela per square meter ($\mathrm{cd} \cdot \mathrm{m}^{-2} \equiv \mathrm{Nit}$).

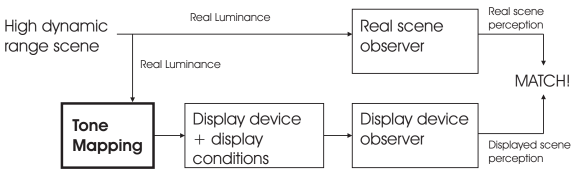

Figure 1. The goal of the tone mapping process - matching the perception of original scene and reproduced image [21].

means of *adaptation mechanisms*. Generally speaking, tone mapping methods try to mimic this retentive behaviour of the human vision. Several formulae were proposed in history to determine the performance of human vision, but we need to be careful in interpreting these formulae. In principle, we can distinguish *threshold* data, where a stimulus is just barely visible, and *suprathreshold* data, valid above a threshold. In the luminance domain, the **threshold**-versus-intensity (TVI) functions show the relationship between just noticeable differences (JND) in intensity and the background illumination level. Over a wide range of illumination levels, the visual system obeys *Weber's law*, which states that the size of the JND is a constant proportion of the background level. On the other hand, Stevens and Stevens [13] proposed a linear scale for brightness that is valid for **suprathreshold** luminance levels. Following Stevens, brightness is expressed in brils, where one *bril* is equivalent to the brightness induced by a 1 second exposure to a 5° white target of $\frac{1}{\pi \times 10^2}\ \mathrm{cd} \cdot \mathrm{m}^{-2}$ (i.e., 1 microlambert). They demonstrated that subjective brightness grows as a power function of luminance (*power law*). See the survey by Ferwerda [10] for more information on human aspects of spatial vision.

3. Naturalness

In this section, we motivate and describe a measure which is useful to determine the performance of a particular tone mapping operator.

Viewing a tone-mapped image should produce a subjective experience that corresponds well with viewing the real scene, as illustrated in Fig. 1. This means that the reproduced image should look as *natural* as possible, i.e., the naturalness is the goal of tone mapping methods. However, naturalness has not been defined so far. If we want to assess the performance of a particular tone mapping operator, we must define the naturalness of an image in some sense. Unfortunately, naturalness is a very subjective quantity.

In order to be able to asses the naturalness of an output image, we propose to define the naturalness as a result of a network of other image attributes, namely the *brightness, contrast, colour reproduction, reproduction of details, simulation of glare, visual acuity* and *artifacts*, see Fig. 3. This implies that the naturalness need not be judged directly, but can be seen as a composition of these attributes. If we are able to measure these attributes somehow, and quantify the relationships between them, we are also able to assess the naturalness as a function of these attributes.

Our goal is therefore to find out how different image attributes influence the naturalness. Since some of the proposed attributes are not independent (as we explain later), we also show a scheme of relationships between them (Fig. 3), and verify this scheme by means of perceptual subjective study.

4. Image attributes

As opposed to radiometric or photometric units, the image attributes used in tone mapping are not well defined and consistent. In this section we describe particular image attributes for tone mapping, and we list typical tone mapping methods that attempt to reproduce them correctly.

Brightness is defined by the CIE as the *attribute of a visual sensation according to which an area appears to emit more or less light.* In fact, brightness is a quantity that measures the subjective sensation produced by a particular luminance, i.e., the brightness is the *perceived luminance*.

Tone mapping methods aim to preserve the overall impression of brightness. Stevens and Stevens proposed an expression for the apparent brightness, see Section 2. Although the expression is convenient for simple targets, the overall brightness of an image is more complex. Thanks to the light adaptation of the human visual system, people are quite insensitive to absolute levels of illumination. The perceived dimension of brightness is due largely to the *contrast* of one region with a surrounding region [13]. When two small adjacent patchs of the retina are stimulated by light, each patch not only responds to the light, but also *inhibits* the response of its neighbor. By providing a dark surrounding to any patch, we reduce the inhibition it receives, and its response to the same amount of light is increased. This phenomena is called *simultaneous contrast* and depends strongly on the *semantics* of the scene, see Fig. 2.

Tumblin and Rushmeier [21] proposed an operator that attempts to preserve the overall impression of brightness using a mapping function that is based on the model by Stevens and Stevens. This mapping function matches the brightness of a real world luminance to the brightness of a display luminance. Bright scenes therefore appear bright, while dark ones appear dark.

Contrast: image contrast is defined in different ways, but it is usually related to variations in image luminance. *Michelson's definition* [16] is as follows: $C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$, where

Figure 2. The simultaneous contrast effect: the two small rectangles are the same shade of gray, but the left one appears lighter than the right one.

$L_{\max}$ and $L_{\min}$ are the maximum and minimum *luminance* values, respectively. Michelson's definition is sometimes used when assessing the *global contrast* of the image. However, using it for global contrast is rather deceptive, because it is strongly affected by possibly random, irrelevant pixels with extreme values.

Weber's definition is used to measure the *local contrast* of a single target against a uniform background. The definition is as follows: $C = \frac{\Delta L}{L}$, where ΔL is the difference of the target luminance from the background luminance. Weber's definition is sometimes used to measure local contrast in the image, but if there are many targets present, the initial assumptions do not hold and interpreting the result is rather tricky. Peli [16] proposed another definition of local contrast that is suitable for *complex images*. Peli's *local band-limited contrast* assigns a contrast value to every point in the image as a function of the spatial frequency band. For each frequency band, the contrast is defined as the ratio of the bandpass-filtered image at that frequency to the low-pass image filtered to an octave below the same frequency. This produces a multiscale representation of the effective image contrast.

Ward's [23] initial tone mapping operator focuses on the preservation of *perceived contrast*. This method transforms input luminance to output luminance using a constant scaling factor. Since Ward's operator scales image intensities by a constant, it does not change scene contrasts for display. Almost the same principle of contrast preservation is exploited also in the other operators [11, 24]. Advanced local tone mapping methods (e.g., the method by Reinhard et al. [18] or by Ashikhmin [1]) are based on a multi-resolution decomposition of the image and define contrast in a way similar to Peli.

Reproduction of colours: the sensation of colour is an important aspect of the human visual system, and correct reproduction of colours can help to increase the realistic appearance of the output image. One important feature of the human visual system is the capacity to see the level of colours in a bright environment. This ability, measured as colour sensitivity, is reduced in dark environments, as the light sensitive rods take over for the colour-sensitive cone system. As the luminance level is raised, the cone system becomes active and colours begin to be seen beginning with the long wavelength reds and progressing toward the middle wavelength greens. Only at relatively high luminances, short wavelength blue targets begin to appear coloured.

The tone mapping operator by Ferwerda et al. [11] captures changes in threshold colour appearance by using separate TVI functions for rods and cones and interpolation for the mesopic luminance range. Ward et al. [24] used a very similar approach in their work. Pattanaik et al. [15] proposed a comprehensive multi-scale model that accounts for both changes in threshold colour discriminability and suprathreshold colourfulness. Using opponent colour processing, the model is able to handle changes in chromatic and luminance-level adaptation as well. In their recent work, Reinhard and Devlin [17] adapted a computational model of photoreceptor behavior that incorporates a chromatic transform that allows the white point to be shifted.

Reproduction of details: conventional output devices have a limited dynamic range and thus a limited span of reproducible subtle contrasts, which translates directly in the amount of reproducible detail. The reproduction of details is critical mainly in very dark and bright areas, because truncation of values occurs most frequently in these areas. The simplest methods to adjust scene intensities for display will usually reduce or destroy important details and textures. On the other side, the effort to reproduce details well is a potential cause of *artifacts.*

Several tone mapping operators focus especially on the reproduction of details. Tumblin and Turk's LCIS operator [22] builds a hierarchy using multiple instances of a low curvature image simplifier. A high detail, low contrast image is constructed from this hierarchy by compressing only the large features and adding back all small details. The idea of compressing just the large features and then adding back subtle non-compressed details is also used in the operators based on the bilateral [7] and trilateral filter [3].

A rather different approach was presented by Ward [24]. Ward's operator based on histogram adjustment aims to preserve *visibility*, where visibility is said to be preserved if we can see an object on the display if and only if we can see it in the real scene. Ward's operator does not strive to reproduce all the details available, but exploits the limitations of human vision to reproduce just the *visible* details.

Visual acuity simulation: visual acuity is the ability of the human visual system to resolve spatial detail. The visual acuity diminishes in dark environments, since cones are not responding to such low light levels. We can simulate this phenomenon to enhance the *naturalness* of the image. Shaler [10] measured the relationship between background luminance and foveal acuity. The original data by

Shaler can be approximated with the following functional fit: $R = 17.25 \cdot \arctan(1.4 \log_{10}(L_a) + 0.35) + 25.72$, where R is the visual acuity in cycles/degree and L_a is the local adaptation luminance in $\mathrm{cd} \cdot \mathrm{m}^{-2}$. This function can be used to predict the visibility of scene *details* at different levels of illumination.

Ferwerda et al. [11] proposed a global blurring function based on a single adaptation level to simulate changes in spatial acuity. Ward et al. [24] applied a similar blurring function locally, according to the foveal adaptation. Using such a method, it is possible to extend virtually any operator to take visual acuity into account.

Glare simulation: owing to the scattering of light in the human eye, and due to diffraction in the lens, effects called *bloom* (the "glow" around bright objects) and *flare lines* (radial streaks emanating from the center of the light source) are seen around very bright objects. Another effect called *lenticular halo* (a set of coloured, concentric rings, surrounding the light source and distal to the ciliary corona) is due to the diffraction of the circular optical grating formed by the radial fibers at the periphery of the crystalline lens. All these real-world phenomena are commonly referred to as *glare effects* [20].

Since the dynamic range of traditional output devices is not sufficient to evoke such phenomena, we must simulate the human response artificially to improve the *naturalness* of the image. Several works on the simulation of glare effects [20, 12] have been published, and some tone mapping operators [2, 24] have included glare simulation as well. Using subjective testing results, Spencer et al. [20] verified that the simulation of glare can substantially increase the apparent *brightness* of light sources in digital images.

Artifacts: as a consequence of tone mapping, artifacts may appear in the output image. The artifacts are degrading the quality and therefore the *naturalness* of the output image.

Simple local tone mapping operators [2, 19] typically exhibit halo artifacts. These *halo artifacts* are caused by contrast reversals, which often happens for small bright features or sharp high-contrast edges, where a bright feature causes strong attenuation of the neighbouring pixels, surrounding the feature or high-contrast edge with a noticeable dark band or halo.

5. Attribute relationships

In order to clarify the relationships between the attributes, we propose the scheme shown in Fig. 3. As we can see from this scheme, the **naturalness**, our measure of overall image quality, is determined by all the attributes. It depends strongly on the overall perceived *brightness*, i.e., dark scenes should appear dark, bright scenes should appear bright. Apparent *contrast* should also be reproduced well to make the result natural. The reproduction of *details* or rather the reproduction of *visibility* of objects is certainly essential to make the output image appear natural. Furthermore, since we are typically facing a limited display gamut, the reproduction of *colour* is an important factor for naturalness as well. The simulation of *visual acuity* loss can significantly improve the naturalness of dim or night scenes, while the simulation of *glare* can enhance the naturalness of the dark scenes with strong light sources. There is no doubt that the presence of ugly *artifacts* degrades naturalness.

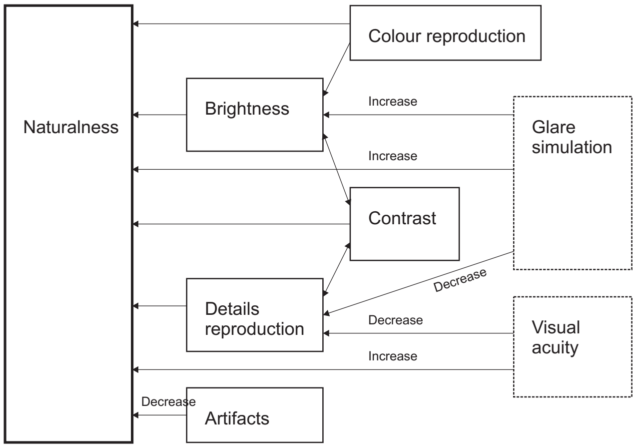

Figure 3. The relationships between image attributes.

Brightness: the perception of brightness is affected greatly by the *contrast arrangement* (i.e., by the semantics of an image, see Fig. 2). Fairchild [8] studied the effect of image contrast on the perceived brightness and concluded that the brightness sometimes increases with *contrast*. The perception of brightness is modified by the *colour representation* as well. The simulation of colour appearance at scotopic levels of illumination can substantially change the perceived brightness. Finally, the *simulation of glare* plays an important role for the brightness perception. The glare simulation increases the apparent brightness of light sources.

The reproduction of **details** is strongly affected by the simulation of the *visual acuity*. Since there are available data that represent the visual acuity (e.g., Shaler's curve), these data place limits on the reproduction of fine details, and may also be utilized to verify the naturalness of detail reproduction. Furthermore, the *visibility preservation* diminishes the reproduced details using a threshold function (e.g., the threshold versus intensity curve, TVI). The simulated *glare* can obscure otherwise reproducible details

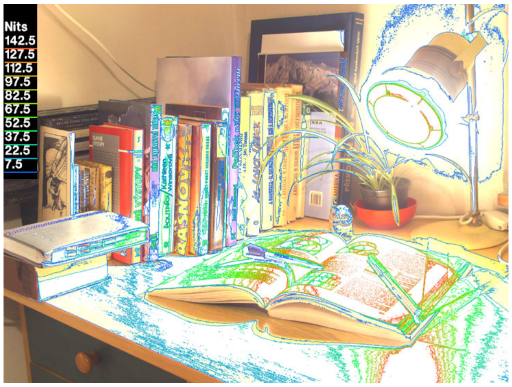

Figure 4. Iso contour plot of the test scene. Maximum luminance is 100195 cd/m^2, minimum luminance is 0.0988018 cd/m^2.

near strong light sources.

In the scheme of relationships (Fig. 3), we can identify attributes that represent **limitations** of the human visual system (depicted in dashed boxes): the simulations of glare and visual acuity and (in part) the reproduction of colour (in the sense of simulation of the scotopic vision). These attributes enhance the naturalness of the output image, but are not desirable when the goal is different, for example when we aim to reproduce as many details as possible.

6. Subjective perceptual study

We have conducted a subjective perceptual study to encourage the proposed idea of the image naturalness measure and to uncover the correlations to the image attributes.

6.1. Subjective testing setup

In the experiment, we wanted to simultaneously present an original HDR image and the tone mapped image to a human observer and ask him to rate the naturalness and the other image attributes. Since HDR output display devices are still unattainable or unsatisfactory, we have arranged a typical *real* indoor HDR scene, see Fig. 4. We acquired the series of 15 images of the scene using a digital camera with varying exposition from a locked-down tripod. The HDR radiance map was recovered from the recorded series using the method by Debevec and Malik [4]. The HDR map was then transformed using 14 essential tone mapping operators to the 14 output LDR images [2]. We involved following tone mapping operators into our experiment – the methods proposed by: Ashikhmin [1], Drago et al. [6], Durand and Dorsey [7], Fattal et al. [9], Chiu et al. [2], Choudhury and Tumblin [3], Tumblin and Turk [22], Pattanaik and Yee [14], Reinhard et al. [18], Schlick [19], Tumblin and Rushmeier's revised operator [21], Ward [23], Ward et al. [24], and a linear mapping.

The sequence of the 14 LDR tone mapped images presented the input visual stimuli for an observer. In the perceptual test arrangement the subject was able to observe both the real scene and a LDR image of the scene displayed on the calibrated CRT monitor. The total number of 9 subjects were asked to express the naturalness, the overall brightness, the overall contrast, the reproduction of details, the visibility, and the reproduction of colours for a particular LDR image by ratings (on the scale 0–10, where 0 represents the worst result, while 10 is the best). Subjects had normal or corrected-to normal vision and were non-experts in the field of computer graphics.

6.2. Test results

Absolute average values of subjective opinion scores judge the following methods to produce the most natural LDR images for our input HDR stimuli (average scores in parentheses, 10 being the best): Choudhury and Tumblin [3] (6.83), Ward et al. [24] (7.66), Tumblin and Turk [22] (7.83), and the best rated was: Reinhard et al. [18] (7.98).

Next, we used the standard (Pearson) correlations for the numerical evaluation of the subjective data. We computed the correlation coefficients between the naturalness and the other image attributes for each of the 14 LDR images (corresponding to a particular tone mapping operator). The correlation coefficients were then averaged and the results summarized in Table 1. The values of correlation coefficients signify quite a strong relations between the naturalness and the image attributes, particularly the overall *brightness* and the overall *contrast*. This result supports the proposed scheme depicted in Fig. 3.

We are acquainted with a perceptual experiment by Yoshida et al. [25]. The authors evaluated 7 tone mapping operators using a direct comparison between the appearance of two real architectural scenes and images of these two scenes displayed on a monitor. The results of the Yoshida's experiment exhibit rather weaker correlations of brightness and detail reproductions to naturalness. Since both ours and Yoshida's experiment has a limited span of input scenes, more thorough and extensive testing is needed to confirm the outlined results. The discrepancy in results is probably caused by the number and nature of input LDR images as well, since we use much more tone mapping operators in our experiment.

---

[2] All the images are available on the web pages of the project: `http://www.cgg.cvut.cz/~cadikm/tmo`

	Contrast	**Brightness**	**Repr. of details**	**Visibility**	**Repr. of colours**
Naturalness	0.7402	0.8136	0.6245	0.6246	0.6156

Table 1. Average correlation coefficients between image naturalness and other image attributes.

7. Conclusions

The field of tone mapping assumes thorough knowledge of various image attributes and approaches from diverse scientific areas. In this paper, we have presented an overview of image attributes for tone mapping. Since the attributes are intimately related, we have proposed a scheme of relationships between them. Moreover, we have proposed a measure called *naturalness*, which can be expressed as a weighted sum of these attributes. We have verified the proposed ideas by means of the subjective psychophysical experiments.

The presented overview of image attributes is helpful for getting into the tone mapping field, or when implementing or developing a new tone mapping operator. On the other hand, the identification of the relationships between the attributes is very useful for the subjective comparison of tone mapping methods. It also simplifies the comparison process by reducing the actual number of attributes that can be used to evaluate a tone mapping operator. Finally, it represents the initial effort to the design a truthful, objective comparison metric for high dynamic range images.

References

[1] M. Ashikhmin. A tone mapping algorithm for high contrast images. In *13th Eurographics Workshop on Rendering*. Eurographics, 2002.

[2] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Proc. of GI*, pages 245– 253, 1993.

[3] P. Choudhury and J. Tumblin. The trilateral filter for high contrast images and meshes. In *EGRW ’03: Proc. of 14th Eurographics workshop on Rendering*, pages 186–196, 2003.

[4] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In T. Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, volume 31 of *Annual Conference Series*, pages 369–378. AW, Aug. 1997.

[5] K. Devlin, A. Chalmers, A. Wilkie, and W. Purgathofer. Star: Tone reproduction and physically based spectral rendering. In D. Fellner and R. Scopignio, editors, *State of the Art Reports, Eurographics 2002*, pages 101–123, 2002.

[6] F. Drago, K. Myszkowski, T. Annen, and N.Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22(3), 2003.

[7] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In *Proc. of 29th conf. on C. G. and inter. techn.*, pages 257–266. ACM Press, 2002.

[8] M. D. Fairchild. *Color Appearance Models*. Reading, Mass.: Addison-Wesley, 2nd edition, 2005.

[9] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In *Proc. of 29th conf. on C. G. and inter. techn.*, pages 249–256. ACM Press, 2002.

[10] J. A. Ferwerda. Fundamentals of spatial vision. In V. Interrante, editor, *Applications of visual perception in computer graphics*, pages 1– 27. Course 32, SIGGRAPH ’98, 1998.

[11] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual adaptation for realistic image synthesis. *Computer Graphics*, 30:249–258, 1996.

[12] M. Kakimoto, K. Matsuoka, T. Nishita, T. Naemura, and H. Harashima. Glare generation based on wave optics. In *Pacific Conference on Computer Graphics and Applications*, pages 133–142, 2004.

[13] S. E. Palmer. *Vision science – photons to phenomenology*. The MIT Press, Cambridge, 3rd edition, 2002.

[14] S. Pattanaik and H. Yee. Adaptive gain control for high dynamic range image display. In *SCCG ’02: Proc. of 18th spring conference on C. G.*, pages 83–87. ACM Press, 2002.

[15] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proc. of 25th conf. on C. G. and inter. techn.*, pages 287–298. ACM Press, 1998.

[16] E. Peli. Contrast in complex images. *Journal of the Optical Society of America A*, 7(10):2032–2040, October 1990.

[17] E. Reinhard and K. Devlin. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, 2005.

[18] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. In *Proc. of 29th conf. on C. G. and inter. techn.*, pages 267–276, 2002.

[19] C. Schlick. An adaptive sampling technique for multidimensional ray tracing. In *Photorealistic Rendering in Computer Graphics*, Springer Verlag, pages pp. 21–29, 1994.

[20] G. Spencer, P. Shirley, K. Zimmerman, and D. P. Greenberg. Physically-based glare effects for digital images. In *Proc. of 22nd conf. on C. G. and inter. techn.*, pages 325–334, 1995.

[21] J. Tumblin and H. Rushmeier. Tone reproduction for realistic images. *IEEE Comput. Graph. Appl.*, 13(6):42–48, 1993.

[22] J. Tumblin and G. Turk. Low curvature image simplifiers (LCIS). In *SIGGRAPH 99 Conference Proceedings*, Annual Conference Series, pages 83–90. Addison Wesley, 1997.

[23] G. Ward. A contrast-based scalefactor for luminance display. *Graphics Gems IV*, pages 415– 421, 1994.

[24] G. Ward Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. on Vis. and C. G.*, 3(4):291–306, 1997.

[25] A. Yoshida, V. Blanz, K. Myszkowski, and H.-P. Seidel. Perceptual evaluation of tone mapping operators with real-world scenes. *Human Vision & Electronic Imaging X, SPIE*, 2005.