Photographic Tone Reproduction for Digital Images

Erik Reinhard
University of Utah

Michael Stark
University of Utah

Peter Shirley
University of Utah

James Ferwerda
Cornell University

Abstract

A classic photographic task is the mapping of the potentially high dynamic range of real world luminances to the low dynamic range of the photographic print. This tone reproduction problem is also faced by computer graphics practitioners who map digital images to a low dynamic range print or screen. The work presented in this paper leverages the time-tested techniques of photographic practice to develop a new tone reproduction operator. In particular, we use and extend the techniques developed by Ansel Adams to deal with digital images. The resulting algorithm is simple and produces good results for a wide variety of images.

CR Categories: I.4.10 [Computing Methodologies]: Image Processing and Computer Vision—Image Representation

Keywords: Tone reproduction, dynamic range, Zone System.

1 Introduction

The range of light we experience in the real world is vast, spanning approximately ten orders of absolute range from star-lit scenes to sun-lit snow, and over four orders of dynamic range from shadows to highlights in a single scene. However, the range of light we can reproduce on our print and screen display devices spans at best about two orders of absolute dynamic range. This discrepancy leads to the *tone reproduction* problem: how should we map measured/simulated scene luminances to display luminances and produce a satisfactory image?

A great deal of work has been done on the tone reproduction problem [Matkovic et al. 1997; McNamara et al. 2000; McNamara 2001]. Most of this work has used an explicit perceptual model to control the operator [Upstill 1985; Tumblin and Rushmeier 1993; Ward 1994; Ferwerda et al. 1996; Ward et al. 1997; Tumblin et al. 1999]. Such methods have been extended to dynamic and interactive settings [Ferwerda et al. 1996; Durand and Dorsey 2000; Pattanaik et al. 2000; Scheel et al. 2000; Cohen et al. 2001]. Other work has focused on the dynamic range compression problem by spatially varying the mapping from scene luminances to display luminances while preserving local contrast [Oppenheim et al. 1968; Stockham 1972; Chiu et al. 1993; Schlick 1994; Tumblin and Turk 1999]. Finally, computational models of the human visual system can also guide such spatially-varying maps [Rahman et al. 1996; Rahman et al. 1997; Pattanaik et al. 1998].

Figure 1: *A high dynamic range image cannot be displayed directly without losing visible detail using linear scaling (top). Our new algorithm (bottom) is designed to overcome these problems.*

Using perceptual models is a sound approach to the tone reproduction problem, and could lead to effective hands-off algorithms, but there are two problems with current models. First, current models often introduce artifacts such as ringing or visible clamping (see Section 4). Second, visual appearance depends on more than simply matching contrast and/or brightness; scene content, image medium, and viewing conditions must often be considered [Fairchild 1998]. To avoid these problems, we turn to photographic practices for inspiration. This has led us to develop a tone reproduction technique designed for a wide variety of images, including those having a very high dynamic range (e.g., Figure 1).

2 Background

The tone reproduction problem was first defined by photographers. Often their goal is to produce realistic "renderings" of captured scenes, and they have to produce such renderings while facing the

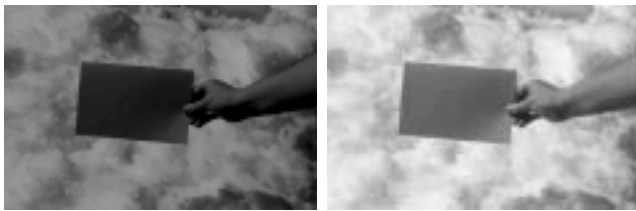Figure 2: *A photographer uses the Zone System to anticipate potential print problems.*

Figure 3: *A normal-key map for a high-key scene (for example containing snow) results in an unsatisfactory image (left). A high-key map solves the problem (right).*

Figure 4: *The mapping from scene zones to print zones. Scene zones at either extreme will map to pure black (zone 0) or white (zone X) if the dynamic range of the scene is eleven zones or more.*

Middle-grey: This is the subjective middle brightness region of the scene, which is typically mapped to print zone V.

Dynamic range: In computer graphics the dynamic range of a scene is expressed as the ratio of the highest scene luminance to the lowest scene luminance. Photographers are more interested in the ratio of the highest and lowest luminance regions where *detail* is visible. This can be viewed as a subjective measure of dynamic range. Because zones relate logarithmically to scene luminances, dynamic range can be expressed as the difference between highest and lowest distinguishable scene zones (Figure 4).

Key: The key of a scene indicates whether it is subjectively light, normal, or dark. A white-painted room would be high-key, and a dim stable would be low-key.

Dodging-and-burning: This is a printing technique where some light is withheld from a portion of the print during development (dodging), or more light is added to that region (burning). This will lighten or darken that region in the final print relative to what it would be if the same development were used for all portions of the print. In traditional photography this technique is applied using a small wand or a piece of paper with a hole cut out.

A crucial part of the Zone System is its methodology for predicting how scene luminances will map to a set of print zones. The photographer first takes a luminance reading of a surface he perceives as a middle-grey (Figure 2 top). In a typical situation this will be mapped to zone V, which corresponds to the 18% reflectance of the print. For high-key scenes the middle-grey will be one of the darker regions, whereas in low-key scenes this will be one of the lighter regions. This choice is an artistic one, although an 18% grey-card is often used to make this selection process more mechanical (Figure 3).

Next the photographer takes luminance readings of both light and dark regions to determine the dynamic range of the scene (Figure 2 bottom). If the dynamic range of the scene does not exceed nine zones, an appropriate choice of middle grey can ensure that all textured detail is captured in the final print. For a dynamic range of more than nine zones, some areas will be mapped to pure black or white with a standard development process. Sometimes such loss of detail is desirable, such as a very bright object being mapped to pure white (see [Adams 1983], p. 51). For regions where loss of detail is objectionable, the photographer can resort to dodging-and-burning which will locally change the development process.

The above procedure indicates that the photographic process is difficult to automate. For example, determining that an adobe building is high-key would be very difficult without some knowledge

limitations presented by slides or prints on photographic papers. Many common practices were developed over the 150 years of photographic practice [London and Upton 1998]. At the same time there were a host of quantitative measurements of media response characteristics by developers [Stroebel et al. 2000]. However, there was usually a disconnect between the artistic and technical aspects of photographic practice, so it was very difficult to produce satisfactory images without a great deal of experience.

Ansel Adams attempted to bridge this gap with an approach he called the Zone System [Adams 1980; Adams 1981; Adams 1983] which was first developed in the 1940s and later popularized by Minor White [White et al. 1984]. It is a system of "practical sensitometry", where the photographer uses measured information in the field to improve the chances of producing a good final print. The Zone System is still widely used more than fifty years after its inception [Woods 1993; Graves 1997; Johnson 1999]. Therefore, we believe it is useful as a basis for addressing the tone reproduction problem. Before discussing how the Zone System is applied, we first summarize some relevant terminology.

Zone: A zone is defined as a Roman numeral associated with an approximate luminance range in a scene as well as an approximate reflectance of a print. There are eleven print zones, ranging from pure black (zone 0) to pure white (zone X), each doubling in intensity, and a potentially much larger number of scene zones (Figure 4).
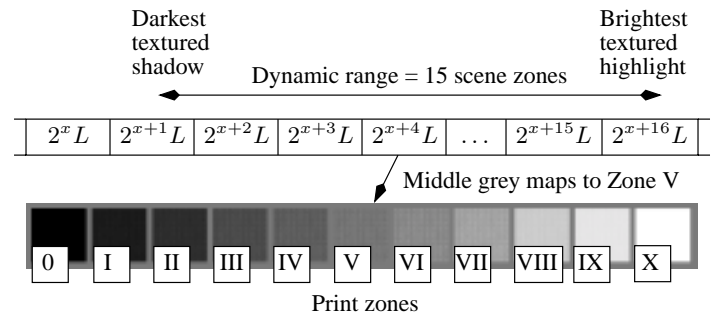
about the adobe's true reflectance. Only knowledge of the geometry and light inter-reflections would allow one to know the difference between luminance ratios of a dark-dyed adobe house and a normal adobe house. However, the Zone System provides the photographer with a small set of subjective controls. These controls form the basis for our tone reproduction algorithm described in the next section.

The challenges faced in tone reproduction for rendered or captured digital images are largely the same as those faced in conventional photography. The main difference is that digital images are in a sense "perfect" negatives, so no luminance information has been lost due to the limitations of the film process. This is a blessing in that detail is available in all luminance regions. On the other hand, this calls for a more extreme dynamic range reduction, which could in principle be handled by an extension of the dodging-and-burning process. We address this issue in the next section.

3 Algorithm

The Zone System summarized in the last section is used to develop a new tone mapping algorithm for digital images, such as those created by rendering algorithms (e.g., [Ward Larson and Shakespeare 1998]) or captured using high dynamic range photography [Debevec and Malik 1997]. We are not trying to closely mimic the actual photographic process [Geigel and Musgrave 1997], but instead use the basic conceptual framework of the Zone System to manage choices in tone reproduction. We first apply a scaling that is analogous to setting exposure in a camera. Then, if necessary, we apply automatic dodging-and-burning to accomplish dynamic range compression.

3.1 Initial luminance mapping

We first show how to set the tonal range of the output image based on the scene's key value. Like many tone reproduction methods [Tumblin and Rushmeier 1993; Ward 1994; Holm 1996], we view the log-average luminance as a useful approximation to the key of the scene. This quantity $\bar{L}_w$ is computed by:

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

where $L_w(x,y)$ is the "world" luminance for pixel (x,y), N is the total number of pixels in the image and δ is a small value to avoid the singularity that occurs if black pixels are present in the image. If the scene has normal-key we would like to map this to middle-grey of the displayed image, or 0.18 on a scale from zero to one. This suggests the equation:

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

where $L(x,y)$ is a scaled luminance and $a = 0.18$. For low-key or high-key images we allow the user to map the log average to different values of a. We typically vary a from 0.18 up to 0.36 and 0.72 and vary it down to 0.09, and 0.045. An example of varying is given in Figure 5. In the remainder of this paper we call the value of parameter a the "key value", because it relates to the key of the image after applying the above scaling.

The main problem with Equation 2 is that many scenes have predominantly a normal dynamic range, but have a few high luminance regions near highlights or in the sky. In traditional photography this issue is dealt with by compression of both high and low luminances. However, modern photography has abandoned these "s"-shaped transfer curves in favor of curves that compress mainly the
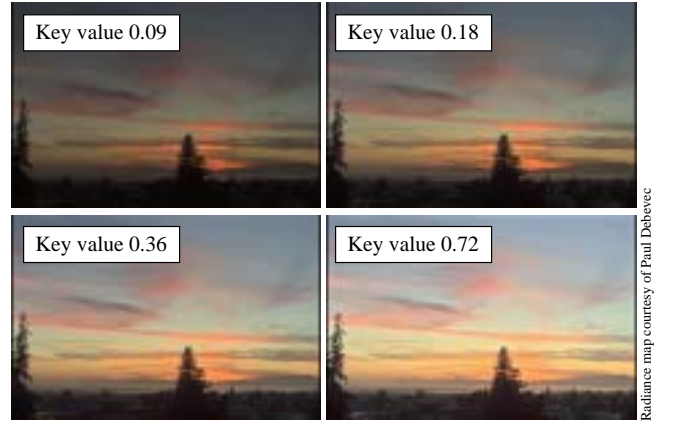

Figure 5: *The linear scaling applied to the input luminance allows the user to steer the final appearance of the tone-mapped image. The dynamic range of the image is 7 zones.*

high luminances [Mitchell 1984; Stroebel et al. 2000]. A simple tone mapping operator with these characteristics is given by:

$$L_d(x,y) = \frac{L(x,y)}{1 + L(x,y)}. \qquad (3)$$

Note that high luminances are scaled by approximately $1/L$, while low luminances are scaled by 1. The denominator causes a graceful blend between these two scalings. This formulation is guaranteed to bring all luminances within displayable range. However, as mentioned in the previous section, this is not always desirable. Equation 3 can be extended to allow high luminances to burn out in a controllable fashion:

$$L_d(x,y) = \frac{L(x,y)\left(1 + \frac{L(x,y)}{L_{\text{white}}^2}\right)}{1 + L(x,y)} \qquad (4)$$

where L_{white} is the smallest luminance that will be mapped to pure white. This function is a blend between Equation 3 and a linear mapping. It is shown for various values of L_{white} in Figure 6. If L_{white} value is set to the maximum luminance in the scene $L_{\max}$ or higher, no burn-out will occur. If it is set to infinity, then the function reverts to Equation 3. By default we set L_{white} to the maximum luminance in the scene. If this default is applied to scenes that have a low dynamic range (i.e., $L_{\max} < 1$), the effect is a subtle contrast enhancement, as can be seen in Figure 7.

The results of this function for higher dynamic range images is shown in the left images of Figure 8. For many high dynamic range images, the compression provided by this technique appears to be sufficient to preserve detail in low contrast areas, while compressing high luminances to a displayable range. However, for very high dynamic range images important detail is still lost. For these images a local tone reproduction algorithm that applies dodging-and-burning is needed (right images of Figure 8).

3.2 Automatic dodging-and-burning

In traditional dodging-and-burning, all portions of the print potentially receive a different exposure time from the negative, bringing "up" selected dark regions or bringing "down" selected light regions to avoid loss of detail [Adams 1983]. With digital images we have the potential to extend this idea to deal with very high dynamic range images. We can think of this as choosing a key value for every pixel, which is equivalent to specifying a local a in Equation 2.
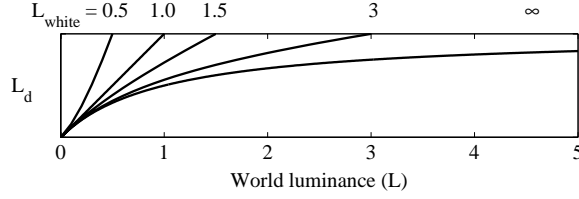
Figure 6: *Display luminance as function of world luminance for a family of values for L_{white}.*

Figure 7: *Left: low dynamic range input image (dynamic range is 4 zones). Right: the result of applying the operator given by Equation 4.*

This serves a similar purpose to the local adaptation methods of the perceptually-driven tone mapping operators [Pattanaik et al. 1998; Tumblin et al. 1999].

Dodging-and-burning is typically applied over an entire region bounded by large contrasts. For example, a local region might correspond to a single dark tree on a light background [Adams 1983]. The size of a local region is estimated using a measure of local contrast, which is computed at multiple spatial scales [Peli 1990]. Such contrast measures frequently use a center-surround function at each spatial scale, often implemented by subtracting two Gaussian blurred images. A variety of such functions have been proposed, including [Land and McCann 1971; Marr and Hildreth 1980; Blommaert and Martens 1990; Peli 1990; Jernigan and McLean 1992; Gove et al. 1995; Pessoa et al. 1995] and [Hansen et al. 2000]. After testing many of these variants, we chose a center-surround function derived from Blommaert's model for brightness perception [Blommaert and Martens 1990] because it performed the best in our tests. This function is constructed using circularly symmetric Gaussian profiles of the form:

$$R_i(x, y, s) = \frac{1}{\pi(\alpha_i s)^2} \exp\left(-\frac{x^2 + y^2}{(\alpha_i s)^2}\right). \qquad (5)$$

These profiles operate at different scales s and at different image positions (x, y). Analyzing an image using such profiles amounts to convolving the image with these Gaussians, resulting in a response V_i as function of image location, scale and luminance distribution L:

$$V_i(x, y, s) = L(x, y) \otimes R_i(x, y, s). \qquad (6)$$

This convolution can be computed directly in the spatial domain, or for improved efficiency can be evaluated by multiplication in the Fourier domain. The smallest Gaussian profile will be only slightly larger than one pixel and therefore the accuracy with which the above equation is evaluated, is important. We perform the integration in terms of the error function to gain a high enough accuracy without having to resort to super-sampling.

The center-surround function we use is defined by:

$$V(x, y, s) = \frac{V_1(x, y, s) - V_2(x, y, s)}{2^\phi a/s^2 + V_1(x, y, s)} \qquad (7)$$

Radiance map courtesy of Greg Ward (top) and Cornell Program of Computer Graphics (bottom)
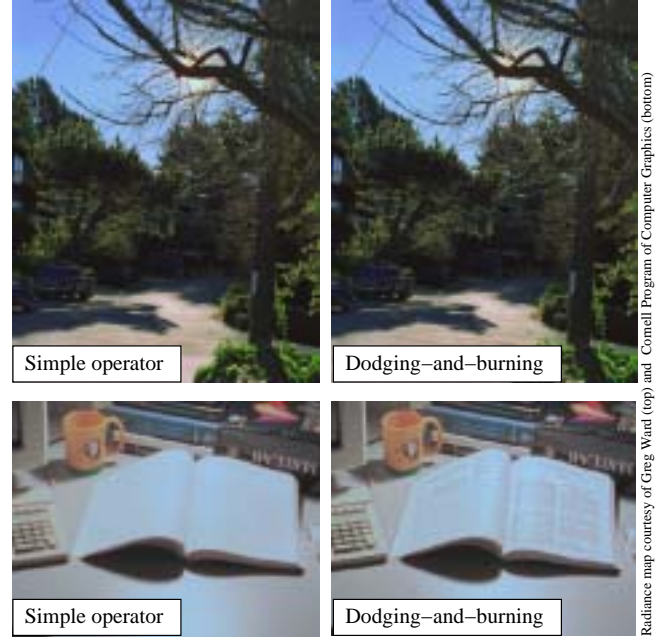
Figure 8: *The simple operator of Equation 3 brings out sufficient detail in the top image (dynamic range is 6 zones), although applying dodging-and-burning does not introduce artifacts. For the bottom image (dynamic range is 15 zones) dodging-and-burning is required to make the book's text visible.*

where center V_1 and surround V_2 responses are derived from Equations 5 and 6. This constitutes a standard difference of Gaussians approach, normalized by $2^\phi a/s^2 + V_1$ for reasons explained below. The free parameters a and ϕ are the key value and a sharpening parameter respectively.

For computational convenience, we set the center size of the next higher scale to be the same as the surround of the current scale. Our choice of center-surround ratio is 1.6, which results in a difference of Gaussians model that closely resembles a Laplacian of Gaussian filter [Marr 1982]. From our experiments, this ratio appears to produce slightly better results over a wide range of images than other choices of center-surround ratio. However, this ratio can be altered by a small amount to optimize the center-surround mechanism for specific images.

Equation 7 is computed for the sole purpose of establishing a measure of locality for each pixel, which amounts to finding a scale s_m of appropriate size. This scale may be different for each pixel, and the procedure for its selection is the key to the success of our dodging-and-burning technique. It is also a deviation from the original Blommaert model [Blommaert and Martens 1990]. The area to be considered local is in principle the largest area around a given pixel where no large contrast changes occur. To compute the size of this area, Equation 7 is evaluated at different scales s. Note that $V_1(x, y, s)$ provides a local average of the luminance around (x, y) roughly in a disc of radius s. The same is true for $V_2(x, y, s)$ although it operates over a larger area at the same scale s. The values of V_1 and V_2 are expected to be very similar in areas of small luminance gradients, but will differ in high contrast regions. To choose the largest neighborhood around a pixel with fairly even luminances, we threshold V to select the corresponding scale s_m. Starting at the lowest scale, we seek the first scale s_m where:

$$|V(x, y, s_m)| < \epsilon \qquad (8)$$

is true. Here ϵ is the threshold. The V_1 in the denominator of Equa-
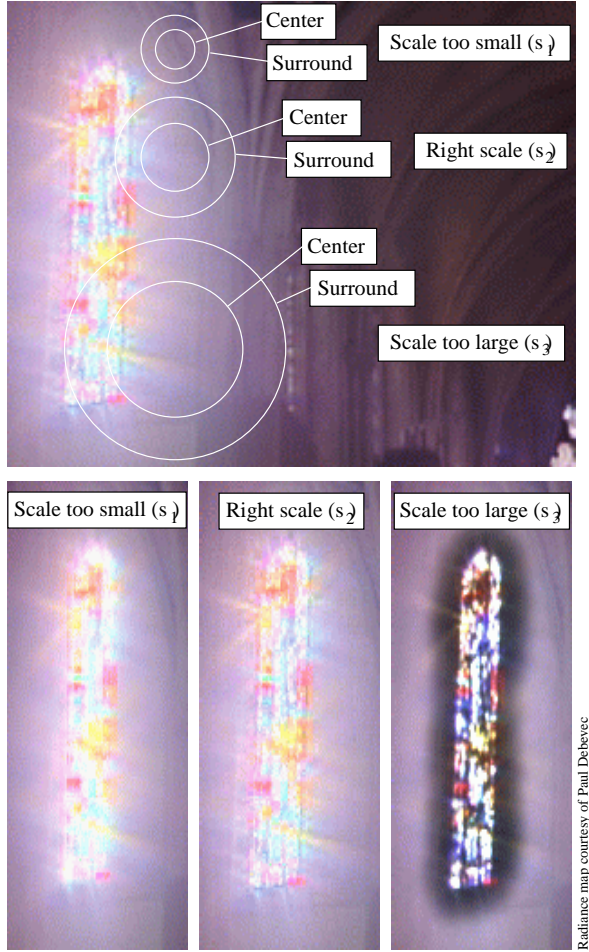
Figure 9: *An example of scale selection. The top image shows center and surround at different sizes. The lower images show the results of particular choices of scale selection. If scales are chosen too small, detail is lost. On the other hand, if scales are chosen too large, dark rings around luminance steps will form.*

tion 7 makes thresholding V independent of absolute luminance level, while the $2^{\phi}a/s^2$ term prevents V from becoming too large when V approaches zero.

Given a judiciously chosen scale for a given pixel, we observe that $V_1(x, y, s_m)$ may serve as a local average for that pixel. Hence, the global tone reproduction operator of Equation 3 can be converted into a local operator by replacing L with V_1 in the denominator:

$$L_d(x, y) = \frac{L(x, y)}{1 + V_1(x, y, s_m(x, y))} \qquad (9)$$

This function constitutes our local dodging-and-burning operator. The luminance of a dark pixel in a relatively bright region will satisfy $L < V_1$, so this operator will decrease the display luminance L_d, thereby increasing the contrast at that pixel. This is akin to photographic "dodging". Similarly, a pixel in a relatively dark region will be compressed less, and is thus "burned". In either case the pixel's contrast relative to the surrounding area is increased. For this reason, the above scale selection method is of crucial importance, as illustrated in the example of Figure 9. If s_m is too small, then V_1 is close to the luminance L and the local operator reduces to our global operator (s_1 in Figure 9). On the other hand, choosing

s_m too large causes dark rings to form around bright areas (s_3 in the same figure), while choosing the scale as outlined above causes the right amount of detail and contrast enhancement without introducing unwanted artifacts (s_2 in Figure 9).

Using a larger scale s_m tends to increase contrast and enhance edges. The value of the threshold ϵ in Equation 8, as well as the choice of ϕ in Equation 7, serve as edge enhancement parameters and work by manipulating the scale that would be chosen for each pixel. Decreasing ϵ forces the appropriate scale s_m to be larger. Increasing ϕ also tends to select a slightly larger scale s_m, but only at small scales due to the division of ϕ by s^2. An example of the effect of varying ϕ is given in Figure 10.

A further observation is that because V_1 tends to be smaller than L for very bright pixels, our local operator is not guaranteed to keep the display luminance L_d below 1. Thus, for extremely bright areas some burn-out may occur and this is the reason we clip the display luminance to 1 afterwards. As noted in section 2, a small amount of burn-out may be desirable to make light sources such as the sun look very bright.

In summary, by automatically selecting an appropriate neighborhood for each pixel we effectively implement a pixel-by-pixel dodging and burning technique as applied in photography [Adams 1983]. These techniques locally change the exposure of a film, and so darken or brighten certain areas in the final print.

4   Results

We implemented our algorithm in C++ and obtained the luminance values from the input R, G and B triplets with $L = 0.27R + 0.67G + 0.06B$. The convolutions of Equation 5 were computed using a Fast Fourier Transform (FFT). Because Gaussians are separable, these convolutions can also be efficiently computed in image space. This is easier to implement than an FFT, but it is somewhat slower for large images. Because of the normalization by V_1, our method is insensitive to edge artifacts normally associated with the computation of an FFT.

The key value setting is determined on a per image basis, while unless noted otherwise, the parameter ϕ is set to 8.0 for all the images in this paper. Our new local operator uses Gaussian profiles s at 8 discrete scales increasing with a factor of 1.6 from 1 pixel wide to 43 pixels wide. For practical purposes we would like the Gaussian profile at the smallest scale to have 2 standard deviations overlap with 1 pixel. This is achieved by setting the scaling parameter α_1 to $1/2\sqrt{2} \approx 0.35$. The parameter α_2 is 1.6 times as large. The threshold ϵ used for scale selection was set to 0.05.

We use images with a variety of dynamic ranges as indicated throughout this section. Note that we are using the photographic definition of dynamic range as presented in Section 2. This results in somewhat lower ranges than would be obtained if a conventional computer graphics measure of dynamic range were used. However, we believe the photographic definition is more predictive of how challenging the tone reproduction of a given image is.

In the absence of well-tested quantitative methods to compare tone mapping operators, we compare our results to a representative set of tone reproduction techniques for digital images. In this section we briefly introduce each of the operators and show images of them in the next section. Specifically, we compare our new operator of Equation 9 with the following.

Stockham's homomorphic filtering Using the observation that lighting variation occurs mainly in low frequencies and humans are more aware of albedo variations, this method operates by downplaying low frequencies and enhancing high frequencies [Oppenheim et al. 1968; Stockham 1972].

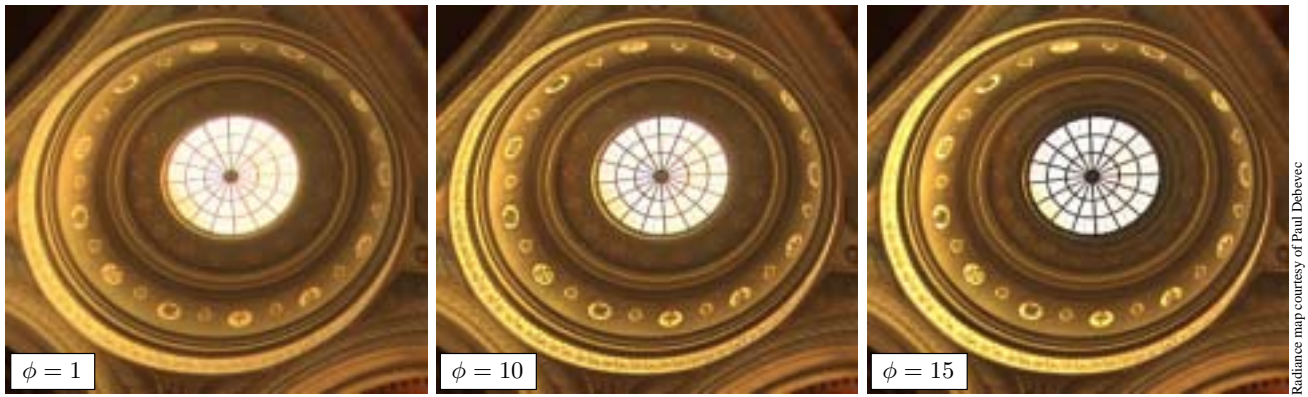**Tumblin-Rushmeier's brightness matching operator** . A model of brightness perception is used to drive this global operator.

Figure 10: *The free parameter ϕ in Equation 7 controls sharpening.*

We use the 1999 formulation [Tumblin et al. 1999] as we have found it produces much better subjective results to the earlier versions [Tumblin and Rushmeier 1991; Tumblin and Rushmeier 1993].

Chiu's local scaling A linear scaling that varies continuously is used to preserve local contrast with heuristic dodging-and-burning used to avoid burn-out [Chiu et al. 1993].

Ward's contrast scale factor A global multiplier is used that aims to maintain visibility thresholds [Ward 1994].

Ferwerda's adaptation model This operator alters contrast, color saturation and spatial frequency content based on psychophysical data [Ferwerda et al. 1996]. We have used the photopic portion of their algorithm.

Ward's histogram adjustment method This method uses an image's histogram to implicitly segment the image so that separate scaling algorithms can be used in different luminance zones. Visibility thresholds drive the processing [Ward et al. 1997]. The model incorporates human contrast and color sensitivity, glare and spatial acuity, although for a fair comparison we did not use these features.

Schlick's rational sigmoid This is a family of simple and fast methods using rational sigmoid curves and a set of tunable parameters [Schlick 1994].

Pattanaik's local adaptation model Both threshold and suprathreshold vision is considered in this multi-scale model of local adaptation [Pattanaik et al. 1998]. Chromatic adaptation is also included.

Note that the goals of most of these operators are different from our goal of producing a subjectively satisfactory image. However, we compare their results with ours because all of the above methods do produce subjectively pleasing images for many inputs. There are comparisons possible with many other techniques that are outside the scope of this evaluation. In particular, we do not compare our results with the first perceptually-driven works [Miller et al. 1984; Upstill 1985] because they are not widely used in graphics and are similar to works we do compare with [Ward 1994; Ferwerda et al. 1996; Tumblin et al. 1999]. We also do not compare with the multiscale-Retinex work because it is reminiscent of Pattanaik's local adaptation model, while being aimed at much lower contrast reductions of about 5:1 [Rahman et al. 1996]. Holm has a complete implementation of the Zone System for digital cameras [Holm 1996], but his contrast reduction is also too low for our purposes.

Pattanaik

New operator

Figure 13: *Desk image (dynamic range is 15 zones).*

Next, we do not compare with the "layering" method because it requires albedo information in addition to luminances [Tumblin et al. 1999]. Finally, we consider some work to be visualization methods for digital images rather than true tone mapping operators. These are the LCIS filter which consciously allows visible artifacts in exchange for visualizing detail [Tumblin and Turk 1999], the mouse-driven foveal adaptation method [Tumblin et al. 1999] and Pardo's multi-image visualization technique [Pardo and Sapiro 2001].

The format in which we compare the various methods is a "knock-out race" using progressively more difficult images. We take this approach to avoid an extremely large number of images. In Figure 11 eight different tone mapping operators are shown side by
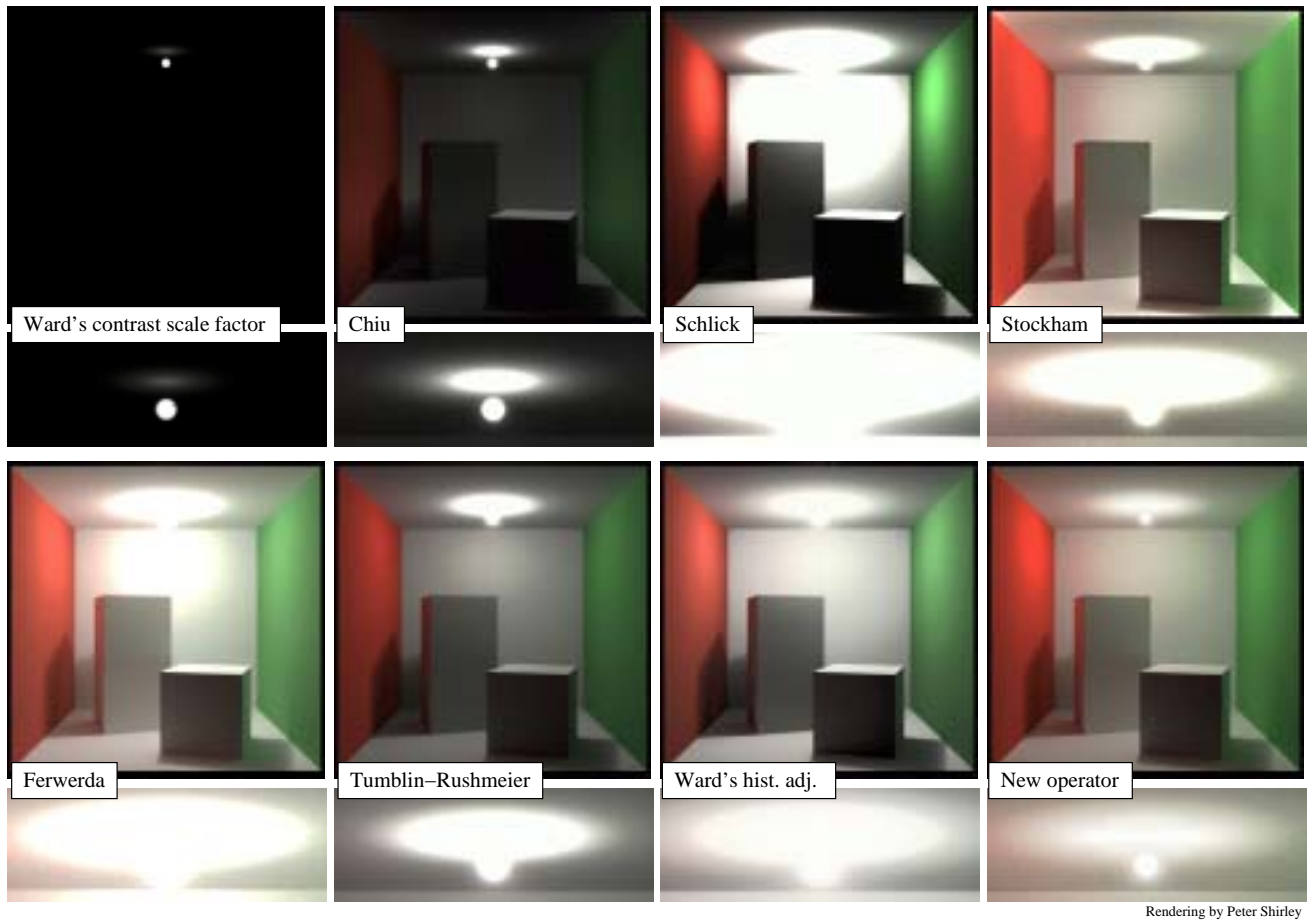
Ward's contrast scale factor

Chiu

Schlick

Stockham

Ferwerda

Tumblin–Rushmeier

Ward's hist. adj.

New operator

Rendering by Peter Shirley

Figure 11: *Cornell box high dynamic range images including close-ups of the light sources. The dynamic range of this image is 12 zones.*

Schlick

Stockham

Ferwerda

Tumblin–Rushmeier

Ward's hist. adj.
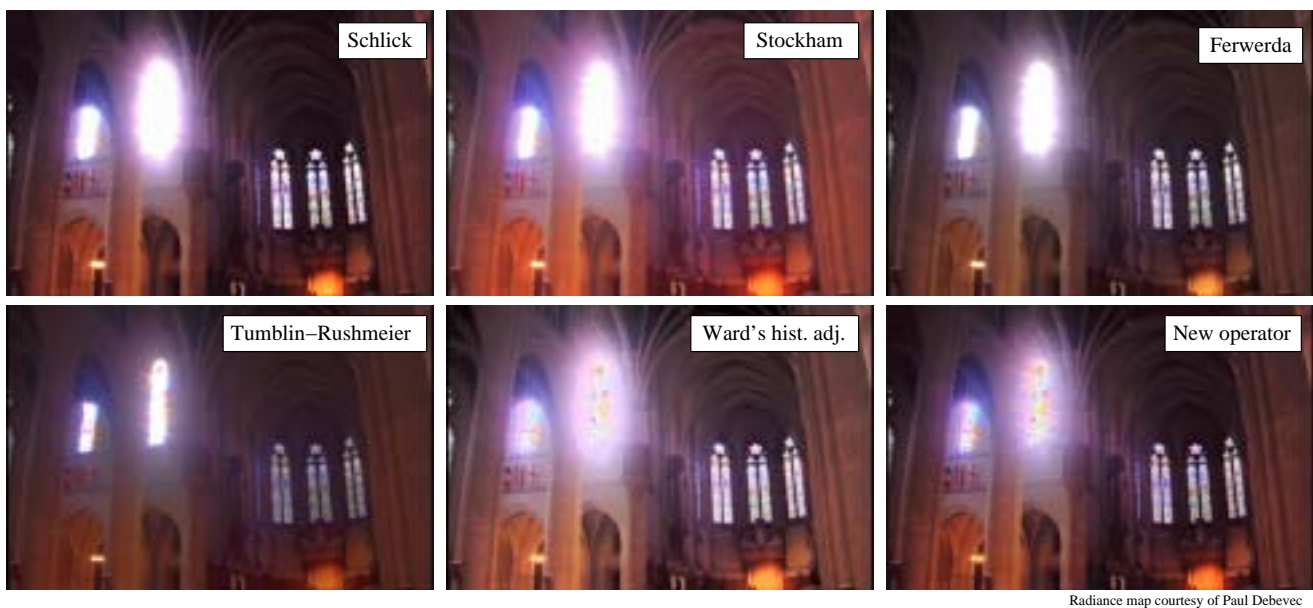
New operator

Radiance map courtesy of Paul Debevec

Figure 12: *Nave image with a dynamic range of 12 zones.*

side using the Cornell box high dynamic range image as input. The model is slightly different from the original Cornell box because we have placed a smaller light source underneath the ceiling of the box so that the ceiling receives a large quantity of direct illumination, a characteristic of many architectural environments. This image has little high frequency content and it is therefore easy to spot any deficiencies in the tone mapping operators we have applied. In this and the following figures, the operators are ordered roughly by their ability to bring the image within dynamic range. Using the Cornell box image (Figure 11), we eliminate those operators that darken the image too much and therefore we do not include the contrast based scaling factor and Chiu's algorithm in further tests.

Similar to the Cornell box image is the Nave photograph, although this is a low-key image and the stained glass windows contain high frequency detail. From a photographic point of view, good tone mapping operators would show detail in the dark areas while still allowing the windows to be admired. The histogram adjustment algorithm achieves both goals, although halo-like artifacts are introduced around the bright window. Both the Tumblin-Rushmeier model and Ferwerda's visibility matching method fail to bring the church window within displayable range. The same is true for Stockham style filtering and Schlick's method.

The most difficult image to bring within displayable range is presented in Figures 1 and 13. Due to its large dynamic range, it presents problems for most tone reproduction operators. This image was first used for Pattanaik's local adaptation model [Pattanaik et al. 1998]. Because his operator includes color correction as well as dynamic range reduction, we have additionally color corrected our tone-mapped image (Figure 13) using the method presented in [Reinhard et al. 2001]. Pattanaik's local adaptation operator produces visible artifacts around the light source in the desk image, while the new operator does not.

The efficiency of both our new global (Equation 3, without dodging-and-burning) and local tone mapping operators (Equation 9) is high. Timings obtained on a 1.8 GHz Pentium 4 PC are given in Table 1 for two different image sizes. While we have not counted any disk I/O, the timings for preprocessing as well as the main tone mapping algorithm are presented. The preprocessing for the local operator (Equation 9) consists of the mapping of the log average luminance to the key value, as well as all FFT calculations. The total time for a 512^2 image is 1.31 seconds for the local operator, which is close to interactive, while our global operator (Equation 3) performs at a rate of 20 frames per second, which we consider real-time. Computation times for the 1024^2 images is around 4 times slower, which is according to expectation.

We have also experimented with a fast approximation of the Gaussian convolution using a multiscale spline based approach [Burt and Adelson 1983], which was first used in the context of tone reproduction by [Tumblin et al. 1999], and have found that the computation is about 3.7 times faster than our Fourier domain implementation. This improved performance comes at the cost of some small artifacts introduced by the approximation, which can be successfully masked by the high frequency content of the photographs. If high frequencies are absent, some blocky artifacts become visible, as can be seen in Figure 14. On the other hand, just like its FFT based counter-part, this approximation manages to bring out the detail of the writing on the open book in this figure as opposed to our global operator of Equation 3 (compare with the left image of Figure 8). As such, the local FFT based implementation, the local spline based approximation and the global operator provide a useful trade-off between performance and quality, allowing any user to select the best operator given a specified maximum run-time.

Finally, to demonstrate that our method works well on a broad range of high dynamic range images, Figure 15 shows a selection of tone-mapped images using our new operator. It should be noted
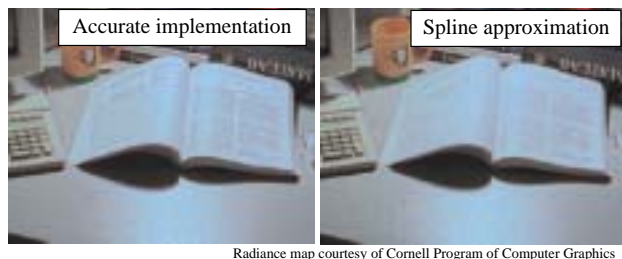

Radiance map courtesy of Cornell Program of Computer Graphics

Figure 14: *Compare the spline based local operator (right) with the more accurate local operator (left). The spline approach exhibits some blocky artifacts on the table, although this is masked in the rest of the image.*

Algorithm	Preprocessing	Tone Mapping	Total
Image size: 512×512			
Local	1.23	0.08	1.31
Spline	0.25	0.11	0.36
Global	0.02	0.03	0.05
Image size: 1024×1024			
Local	5.24	0.33	5.57
Spline	1.00	0.47	1.47
Global	0.70	0.11	0.18

Table 1: *Timing in seconds for our global (Equation 3) and local (Equation 9) operators. The middle rows show the timing for the approximated Gaussian convolution using a multiscale spline approach [Burt and Adelson 1983].*

that most of the images in this figure present serious challenges to other tonemapping operators. Interestingly, the area around the sun in the rendering of the landscape is problematic for any method that attempts to bring the maximum scene luminance within a displayable range without clamping. This is not the case for our operator because it only brings textured regions within range, which is relatively simple because, excluding the sun, this scene only has a small range of luminances. A similar observation can be made for the image of the lamp on the table and the image with the streetlight behind the tree.

5 Summary

Photographers aim to compress the dynamic range of a scene in a manner that creates a pleasing image. We have developed a relatively simple and fast tone reproduction algorithm for digital images that borrows from 150 years of photographic experience. It is designed to follow their practices and is thus well-suited for applications where creating subjectively satisfactory and essentially artifact-free images is the desired goal.
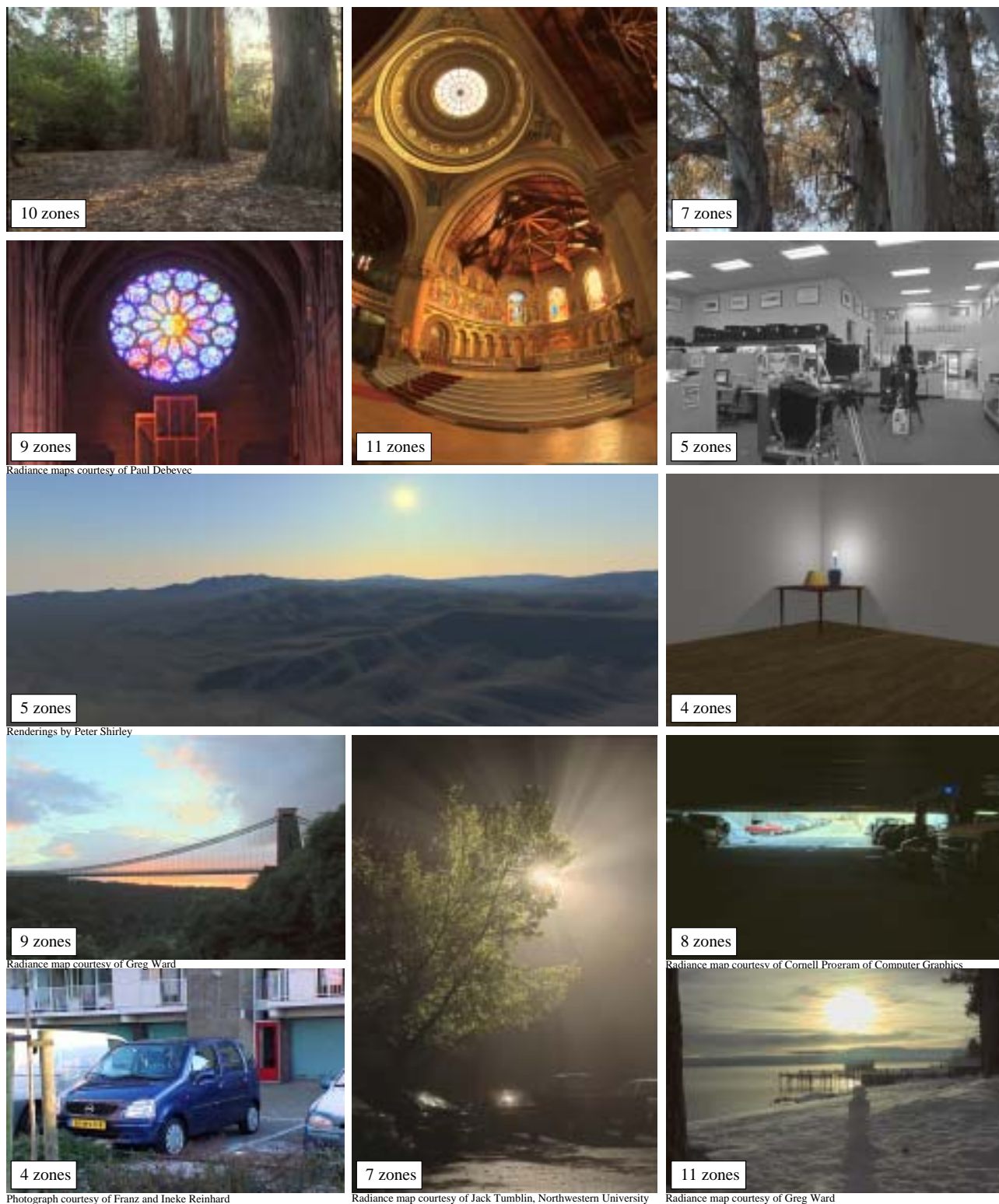
Acknowledgments

Figure 15: *A selection of high and low dynamic range images tone-mapped using our new operator. The labels in the figure indicate the dynamic ranges of the input data.*

References

ADAMS, A. 1980. *The camera*. The Ansel Adams Photography series. Little, Brown and Company.

ADAMS, A. 1981. *The negative*. The Ansel Adams Photography series. Little, Brown and Company.

ADAMS, A. 1983. *The print*. The Ansel Adams Photography series. Little, Brown and Company.

BLOMMAERT, F. J. J., AND MARTENS, J.-B. 1990. An object-oriented model for brightness perception. *Spatial Vision 5*, 1, 15–41.

BURT, P. J., AND ADELSON, E. H. 1983. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics 2*, 4, 217–236.

CHIU, K., HERF, M., SHIRLEY, P., SWAMY, S., WANG, C., AND ZIMMERMAN, K. 1993. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface '93*, 245–253.

COHEN, J., TCHOU, C., HAWKINS, T., AND DEBEVEC, P. 2001. Real-Time high dynamic range texture mapping. In *Rendering techniques 2001*, S. J. Gortler and K. Myszkowski, Eds., 313–320.

DEBEVEC, P. E., AND MALIK, J. 1997. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH 97 Conference Proceedings*, Addison Wesley, T. Whitted, Ed., Annual Conference Series, ACM SIGGRAPH, 369–378.

DURAND, F., AND DORSEY, J. 2000. Interactive tone mapping. In *Eurographics Workshop on Rendering*, 219–230.

FAIRCHILD, M. D. 1998. *Color appearance models*. Addison-Wesley, Reading, MA.

FERWERDA, J. A., PATTANAIK, S., SHIRLEY, P., AND GREENBERG, D. P. 1996. A model of visual adaptation for realistic image synthesis. In *SIGGRAPH 96 Conference Proceedings*, Addison Wesley, H. Rushmeier, Ed., Annual Conference Series, ACM SIGGRAPH, 249–258.

GEIGEL, J., AND MUSGRAVE, F. K. 1997. A model for simulating the photographic development process on digital images. In *SIGGRAPH 97 Conference Proceedings*, Addison Wesley, T. Whitted, Ed., Annual Conference Series, ACM SIGGRAPH, 135–142.

GOVE, A., GROSSBERG, S., AND MINGOLLA, E. 1995. Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience 12*, 1027–1052.

GRAVES, C. 1997. *The zone system for 35mm photographers*, second ed. Focal Press.

HANSEN, T., BARATOFF, G., AND NEUMANN, H. 2000. A simple cell model with dominating opponent inhibition for robust contrast detection. *Kognitionswissenschaft 9*, 93–100.

HOLM, J. 1996. Photographics tone and colour reproduction goals. In *CIE Expert Symposium '96 on Colour Standards for Image Technology*, 51–56.

JERNIGAN, M. E., AND MCLEAN, G. F. 1992. Lateral inhibition and image processing. In *Non-linear vision: determination of neural receptive fields, function, and networks*, R. B. Pinter and B. Nabet, Eds. CRC Press, ch. 17, 451–462.

JOHNSON, C. 1999. *The practical zone system*. Focal Press.

LAND, E. H., AND MCCANN, J. J. 1971. Lightness and retinex theory. *J. Opt. Soc. Am. 63*, 1, 1–11.

LONDON, B., AND UPTON, J. 1998. *Photography*, sixth ed. Longman.

MARR, D., AND HILDRETH, E. C. 1980. Theory of edge detection. *Proceedings of the Royal Society of London, B 207*, 187–217.

MARR, D. 1982. *Vision, a computational investigation into the human representation and processing of visual information*. W H Freeman and Company, San Fransisco.

MATKOVIC, K., NEUMANN, L., AND PURGATHOFER, W. 1997. A survey of tone mapping techniques. In *13th Spring Conference on Computer Graphics*, W. Straßer, Ed., 163–170.

MCNAMARA, A., CHALMERS, A., AND TROSCIANKO, T. 2000. STAR: Visual perception in realistic image synthesis. In *Eurographics 2000 STAR reports*, Eurographics, Interlaken, Switzerland.

MCNAMARA, A. 2001. Visual perception in realistic image synthesis. *Computer Graphics Forum 20*, 4 (December), 211–224.

MILLER, N. J., NGAI, P. Y., AND MILLER, D. D. 1984. The application of computer graphics in lighting design. *Journal of the IES 14* (October), 6–26.

MITCHELL, E. N. 1984. *Photographic Science*. John Wiley and Sons, New York.

OPPENHEIM, A. V., SCHAFER, R., AND STOCKHAM, T. 1968. Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE 56*, 8, 1264–1291.

PARDO, A., AND SAPIRO, G. 2001. Visualization of high dynamic range images. Tech. Rep. 1753, Institute for Mathematics and its Applications, University of Minnesota.

PATTANAIK, S. N., FERWERDA, J. A., FAIRCHILD, M. D., AND GREENBERG, D. P. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *SIGGRAPH 98 Conference Proceedings*, Addison Wesley, M. Cohen, Ed., Annual Conference Series, ACM SIGGRAPH, 287–298.

PATTANAIK, S. N., TUMBLIN, J., YEE, H., , AND GREENBERG, D. P. 2000. Time-dependent visual adaptation for fast realistic display. In *SIGGRAPH 2000 Conference Proceedings*, Addison Wesley, K. Akeley, Ed., Annual Conference Series, ACM SIGGRAPH, 47–54.

PELI, E. 1990. Contrast in complex images. *J. Opt. Soc. Am. A 7*, 10 (October), 2032–2040.

PESSOA, L., MINGOLLA, E., AND NEUMANN, H. 1995. A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Research 35*, 15, 2201–2223.

RAHMAN, Z., JOBSON, D. J., AND WOODELL, G. A. 1996. A multiscale retinex for color rendition and dynamic range compression. In *SPIE Proceedings: Applications of Digital Image Processing XIX*, vol. 2847.

RAHMAN, Z., WOODELL, G. A., AND JOBSON, D. J. 1997. A comparison of the multiscale retinex with other image enhancement techniques. In *IS&T's 50th Annual Conference: A Celebration of All Imaging*, vol. 50, 426–431.

REINHARD, E., ASHIKHMIN, M., GOOCH, B., AND SHIRLEY, P. 2001. Color transfer between images. *IEEE Computer Graphics and Applications 21* (September/October), 34–41.

SCHEEL, A., STAMMINGER, M., AND SEIDEL, H.-P. 2000. Tone reproduction for interactive walkthroughs. *Computer Graphics Forum 19*, 3 (August), 301–312.

SCHLICK, C. 1994. Quantization techniques for the visualization of high dynamic range pictures. In *Photorealistic Rendering Techniques*, Springer-Verlag Berlin Heidelberg New York, P. Shirley, G. Sakas, and S. Müller, Eds., 7–20.

STOCKHAM, T. 1972. Image processing in the context of a visual model. *Proceedings of the IEEE 60*, 7, 828–842.

STROEBEL, L., COMPTON, J., CURRENT, I., AND ZAKIA, R. 2000. *Basic photographic materials and processes*, second ed. Focal Press.

TUMBLIN, J., AND RUSHMEIER, H. 1991. Tone reproduction for realistic computer generated images. Tech. Rep. GIT-GVU-91-13, Graphics, Visualization, and Useability Center, Georgia Institute of Technology.

TUMBLIN, J., AND RUSHMEIER, H. 1993. Tone reproduction for computer generated images. *IEEE Computer Graphics and Applications 13*, 6 (November), 42–48.

TUMBLIN, J., AND TURK, G. 1999. LCIS: A boundary hierarchy for detail-preserving contrast reduction. In *Siggraph 1999, Computer Graphics Proceedings*, Addison Wesley Longman, Los Angeles, A. Rockwood, Ed., Annual Conference Series, 83–90.

TUMBLIN, J., HODGINS, J. K., AND GUENTER, B. K. 1999. Two methods for display of high contrast images. *ACM Transactions on Graphics 18 (1)*, 56–94.

UPSTILL, S. 1985. *The Realistic Presentation of Synthetic Images: Image Processing in Computer Graphics*. PhD thesis, University of California at Berkeley.

WARD, G., RUSHMEIER, H., AND PIATKO, C. 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics 3*, 4 (December).

WARD LARSON, G., AND SHAKESPEARE, R. A. 1998. *Rendering with Radiance*. Morgan Kaufmann Publishers.

WARD, G. 1994. A contrast-based scalefactor for luminance display. In *Graphics Gems IV*, P. Heckbert, Ed. Academic Press, Boston, 415–421.

WHITE, M., ZAKIA, R., AND LORENZ, P. 1984. *The new zone system manual*. Morgan & Morgan, Inc.

WOODS, J. C. 1993. *The zone system craftbook*. McGraw Hill.