The Human Visual System and HDR Tone Mapping

06

The dynamic range of illumination in a real-world scene is high — on the order of 10,000 to 1 from highlights to shadows, and higher if light sources are directly visible. A much larger range of illumination can also occur if the scene includes both an outdoor area illuminated by sunlight and an indoor area illuminated by interior light (see, for example, Figure 6.1). Using techniques discussed in Chapter 4, we are able to capture this dynamic range with full precision. Unfortunately, most display devices and display media available to us come with only a moderate absolute output level and a useful dynamic range of less than 100 to 1. The discrepancy between the wide ranges of illumination that can be captured and the small ranges that can be reproduced by existing displays makes the accurate display of the images of the captured scene difficult. This is the HDR display problem, or HDR tone-mapping problem. We introduce the tone-mapping problem in this chapter and discuss individual solutions in detail in the following two chapters.

6.1 TONE-MAPPING PROBLEM

For a display to exhibit realism, the images should be faithful visual representations of the scenes they depict. This is not a new problem. Artists and photographers have been addressing this problem for a long time. The core problem for the artist (canvas), photographer (positive print), and us (display device) is that the light intensity level in the environment may be completely beyond the output level reproduced by the display medium. In addition, the contrast experienced in a real

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 2

FIGURE 6.1 *Image depicting both indoor and outdoor areas. The different lighting conditions in these areas gives rise to an HDR. (Image courtesy of the Albin Polasek Museum, Winter Park, Florida.)*

environment may greatly exceed the contrast range that can be reproduced by those display devices. Appearance of a scene depends upon the level of illumination and the contrast range [31]. Some commonly noticed examples are that scenes appear more colorful and contrasty on a sunny day, colorful scenes of the day appear gray during night, and moonlight has a bluish appearance. Hence, simple scaling or compression of the intensity level and the contrast range to fit them into the display limits is not sufficient to reproduce the accurate visual appearance of the scene. Tumblin and Rushmeier [131] formally introduced this problem and suggested the use of visual models for solving this problem (see Figure 6.2 for a pictorial outline). Ever since, developing tone-mapping algorithms that incorporate visual models has

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 3

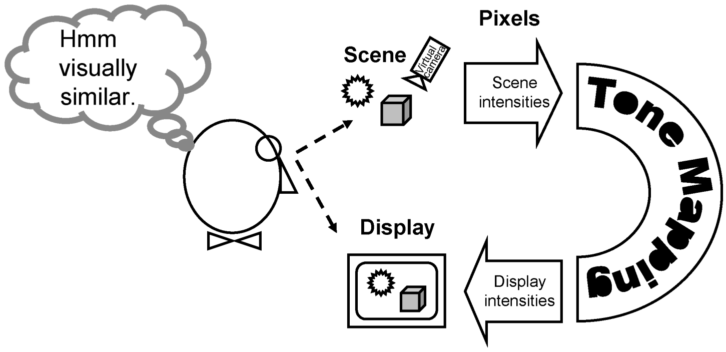


FIGURE 6.2 *Pictorial outline of the tone-mapping problem. The ultimate goal is a visual match between the observed scene and the tone-mapped version of the captured HDR image on the display.*

been an active area of research within the computer graphics and digital-imaging communities.

Reproducing the visual appearance is the ultimate goal in tone mapping. However, defining and quantifying visual appearance itself is not easy and is a current research topic [31]. Instead of delving deep into appearance-related issues in tone mapping, in this chapter we address one basic issue for realistic display of HDR images. First, the HDR must be reduced to fit the display range. This can be achieved by simple scaling of the image. However, such simple scaling often generates images with complete loss of detail (contrast) in the resulting display (Figure 6.3). That gives us a seemingly simpler problem to solve: how to compress the dynamic range of the HDR image to fit into the display range while preserving detail.

The human visual system deals with a similar problem on a regular basis. The signal-to-noise ratio of individual channels in the visual pathway (from retina to brain) is about 32 to 1, less than 2 orders of magnitude [19,55]. Even with this dynamic range limitation, the human visual system functions well: it allows us

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 4

FIGURE 6.3 *HDR image depicting both indoor and outdoor areas. Linear scaling was applied to demonstrate the lack of detail afforded by linear scaling. (Image courtesy of the Albin Polasek Museum, Winter Park, Florida.)*

to function under a wide range of illumination, and allows us to simultaneously perceive the detailed contrast in both the light and dark parts of an HDR scene. Thus, if the goal is to match this perceived realism in the display of HDR images it is important to understand some of the basics of the human visual system. Hence, this chapter focuses on aspects of the human visual system relevant to HDR imaging. We show that most tone-mapping algorithms currently available make use of one of a small number of visual models to solve the HDR problem.

The material described in the following sections has been distilled from the psychophysics and electrophysiology literature, wherein light is variously measured as quanta, intensity, luminance, radiance, or retinal illuminance. To avoid confusion, wherever possible we will use the term *luminance*. If any unit other than luminance is required, we provide the units in which they originally appeared in the literature.

6.2 HUMAN VISUAL ADAPTATION

A striking feature of the human visual system is its capacity to function over the huge range of illumination it encounters during the course of a day. Sunlight can be as much as a million times more intense than moonlight. The intensity of starlight can be one-thousandth of the intensity of moonlight. Thus, the effective range of illumination is more than a billion to one [135]. The dynamic range simultaneously available in a single scene at a given time is much smaller, but still hovers at about four orders of magnitude.

The visual system functions in this range by adapting to the prevailing conditions of illumination. Thus, adaptation renders our visual system less sensitive in daylight and more sensitive at night. For example, car headlights that let drivers drive at night go largely unnoticed in daylight, as shown in Figure 6.4.

In psychophysical studies, human visual adaptation is evaluated by measuring the minimum amount of incremental light by which an observer distinguishes a test

FIGURE 6.4 *Although the headlights are on in both images, during daylight our eyes are less sensitive to car headlights than at night.*

object from the background light. This minimum increment is called a *visual threshold* or *just-noticeable difference* (JND). In a typical threshold measurement experiment, a human subject stares at a wide blank screen for a sufficient amount of time to adjust to its uniform background intensity, I_b. Against this uniform background a small test spot of intensity $I_b + \Delta I$ is flashed. This test spot is called the *stimulus*. The increment ΔI is adjusted to find the smallest detectable ΔI_b. The value of this threshold depends on the value of the background, as shown in Figure 6.5. This figure plots typical threshold versus intensity (TVI) measurements at various

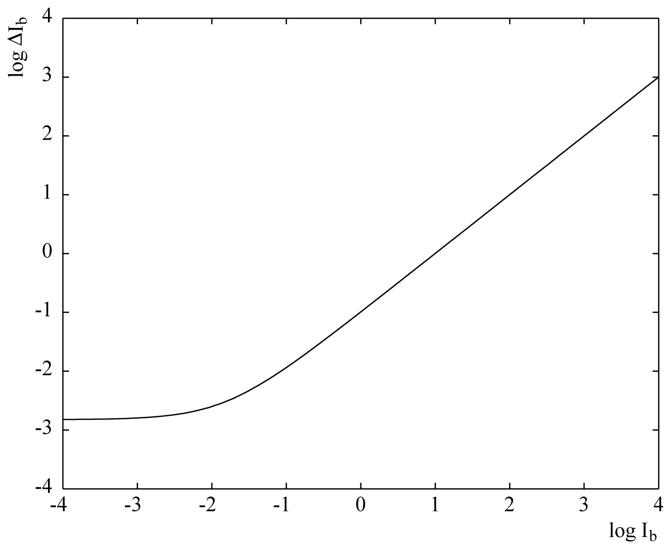


FIGURE 6.5 *Threshold versus intensity (TVI) relation. The plot on the right shows the visual threshold ΔI_b at different background intensities I_b.*

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 7

background intensities. Over much of the background intensity range, the ratio

$$\frac{\Delta I_{\mathrm{b}}}{I_{\mathrm{b}}}$$

is nearly constant, a relation known for over 140 years as Weber's law. The value of this constant fraction is about 1% [135], which can vary with the size of the test spot and the duration for which the stimulus is shown. The constant nature of this fraction suggests that visual adaptation acts as a normalizer, scaling scene intensities to preserve our ability to sense contrasts within scenes.

Visual adaptation to varying conditions of illumination is thought to be possible through the coordinated action of the pupil, the rod-cone system, photochemical reactions, and photoreceptor mechanisms. The role of each of these factors is discussed in the following sections.

6.2.1 THE PUPIL

After passing through the cornea and the aqueous humor, light enters into the visual system through the pupil, a circular hole in the iris (Figure 6.6) [38,48,49]. One of the mechanisms that allows us to adapt to a specific lighting condition is regulation of the amount of light that enters the eye via the size of the opening of the pupil. In fact, the pupil changes its size in response to the background light level. Its diameter changes from a minimum of about 2 mm in bright light to a maximum of about 8 mm in darkness. This change accounts for a reduction in light intensity entering the eye by only a factor of 16 (about 1 log unit). In a range of about 10 billion to 1, the intensity regulation by a factor of 16 is not very significant. Hence, the pupil's role in visual adaptation may be ignored for the purpose of tone reproduction.

6.2.2 THE ROD AND CONE SYSTEMS

Light that has passed through the pupil travels through the lens and the vitreous body before reaching the retina, where it is reflected from a pigmented layer of cells before being absorbed by photoreceptors. The latter convert light into neural signals before they are relayed to other parts of the visual system. The human retina

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 8

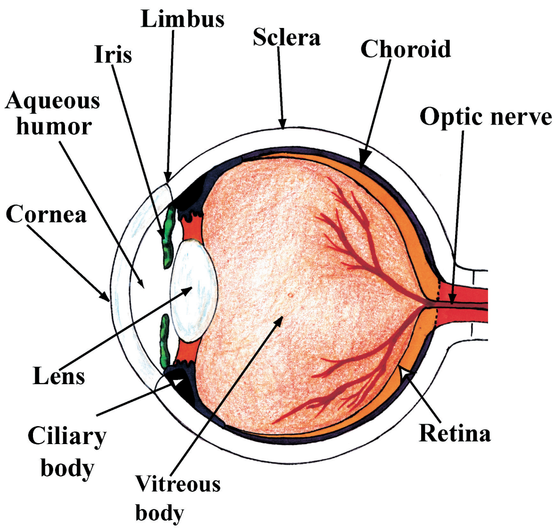


FIGURE 6.6 *Schematic diagram of the human eye and its various components. (Image courtesy of Karen Lefohn.)*

has two distinct types of photoreceptors: rods and cones. Rods are very sensitive to light and are responsible for vision from twilight illumination to very dark lighting conditions. Cones are relatively less sensitive and are responsible for vision in daylight to moonlight. Depending on whether the vision is mediated by cones or rods, illumination is broadly divided respectively into *photopic* and *scotopic* ranges. The boundary between photopic and scotopic is fuzzy, with some overlap occurring. A range of illumination between indoor light to moonlight in which both rods and cones are active is referred to as the *mesopic* range. Rods and cones divide the huge range of illumination into approximately two smaller ranges, and individually adapt to this range.

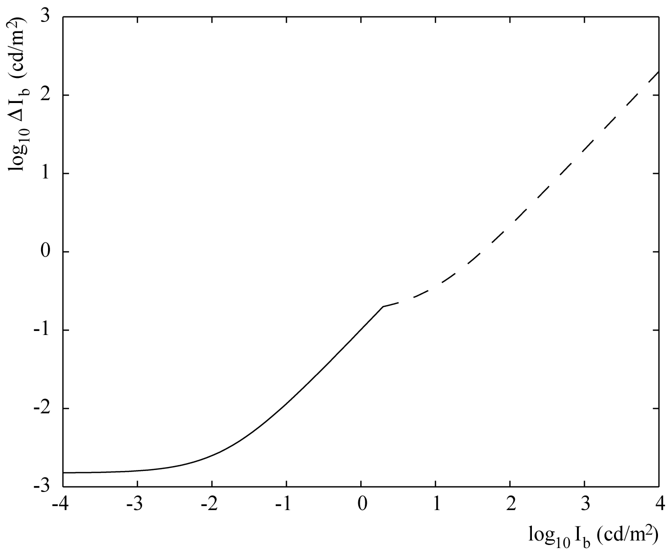


FIGURE 6.7 *TVI curves for rods and cones.*

The manifestation of adaptation of rods and cones in their respective ranges of illumination is shown in the TVI plots in Figure 6.7. The solid line corresponds to the thresholds for rods, and the dashed line corresponds to the threshold for cones. In scotopic illumination conditions, rods are more sensitive than cones and have a much lower threshold, and the vision in those illumination conditions is mediated by the rod system.

As illumination is increased, cones become increasingly sensitive (demonstrated by the crossover of the cone and rod TVI curves). At higher illuminations, rods begin to saturate and eventually the rod system becomes incapable of discriminating between two lights that differ in intensity by as much as a factor of one hundred [51]. The equation of the rod curve shown in Figure 6.7 is

$$\Delta I_{\mathrm{b}} = 0.1(I_{\mathrm{b}} + 0.015),$$

and the equation describing the cone curve is

$$\Delta I_{\mathrm{b}} = 0.02(I_{\mathrm{b}} + 8).$$

These equations are due to Rushton and MacLeod and fit their threshold data [111]. In this case, the equations are given in trolands (td), which is a measure of retinal illuminance. A value of 1 td is obtained when a surface with a luminance of 1 $\mathrm{cd/m^2}$ is viewed through a pupil opening of 1 $\mathrm{mm^2}$. Thus, trolands are given as luminance times area of the pupil.

The diameter d of the circular pupil as a function of background luminance may be estimated [149]. Moon and Spencer [83] propose the following relation between luminance and pupil diameter.

$$d = 4.9 - 3\tanh\bigl(0.4(\log L + 1.0)\bigr)$$

Alternatively, de Groot and Gebhard [89] estimate the pupil diameter to be

$$\log d = 0.8558 - 4.01 \times 10^{-4}(\log L + 8.6)^3.$$

In both of these equations the diameter d is given in mm, and L is the luminance in $\mathrm{cd/m^2}$.

The role of rods and cones in adaptation is important, and deserves consideration when dealing with intensities of extreme dynamic range. However, the individual operating ranges of rods and cones are still very large (a million to one). Thus, additional processes must play a significant role in their adaptation.

6.2.3 PHOTO-PIGMENT DEPLETION AND REGENERATION

Light is absorbed by the rod and cone photoreceptors through a photochemical reaction. This reaction breaks down photosensitive pigments and temporarily renders them insensitive — a process called bleaching. The pigments are regenerated in a relatively slow process. Thus the visual adaptation as a function of light intensity could be attributed to the depletion and regeneration of photo-pigment. Rod photo-pigments are completely depleted when exposed to light intensity above the mesopic range. It is believed that this depletion renders rods inoperable in the photopic range.

However, cone photo-pigments are not significantly depleted even in bright sunlight, but as demonstrated in the TVI relationship the sensitivity of the cones continues to diminish as a function of background intensity. This lack of correlation between photo-pigment concentration and visual sensitivity, as well as other experimental evidence, suggests that unless virtually all pigments are bleached the visual adaptation to different illumination conditions cannot be completely attributed to photo-pigment concentration [19].

6.2.4 PHOTORECEPTOR MECHANISMS

Photoreceptors convert absorbed light energy into neural responses. Intercellular recordings show that the response characteristics of rods and cones have the following behavior.[1] Compared to the broad range of background light intensities over which the visual system performs, photoreceptors respond linearly to a rather narrow range of intensities. This range is only about 3 log units, as shown in Figure 6.8. The log-linear plot in this figure of the intensity-response function is derived from measurements of the response of dark-adapted vertebrate rod cells on brief exposures to various intensities of light [19].

The response curve of cones follows the same shape as the response curve of rod cells. However, because of the higher sensitivity of rod cells to light the response curve for the cones appears to the right on the log intensity axis. Figure 6.9 shows the response curves for both rods and cones.

1 Electrophysiology is a field of study that may be used to detect the response of individual cells in the human visual system. Whereas the visual system is stimulated with a pattern of light, single-cell recordings are made whereby a thin electrode is held near the cell (extracellular recordings) or inside the cell (intercellular recordings), thus measuring the cell's electrical behavior [92].

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 12

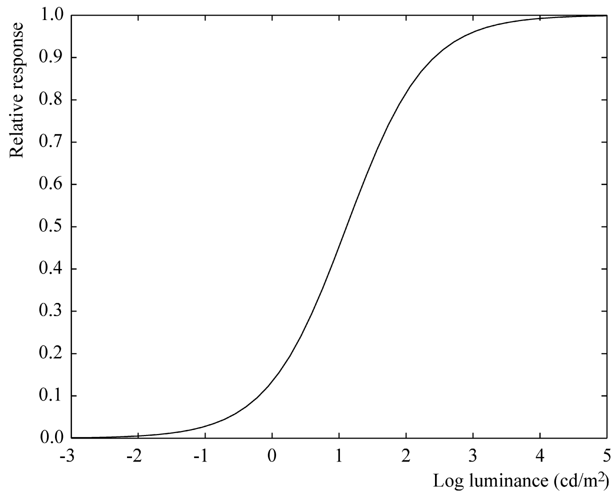


FIGURE 6.8 *Response of dark-adapted vertebrate rod cells to various intensities. The intensity axis in the image is shown in arbitrary units [19].*

The response curves for both rods and cones can be fitted with the following equation.

$$\frac{R}{R_{\max}} = \frac{I^n}{I^n + \sigma^n}. \tag{6.1}$$

Here, R is the photoreceptor response ($0 < R < R_{\max}$), $R_{\max}$ is the maximum response, I is light intensity, and σ is the semisaturation constant (the intensity that causes the half-maximum response). Finally, n is a sensitivity control exponent that has a value generally between 0.7 and 1.0 [19].

This equation, known as the Michaelis–Menten equation (or Naka–Rushton equation), models an S-shaped function (on a log-linear plot) that appears re-

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 13

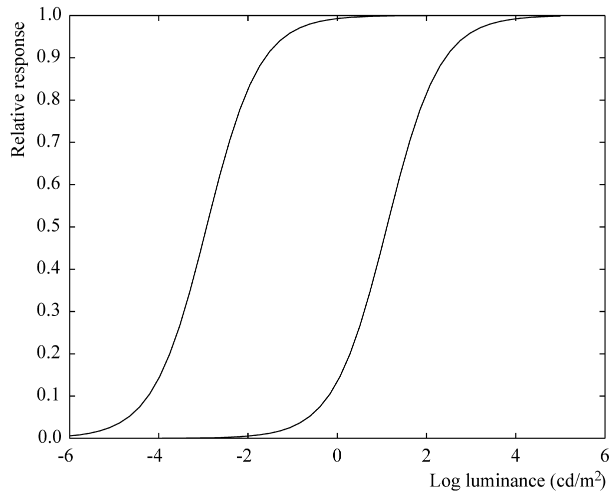


FIGURE 6.9 *Response of dark-adapted rod and cone cells to various intensities in arbitrary units.*

peatedly in both psychophysical experiments [1,50,134,147] and widely diverse direct-neural measurements [19,41,44,65,87,133]. The role of σ in Equation 6.1 is to control the position of the response curve on the (horizontal) intensity axis. It is thus possible to represent the response curves of rods and cones shown in Figure 6.9 by simply using two different values of σ, say σ_{rod} and σ_{cone}, in Equation 6.1.

Photoreceptor Adaptation The response curves shown in Figures 6.8 and 6.9 demonstrate that when the dark-adapted photoreceptor is exposed to a brief light of moderately high intensity the response reaches its maximum and the photoreceptor is saturated. The photoreceptor loses sensitivity to any additional light intensity. This initial saturation of the photoreceptor matches with our visual experience of blinding brightness when exposed to light about a hundred or more times more intense than the current background intensity. However, this initial experience does not continue for long. If exposed to this high background intensity for a while, the human visual system adapts to this new environment and we start to function normally again.

Measurements have shown that if photoreceptors are exposed continuously to high background intensities the initial saturated response does not continue to remain saturated. The response gradually returns toward the dark-adapted resting response, and the photoreceptor's sensitivity to incremental responses is gradually restored. Figure 6.10 shows the downward shift in the measured response at two different background intensities (shown in vertical lines). An interesting observation is that the response never completely returns to the resting response. Rather, it stabilizes on a plateau. Figure 6.10 shows the plateau curve (lower curve) for a range of background intensities. In addition to the restoration of sensitivity, the intensity-response curve measured at any given background intensity shows a right shift of the response-intensity curve along the horizontal axis, thus shifting the narrow response range to lie around the background intensity. The shifted curves are shown in Figure 6.11.

Independent measurements have verified that the shapes of the intensity-response curves at any background are independent of the background. However, with background intensity the position of the response function shifts horizontally along the intensity axis. This shift indicates that given sufficient time to adapt the visual system always maintains its log-linear property for about 3 log units of intensity range around any background. This shift is also modeled by the Michaelis–Menten equation by simply increasing the value of the semisaturation constant σ as a function of the background intensity. This yields the modified equation

$$\frac{R}{R_{\mathrm{max}}} = \frac{I^n}{I^n + \sigma_{\mathrm{b}}^n}, \tag{6.2}$$

where σ_{b} is the value of the half-saturation constant that takes different values for different background intensities, I_{b}. Thus, the photoreceptor adaptation modeled by the Michaelis–Menten equation provides us with the most important mechanism of adaptation.

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 15

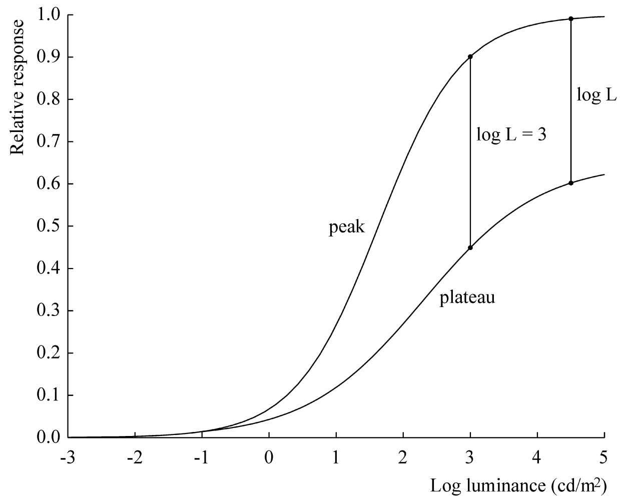


FIGURE 6.10 *Recovery of response after a long exposure to background intensities [19].*

Response–threshold Relation The observed linear relationship between the visual threshold and background intensity (the TVI relationship from Section 6.2) can be derived from the cellular adaptation model. (See Figure 6.12 for an intuitive derivation.) For this derivation we assume that the threshold ΔI_b is the incremental intensity required to create an increase in cellular response by a small criterion amount δ [45,134]. Based on this assumption, we derive ΔI_b from the response equation as follows. Rearranging Equation 6.2 yields

$$I = \sigma_\mathrm{b} \left(\frac{R}{R_\mathrm{max} - R} \right)^{\frac{1}{n}}.$$

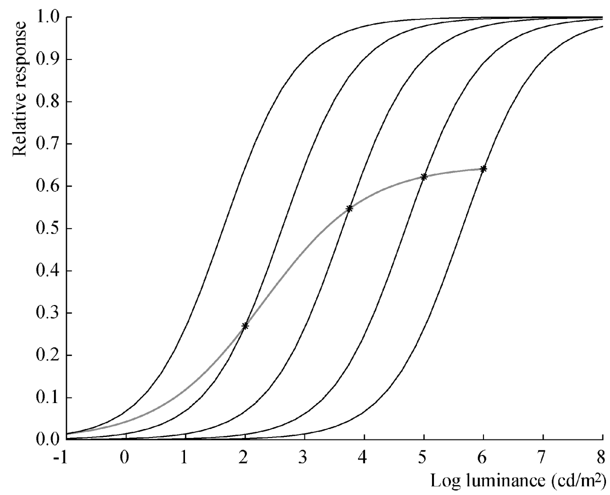
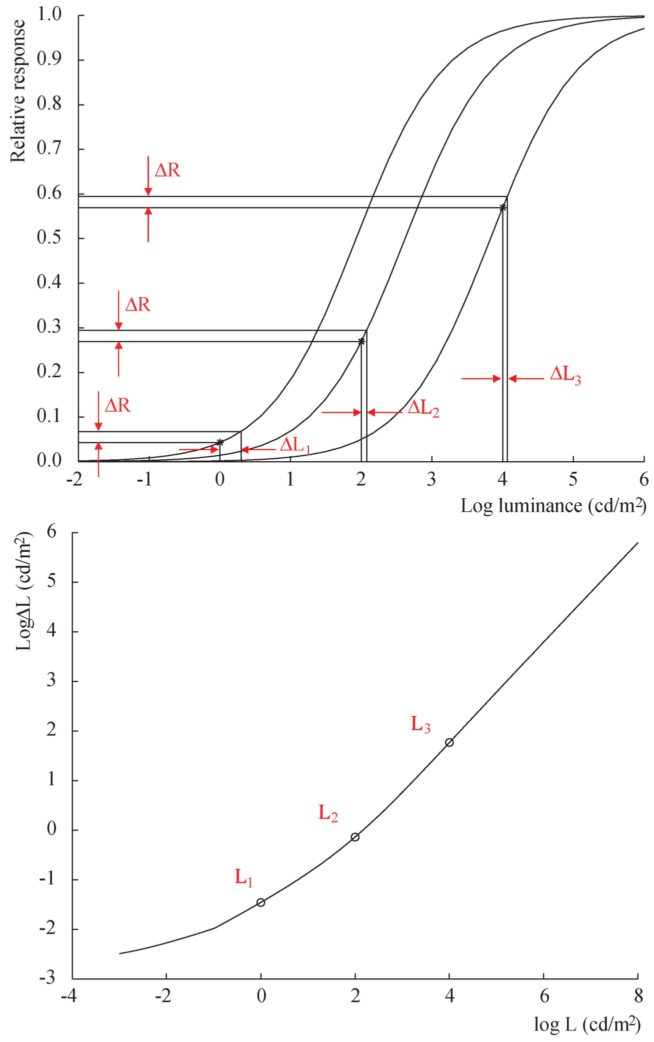


FIGURE 6.11 *Photoreceptor response adaptation to different background intensities. The plateau of Figure 6.10 is shown in gray. It represents the locus of the photoreceptor response to the background intensity itself.*

FIGURE 6.12 *Pictorial illustration of the response-threshold relation. The figure at the top plots three photoreceptor response functions at three different background luminances (L_1, L_2, L_3 from left to right) about 2 log units apart from each other. The response to the background luminance itself is shown by the $*$ symbol on the plots. ΔR is a small and fixed amount of response above the response to the background luminance and ΔL_is are the increments in luminance required to cause the same ΔR response change. The figure at the bottom plots the ΔL_i values as a function the background luminances L_i. The three values corresponding to background luminances L_1, L_2, and L_3 are shown in solid dots. The curve passing through the plotted ΔL_i values has a shape similar to the TVI function shown in the earlier part of this chapter.*



Relative response
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
ΔR
ΔR
ΔR
ΔL1
ΔL2
ΔL3
-2
-1
0
1
2
3
4
5
6
Log luminance (cd/m2)
LogΔL (cd/m2)
6
5
4
3
2
1
0
-1
-2
-3
-4
-2
0
2
4
6
8
L1
L2
L3
log L (cd/m2)

By differentiating this expression with respect to R, we get

$$\begin{aligned}\frac{\mathrm{d}I}{\mathrm{d}R} &= \sigma_\mathrm{b} \cdot \frac{1}{n} \cdot \left(\frac{R}{R_\mathrm{max} - R}\right)^{\frac{1}{n}-1} \left(\frac{R_\mathrm{max}}{(R_\mathrm{max} - R)^2}\right) \\ &= \sigma_\mathrm{b} \cdot \frac{1}{n} \cdot \frac{R_\mathrm{max}}{(R_\mathrm{max} - R)^{\frac{n+1}{n}}} R^{\frac{1-n}{n}}.\end{aligned}$$

This gives an expression for the incremental intensity (i.e., $\mathrm{d}I$) required to increase the response of the system by $\mathrm{d}R$. If we assume that the criterion response amount δ for the threshold condition is small enough, from the previous equation it is possible to compute the expression for ΔI as

$$\begin{aligned}\frac{\Delta I}{\delta} &\cong \frac{\mathrm{d}I}{\mathrm{d}R} \\ &= \sigma_\mathrm{b} \cdot \frac{1}{n} \cdot \frac{R_\mathrm{max}}{(R_\mathrm{max} - R)^{\frac{n+1}{n}}} R^{\frac{1-n}{n}}.\end{aligned}$$

Note that in all these equations R is the response of the cellular system exposed to intensity I, which may be different from the background intensity I_b to which the system is adapted. For threshold conditions, we can write $R = R_\mathrm{b} + \delta$, where R_b is the plateau response of the system at the background intensity I_b. Thus,

$$\Delta I = \delta \cdot \sigma_\mathrm{b} \cdot \frac{1}{n} \cdot \frac{R_\mathrm{max}}{(R_\mathrm{max} - R_\mathrm{b} - \delta)^{\frac{n+1}{n}}} (R_\mathrm{b} + \delta)^{\frac{1-n}{n}}.$$

For dark-adapted cells, the response of the system $R_\mathrm{b} = 0$. Thus, the expression of the threshold under a dark adaptation condition is

$$\Delta I_\mathrm{dark} = \delta \cdot \sigma_\mathrm{dark} \cdot \frac{1}{n} \cdot \frac{R_\mathrm{max}}{(R_\mathrm{max} - \delta)^{\frac{n+1}{n}}} \delta^{\frac{1-n}{n}}.$$

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 19

The relative threshold, $\Delta I/\Delta I_{\mathrm{dark}}$, for adaptation at any other background intensity I_{b} is

$$\begin{aligned}
\frac{\Delta I}{\Delta I_{\mathrm{dark}}} &= \frac{\sigma_{\mathrm{b}}}{\sigma_{\mathrm{dark}}} \cdot \left(\frac{R_{\mathrm{b}}+\delta}{\delta}\right)^{\frac{1-n}{n}} \left(\frac{R_{\max}-\delta}{R_{\max}-R_{\mathrm{b}}-\delta}\right)^{\frac{n+1}{n}} \\
&\approx \frac{\sigma_{\mathrm{b}}}{\sigma_{\mathrm{dark}}} \cdot \left(\frac{\delta}{R_{\mathrm{b}}}\right)^{\frac{n-1}{n}} \left(\frac{R_{\max}}{R_{\max}-R_{\mathrm{b}}}\right)^{\frac{n+1}{n}} \\
&= \frac{\sigma_{\mathrm{b}}}{\sigma_{\mathrm{dark}}} \cdot \left(\frac{\delta}{R_{\max}}\right)^{\frac{n-1}{n}} \left(\frac{I_{\mathrm{b}}^n+\sigma_{\mathrm{b}}^n}{I_{\mathrm{b}}^n}\right)^{\frac{n-1}{n}} \left(\frac{I_{\mathrm{b}}^n+\sigma_{\mathrm{b}}^n}{\sigma_{\mathrm{b}}^n}\right)^{\frac{n+1}{n}} \\
&= \frac{1}{\sigma_{\mathrm{dark}}} \cdot \left(\frac{\delta}{R_{\max}}\right)^{\frac{n-1}{n}} \frac{(I_{\mathrm{b}}^n+\sigma_{\mathrm{b}}^n)^2}{I_{\mathrm{b}}^{n-1}\sigma_{\mathrm{b}}^n}.
\end{aligned}$$

For $n = 1$ and $I_{\mathrm{b}} = \sigma_{\mathrm{b}}$, $\frac{\Delta I}{\Delta I_{\mathrm{dark}}}$ is directly proportional to I_{b}. This relation is in agreement with the Weber relation seen in TVI measurements. Thus, Weber's law may be considered as a behavioral manifestation of photoreceptor adaptation. The preceding discussion of the various mechanisms of visual adaptation affords the following conclusions.

- Photoreceptor adaptation plays a very important role in visual adaptation. An appropriate mathematical model of this adaptation (for example, Equation 6.2) can be effectively used to tone map HDR images. The TVI relation can be derived from the photoreceptor adaptation model, and hence can be used as an alternate mathematical model for tone mapping.
- The rod and cone combination extends the effective range over which the human visual system operates. Depending on the range of intensities present in an image, the appropriate photoreceptor system or combination of them may be chosen to achieve realistic tone mapping.

6.3 VISUAL ADAPTATION MODELS FOR HDR TONE MAPPING

Figure 6.13 outlines a basic framework for HDR tone mapping using models of visual adaptation. The two key features of the framework are forward and inverse adaptation models. The forward adaptation model will process the scene luminance values and extract visual appearance parameters appropriate for realistic tone mapping. The inverse adaptation model will take the visual appearance parameters and the adaptation parameters appropriate to the display viewing condition and will output the display luminance values. Either of the visual adaptation models discussed in the previous section (photoreceptor adaptation model or threshold adaptation model) may be used for forward and inverse adaptation. Most tone-mapping algorithms available today make use of one of these models. To achieve the goal of realistic HDR compression, these algorithms use photoreceptor responses or JNDs

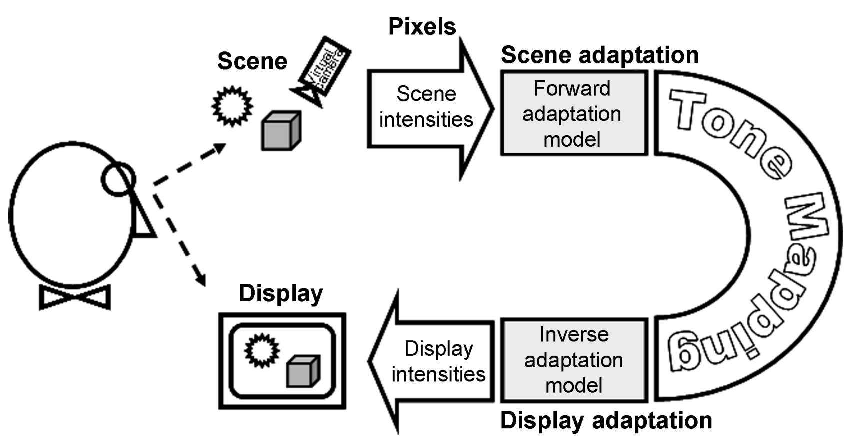


FIGURE 6.13 Framework for solving tone-mapping problem using visual adaptation.

as the correlates of the visual appearance. In this section, we explore various algorithms and show their relation to the visual adaptation models discussed in the previous section.

6.3.1 PHOTORECEPTOR ADAPTATION MODEL FOR TONE MAPPING

This section brings together a large number of tone-mapping algorithms. The common relationship between them is the use of an equation similar to the photoreceptor adaptation equation (Equation 6.2) presented in Section 6.2. In the following paragraphs we only show the similarity of the equation used to the photoreceptor adaptation equation, and defer the discussion of the details of these algorithms to the following two chapters. Here we show the actual form of the equations used in the algorithms, and where required rewrite them such as to bring out the similarity with Equation 6.2. In their rewritten form they are functionally identical to their original forms. It is important to note that although these equations may be derived from the same adaptation equations they largely differ in their choice of the value of the parameters, and only a few of them specifically claim the algorithm to be based on visual adaptation. Thus, all but a few of these algorithms (see [94,95]) ignore the inverse adaptation. They use Equation 6.2 because of its several desirable properties. These properties are as follows.

- Independent of input intensity, the relative response is limited to between 0 and 1. Thus, the relative response output can be directly mapped to display pixel values.
- The response function shifts along the intensity axis in such a way that the response of the background intensity is well within the linear portion of the response curve.
- The equation has a near linear response to the intensity in the log domain for about 4 log units. The intensity ranges of most natural scenes without any highlights or directly visible light sources do not exceed 4 log units. Thus, such scenes afford an approximately logarithmic relation between intensity and response.

Rational Quantization Function Schlick used the following mapping function for computing display pixel values from pixel intensity (I) [113].

$$
\begin{aligned}
F(I) &= \frac{pI}{pI - I + I_{\text{max}}} && \text{[Original form]} \\
&= \frac{I}{I + \frac{I_{\text{max}} - I}{p}}. && \text{[Rewritten form]}
\end{aligned}
$$

Here I_{max} is the maximum pixel value, and p takes a value in the range $[1, \infty]$. We can directly relate this equation to Equation 6.2 by substituting 1 for n and $\frac{I_{\text{max}} - I}{p}$ for σ_{b} in that equation. Note that the value of σ_{b} depends on the value of I itself, which may be interpreted as if the value of every pixel served as the background intensity in the computation of the cellular response.

Gain Control Function Pattanaik et al. introduced a gain control function for simulating the response of the human visual system and used this gain-controlled response for tone mapping [94]. They proposed two different equations for modeling the response of rod and cone photoreceptors. The equations are

$$
\begin{aligned}
F_{\text{cone}}(I) &= \frac{I}{c_1 (I_{\text{b}} + c_2)^n} \quad \text{and} \\
F_{\text{rod}}(I) &= \frac{r_1}{r_2 (I_{\text{b}}^2 + r_1)^n} \frac{I}{r_3 (I_{\text{b}} + r_4)^n},
\end{aligned}
$$

where the cs and rs are constants chosen to match certain psychophysical measurements. In their formulation, the Is represent light intensity of the image pixels of successively low-pass filtered versions of the image, and for every level of the image the background intensity I_{b} is chosen as the intensity of the pixel at the next level. These equations have a vague similarity with Equation 6.2 and have been given here for completeness.

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 23

S-shaped Curve Tumblin et al. used an S-shaped curve (sigmoid) as their tone-mapping function [129]. The equation of this curve is

$$
\begin{aligned}
F(I) &= \left[\frac{\left(\frac{I}{I_\mathrm{b}}\right)^n + \frac{1}{k}}{\left(\frac{I}{I_\mathrm{b}}\right)^n + k}\right] \cdot D & \text{[Original form]} \\
&= \left[\frac{I^n}{I^n + kI_\mathrm{b}^n} + \frac{I_\mathrm{b}^n}{k(I^n + kI_\mathrm{b}^n)}\right] \cdot D, & \text{[Rewritten form]}
\end{aligned}
$$

where k, D, and n are the parameters for adjusting the shape and size of the S-shaped curve. According to the authors, this function is inspired by Schlick's quantization function, shown previously. The rewritten equation has two parts. The first part is identical to Equation 6.2. The second part of the equation makes it an S-shaped function on a log-log plot.

Photoreceptor Adaptation Model Pattanaik et al. [95] and Reinhard and Devlin [108] made explicit use of Equation 6.2 for tone mapping. Pattanaik et al. used separate equations for rods and cones to account for the intensity in scotopic and photopic lighting conditions. The σ_b values for rods and cones were computed from the background intensity using

$$
\begin{aligned}
\sigma_{\mathrm{b,rod}} &= \frac{c_1 I_{\mathrm{b,rod}}}{c_2 j^2 I_{\mathrm{b,rod}} + c_3 (1 - j^2)^4 I_{\mathrm{b,rod}}^{1/6}} \\
\sigma_{\mathrm{b,cone}} &= \frac{c_4 I_{\mathrm{b,cone}}}{k^4 I_{\mathrm{b,cone}} + c_5 (1 - k^4)^2 I_{\mathrm{b,cone}}^{1/3}},
\end{aligned}
$$

where

$$
\begin{aligned}
j &= \frac{1}{c_6 I_{\mathrm{b,rod}} + 1} \\
k &= \frac{1}{c_7 I_{\mathrm{b,cone}} + 1}
\end{aligned}
$$

and $I_{\mathrm{b,rod}}$, $I_{\mathrm{b,cone}}$ are respectively the background intensities for the rods and cones. Pattanaik and Yee extended the use of these functions to tone map HDR images [96]. Reinhard and Devlin provided the following much simpler equation for computing σ_{b} at a given background intensity.

$$\sigma_{\mathrm{b}} = (f I_{\mathrm{b}})^m.$$

Here, f and m are constants and are treated as user parameters in this tone-mapping algorithm.

Photographic Tone-mapping Function The photographic tone-mapping function used by Reinhard et al. [106,109] is very similar to Equation 6.2. The equation can be written in the following form.

$$\begin{aligned} F(I) &= \frac{a\dfrac{I}{I_{\mathrm{b}}}}{1 + a\dfrac{I}{I_{\mathrm{b}}}} && \text{[Original form]} \\ &= \frac{I}{I + \dfrac{I_{\mathrm{b}}}{a}} && \text{[Rewritten form]} \end{aligned}$$

Here, a is a scaling constant appropriate to the illumination range (key) of the image scene.

6.3.2 THRESHOLD VERSUS INTENSITY MODEL FOR TONE MAPPING

In the previous section we have shown the relationship between the TVI model and the photoreceptor adaptation model. Thus, it is obvious that the TVI model can be used for tone reproduction. Ward's [139] tone-mapping algorithm is the first to make use of the TVI model. In his algorithm, Ward used a JND (just-noticeable difference), the threshold ΔI_{b} at any background I_{b}, as a unit to compute the correlate of the visual appearance parameter. From the scene pixel luminance I_{scene} and the scene background luminance $I_{\mathrm{b,scene}}$, Ward computed the ratio $\frac{I - I_{\mathrm{b,scene}}}{\Delta I_{\mathrm{b,scene}}}$.

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 25

This ratio represents the number of JNDs by which the pixel differs from the background. Using the display background luminance $I_{\mathrm{b,display}}$, and display adaptation threshold $\Delta I_{\mathrm{b,scene}}$ he inverted the JNDs to compute the display pixel luminance. The inversion expression is as follows.

$$I_{\mathrm{display}} = \mathrm{JNDs} \times \Delta I_{\mathrm{b,display}} + I_{\mathrm{b,display}} \tag{6.3}$$

Ferwerda et al. [35] later adapted this concept to compute JNDs specific to rods and cones for the purpose of tone-mapping images with a wide range of intensities. If the background intensity is locally adapted, the log-linear relationship of the threshold-to-background intensity provides the necessary range compression for HDR images. The issue of local versus global adaptation is discussed in the next section.

6.4 BACKGROUND INTENSITY IN COMPLEX IMAGES

In the previous sections we introduced two important adaptation models: the photoreceptor response model and the TVI model. Both of these adaptation models require knowledge of the background intensity I_{b}. For any use of either of these models in tone reproduction, I_{b} has to be computed from the intensity of the image pixels. In this section we describe various methods commonly used to estimate I_{b} from an image.

6.4.1 IMAGE AVERAGE AS I_{b}

The average of the intensity of the image pixels is often used as the value of I_{b}. The average could be the arithmetic average

$$\frac{1}{n}\sum_{i=1}^{n} I_i$$

or geometric average

$$\prod_{i=1}^{n} (I_i + \varepsilon)^{\frac{1}{n}},$$

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 26

where n in the equations is the total number of pixels in the image, and ε (an arbitrary small increment), is added to the pixel intensities to take into account the possibility of any zero pixel values in the image. The geometric average can also be computed as

$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\log(I_i+\varepsilon)\right),$$

where the exponent $\frac{1}{n}\sum_{i=1}^{n}\log(I_i+\varepsilon)$ is the log average of the image pixels.

In the absence of any knowledge of the actual scene, one of these image averages is probably the most appropriate estimate of I_{b} for most images. A visual adaptation model using such an average is referred to as a global adaptation, and the tone-mapping method is referred to as global tone mapping. The geometric average is often the preferred method of average computation. This is largely because (1) the computed background intensity is less biased toward outliers in the image and (2) the relationship between intensity and response is log-linear.

6.4.2 LOCAL AVERAGE AS I_{b}

In images with a very high dynamic range, the intensity change from region to region can be drastic. Hence, the image average (also called global average) is not sufficiently representative of the background intensity of the entire image. The proper approach in such cases would be to segment the image into regions of LDR and use the average of pixels in each region. Yee and Pattanaik's work shows that such segmentation in natural images is not always easy, and that tone mapping using the local average from regions obtained using existing segmentation techniques may introduce artifacts at region boundaries [150].

An alternative and popular approach is to compute a local average for every pixel p in the image from its neighboring pixels. The various techniques under this category include box filtering and Gaussian filtering. These techniques are easily computed. The computation may be expressed as

$$I_{\mathrm{b},p}=\frac{1}{\sum\limits_{i\in\Omega}w(p,i)}\sum_{i\in\Omega}w(p,i)I_i. \tag{6.4}$$

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 27

For Gaussian filtering

$$w(p,i) = \exp\left(-\frac{\|p-i\|^2}{s^2}\right).$$

For box filtering, it is expressed as

$$w(p,i) = \begin{cases} 1 & \text{for } \|p-i\| < s, \\ 0 & \text{otherwise.} \end{cases}$$

In these equations, Ω represents all pixels of the image around p, $\|.\|$ is the spatial distance function, and s is a user-defined size parameter in these functions. Effectively, the value of s represents the size of a circular neighborhood around the pixel p that influences the average value.

Although for most pixels in the image the local average computed in this fashion is representative of the background intensity, the technique breaks down at HDR boundaries. This is due to the fact that the relatively large disparity in pixel intensities in the neighborhood of the boundary biases the average computation. Thus, the background intensity computed for pixels on the darker side of the boundary is positively biased, and those computed for the pixels on the brighter side are negatively biased. This biasing gives rise to halo artifacts in the tone-mapped images. Figure 6.14 highlights the problem. The image shown is computed using local box-filtered values for the background intensity. Note the dark band on the darker side of the intensity boundary. Although not noticeable, similar bright banding exists on the brighter side of the boundary.

This problem can be avoided by computing the average from only those pixels whose intensities are within a reasonable range of the intensity of the pixel under consideration. The tone-mapped image in Figure 6.15 shows the result using background intensity from such an adaptive computational approach. There is a significant improvement in the image quality, but at an increased cost of computation. Two such computational approaches are discussed in the following sections.

Local Average Using Variable Size Neighborhood In this approach, the size parameter s in Equation 6.4 is adaptively varied. Reinhard et al. and Ashikhmin

FIGURE 6.14 *Halo artifacts associated with the use of I_b computed by local averaging. The artifacts are most noticeable at the illumination boundaries.*

FIGURE 6.15 *Tone mapping using adaptive local averaging.*

simultaneously proposed this very simple algorithm [6,109]. Starting from a value of s equal to 1, they iteratively double its value until the pixels from across the HDR boundary start to bias the average value. They assume that the average is biased if it differs from the average computed with the previous size by a tolerance amount. They use this s in Equation 6.4 for computing their local average.

Local Average Using Bilateral Filtering In this approach, the size parameter s remains unchanged, but the pixels around p are used in the average summation only if their intensity values are similar to the intensity of p. The similarity can be user defined. For example, the intensities may be considered similar if the difference or the ratio of the intensities is less than a predefined amount. Such an approach may be implemented by filtering both in spatial and intensity domains. The name "bilateral" derives from this dual filtering. The filter can be expressed as

$$I_{\mathrm{b},p} = \frac{1}{\sum_{i \in \Omega} w(p,i) g(I_p, I_i)} \sum_{i \in \Omega} w(p,i) g(I_p, I_i) I_i, \tag{6.5}$$

where $w()$ and $g()$ are the two weighting functions that take into account the dual proximity. The forms of these weighting functions can be similar, but their parameters are different: for $g()$ the parameters are the intensities of the two pixels, and for $w()$ the parameters are the positions of the two pixels.

Durand and Dorsey use Gaussian functions for both domains [23]. Pattanaik and Yee use a circular box function for $w()$, and an exponential function for $g()$ [96]. Choudhury and Tumblin have proposed an extension to this technique to account for gradients in the neighborhood. They named their extension "trilateral filtering" [10].

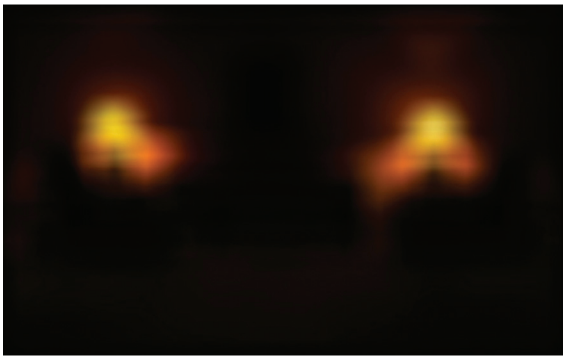
Figure 6.16 shows the linearly scaled version of the original HDR image and the images assembled from intensities computed for each pixel using some of the adaptive local adaptation techniques discussed in this section.

6.4.3 MULTISCALE ADAPTATION

Although the use of local averages as the background intensity is intuitive, the choice of the size of the locality is largely ad hoc. In this section we provide some empirical

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 30

(a)

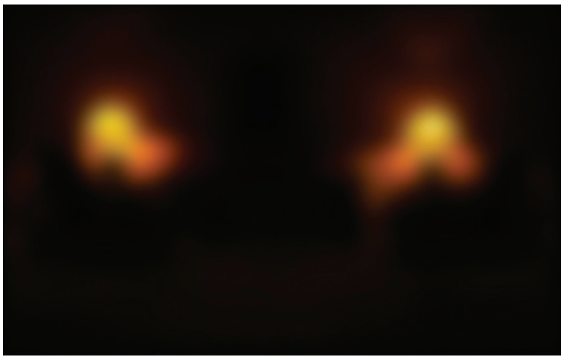


(b)

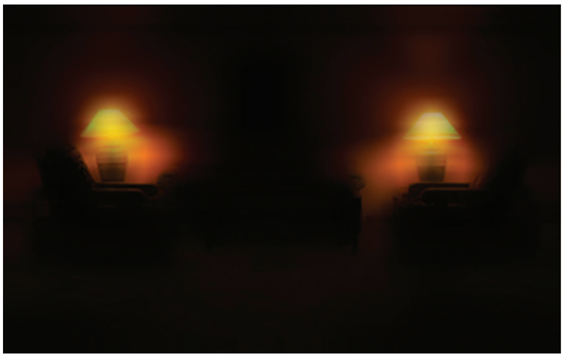
(continued on next page)

FIGURE 6.16 Local averages for a sample HDR image (a). Images (b) and (c) were computed using Equation 6.4, and images (d) and (e) were computed using Equation 6.4. Equal weighting is used for images (b) and (d) and Gaussian weighting for images (c) and (e). $g()$ for image (d) is from Pattanaik [96] and for image (e) from Durand and Dorsey [23]. (HDR image courtesy Columbia University.)

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 31



(c)



(d)

(*continued on next page*)

FIGURE 6.16 (Continued.)

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 32

(*e*)

FIGURE 6.16 (*Continued.*)

support for the use of local averages and the associated importance to the size of the locality.

Physiological and psychophysical evidence indicates that the early stages of visual processing can be described as the filtering of the retinal image by bandpass mechanisms sensitive to patterns of different scales [146]. These bandpass mechanisms adapt independently to the average intensity within a region of a scene defined by their spatial scale. In a complex scene, this average will be different at different scales and thus the mechanisms will all be in different states of adaptation. Thus, to correctly account for the changes in vision that occur with changes in the level of illumination we need to consider local adaptation at different spatial scales within HDR environments. Peli suggests that an appropriate way of characterizing the effects of local adaptation on the perception of scenes is to use low-pass images that represent the average local luminance at each location in the image at different spatial scales [98]. Reinhard et al. [109] and Ashikmin [6] use this multiscale approach to adaptively decide the effective neighborhood size. Pattanaik et al.'s [94]

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 33

multiscale adaptation also demonstrates the usefulness of the multiscale nature of the visual system in HDR tone mapping.

6.5 DYNAMICS OF VISUAL ADAPTATION

In earlier sections we discussed the adaptation of the visual system to background intensity. However, visual adaptation is not instantaneous. In the course of the day, light gradually changes from dim light at dawn to bright light at noon, and back to dim light at dusk. This gradual change gives the visual system enough time to adapt, and hence the relatively slow nature of visual adaptation is not noticed. However, any sudden and drastic change in illumination, from light to dark or dark to light, makes the visual system lose its normal functionality momentarily. This loss of sensitivity is experienced as total darkness during a light-to-dark transition, and as a blinding flash during a dark-to-light transition. Following this momentary loss in sensitivity, the visual system gradually adapts to the prevailing illumination and recovers its sensitivity. This adaptation is also experienced as a gradual change in perceived brightness of the scene.

The time course of adaptation, the duration over which the visual system gradually adapts, is not symmetrical. Adaptation from dark to light, known as *light adaptation*, happens quickly (in a matter of seconds), whereas *dark adaptation* (adaptation from light to dark) occurs slowly (over several minutes). We experience the dark-adaptation phenomenon when we enter a dim movie theater for a matinee. Both adaptation phenomena are experienced when we drive into and out of a tunnel on a sunny day. The capability of capturing the full range of light intensities in HDR images and video poses new challenges in terms of realistic tone mapping of video-image frames during the time course of adaptation.

In Section 6.2 we argued that vision is initiated by the photochemical interaction of photons with the photo-pigments of the receptor. This interaction leads to bleaching and hence to loss of photo-pigments from receptors. The rate of photon interaction and hence the rate of loss in photo-pigments is dependent on the intensity of light, on the amount of photo-pigment present, and on photosensitivity. A slow chemical regeneration process replenishes lost photo-pigments. The rate of regeneration depends on the proportion of bleached photo-pigments and on the time constant of the chemical reaction.

From the rate of bleaching and the rate of regeneration it is possible to compute the equilibrium photo-pigment concentration for a given illumination level. Because the rate of photon interaction is dependent on the amount of photo-pigments present, and because the bleaching and regeneration of bleached photo-pigments are not instantaneous, visual adaptation and its time course were initially thought to be directly mediated by the concentration of unbleached photo-pigments present in the receptor.

Direct cellular measurements on isolated and whole rat retinas by Dowling ([19] Chapter 7) show that dark adaptation in both rods and cones begins with a rapid decrease in threshold followed by a slower decrease toward the dark-adaptation threshold. The latter slow adaptation is directly predicted by photo-pigment concentrations, whereas the rapid adaptation is attributed almost entirely to a fast neural adaptation process that is not well understood.

The Michaelis–Menten equation (Equation 6.1) models the photoreceptor response and accounts for visual adaptation by changing the σ_{b} value as a function of background intensity. Photoreceptor adaptation and pigment bleaching have been proposed to account for this change in σ_{b} value. Valeton and van Norren have modeled the contribution of these two mechanisms to the increase in σ_{b} with

$$\sigma_{\mathrm{b}} = \sigma_{\mathrm{dark}} \sigma_{\mathrm{b,neural}} \sigma_{\mathrm{b,bleach}}, \tag{6.6}$$

where σ_{dark} is the semisaturation constant for dark conditions, $\sigma_{\mathrm{b,neural}}$ accounts for the loss in sensitivity due to neural adaptation, and $\sigma_{\mathrm{b,bleach}}$ accounts for the loss in sensitivity due to loss of photo-pigment [133]. The value of $\sigma_{\mathrm{b,bleach}}$ is inversely proportional to the fraction of unbleached photo-pigments at the background light.

Pattanaik et al. extended the use of the adaptation model to compute the time course of adaptation for simulating visual effects associated with a sudden change in intensities from dark to light, or vice versa [95]. They use a combination of Equations 6.1 and 6.6 to carry out the simulation, as follows.

$$\frac{R}{R_{\max}} = \frac{I^n}{I^n + \sigma_{\mathrm{b}}^n(t)}$$

$$\sigma_{\mathrm{b}}(t) = \sigma_{\mathrm{dark}} \sigma_{\mathrm{b,neural}}(t) \sigma_{\mathrm{b,bleach}}(t)$$

reinhard v.2005/03/22 Prn:14/06/2005; 9:42 F:reinhard06.tex; VTEX/SL p. 35

Here, time-dependent changes of $\sigma_{\mathrm{b,neural}}$ and $\sigma_{\mathrm{b,bleach}}$ are modeled with exponential decay functions.

6.6 SUMMARY

This chapter proposed the view that the modeling of human visual adaptation is key to realistic tone mapping of HDR images. We saw that photoreceptor adaptation is the most important factor responsible for visual adaptation, with Equation 6.1 being the mathematical model for this adaptation. The relation between various tone-mapping algorithms and the photoreceptor adaptation model was made evident. Background intensity is a key component in this model. Some of the commonly used methods for computing this background intensity in images were discussed. We also saw the usefulness of a human visual model in realistic simulation of visual effects associated with the wide range of real-life illuminations. Whereas this chapter explored the similarities between several current tone-reproduction operators, the following two chapters discuss their differences and present each tone-reproduction operator in detail.