# Perceptually Based Tone Mapping of High Dynamic Range Image Streams

Piti Irawan, James A. Ferwerda, and Stephen R. Marschner

Department of Computer Science and Program of Computer Graphics, Cornell University

**Abstract**

*This paper presents a new perceptually based tone mapping operator that represents scene visibility under time-varying, high dynamic range conditions. The operator is based on a new generalized threshold model that extends the conventional threshold-versus-intensity (TVI) function to account for the viewer's adaptation state, and a new temporal adaptation model that includes fast and slow neural mechanisms as well as photopigment bleaching. These new visual models allow the operator to produce tone-mapped image streams that represent the loss of visibility experienced under changing illumination conditions and in high dynamic range scenes. By varying the psychophysical data that the models use, we simulate the differences in scene visibility experienced by normal and visually impaired observers.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation-Display Algorithm. I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism. Keywords: Tone mapping, perception, adaptation, low vision.

## 1. Introduction

The variations in light levels we experience in the world are vast. For example, the average luminance in a outdoor scene can be 100 million times greater during the day than at night. The dynamic range of luminances can also be large, with ratios on the order of 10,000:1 from highlights to shadows. Luminance levels can change dramatically over time and from place to place.

Vision functions across these changes through a variety of adaptation mechanisms that include the pupil, the rod and cone receptors, photopigment bleaching and regeneration, and neural gain controls. However, vision is not equally good under all conditions. In particular, older people and those with visual disorders may be profoundly impaired by the low intensity, high dynamic range, and rapidly changing illumination conditions we often experience. Disabling glare and loss of visibility are frequent complaints of people with so-called *low vision*.

We have been working to develop a low vision simulator that shows people with normal vision what the world looks like to people with visual impairments. Such a simulator would have important applications in medical training, disability determination, and design of accessible environ-

ments. A key component of this simulator is a perceptually based tone reproduction operator that creates images that accurately predict visibility in real or modeled scenes-.

In this paper we describe a new tone reproduction operator created for this low vision simulator. The operator seeks to accurately represent scene visibility under time-varying, high dynamic range conditions. To achieve this we developed new models of temporal adaptation and contrast visibility. By varying the psychophysical data the adaptation models use, we are able to simulate the differences in scene visibility experienced by normal and visually impaired observers under changing illumination conditions. In the following sections we first review the properties of existing tone reproduction operators, then describe how we have extended two advanced perceptually-based operators, [WLRP97] and [PTYG00], and combined them with our new adaptation model to achieve our goals.

## 2. Background
### 2.1. Tone mapping

The development of techniques for high dynamic range (HDR) image capture and synthesis have made tone mapping an important problem in computer graphics (reviewed in [DCWP02]). The fundamental problem is how to map the
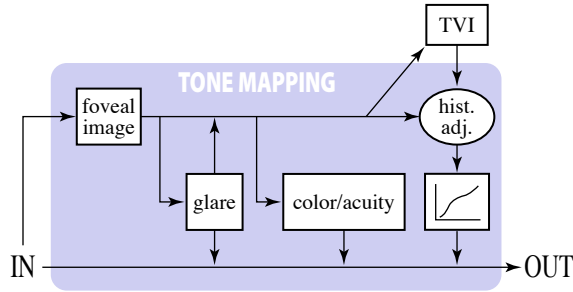
**Figure 1:** *Ward's visibility preserving operator.*

large range of intensities found in an HDR image into the limited range generated by a conventional display device. Different tone mapping operators have been introduced in the research literature, and a taxonomy is emerging that allows the operators to be classified in a broader context.

### 2.1.1. A taxonomy of tone mapping operators

A primary distinction is whether an operator is *global* or *local*. Global operators apply a single mapping function to all pixels of the image, whereas local operators modify the mapping depending on the characteristics of different parts of the image. A second important distinction is between *empirical* and *perceptually based* operators. Empirical operators seek to meet criteria such as dynamic range compression, detail preservation, or freedom from artifacts. On the other hand, perceptually based operators strive to produce images that are predictive visual simulations of scene appearance. A third distinction is between *static* and *dynamic* operators. Most existing operators are static, designed to process still images. In contrast, dynamic operators are explicitly designed for processing image streams.

### 2.1.2. Tone mapping operators for vision simulation

Since our overall goal is to create a real-time vision simulator that is capable of accurately representing how high dynamic range scenes appear to people with low vision, we need a tone mapping operator that is both perceptually based and dynamic. Whether the operator is global or local is of secondary importance. To our knowledge no existing operators meet these requirements, so we developed a new one.

To create this operator we took substantial inspiration from the perceptually based operators of Ward et al. [WLRP97] and Pattanaik et al. [PTYG00]. Neither of these operators alone can serve our purposes, but between them we found components that provided a good foundation for our work.

### 2.1.3. Ward's visibility preserving operator

The goal of Ward's operator is to produce displayable images that accurately represent the threshold visibility of scene features. To do this, the operator uses histogram adjustment constrained by a model of visual adaptation.

The main components of Ward's operator are illustrated in Figure 1. In the first stage the input image is downsampled to create a low-resolution "foveal image" in which each pixel
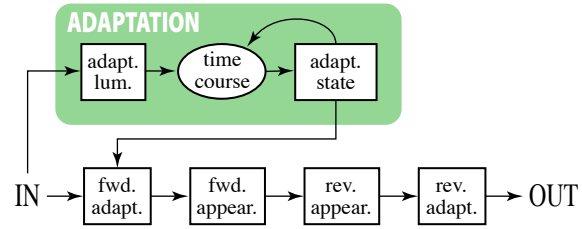


**Figure 2:** *Pattanaik's time dependent operator.*

covers $1°$ of visual angle. Next, a visual model that simulates the effects of veiling glare and low-luminance color and acuity loss is applied. At the heart of the operator is a histogram adjustment algorithm, in which the cumulative distribution function of the histogram of the foveal image, constrained in slope by the human threshold versus intensity (TVI) function (see Figure 4), is used to define the mapping from scene luminances to display luminances.

Ward's operator is an elegant solution to the perceptually based tone mapping problem, and we use significant components of it in our new operator. However, it is a static operator, designed to process individual images, and it is unable to correctly represent changes in scene visibility over time. To address this issue, we looked to Pattanaik's time dependent operator that incorporates a model of the temporal dynamics of vision.

### 2.1.4. Pattanaik's time dependent operator

The goal of Pattanaik's operator is to process a stream of input images of a scene and produce an output stream that simulates the changes in visual appearance caused by variations in scene luminance. To accomplish this, the operator merges components of an advanced color appearance model with a physiologically based model of the temporal dynamics of visual adaptation.

Figure 2 shows the main components of Pattanaik's operator. At the highest level the operator consists of a forward and inverse pair of visual models. First, an adaptation model transforms input scene intensities into retina-like responses. These responses are then transformed by a simplified version of Hunt's color appearance model [Hun95], to produce a representation of the scene's suprathreshold "whiteness/blackness" and "colorfulness" appearance correlates. To complete the tone mapping process, the adaptation state of the display observer is determined, and the inverse appearance model transforms the appearance correlates into display values that are calculated to produce corresponding responses in the display observer.

The time-dependent features of the operator derive from the forward adaptation model. The characteristics of the model are illustrated in Figure 3 where the curves show the S-shaped response profile of the combined rod and cone system at different luminance levels. At any given level the system only has a linear response range of 2 to 3 log units, so inputs above or below this range will be subject to response compression. Figure 3(a) shows the response of the system
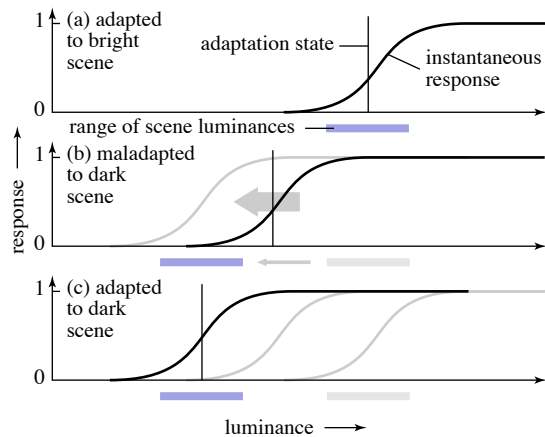
**Figure 3:** *Adaptation over time in Pattanaik's operator.*



**Figure 4:** *Threshold-versus-intensity functions. Based on data from [Bla46].*

to a scene illuminated at daylight levels. Note that through adaptation, the linear response range is nearly centered over the scene luminance range, so sensitivity (and therefore visibility) will be good and there will be little compression of response. Figure 3(b) shows the situation an instant after the illumination in the scene has suddenly dropped. While scene luminance range is now much lower, constraints on the speed of adaptation only allow the system to partially adjust its sensitivity, so responses will be severely compressed and much of the scene will be invisible. However, as Figure 3(c) shows, given enough time the system will continue to adapt, sensitivity will return, and visibility will be (at least partially) restored.

In the visual system both neural and photochemical adaptation mechanisms are responsible for altering the sensitivity of the rod and cone systems and shifting the response profiles across the luminance range. Neural adaptation is a fast and symmetric process that can alter sensitivity within milliseconds, but the magnitude of its effect is limited. Photopigment bleaching and regeneration, on the other hand, can have a much greater impact on sensitivity, but it is an asymmetric process, with potentially rapid bleaching followed by relatively slow regeneration and recovery of sensitivity. Pattanaik models the temporal dynamics of adaptation with four low-pass exponential filters; two each for the neural and photochemical mechanisms in the rod and cone systems.

Pattanaik's operator stands as the most advanced perceptually based operator for tone mapping image streams. However, with respect to our goals it is limited for two reasons: 1) it uses a suprathreshold color appearance model rather than a threshold visibility model so the images produced may not accurately represent visibility; and 2) it produces a simple S-shaped global mapping function so high dynamic range scenes may not be mapped correctly. However, its temporal adaptation model addresses our need to correctly represent the appearance of dynamic scenes, so with some significant modifications, we will incorporate this component in our new operator.
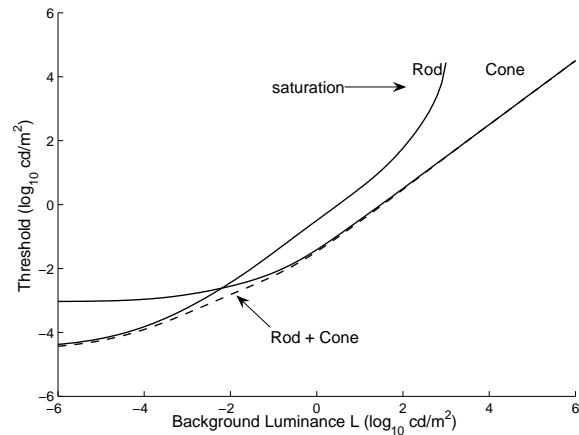
## 2.2. Adaptation and maladaptation

The purpose of visual adaptation is to optimize visual sensitivity with respect to prevailing levels of stimulation. Physiological studies of adaptation have tended to focus on changes in the response properties of individual photoreceptors [NR66], while psychophysical studies have focused on changes in the visibility of patterns imaged on particular regions of the retina [Bla46]. While these studies have provided fundamental knowledge about adaptation, the highly controlled conditions used in the experiments are not representative of normal vision.

Under normal conditions vision is an active process. We have mobile eyes that are constantly scanning the visual field with a variety of eye movements. An important consequence of eye movement is that the retinal image is constantly changing, so any particular retinal location is receiving continuously varying levels of stimulation. Because adaptation takes time, under natural conditions the visual system is rarely fully adapted the way it is in laboratory experiments, and the result of this *maladaptation* is that sensitivity, and therefore visibility, will typically be less than optimal.

### 2.2.1. Measuring adaptation and maladaptation

Adaptation is often described with the threshold versus intensity (TVI) functions, which give the threshold $\Delta L$, required to create a visible contrast at various background levels $L$. The solid curves in Figure 4 show representative TVI functions for the rod and cone systems. Note the both curves are flat at extremely low luminance levels and become linear over the range where the visual system adapts well. The rod curve bends upward for high luminances due to saturation, because the rod system has a limited ability to adapt to brighter conditions.

Classically, the TVI functions are measured using spot-on-background patterns. An observer is adapted to a circular background field of a particular luminance ($L$), and then tested to see how much more intense ($\Delta L$) a central spot
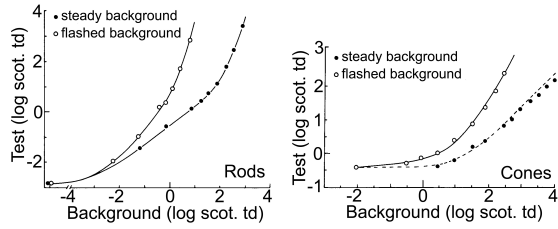
**Figure 5:** *TVI (filled) and probe-on-flash (open) curves for the rod and cone systems [WECH*90].*

has to be to be visible. By repeating this experiment for a range of background luminances the TVI functions can be described.

It is important to understand that because the observers are fully adapted to the backgrounds, the TVI functions measure the optimal sensitivity of the visual system at the tested luminance levels. As described above, under natural conditions, the visual system will only rarely achieve this sensitivity and will generally be *maladapted* to some degree. This has important consequences for visibility and other measures of visual performance.

Psychophysicists have characterized the effects of maladaptation on visual thresholds with *probe-on-flash* experiments [WECH*90]. The observer again adapts to a steady background, but instead of testing for visibility of a spot against that background, a circular spot and annulus pattern are briefly flashed, and the threshold for seeing the spot against the flashed annulus is measured. The short duration of the flash bypasses the normal adaptation processes to allow the measurement of threshold sensitivity at luminance levels away from the background adapting level. The results of typical probe-on-flash studies are indicated by the open circles in Figure 5. Note how the probe-on-flash curves diverge from the TVI curves (filled circles), indicating that due to maladaptation thresholds are higher than would be predicted from TVI experiments. Because of the active nature of the eye, maladaptation like this is a constant condition of human vision and has a significant impact on visibility under real-world conditions. For this reason we account for the effects of maladaptation in our operator.

## 3. A new operator for vision simulation

Recall that our goal is to develop a low vision simulator that produces images that show people with normal vision what the world looks like to people with visual impairments, and that a key component of this simulator is a perceptually based tone reproduction operator that can handle high dynamic range image streams. Like Ward's operator, our new operator should preserve visibility while compressing high dynamic range scenes for display. Like Pattanaik's operator, it should model the viewer's changing adaptation state and include the effects of maladaptation in the output images.

Producing an operator that combines these features is not a simple matter of bolting together the parts, because the two operators have different goals and incompatible notions
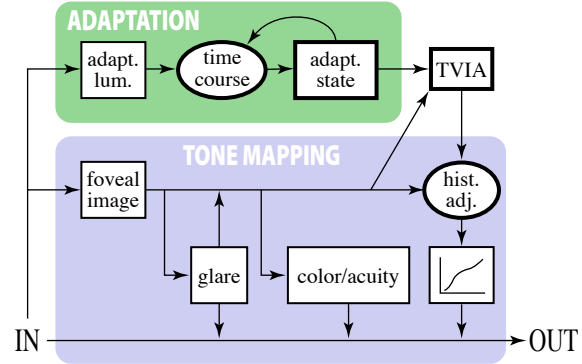


**Figure 6:** *Block diagram of our new operator, which combines components from Ward's and Pattanaik's operators with a new threshold visibility model, the TVIA. Components outlined with heavy lines have new contributions.*

of adaptation. Since we need a threshold visibility model, we used Ward's TVI-constrained histogram equalization as the basis for the operator. But the TVI function Ward uses is based on experiments in which observers were optimally adapted. In order for a Pattanaik-like temporal adaptation model to make sense in the histogram equalization framework, a generalized TVI function without this steady-state assumption is needed, but to our knowledge no one has proposed such a function. We had to return to the psychophysics literature and develop a more advanced threshold model that predicts the contrast threshold as a function of both the stimulus luminance and the observer's adaptation state.

### 3.1. Overview of new operator

The essential components of our new operator are shown in Figure 6. Ward's approach forms the basis for our new operator, with modifications to the histogram adjustment procedure. However, where Ward uses a static TVI-based threshold model we introduce a new dynamic threshold model, which we call the TVIA function, for *threshold versus intensity and adaptation*. To provide the adaptation state we use a time-course model based on exponential filters similar to Pattanaik's, but we extend the model to account for more temporal adaptation effects.

The TVIA model is based on S-shaped response functions inspired by Naka and Rushton's measurements [NR66], as are the response curves from Hunt's model that were used by Pattanaik. However, our use of the response curves is entirely different from Pattanaik's: we use small differentials along the curves to derive thresholds that drive the histogram adjustment process, rather than using the curves directly to compute appearance.

### 3.2. Histogram adjustment

See Figure 7 for definitions of symbols used in this section.

The simplest way to map world luminance to display luminance is by a linear function with constant $L_d/L_w$, but this only works if the display luminance range spans a larger

| | |
|---|---|
| $L_w$ | = world luminance (in $cd/m^2$) |
| $L_{w_i}$ | = world luminance for histogram bin $b_i$ |
| $L_{wmin}$ | = minimum world luminance for scene |
| $L_{wmax}$ | = maximum world luminance for scene |
| $L_d$ | = display luminance (in $cd/m^2$) |
| $L_{d_i}$ | = display luminance for histogram bin $b_i$ |
| $L_{dmin}$ | = minimum display luminance for scene |
| $L_{dmax}$ | = maximum display luminance for scene |
| $N$ | = number of histogram bins |
| $T$ | = total number of adaptation samples |
| $f(b_i)$ | = frequency count for histogram bin $b_i$ |
| $\Delta L(L_a)$ | = "just noticeable difference" for adaptation level $L_a$ |

**Figure 7:** *Symbols used in histogram adjustment.*

dynamic range than the world luminance values. For HDR scenes, a global tone mapping operator must selectively allocate the available display luminance values to world luminance values. As introduced in Subsection 2.1.3, Ward uses histogram adjustment to solve this problem.

**3.2.1. Ward's histogram adjustment method**
Since allocating larger display luminance ranges means less contrast compression, it is reasonable to assign more values to world luminances that occur frequently in the image. That is, we want to compress contrast in sparsely populated regions of the image's histogram, thereby conserving the available luminance range for the densely populated regions.

Naive histogram adjustment, however, may magnify contrast in well-populated regions, violating the goal of preserving visibility. Ward avoids this problem by limiting the slope of the mapping function to the ratio of contrast visibility thresholds for the display and world observers:

$$\frac{dL_d}{dL_w} \leq \frac{\Delta L(L_d)}{\Delta L(L_w)} \qquad (1)$$

With this constraint, two world luminances that are not visibly different will map to two display luminances that are also not visibly different. This is the sense in which Ward's operator (and ours) preserves contrast visibility.

From (1) Ward derives a constraint on the values of the histogram that will be used for histogram adjustment:

$$f(b_i) \leq \frac{T}{N} \cdot \frac{\log(L_{wmax}) - \log(L_{wmin})}{\log(L_{dmax}) - \log(L_{dmin})} \cdot \frac{\Delta L(L_{d_i})/L_{d_i}}{\Delta L(L_{w_i})/L_{w_i}} \qquad (2)$$

Ward simply truncates the histogram to ensure that this condition is met. However, this changes the total number of adaptation samples $T$ and, worse, changes the mapping function, which in turn changes $L_{d_i}$ and $\Delta L(L_{d_i})$, creating a nonlinear problem. Ward iteratively truncates counts and recomputes the ceilings until a termination tolerance is reached.

**3.2.2. A new, temporally continuous, method**
The human eye continuously adapts to the brightness of its surrounding. However, we find that Ward's iteration can behave discontinuously; that is, a small change in the input can result in a sudden change in the mapping function. Disconti-

nuities also occur as the operator switches between the low and high dynamic range modes.

Instead of the iterative procedure, we choose instead to redistribute counts that exceed the ceiling. Rather than simply truncating the histogram, we keep track of the total of the truncated counts and redistribute that total to the other bins, taking care not to fill them past the ceiling. Any reasonable method for doing this may be used; we used a procedure that redistributes the trimmings proportional to the existing counts and distributes any excess uniformly.

For this to work we have to ensure that the sum of all the ceilings is not less than $T$. That is:

$$\sum_{i=1}^{N} \frac{T}{N} \cdot \frac{\log(L_{wmax}) - \log(L_{wmin})}{\log(L_{dmax}) - \log(L_{dmin})} \cdot \frac{\Delta L(L_{d_i})/L_{d_i}}{\Delta L(L_{w_i})/L_{w_i}} \geq T \quad (3)$$

If this constraint is not met, it signals that we are in a low dynamic range condition: the dynamic range of the scene is lower than that of the display, so that no response compression is needed. Ward handles this case by switching to a separate linear mapping mode, but in our method we instead simply reduce $\log(L_{dmax}) - \log(L_{dmin})$ so that (3) is satisfied exactly, then proceed as usual. We thereby handle low and high dynamic range in a unified and continuous way. Also unlike the previous method, we maintain the contrast preservation property in the low dynamic range case, which is important when the operator is simulating severely impaired visibility due to low vision or extreme maladaptation (such as walking from a sunlit street into a dark theater).

It is difficult to evaluate (3) directly because $L_{d_i}$ depends on the (unknown) mapping function used. We remove this dependency by making the simplifying but reasonable assumption that the display luminance range lies in the region governed by Weber's Law, which means $\Delta L(L_{d_i})/L_{d_i}$ is a known constant.

If the operator decides to use less display dynamic range than is available, we have to choose which portion of the available range to use. We compare the highest and lowest responses generated by the world luminances and the display luminances to determine which portion of the available range to be allocated. However, one can freely choose another way to do this without affecting the essential function of the tone mapping operator.

**3.3. A generalized threshold model (TVIA)**
In Ward's operator the luminance threshold $\Delta L(L_a)$ used in Subsection 3.2 comes from a TVI function (Figure 4). As we outlined at the beginning of Section 3, this implicitly assumes the eye is optimally adapted at all times. To eliminate this assumption we generalize the TVI function $\Delta L(L_a)$ to a function of two variables that depends separately on the stimulus luminance and the adaptation state. We denote this TVIA function as $\Delta L(L, \sigma(L_a))$: it gives the "just noticeable difference" when the eye is looking at luminance level $L$ while adapted to luminance level $L_a$. Here, $\sigma(L_a)$ represents
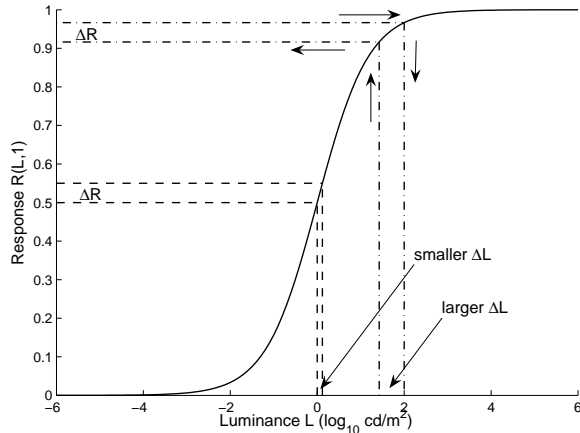
**Figure 8:** *Response function at $\sigma = 1$. The farther a luminance is from the adaptation state (in this case $\sigma(L_a) = 1$), the larger the luminance difference required to produce the same change in response (R).*

the adaptation state that is reached when we are fully adapted to $L_a$.

The TVI and probe-on-flash experiments provide information about the thresholds of the visual system, but data is not available for all luminance levels and adaptation states. In particular, while many TVI curves are available, probe-on-flash data is only available for a few isolated adaptation levels. To extend the TVI to the TVIA, we need to develop a new threshold model that is physiologically plausible, consistent with the known psychophysical data, and capable of supplying thresholds for any combination of luminance level and adaptation state. Our TVIA model is derived from a simple physiological model of visual response, and it is calibrated by requiring that it match the TVI function when $L_a = L$.

### 3.3.1. Visual response model

Based on their studies of photoreceptor response, Naka and Rushton [NR66] proposed the following function to describes the response $R$ of retinal photoreceptors:

$$R(L, \sigma(L_a)) = \frac{L^n}{L^n + \sigma(L_a)^n} \qquad (4)$$

Values for $n$ between 0.7 and 1 have been reported in the literature. [BW70] [NW74]

$R(L, \sigma(L_a))$ is the response generated by looking at luminance $L$ in adaptation state $\sigma(L_a)$. For constant $\sigma$, $R(L, \sigma)$ is a sigmoid curve centered at $L = \sigma$ (the curve for $\sigma(L_a) = 1$ is shown in Figure 8). Because the curve is steepest near the center, the response is most sensitive to changes in luminance when $L$ is near $\sigma$, and when $L$ is far from $\sigma$, the response is nearly constant due to the compressive nature of the curve. Because the photoreceptors are unable to signal luminance differences far removed from $\sigma$, the system adapts by changing $\sigma$, which shifts the response function left or right to lower or higher luminance levels.

Given this response model, we assume that under all conditions there is a single criterion response $\Delta R$, which is the smallest amount of additional response that is needed in order to produce a just noticeable difference (JND).

### 3.3.2. Response threshold and luminance threshold

Since the criterion response $\Delta R$ is the amount of additional response that is needed in order to perceive a difference in luminance, adding $\Delta L(L, \sigma(L_a))$ to $L$ should increase the response generated by exactly $\Delta R$.

$$R(L, \sigma(L_a)) + \Delta R = \frac{(L + \Delta L(L, \sigma(L_a)))^n}{(L + \Delta L(L, \sigma(L_a)))^n + \sigma(L_a)^n} \qquad (5)$$

Rearranging Equation (5), we get:

$$\Delta L(L, \sigma(L_a)) = \sigma(L_a) \left( \frac{R(L, \sigma(L_a)) + \Delta R}{1 - (R(L, \sigma(L_a)) + \Delta R)} \right)^{\frac{1}{n}} - L \qquad (6)$$

Since no amount of $\Delta L(L, \sigma(L_a))$ can increase response above 1, $R(L, \sigma(L_a)) + \Delta R > 1$ means the visual system reaches saturation and is unable to discriminate luminance values above $L$ at current adaptation state $\sigma(L_a)$. The steps of computing $\Delta L(L, \sigma(L_a))$ are as follows:

1. Compute $R(L, \sigma(L_a))$.
2. Add $\Delta R$ to this value.
3. If $R(L, \sigma(L_a)) + \Delta R > 1$, $\Delta L(L, \sigma(L_a)) = \infty$.
4. Otherwise, compute $\Delta L(L, \sigma(L_a))$ (Equation (6)).

To find $\Delta R$, we use the known luminance thresholds from the TVI function. We start by assuming that $\sigma(L) = L$ for all $L$. By doing so, the response generated by looking at luminance $L$ while adapted to $L$ is $R(L, L)$ and the response generated by adding a just noticeable luminance increment to $L$ is $R(L + \Delta L(L), L)$. The difference between these two values is the amount of additional response necessary to cause a perceived difference in luminance. We take the smallest such value as $\Delta R$, then adjust the function $\sigma$ to match the rest of the TVI curve.

### 3.3.3. Defining the adaptation state $\sigma$

Recall that adaptation corresponds to lowering or raising the value of $\sigma(L_a)$, which shifts the response function left or right. In our discussion so far we have assumed that we are given $\sigma(L_a)$, but in fact the function $\sigma$ is uniquely determined by the TVI function, because the TVIA generalizes the TVI and must match for $L = L_a$. We model changes in relative thresholds using incomplete adaptation. For luminances in the Weber range, where the TVI curve gives relative thresholds close to the minimum, $\sigma(L_a)$ is close to $L_a$; where the TVI curve gives higher thresholds, $\sigma(L_a)$ positions the sigmoid so that $L_a$ is not at the center, resulting in a higher threshold.

To find $\sigma$, we look at $\Delta L(L, \sigma(L))$, which is the contrast visibility threshold when the eye is looking at luminance level $L$ while optimally adapted to that same luminance level. Since it is just noticeable, adding $\Delta L(L, \sigma(L))$ to $L$ should in-
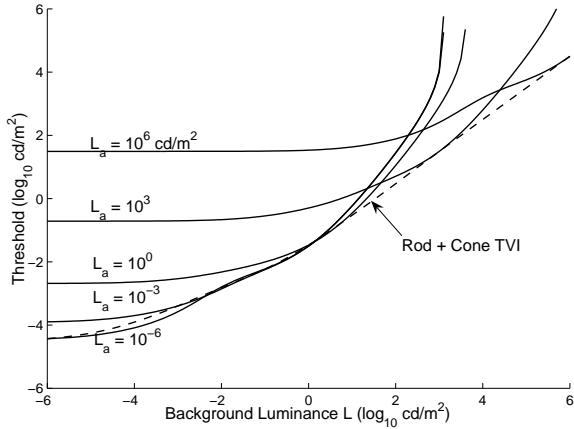
**Figure 9:** *Slices of the TVIA, our general threshold function, for several values of the adaptation luminance $L_a$. The curves increase for luminances away from $L_a$. The TVI function is approximately the lower envelope of these curves.*

crease the response generated by exactly $\Delta R$. That is:

$$R(L, \sigma(L)) + \Delta R = \frac{(L + \Delta L(L, \sigma(L)))^n}{(L + \Delta L(L, \sigma(L)))^n + \sigma(L)^n} \quad (7)$$

The equation above is just Equation (5) with $L_a = L$. From the TVI function, we know $\Delta L(L, \sigma(L))$ for every $L$. Now we can numerically solve for $\sigma(L)$ to obtain the adaptation state that the eye is in when fully adapted to luminance $L$. Graphically, this calculation is equivalent to shifting the response function left and right until the difference between $R(L, \sigma(L))$ and $R(L + \Delta L(L, \sigma(L)), \sigma(L))$ is exactly $\Delta R$.

We are now able to obtain the general TVIA function. Slices of the function for several fixed values of $L_a$ are shown in Figure 9. Luminances that are near the adaptation luminance fall near the middle of the response function, resulting in small thresholds; luminances far from the adaptation luminance fall at the ends, resulting in large relative thresholds. The TVI is approximately the lower envelope of these curves. The individual curves in the TVIA function also provide a good qualitative match to the changes in the contrast thresholds measured in the probe-on-flash experiments (quantitative variation is to be expected from differences in the experimental conditions). The TVIA function we have derived is a new general threshold model that allows us to predict contrast sensitivity for any combination of adaptation state and background luminance level.

### 3.4. Adaptation over time

In order to process time-varying scenes, we need a way to generate the adaptation state $\sigma$ required by the TVIA model for every frame, based on the changes in overall illumination in the input stream. The primary constraint is that after a long period of exposure to a constant adaptation luminance, $\sigma$ must converge to the $\sigma(L_a)$ derived in the previous section.

#### 3.4.1. Temporal dynamics of adaptation

Psychologists have described four mechanisms that control the adaptation state: pupil size (which we omit because of its relatively small effect), photoreceptor pigment bleaching, slow neural adaptation, and fast neural adaptation [HF86]. As Pattanaik et al. observed, it is important to track these different mechanisms separately because each adapts on a different time scale. For this reason we will break the steady-state function $\sigma(L_a)$ up into a product of terms, one for each adaptation mechanism, then define temporal behavior for each.

Therefore, we need to define $\sigma_b$ (adaptation due to pigment bleaching), $\sigma_c$ (slow neural adaptation), and $\sigma_n$ (fast neural adaptation) such that $\sigma(L_a) = \sigma_b(L_a) \cdot \sigma_c(L_a) \cdot \sigma_n(L_a)$.

Pigment bleaching is a well-studied process with known steady-state and temporal behavior. We denote the fraction of unbleached pigment left in the receptors after looking at luminance $L$ for a long time $p(L)$.

We assume that the amount of signal transmitted by receptors is proportional to $L \cdot p(L)$ [HHC37]. Scaling the luminance down is equivalent to shifting the sigmoid-shaped response curve rightwards by the same factor, so we have:

$$\sigma_b(L_a) = \frac{1}{p(L_a)} \quad (8)$$

where $p(L_a)$ is the fraction of unbleached pigment for a viewer fully adapted to $L_a$. In steady state the value of $p(L)$ follows the following formula.

$$p(L) = \frac{I_0}{I_0 + L} \quad (9)$$

where $I_0$ is around $10^4 cd/m^2$ [HF86].

Dividing $\sigma$ with $\sigma_b$ yields a range of about 4 log units. We attribute half of the remaining adaptation to each of the two neural adaptation mechanisms. The formulae for $\sigma_n$ are obtained from fitting the $\sigma$ curve with a sigmoid with magnitude of 2 log $cd/m^2$. With both $\sigma_b$ and $\sigma_n$ known, we can obtain the formulae for $\sigma_c$ by dividing $\sigma$ by $\sigma_n \cdot \sigma_b$ and fitting a curve to the values. The resulting formulae for cone system are as follows:

$$\log_{10}(\sigma_n(L_a)) = \frac{2.027L^{0.6406}}{L^{0.6406} + 5.859^{0.6406}} + 0.01711 \quad (10)$$

$$\log_{10}(\sigma_c(L_a)) = \frac{1.929L^{0.8471}}{L^{0.8471} + 1048^{0.8471}} + 0.01820 \quad (11)$$
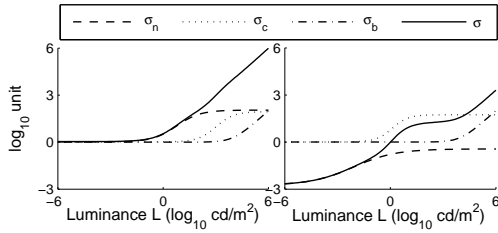
And for the rod system:

$$\log_{10}(\sigma_n(L_a)) = \frac{2.311L^{0.3604}}{L^{0.3604} + 0.008061^{0.3604}} - 2.749 \quad (12)$$

$$\log_{10}(\sigma_c(L_a)) = \frac{1.735L^{0.9524}}{L^{0.9524} + 1.277^{0.9524}} + 0.005684 \quad (13)$$

The steady state values for all the $\sigma$s are plotted in Figure 10. The flat region between $10^1 cd/m^2$ and $10^3 cd/m^2$ for the

**Figure 10:** *Steady-state adaptation state σ and its components across the luminance range for cone system (left) and rod system (right). The three components represent photopigment bleaching (σ_b) and fast (σ_n) and slow (σ_c) neural adaptation; σ is the product of the three.*

rods is where rod saturation occurs in the TVI function (Figure 4): the value of σ falls behind the increase in brightness, resulting in loss of function in the rods. (The curve does eventually unflatten because of pigment bleaching, but not enough to be useful.) In the cone system, saturation does not occur because pigment bleaching takes over before neural adaptation stops.

The time course of pigment bleaching and regeneration after a change of luminance at time $t = 0$ from $L_0$ to $L_a$ is known to follow the equation:

$$p = p(L_a) + (p_0 - p(L_a)) \cdot e^{\frac{-t}{t_0 \cdot p(L_a)}} \qquad (14)$$

Because in the dark $p(L_a) = 1$ and in a bright surrounding $p(L_a) < 1$, pigment bleaching happens faster than pigment regeneration. The time constant $t_0$ is 110 seconds for cones and 400 seconds for the rods [Alp97] [HA73].

We follow [PTYG00] in describing the time course of neural adaptation using simple exponential decay function:

$$L = L_a + (L_0 - L_a) \cdot e^{\frac{-t}{t_0}} \qquad (15)$$

Unlike pigment bleaching and regeneration, neural adaptation is symmetric. Following [PTYG00], we set $t_0$ to 0.08 seconds for the cones and 0.15 seconds for the rods.

We implement these equations for an image stream by maintaining an adaptation luminance for each neural mechanism, the fraction of unbleached cone pigment, and the fraction of unbleached rod pigment. At each frame we update the adaptation state by using Equations 14 and 15 with $t$ set to the frame duration. $L_a$ is set to the arithmetic mean of the foveal image.

Slow neural adaptation is slower than fast neural adaptation, but faster than pigment bleaching and regeneration. We determined time constants for slow neural adaptation by matching the model to published dark adaptation data [Hai41]. Figure 13 shows the time course of dark adaptation. The first drop in threshold is due to fast neural adaptation, while the subsequent adaptation is dictated by cone system before the more sensitive rod system takes over.

### 3.4.2. Adaptation during fixations
An additional modification is required to make this threshold model suitable for tone mapping images. If we assume a
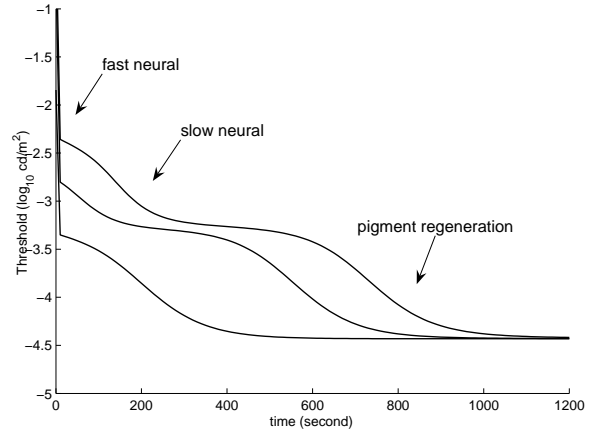


**Figure 11:** *Modeling dark adaptation. The top curve starts with the eye adapted to 15000 cd/m²; middle curve 1500 cd/m²; bottom curve 12.5 cd/m². Three parts of adaptation can be seen: fast neural adaptation, followed by the slow neural adaptation, and trailed by the pigment regeneration process.*

fixed adaptation state for the entire image, we will severely overestimate thresholds, because in reality a viewer looking at an image exhibits saccadic eye movements, fixating for approximately 200-300 milliseconds in one area before jumping to another [SB02]. During these fixations the visual system adapts significantly, so using a fixed adaptation state is unrealistic. To account for this partial adaptation we keep track of the average adaptation state over the image, then compute a temporary adaptation state for each luminance level $L$ as if the observer started in the average state and then adapted for a fixation time $t_f$ to luminance $L$.

Using partial adaptation always results in increased visibility compared to a single global adaptation, and our model always results in decreased visibility compared to Ward, in which the eye is assumed fully adapted for every luminance level. The fixation time $t_f$ acts as a parameter to continuously change the model from global adaptation, when $t_f = 0$, to full local adaptation (Ward's approach), when $t_f = \infty$. We use $t_f = 1/3$ second to simulate partial adaptation due to saccadic eye movements.

Figure 12 shows Ward's bathroom scene as seen by an observer who is adapted to luminances much higher and much lower than the prevailing luminance. With $t_f = \infty$ there is no concept of adaptation state and the images are identical; with $t_f = 0$ there is substantial loss of visibility in both cases; and with $t_f = 1/3$ sec the observer is able to adapt partially. Because light adaptation is much faster for large changes in luminance than dark adaptation, visibility is good in the case of too-low adaptation luminance but poor in the case of too-high adaptation luminance.

The partial adaptation approach completes our adaptation model. Together with the TVIA and the improved histogram
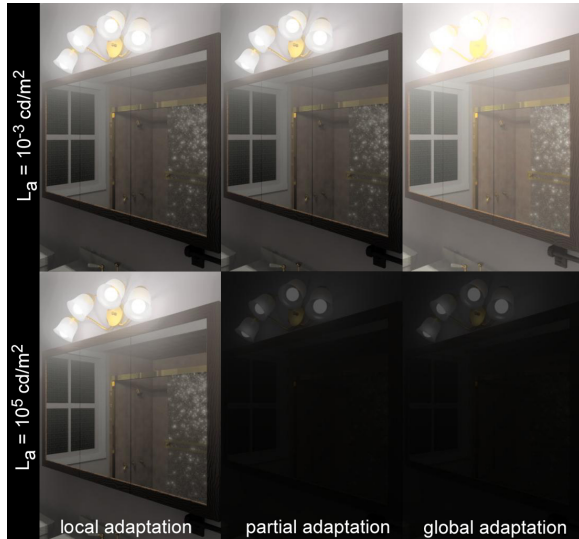
**Figure 12:** *Partial adaptation. The average luminance of the scene is $10^1$ cd/$m^2$; observer is adapted to $10^{-3}$ cd/$m^2$ (top row) and $10^5$ cd/$m^2$ (bottom row). The columns show the effect of local, partial, and global approaches to calculating adaptation state.*

adjustment method we have described all the components of our new operator.

## 4. Results

We have now achieved our goal of creating a new perceptually-based tone mapping operator for high dynamic range image streams. In this section we will first demonstrate the basic capabilities of the operator and then show how we have used it to create a low vision simulator that produces images that predict the effects of aging on visual performance.

We implemented the operator in MATLAB, and running on a standard desktop PC (1.7 GHz, 512 Mb, Windows XP) it processes the 1080 by 720 images in Figures 13, 15, and 16 in 5 to 10 seconds without acuity processing. Simulating acuity changes adds up to 30 seconds to the processing time for dim scenes but could easily be implemented more efficiently.

### 4.1. Dynamic range mapping; simulating visibility

One goal we set for our operator is to be able to tone map scenes of arbitrary dynamic range. Figure 13 shows a stair tower scene. A high dynamic range image of the scene was captured using a Canon D30 digital camera and the techniques described in [DM97]. With the door open the scene has a dynamic range of 723:1, a maximum luminance of 2892 $cd/m^2$ and an average (arithmetic mean) of 280 $cd/m^2$. With the door closed the scene's dynamic range decreases dramatically to 28:1 and the maximum and average luminances are 0.5 $cd/m^2$ and 0.2 $cd/m^2$ respectively.

With the door open the scene's dynamic range exceeds the dynamic range of a conventional display so a simple linear tone mapping operator will not be sufficient. The small graph on the left shows the original luminance histogram of the "door open" image (gray line), the ceiling curve for the histogram bins (dotted line), and the histogram after redistribution (black line). The small graph on the right shows the resulting non-linear visibility preserving mapping function. Note the effects of glare near the edges of the door frame and the relatively low visibility of the dimly illuminated interior regions.

When the door is closed, both the scene dynamic range and average luminance level change dramatically. Because the scene range is now smaller than the display range it is possible for the operator to use a linear mapping function, but because the luminance levels are low, the function should use less than the full display range to simulate the limits of contrast visibility under these conditions. The right side of Figure 13 shows the tone mapped image, the original and adjusted histograms, and the mapping function. Note the loss of color saturation produced by the operator's visual model. Acuity changes have also been computed, but the detail losses at this level of illumination are are too small to be visible in the images.

### 4.2. Handling image streams; time course effects

A second goal we set to achieve was to develop an operator that can efficiently process image streams and simulate the transient visual effects caused by changes in scene luminance levels. Figure 14 presents some stills from the accompanying video that show our operator applied to Pattanaik et al.'s tunnel sequence. As in the original, roadway luminance is 5000 $cd/m^2$ outside the tunnel and 5 $cd/m^2$ inside. Note the poor interior visibility at the tunnel entrance, the partial recovery of visibility in the interior, the glare and corresponding losses when approaching the exit, and the final recovery. Keen observers may notice differences the images in our sequence and Pattanaik et al.'s. These are due to 1) the different perceptual-matching criteria used by our threshold visibility model and their suprathreshold color appearance model; and 2) our ability to handle the animation's high dynamic range images.

### 4.3. Simulating low vision

Finally, recall that our overarching goal is to create a low vision simulator that can show people with normal vision what the world looks like to people with visual impairments. By modifying the psychophysical data used in our new operator's visual models (glare, adaptation, acuity, color, etc.) we can simulate how scenes appear to observers with different visual abilities. In this section, we demonstrate this capability by simulating the changes of vision with age.

Three important changes in vision that occur over the life span are increases in glare, decreases in contrast sensitivity, and slowing of the time course of dark adaptation. To account for the changes in light scattering in the aging eye we use the following model described by [Vos84], where age $A$
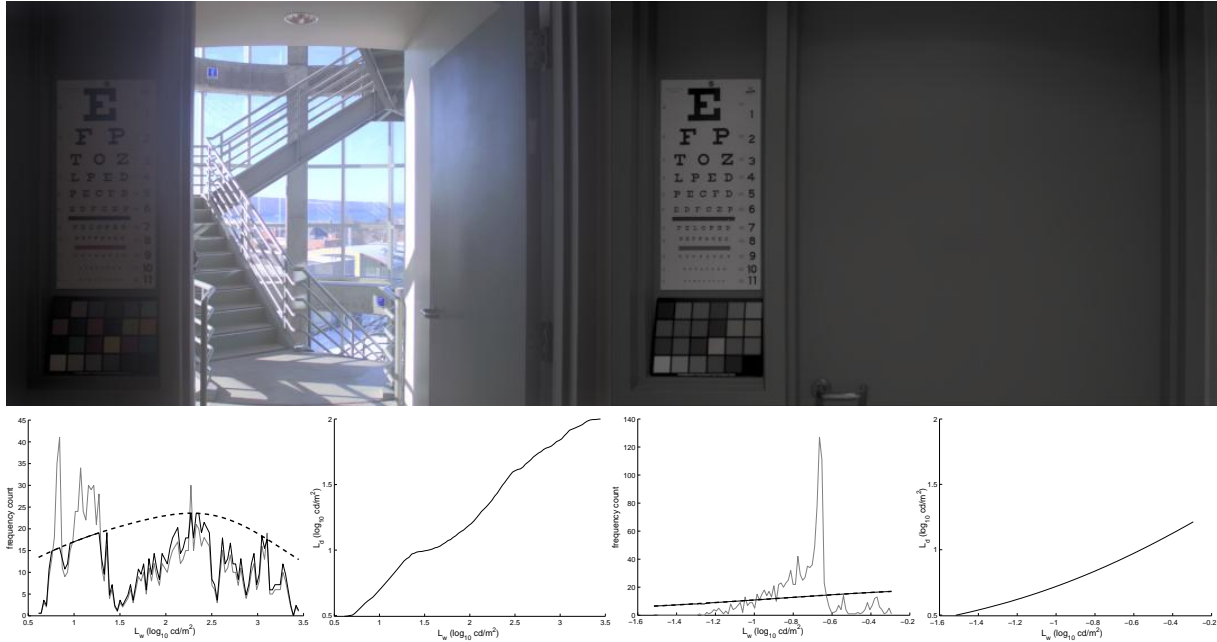
**Figure 13:** *Visibility preserving tone-mapping of a high dynamic range scene (left) and a low dynamic range scene (right).*



**Figure 14:** *Simulating the effects of the time course of adaptation. New operator applied to Pattanaik's tunnel image sequence.*

modifies the parameter $V$ that represents the percentage of light energy scattered by the glare filter in Ward's operator.

$$V = 0.08 \cdot \frac{10 + 5 \cdot 10^{-7} \cdot A^4}{10 + 5 \cdot 10^{-7} \cdot 20^4} \qquad (16)$$

Losses in contrast sensitivity with age are modeled by the following threshold elevation equation based on measurements by [SZT*97] and [JOCF98], where $A$ is the observer's age in decades. Similar effects of age were described earlier by [BB71].

$$\text{Threshold} = 10^{0.0085 \cdot max((A-20),0)} \cdot (\text{Threshold at age 20})$$
$$(17)$$

Finally, studies by [CB92] and [JOJ99] have shown that the slowing of the time course of dark adaptation with age is largely due to changes in the rates of photopigment re-

generation. We model this by increasing the corresponding rod and cone time constants in our adaptation model by 20.4 sec/decade and 12.6 sec/decade respectively in accordance with their measurements.

Figures 15 and 16 show sequences generated by our low vision simulator. The upper and lower rows simulate the differences in glare and dark adaptation for young and old observers (20 years and 70 years respectively). The "before" image shows the high dynamic range "door open" scene. Note the differences in glare and in contrast visibility at both the high and low ends of the luminance range. The 0, 5, and 10 minute images show differences in the rate and degree of dark adaptation when the door in the "before" image is closed. The 20 year old is well adapted after 5 minutes while visibility for the 70 year old is still poor after 10 minutes.

Figure 16 similarly illustrates differences in light adapta-
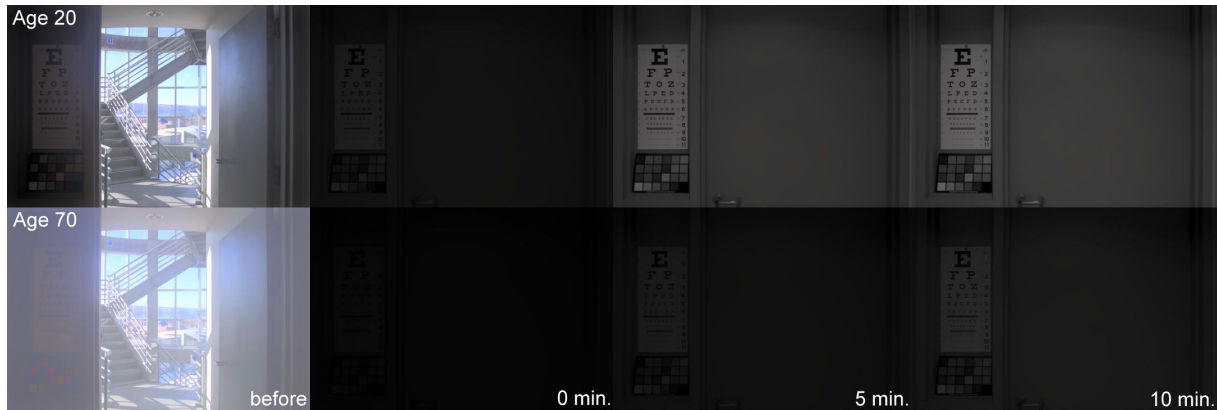
**Figure 15:** *Simulating the effects of glare and age on dark adaptation.*



**Figure 16:** *Simulating the effects of glare and age on light adaptation.*

tion in young and old observers. Note that when the door is opened, the 20 year old observer rapidly adapts to the higher luminance levels and dynamic range while the older observer shows significant losses in visibility at this transition and overall poorer adaptation to the new conditions.

## 5. Conclusion and future work

We have presented a new tone mapping operator for high dynamic range scenes, based on a new model for contrast visibility under varying adaptation, that fills an important gap in the field. The new operator is both perceptually based and dynamic: it quantitatively reproduces contrast visibility while at the same time accounting for changes in the observer's adaptation state over time. It includes three new components: the TVIA, our new contrast visibility model; an improved histogram adjustment procedure that is temporally continuous, and therefore is suitable for processing dynamic image sequences; and a generalized model for the time course of adaptation that accounts for the contributions of three adaptation mechanisms. We have shown how this

new operator can solve the problem that originally motivated the work: simulating visibility in time-varying, high dynamic range scenes for observers with low vision.

Time-varying adaptation has fundamental implications for the goals of perceptually based tone mapping, because observers are constantly and significantly adapting as they look around a high dynamic range scene. The observer cannot be optimally adapted to every part of the scene, which was the previously accepted assumption. On the other hand, substantial adaptation does happen from one part of the image to another. We have shown how to allow a realistic amount of adaptation to reflect the very good abilities of normal observers in moderately high dynamic ranges while still modeling loss of visibility in observers who have limited vision or are very maladapted.

Our work on tone mapping and vision simulation opens up a number of areas for future work. Our operator is global in nature, limiting its performance in scenes that have content over a continuous range of luminance levels. The same visual models could be adapted to local tone mapping opera-

tors, giving more flexibility. It would be useful to implement the operator efficiently in graphics hardware and, eventually, integrate it into a high dynamic range camera system to create a portable, real time low vision simulator.

Our model for vision under maladaptation is based solely on contrast thresholds. While it seeks to accurately represent the threshold visibility of scene features, it provides no guarantee of matching appearance at suprathreshold levels. The color and acuity aspects of the model, which we have not changed from previous work, could also be made more accurate, and more types of visual impairments (e. g. glaucoma, cataracts, or macular degeneration) could be added.

## Acknowledgements

## References

[Alp97]  ALPERN M.: Rhodopsin kinetics in the human eye. *Journal of Physiology 217* (1997), 447–471.

[BB71]  BLACKWELL O. M., BLACKWELL H. R.: Visual performance data for 156 normal observers of various ages. *Journal of Illuminating Engineering Society 1* (1971), 3–13.

[Bla46]  BLACKWELL H. R.: Contrast thresholds of the human eye. *Journal of the Optical Society of America 36* (1946), 624–643.

[BW70]  BOYNTON R. M., WHITTEN D. N.: Visual adaptation in monkey cones: Recordings of late receptor potentials. *Science 170* (1970), 1423–1426.

[CB92]  COILE D. C., BAKER H. D.: Foveal dark adaptation, photopigment regeneration, and aging. *Visual Neuroscience 8* (1992), 27–39.

[DCWP02]  DEVLIN K., CHALMERS A., WILKIE A., PURGATHOFER W.: STAR: Tone reproduction and physically based spectral rendering. In *State of the Art Reports, Eurographics 2002* (2002), pp. 101–123.

[DM97]  DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (1997), pp. 369–378.

[GWWH03]  GOODNIGHT N., WANG R., WOOLLEY C., HUMPHREYS G.: Interactive time-dependent tone mapping using programmable graphics hardware. In *Proceedings of the Eurographics Workshop on Rendering 2003* (2003), pp. 26–37.

[HA73]  HOLLINS M., ALPERN M.: Dark adaptation and pigment regeneration in human cones. *Journal of General Physiology 62* (1973), 430–447.

[Hai41]  HAIG C.: The course of rod dark adaptation as influenced by the intensity and duration of pre-adaptation to light. *Journal of General Physiology 24* (1941), 735–751.

[HF86]  HOOD D. C., FINKELSTEIN M. A.: Sensitivity to light. In *Handbook of Perception and Human Performance: Sensory Processes and Perception*, Boff K. R., Kaufman L.,, Thomas J. P., (Eds.). John Wiley & Sons, Inc., 1986, ch. 5.

[HHC37]  HECHT S., HAIG C., CHASE A. M.: The influence of light adaptation on the subsequent dark adaptation of the eye. *Journal of General Physiology 20* (1937), 831–850.

[Hun95]  HUNT R. W. G.: *The reproduction of color*, 5 ed. Fountain Press, 1995.

[JOCF98]  JACKSON G. R., OWSLEY C., CORDLE E. P., FINLEY C. D.: Aging and scotopic sensitivity. *Vision Research 38* (1998), 3655–3662.

[JOJ99]  JACKSON G. R., OWSLEY C., JR G. M.: Aging and dark adaptation. *Vision Research 39* (1999), 3975–3982.

[NR66]  NAKA K. I., RUSHTON W. A. H.: S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology 185* (1966), 587–599.

[NW74]  NORMAN R. A., WERBLIN F. S.: Control of retinal sensitivity. i. light and dark adaptation of vertebrate rods and cones. *Journal of General Physiology 63* (1974), 37–61.

[PTYG00]  PATTANAIK S. N., TUMBLIN J., YEE H., GREENBERG D. P.: Time-dependent visual adaptation for fast realistic image display. In *SIGGRAPH '00* (2000), pp. 47–54.

[PY02]  PATTANAIK S. N., YEE H.: Adaptive gain control for high dynamic range image display. In *Proceedings of Spring Conference on Computer Graphics 2002* (2002).

[SB02]  SEKULER R., BLAKE R.: *Perception*, 4 ed. New York, McGraw-Hill, 2002.

[SZT*97]  STURR J. F., ZHANG L., TAUB H. A., HANNON D. J., JACKOWSKI M. M.: Psychophysical evidence for losses in rod sensitivity in the aging visual system. *Vision Research 37*, 4 (1997), 475–481.

[Vos84]  VOS J. J.: Disability Glare - a state of the art report. *CIE Journal 3*, 2 (1984), 39–53.

[WECH*90]  WALRAVEN J., ENROTH-CUGELL C., HOOD D., MACLEOD D., SCHNAPF J.: The control of visual sensitivity. In *Visual perception: the neurophysiological foundations.*, Spillman L., Werner J., (Eds.). San Diego: Academic Press, 1990, ch. 5.

[WLRP97]  WARD-LARSON G. W., RUSHMEIER H., PIATKO C.: A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics 3*, 4 (/1997), 291–306.

[YP03]  YEE H., PATTANAIK S. N.: Segmentation and adaptive assimilation for detail-preserving display of high-dynamic range images. *The Visual Computer 19*, 7-8 (2003), 457–466.