Time-Dependent Visual Adaptation For Fast Realistic Image Display

Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, Donald P. Greenberg

Program of Computer Graphics, Cornell University

ABSTRACT

Human vision takes time to adapt to large changes in scene intensity, and these transient adjustments have a profound effect on visual appearance. This paper offers a new operator to include these appearance changes in animations or interactive real-time simulations, and to match a user's visual responses to those the user would experience in a real-world scene.

Large, abrupt changes in scene intensities can cause dramatic compression of visual responses, followed by a gradual recovery of normal vision. Asymmetric mechanisms govern these time-dependent adjustments, and offer adaptation to increased light that is much more rapid than adjustment to darkness. We derive a new tone reproduction operator that simulates these mechanisms. The operator accepts a stream of scene intensity frames and creates a stream of color display images.

All operator components are derived from published quantitative measurements from physiology, psychophysics, color science, and photography. Kept intentionally simple to allow fast computation, the operator is meant for use with real-time walk-through renderings, high dynamic range video cameras, and other interactive applications. We demonstrate its performance on both synthetically generated and acquired "real-world" scenes with large dynamic variations of illumination and contrast.

CR Categories: I.3.3 **[Computer Graphics]**: Picture/image generation – *Display algorithms;* I.4.3 **[Image Processing and Computer Vision]**: Enhancement – *Filtering*.

Keywords: Rendering, realistic image display, time course of adaptation, background intensity, adaptation model.

1. INTRODUCTION

The human visual system can accept a huge range of scene intensities (from about 10^{-6} to 10^{+8} cd/m^2, or 14 $\log_{10}$ units), because it continually adjusts to the available light in any viewed scene. These adjustments for viewed intensity, known as visual adaptation, occur almost entirely within the retina [4]. Surprisingly, the eye's iris diameter only mildly affects adaptation; its 2-8mm adjustment range varies retinal illumination by only about 1 $\log_{10}$ unit [28].

Adaptation and its changes over time have profound effects on the visual appearance of any viewed scene. Continual adjustment helps keep the visual system acutely sensitive to scene content over a wide range of illumination, but adaptation also tends to hide or obscure any very slow changes in scene intensity or spectral content. For example, on an overcast late afternoon, an automobile driver may not notice the loss of daylight; adaptation can hide the gradual lighting reduction until another car's headlights reveal the darkness. Adaptation can also exaggerate large, rapid changes in scene intensity. For example, on a sunny day (~10^{+4} cd/m^2), people entering a dim motion picture theater (~10^{-1} cd/m^2) may see only blackness and the movie screen. Their vision is restored after tens of minutes of adaptation, and they may even see popcorn spilled under the seats (~10^{-3} cd/m^2). On leaving, people see only blinding whiteness, but in a few seconds their adaptation restores the normal appearance of a sunny day.

This paper offers a practical, accurate, and fully automatic way to reproduce similar visual experiences caused by time-dependent adaptation, even when scene intensity changes greatly exceed display device abilities. We present a new time-dependent *tone reproduction operator* that can rapidly create readily displayable color image sequences from any desired input scene, either static or dynamic, real or synthetic. The operator is simple, uses global rather than local adaptation models, and may be robust enough for real-time use with interactive renderings, with output from high dynamic range video cameras, or for rapid evaluations of lighting designs.

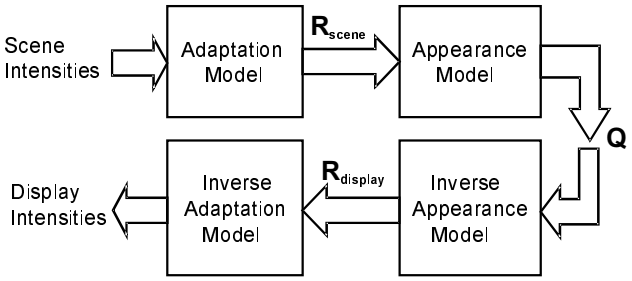

Figure 1: Tone Reproduction Operator Overview.

Our new operator follows the tone reproduction framework proposed by Tumblin and Rushmeier [19], and is built from a forward and inverse instance of a pair of perceptual models, as shown in Figure 1. The *adaptation model* transforms viewed

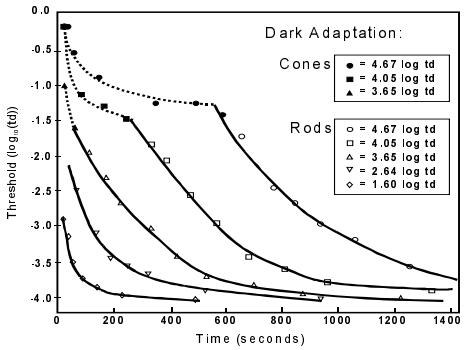

Figure 2: Increment Thresholds During Dark Adaptation. Ability to sense small illumination changes develops slowly in the dark. Test subjects were first adapted thoroughly to a uniform background intensity. Experimenters removed the background light and then periodically measured the test subject's thresholds in darkness. Five curves show results from five widely-spaced initial background intensities. Both rod and cone thresholds fall asymptotically, but cones (dashed lines, filled symbols) adapt to darkness more rapidly, and dominate threshold measurements until slower rod thresholds (solid lines, open symbols) can fall below them ([10], Table III).

scene intensities to retinal-response-like vectors **R**, and the *appearance model* converts **R** to appearance vectors **Q** that express correlates of "whiteness/blackness" and "colorfulness." The upper model pair computes viewed scene appearance, and the lower pair of inverse models computes display intensities that match the scene appearance. Our forward adaptation model is an abbreviated version of an authoritative, sophisticated model of static color vision by R.W.G. Hunt and colleagues [15], augmented with exponential filters for time-dependent adaptation mechanisms. After reviewing related work in Sections 2 and 3, Section 4 presents detailed derivations of the entire operator, and Section 5 demonstrates its performance on time-varying scenes.

2. PREVIOUS WORK

There is a wealth of published psychophysical work measuring the performance of the human visual system as a function of steady background intensity [6, 8, 14, and 28]. These books and papers provide data about how thresholds increase, visual acuity improves, and motion, color, and contrast sensitivity increases with additional available light. During the last decade, computer graphics researchers [5, 17, 19, 20, 25, 26 and others] have used these data to construct more perceptually accurate scene-to-display mappings. Published mapping methods compute displayed images with improved appearance for a wide range of scene intensities as they might appear under static (steady-state) viewing conditions. These models capture many of the stationary light-dependent aspects of viewed scene appearance.

However, we found surprisingly little published work on time-dependent models of visual adaptation that is suitable for computer graphics. Though several authors, such as Graham & Hood [9], Sperling & Sondhi [18], Walraven & Valeton [23] and Wilson [27] have published extensive models of adaptation processes, the work primarily addressed psychophysical threshold experiments rather than the appearance of arbitrary image sequences. Accordingly, these models do not address appearance effects or consider the problems of displaying computed results. In the computer graphics literature, only Ferwerda [5] offers any time-dependent method, but his simple and clever model is not in accordance with psychophysical data on response compression, and the method is restricted to step-like scene intensity changes.

The work presented here is novel in three ways. First, the model is general: it accepts time sequences of arbitrary scene intensities. Next, it captures the appearance of widely varying amounts of adaptation, and includes both bleaching and network effects. Finally, the model is firmly grounded in published research results from psychophysics, physiology, and color science.

3. BACKGROUND

Almost all known adaptation mechanisms are found within the retina, and each mechanism has its own time course. In addition to the mild effect of pupil diameter adjustments, the combined effects of receptor types, photopigment bleaching and retinal connection networks explain the huge span of human vision [4].

The human retina holds two types of photoreceptor cells. Cones sense color and respond well in dim to bright light ($\sim 10^{-1}$ to $\sim 10^{+8}$ cd/m^2), and rods respond best between darkness and moderate light ($\sim 10^{-6}$ to $\sim 10^{+1}$ cd/m^2), but are blinded by saturation above $\sim 10^{+2}$ cd/m^2 [14]. Within their response ranges, receptors react when one of its "visual pigment" molecules captures a photon. The captured photon triggers a complex cascade of reactions known as "bleaching" that desensitizes the molecule. Bright light rapidly reduces a receptor's usable photopigment concentration, but slow retinal mechanisms restore it [4]. Photopigment concentration sets an upper limit on receptor sensitivity, and simple rate equations can predict reasonably well how these concentrations change with time and light [14].

Unlike film or television camera sensors, individual receptor cells share interdependent signals. Two more neural cell layers in the retina process these signals (see [4] Chapter 4 for a masterful summary) and their interactions strongly affect adaptation and its time course. Extensive psychophysical experiments (see [12] and [9]) have revealed rapid multiplicative and subtractive adaptation mechanisms, and more may exist within the retina ([24], pg. 76).

3.1 Adaptation Measurements and Models

Adaptation processes greatly complicate visual response function measurements because varying the test stimulus may cause adaptation that changes the response function as well. At least two very different approaches to this problem are common in the vision research literature. Psychophysicists often measure a test subject's ability to see test stimuli made so small, fast, or weak that adaptation does not change significantly, and physiologists measure the underlying biological mechanisms responsible for adaptation and light sensitivity. Both approaches offer only partial explanations of how adaptation affects visual appearance.

Increment threshold tests may offer the simplest measurements of adaptation. Test subjects first stare at a wide blank screen for enough time to adjust to its uniform "adapting intensity" I_a. Against this background, psychophysicists then show a small test spot of intensity $I_a+\Delta I$ and quickly find the smallest detectable ΔI. For moderate spot sizes and all I_a greater than about 10^{-4} cd/m^2, larger adapting backgrounds I_a cause larger increment thresholds ΔI. Over much of this range ($\Delta I/I_a$) is nearly constant, a relation known for over 140 years as the Weber-Fechner fraction. This fraction suggests adaptation acts as a normalizer, scaling scene intensities to preserve our ability to sense contrasts within it.

A sudden change to background I_a temporarily disrupts the simple monotonic (ΔI vs. I_a) function. Figure 2 shows how $\log_{10}(\Delta I)$ changes over time when test subjects are suddenly plunged into complete darkness after thoroughly adapting to one of five I_a intensities. Rods dominate retinal response while adapting from

dim light (I_a=2.64 $\log_{10}$(td[1])) to darkness, and ΔI falls asymptotically to its dark-adapted value in just a few minutes. The time course of cone adaptation becomes important in the transition from bright light (4.67 $\log_{10}$(td)) to darkness. First ΔI drops quickly but pauses near the minimum cone threshold (about –1 log(td)), then as rod thresholds finally fall below cone thresholds, ΔI slowly approaches the rod dark-adapted value.

Direct cellular measurements on isolated and whole rat retinas by Dowling (1963), Weinstein, *et al.,* (1967) and others (see summary in [4], Chapter 7) offer further help. Their work showed dark adaptation in both rods and cones begins with a rapid decrease in threshold governed almost entirely by retinal network interconnections, but this fall is limited to a level directly predicted by photopigment concentrations. More recently, works by [1], [12] and [9] suggest these neural processes are complete in about 200mS for non-bleaching changes in adaptation.

Figure 3 shows increment threshold changes for a fully dark-adapted observer exposed to bright background light [2]. The entire light-adaptation process is much faster than dark adaptation, with a markedly different effect on thresholds. At the onset of the bright adapting light, ΔI jumps immediately to a very high value, then quickly settles back towards its static value. Light adaptation also includes both a fast neural component [1, 9] and a slower, pigment-limited process. Though initially puzzling, these threshold behaviors are reasonably well explained by examining retinal mechanisms.

Most retinal cells vary their response only within a range of intensities that is very narrow if compared against the entire range of vision. Adaptation processes dynamically adjust these narrow response functions to conform better to the available light. Direct cellular measurements of response functions for cone, rod, and bipolar cells [4] and firing rates for sustained ON-center retinal ganglia [24] closely follow:

$$R(I) = R_{\max} \frac{I^n}{I^n + \sigma^n} \qquad (1)$$

an S-shaped curve (see multiple examples in Figure 5) where I is light intensity, R is neural response ($0 < R < R_{\max}$), semi-saturation constant σ is the I value that causes the half-maximum response, and n is a sensitivity control similar to gamma for video, film, and CRTs. Introduced by Naka and Rushton in 1966 to describe fish S-potentials [14], this hyperbolic function appears repeatedly in both psychophysical experiments with flashed test stimuli [13, 1, 23, 27] and widely diverse, direct neural measurements [4, 7, 8, 22]. Psychophysical experiments modeling adaptation and saturation in rods [1] and cones [14] using Equation 1 show both $R_{\max}$ and σ depend on both I and time.

Equation 1 helps explain why threshold values differ so markedly during dark and light adaptation in Figure 2 and Figure 3. Suppose we choose values for R_{max} and σ to describe the visual response of a light-adapted observer, and make the simple assumption (as did [9],[14]) that thresholds measure some small fixed increment in response value R. Sudden darkness will not immediately change the viewer's response function, and though most scene intensities will fall well below σ in Equation 1, at first the threshold value ΔI is only weakly affected (for I near zero, $R \cong (I/\sigma)^n$; to change R by some small constant amount requires a ΔI value that is nearly unaffected by I). Over time, adaptation will

[1]Trolands (td) merge scene intensity with pupil area to estimate retinal illuminance: td = (Intensity in cd/m^2)·(pupil area in mm^2).

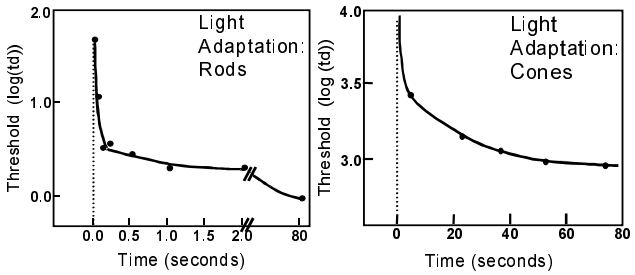

Figure 3: Increment Thresholds During Light Adaptation. Dark-adapted observers suddenly exposed to bright background light of 0.5 log Trolands for rods (left [1]) or to 3.7 log Trolands for cones (right [2]) initially experience very high thresholds, but these rapidly decay back to their static values.

gradually reduce the threshold by adjusting σ and R_{max} to their dark-adapted values as shown in Figure 2. However, once fully dark-adapted, suddenly large I values ($I >> \sigma$) cause large response values near R_{max}, and any noticeable increase in this response requires a huge additional ΔI, a property known as "response compression" [23]. Only when adaptation brings σ nearer to the new, much brighter I values will the response and threshold fall to their static values.

Beginning in 1980, R. W. G. Hunt and colleagues have assembled and continually refined an intricate mathematical model of human color vision for use in printing, photography, and video [15]. His model is a masterful synthesis of published data from psychophysics, physiology and color science, and is suitable for critical evaluations of many forms of color image reproduction. Briefly, Hunt's model uses Equation 1 to estimate rod and cone responses to a viewed image, along with careful modeling of color response, to include numerous important subtleties. We will use parts of Hunt's model in Section 4.1.

3.2 Display Interpretation

Light-adapted response alone does not entirely explain the visual appearance of displayed images: humans easily accept and understand reduced-contrast image renditions printed in newspapers or seen on poorly adjusted CRTs under high ambient light levels. We prefer higher contrasts, as confirmed by both formal studies and advertisements for film, CRTs and printers, but do not require them. For example, bright office light has limited CRT display contrasts to only 18:1 at the computer used to write this text. Nevertheless, the displayed text and figures appear as dark, rich black against a clean, paper-white background.

Extensive studies of visual appearance and preference in photographic prints and transparencies by Jones, Nelson, Condit, Bartelson, Breneman and others (summary in [16]) offer some useful insights. Each surmised that viewers estimate scene intensities by comparing display intensities against mental estimates of *reference white* and possibly *reference black.* These values describe the display intensities needed to represent scene objects with very high and very low diffuse reflectances viewed under prevalent scene lighting. Hunt uses both reference white and reference black in his model, and we closely follow his work.

4. DERIVING THE OPERATOR

This section provides a complete description of all parts of our time-dependent tone reproduction operator. The operator makes a displayable sequence of RGB images from an input sequence of scene values expressed in cd/m^2 or similar units, even if the input values are not displayable. The operator creates one display image for each frame of input data, and keeps only a handful of

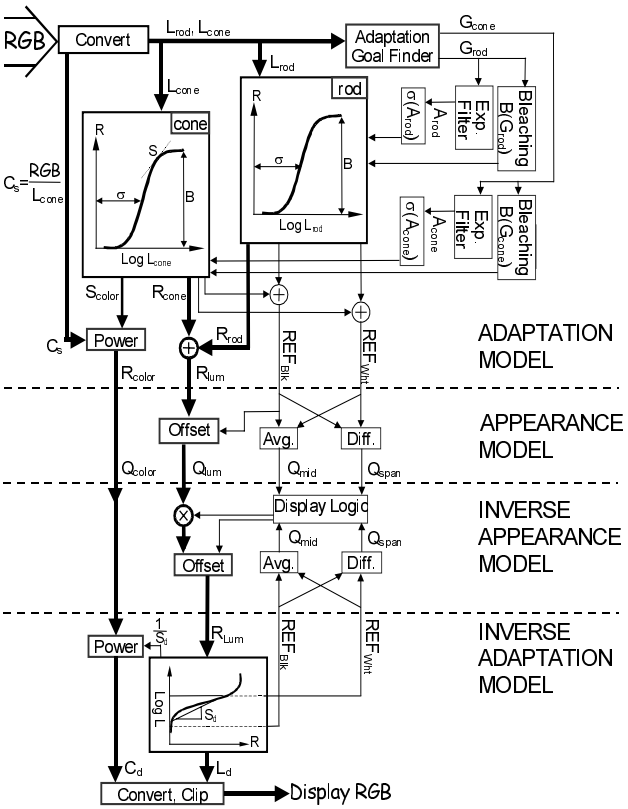

Figure 4: Detailed Tone Reproduction Block Diagram. The adaptation and appearance models of Equations 2-8 are diagrammed and assembled here according to the scheme shown in Figure 1. Thick lines carry pixel-by-pixel image data; thin lines convey scalars or vectors that apply to all pixels in an image but may vary over time. Dotted line divisions denote major sections shown in Figure 1.

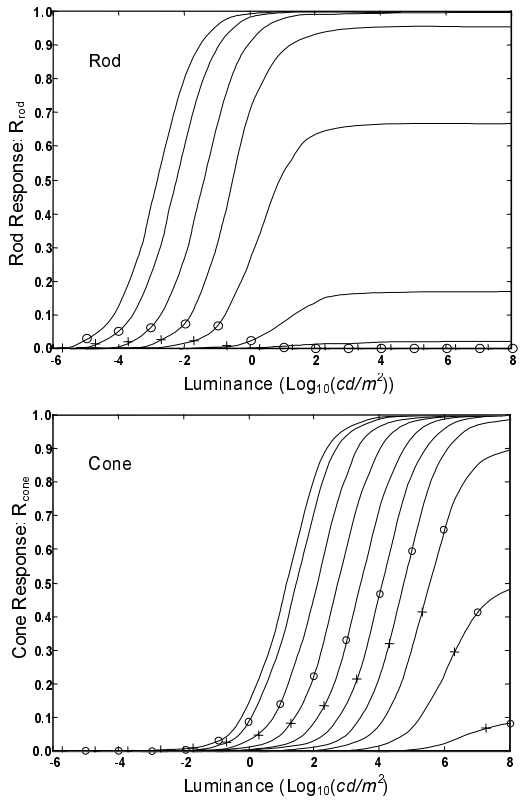

Figure 5: Model of Rod & Cone Response. These plots of R_{rod} and R_{cone} vs. luminances L_{rod} and L_{cone} were drawn with fixed adaptation luminance amounts $A_{cone} = A_{rod} = 2 \cdot 10^{-5}$, $2 \cdot 10^{-4}$, … , $2 \cdot 10^{+6}$, $2 \cdot 10^{+7} cd/m^2$. Crosses and circles mark response to adaptation luminance and "reference white" respectively. Note that rods adapted above about 10 cd/m^2 are saturated, with little or no response, as are cone above 10^{+5} cd/m^2.

time-dependent state variables from frame to frame; multiple frame buffers are not required. By following the outline given in Figure 1, we construct the complete operator as diagrammed in Figure 4. The next two subsections describe the operator's adaptation and appearance models.

4.1 Adaptation Model

The adaptation model acts as an idealized, film-like eye with uniform resolution and no localized differences in adaptation. At any instant, the same function governs response to light at all points in the scene. For each scene pixel, our model computes retina-like response signals R_{rod} and R_{cone} for rod and cone luminance and response vector R_{color} for color information.

Our adaptation model is a judicious simplification of Hunt's static model of color vision [15] that adds new, time-dependent adaptation components. These four components separately mimic the fast neural adaptation attributed to retinal interconnections (network) and the much slower process of photopigment bleaching and regeneration in both rods and cones.

4.1.1 Static Response

As shown in Figure 4, we begin by converting scene RGB values or radiances into luminance values for rods and cones (CIE standard Y', Y), labeled L_{rod} and L_{cone} respectively. We use only color-ratio components (red/L_{cone}, green/L_{cone}, blue/L_{cone}) for color.

As in Hunt's model (see [15], pg. 712, 721), we compute both rod and cone luminance responses using Equation 2, with the maximum response R_{max} given by a photopigment bleaching term B:

$$R_{rod} = B_{rod} \frac{L_{rod}^n}{L_{rod}^n + \sigma_{rod}^n}, \quad R_{cone} = B_{cone} \frac{L_{cone}^n}{L_{cone}^n + \sigma_{cone}^n} \qquad \textbf{(2)}$$

As before, R is the response to luminance L, and both B and σ are determined by adaptation to overall scene luminance. Our model is primarily concerned with variable responses to luminance, and discards the sophisticated color calculations performed by Hunt's model. Instead, we only approximate the amount of color compression caused by Equation 2 at L_{cone} and apply it to the color ratio vector **C** to make color response vector $\mathbf{R}_{color}$:

$$R_{color} = \left((\text{red}, \text{green}, \text{blue}) / L_{cone} \right)^{S_{color}},$$

where:

$$S_{color} = \frac{dR_{cone}}{d \log(L_{cone})} = \frac{n \cdot B_{cone} L_{cone}^n \sigma_{cone}^n}{\left(L_{cone}^n + \sigma_{cone}^n \right)^2} \qquad \textbf{(3)}$$

Hunt slightly modified the direct cellular measurements of Valeton and Van Norren [22] to restore pupil area effects removed from their data. He decreased *n* slightly to *n* = 0.73 and broadened the response range by about one $\log_{10}$ unit. Following Hunt's suggestion to define "reference white" as five times the adaptation luminance, the half-saturation parameters σ for rods and cones become:

$$\sigma_{rod} = \frac{2.5874 A_{\text{rod}}}{19000 j^2 A_{\text{rod}} + 0.2615 (1 - j^2)^4 A_{\text{rod}}^{1/6}} \qquad \textbf{(4)}$$

$$\sigma_{cone} = \frac{12.9223\, A_{\text{cone}}}{k^4 A_{\text{cone}} + 0.171(1-k^4)^2 A_{\text{cone}}^{1/3}} \qquad \textbf{(5)}$$

where: $j = \frac{1}{5\times 10^5 A_{\text{rod}} + 1}$ and $k = \frac{1}{5A_{\text{cone}} + 1}$.

Hunt's bleaching parameters for rods and cones are:

$$B_{cone} = \frac{2\times 10^6}{2\times 10^6 + A_{\text{cone}}} \quad \text{and} \quad B_{rod} = \frac{0.04}{0.04 + A_{\text{rod}}} \qquad \textbf{(6)}$$

Note B_{rod} shrinks rapidly towards zero for A_{rod}>>1 cd/m^2 to mimic rod saturation. The resulting static response model produces the curves for rods and cones shown in Figure 5. Our dynamic model will vary only the horizontal position (σ) and amplitude (B) of these curves.

Unlike Hunt's static values, we use separately computed, time-varying adaptation amounts (A_{rod}, A_{cone}) to compute σ, and bleaching and regeneration kinetics of pigments to compute B values for rods and cones. We rename Hunt's static (A_{rod}, A_{cone}) values as (G_{rod},G_{cone}), the 'goal' adaptation values eventually reached if the current scene is held fixed. In the next section, the goal amounts computed for each frame drive calculations of time-varying adaptation effects.

4.1.2 Dynamic Response

Hunt's model assumes scene viewers have achieved a static, steady state of adaptation where A_{rod} = G_{rod} and A_{cone} = G_{cone}. Typically, adaptation is measured by the amount of light required for a viewer to reach the same state while staring at a uniform blank background. Our new time-dependent or dynamic model will eventually reach this same state given enough time, but to model transient effects we use four separate time-dependent terms. 'A' terms describe fast, symmetric neural effects and are used to compute the σ values of Equations 4 and 5. 'B' terms model slower asymmetric effects from pigment bleaching, regeneration and saturation effects, and set response amplitudes as R_{max} did in Equation 1.

To compute 'A' and 'B' values, we first find the steady-state or goal adaptation values G_{rod} and G_{cone} for the current input frame of scene data. Several methods are plausible, and the best choice may depend on the application. In accordance with Hunt's 'reference white' values, we chose G values as one-fifth of the paper-white reflectance patch in the Macbeth chart for the image sequence on this paper's title page, but, in the movie sequence excerpted in Figure 7, we used the 1-degree foveal weighting method found in Ward-Larson *et al.*[26], and directed the foveal center to the roadway surface as a driver might. Users may also wish to aim this foveal weighting interactively as was done by Tumblin *et al.* [21]. Either method is a valid choice for the block labeled "Adaptation Goal Finder" in Figure 4.

Exponential decay functions are often used to model temporal processing of the visual system (see [18, 7, 1, 23, 12, 9]). Following this tradition, we chose to model the four adaptation signals A_{rod}, A_{cone}, B_{rod}, and B_{cone} with two forms of exponential smoothing filters applied to the adaptation goal signals G_{rod} or G_{cone} computed for every frame. The outputs of these filters are smoothed, delayed versions of their inputs, and Figure 6 illustrates our discrete implementation for both types.

We compute the fast, neurally-driven adaptation values A_{rod} and A_{cone} from goal values G_{rod} and G_{cone} respectively using simple fixed exponential filters where J and K functions are a fixed scale factor F: J(x)=K(x)=Fx. The response of these filters to a unit-height step-like input as in Figure 6 is given by $\left(1-e^{-t/t_0}\right)$, where *t* is time and t_0 is the "time constant." To find the constant F, just apply the discrete time-step size T to get $F = \left(1-e^{-T/t_0}\right)$. We chose t_0 values by curve-fitting to the dark-adaptation time course data from [3] after discounting regeneration effects on measured thresholds, yielding $t_{0,rod}$= 150mS and $t_{0,cone}$= 80mS. These data were measured on stimuli that caused significant bleaching, and our independent estimates agree reasonably well with non-bleaching measurements published by [1], [12] and [9]. We do not distinguish between multiplicative and subtractive adaptation in our model because the former is usually complete within one or two frame times.

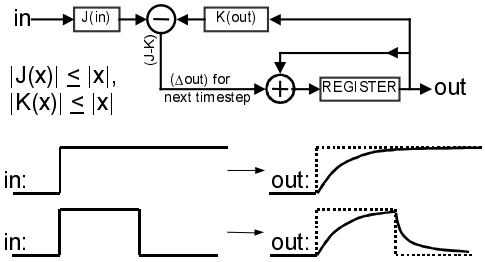

Figure 6: Discrete Exponential Smoothing Filter.

We compute time-dependent bleaching factors B_{rod}, B_{cone} by extending Hunt's static expressions in Equation 6 to include pigment kinetics (see [14], page 5-55, eqns. 10-17). The exponential filter of Figure 6 supplies the kinetics by its J and K functions. The J() functions describe the pigment depletion rate that depends on both the current amounts of light (G) and on the pigment concentration (B). The K() functions describe the competing process of pigment regeneration, which depends only on the pigment concentration (B). For rods, let *in* = G_{rod} (current value) and *out* = B_{rod}, (previous result), then find B_{rod} for the next time-step using:

$$J_{rod}(in) = T\cdot\frac{out}{16}\cdot in; \quad K_{rod}(out) = T\cdot\frac{1-out}{\tau_{rod}}. \qquad \textbf{(7a)}$$

Similarly for cones, let *in* = G_{cone} (current value) and *out* = B_{cone} (previous result) then find B_{cone} for next timestep using:

$$J_{cone}(in) = T\cdot\frac{out}{2.2\times 10^8}\cdot in; \quad K_{cone}(out) = T\cdot\frac{1-out}{\tau_{cone}}. \qquad \textbf{(7b)}$$

A published consensus on regeneration time constants is τ_{cone}=110 seconds for cones and τ_{rod}=400 seconds for rods [14]. Be sure to use cd/m^2 units in Equations 4-7 to agree with all constants.

4.2 Visual Appearance Model

Our visual appearance model is simple but extremely important to the behavior of the overall tone reproduction operator. As seen in Figure 5, even after adaptation, the response to a bright scene can be much stronger or weaker than any response achievable by a dim CRT display or a photographic print. Our appearance model assumes humans can assign equivalent appearance to dim displays and very bright or very dark scenes by a simple linear mapping of visual responses. The model determines "reference white" and "reference black" responses from among the current visual responses, and judges the appearance of any visual response against these reference standards. As shown Figure 4, our appearance model computes luminance appearance values Q_{Lum} by subtracting reference black response REF_{blk} from R_{lum}, where R_{lum} = R_{rod} + R_{cone}. We follow Hunt's suggestion and determine reference white as five times the current adaptation level and reference

black as 1/32 the intensity of reference white. For our time-dependent adaptation model, we find the response to reference white and black as:

$$REF_{wht} = R_{rod}\Big|_{L_{rod}=5\cdot A_{rod}} + R_{cone}\Big|_{L_{cone}=5\cdot A_{cone}} \quad \textbf{(8)}$$

$$REF_{blk} = R_{rod}\Big|_{L_{rod}=\frac{5}{32}\cdot A_{rod}} + R_{cone}\Big|_{L_{cone}=\frac{5}{32}\cdot A_{cone}}$$

(vertical line means "evaluated when"). We also compute the width and midrange of visual response as $Q_{span} = (REF_{wht}\text{-}REF_{blk})$ and $Q_{mid} = 0.5(REF_{wht} + REF_{blk})$. Color appearance values Q_{color} are set by R_{color} values of the scene.

4.3 The Tone Reproduction Operator

We can now assemble a new, time-dependent, tone reproduction operator by devising an inverse appearance and adaptation model to convert scene appearance values Q_{color}, Q_{Lum}, Q_{mid} and Q_{span} backwards into display intensity or RGB values as shown in Figure 4. We will explain these last models in reverse for clarity.

The inverse adaptation model finds display RGB values for a given set of visual response values R_{lum} and R_{color}. For simplicity, we assume the display device gamma is 1.0, forcing proportionality between RGB and display intensity values. Compared to the input range of the human visual system, the output range of most displays is quite small and usually cannot cause large changes in the visual adaptation values we compute for the display observer. (Some exceptions exist; brilliant video projectors viewed in an otherwise dark room can change viewer adaptation dramatically). For simplicity, we assume display observers have fixed, steady-state adaptation amounts. For the results shown in this paper, we assume a typical CRT display in ordinary office lighting, and set

$$A_{rod} = A_{cone} = L_{display} = 25\ cd/m^2,$$
$$REF_{wht} = 125\ cd/m^2,\ REF_{blk} = 4\ cd/m^2,$$
$$\sigma_{rod} = 722\ cd/m^2,\ \sigma_{cone} = 646\ cd/m^2,$$
$$B_{cone} = 1,\ B_{rod} = 0.0016.$$

We compute display luminance using these constants and the inverse of Equation 2. For simplicity, we do not compute Equation 3 for color, but instead compute a constant display S_d value from a forward-difference estimate of the slope of Equation 2 measured between the display REF_{wht} and display REF_{blk}. For our typical CRT with a maximum contrast of 32:1, we set $S_d = 0.1383$. We raise color appearance value Q_{color} to the power $1/S_d$ to convert it to display color ratio C_d.

The inverse appearance model is only slightly more complicated. The model attempts to do the least harm to visual appearance in translation to the display. We assume display observers will accept minimal amounts of response offset and compression, but will object to any temporal discontinuities or response exaggerations. We also assume the display minimum and maximum values will evoke REF_{blk} and REF_{wht} responses in the viewer, and our inverse observer model attempts to map scene appearance values Q to a display observer's response values with as little distortion as possible using the following rules:

1. IF the display can directly reproduce scene visual responses, do so. Exactly cancel the offset to R_{Lum} that was applied by the forward appearance model. ELSE
2. IF scene Q_{span}>display Q_{span}, compress and offset scene Q_{Lum} to match scene REF_{wht} and REF_{blk} to display REF_{wht} and REF_{blk}. ELSE
3. IF scene Q_{mid}>display Q_{mid}, offset scene Q_L downwards only enough to ensure scene $REF_{wht} \le$ display REF_{wht}. ELSE
4. Offset scene Q_{Lum} upwards only enough to ensure display $REF_{blk} \le$ scene REF_{blk}.

5. RESULTS

We have constructed two examples to demonstrate the performance of our time-dependent tone reproduction operator on both real-world and synthetically-generated scene intensity data. The strip of images across the top of the title page of this paper shows display images computed by our operator from two photographs of the same scene under widely different illumination conditions. The first frame in the sequence (t=0) shows predicted scene appearance for an observer statically adapted to the moonlight illumination from the side of the scene, where $A_{rod} = A_{cone} = 0.01$ cd/m^2. Immediately after this frame, the scene was suddenly lit by brilliant overhead illumination equivalent to mid-day sun: adaptation goal values are $G_{rod} = G_{cone} = 1000 cd/m^2$. In the next frame (T=30mS) the lighting has changed, but the scene viewer's adaptation state has not; a combination of clipping and response compression produces a displayed image that is almost entirely white. In subsequent frames, rapid retinal network adaptation increases rod and cone σ values to reduce response compression and restore the colorful appearance of daylight illumination.

In the second example, we simulate driving through a long highway tunnel on a sunny day, and Figure 7 shows frames from a videotape that accompanies this submission. As the daylight-adapted driver enters the tunnel, scene lighting falls quickly from about 5,000 cd/m^2 to 5 cd/m^2, and the driver is temporarily blinded while driving at highway speeds due to response compression. The driver's vision is again disrupted, though only very briefly, on leaving the tunnel.

Graphs below each video frame show the time-varying scene-to-display mapping applied by the tone reproduction operator. The effects of the exponential filters used to drive σ_{rod} and σ_{cone} are evident in slower adjustments on entering than on leaving the tunnel, and are demonstrated by the shift in the scene-to-display graph in response to large changes in scene lighting.

6. CONCLUSION AND FUTURE WORK

We have presented a simple time-dependent tone reproduction operator to reproduce the appearance of scenes that evoke changes to visual adaptation. Though the operator uses a broad range of published data, its global model of adaptation does not require extensive processing and is suitable for use in real-time applications. The operator is entirely automatic. We have demonstrated its effectiveness on both real-world and synthetic sources, and in both still and moving image sequences.

Though the results are simple to compute and have a pleasing and plausible appearance, tremendous opportunities for further improvements remain. An obvious refinement would include more of R. W. G. Hunt's model of static color vision and provide dynamic color adaptation. The operator could be improved by adding other secondary effects of adaptation, such as after-images, noise processes, and even loss of acuity under low-light conditions as already addressed by Ferwerda [5] and Ward [26]. More substantially, local adaptation effects are a vitally important part of visual appearance, and multiple instances of the tone reproduction function developed here might be applied to localized components of the scene if the proper scene decomposition could be

found. Finally, a time-varying inverse adaptation model might further increase the accuracy of the displayed images.

ACKNOWLEDGEMENTS

We thank SuAnne Fu for designing the tunnel model used in our illustration, and Jonathan Corson-Rikert and Peggy Anderson for carefully proof-reading the paper.

This work was supported by the NSF Science and Technology Center for Computer Graphics and Scientific Visualization (ASC-8920219) and by the MRA parallel global illumination project ASC-9523483, and performed using equipment generously donated by Hewlett-Packard and Intel Corporation.

REFERENCES

[1] Adelson, E. H. (1982). Saturation and adaptation in the rod system. Vision Res., 22, 1299-1312.

[2] Baker, H. D. (1949). The course of foveal light adaptation measured by the threshold intensity increment. Journal of the Optical Society of America, 39, 172-179.

[3] Crawford, B.H. (1937). Change of visual sensitivity with time. Proc. of the Royal Soc., B123, 69-89.

[4] Dowling, J. E. (1987). The Retina: An approachable part of the brain. Cambridge: Belknap.

[5] Ferwerda, J. A., Pattanaik, S.N., Shirley, P. and Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis, SIGGRAPH 96, 249-258.

[6] Finkelstein, M. A., Harrison, M., and Hood, D.C. (1990). Sites of sensitivity control within a long wavelength cone pathway. Vision Research, 30, 1145-1158.

[7] Geisler, W. S. (1981). Effects of bleaching and backgrounds on the flash response of the cone system. Journal of Physiology, 312, 413-434.

[8] Graham, N. (1989). Visual pattern analyzers, New York: Oxford University Press.

[9] Graham, N. and Hood, D. C. (1992). Modeling the dynamics of light adaptation: the merging of two traditions. Vision Research, 32, 1373-1393.

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

[11] Havard, J. (1991). New techniques and technology for tunnel lighting. Public works, November 1991, 122, 66-68.

[12] Hayhoe, M. M., Benimoff, N. I. and Hood, D. C. (1987). The time-course of multiplicative and subtractive adaptation process. Vision Research, 27, 1981-1996.

[13] Hood, D. C. and Finkelstein, M. A. (1979). Comparison of changes in sensitivity and sensation: implications for the response-intensity function of the human photopic system. Journal of Experimental Psychology: Human Perceptual Performance. 5, 391-405.

[14] Hood, D. C. and Finkelstein, M. A. (1986). Sensitivity to light. In Boff, K. R., Kaufman, L. R. and Thomas, J. P. (ed.), Handbook of Perception & Human Performance, Chapter 5, New York: Wiley.

[15] Hunt, R. W. G. (1995). The Reproduction of Colour, Chapter, Fountain Press, England.

[16] Nelson, C. N.(1966). The theory of tone reproduction, in James, T. H. (ed.), The Theory of the Photographic Process, 3rd ed. Chap. 22, 464-498, New York.

[17] Pattanaik, S. N., Ferwerda, J. A., Fairchild, M. and Greenberg D. P. (1998). A multiscale model of adaptation & spatial vision for realistic image display. Proceedings of SIGGRAPH 98, 287-298.

[18] Sperling, G. and Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. Journal of the Optical Society of America, 58, 1133-1145.

[19] Tumblin, J. and Rushmeier, H. (1993). Tone reproduction for realistic images, IEEE Computer Graphics and Applications, 13(6), 42-48.

[20] Tumblin, J. and Turk, G. (1999). LCIS: A boundary hierarchy for detail-preserving contrast reduction. Proceedings of SIGGRAPH 99, 83-90, Los Angeles.

[21] Tumblin, J. Hodgins, J., and Guenter, B. (1999) Two methods for display of high contrast images. ACM Transactions on Graphics, 18(1), 56-94.

[22] Valeton, J. M. and van Norren, D. (1983). Light adaptation of primate cones: An analysis based on extracellular data. Vision Research, 23, 1539-1547.

[23] Walraven, J. and Valeton, J. M. (1984). Visual adaptation and response saturation. In van Doorn, A. J., van de Grind W. A. and Koenderink J. J. (Ed.), Limits of Perception, The Netherlands:VNU Science Press.

[24] Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A., and Schnapf, J.L. (1990) The Control of Visual Sensitivity. In Spillmann, L., and Werner, J. A. (ed.), Visual Perception: The Neurophysiological Foundations, Chapter 5, San Diego CA: Academic Press.

[25] Ward, G. (1994). A contrast-based scale-factor for luminance display. In Heckbert, P. S. (Ed.), Graphics Gems IV, Boston: Academic Press Professional.

[26] Ward-Larson, G., Rushmeier, H. and Piatko, C. (1997) A visibility matching tone reproduction operator for high dynamic range scenes. IEEE Transactions on Visualization and Computer Graphics, 3(4), 291-306.

[27] Wilson, R.H. and Kim, J. (1998). Dynamics of a divisive gain control in human vision, Vision Research, 38, 2735-2741.

[28] Wyszecki, G., and Stiles, W. S. (1982). Color Science. New York: Wiley.

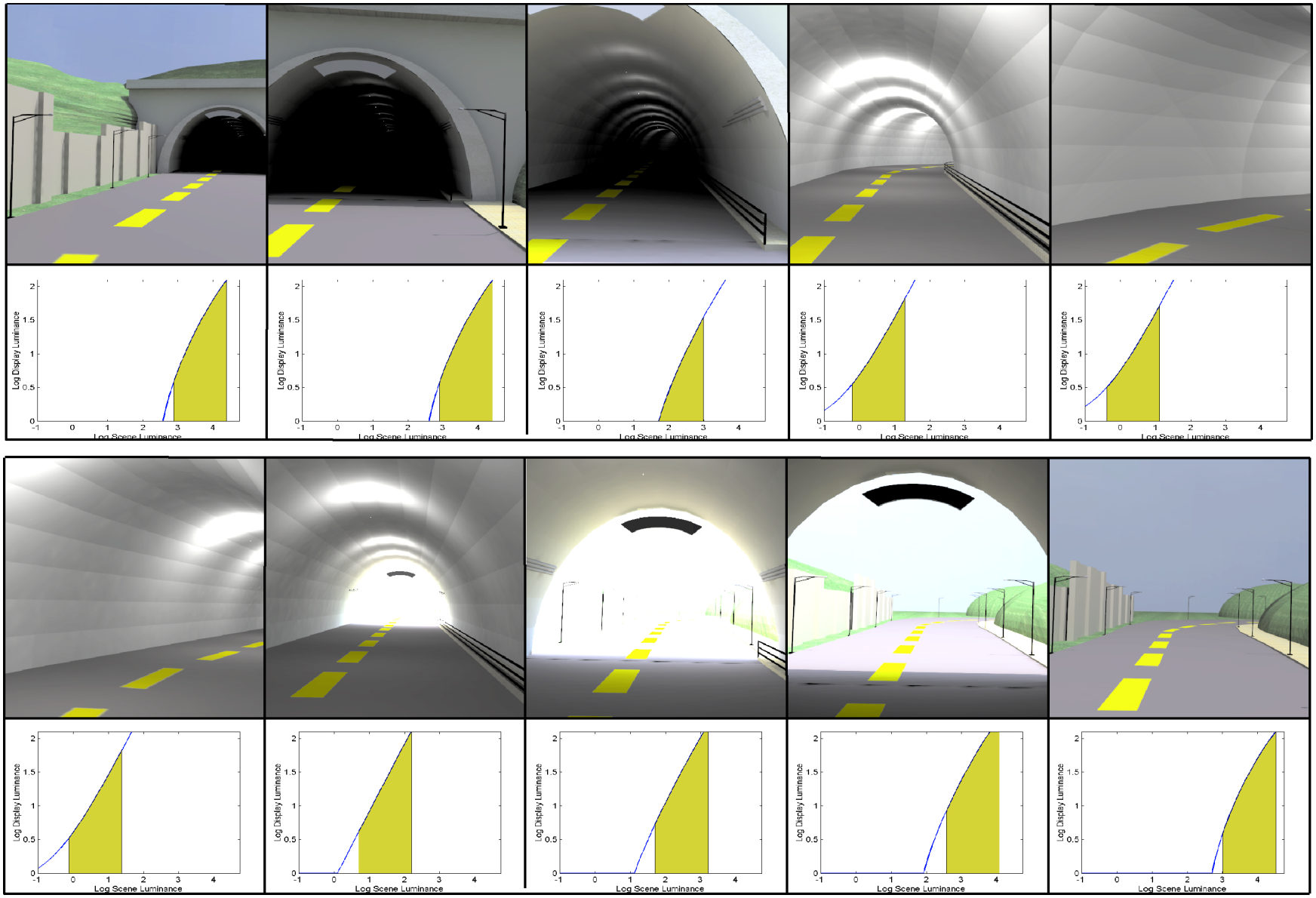

Figure 7: Tunnel Lighting Predictions Selected frames from a video sequence (available on Proceedings videotape) that combines our operator with global illumination solutions are used here to predict the appearance and safety of highway tunnel lighting. Roadway intensity varies between 5 and 5,000 cd/m^2. Tunnel designers ordinarily provide strong lighting just inside tunnel entrances to allow daytime drivers sufficient time at highway speeds to adapt to dim interior lighting [11], but excluding the lights here causes a dangerous momentary loss of vision while driving. Graphs below each frame show the time-varying, scene-to-display mapping curves computed by our tone reproduction operator; note the time-varying position and shape of the curves, and the pronounced but temporary response compression at the ends of the tunnel.