A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display

Sumanta N. Pattanaik James A. Ferwerda Mark D. Fairchild* Donald P. Greenberg
Program of Computer Graphics†, Cornell University

Abstract

In this paper we develop a computational model of adaptation and spatial vision for realistic tone reproduction. The model is based on a multiscale representation of pattern, luminance, and color processing in the human visual system. We incorporate the model into a tone reproduction operator that maps the vast ranges of radiances found in real and synthetic scenes into the small fixed ranges available on conventional display devices such as CRT's and printers. The model allows the operator to address the two major problems in realistic tone reproduction: wide absolute range and high dynamic range scenes can be displayed; and the displayed images match our perceptions of the scenes at both threshold and suprathreshold levels to the degree possible given a particular display device. Although in this paper we apply our visual model to the tone reproduction problem, the model is general and can be usefully applied to image quality metrics, image compression methods, and perceptually-based image synthesis algorithms.

CR Categories: I.3.0 [Computer Graphics]: General;

Keywords: realistic imaging, visual perception, tone reproduction, adaptation, spatial vision

1 INTRODUCTION

The range of light we encounter in natural scenes is vast. The *absolute level* of illumination provided by direct sunlight can be 100 million times more intense than starlight. The *dynamic range* of light energy can also be large, on the order of 10,000 to 1 from highlights to shadows, or higher if light sources are visible.

Although physically-based rendering methods and new techniques that utilize the output of digital cameras [Debevec97] now allow us to produce *radiance maps* that accurately represent the wide variations of light energy in scenes, neither of these methods specify how to realistically display these images on conventional electronic and print-based media which have only moderate output levels and typical dynamic ranges of less than 100 to 1.

Recently graphics researchers have started to address this issue by developing *tone reproduction operators* that map scene radiances to display outputs with the goal of producing a visual match between the scene and the display. There are two major problems to be solved in realistic tone reproduction:

- to find an operator that maps the vast ranges of radiances found in scenes into the range that can be produced by a given display device.
- to be certain that this operator produces images that match our perceptions of the scenes.

The critical element that links these two problems is the *visual model* used in the tone reproduction operator. Visual models are used to relate the perceptual responses of a scene observer to the responses of the display observer in order to specify a mapping that produces a visual match between the scene and the display. A central issue is that different tone reproduction operators have made use of different visual models to determine what constitutes a match.

Tumblin and Rushmeier's [1993] operator is based on Stevens' [1961] model of *brightness and contrast perception* illustrated in Figure 1b. The operator attempts to produce images that capture the changes in *suprathreshold brightness and apparent contrast* that occur with changes in the level of illumination. Ward [1994] introduced an operator based on a model of *contrast sensitivity* derived from threshold vs. intensity (TVI) functions similar to those shown in Figure 1a. Its goal is to match the *threshold visibility* of features in the image to features in the scene. Ferwerda [1996] developed an operator based on a model of adaptation that like Ward's matches threshold visibility, but also accounts for the changes in *visual acuity* and *color discriminability* that occur with the changes in the level of illumination.

Both threshold and suprathreshold models of vision capture important aspects of our visual experience, and a realistic tone reproduction operator should produce a mapping that matches both aspects. Unfortunately, threshold models don't scale well to predict suprathreshold appearance, and suprathreshold models don't accurately predict visual thresholds.

Recently much effort has been devoted to developing tone reproduction operators for *high dynamic range* scenes. Chiu [1993], Schlick [1995], and Jobson [1996] introduced spatially-varying operators that compress high dynamic range scenes into the limited range available on display devices, but the ad-hoc visual models they incorporate limits what can be said about the visual fidelity of the mappings. Tumblin [1997] has recently introduced an operator for high dynamic range scenes based on a model of *perceptual constancy*. Although this operator produces attractive images, the model it uses is not quantitative, and therefore the operator can't predict whether an image will be a visual match to a scene. Finally, Ward-Larson [1997] has introduced an operator that extends the work of Ward [1994] and Ferwerda [1996] with a model of *local adaptation*, to produce a threshold-based operator that can handle

*On sabbatical leave from: Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY 14623-5604, USA. Web address: http://www.cis.rit.edu/people/faculty/fairchild

†580 Rhodes Hall, Cornell University, Ithaca, NY 14853, USA.
E-mail addresses: {sumant,jaf,mdf,dpg}@graphics.cornell.edu
Web address: http://www.graphics.cornell.edu.

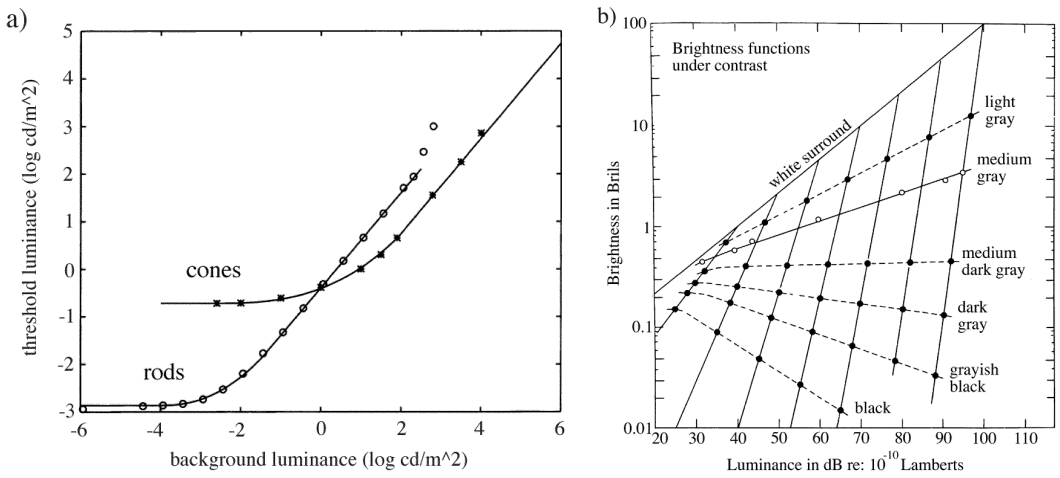

Figure 1: Threshold and suprathreshold models of vision: a) Threshold vs. intensity (TVI) functions for the rod and cone systems. The curves plot the smallest threshold increment ΔL necessary to see a spot against a uniform background with luminance L. b) Stevens' model of suprathreshold brightness and apparent contrast. The curves plot the changes in brightness and apparent contrast of gray targets and a white surround as the level of illumination rises (1 Bril = apparent brightness of a target with a luminance of 1μLambert). Adapted from [Ferwerda96, Stevens61].

high dynamic range scenes, and also match the changes in threshold visibility, visual acuity and color discriminability that occur with changes in the level of illumination.

Although the innovations introduced in each of these operators represent significant advances toward addressing the two fundamental problems of realistic tone reproduction, overall the problems have been attacked piecemeal. Thus some operators can only handle achromatic scenes, or scenes illuminated at daylight levels. Others can handle wide absolute ranges of illumination in chromatic scenes, but can't handle high dynamic ranges. Still others can handle full ranges of scene radiances, but can't guarantee that the images will match the scenes in any meaningful way. Finally, those that do produce visual matches differ on whether they match threshold measures like visibility and visual acuity, or suprathreshold measures like brightness and apparent contrast. Since the images these operators produce depend upon the visual models they incorporate, a comprehensive solution to the problems of realistic tone reproduction requires a more complete model of visual perception.

In this paper we develop a computational model of adaptation and spatial vision for realistic tone reproduction. The model incorporates a multiscale representation of luminance, pattern, and color processing in human vision that accounts for the changes in threshold visibility, visual acuity, color discriminability, and suprathreshold brightness, colorfulness, and apparent contrast that occur with changes in scene illumination. We incorporate the model into a tone reproduction operator that maps the vast ranges of radiances found in real and synthetic scenes into the small fixed ranges available on conventional display devices such as CRT's and printers. The model allows the operator to address the two major problems in realistic tone reproduction: images of wide absolute range and high dynamic range scenes can be displayed; and the displayed images match our perceptions of the scenes at both threshold and suprathreshold levels to the limits possible on a given display.

2 BACKGROUND

2.1 Adaptation and Visual Thresholds

The range of light we encounter in natural scenes is vast, but the responsive range of the neurons that make up the visual system is small. Our visual system copes with this vast range of illumination through *adaptation.* Although adaptation allows the visual system to function over a wide range of illumination levels, this does not mean that we see equally well at all levels. At low, *scotopic* levels our eyes are very sensitive and can detect small luminance differences, however visual acuity and the ability to discriminate colors are both poor. In contrast, at high, *photopic* levels, we have sharp color vision, but absolute sensitivity is low and luminance differences have to be large to be detectable. To produce realistic images that capture the visual appearance of scenes we need to understand these adaptation-related changes in vision.

The effects of adaptation on visual sensitivity have been measured in threshold experiments. Figure 1a shows the results of a threshold experiment that measured the changes in visibility that occur with changes in the level of illumination. The curves plot the smallest luminance increment ΔL that can be detected at a particular background luminance L and are known as threshold-vs.-intensity (TVI) functions. The two curves show the TVI functions for the rod and cone systems.

Over a wide range of background luminances, the size of the threshold increment increases in proportion to the background luminance making the functions linear on a log-log scale. This linear relationship $\Delta L = kL$ is known as *Weber's law* and indicates that the visual system has constant contrast sensitivity since the *Weber contrast* $\Delta L/L$ is constant over this range.

Constant contrast sensitivity is a desirable attribute for the visual system to have, since contrast in the retinal image is a function of surface reflectances and is invariant with respect to changes in the level of illumination. This "discounting of the illuminant" through adaptation is a major factor in perceptual constancy which underlies our ability to recognize objects under different illumination conditions [Shapley84]. Weber-like adaptation processes within the different cone systems (known as von Kries adaptation [Wyszecki82]) can also help explain chromatic adaptation and the perceived constancy of surface colors as the chromaticity of the illuminant changes.

2.2 Threshold Models of Spatial Vision

Although the TVI functions shown in Figure 1a give us useful information about the changes in visual sensitivity that occur with

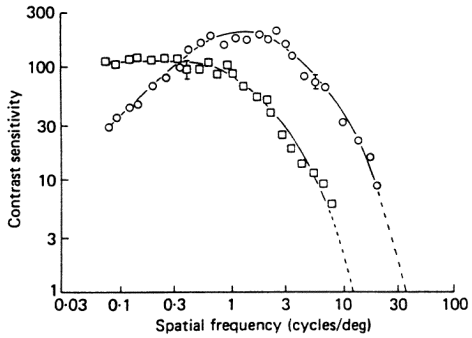

Figure 2: Threshold contrast sensitivity as a function of spatial frequency for a monochromatic luminance grating ($\circ; green; 526nm$) and a isoluminant chromatic grating ($\square; red/green; 602, 526nm$). Adapted from [Mullen85].

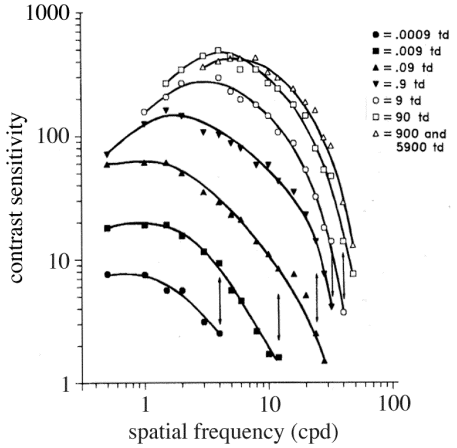

Figure 3: Contrast sensitivity functions for sinusoidal gratings illuminated at different mean luminance levels. Levels are specified in Troland (Td) units of retinal illuminance (Trolands = luminance in cd/m^2 x pupil area). Adapted from [vanNes67].

changes in the level of illumination, it's difficult to understand how to generalize from the results of these studies on the detectability of spots on uniform backgrounds to predicting the visibility of real objects (*e.g.* tanks, tumors) in complex scenes. In the 1960's vision researchers began to measure visual sensitivity for sinusoidal grating patterns to understand the properties of spatial vision.

The contrast sensitivity functions (CSF's) shown in Figure 2 plot visual sensitivity for detecting sinusoidal gratings as a function of their spatial frequency. Here sensitivity is defined as (1/threshold contrast) using the *Michaelson* definition of contrast: $(L_{max} - L_{min})/(L_{max} + L_{min})$ where L_{max} and L_{min} are the luminances at the peaks and troughs of the gratings [Laming91].

There is substantial evidence that the responses of the rod and cone photoreceptors are organized at an early stage in the visual system into responses in an achromatic channel sensitive to luminance variations and two chromatic channels, one sensitive to variations along a red/green axis and the other sensitive along a yellow/blue axis [Hurvich81]. The two curves in Figure 2 show the CSF's of the achromatic and red/green chromatic channels.

There are several things to notice about the CSF's. The first is that the spatial frequency response of the achromatic channel ($\circ$) has the characteristics of a bandpass filter. Contrast sensitivity is highest for gratings with frequencies around 2 to 4 cycles/degree of visual angle (cpd) and sensitivity drops for both higher and lower spatial frequencies. On the other hand, the spatial frequency response of the red/green chromatic channel ($\square$) has the characteristic of a lowpass filter. Sensitivity is good for low spatial frequencies, but declines at higher frequencies. The contrast sensitivity function of the yellow/blue channel shows a similar pattern of response after correction for chromatic aberration.

The high frequency cutoffs of the CSF's indicate the limits of spatial resolution in the two channels. The achromatic channel has a cutoff at approximately 30 cpd which is in good correspondence with the limits of visual acuity measured in clinical tests. The high frequency cutoff for the chromatic channels is only around 11 cpd. This means that the chromatic channels have much lower spatial resolution than the achromatic channel.

The contrast sensitivity functions have been widely used to model the visual system's response to complex objects. If the image of an object can be described in terms of its sinusoidal Fourier components, then the visibility of that object can be measured by applying the contrast sensitivity function to the components. When the components are above threshold the object will be seen, when they're below threshold it will be invisible.

This approach to predicting the visibility of complex objects has been widely used, but there is a severe limit on its generality that is often overlooked which will lead to gross errors in the predictions, namely that all the grating patterns used to measure the CSF's have the same mean luminance. While the contrast sensitivity functions show how sensitivity varies with spatial frequency, they do not take into account the changes in sensitivity caused by adaptation. To account for changes in the visibility of real objects in real scenes, we need to understand the interactions of adaptation with threshold spatial vision.

2.3 Adaptation and Threshold Spatial Vision

The results of a classic study on the effects of adaptation on threshold spatial vision are shown in Figure 3. van Nes [1967] measured contrast sensitivity functions for achromatic gratings illuminated at a wide range of different levels. Each curve in the graph represents the CSF measured at a particular luminance level.

There are several things to notice in the graph. The first is that overall, contrast sensitivity improves with the level of illumination. Peak contrast sensitivity changes from a value of 8 (threshold contrast of 12%) at an illumination level of 0.0009 Trolands (Td) to a value of 500 (threshold contrast of 0.2%) at 5900 Td.

The next thing to notice is that the shape of the CSF changes from being lowpass at the lowest illumination levels to being bandpass at higher levels. This reflects the transition from rod vision in the scotopic range to cone vision at photopic levels.

The final thing to notice is that as the level of illumination increases, the high frequency cutoff of the CSF moves to higher and higher spatial frequencies. This corresponds to the improvement in spatial resolution and visual acuity that we experience at higher luminance levels. The cutoff changes from about 4 cpd at 0.0009 Td to about 50 cpd at 5900 Td which corresponds to an improvement in acuity from around 20/150 at the lowest level to almost 20/10 at the highest.

The curves in Figure 3 show the effects of adaptation on spatial contrast sensitivity in the achromatic channel of the visual system. Data from van der Horst [1969] shows a similar pattern of results in the chromatic channels.

These data begin to give us a clearer picture of the interactions between adaptation and threshold spatial vision. From these data we can begin to understand in a unified framework, the changes in

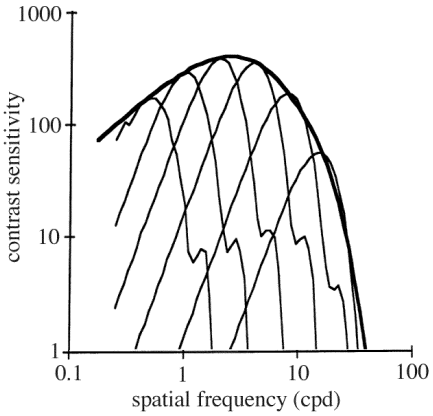

Figure 4: Multiscale bandpass mechanisms underlying the contrast sensitivity functions. Adapted from [Lubin95].

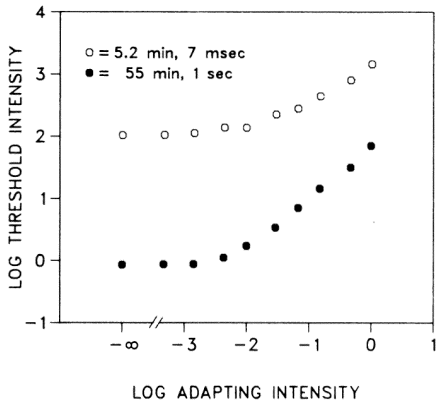

Figure 5: Threshold-vs.-intensity functions for spot patterns with different spatial and temporal parameters. Adapted from [Barlow72].

visibility, acuity, and color discrimination that occur with changes in the level of illumination. However there is one more aspect of these interactions that we need to investigate to have a more complete understanding, and this is the effect of local adaptation.

2.4 Local Adaptation and Multiscale Models of Threshold Spatial Vision

As we look from place to place in a scene our eyes adapt locally to the prevailing conditions of illumination. This local adaptation greatly enhances our ability to see in high dynamic range scenes where some portions of the scene are brightly illuminated and others are in shadow. How does local adaptation work?

Physiological and psychophysical evidence now indicates that the early stages of visual processing can be described as the filtering of the retinal image by visual mechanisms sensitive to patterns of different scale whose response characteristics are bandpass in the spatial frequency domain [Wilson91]. These multiple mechanisms are sensitive to different ranges of spatial frequencies, and the CSF's that are measured in psychophysical experiments are the envelope of these mechanism sensitivities. Figure 4 shows the achromatic CSF described in this way. Losada [1994] has shown that the chromatic CSF's can be described in a similar way.

Now if we look back at van Nes's data (Figure 3) on the changes in spatial contrast sensitivity that occur with changes in the level of illumination, it can be seen that the CSF curves don't simply shift upwards with increasing illumination, but change shape as well. This is a reflection of the fact that these bandpass mechanisms adapt to the average luminance within a region of a scene defined by their spatial scale. In a complex scene, this average is going to be different at different scales so the mechanisms will all be in different states of adaptation.

Thus local adaptation is not only spatially local within different regions of the visual field, but is also local with respect to the scale and spatial frequency filtering characteristics of the bandpass mechanisms involved in early visual processing. Therefore, to correctly account for the changes in visual sensitivity that occur with changes in the level of illumination, we need to describe the effects of local adaptation at different spatial scales.

Peli [1990] has suggested that an appropriate way to characterize the effects of local adaptation is to determine the *band-limited local contrast* at each location in the scene. Band-limited local contrast is calculated by first filtering the retinal image into a set of bandpass images defined by the filter characteristics of the visual mechanisms, and then dividing the signals in these images by lowpass images that represent the average local luminance at each location in the image at different spatial scales. This produces a multiscale representation of the image where the signals in each band represent the effective contrasts at each scale, having taken the effects of local adaptation into account. Both Peli and Lubin [1995] have shown that this kind of representation corresponds well with perceived threshold contrasts in complex images.

2.5 Suprathreshold Models of Vision

Threshold models of vision allow us to define the borders between the visible and the invisible. These models have a long and useful history in applied vision research, but because threshold models only define the limits of vision, they don't really tell us much about ordinary "seeing" where the contrasts, sizes of spatial details, and color saturations are typically well above threshold. To characterize how changes in the level of illumination affect the everyday appearances of objects in scenes, suprathreshold models of vision are needed.

Stevens' [1961] model of brightness and apparent contrast shown in Figure 1b summarizes much of what is known about the intensity dependence of surface appearance at suprathreshold levels. Stevens had subjects estimate the apparent brightnesses of gray patches seen against a white surround at different illumination levels. The brightness of the surround increased as a power function (exponent 0.33) of its luminance. The brightnesses of the gray patches either increased, decreased or remained constant depending on their contrast with respect to the surround. Overall, the diverging curves quantify a familiar aspect of our visual experience: as we turn up the light, the world becomes more vivid. Whites become brighter, blacks become deeper and the whole range of apparent contrasts expands. Although Stevens only tested achromatic surfaces, Hunt [1995] has measured a related set of phenomena for chromatic displays, where the vibrancy and *colorfulness* of colored surfaces increases at higher levels of illumination.

While these suprathreshold changes in brightness, colorfulness, and apparent contrast are certainly true to our everyday experience, it is difficult to reconcile these results with the predictions of threshold models and Weber's law which show that adaptation produces a visual system with constant contrast sensitivity, and imply that apparent suprathreshold contrasts should be constant over changes in the level of illumination within the Weber range. Differences in the TVI functions for different kinds of visual stimuli suggest a solution to this conundrum.

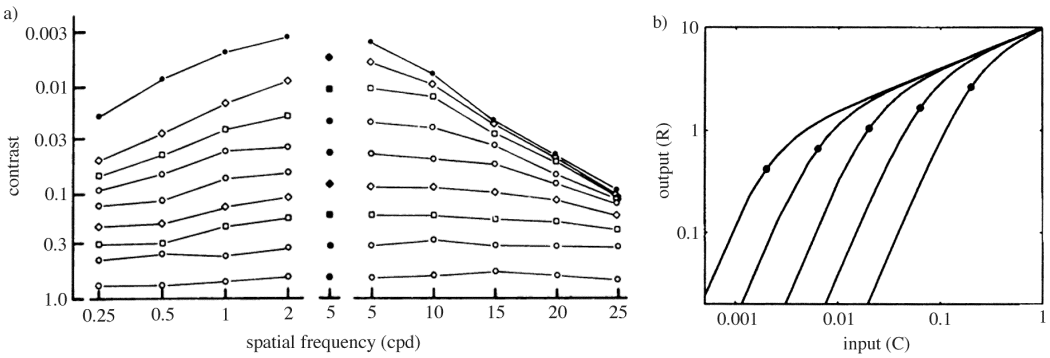

Figure 6: Suprathreshold contrast constancy and non-linear contrast transducers in human vision. Adapted from [Georgeson75, Watson97b].

2.6 Adaptation and Suprathreshold Vision

Figure 5 shows photopic TVI functions measured by Barlow [1972] for incremental spot patterns with different spatial and temporal characteristics. The lower curve shows thresholds measured for a large spot presented in a long exposure. The upper curve shows thresholds for a small, briefly flashed spot. There are two important differences between the curves. First, threshold values for the large/long spot are everywhere lower than thresholds for the small/brief spot. Second, although the slope of the large/long TVI curve follows Weber's law at higher background levels, the short/brief TVI has a lower slope indicating sub-Weber behavior.

The significance of the differences between these two TVI's is that low threshold values, and the constant contrast sensitivity implied by Weber's law, are only obtained under optimal conditions in laboratory experiments, such as those shown in the lower curve in this experiment or those given in the experiments that measured the rod and cone TVI's shown in Figure 1a. The TVI function for the small/brief spot is more like what we should expect under natural conditions where our eyes are continually moving across a complex scene and both the visual stimuli and our state of adaptation will be changing rapidly. Here threshold sensitivity is limited by both incomplete adaptation and quantal fluctuations in the stimulus, making thresholds higher-than-optimal, but also producing a TVI with a sub-Weber slope, where threshold contrast sensitivity continues to improve at higher illumination levels because the magnitude of the visual response to a constant physical contrast increases as the level of illumination rises. The insights provided by this experiment are the key that allow us to unify the predictions of threshold and suprathreshold models in one coherent framework. To complete the picture we've been developing of a model that can explain the interactions of adaptation and spatial vision at threshold and suprathreshold levels, we need to understand the properties of suprathreshold spatial vision.

2.7 Adaptation and Suprathreshold Spatial Vision

A CSF describes how threshold contrast sensitivity varies for sinusoidal gratings of different spatial frequencies. The upper curve in Figure 6a shows an achromatic CSF measured by Georgeson [1975]. This typical curve shows that we are most sensitive to gratings with frequencies around 4-5 cpd and that sensitivity falls off at both higher and lower frequencies. This CSF was measured as part of a suprathreshold *contrast matching* experiment. In this experiment subjects matched the apparent contrast of different frequency test gratings to the apparent contrast of a standard grating of 5 cpd. In separate trials, the physical contrast of the standard was varied from less than 1% to more than 75%. The lower curves in Figure 6a summarize the results.

At low standard contrasts, the matches followed the form of the threshold CSF. High and low frequency gratings had to have higher physical contrast to have the same apparent contrast as the standard. But at standard contrasts above about 20% the curves flattened out. Gratings matched in apparent contrast when they had the same physical contrast. Georgeson called this phenomenon *contrast constancy*.

The differences in the shapes of the threshold CSF and the suprathreshold contrast matching functions indicates the existence of nonlinear contrast transduction processes in the visual system. Brady [1995] has suggested that these contrast nonlinearities reflect differences in the signal/noise characteristics of the bandpass visual mechanisms which can explain both the curvature of the threshold CSF and the flatness of the suprathreshold contrast matching functions.

Watson and Solomon [1997b] have developed a model of these contrast nonlinearities as part of their work on the closely related phenomenon of visual masking. The transducer functions for a set of hypothetical bandpass mechanisms are shown in Figure 6b. The horizontal axis indicates the input contrast, the vertical axis plots the response. At low input contrast levels, the transducers all have different cutoffs. The differences in these cutoffs imply that the mechanisms all have different sensitivities, since at any particular low level, the input signal will be above the cutoffs for some mechanisms and below for others. This characteristic of the transducer functions accounts for CSF-like responses at low contrast levels. But at higher input contrast levels, the transducer functions converge, which means that given the same input contrast, all the mechanisms will produce the same response. This characteristic accounts for the contrast constancy observed at higher contrast levels. The action of these transducer functions, which mimics the contrast nonlinearity in the visual system, provides a coherent framework for understanding threshold and suprathreshold spatial vision.

2.8 Summary

In the previous sections we have outlined a coherent framework for understanding the changes in vision that occur with changes in the level of illumination in scenes. The framework relates research on adaptation with research on spatial vision to provide a unified view of variations in threshold performance and suprathreshold appearance at different illumination levels. The framework allows us to account for the changes in threshold visibility, visual acuity, and color discrimination, and suprathreshold brightness, colorfulness, and apparent contrast that occur under different illumination conditions. The key features of the framework are:

- Multiscale processing of the retinal image by visual mechanisms in the rod and cone pathways with bandpass spatial

frequency response characteristics.

- Adaptation processes operating within these bandpass mechanisms that act as "luminance gain controls" to produce visual signals that are primarily correlated with scene contrasts, but are still luminance dependent and increase in magnitude with increasing luminance. Independent adaptation processes within the bandpass mechanisms in the rod and cone pathways account for "local" adaptation, chromatic adaptation, and changes in sensitivity over the scotopic to photopic range.

- Organization of the bandpass mechanisms in the rod and cone pathways into an achromatic and two chromatic channels with different spatial frequency response characteristics.

- Nonlinear transducers operating on the adapted outputs of the mechanisms in these achromatic and chromatic channels that scale the visual signals to produce CSF-like response at threshold levels, and contrast constancy at suprathreshold levels.

In the following section we will develop a computational model of vision based on this framework and apply it to the problem of realistic tone reproduction. The unique features of the model will allow our tone reproduction operator to address the two major problems in realistic tone reproduction: images of wide absolute range and high dynamic range scenes can be displayed on conventional display devices; and these images should match our perceptions of the scenes at threshold and suprathreshold levels.

3 THE COMPUTATIONAL MODEL

3.1 Overview

We will now draw on the psychophysical framework outlined in the previous section to develop a computational model of adaptation and spatial vision for realistic tone reproduction. Figure 7 provides a flow chart of each major step in this computational model. The model has two main parts: the *Visual model* and the *Display model*. The visual model processes an input image to encode the perceived contrasts for the chromatic and achromatic channels in their band-pass mechanisms. The display model then takes this encoded information and reconstructs an output image. The model must be inverted in order to produce equivalent appearances under the viewing conditions of the display device. This procedure does not "undo" the processes of the model since the thresholding and saturation procedures are accomplished and the gain control parameters differ for the original scene and the display. The reconstruction process creates an output image that reproduces the visual appearance of the input image to the limits possible on a given display device. The specific computational procedures that were used to implement each step of the model are described below. A pictorial representation of the signals at each stage of the model is presented in Figure 8.

3.2 Input Image Preprocessing

Prior to applying the visual model, certain preprocessing steps are required to assure that the image data appropriately correspond to the information accessible to the early stages of the human visual system. First, the image must be spatially sampled such that the band pass signals represent appropriate spatial frequencies. For this implementation, band pass mechanisms with peak spatial frequencies of 16, 8, 4, 2, 1, and 0.5 cpd were required. The spatial sampling necessary to produce these band-pass signals depends upon the Gaussian filters chosen for the image decomposition. With the

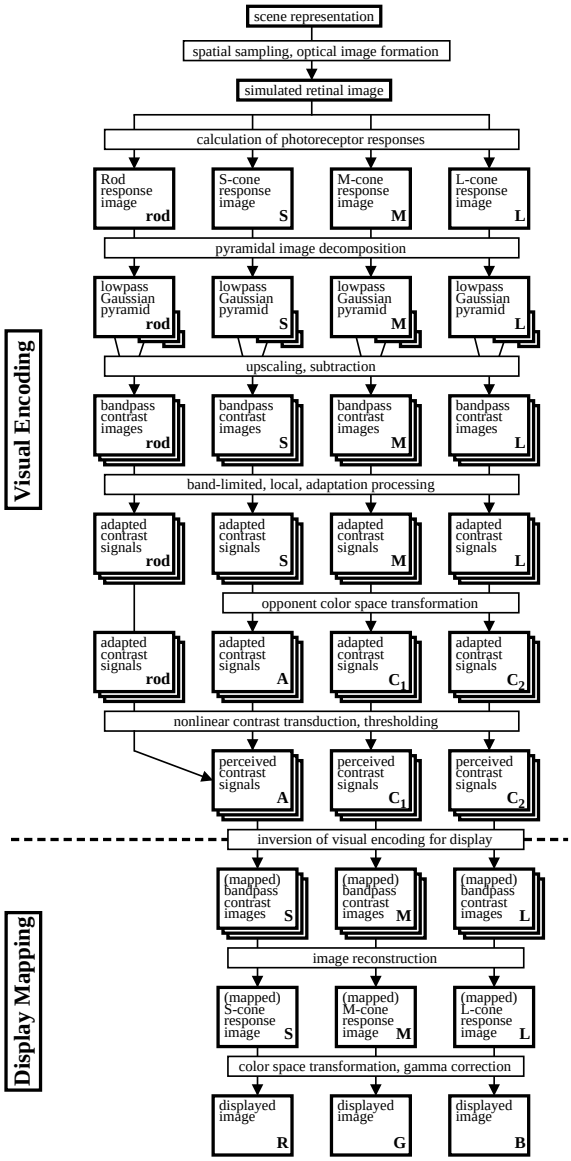

Figure 7: Flow chart of the computational model of adaptation and spatial vision for realistic tone reproduction.

filter described in Section 3.3 it is necessary to sample the image at a rate of 130 pixels/degree.

The next step is to introduce compensations for optical point-spread in the eye and disability glare. Optical point-spread is incorporated via a convolution with a function described by Westheimer [1986] and glare effect is introduced via convolutions with functions described by Spencer [1995].

The image must then be spectrally sampled to represent the visual system's initial photoreceptor responses. This is accomplished by integrating the spectral radiance distribution for each pixel after multiplication by the spectral responsivities of the long-, middle- and short-wavelength sensitive cones (LMS) and the rods. We use the Hunt-Pointer-Estevez [Fairchild98] responsivities for the cones and the CIE scotopic luminous efficiency function, $V'(\lambda)$, [Wyszecki82] for the rods.

In many applications, a spectral radiance image is not available. In such cases, the cone signals can be calculated as a linear transform of the CIE 1931 XYZ tristimulus values as shown in Equa-

tion 1.

$$\begin{vmatrix} L \\ M \\ S \end{vmatrix} = \begin{vmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} \quad (1)$$

However, it is impossible to obtain the proper rod signals. We derived a linear transform of XYZ tristimulus values as a rough approximation to the rod signal via linear regression of the color matching functions and the $V'(\lambda)$ curve. The resulting transform is given in Equation 2 where R represents the rod response for a pixel.

$$R = -0.702X + 1.039Y + 0.433Z \quad (2)$$

Since it is possible to obtain negative values of R when Equation 2 is applied to some saturated colors, it must be clipped to zero. We chose a simple linear transformation for this approximation since it scales over any range of luminance levels.

Finally the input signals must be calibrated prior to input to the visual transforms. We chose to calibrate the model such that the LMS cone signals and the rod signal are all equal to unity for an equal-radiance spectrum at a luminance of 1.0 cd/m^2.

3.3 Spatial Decomposition

The 4 images representing the calibrated photoreceptor responses are then subjected to spatial processing. The first step is to carry out the spatial decomposition of these images. We carry out this decomposition by the Laplacian pyramid (difference-of-Gaussian pyramid) approach proposed by Burt and Adelson [1983]. This approach guarantees the construction of a non-negative low-pass image in high dynamic range situations, and is perfectly invertible. We first calculate a Gaussian pyramid using a 5 tap filter (with 1D weights: .05 .25 .4 .25 .05) [Burt83]. Each level of the Gaussian pyramid represents a low-pass image limited to spatial frequencies half of those of the next higher level. Our Gaussian pyramid has 7 levels.

Each level of the Gaussian pyramid is then upsampled such that each image is returned to the size of the initial image. Difference-of-Gaussian images are then calculated by taking the image at each level and subtracting the image from the next lower level. This results in 6 levels of band-pass images with peak spatial frequencies at 16, 8, 4, 2, 1, and 0.5 cpd. These images can be thought of as representations of the signals in six band-pass mechanisms in the human visual system. The lowest-level low pass image is retained since it must be used to reconstruct the image for reproduction applications.

3.4 Gain Control

The difference-of-Gaussian images are then converted to adapted contrast signals using a luminance gain control. The gains are set using TVI-like functions that represent the increment thresholds of the rod and cone systems and the growth in response required to allow perceived contrast to increase with luminance level (sub-Weber's law behavior). The gain functions are given for the cones in Equation 3 and the rods in Equation 4.

$$G_{cone}(I) = \frac{1}{0.555(I+1.0)^{0.85}} \quad (3)$$

$$G_{rod}(I) = \left[\frac{10}{I^2+10}\right]\left[\frac{1}{0.908(I+0.001)^{0.85}}\right] \quad (4)$$

In the above equations, I represents the rod or cone signal that is used to set the level of adaptation and $G(I)$ is the gain-control factor. Equations 3 and 4 were derived to match available psychophysical TVI and brightness matching data. The constraints in their derivation were that both the rod and cone gains were set equal to 1.0 at a 1.0 cd/m^2, the absolute thresholds would be around 1.0 cd/m^2 for cones and 0.001 cd/m^2 for rods, the ultimate slopes of the functions would be 0.85 for sub-Weber's Law behavior, and the rods would saturate, losing 50% of their responsivity at roughly 3 cd/m^2. In our model, each pixel in a given difference-of-Gaussian image is multiplied by the gain derived from the corresponding pixel in the lower-level low-pass image that was used in its derivation. This is illustrated in Equation 5.

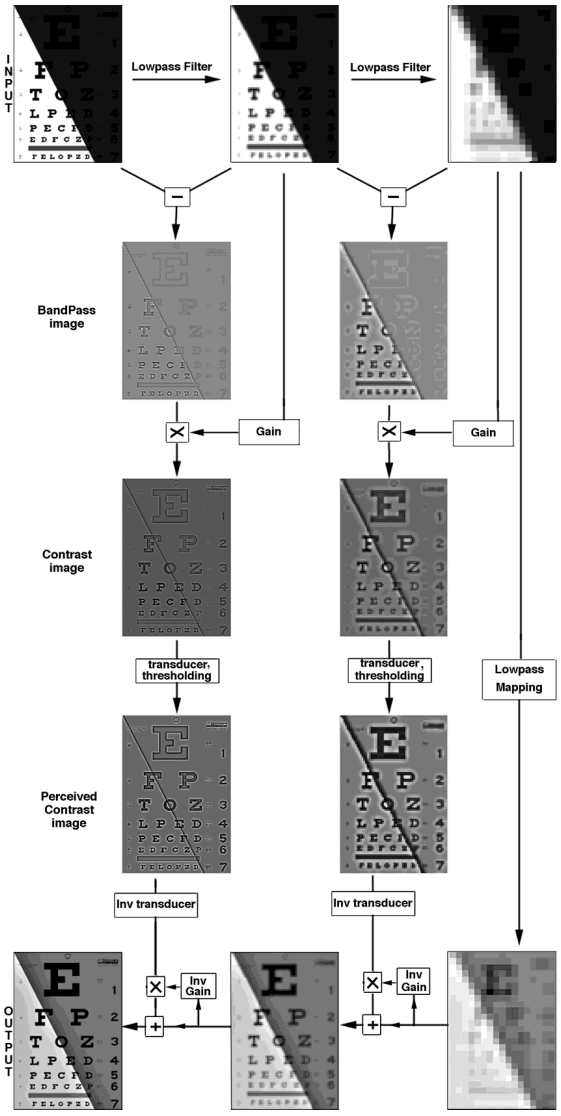

Figure 8: Pictorial representation of the computational model. Note that only 2 out of the 6 spatial mechanisms of one of the channels have been shown for the purpose of illustration. Original image is a Snellen chart with a 30:1 shadow boundary.

$$ACI_n = G\left(LP_{n+1}\right)\left[LP_n - LP_{n+1}\right] \quad (5)$$

ACI_n is the adapted contrast image at level n and LP represents the various low-pass images. The adapted contrast images are analogous to the contrast images that Peli [1990] obtained. However, in our model the magnitude of these images is a function of luminance level as specified by the gain control functions. This is necessary to allow prediction of luminance-dependent appearance

effects. The luminance gain controls are applied in the same manner to each of the difference-of-Gaussian images for each of the photoreceptors. Equation 3 is used to calculate the gains for each of the cones and Equation 4 is used for the rods. Note that performing the gain control at this point in the model allows proper prediction of chromatic-adaptation effects.

3.5 Opponent Color Processing

The next stage of the model is to transform the adapted contrast images for the cones into opponent signals. We use the transform of Hunt [1995] that has also been recently adopted in the CIE color appearance model, CIECAM97s [Fairchild98] as given in Equation 6.

$$\begin{vmatrix} A \\ C_1 \\ C_2 \end{vmatrix} = \begin{vmatrix} 2.0 & 1.0 & 0.05 \\ 1.0 & -1.09 & 0.09 \\ 0.11 & 0.11 & -0.22 \end{vmatrix} \begin{vmatrix} L \\ M \\ S \end{vmatrix} \quad (6)$$

In the above equation, L, M, S represent the cone signals and A, C_1, C_2 represent luminance, red-green, and yellow-blue opponent signals respectively. This transform is applied without modification to the adapted contrast signals to obtain adapted contrast signals in an opponent color space. This transformation is necessary to model differences in the spatial processing of luminance and chromatic signals. At this stage, the rod images are retained separately since their spatial processing attributes also differ from the cones.

3.6 Adapted Contrast Transducers

The adapted contrast signals are then passed through contrast transducer functions similar to those described by Watson and Solomon [1997]. Different transducer functions are applied to each spatial frequency mechanism in order to model psychophysically derived human spatial contrast sensitivity functions. For example, the transducer for the 16 cpd achromatic mechanism has a higher threshold than the transducer for the 4 cpd achromatic mechanism since we are less sensitive to the higher spatial frequencies. The transducers are also different for the chromatic channels to represent their lower sensitivities and low-pass, rather than band-pass nature. Finally, the rod system has a distinct set of transducers to represent its unique spatial characteristics. At high contrast levels the transducer functions converge to a common square-root form to properly represent perceived contrast constancy and introduce a compressive nonlinearity typically found in color appearance models. The functional form of our transducer functions vary from that proposed by Watson and Solomon [1997] since their function was not analytically invertible and it is necessary to invert our model for image reproduction applications. We chose to use a two-part function consisting of two power functions in order to replicate the two regions of distinct slope in the transducer functions. The contrast transducers used in our model are given by Equations 7 and 8 for the cones and Equation 9 for the rods.

$$T_{\text{cone,Achromatic}}(c) = \begin{cases} 22.4\,(c/0.536)^{\frac{1}{2}} & \text{if } c \geq 0.536 \\ 22.4\,(c/0.536)^{p} & \text{otherwise.} \end{cases} \quad (7)$$

$$T_{\text{cone,Chromatic}}(c) = \begin{cases} 22.4\,(c/0.176)^{\frac{1}{2}} & \text{if } c \geq 0.176 \\ 22.4\,(c/0.176)^{p} & \text{otherwise.} \end{cases} \quad (8)$$

$$T_{\text{rod}}(c) = \begin{cases} 22.4\,(c/0.0335)^{\frac{1}{2}} & \text{if } c \geq 0.0335 \\ 22.4\,(c/0.0335)^{p} & \text{otherwise.} \end{cases} \quad (9)$$

In the above equations, c represents the adapted contrast signals (ACI's) and $T(c)$ represents the output of the transducers. The exponent, p, in Equations 7, 8 and 9 differs for each spatial frequency mechanism as given in the following table.

Peak(cpd)	.5	1.0	2.0	4.0	8.0	16.0
p for A	1.93	1.35	1.15	1.04	1.15	1.40
p for C_1&C_2	1.93	1.93	2.35	2.85	-	-
p for Rod	3.39	3.39	4.50	7.64	-	-

If c is negative, the absolute value of c is taken and then the negative sign is replaced after transformation.

Equations 7 through 9 were derived by specifying the desired thresholds for the various spatial and chromatic mechanisms at 1000 cd/m^2 for the cones and 0.5 cd/m^2 for the rods. The lower parts of the functions were forced to pass through 1.0 for the desired threshold contrast. At sinusoidal contrasts greater than 5% (at these calibration luminance levels) the functions converge to the square-root form that produces contrast constancy. The square root is derived to mimic the compressive nonlinearities typically found in color appearance models. These transducers do produce CSF behavior that changes with luminance as illustrated in Figure 3 since the input adapted contrast signals vary with luminance due to the sub-Weber gain control functions. Transducers for the chromatic cone channels and the rod channel do not exist for spatial frequencies above those represented by the mechanism centered at 4 cpd since these systems cannot resolve contrast at higher spatial frequencies. (Note that the 4 cpd mechanism carries information out to about 16 cpd which is thus the acuity limit of the mechanism.) The contrast transducer functions are calibrated such that psychophysical contrast sensitivity function data are modeled and sinusoidal contrasts above about 5% produce contrast constancy as a function of spatial frequency and sinusoidal contrasts of 100% produce transducer output of approximately 100. It should be recalled that in our model, these levels will be luminance dependent. The contrast transducer functions are also designed such that contrasts that are below threshold have an output level with magnitude less than 1.0.

One of the key functions of the transducers is to set the threshold level such that image content that is imperceptible for a given set of viewing conditions can be removed. To accomplish this, the output of the transducer functions is thresholded such that all absolute values less than 1.0 are set to 0.0. An alternative approach would be to replace all absolute values less than 1.0 with a random number between 0.0 and 1.0. This might better replicate the appearance of visual noise at low contrast levels.

In addition to establishing thresholds, the transducer functions are used to model saturation of the visual neurons that signal contrast. Thus, the transducer functions are limited to maximum values of 50 to simulate the typical physiological dynamic range. [Hunt95] Since the contrast mechanisms are bipolar, this represents a 100:1 dynamic range in each spatial mechanism and therefore even larger perceptual dynamic ranges in fully reconstructed images. This saturation is also not a severe restriction on the image content since the gain control mechanisms already accomplish a high degree of dynamic-range compression.

3.7 Combination of Rod and Cone Signals

Up to this point in the model it is necessary to keep the rod signals separate in order to appropriately integrate their unique adaptation and spatial vision properties. After the contrast transducers, the rod and cone signals can be combined to produce signals that represent the three-dimensional color appearances of the input image. We assume that the rods contribute only to the luminance signal and thus combine the A signal from the cones with the rod signal, denoted A_{rod}, to produce a total achromatic signal, A_{total}, using Equation 10.

$$A_{total} = A_{cone} + A_{rod}/7 \quad (10)$$

The differential weighting of the rod and cone signals is a result of the model calibration necessary to establish the rod and cone gain

controls and transducer functions. It results in a total achromatic output that is monotonic with luminance.

At this stage in the model we have three channels representing achromatic, red-green, and yellow-blue apparent contrast for 6 band-pass mechanisms. These signals model threshold behavior, in that any contrast signals that could not be perceived have been eliminated at the contrast transducer functions. They also represent suprathreshold appearance since the contrast signals grow with luminance and the chromatic channels will become zero at luminance levels below the cone threshold. At this stage, the model has also accomplished a significant level of dynamic-range compression since the contrast signals range only 2 orders of magnitude (1 to around 100) for luminance differences ranging over 10 orders of magnitude. This compression is accomplished by both the gain control functions and the nonlinear transducers.

3.8 Treatment of the Low Pass Image

The lowest level low-pass image from the upsampled Gaussian pyramid must be retained in order to reconstruct an image from the adapted contrast images that have been passed through the model (each a band-pass image). To this point, we have not discussed the application of the visual model to this low pass image. The best approach to processing the low-pass image depends on the application. For simple images of low dynamic range (*e.g.*, less than 50:1), an appropriate treatment of the low-pass image is to multiply it by a constant gain factor derived from the image mean. This technique will do little to compress the range of high-dynamic range images since the contrast within the low pass image will be preserved. An alternative that produces maximum dynamic-range compression is to multiply each pixel in the low-pass image by a gain factor derived from the pixel itself. (The gain factors are derived using Equations 3 and 4.) Techniques intermediate between these two might produce optimal image reproductions for various applications. The above treatment of the low-pass image is consistent with the full model of visual perception and can be thought of as a treatment of the effects of eye movements on the perception of a scene. In the extreme case of adapting the low pass image to its own values, the treatment mimics the visual response assuming that observer fixated on each and every image location and judged them completely independent of one another. For the other extreme case of adapting the low pass image using the mean signal, the treatment simulates completely random and continuous eye movements uniformly distributed across the scene. Intermediate treatments between these two extremes might more accurately model real world eye movements which are scene dependent and represent some average between fixating each image element of interest and randomly viewing all locations in a scene.

Transducer functions are necessary for the low pass image as well since the rod and cone information is combined after the transducer stage. We have adopted low-pass transducers that are simple power functions based on typical practice in color appearance modeling. [Fairchild98] The scaling of the low-pass transducers is set to preserve equal magnitude of signals for the low-pass and band-pass model output for a sinusoidal grating. The low-pass transducers are given in Equations 11, 12 and 13 for the achromatic and chromatic cone signals and rod signals respectively.

$$T_{LP\_cone,Achromatic}(LP) = 30.5\,(LP)^{\frac{1}{2}} \quad (11)$$

$$T_{LP\_cone,Chromatic}(LP) = 53.3\,(LP)^{\frac{1}{2}} \quad (12)$$

$$T_{LP\_rod}(LP) = 122\,(LP)^{\frac{1}{2}} \quad (13)$$

T represents the output of the low-pass transducers and LP represents the pixel values in the adapted, opponent-transformed, low-pass image.

3.9 Image Reconstruction for Display

The output of the visual model consists of appearance signals in an achromatic and two chromatic channels and six spatial band-pass mechanisms plus a low-pass image. We now take these appearance signals backward through the model to recreate cone signals (and ultimately device signals such as RGB or CMYK) that replicate the full color appearance of the image on a photopic, trichromatic display device such as a CRT display.

The first step of the inversion process is to go through the inverse of the transducer functions given in Equations 7, 8, 11 and 12. The $\mathrm{AC_1C_2}$ signals are then transformed to adapted LMS cone signals using the inverse of the matrix transformation given in Equation 6. At this point we have adapted contrast signals that have been subjected to the appropriate visibility thresholding and saturation by the contrast transducer functions.

The next step is to reverse the gain control process for the viewing conditions of the output display. This begins by determining the gain control factors for the mean luminance of the target display device using Equation 3. The adapted low-pass images are then divided by the display-mean gain control factors to produce images that represent the appropriate LMS cone responses for the display low-pass image. This display low-pass image is then used to begin the process of reconstruction of a full-resolution image from the six adapted contrast signal images.

Gain control factors are calculated for each pixel of the display low-pass image using Equation 3 and these are used to scale the lowest frequency (0.5 cpd peak) adapted contrast signal image back to the display. This image is then added to the low-pass image to produce a new low-pass image that includes the contrast information from the 0.5 cpd image. This image is then used to calculate gain control factors that are applied to the next level (1.0 cpd peak). The resulting image is again added to the new low-pass image to generate yet another low-pass image that incorporates the information from both the 0.5 and 1.0 cpd mechanisms. The process is repeated at each level until all of the spatial frequency mechanisms have been scaled to the display and added to the output image.

At this point in the reconstruction we have an LMS image representing the cone signals that are desired when viewing the output display. These must be converted to signals appropriate for the given display device using typical device-independent color imaging procedures. For a CRT this involves a linear transform from LMS to CIE XYZ (inverse of Equation 1) followed by a second linear transform from CIE XYZ to device RGB. The second transform is defined by the CIE tristimulus values of the display primaries. At this point we have linear RGB signals that must either be transformed through the inverse of the CRT display transfer function (often referred to as gamma correction) or displayed on a system with linearizing video look-up tables. A similar, although more complex, process is required for printing devices.

Finally, it is not uncommon that the desired display colors simply cannot be produced on a given device (*i.e.*, they are out of gamut). This includes the mapping of the desired luminance level and range into that of the display. There are a wide variety of techniques that have been suggested to address this issue the details of which are beyond the scope of this paper. The approach that is taken depends on the particular application. Some of the issues that are encountered with this particular model are discussed in the next section with respect to the rendering on paper and on CRT monitor of the various example images.

Figure 9: Application of the model to a wide range of illumination levels.

4 APPLYING THE MODEL

4.1 Wide Absolute Range

The series of images in Figure 9 illustrate application of the model to a wide range of luminance levels spanning six orders of magnitude from 0.1 to 10,000 cd/m^2. These images were created using the model as described in section 3 with the low-pass images adapted to the mean luminance of the input images. The size of the original image was about $15^\circ \times 10^\circ$. For image reconstruction as printed images, a mean adapting luminance of 700 cd/m^2 was assumed. This is approximately the luminance of a standard print viewing booth. Thus this series of images should provide faithful reproductions of the visual impression at the various luminance levels when the printed figure is viewed at a mean luminance of 700 cd/m^2. The gamut-mapping selected for this demonstration was a linear scaling that placed the white areas of the 1000 cd/m^2 image at the paper white. While the model can be applied successfully over a wider absolute range, it is impossible to reproduce the results within the limited dynamic range (approximately 50:1) of the printed images unless a variable scaling is used.

Features to note in Figure 9 include: the decrease in luminance contrast and colorfulness as luminance is decreased, the loss of color vision upon the transition from cone to rod vision below 1 cd/m^2, the decrease in spatial acuity with decrease in luminance, and the changes in relative visibility of various colors and patterns. The Purkinje shift (blue to gray and red to black) is also correctly predicted upon changes from photopic to scotopic luminance levels. All of these features illustrate that the model has appropriately encoded aspects of threshold visibility and suprathreshold appearance over a wide range of luminance levels.

4.2 Chromatic Adaptation

Figure 10 shows the unique feature of this model that it can handle changes in chromatic, as well as luminance-level, adaptation. The top row of images illustrate a scene illuminated by a very reddish light source, a nearly-white incandescent light source, and a very blue light source as they would be rendered by a system incapable of chromatic adaptation. The shift in color balance of the reproduced prints is objectionable since the human visual system largely compensates for these changes in illumination color through its mechanisms of chromatic adaptation. Since our model treats gain control in each of the classes of cone photoreceptors independently, it is capable of predicting changes in chromatic adaptation similar to those that would be predicted by a von Kries model. However, due to the nature of the gain control functions used to obtain increases in contrast and colorfulness with luminance, the degree of chromatic adaptation predicted by the model is less than 100% complete.

The bottom row of images illustrate the output of the visual model when the low-pass images are adapted to the mean signal levels in the image and the reconstructed images are created assuming adaptation to an equal-energy white. All of the computations were completed at a mean luminance of 50 cd/m^2. The gamut-mapping selected for these images was a linear scaling that mapped 100 cd/m^2 in the reconstructed image to the monitor white. 100 cd/m^2 is approximately the maximum luminance of a display monitor. The sizes of the original images were $10^\circ \times 8^\circ$. These images illustrate that the model almost completely accounts for the changes in illumination color. However, as expected the reproduced appearance from the reddish light source retains a slight reddish cast while the reproduction from the bluish light source retains a slight bluish cast. These reproductions match our perceptions of changes in illumination color and replicate the incomplete nature of chromatic adaptation that is widely recognized in the color science literature. [Fairchild98]

4.3 High Dynamic Range

Figure 11 illustrates application of the model to the tone mapping of high-dynamic range images. The original images have areas of detail that are in high illumination levels and other areas that are in low illumination levels. The left most image in Figure 11 is a global illumination rendering. The other two were constructed from successive photographic exposures using the technique of Debevec and

Figure 10: Illustration of chromatic adaptation.

Malik [1997]. To provide the higher degree of compression necessary for high-dynamic-range mapping, the low pass image was adapted to itself. The reproduced images were reconstructed for display at a mean luminance of 50 cd/m^2.

The images on the top row of Figure 11 are linear mappings of the original high-dynamic range images into the limited dynamic range of the output device. The original images had luminance ranges of approximately 10,000:1. The images on the bottom row represent the mapping obtained by application of the visual model. In Figure 11 it is clear that far more detail can be observed in both shadow and highlight regions when the images are mapped using the visual model.

5 CONCLUSIONS AND FUTURE WORK

In this paper we have introduced a new visual model for realistic tone reproduction. The model is based on a multiscale representation of luminance, pattern, and color processing in the human visual system, and provides a coherent framework for understanding the effects of adaptation on spatial vision. The model allows us to account for the changes in threshold visibility, visual acuity, and color discrimination, and suprathreshold brightness, colorfulness and apparent contrast that occur with changes in the level of illumination in scenes.

We have applied the visual model to the problem of realistic tone reproduction and have developed a tone reproduction operator that addresses the two major problems in realistic tone reproduction: images of wide absolute range and high dynamic range scenes can now be displayed on conventional display devices like CRTs and printers; and these images are faithful visual representations that match our perceptions of the scenes at threshold and suprathreshold levels to the limits possible on a given display. This work should have major impact on the field of digital imaging. Scenes that could never be reproduced before can now be imaged with high visual fidelity.

Beyond the clear applications of this work in realistic tone reproduction, the visual model presented in this paper can be usefully applied in a variety of other areas in digital imaging where the characteristics of threshold and suprathreshold vision are important. Potential application areas include: image quality metrics; image coding and compression methods; perceptually-based image synthesis algorithms; image-based rendering; and advanced display system design.

There is still much work to be done in this area. First, this is a static model of vision. Future models should incorporate knowledge about the temporal aspects of visual processing in order to allow both dynamic scenes, and scenes where the level of illumination is dynamically changing to be properly displayed. Second, we hope to integrate our visual model with ongoing work in the color science community on appearance models [Fairchild98] for predicting how images look under different viewing conditions. We should also draw on the testing methods developed by researchers in this community to verify that our images are in fact good visual matches to actual scenes. Finally, a number of researchers [Gilchrist77, Adelson93] in the vision community have shown that threshold and suprathreshold properties of scenes aren't simply a function of the two dimensional patterns of luminances in the retinal image, but also depend upon our perceptions of the spatial arrangement of surfaces and illumination in three dimensions. Future work should address how these 3D issues affect our perceptions of scenes and influence the development of operators for realistic tone reproduction.

ACKNOWLEDGEMENTS

Special thanks to David Hart, Steve Marschner, Hurf Sheldon, Chris Williams, Ben Trumbore, Eric Lafortune and Dan Kartch for their help in preparing this paper.

This work was supported by the NSF Science and Technology Center for Computer Graphics and Scientific Visualization (ASC-8920219) and by NSF grant ASC-9523483 and was performed on workstations generously provided by the Hewlett-Packard Corporation.

References

[Adelson93] Adelson, E.H. (1993) Perceptual organization and judgment of brightness. Science, 262, 2042-2044.

[Barlow72] Barlow, H.B. (1972) Dark and Light Adaptation: Psychophysics. In D. Jameson and L. Hurvich (Eds.), *Handbook of Sensory Physiology*, V. 7(4), New York: Springer, 2-27.

Figure 11: Tone mapping of high-dynamic range images. The images on the top are linear mappings of the original high-dynamic range images. The images on the bottom are the mapping obtained by application of the visual model.

[Brady95] Brady, N. and Field, D.J. (1995) What's Constant in Contrast Constancy? The Effects of Scaling on the Perceived Contrast of Bandpass Patterns. Vision Res., 35(6), 739-756.

[Burt83] Burt, P.J., and Adelson, E.H. (1983) The Laplacian Pyramid as a Compact Image Code. IEEE Transaction on Communication, 31(4), 532-540.

[Chiu93] Chiu, K., Herf, M., Shirley, P., Swamy, S., Wang, C., and Zimmerman, K. (1993) Spatially Nonuniform Scaling Functions for High Contrast Images. Proceedings Graphics Interface 93, 245-254.

[Debevec97] Debevec, P.E. and Malik, J. (1997) Recovering High Dynamic Range Radiance Maps from Images. Proceedings SIGGRAPH 97, 369-378.

[Fairchild98] Fairchild, M.D. (1998) *Color Appearance Models*. Reading, MA: Addison-Wesley.

[Ferwerda96] Ferwerda, J.A., Pattanaik, S.N., Shirley, P., and Greenberg, D. (1996) A Model of Visual Adaptation for Realistic Image Synthesis. Proceedings SIGGRAPH 96, 249-258.

[Georgeson75] Georgeson, M.A. and Sullivan, G.D. (1975) Contrast Constancy: Deblurring in Human Vision by Spatial Frequency Channels. J. Physiol., 252, 627-656.

[Gilchrist77] Gilchrist, A.L. (1977) Perceived Lightness Depends on Perceived Spatial Arrangement. Science, 195, 185-187.

[Hunt95] Hunt, R.W.G. (1995) *The Reproduction of Color*. 5th edition, Kingston-upon-Thames, England: Fountain Press.

[Hurvich81] Hurvich, L. (1981) *Color Vision*. Sunderland, MA: Sinauer Assoc.

[Jobson96] Jobson, D.J., Rahman, Z., and Woodell, G.A. (1996) Retinex Image Processing: Improved Fidelity to Direct Visual Observation. Proceedings 4th Color Imaging Conference, Society for Imaging Science and Technology, 124-126.

[Laming91] Laming D. (1991) Contrast Sensitivity. In J.J. Kulikowski, V. Walsh, and I.J. Murray (Eds.) *Limits of Vision*, Vol. 5, Vision and Visual Dysfunction. Boca Raton, FL, CRC Press, 35-43.

[Lubin95] Lubin, J. (1995). A Visual Discrimination Model for Imaging System Design and Evaluation. In E. Peli (Ed.) *Vision Models for Target Detection*. Singapore, World Scientific, 245-283.

[Losada94] Losada, M.A., and Mullen, K.T. (1994) The Spatial Tuning of Chromatic Mechanisms Identified by Simultaneous Masking. Vision Res., 34(3), 331-341.

[Mullen85] Mullen, K.T. (1985) The Contrast Sensitivity of Human Color Vision to Red-Green and Blue-Yellow Chromatic Gratings. J. Physiol., 359, 381-400.

[Peli90] Peli, E. (1990) Contrast in Complex Images. J. Opt. Soc. Am. A, 7(10), 2032-2040.

[Schlick95] Schlick, C. (1995) Quantization Techniques for High Dynamic Range Pictures. In G. Sakas, P. Shirley, and S. Mueller, (Eds.), *Photorealistic Rendering Techniques*, Berlin: Springer-Verlag, 7-20.

[Shapley84] Shapley, R. and Enroth-Cugell, C. (1984) Visual Adaptation and Retinal Gain Controls. In N. Osborne and G. Chader (Eds.). *Progress in Retinal Research*, V. 3, Oxford: Pergamon Press., 263-347.

[Spencer95] Spencer, G., Shirley, P., Zimmerman, K., Greenberg, D. P. (1995) Physically-based Glare Effects for Computer Generated Images. Proceedings SIGGRAPH 95, 325-334.

[Stevens61] Stevens, S.S. (1961) To Honor Fechner and Repeal His Law. Science, 133, 13 Jan., 80-86.

[Tumblin93] Tumblin, J., and Rushmeier, H. (1993) Tone Reproduction for Realistic Images, IEEE Computer Graphics and Applications, 13(6), 42-48.

[Tumblin97] Tumblin, J., Hodgkins, J. and Guenter, B. (1997) Display of High Contrast Images Using Models of Visual Adaptation. Visual proceedings SIGGRAPH 97, 154.

[vanNes67] van Nes F.L. and Bouman M.A. (1967) Spatial Modulation Transfer in the Human Eye. J. Opt. Soc. Am., 57, 401-406.

[vanderHorst69] van der Horst, G.J.C. and Bouman, M.A. (1969) Spatiotemporal Chromaticity Discrimination. J. Opt. Soc. Am., 59(11), 1482-1488.

[Ward94] Ward, G. (1994) A Contrast-based Scalefactor for Luminance Display. In P.S. Heckbert (Ed.), Graphics Gems IV, Boston: Academic Press Professional.

[Ward-Larson97] 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.

[Watson97] Watson, A.B. and Solomon, J.A. (1997) Model of Visual Contrast Gain Control and Pattern Masking. J. Opt. Soc. Am. A, 14(9), 2379-2391.

[Westheimer86] Westheimer, G. (1986) The Eye as an Optical Instrument. In K. Boff, L. Kaufman, and J. Thomas (Ed.), *Handbook of Perception and Human Performance*, New York: Wiley and Sons.

[Wilson91] Wilson, H.R. (1991). Psychophysical Models of Spatial Vision and Hyperacuity. in D. Regan (Ed.) *Spatial Vision*, Vol. 10, Vision and Visual Dysfunction. Boca Raton, FL, CRC Press, 64-81.

[Wyszecki82] Wyszecki G., and Stiles W.S. (1982) Color Science: Concepts and Methods, Quantitative Data and Formulae (2nd edition). New York: Wiley.