Multiscale model of Adaptation, Spatial Vision and Color Appearance

Sumanta N. Pattanaik[1]   **Mark D. Fairchild**[2]   **James A. Ferwerda**[1]
Donald P. Greenberg[1]
[1]*Program of Computer Graphics, Cornell University, Ithaca, NY-14853, USA*
[2]*Center for Imaging Science, RIT, Rochester, NY 14623-5604, USA.*

Abstract

In this paper we present a multiscale color appearance model which simulates luminance, pattern and color processing of the human visual system to accurately predict the color appearance attributes of spectral stimuli in complex surroundings under a wide range of illumination and viewing conditions.

1. Introduction

The aim of a color appearance model is to predict various visual phenomena which simple tristimulus colorimetry can not adequately describe. Colorimetry simply predicts whether two visual stimuli of different spectral power distributions will match in color when viewed under identical visual conditions. This matching is defined by the spectral responsivities of the photoreceptors in the visual system. If the signals from the three cone types are equal for two visual stimuli, then they match in color when seen in the same conditions.

The amplitude of visual stimuli that we encounter in natural scenes is vast. While the responsive range of the visual photoreceptors is small, the visual system functions over this vast range with reasonable ease. It is believed that visual system adapts to the widely varying amplitudes of visual stimulus by *adaptive gain control*. This mechanism controls the relationship between the photoreceptor signals and the amplitude of spectral stimulus, by turning down the gain when the stimulus amplitude is high and by turning up the gain when the stimulus amplitude is low. The gain control is independent for the different type of photoreceptors and hence explains the visual system's capability to adjust to both wide ranges of illumination and varying colors of illumination in order to approximately preserve the appearance of the object colors. The non-linear behavior of these mechanisms results in the increase in colorfulness (Hunt effect) and increase in apparent contrast (Stevens effect) with increasing illumination. Most of the color appearance models [9, 7] available today incorporate some form of adaptive gain control.

It is also known that the appearance of a visual stimulus depends not only on the stimulus itself but also on the other stimuli that are nearby in space. The spatial configuration of the viewing field seems to play a critical importance in the perceived appearance of a stimulus. *Induction, crispening, and spreading* are three important and easy to observe appearance phenomena that are directly related to the spatial surrounding of the stimulus. To account for these effects many color appearance models require a viewing field specification. In this specification the viewing field is divided into as many as four components [7]: stimulus, proximal field, background and surround. The effects of stimuli in these various components of the viewing field are incorporated into the computational model. However, all of these models assume homogeneous stimulus field in each of the components and hence are inadequate for use in complex scenes.

Physiological and psychophysical evidences indicate that the photoreceptor response image is filtered by visual mechanisms sensitive to patterns of different scale, and the response characteristics of these mechanisms are bandpass in the spatial frequency domain [14]. Most of the appearance phenomena discussed in the preceding paragraph can be explained as consequences of this multiscale processing in the visual system. As in the adaptive gain control mechanism, nonlinearities in the responses of the bandpass mechanisms result in nonlinear variations in the appearance of a stimulus as a function of the surrounding stimuli.

In the vision community, many researchers [5, 10] have proposed and successfully applied multiscale visual models for predicting visibility, masking and other related phenomena. However, most of these models assume achromatic stimuli within a limited luminance range and hence make very little prediction of color appearance. In this paper we introduce a new multiscale visual model that not only accounts for the effects of adaptation and spatial vision, but also correctly predicts color appearance attributes under a wide variety of conditions in complex scenes. In the following sections we present this model.

Figure 1: *Flow chart of the computational model of Color Appearance.*

2. The Computational Model

Figure 1 provides a flow chart of each major step in the computational model. The model processes an input image to encode the perceived contrasts for the chromatic and achromatic channels in their band-pass mechanisms. Correlates of brightness, lightness, colorfulness, chroma, hue and saturation are derived from the encoded visual representation.

2.1. Input Image Preprocessing

First, the image is spatially sampled such that the band pass signals represent appropriate spatial frequencies. Then compensations are introduced for optical point-spread and scattering in the eye. The image is then spectrally sampled to represent the visual system's initial photoreceptor responses.

2.2. Spatial Decomposition

The four images representing the photoreceptor responses are then subjected to spatial decomposition. We chose to use the Laplacian pyramid approach proposed by Burt and Adelson [3]. We first calculated a seven level Gaussian pyramid using a five tap filter. Each level of this Gaussian pyramid represents a low-pass image limited to spatial frequencies half of those of the next higher level.

The Gaussian pyramid is then upsampled such that image in each level in the upsampled pyramid is a low-pass version of the corresponding image in the original pyramid. The upsampled pyramid has six levels. Difference-of-Gaussian pyramid is then calculated by subtracting the upsampled pyramid from the original pyramid. 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 (7th) low pass image of the original pyramid is retained for separate processing.

2.3. Gain Control

The difference-of-Gaussian pyramid is then converted to adapted contrast signals using a luminance gain control. The gains are set using TVI-like functions shown in Figure 2. The functions shown have sub-Weber's law behavior [4] which allows perceived contrast to increase with luminance level. Each pixel in a given difference-of-Gaussian image is multiplied by the gain derived from the corresponding pixel in the upsampled pyramid.

The resulting adapted contrast pyramid images are analogous to the contrast images that Peli [12], Lubin [10] and Brill [2] obtained. However, in our model the magnitude
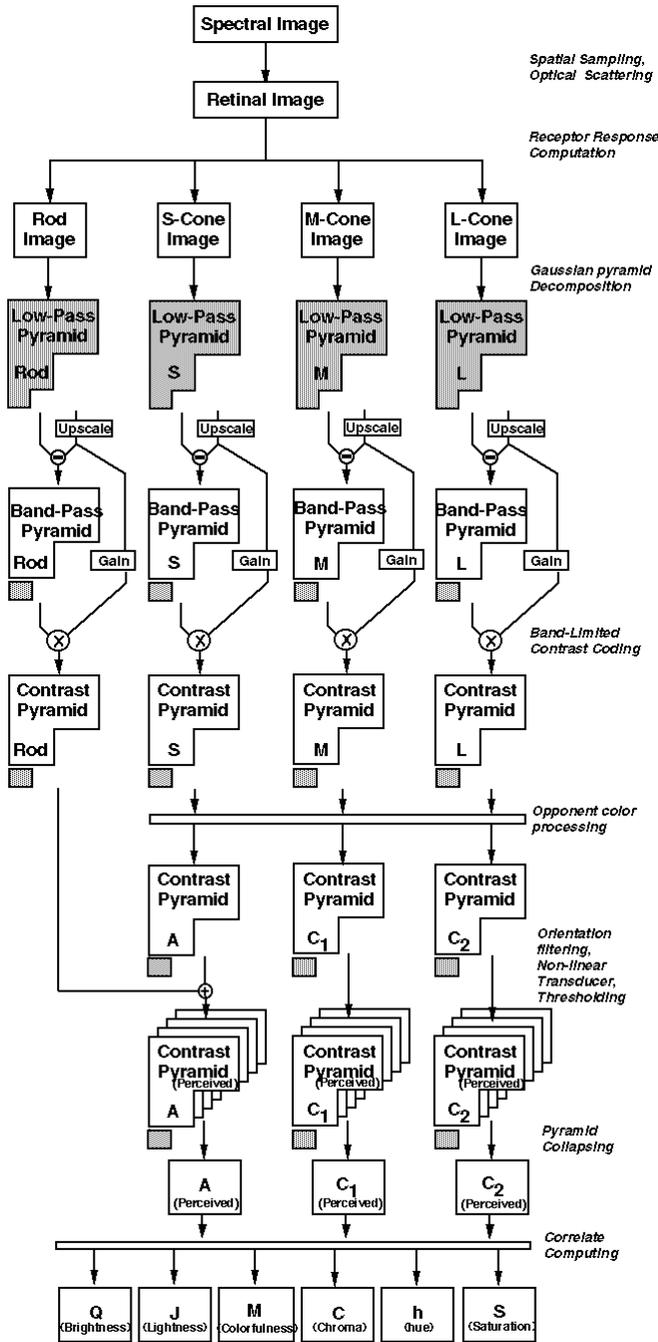
Figure 2: *Gain functions for cones and rods.*

Figure 3: *Contrast transducer functions. (a) for Cone achromatic mechanisms, (b) for Cone chromatic mechanisms; and (c) for Rod achromatic mechanisms.*

of these images is a function of the luminance level specified by the gain control functions. This is necessary to allow prediction of luminance-dependent appearance effects. These luminance gain controls are applied in the same manner to the difference-of-Gaussian pyramid for each of the photoreceptors. This allows prediction of chromatic adaptation effects.

2.4. Opponent Color Processing

In the next stage of the model the adapted contrast signals for the cones are transformed into opponent signals. This transformation is necessary to model differences in the spatial processing of achromatic and chromatic signals [13]. At this stage, the rod images are retained separately since their spatial processing attributes are different from the cones.

2.5. Orientation Filtering, Contrast Transducers And Thresholding

The adapted contrast signals are then processed by oriented band pass filters to simulate the orientation tuning of the visual system. These filtered signals are then passed through contrast transducer functions. Different transducer functions are applied to each spatial frequency mechanism in order to model the human spatial contrast sensitivity functions. 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 is processed through a distinct set of transducers to represent its unique spatial characteristics. At high contrast levels ($>$ 5%) the transducer functions converge to a common square-root form to properly represent perceived contrast constancy [1] and introduce a compressive non-linearity typically found in masking experiments and color appearance models. The contrast transducers used in our
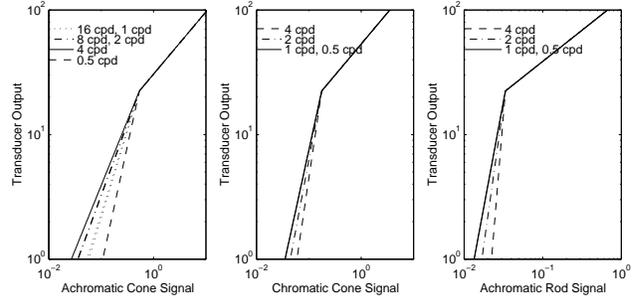
model are illustrated in Figure 3. The contrast transducer functions are also designed such that contrasts that are below threshold have an output level less than 1.0. In the output of the transducer functions all values less than 1.0 are set to 0.0.

2.6. Combination Of Rod And Cone Signals

After the contrast transducers, the rod and cone signals are combined (weighted combination). We assume that the rods contribute only to the luminance signal and thus combine the achromatic signal from the cones with the rod signal.

At this stage in the model we have three channels representing achromatic, red-green, and yellow-blue apparent contrast for the oriented band-pass mechanisms. These signals model threshold behavior, in that any contrast signals that could not be perceived have been eliminated by the contrast transducer functions. They also model suprathreshold appearance since the contrast signals grow with luminance and the signals from chromatic channels become zero at luminance levels below the cone threshold.

2.7. Treatment Of The Low Pass Image

The lowest level low-pass image from the original Gaussian pyramid is also processed through a gain control mechanism similar to the gain control of the band-pass images and a low-pass specific non-linear transducer.

2.8. Computation Of Correlates Of Color Appearance Attributes

The output of the model consists of appearance signals in an achromatic and two chromatic channels and six spatial band-pass mechanisms plus a low-pass image. Images are reconstructed from these signals to create a color-appearance map that encodes the apparent color of each pixel in the image for its particular viewing conditions.
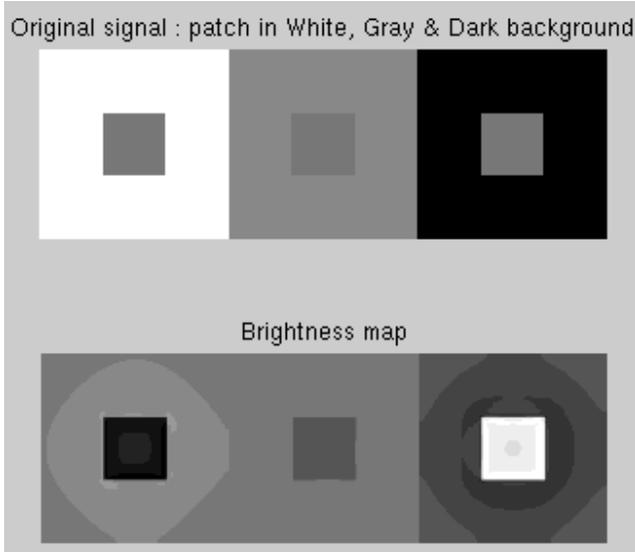
Figure 4: Prediction of induction effect.

The correlates of the appearance attributes are computed from this color appearance map. Difference metrics in these appearance dimensions can be used to derive image quality metrics.

Figure 5: Prediction of crispening effect.

3. Predictions Of The Model

3.1. Induction, Crispening, Spreading And Stevens Effect

Figure 4 predicts the induction effect. The top images show a gray patch on white, gray and dark background. The bottom images show the brightness maps of those images obtained from our model. The maps indicate that the model correctly predicts the change in brightness of the gray patch as a function of background luminance.

Figure 5 predicts the crispening effect. Crispening is the apparent increase in the magnitude of color difference when background on which two stimuli are compared are similar to the stimuli themselves[7]. The top images correspond to a pair of gray patches on three different background. The bottom images illustrate the brightness map. The images correctly predict larger brightness differences (crispening) for the gray patches on the gray surround as compared to the gray patches on white and dark surrounds.

The images in Figure 6 predict contrast changes at a wide range of luminance levels spanning 9 orders of magnitude from 0.001 to $100,000\,\text{cd/m}^2$. The images are brightness maps of a simple scene containing 4 patches of varying reflectances (10%, 30%, 70% and 90%) on a background of uniform reflectance ($50\,\%$). As can be seen from
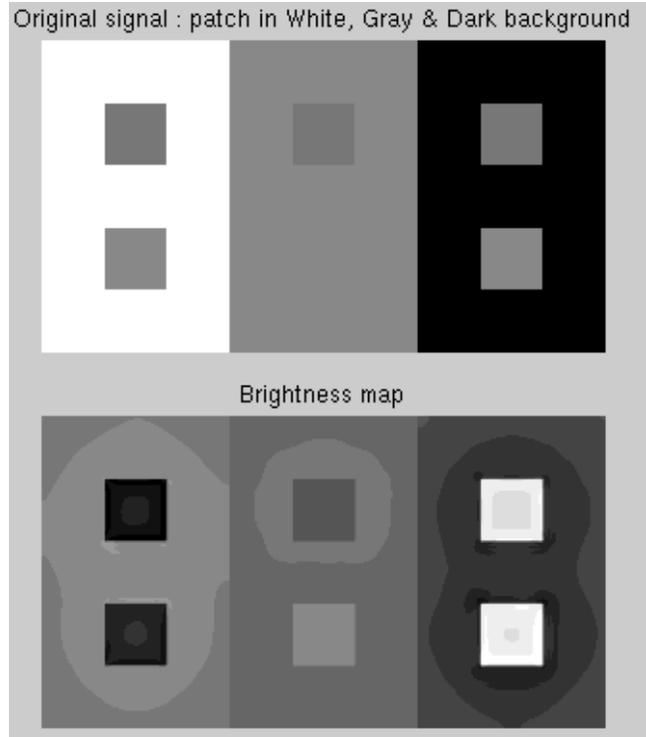
the Figure 6, contrast increases with the increase in luminance.

Figure 7 predicts the spreading effect. Spreading is the apparent mixture of a color stimulus with its surround. The image on the left shows the stimuli input to the model. The spread of color is well captured in the hue map shown on the right.

3.2. Chromatic Adaptation

Figure 8 shows the effect of chromatic adaptation. The top row of images shows a scene illuminated by a nearly-white incandescent light source, a very reddish 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. The middle row shows the rendering from a tone mapping system [11] based on the visual processing carried out in our model. As 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
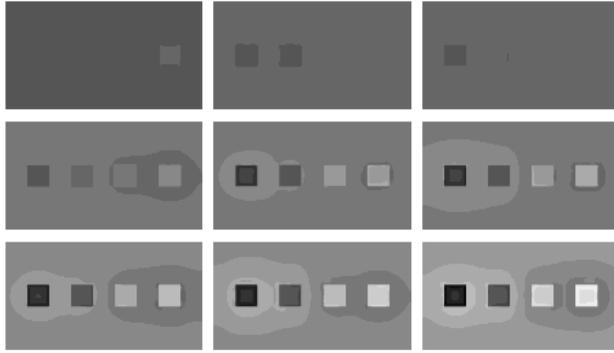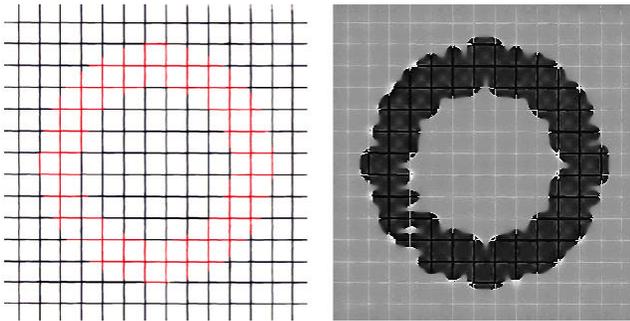
Figure 6: Prediction of Stevens effect.

Figure 7: Prediction of spreading effect.

Figure 8: Illustration of chromatic adaptation.

adaptation predicted by the model is less than 100% complete. The last row of images illustrate the surround effect on the output of the model. The chromatic adaptation is much less in this case because of the gray surround. These images simulate the yellowish appearance of an illuminated window at dusk or a bluish CRT display viewed at a distance. 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. [7]

In a recent experiment [8] corresponding colors data were collected using complex images and comparisons between prints under an illuminant D50 simulator and CRT displays with both illuminant D50 and D65 white points. The results were used to compare the performance of various chromatic adaptation transforms and color appearance models. The multiscale adaptation model described in this paper performed as well as the best models (including CIELAB, von Kries and modified forms of RLAB, ZLAB, and CIECAM97s) and significantly better than other models.

3.3. High Dynamic Range Image Reproduction

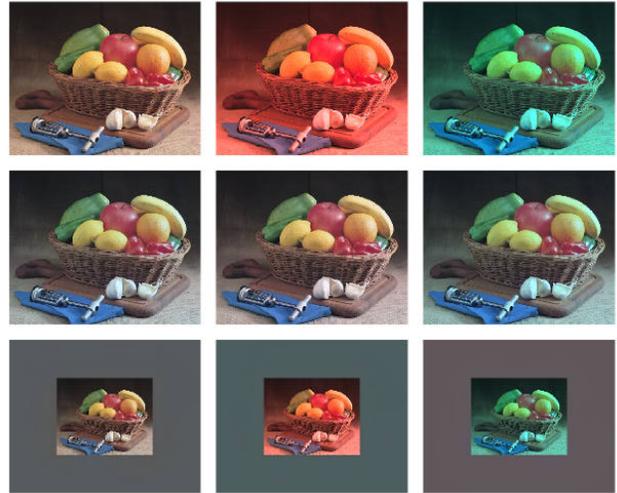Figure 9 illustrates application of the model to the tone mapping of high-dynamic range images. The image on

the top of Figure 9 is linear mapping of the original high-dynamic range image into the limited dynamic range of the output device. The original image had a luminance level of approximately 10,000 cd/m^2 in the outside areas and 10 cd/m^2 in the indoor areas. The image on the bottom represents the mapping obtained by inverting the visual representation of the image derived by our model for the viewing conditions of the output display. In Figure 9 it is clear that far more detail can be observed both inside and outside the parking garage when the image is mapped using the visual model.

4. Future Work

We have calibrated our model to correctly predict psychophysical measurement data available in the literature. We plan to further validate the predictions of the model against perceptual experiments involving color images. Our ultimate aim is to use the predictions of this model to develop an *image quality metric* to verify the fidelity in cross-media color reproduction and the perceived quality of images.

References

[1] Brady, N., Field, D.J. (1995). What's Constant in Contrast Constancy ? The effect of scaling on the Perceived Contrast of Bandpass Patterns. Vision Research, Vol 35(6), 739-756.

[2] Brill, M. H. (1997). Color Management: New roles for color transforms. Proceedings of 5th CIC Conference, pp. 78-82.

Figure 9: Application of the model to the tone mapping of high-dynamic range images. Image of a parking lot constructed from successive multiple photographic exposures [6].

[8] Fairchild, M.D. and Johnson G. (1998). Color Appearance Reproduction: Visual Data and Predictive Modeling. Draft Manuscript, not yet submitted.

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

[10] Lubin, J. (1995). A visual discrimination model for imaging system design and evaluation. In *Vision Models for Target Detection* ed. E. Peli, World Scientific, Singapore, pp. 245-283.

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

[12] Peli, E. (1990) Contrast in complex images. J. Optical Society of America (A), 7(10), 2032-2040.

[13] Poirson, A.B. and Wandell, B.A. (1993). The appearance of colored patterns: pattern-color separability. Journal of Optical Society of America (A), 1993, Vol. 10 (12), pp. 2458-2470.

[14] 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.

[3] Burt, P. J., Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. IEEE Transaction on Communication, Vol. COM-31, No. 4, 532-540.

[4] Chen, B., MacLeod, D.I.A., Stockman, A. (1987). Improvement of human vision under bright light: Grain or gain ? J. Physiology, 394, 41-66.

[5] Daly, S. (1993). The visible differences predictor: an algorithm for assessment of image fidelity. In *Digital Images and Human Vision* ed. A. B. Watson, MIT Press, Cambridge, MA, 179-206.

[6] Debevec, P.E. and Malik, J. (1997) Recovering high dynamic range radiance maps from images. Proceedings SIGGRAPH 97, 369-378.

[7] Fairchild, M.D. (1998) *Color appearance models*. Reading, MA: Addison-Wesley.