On Contrast Sensitivity in an Image Difference Model

Garrett M. Johnson and Mark D. Fairchild
Munsell Color Science Laboratory, Center for Imaging Science
Rochester Institute of Technology, Rochester New York

Abstract

The importance of the human contrast sensitivity function (CSF) in a color image difference metric is examined. The contrast sensitivity function is used to extend traditional color difference equations for use with complex image stimulus. Three CSF models of various complexities are evaluated using psychophysical image difference data. Additionally, two models of spatial frequency adaptation are evaluated. These models adjust the peak sensitivity and general shape of the contrast sensitivity functions to better correlate with complex stimuli.

Introduction

A computational metric of overall perceived image quality seems tantalizingly close, yet somehow remains just out of our grasp. The human visual system, which governs perceived quality, is an extraordinarily complicated system. Image quality must ultimately be a function both of the imaging system itself, as well as the perceptions of the observers who view the images. Adding complexity to this goal, quality cannot be described as being purely physiological, as cognitive or subjective mechanisms such as preference must also be considered.

In order to perceive a change in quality between images, one must first be able to see a difference in the images. If two images appear identical, then obviously they must have the same quality. Thus, the first goal in image quality modeling is the creation of a viable color image difference metric. From the difference metric, one can then develop correlates to image quality scales, often through psychophysical experimentation.

The authors recently presented a framework for a modular color image difference metric.[1] This framework is inspired by the spatial extension to CIELAB, known as S-CIELAB.[2] The underlying bases for these models are the CIE metrics for perceived color difference, of which the state of the art is CIEDE2000.[3] Building on the foundation of S-CIELAB, several modules for a color image difference metric were described.[1] Among these modules were models of contrast sensitivity, local attention, and local and global contrast. It is the first of these modules, that of the contrast sensitivity function, that is examined in more detail here.

Contrast Sensitivity Functions

The level of contrast necessary to elicit a perceived response by the human visual system is known as the contrast threshold. The inverse of the threshold is known as the contrast sensitivity.[4] It is well known that the contrast sensitivity of the visual system varies as a function of spatial frequency. This relationship between contrast sensitivity and spatial frequency is known as the contrast sensitivity function (CSF) of the human visual system. Contrast sensitivity differs for achromatic and chromatic stimuli. It is generally well understood that achromatic contrast sensitivity can be described with a band pass spatial filter peaking around 3-4 cycles per degree of visual angle (cpd). The human visual system is much less sensitive to chromatic contrast at high frequencies, and as such is typically described with low-pass behavior in the isoluminant chromatic channels. The focus in this paper is on the achromatic contrast sensitivity functions.

The contrast sensitivity function, and its inverse the contrast threshold function, describes the contrast levels at a given spatial frequency necessary to elicit a perceptual response (for a given spatial pattern, luminance level, and temporal frequency). For image quality and difference metrics, the CSF is often used to modulate frequencies that are less perceptible. For this reason, the CSF is often erroneously referred to as the modulation transfer function (MTF) of the human visual system. While similar in nature to an MTF, specification of a CSF makes no implicit assumption that the human visual system behaves as a linear system.[5] A better description, as given by Barten[6] and Daly[5], is to refer to the contrast sensitivity function as a threshold modulation function that normalizes all frequencies such that they have equal contrast thresholds.

There are many models of contrast sensitivity that can be used to generate threshold modulation functions for image difference calculations. The spatial preprocessing of the S-CIELAB model uses separable convolution kernels to "blur" out spatial frequencies that cannot be perceived.[2] This technique is computationally efficient, but is not necessarily the most precise method.[1] Three other CSF models, of varying complexity, are tested here.. These include the Movshon[7], Barten[6], and Daly[5] models.

Movshon CSF

The simplest contrast sensitivity function examined here is a three parameter exponential function, first described by Movshon and Kiorpes.[7] The form of the equation is shown below, in Equation 1, where f represents spatial frequency in cycles per degree (CPD).

$$csf(f) = a \cdot f \cdot e^{-b \cdot f} \tag{1}$$

This equation can be fitted to existing luminance contrast sensitivity data, if desired. For general use, the values of 75, 0.2, and 0.8 can used for *a, b,* and *c* respectively. This three-parameter equation forms the general band-pass shape desired for the achromatic contrast sensitivity function. For use in a color image difference metric, the two-dimensional form of the equation is used. The result is an isotropic function.

The relatively simple form of this model is both its strength and weakness. This function is the same for all viewing conditions, unless the parameters are specifically fit to existing data. It is generally assumed that viewing conditions can greatly alter the contrast sensitivity function. This is especially the case for luminance level, which is known to "flatten" the shape of the contrast sensitivity function. To better predict changes in viewing condition, a more complicated function is necessary.

Barten CSF

A more complicated model of the contrast sensitivity function was described by Barten for use in the SQRI image quality model.[6] The contrast sensitivity model begins with the optical MTF of the human eye, which is expressed as a Gaussian function. The MTF is then modified with models of photon and neural noise, and lateral inhibition. The result is an isotropic band-pass shape that is a function of luminance level, pupil diameter, and image size. While more complicated in design than the Movshon model, it possesses greater flexibility for varying viewing conditions. The text by Barten describes in great detail the development and form of this CSF.[6]

Daly CSF

Daly described another form of CSF for use in the Visible Difference Predictor.[5] This model is a function of many parameters, including radial spatial frequency, orientation, luminance levels, image size, image eccentricity, and viewing distance. The result is an anisotropic band-pass function that represents greater sensitivity to horizontal and vertical spatial frequencies than to diagonal frequencies. This corresponds well with the known behavior of the human visual system (the oblique effect).

DC Normalization

The band-pass-function nature of the achromatic CSF leads to interesting problems when used in conjunction with traditional CIE color difference formulas. Existing formula, such as CIEDE2000, do an admirable job of predicting the perceived color difference of uniform patches.

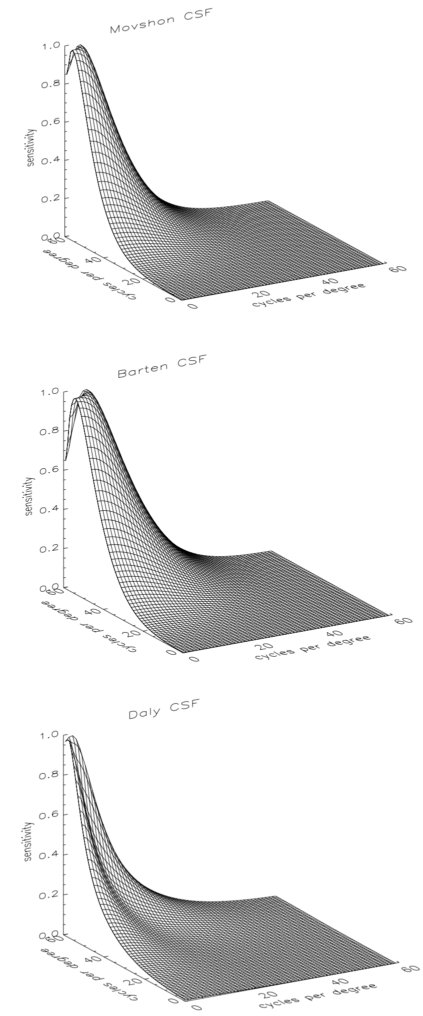

Figure 10. Two-Dimensional Representation of Contrast Sensitivity Functions.

Ideally an image difference metric would reduce to a traditional color difference when uniform patches are used as input stimuli. In order to do this, it is crucial to maintain the integrity of the DC component of the image. Normalizing the CSF so that the DC component is 1.0, results in a CSF that is greater than 1.0 at certain frequencies. This is analogous to a gain increase for certain frequencies, and a modulation for others. This technique was previously shown to produce desirable results in prediction image differences.[1] Figure 2 shows one-dimensional slices of the Movshon, Barten, and Daly CSFs normalized to maintain DC integrity, at a luminance level of 150 cd/m^2 and a viewing distance of 1.5 meters. The Barten CSF modulates the DC far more than the others, resulting in a larger gain for certain frequencies.

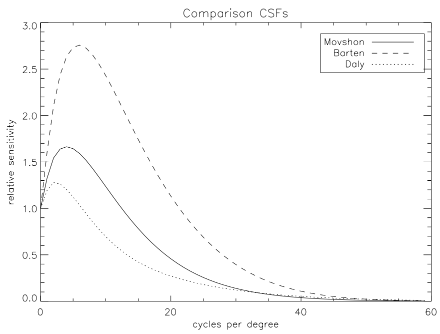

Figure 2. Comparison of DC maintaining Contrast Sensitivity Functions

Color Image Difference Metric Workflow

The above contrast sensitivity functions were then added into the modular color image difference framework. The general workflow of this model is as follows:[1]

- *Convert image pairs into device independent coordinates such as CIE XYZ or LMS cone responses through careful characterization of image system and viewing conditions.*
- *Transform to opponent-color channels, Achromatic, red-green, yellow-blue.*
- *Fourier transform of opponent channels to obtain frequency representation.*
- *Filter frequency representation by multiplying with CSFs*
- *Inverse transform back into opponent space, and then into CIE XYZ coordinates.*
- *Calculate CIELAB values for each pixel.*
- *Calculate a pixel-by-pixel CIEDE2000 color difference to create an error image.*
- *Use image statistics such as mean, std. dev, median to obtain a statistical error measurement.*

Experimental Comparisons

Experimental validation is necessary to gain an understanding on the importance of the contrast sensitivity function for use in a color image difference metric. This experimentation also is useful for revealing the differences between the three CSF models described above. The prediction of each of these models was compared with data generated from a large scale psychophysical experiment examining perceived image sharpness.[8] This experiment resulted in an interval scale of perceived sharpness differences between original images, and 71 manipulations of the images. Among these manipulations were changes in pixel resolution, image contrast, additive noise, and spatial sharpening. The adapting luminance was taken to be 150 cd/m^2, with a viewing distance of 1.5 meters. The resolution of the display was 90 cycles per degree of visual angle. The color image difference metric was then used to calculate a perceived difference between the original and manipulated images. The relationship between the predicted error, and the experimental interval scale is then examined. Ideally, the image difference metric would be highly correlated with the experimental results.

The differences between the three contrast sensitivity functions can be examined by plotting the predicted results of each function against the other functions. Figure 3 illustrates the relationship between the model predictions of the three different CSFs. The predicted values plotted are the mean CIEDE2000 errors between the original image, and the 71 manipulations, for a single scene.

Figure 3 shows the close relationship the three contrast

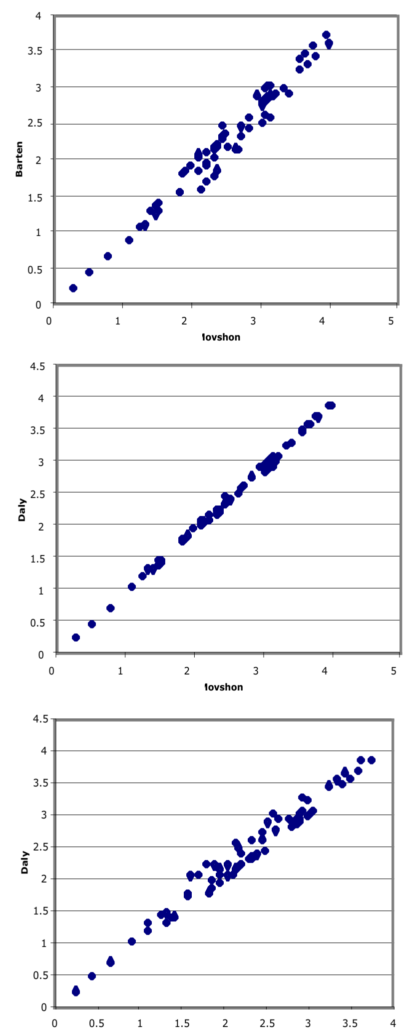

Figure 3. Relationship between mean CIEDE2000 experimental predictions of three contrast sensitivity functions. The top panel illustrates the relationship between the Movshon and Barten CSF, the middle between the Movshon and Daly, and the Lower between the Daly and Barten.

sensitivity functions have with each other. The Movshon and Daly functions are very close, for this particular viewing condition, as seen by the near linear relationship between mean CIEDE2000 error predictions in the middle panel.

The predicted mean CIEDE2000 color difference using each of the CSFs can be plotted against the

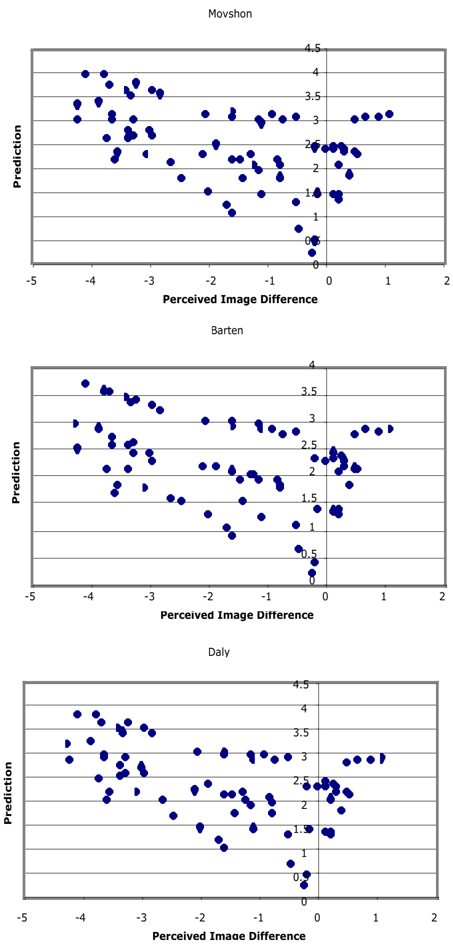

Figure 4. Comparison between model predictions and experiment difference scale. The top panel is the Movshon model, middle is Barten, Lower is Daly.

experimental interval scale. This is shown in Figure 4.

There are several interesting points revealed in Figure 4. All the contrast sensitivity functions are able to predict the general "V" shaped trend (which is the desired result), centered about the origin. The Daly and Movshon functions have very similar plots, as illustrated by the top and bottom panels. The Barten function produced a greater spread, and also introduced a distinct separation. The main difference between the functions, as shown in Figure 2, is the location of the peak, and the amount of gain resulting from the DC normalization.

Spatial Frequency Adaptation

There are several techniques for measuring the contrast sensitivity function of a human observer. Most often it is through the use of a simple grating patern.[4] This pattern is usually flashed temporally, to prevent the observer from adapting to the spatial frequencies being tested. Spatial frequency adaptation, similar to chromatic adaptation, results in a decrease in sensitivity based on the adapting frequency.

While spatial-frequency adaptation is not desired when measuring the fundamental nature of the human visual system, it is a fact of life in real world imaging situations. Models of spatial-frequency adaptation that alter the nature of the contrast sensitivity function might be better suited for use with complex image stimuli than those designed to predict simple gratings. Two such models of spatial frequency adaptation are presented below.

Natural Scene Assumption

It is often assumed that the occurrence of any given frequency in the "natural world" is inversely proportional to the frequency itself. This is known as the *1/f* approximation. If this assumption is held to be true, then the contrast sensitivity function for natural scenes should be decreased more for the lower frequencies, and less for higher. Equation 2 shows a simple "von Kries" type of adaptation based on this natural world assumption.

$$csf_{adapt}(f) = \frac{csf(f)}{1/f} = f \cdot csf(f) \qquad (2)$$

In practice this type of spatial adaptation modulates the low frequencies too much, and places too high an

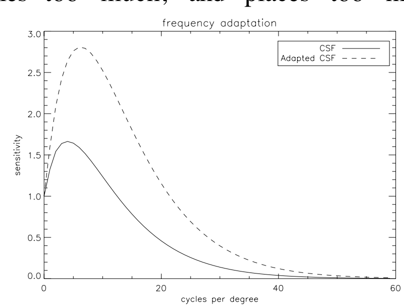

Figure 5. Effect of spatial frequency adaptation on the contrast sensitivity

emphasis on the higher frequencies. A nonlinear compressive relationship with frequency is better suited for imaging applications. This is shown in Equation 3.

$$csf_{adapt}(f) = \frac{csf(f)}{\left(1/f\right)^{1/3}} = f^{1/3} \cdot csf(f) \qquad (3)$$

The effect of this frequency adaptation is an overall shifting of the peak of the contrast sensitivity, as well as an increase in the gain of certain frequencies. Figure 5 illustrates the effect of this spatial frequency adaptation on a contrast sensitivity function.

This type of adaptation model is applied independently of the contrast sensitivity function. The three CSFs from above were modified using Equation 3, and used to predict the experimental results. These predictions are shown in Figure 6.

When compared with Figure 4, these plots show considerably tighter grouping. This indicates that spatial frequency adaptation might play a key role in viewing complex image stimuli.

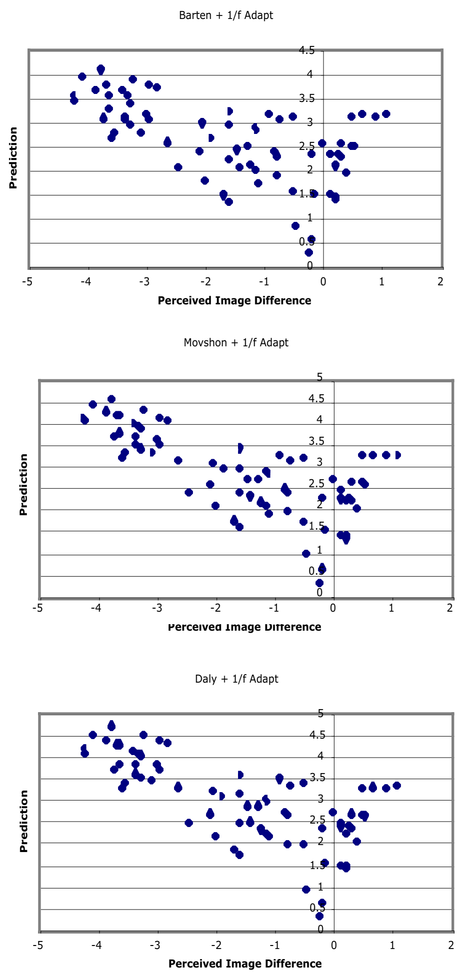

Figure 6. Experimental predictions for three spatially adapted contrast sensitivity functions

Image Dependent Spatial Adaptation

A more complicated approach to spatial frequency adaptation involves adapting to the frequencies present in the image itself, rather than making assumptions about the natural world. The general form of this type of adaptation is shown if Equation 4. In this case, the contrast sensitivity is modulated by the percentage of occurrence any given frequency in an image. This frequency of occurrence is similar to a frequency histogram.

$$csf_{adapt}(f) = \frac{csf(f)}{histogram(f)} \tag{4}$$

Again, this idea is more difficult in practice. For most images, the DC component represents the majority of the frequencies present. This tends to overwhelm all other frequencies, and results in complete modulation of the DC component, and gross exaggeration of the very high frequencies. One way around this problem is to clip the percentage of the DC component to a maximum 10 percent contribution. This is illustrated in Equation 5.

$$histogram(f) \approx fft(image) < 10\% \tag{5}$$

The resulting function is still incredibly noisy, and prone to error as some frequencies have very small contributions. These small contributions correspond to very large increases in contrast sensitivity when normalized using Equation 4. This problem can be eliminated by smoothing the entire range of frequencies using a statistical filter, such as a Lee filter. These filters compute statistical expected values based on a local neighborhood, and are specifically designed to eliminate noise.

The results from the image dependent spatial-frequency adaptation show significant promise. The plots in Figure 7 show a much tighter grouping, when compared to the plots in Figures 4 and 6. The Daly model seemed to benefit the greatest from the image dependent adaptation, as shown from the narrow spread in the bottom panel of Figure 7. The previous model predictions from the Daly and Movshon model were very close, as illustrated in Figure 2. The difference between to two models at the specific viewing conditions used for this experiment essentially reduces to the anisotropic nature of the Daly model. This reduction in sensitivity in the diagonal frequencies corresponds well with the experimental perceived image differences. Perhaps this is because the digital images used in the experiment also have less frequency information in the diagonal directions. This would cause those frequencies to be less adapted, and in turn cause an increase in sensitivity for the isotropic contrast sensitivity functions.

Conclusions

The human contrast sensitivity function (CSF) describes the contrast levels at a given spatial frequency necessary to elicit a perceptual response, for a given spatial pattern, luminance level, and temporal frequency. Three CSF models of varying complexity were tested for use in a color image difference metric. These functions are used as a pre-processing filter in conjunction with traditional CIE color difference formulas, such as CIEDE2000.

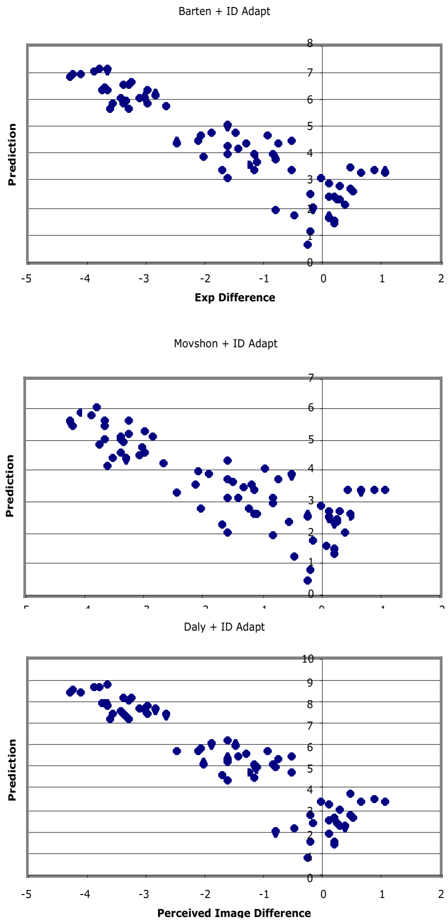

Figure 7. Experimental predictions for three contrast sensitivity functions using image dependent spatial frequency adaptation.

The three CSF models were used to predict experimental image difference scales. The models predicted similar results, despite their varying complexities. It should be noted that these results were obtained with a single viewing condition. Two of the models change behavior based on the viewing conditions, so it is possible that differences will emerge. Future research should examine the behavior across a wider range of conditions.

The CSFs were then adjusted using two models of spatial frequency adaptation. A simple model using an average scene assumption was demonstrated to improve experimental prediction for all contrast sensitivity functions. An image dependent spatial frequency-adaptation model further improved predicted results. The Daly CSF showed the most significant improvement when used with the image dependent adaptation model. This is most likely due to the anisotropic nature of the Daly CSF, which behaves similarly to the human visual system.

Acknowledgements

The authors thank Kazuhiko Takemura and Fuji Photo Film for their support of this research work.

References

1. G.M. Johnson and M.D. Fairchild, "Darwinism of Color Image Difference Models," *Proc. of IS&T/SID 9th Color Imaging Conference*, (2001).
2. X.M. Zhang and B.A. Wandell, "A spatial extension to CIELAB for digital color image reproduction," *Proc. of the SID Symposiums*, (1996).
3. M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 Colour Difference Formula," *Color Research and Applications*, 26 (2000).
4. B.A. Wandell, *Foundations of Vision*, Sinear Associates, Sunderland, MA (1995).
5. S. Daly, The Visible Differences Predictor: An algorithm for the assessment of image fidelity, Ch. 13 in *Digital Images and Human Vision*, A.B. Watson Ed., MIT Press, Cambridge, MA (1993).
6. P.G.J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*, SPIE, Bellingham, WA (1999).
7. T. Movshon and L. Kiorpes, "Analysis of the development of spatial sensitivity in monkey and human infants," *JOSA A*. vol5, (1998).
8. G.M. Johnson and M.D. Fairchild, "Sharpness Rules," *Proc of IS&T 8th Color Imaging Conference*, (2000).

Biography

Garrett M. Johnson is a Color Scientist at the Munsell Color Science Laboratory, as well as a Ph. D. Student in the Center for Imaging Science, at the Rochester Institute of Technology. He has a BS in Imaging Science and a MS in Color Science, both from RIT. His research interests include color image difference and color appearance modelling, image synthesis, and computer graphics. He is a member of ISCC, IS&T, and ACM-SIGGRAPH.

Mark D. Fairchild is Director of the Munsell Color Science Laboratory and Professor of Color Science and Imaging Science at Rochester Institute of Technology. He received his B.S. and M.S. degrees in Imaging Science from RIT and his Ph.D. in Vision Science from the University of Rochester. Dr. Fairchild is active in ISCC, IS&T, CIE, SID, OSA, and ACM-SIGGRAPH. He served as chair of CIE TC1-34, which formulated the CIECAM97s color appearance model, and is currently a member of several other CIE committees and the CIE reporter on color appearance models. He received the 1995 Bartleson Award from the Color Group (Great Britain) for his research in color appearance and other aspects of color science