Analysis of the retinex theory of color vision

David H. Brainard and Brian A. Wandell

Department of Psychology, Stanford University, Stanford, California 94305

Received August 16, 1985; accepted June 11, 1986

If color appearance is to be a useful feature in identifying an object, then color appearance must remain roughly constant when the object is viewed in different contexts. People maintain approximate color constancy despite variation in the color of nearby objects *and* despite variation in the spectral power distribution of the ambient light. Land's retinex algorithm is a model of human color constancy. We analyze the retinex algorithm and discuss its general properties. We show that the algorithm is too sensitive to changes in the color of nearby objects to serve as an adequate model of human color constancy.

INTRODUCTION

Introduction to Color Constancy

If color appearance is to be a useful feature in identifying an object, then color appearance must remain roughly constant when the object is viewed in different contexts. People maintain object color appearance despite variation in the color of nearby objects *and* despite variation in the spectral power distribution of the ambient light.[1-4]

Historically, changes in the color appearance of an object caused by variation in the surface reflectance functions of surrounding objects have been called simultaneous contrast, whereas changes in color appearance caused by variation in the spectral power distribution of the ambient light have been called failures of color constancy. *Both* of these effects reduce the usefulness of color appearance as a feature for identifying objects. In our view, *color constancy* should be defined as the maintenance of color appearance despite variation in the color of nearby objects *and* despite variation in the spectral power distribution of the ambient light. Although human color vision does not maintain perfect color constancy, human performance is better than that of any currently available man-made systems.

The fundamental difficulty in designing a color constant system arises becase the light in the visual image confounds two factors: the spectral power distribution of the ambient light and the surface reflectance of the objects in the scene. The problem of separating these two factors cannot be solved for all possible viewing conditions. For example, if there is only a single unknown object in the image illuminated by an unknown light source, no algorithm can correctly determine the surface reflectance of the object. It follows that all color constant algorithms must use information obtained from light reflected from several different objects in the scene. We therefore think it natural to broaden the definition of color constancy to include preservation of appearance across variation in the surface reflectance of nearby objects.

A number of algorithms have been proposed for the purpose of modeling human color constancy or to achieve approximate color constancy.[3-7] Only the more recent algorithms[5-7] make explicit the class of viewing contexts for which the algorithm will be color constant.

The retinex algorithm proposed by Land and McCann[4] and Land[8] is important because it was the first attempt at developing a computational model for human color constancy. There is widespread current interest in the retinex algorithm.[8-17] The experiments that lead to the development of the retinex algorithm and the algorithm itself have been widely discussed in the literature.[6,18-23] In spite of this, the retinex algorithm is not generally well understood.

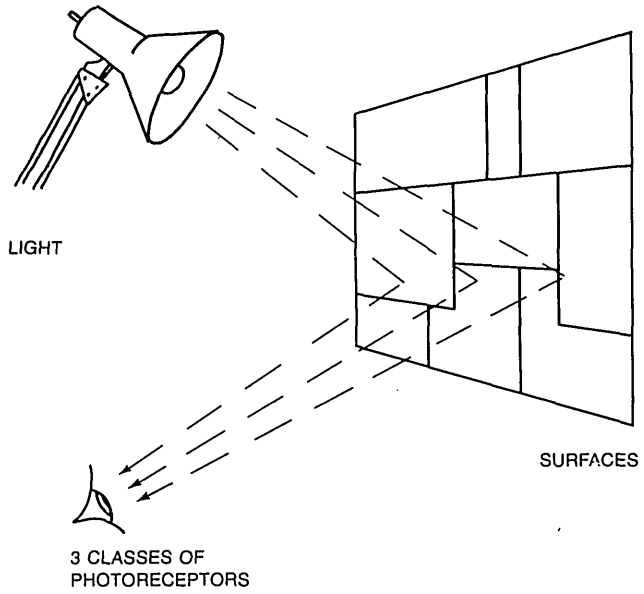
In this paper we quantitatively describe the problem of color constancy. We then describe Land's most recent retinex algorithm.[8,17] We present some simple results that describe the general behavior of the retinex algorithm. We extend our analysis by describing a computer simulation of the algorithm. The simulation demonstrates that the retinex algorithm is a poor model of human color constancy. The dramatic failures of retinex as a model for human vision arise because the assignment of a color value to a surface is far too dependent on the composition of the other surfaces in the image. In Appendix A we discuss other versions of retinex, suggested by McCann and Houston[11] and Horn.[24]

Formulation of Color Constancy

The retinex algorithm has been proposed for and tested on only a limited class of viewing contexts. We will formulate the problem of color constancy with respect to this simplified model of the natural environment. The viewing context is illustrated in Fig. 1. An observer looks at a flat two-dimensional surface. The materials on the surface are matte, and they reflect the ambient light toward a normal color observer, who has three classes of photoreceptors. In our formulation we describe the values of all functions of wavelength by their values at a discrete number of sample wavelengths λ_n for $n = 1, N$. In particular, we characterize the spectral power distribution of the ambient light by the function $E(\lambda_n)$ and the reflectance at a point x on the surface by the function $S^x(\lambda_n)$. The light arriving at the eye is called the *color signal*. It is equal to the product of the spectral power distribution of the ambient light and the surface reflectance function, $C^x(\lambda_n) = E(\lambda_n)S^x(\lambda_n)$. The color signal defines all the image information that is available at the eye to make judgments concerning the surface reflectance at different points on the surface.

The observer has three photoreceptor arrays that spatially

0740-3232/86/101651-11$02.00



Mondrian Viewing Context

Fig. 1. A flat surface of colored materials is illuminated by a light source.

sample the color signal. The response of the photoreceptors is computed from the color signal and the spectral sensitivity of the photopigment in the kth receptor class, $R_k(\lambda_n)$:

$$\rho_k{}^x = \sum_{n=1}^{N} C^x(\lambda_n)R_k(\lambda_n)$$

$$= \sum_{n=1}^{N} E(\lambda_n)S^x(\lambda_n)R_k(\lambda_n). \quad (1)$$

This equation can be written as a matrix product of the form

$$\rho^x = \Lambda_E \sigma^x \quad (2)$$

in which the nth entry of the column vector σ^x is the surface reflectance at λ_n, $S^x(\lambda_n)$. The matrix Λ_E is $3 \times N$, and its k, nth entry is $E(\lambda_n)R_k(\lambda_n)$. The entries of the matrix depend only on the spectral power distribution of the ambient light (a physical variable) and the receptor spectral sensitivities (a fixed value). We emphasize the fact that the matrix depends only on the spectral power distribution of the ambient light by calling the matrix the *lighting matrix*.

The problem of color constancy can be expressed as the following challenge: Beginning with the spatial array of data on the left-hand side of Eq. (2), recover the spatial array of surface reflectance functions on the right-hand side of the equation. An algorithm with perfect color constancy uses the photoreceptor values on the left-hand side to calculate color estimates at each pixel that are independent of *both* the spectral power distribution of the ambient light (i.e., the lighting matrix Λ_E) and the other surfaces comprising the scene (i.e., the vectors σ^x).

The class of viewing contexts used to evaluate the retinex is limited in various ways. The spatial distribution of the surface reflectances on the two-dimensional surface is restricted to be composed of an array of rectangular, matte, colored papers. Land refers to this spatial distribution of surfaces as a Mondrian. This means that there is generally a clear signal in the image data at each of the boundaries between the differently colored papers. Other limitations of this viewing context are that there are no secondary reflections among surfaces, there is no specular reflectance, and the surfaces are restricted to lie in a plane.

In this paper, we study the performance of the retinex algorithm for the simple case in which the intensity of the light is spatially uniform across the surfaces. This means that it is not necessary to discount spatial variation in the intensity of the ambient light. The problem of discounting such spatial variation across a single scene is logically distinct from the problem of discounting variation in the spectral power distribution of the ambient light from one scene to another. We discuss this point more fully in Appendix A in connection with our analysis of Ref. 24.

THE RETINEX ALGORITHM

There are several published variants of the retinex theory.[4,8,11,25] Here we consider the algorithm first described in Ref. 8. All the published versions of retinex share basic underlying principles, and in Appendix A we analyze the McCann–Houston algorithm[11] and an algorithm presented by Horn.[24]

The input into the retinex algorithm is an array of photoreceptor responses for each location in the image. We can think of this input as being three separate arrays of data, one for each receptor class. Each of these spatial arrays contains the receptor responses for a single receptor class for each location in the image. We call each separate spatial location in the image a *pixel* and use the superscript x to denote a particular pixel in the spatial array. The algorithm transforms the spatial array of photoreceptor responses in the kth class, $\rho_k{}^x$, into a corresponding spatial array of values, $l_k{}^x$. Land calls the individual values, $l_k{}^x$, *lightness values*. At each pixel, one lightness value is computed for each photoreceptor class. A central principle of the retinex algorithm is that the lightness values at pixel x are calculated independently for each receptor class. Thus we need describe the computation of only one of the spatial arrays of lightness values. This permits us to supress the subscript k that indicates the photoreceptor class.

The Lightness Computation

The algorithm estimates the spatial array of lightness values for a single receptor class by computing a series of *paths*. Each path is computed as follows. We select a starting pixel x_1. We then randomly select a neighboring pixel x_2. We calculate the difference of the logarithms of the sensor responses at the two positions. This value is added into an accumulator register for position x_2 such that

$$A(x_2) \leftarrow A(x_2) + \log(\rho^{x_2}) - \log(\rho^{x_1}). \quad (3)$$

In addition, a counter register $N(x_2)$ for position x_2 is incremented to indicate that a path has crossed this position. At the start of the computation all accumulators and counters are initialized to 0.

The path calculation proceeds iteratively with the random selection of a neighbor of pixel x_2. In general, the accumulator of position x_i on this path is updated by

$$A(x_i) \leftarrow A(x_i) + \log(\rho^{x_i}) - \log(\rho^{x_1}), \tag{4}$$

and the corresponding counter register $N(x_i)$ is incremented. Note that the sensor response of the first element of the path plays a special role in the accumulation for that path calculation: It is used as a normalizing term at every point on the path. The path starting at x_1 need not pass through all the positions in the image, and it may pass through some positions more than once. The number of positions traversed by a path is a parameter N_{pl} of the algorithm. (The subscript pl stands for path length.)

After the first path has been computed, the procedure is repeated for a new path that starts at another randomly chosen position. The number of paths in a complete calculation of the algorithm followed is also a parameter, N_{np}. (The subscript np stands for number of paths.)

After all paths have been completed, the lightness value for each pixel x is computed by simply dividing the accumulated values in $A(x)$ by the contents of the corresponding counter register, $N(x)$.

The full retinex algorithm described in Ref. 8 includes an additional thresholding operation that occurs during the path calculation. The thresholding operation compares the sensor response at adjacent pixels along the path. If the value at pixel x_{i+1} is sufficiently close to the sensor response at the previous pixel, x_i, then the accumulator value at x_{i+1} is updated by using the sensor response at x_i. The intention of the thresholding operation is to "remove the effects of nonuniform illumination over the scene" (Ref. 8, p. 5165). In this paper we will analyze retinex only for images with spatially uniform ambient light distributions. Thus the thresholding operation has no effect on any of our calculations, and we will not address this aspect of the algorithm (but see Ref. 23).

The Meaning of the Lightness Values

The purpose of the retinex algorithm is to compute lightness values that will be invariant under changes of viewing context, much as human performance is roughly invariant under similar changes. At each pixel x, the lightness triplet should depend only on the surface reflectance $S^x(\lambda_n)$ and not on the spectral power distribution of the ambient light or the surface reflectance functions of the other papers comprising the Mondrian. In the following sections we investigate both the extent to which the lightness values are color constant and whether the retinex algorithm in fact models human vision.

ANALYSIS OF THE RETINEX ALGORITHM

The retinex algorithm is stochastic, as both the starting pixels of the paths and the path routes are chosen randomly. To understand the retinex algorithm, we examine its convergence properties.

Consider the kth receptor class. The lightness value $l_k{}^x$ for pixel x is the average of the contributions to $A(x)$ from the paths that passed through pixel x. Each time a path passes through pixel x, $A(x)$ is incremented by the log of the ratio of the receptor response at x to the receptor response at the starting pixel of the path. Let the total number of times a path has crossed pixel x be N. If we denote the starting points for the paths passing through pixel x as $x_1, \ldots, x_N$, then

$$\begin{aligned} l_k{}^x &= \frac{\log\left(\frac{\rho_k{}^x}{\rho_k{}^{x_1}}\right) + \ldots + \log\left(\frac{\rho_k{}^x}{\rho_k{}^{x_N}}\right)}{N} \\ &= \log\left[\frac{\rho_k{}^x}{(\rho_k{}^{x_1} \ldots \rho_k{}^{x_N})^{1/N}}\right] \\ &= \log\left[\frac{\rho_k{}^x}{\hat{G}_k(x)}\right], \end{aligned} \tag{5}$$

where $\hat{G}_k(x)$ is the geometric mean of the receptor responses of pixels $x_1, \ldots, x_N$ that are the path starting locations corresponding to each crossing of pixel x. Note that $\hat{G}_k(x)$ may depend on the spatial position of pixel x.

If the number of paths N_{np} is small, the random variable $\hat{G}_k(x)$ will have large variance, as will the computed lightness values. This is an undesirable property since for a small number of paths the algorithm will not return stable color estimates when run repeatedly using a fixed input. As the number of paths grows large, $\hat{G}_k(x)$ will converge to a limiting value, $G_k(x)$, and hence the lightness values will be stable. Therefore we may supress the parameter N_{np} and examine the properties of the algorithm under the assumption that N_{np} is large enough so that the lightness values may be computed by their value in the limit as N_{np} grows without bound.

The number of times that a path starting at any pixel x_i crosses pixel x is a random variable, as is the total number of paths N that cross pixel x. Denote the expected number of crossings of pixel x by a path starting at location x_i by $E(x|x_i)$. If the path starting points are selected with equal probability among all pixels, then, as N_{np} tends to infinity, the value of $\hat{G}_k(x)$ is given by the equation

$$\log G_k(x) = \frac{1}{E(N)}\left[\sum_{\text{all pixels}} E(x|x_i)\log(\rho_k{}^{x_i})\right], \tag{6}$$

where $E(N)$ is the expected number of crossings through pixel x, $E(N) = \Sigma_{\text{all pixels}} E(x|x_i)$. In this case, we can compute the lightness values as

$$l_k{}^x = \log\left[\frac{\rho_k{}^x}{G_k(x)}\right], \tag{7}$$

where $G_k(x)$ is given by Eq. (6).

Now we must address the question of how to compute the values of the $E(x|x_i)$ as a function of the parameter N_{pl}. The authors of the retinex algorithm have not provided a precise description of the statistical properties of the path-generation process.[26] We have chosen to model the path-generation process by using a Markov process. On each step the path is equally likely to move from its current position to any of the immediately adjacent pixels. There are eight adjacent pixels for points in the center of the image, five adjacent pixels for points at the edge, and three adjacent pixels at the corners. The choice of a Markov process simplifies the

analysis of the dependence of the algorithm's behavior on the path length. This choice seems consistent with the intention of the authors, though different in detail.[27] When the Markov path-generation procedure is used, the probability that the path is at pixel x on on step $n + 1$ depends only on the path location on step n. We can calculate the probability, $P(n = x|1 = x_i)$, that a path will be at pixel x on the nth step, given that the path started at pixel x_i. For a given path length N_{pl}, the number of times that we expect a path starting at x_i to pass through pixel x is given by

$$E(x|x_i) = \sum_{n=1}^{N_{pl}} P(n = x|1 = x_i).$$

We have computed the $E(x|x_i)$ for a pixel in the sixth row and sixth column of a Mondrian of spatial dimensions 10×10, for several values of N_{pl}. For very small path length ($N_{pl} < 5$), only the neighbors of the pixel contribute to $G_k(x)$. For $N_{pl} = 25$, the neighbors of the pixel contribute six times as frequently as remote pixels. For $N_{pl} = 200$, the neighbors contribute 1.25 times as frequently as remote pixels. When the path length is large ($N_{pl} \gg 200$), the contributions from all pixels are approximately equal. In this case $G_k(x)$ is simply the geometric mean of the receptor responses from all pixels in the image and is independent of x. The lightness of any pixel is given by

$$l_k{}^x = \log\left(\frac{\rho_k{}^x}{G_k}\right), \tag{8}$$

where G_k is the geometric mean of all receptor responses for the kth receptor class. Other authors have noted that the retinex algorithms perform some kind of normalization.[4,17,21,22,25] Equation (8) provides the exact expression for this normalization for the Land[8,17] version of the retinex algorithm in the limiting case of large N_{np} and long N_{pl}. In Appendix A, we provide a similar result for the McCann–Houston[11] version of the retinex algorithm.

In order to evaluate the performance of retinex, we must know the appropriate path length. The published versions of retinex are not explicit about how to choose N_{pl}. Land (Ref. 8, p. 5166) writes that contributions to the lightness of a single point in the image are "taken over areas from the entire visual field and not just those nearby; experiments indicate there may be nearly as much contribution from distant areas as nearby ones." This suggests that Land's implementation uses a rather long path length.

In the next section we investigate the performance of the algorithm in the case when N_{pl} is large enough so that we may calculate the lightness values with Eq. (8). Our investigation will show that in this case the lightness values are too dependent on the composition of surfaces in the Mondrian to predict human color sensation.

In earlier work with another varient of the retinex algorithm, McCann *et al.* (Ref. 25, p. 454) used a path length of 200 with a 480-point image array. With this choice of parameter values, all points in the image will not contribute equally to the lightness value calculated for each point. For shorter path lengths, the lightness at each pixel depends to a larger extent on nearby pixels than remote pixels. We take this point up in more detail in the section headed Discussion.

PERFORMANCE OF RETINEX

Introduction

Even within the framework of the Mondrian viewing context with uniform illumination, no algorithm that uses the responses of a small number of receptor classes can be color constant if the lights and surfaces are allowed to have arbitrary spectral power distributions and surface reflectance functions.[21,28] More recent computational models of color constancy[6,7,28,29] specify the classes of surfaces and lights that are within the model's domain of application. The retinex algorithm—formulated much earlier—has not been updated to include such a specification. It is, however, possible to compute the lightness values for any set of lights and surfaces. In this section we discuss the performance of the retinex algorithm for light spectral power distributions and surface reflectance functions that are frequently encountered in the natural environment.

As we have emphasized in our definition of color constancy, there are two ways in which an algorithm can fail. It can fail to return color constant descriptors when the spectral power distribution of the ambient light is varied and it can fail to be color constant when the composition of the surfaces in the Mondrian is varied. We discuss the retinex algorithm's performance with respect to both of these types of image variation.

Dependence of Color Value on Ambient Light

McCann *et al.*[25] have investigated the performance of an earlier version of retinex algorithm when only the spectral power distribution of the ambient light is varied. They used the mixture of three narrow-bandwidth lights as the illuminant in their experiments. Their results indicate that, although the lightness values are not perfectly constant across illuminations, the failures of the algorithm are not dramatic and are of the same order of magnitude as human failures. We have extended their simulations to include broadband illuminants using the Land[8] algorithm. The results of our simulations, not presented here, are consistent with the conclusions drawn by McCann *et al.*[25] The computed lightness values remained roughly constant as the illuminant was varied.

When the composition of the Mondrian is held constant and when the path length is sufficiently long, the retinex algorithm acts by scaling the separate receptor classes. This method of color correction is similar to one widely used in industry. A gray-card reference surface is inserted into the scene, and each receptor class is scaled by its response to the gray card. We pursue this point in more detail in the section headed Discussion.

Dependence of Color Values on Surface Composition

The dependence of the lightness values computed by the retinex algorithm under changes of the composition of the Mondrian has not been addressed in the literature. In order to explore the performance of the retinex algorithm when the composition of the Mondrian is varied, we simulated the algorithm by using a number of different Mondrians.

Simulation Framework

The input to the simulation of retinex is a list of receptor responses for each pixel. We calculated the receptor re-

R8/4	B8/2	G8/2
10YR5/6	Y8/2	P6/8
B2/2	B3/2	B4/2

Blue Mondrian

R8/4	B8/2	G8/2
10YR5/6	Y8/2	P6/8
G6/6	G7/6	G8/6

Green Mondrian

R8/4	B8/2	G8/2
10YR5/6	Y8/2	P6/8
Y6/8	Y7/8	Y8/8

Yellow Mondrian

R8/4	B8/2	G8/2
10YR5/6	Y8/2	P6/8
P7/4	P6/6	P7/6

Purple Mondrian

Test Mondrian Composition

Fig. 2. The spatial composition of the blue, green, yellow, and purple Mondrians used in the simulations of the Land[8] retinex algorithm.

sponses for a Mondrian composed of matte surfaces taken from the Munsell chips, uniformly illuminated by CIE standard daylight D_{65}. (The Munsell chips are a set of colored papers spanning a broad range of natural colors. The chips are widely used for color measurement in science and industry.) To calculate the receptor responses we used the Smith–Pokorny estimates of the three human cone photoreceptors as reported by Boynton.[30] The receptor responses are given by

$$\rho_k{}^x = \sum_{n=1}^{N} E(\lambda_n) S^x(\lambda_n) R_k(\lambda_n). \quad (9)$$

We used the spectral reflectance data of Nickerson[31] and represented these data as the best-fitting linear combination of Cohen's[32] basis functions for the Munsell chips.[22,28] Judd and Wyszecki[33] described the spectral power distribution of illuminant D_{65}.

From Lightness Triplets to Color Names

We used the following method (see Land[8]) to evaluate the dependence of lightness values on changes in the viewing context. We assigned each Munsell chip to a lightness triplet calculated under standard viewing conditions. The standard viewing conditions consisted of a Mondrian of 462 Munsell color papers illuminated by CIE D_{65}. We will refer to this Mondrian as the standard Mondrian. To evaluate the color appearance predicted by the retinex algorithm of a test chip in some different viewing context, we first comput-

Mondrian	STANDARD	BLUE	GREEN	YELLOW	PURPLE
Predicted Color	R8/4	Y8/2	YR6/2	P6/6	10R6/4
Lightness Triplet	0.9640 0.9476 1.0435	0.9077 0.8250 0.7905	0.3166 0.2310 0.3940	0.2968 0.2824 0.8885	0.3248 0.2904 0.2150

Test Chip: R8/4 Light: D_{65}

Mondrian	STANDARD	BLUE	GREEN	YELLOW	PURPLE
Predicted Color	B8/2	N8/	RP6/2	10PB6/8	R6/2
Lightness Triplet	0.9762 1.0844 1.2631	0.9199 0.9618 1.0102	0.3287 0.3679 0.6137	0.3090 0.4193 1.1081	0.3370 0.4272 0.4346

Test Chip: B8/4 Light: D_{65}

Change of Context Simulation Results

Fig. 3. Predicted color values and lightness triplets computed by the Land[8] retinex algorithm for two chips in the standard, blue, green, yellow, and purple Mondrians. The lightness values were computed according to Eq. (8).

ed the lightness triplet in the new viewing context by using Eq. (8):

$$l_k^x = \log\left(\frac{\rho_k^x}{G_k}\right), \tag{10}$$

where G_k is the geometric mean of receptor responses for the kth receptor class. (This is the limiting lightness value for the algorithm when the number of paths is large and the path length is long.) To find the color value for the test chip, we compare the computed lightness triplet with all the lightness triplets calculated under the standard viewing context. We then find the Munsell chip in the standard viewing context whose lightness value is closest (in a Euclidean metric sense) to the test chip's lightness triplet. The color value of this Munsell chip is used as the predicted color value for the test chip.

The 462 Munsell papers sparsely populate the set of perceptually distinct color values. It is possible, therefore, that our method will assign the same color value to two lightness triplets that produce different color sensations. We can be confident, when the method assigns different color values to two lightness triplets, that the corresponding predicted color sensations are noticeably different.

Simulation Results

To evaluate the dependence of the lightness values on composition of the Mondrian, we computed the lightness triplets for six test chips in each of four Mondrians illuminated by CIE D_{65}. The composition of these Mondrians is shown in Fig. 2. These 3 × 3 Mondrians differ only in the composition of the lowest row: Their upper two rows are identical. We will refer to these four Mondrians as the blue, green, yellow, and purple Mondrians, respectively, as these names indicate the hues of the lower chips. We chose these Mondrians because the variation in the composition of the lower row significantly shifts the geometric mean of the receptor responses when they are illuminated by D_{65}. A human observer viewing these Mondrians on a black background illuminated by daylight perceives virtually no change in the appearance of the upper two rows of chips across the four Mondrians. Figure 3 shows the lightness triplets and corresponding predicted color values for two of the test chips, as computed for each of the Mondrians. Both the computed lightness triplets and the corresponding color values change markedly as a function of the composition of the Mondrian. For example, the retinex algorithm predicts that test chip R8/4, which in fact appears pink in the viewing conditions simulated, will range in appearance from beige (Y8/2) in the blue Mondrian to purple (P6/6) in the yellow Mondrian. This range of predicted color values is typical of all the test chips and is not characteristic of human performance. Although human color vision does exhibit simultaneous contrast effects, these effects are small compared with the magnitude of color shift predicted by the retinex algorithm. We conclude that the retinex algorithm does not correctly predict color appearance when the composition of the Mondrian is varied.

DISCUSSION

The goal of an ideal color constancy algorithm is to use the information in the receptor responses to compute color values that depend only on the surface reflectance spectra. Equation (2),

$$\rho^x = \Lambda_E \sigma^x,$$

expresses the relation between the surface reflectance σ^x and the responses ρ^x. To recover σ^x, we multiply the response vector ρ^x by the pseudoinverse of Λ_E[6,7,28,29]:

$$\sigma^x = \Lambda_E^{-1} \rho^x. \tag{11}$$

The matrix Λ_E^{-1} describes the proper correction for the spectral power distribution of the ambient light. The entries of Λ_E^{-1} are independent of the surfaces in the image. The retinex algorithm always attempts to correct for the illuminant by an equation of the form

$$l^x = \Gamma^x \rho^x, \tag{12}$$

where Γ^x is a diagonal matrix. The entries of Γ^x depend on the spectral power distribution of the ambient light, the position x in the image, *and* the other surfaces in the image. As the path length increases, the algorithm converges toward using a single matrix Γ, for all positions x, whose kth diagonal entry approaches $1/G_k$. Recall that the value G_k is the geometric mean of the receptor responses for the kth receptor class.

The simulation results show that, in the case of long path length, the dependence on the surfaces in the image is too strong. The predicted color values of the retinex algorithm vary too widely with changes in the composition of the Mondrian to serve as an adequate model of human color vision.

For shorter path lengths, the lightness triplet calculated for a pixel x depends on the location of the pixel. In this case, the diagonal entries of the matrix Γ^x are a weighted geometric mean of the responses near location x. While reducing the path length will change the set of points responsible for the overdependence of predicted color names on the surfaces in the Mondrian, the qualitative effect will not change. Rather, the same dependence on changes of surfaces as was exhibited in the long-path-length case will be replayed repeatedly across the image on a local scale. We have not discovered a choice of path length that allows the retinex algorithm to predict color names that are color constant.

Another way to understand the retinex algorithm is to compare it with the following simple color correction procedure. To perform a color correction, we place a reference surface in the image. We divide each of the responses of the kth receptor class by the response of the kth receptor class to the reference surface. (The normalized receptor responses parallel the retinex algorithm's lightness triplets.) This method is currently used in industry for correcting the color output of video cameras. Brill and West[21] (see also Ref. 18) have analyzed this method of color correction. They show that the method performs correctly only for a severely limited class of surface reflectance functions and spectral power distributions of the ambient light. When a reference surface is available, the algorithm presented by Buchsbaum[6] will properly correct for changes in the illuminant for a much broader class of surface reflectance functions and spectral power distributions of the ambient light.

The retinex algorithm is equivalent to normalization with

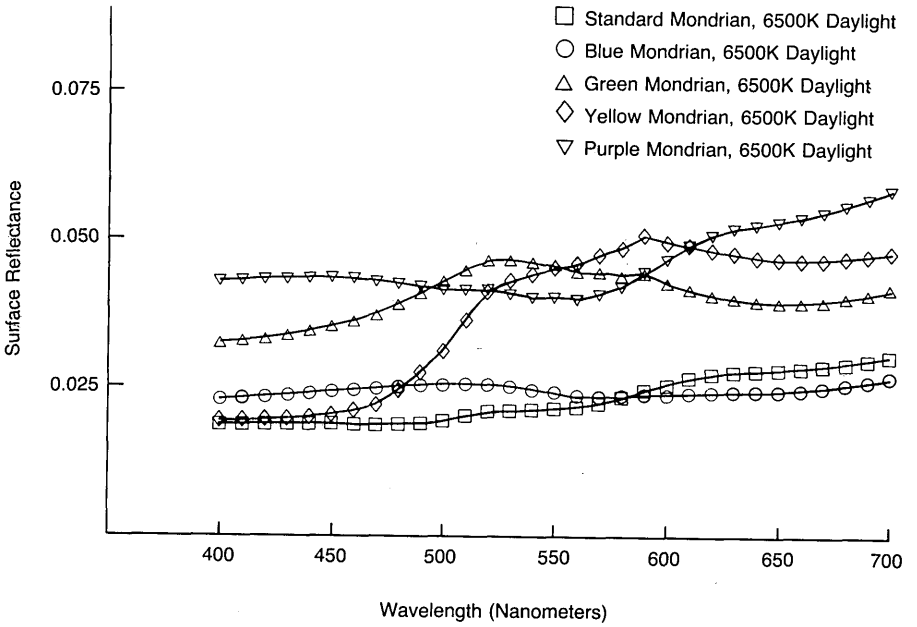


Fig. 4. The effective reference surface for the standard, blue, green, yellow, and purple Mondrians under CIE illuminant D_{65}.

respect to an effective reference surface. Indeed, the reflectance function σ_r^x for this surface can be found simply:

$$\sigma_r^x = \Lambda_E^{-1}\gamma, \tag{13}$$

where γ is the column vector consisting of the G_k. The effective reference surface used by the retinex varies both with the illuminant and with the surfaces in the scene. We have calculated how the effective reference surface depends on both the spectral power distribution of the ambient light and the surfaces in the scene. The effective reference surface is almost constant under changes of illuminant. Figure 4 shows the effective reference surface for the standard, blue, green, yellow, and purple Mondrians under D_{65}. The effective reference surface is markedly different for Mondrians containing different surfaces.

The difficulty with the retinex algorithm is that it normalizes to different reference surfaces for different scenes. It is as if, rather than using a standard gray card for color balancing as is done in industry, we were to use a different reference card for every scene. It is not surprising that the retinex algorithm fails to be color constant with changes in scene composition.

Finally, we note that it is possible to perform an empirical test of the retinex algorithm as a model of human performance without specifying the detailed algorithm parameters. To do this, begin with an arbitrary first image. Calculate the three photoreceptor response arrays for this image. Construct a second image so that its red photoreceptor array is a scaled version of the red photoreceptor array of the first image, so that its green photoreceptor array is a scaled version of the green photoreceptor array of the first image, and so that its blue photoreceptor array is a scaled version of the blue photoreceptor array of the first image. No matter how the paths are constructed, the retinex algorithm will assign identical lightness values to these two images. This is because, within each receptor class, the algorithm's normalization will discount the scale factors that distinguish the two. Thus the retinex algorithm predicts that the two images will have identical color appearance. We are unaware of any empirical tests of this prediction.

CONCLUSION

Historically, work on color constancy has emphasized the need to correct for the spectral power distribution of the ambient light. This emphasis has made it easy to forget that the information reflected from a single surface is insufficient to separate ambient light and surface reflectance and that color constancy algorithms require information from multiple surfaces. It is important to evaluate the performance of such algorithms when the surfaces that comprise the scene are varied. Human vision maintains approximate color constancy despite variation in the spectral reflectance functions of nearby surfaces and despite variation in the spectral power distribution of the ambient light.

The complex set of calculations that define the retinex algorithm is equivalent to a simple normalization. This normalization is not color constant: A color constant algorithm must correct for the ambient light independent of the surfaces in the scene. The retinex algorithm corrects for the light in a manner that depends strongly on the surfaces. This dependence is not characteristic of human color vision. We conclude that retinex is not a color constant algorithm and that it is not an adequate model of human performance.

APPENDIX A: RELATED IMPLEMENTATIONS OF THE RETINEX ALGORITHM

Introduction

In this appendix we review two additional methods proposed for calculating the lightness values for an image. The first

method is described by McCann and Houston[11] and Frankle and McCann.[10] It is very similar to the calculation described by Land. The method differs because it includes an additional nonlinear operation—the reset—that was also used in the original retinex calculations.[4] The second method is described by Horn[24] and cited by Land[8] as being equivalent to Land's method (but see Ref. 15).

Retinex with Reset

In the original retinex algorithm,[4] a special step is present in the path calculation. Consider the path starting at location x_1. If at any point during the path computation the value to be added to the accumulator $A(x)$ at location x exceeds 0, then the path calculation is *reset*. The reset operation has two parts. First, the value added to $A(x)$ is set to 0. Second, for the remainder of the calculations along the path, the value $\rho_k{}^x$ is substituted for the receptor response $\rho_k{}^{x_1}$. A sophisticated computational procedure for the retinex algorithm with reset is described by McCann and Houston[11] and Frankle and McCann.[10] It is not presented in terms of a set of path calculations, although they state that "the lightness calculation is in principle the same as in earlier work: ratio, product, reset, average" (Ref. 10, p. 1005). McCann and Houston compute the value of $\exp(l_k{}^x)$ rather than $l_k{}^x$ itself. In this appendix we use the notation $l_k{}^x$ to refer to the quantity calculated by their procedure rather than to the lightness, as defined by Land.[8] Whether the lightness is expressed in linear or logarithmic units has no bearing on the conclusions of this paper.

The computation is defined by a simple algorithm. For a given receptor class, the algorithm is initialized by setting the lightness value at all locations to 1. The algorithm determines the lightness at location x by using an interactive comparison procedure. The lightness at location x is compared with the lightness at a sequence of comparison locations, c_n. Since we wish to study the convergence properties of the algorithm, we assume that the comparison list is infinitely long and that all locations are repeatedly compared with location x. The lightness at all locations is computed concurrently. The lightness value at x is given by the following recursive equation, which defines the lightness after $n+1$ iterations in terms of the lightness after n iterations:

$$l^x(n+1) = \left\{ l^x(n) f^*\left[\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \right] \right\}^{1/2}, \tag{A1}$$

where $f^*[\ ]$ is a reset function defined by $f^*(a) = a$ for $a < 1$ and $f^*[a] = 1$ otherwise. In the limit, as the algorithm continues through many iterations, the lightness at pixel x is given by

$$l^x = \lim_{n\to\infty} \left\{ l^x(n) f^*\left[\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \right] \right\}^{1/2}. \tag{A2}$$

This limit will converge only if

$$(l^x)^2 = \lim_{n\to\infty} l^x(n) f^*\left[\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \right]. \tag{A3}$$

Further, since we are assuming that $\lim_{n\to\infty} l^x(n) = l^x$, it follows that convergence occurs if and only if

$$l^x = \lim_{n\to\infty} f^*\left[\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \right]. \tag{A4}$$

To understand the convergence of the lightness l^x in general, we begin by analyzing the lightness at the location at which the receptor response is smallest, x_s. For location x_s, it must be that $(\rho^{x_s}/\rho_{c_n}) \le 1$. Moreover, it may be seen from Eq. (A1) that the lightness at any location is always bounded above by 1. This is because the lightness value is initialized to 1, and at each step in the iteration the lightness is multiplied by a value that is less than or equal to 1. It follows that the location with the smallest sensor response will never encounter a reset operation during the calculation of its lightness value because the argument to the reset function is always less than 1. In studying the convergence of the lightness at x_s, we may remove the reset function so that we have

$$l^{x_s} = \lim_{n\to\infty} \frac{\rho^{x_s}}{\rho^{c_n}} l^{c_n}(n). \tag{A5}$$

Since all locations x are repeatedly compared with location x_s, then, if l^{x_s} converges, Eq. (A5) must be true for any comparison location x. If the computation converges, the lightness of any location can be expressed in terms of the lightness at x_s by the equation

$$l^x = \left(\frac{\rho^x}{\rho^{x_s}} \right) l^{x_s}. \tag{A6}$$

Now we consider the location with the largest receptor response, x_l. We will show that for each location x, $l^x \le (\rho^x/\rho^{x_l})$ and $l^x \ge (\rho^x/\rho^{x_l})$. This forces the conclusion that $l^x = (\rho^x/\rho^{x_l})$.

To show $l^x \le (\rho^x/\rho^{x_l})$, we rewrite Eq. (A6) for location x_l:

$$l^{x_s} = l^{x_l} \left(\frac{\rho^{x_s}}{\rho^{x_l}} \right). \tag{A7}$$

Since l^{x_l} is bounded by 1, it follows that $l^{x_s} \le (\rho^{x_s}/\rho^{x_l})$. Using this inequality in Eq. (A6), we conclude that

$$l^x \le \frac{\rho^x}{\rho^{x_l}} \tag{A8}$$

for all possible comparison locations x.

To show $l^x \ge (\rho^x/\rho^{x_l})$, we use a proof by induction with respect to the steps of the iteration. On the first step of the iteration, $l^x(1) = 1 \ge (\rho^x/\rho^{x_l})$ for every location x. Suppose that at step n, $l^x(n) \ge (\rho^x/\rho^{x_l})$ for every location x. Then we have

$$\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \ge \left(\frac{\rho^x}{\rho^{c_n}} \right) \left(\frac{\rho^{c_n}}{\rho^{x_l}} \right) \ge \frac{\rho^x}{\rho^{x_l}}. \tag{A9}$$

Since the reset function is monotonic, we have

$$f^*\left[\frac{\rho^x}{\rho^{c_n}} l^{c_n}(n) \right] \ge f^*\left(\frac{\rho^x}{\rho^{x_l}} \right) = \frac{\rho^x}{\rho^{x_l}}. \tag{A10}$$

We substitute for the reset function in Eq. (A1) to get

$$l^x(n+1) \ge \left[l^x(n) f^*\left(\frac{\rho^x}{\rho^{x_l}} \right) \right]^{1/2} = \left[l^x(n) \frac{\rho^x}{\rho^{x_l}} \right]^{1/2}. \tag{A11}$$

Since $l^x(n) \ge (\rho^x/\rho^{x_l})$ by the induction hypothesis, it follows that $l^x(n+1) \ge (\rho^x/\rho^{x_l})$. We conclude by induction that $l^x \ge (\rho^x/\rho^{x_l})$. It follows that $l^x = (\rho^x/\rho^{x_l})$. Thus, if the computation converges, which we have not shown, the lightness values approach a simple normalization of the receptor response with respect to the response at the largest location for

Mondrian	STANDARD	RED	BLUE	GRAY	GRAY-RED
Predicted Color	R4/4	Y7/6	YR6/8	10YR7/8	GY9/5
Lightness Triplet	−2.1123 −2.2586 −2.1942	−0.9906 −1.0808 −2.0028	−1.0028 −1.2176 −2.2225	−0.6581 −0.9327 −1.9953	0.0000 0.0000 −0.9079

Test Chip: R4/4 Light: D_{65}

Mondrian	STANDARD	RED	BLUE	GRAY	GRAY-RED
Predicted Color	Y4/4	10YR5/6	10Y5/6	Y6/6	GY8/8
Lightness Triplet	−2.1974 −2.1609 −2.9954	−1.5781 −1.5527 −2.5189	−1.5903 −1.6895 −2.7385	−1.2456 −1.4047 −2.5113	−0.5875 −0.4719 −1.4239

Test Chip: Y4/4 Light: D_{65}

Change of Context Simulation Results

Fig. 5. The spatial composition of the red, blue, gray, and gray-red Mondrians used in the simulations of retinex algorithm with reset.

the receptor class. This parallels the computation by Land.[8] Land's calculation is equivalent to normalization with respect to the geometric mean of receptor responses for the receptor class, whereas the retinex with reset calculation is equivalent to normalization with respect to the response at the largest location. Retinex with reset will also be sensitive to changes in the viewing context; in particular, it will be sensitive to any change in the reflectance function of the surface causing the largest receptor response.

We simulated the performance of the McCann–Houston algorithm in the limiting case that the lightness values are given by

$$\exp(l_k^x) = \frac{\rho_k^x}{\rho_k^{x_l}}. \qquad \text{(A12)}$$

Figure 5 shows the four Mondrians that we used in this simulation. These 3 × 3 Mondrians differ only in the composition of the lowest row: Their upper two rows are identical. We will refer to these four Mondrians as the red, blue, gray, and gray-red Mondrians, respectively. We chose these Mondrians because the variation in the composition of the lower row significantly alters the largest receptor responses when they are illuminated by D_{65}. Figure 6 shows the lightness triplets and corresponding predicted color values for two of the test chips, as computed for each of the Mondrians. For the retinex algorithm with reset, the predicted color values also change dramatically with small changes in the Mondrian. We conclude that the reset operation does not significantly reduce the retinex algorithm's overdependence on the composition of the Mondrian. On the contrary, the

R4/4	Y4/4	BG4/4
B2/2	B3/2	B4/2
10YR5/6	R8/4	P6/8

Red Mondrian

R4/4	Y4/4	BG4/4
B2/2	B3/2	B4/2
10YR5/6	B8/2	P6/8

Blue Mondrian

R4/4	Y4/4	BG4/4
B2/2	B3/2	B4/2
10YR5/6	N2/	P6/8

Gray Mondrian

R4/4	Y4/4	BG4/4
B2/2	B3/2	B4/2
10YR5/6	N2/	RP4/4

Gray-red Mondrian

Test Mondrian Composition

Fig. 6. Predicted color values and lightness triplets computed by the McCann–Houston retinex algorithm for two chips in the standard, red, blue, gray, and gray-red Mondrians. The lightness values were computed according to Eq. (A12).

reset operation localizes the dependence to just a few pixels: the pixels with the largest response in each receptor class.

It is possible to run the McCann–Houston[11] algorithm without choosing the comparison pixels from the entire image. In this case, the limiting lightness value for a pixel at location x will be given by normalization to the pixel with the largest receptor response among the set of comparison pixels chosen for location x. We do not believe that limiting the set of comparison pixels can eliminate the overdependence on the composition of the Mondrian. The predicted color for a particular location will remain sensitive to variation in the surface reflectance function at pixels in the comparison set for that location.

Determining Lightness from an Image

The second method that we consider is presented in Horn's[24] paper entitled "Determining lightness from an image." Land[8] cites this as another method of implementing the retinex calculation. Recently, it has been developed further by Blake[15] and Terzopoulos.[16] The central premise of Horn's[24] paper is that the surface reflectance and spectral power distribution of the ambient light may be separated in the image data. The physical property of images that permits this, it is argued, is that the spatial variation in the ambient lighting occurs at a lower rate than the spatial variation of the surface reflectance of the objects.

Horn bases his calculations on two steps. First, he assumes that the image may be written as the product of two functions, object reflectance at a point, which we call $r(x, y)$, and ambient light intensity, $s(x, y)$. He writes the recorded image intensity $p(x, y)$ as

$$p(x, y) = s(x, y)r(x, y) \tag{A13}$$

(Ref. 24, p. 281).

In the case of the uniformly illuminated Mondrian, the spatial distribution of the ambient light has no step edges. The boundaries of the colored papers, on the other hand, define many different step edges of the surface reflectance. Horn assumes that, in general, boundaries between objects will give rise to step edges and changes in the ambient light will give rise only to smoothly varying changes in intensity. Based on this assumption, Horn argues that we may recover the surface reflectance function as follows. First, compute the logarithm of the image-intensity record

$$\log[p(x, y)] = \log[s(x, y)] + \log[r(x, y)]. \tag{A14}$$

Next compute the Laplacian of the resulting image. Since the Laplacian is a bandpass spatial operator, it will take on a relatively larger values when approaching step discontinuities, and it will be small or zero over regions of nearly constant value. By assumption, the larger values of the Laplacian arise from surface reflectance changes, owing to the term $\log[r(x, y)]$. Horn applies a threshold operator to the Laplacian of the logarithm of the image data. The threshold operator sets all values smaller than some criterion level to zero and passes unchanged all values larger than criterion. Horn's claim is that, following thresholding, the effects of the ambient light have been removed.

Horn's method does not make the role of wavelength explicit anywhere in the calculation. By doing so, it is easy to see why the method does not recover the surface reflectance function. We rewrite Eq. (A13) with the spatial coordinates made explicit only in the surface reflectance function:

$$\rho(x, y) = \sum_{n=1}^{N} E(\lambda_n)S(x, y, \lambda_n)R_k(\lambda_n). \tag{A15}$$

When the role of wavelength is made explicit, we find that applying the logarithm to both sides of Eq. (A15) does not yield the additive separation of the surface and light functions required for the use of the Laplacian and thresholding operations. Thus Horn's method cannot eliminate the effects of changing the spectral power distribution of the ambient light.

The method may remove some of the effects of variation in the intensity of the ambient illumination. For example, as noted by a referee, if the ambient light function is separable in space and wavelength so that $E(\lambda_n) = E_1(\lambda)E_2(x, y)$, then

$$\rho(x, y) = E_2(x, y) \sum_{n=1}^{N} E_1(\lambda)S(x, y, \lambda_n)R_k(\lambda_n). \tag{A16}$$

Under this assumption, applying the logarithm to both sides of Eq. (A16) permits an additive separation of the *spatial variation* of the ambient lighting from the spatial variation of the surface reflectance function. However, the *spectral power distribution* of the lighting remains confounded with the surface reflectance functions.

ACKNOWLEDGMENTS

This work was supported by grant no. 2 RO1 EY03164 from the National Eye Institute and grant NCC 2-307 from the NASA-Ames Research Center. Some of the material in this paper is based on work supported under a National Science Foundation Graduate Fellowship. We thank W. Brainard, M. Huang, L. Maloney, J. Nachmias, M. Pavel, R. N. Shepard, and D. Varner for useful discussions.

REFERENCES AND NOTES

1. H. von Helmholtz, *Handbuch der Physiologischen Optik*, 2nd ed. (Voss, Hamburg, 1896).
2. H. Helson, "Fundamental problems in color vision. I. The principle governing changes in hue, saturation and lightness of nonselective samples in chromatic illumination," J. Exp. Psychol. **23,** 439–476 (1938).
3. D. B. Judd, "Hue saturation and lightness of surface colors with chromatic illumination," J. Opt. Soc. Am. **30,** 2–32 (1940).
4. E. H. Land and J. J. McCann, "Lightness and retinex theory," J. Opt. Soc. Am. **61,** 1–11 (1971).
5. M. H. Brill, "A device performing illuminant-invariant assessment of chromatic relations," J. Theor. Biol. **71,** 473–478 (1978).
6. G. Buchsbaum, "A spatial processor model for object color perception," J. Franklin Inst. **310,** 1–26 (1980).
7. L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," J. Opt. Soc. Am. A **3,** 29–33 (1986).
8. E. H. Land, "Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image," Proc. Nat. Acad. Sci. USA **80,** 5163–5169 (1983).
9. E. H. Land, D. H. Hubel, M. S. Livingston, S. H. Hollis, and M. M. Burns, "Colour-generating interactions across the corpus callosum," Nature **303,** 616–618 (1983).
10. J. Frankle and J. J. McCann, "Method and apparatus for lightness imaging," United States Patent No. 4,384,336, 1983.
11. J. J. McCann and K. Houston, "Calculating color sensations from arrays of physical stimuli," IEEE Trans. Systems Man Cybern. **SMC-13,** 1000–1007 (1983).

12. M. Livingston and D. Hubel, "Anatomy and physiology of a color system in the primate visual cortex," J. Neurosci. **4,** 309–356 (1984).
13. D. Ingle, "The goldfish as a retinex animal," Science **227,** 651–654 (1985).
14. S. Grossberg and E. Mingolla, "Neural dynamics of form perception: boundary completions, illusory figures, and neon color spreading," Psychol. Rev. **92,** 173–207 (1985).
15. A. Blake, "Boundary conditions for lightness computation in Mondrian world," Comput. Vis. Graphics Image Process. **32,** 314–327 (1985).
16. D. Terzopoulos, "Image analysis using multigrid relaxation methods," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8,** 129–139 (1986).
17. E. H. Land, "Recent advances in retinex theory," Vision Res. **26,** 7–22 (1986).
18. M. M. Woolfson, "Some new aspects of color perception," IBM J. Res. Dev. **3,** 313 (1959).
19. G. Walls, "Land! Land!" Psychol. Bull. **57,** 29–48 (1960).
20. D. B. Judd, "Appraisal of Land's work on two-primary color projections," J. Opt. Soc. Am. **50,** 254–268 (1960).
21. M. Brill, and G. West, "Contributions to the theory of color," J. Math. Biol. **11,** 337–350 (1981).
22. J. Worthey, "Limitations of color constancy," J. Opt. Soc. Am. A **2,** 1014–1026 (1985).
23. R. M. Shapley, "The importance of contrast for the activity of single neurons, the VEP and perception," Vision Res. **26,** 45–62 (1986).
24. B. K. P. Horn, "Determining lightness from an image," Comput. Graphics Image Process. **3,** 277–299 (1974).
25. J. J. McCann, S. P. McKee, and T. H. Taylor, "Quantitative studies in retinex theory: a comparison between theoretical predictions and observer responses to the color Mondrian experiments," Vision Res. **16,** 445–448 (1976).
26. Partial descriptions of the path-generation process can be found in McCann *et al.* (Ref. 25, p. 454), where they write: "The origin of the path and its direction in the 480-point array were determined by a random-number generator. The paths traveled straight ahead until they reached the perimeter of the target where they either reflected back across the target or traveled along the perimeter. The direction of the reflection from the perimeter was also chosen by the random-number generator." Land (Ref. 8, p. 5165) writes "The signal will proceed ... with freedom to branch within the rule of implied directionality inherent in the concept of a signal radiating from a single source," whereas in Land (Ref. 17, p. 12) the paths are generated by an "optical mouse system [which] is swept along the random pathways" of the image.
27. McCann *et al.*[25] emphasize that the details of the path-generation process are not critical when they write (Ref. 25, p. 453): "The model described in this paper is one of many embodiments using the ratio-multiplication process."
28. L. Maloney, "Computational approaches to color vision," Ph.D. dissertation (Stanford University, Stanford, Calif., 1984).
29. B. A. Wandell, "The synthesis and analysis of color images," IEEE Trans. Pattern Anal. Mach. Vis. (to be published).
30. R. M. Boynton, *Human Color Vision* (Holt, New York, 1979).
31. D. Nickerson, "Spectrophometric data for a collection of Munsell samples," Report, Cotton Division, U.S. Department of Agriculture, Washington, D.C., 1957.
32. J. Cohen, "Dependency of the spectral reflectance curves of the Munsell color chips," Psychonanal. Sci. **1,** 369–370 (1964).
33. D. Judd and G. Wyszecki, *Color in Business, Science, and Industry* (Wiley, New York, 1975).