Biol. Cybernetics 33, 125–135 (1979)

Biological Cybernetics

Why Rods and Cones?

Whitman Richards
Department of Psychology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Abstract. A key problem in vision is to normalize one's lightness scale so that surface reflectances are always assigned the same gray value regardless of the illumination level. The solution requires an assessment of the relation between the strength of the illuminant and the strength of the image signal-information that is not available in the image alone. However, the level of scattered light in the optical system does provide an independent measure of the illuminant strength, and can be used to solve the lightness scale normalization problem. To do this requires a comparison between two imaging systems, each of which respond differently to the internal optical scatter. The rod and cone systems have properties that are ideally suited for such a role.

1. Introduction

Some years ago William Rushton (1962) raised the tantalizing question: "Why rods and cones?". The most common answer begins by pointing out that rods have a lower threshold than cones, and extend our range of vision into darkness, whereas the evolution of cones added spectral analysis to our visual repertoire.

These standard answers seem unsatisfactory to me (and presumably to Rushton also) for the following reasons:

1) Our total dynamic range of photopic vision is about ten million to one. Rod vision only adds a factor of a few hundred to this range, and at wavelengths less than 520 nm. Above 620 nm, cones are more sensitive than rods at the lowest light levels (Wald, 1945). Is this small contribution by rods to night vision thus their most significant function?

2) Rods can participate in color vision (McCann and Benton, 1969). In fact, color vision is possible utilizing rods plus one cone type (Blackwell and Blackwell, 1961). In some animals (i.e., frog), two rod types coexist in the same retina and could subserve color vision. Therefore, although multiple receptor types are necessary for color vision, these receptors need not be cones. Thus, why cones?

Yet rods and cones are structurally quite distinct elements as their names suggest. Perhaps it is these structural differences that provide the clue to their significant roles. Because of the tapering of the cone outer segment, these receptors have a narrow field of view centered on the pupil of the eye (Laties, 1968; Enoch and Laties, 1971). Rods, on the other hand, are not so concerned with the direction of light reaching them, and will respond well to light scattered about within the eye itself. Cones are thus sensitive to the direction of light, whereas rods are not (Stiles and Crawford, 1933).

A second important difference between the two receptor types is that cones recover rapidly from bleaching lights whereas rods do not. Associated with the more sluggish character of the rod system is a greater spatial integrating power, but all at the expense of less sensitivity to light increments (or decrements). For example, when the rod system is isolated, incremental sensitivity is found to be about 6–15 times less than that of the long-wave cones (Stiles, 1959; Aguilar and Stiles, 1954).

Two very distinctive differences between rods and cones are thus 1) cone responses are quite dependent upon the direction of light striking the retina, whereas rods are not, and 2) the sensitivity of the rod system to small changes in (photopic) illumination is inferior to that of the cone systems. These two properties can provide a solution to a key problem of vision: the normalization of lightness scales.

2. The Problem

As we view the world about us, objects and surfaces take on various shades of gray from black to white –

0340-1200/79/0033/0125/$02.20

one critical attribute of colors. If the illumination level is changed as we view a given scene, the grays or lightness values we associate with a given object or surface stay almost the same. This is the phenomenon of brightness constancy. Although the illumination on the retina can change over a factor of one-thousand, the lightnesses of a surface will remain the same after a brief period of adaptation to the new illuminant. Clearly, as Land has stressed so forcefully (1964, 1971), it is the reflectances of the surfaces that the grays or lightness sensations represent to us. For although the flux entering our eyes may change as the illumination changes from bright to dim, the ratios of the intensities associated with each surface will not change, for these ratios are set by the reflectivities of the objects themselves.

How then do we know when the flux striking the retina corresponds to a highly reflecting surface or a poorly reflecting one? Can we merely look at the activity of each individual receptor and reach a correct conclusion? Clearly we cannot, for in dim light a modest receptor activity may be associated with a highly reflecting surface which appears "white", whereas under sunlight the same receptor activity can signal darkness. What, then, is the procedure we use to decide upon the gray value to be assigned to each particular region of retinal activity?

3. Land's Solution

Land's studies (1964, 1971) as well as others before him (see Graham, 1965) show that a comparison must be made between receptor activites at different portions of the retina. Thus, we can deduce that surface A is darker (less reflective) than an adjacent surface B because the receptor activity in region A is less than that of B. But what gray level value do we assign to B? Land solves this problem of normalizing the lightness scale by finding the surface eliciting the greatest receptor activity and designating this as "white" with a reflectance of 1.0. The lightness values of the remaining surfaces are then computed by comparison with this most intense region.

But will this "brightest" region always be 100% reflecting? Why does the rule fail in darkness under rod vision where surfaces having very high reflectances may still appear grayish and not white? And what if the most intense region is not a surface at all, but a light source?

4. Average Firing Rate

Another possible solution to the lightness normalization problem is that an average receptor activity level is computed over a local region. This average receptor activity is often taken as representing a neutral gray sensation similar to that seen in a Ganzfeld. When receptor activities change across a border, the magnitude of this change is compared with the local average in order to yield a lightness sensation. Although not explicitly formulated, this proposal is implicit in many treatments of gray scales (Graham, 1965).

Such a model suffers from several weaknesses. First, how is the area of averaging to be determined? If it is too small, then essentially only the second spatial derivative of intensities are being processed and all uniform areas should appear equally gray, regardless of their reflectances. On the other hand, if a large integrating area is used to determine a reference gray level, then even small illumination gradients should lead to the perception of whiteness in the regions of high retinal illuminance and darkness in the opposite regions. This does not happen. Finally, why should snow scenes or beaches look so bright and white in sunlight, and yet so gray at night? If we were computing our lightness scales using an average of receptor activity for a reference, then these highly reflective scenes should appear gray under all levels of illumination.

If one is to counter by arguing that snow scenes appear white because receptor activity is high, then what is the internal reference that judges "how high"? And specifically, how is this internal standard used to identify a surface that has a known reflectance, in spite of vast changes in illumination level. This is the essence of the lightness normalization problem.

5. Computational Analysis

5.1. Optics and Illuminants

In the natural world, the light seen reflected from an object may vary over a considerable range, depending upon the reflectance of the object, the positions of the light source and viewer, and especially the shape of the object. Horn (1975) and Woodham (1978) have analyzed this problem in considerable detail, and show that the reflectance of an object can be determined up to a constant, provided that the surface has uniform reflectance. To determine the constant and hence the true surface reflectance, additional information is needed, such as the strength of the source relative to the intensity of a point in the scene, or in the image. Once this information has been obtained for one point in an image, then, in principle, the true reflectances at all points in the image can be determined.

To recover the value of the scalar constant is the essence of the lightness normalization problems. Is there any additional information available in the image that has not previously been used and which will allow us to compute the true reflectivity value? One factor present in all images is a certain amount of scattered light, depending upon the quality of the optical system. However, because of scatter within the optical system, no image point can have a zero value. Instead, the intensities at every point are raised by a constant increment, assuming uniform scatter throughout the system. Such increments will not change the gradients on which the Horn type of analysis is based, but the increment will affect the intensity ratios across the boundary between two surfaces of different reflectance (such as a Land Mondrian).

To see that scattered light under some conditions permits a solution to the lightness normalization problem, it is useful to follow an analysis used by Ullman (1976). Consider a Mondrian constructed from flat matte rectangles of different sizes and reflectances (Land, 1971). In the presence of a point light source there will be a gradient of reflectivity across the Mondrian, depending upon the relations between the viewer, the source, and the inclination of the Mondrian. Let these angular dependencies of reflectivity be a function $f(\theta, \phi, \psi)$, which is normalized to 1.

Let i_j be the recorded image intensity for surface J of reflectance ϱ_j. Assume that the transfer function T is known, where T describes the relation between the recorded image intensity i_j and the actual incident intensity e_j. Then with no stray light in the optical system, the resultant measure of image intensities for a source of strength I would be

$$i_j = T(e_j) = T[\varrho_j \cdot I \cdot f(\theta, \phi, \psi)]\,. \tag{1}$$

However, if stray light of the optical system is included, then a constant fraction $\beta(x, y)$ of I must be added, where x, y refer to the image coordinates in the neighborhood of the position identified with i_j. For simplification, we will assume that the source is visible in the image with no other light illuminating the Mondrian, and that $\beta(x, y)$ is known completely and is constant, corresponding to scatter in a whole-field integrating system. Under these very restrictive conditions, the image intensities now become

$$i_j = T(e_j) = T[\varrho_j \cdot I \cdot f(\theta, \phi, \psi) + \beta \cdot I]\,. \tag{2}$$

Consider next a boundary between two abutting rectangles of the Mondrian. At a neighboring region along the boundary we can describe a second set of image intensity values i'_j and from its relation to i_j define an image intensity gradient G_j:

$$i_j - i'_j = T(G_j) = T\{\varrho_j I[f(\theta, \phi, \psi) - f(\theta', \phi', \psi')]\}\,. \tag{3}$$

Note that the scattered light term drops out because it is constant over the neighborhood and therefore does not change the gradient measure.

At the border between two abutting surfaces S_1 and S_2, each of different reflectances ϱ_1 and ϱ_2, we now have the following relations between the two image intensities and gradients:

$$\begin{aligned} \frac{T^{-1}(i_1)}{T^{-1}(i_2)} &= \frac{e_1}{e_2} = \frac{(\varrho_1 \cdot f(\theta, \phi, \psi) + \beta)}{(\varrho_2 \cdot f(\theta, \phi, \psi) + \beta)}, \\ \frac{T^{-1}(i_1 - i'_1)}{T^{-1}(i_2 - i'_2)} &= \frac{G_1}{G_2} = \frac{\varrho_1}{\varrho_2}, \end{aligned} \tag{4}$$

where T^{-1} is the inverse of the transfer function of the recording system.

Thus, if the reflectivity function, f, is known, then the reflectances ϱ_1 and ϱ_2 can be determined because β and T^{-1} are measurable directly, and the values e_1/e_2 and G_1/G_2 can be derived from the image. But Horn (1975) has shown that f is recoverable from the image gradients for surfaces of uniform reflectance, and hence reflectance can be recovered from the image under certain very restrictive conditions. For the Mondrian, the solution is

$$\varrho_j = \frac{\beta G_j (i_j - i_k)}{G_j i_k - G_k i_j}, \tag{5}$$

where S_j and S_k are abutting surfaces.

Unfortunately, in most real scenes and images, the light reflected from adjacent objects provides an important contribution to the illuminance of surfaces, and Eq. (2) is not valid. Furthermore, all sources may not be in the image, nor will the scattered light distribution function $\beta(x, y)$ necessarily be known. For many scenes, however, indirect lighting constitutes the major portion of the illumination of a surface. For this case, an average scene reflectance, $\bar{\varrho}$, must be assumed (or measured) and Eq. (2) must be modified as follows:

$$i_j = T(e_j) = T[\varrho_j \cdot I \cdot f(\theta, \phi, \psi) + \bar{\varrho} \cdot \beta \cdot I]\,. \tag{6}$$

The reflectance ϱ_j of a surface S_j now cannot be determined without knowledge of $\bar{\varrho}$. Yet the brain clearly has a scheme for normalizing gray scales. Is it merely assuming an average reflectance value?

5.2. *Two Optical Systems*

Consider the case where two different optical systems, C and R, process the same image. Let the C system (or image) be essentially free of scatter light, whereas the R image system is not, with its optical system acting like a total integrator of the scattered light. For the R system, therefore, the scattered light will be constant everywhere (although the recording image may be such that only a limited portion of the total field of scatter is sampled). Let c_j be the direct, scatter-free intensity of the C system for a point on surface S_j and let r_j be the image intensity for the same point as seen through the R optical system. Then for the surface *reflectivity* ϱ_j we have

$$T_i^{-1}(c_j')=c_j=\varrho_j\cdot I\,,$$
$$T_r^{-1}(r_j')=r_j=\varrho_j\cdot I+\beta\cdot I\cdot\bar{\varrho}\,, \tag{7}$$

where $\bar{\varrho}$ is the average reflectivity, c_j' and r_j' are respectively the outputs of the C and R imaging systems, and T^{-1} is the inverse of the image intensity transfer function. For the moment, assume that the two transfer functions T_c and T_r are of the same form (i.e., linear, log, exponential, etc.) but differ only by a constant factor α such that $\alpha T_c^{-1}=T_r^{-1}$.

Consider now the following three relations between the image intensities of adjacent surfaces S_j and S_k

$$\frac{c_j}{c_k}=\frac{\varrho_j}{\varrho_k}, \tag{8}$$

$$\frac{r_j}{r_k}=\frac{\varrho_j+\beta\bar{\varrho}}{\varrho_k+\beta\bar{\varrho}}, \tag{9}$$

and for any single surface,

$$\frac{c_i}{r_i}=\frac{\alpha\varrho_i}{\varrho_i+\beta\bar{\varrho}}. \tag{10}$$

Note that Eqs. (8) and (9) are formerly the same as Eq. (4) but now gradient information need not be used.

We now would like to solve for ϱ_i explicitly, yet this is not possible although surprisingly both α and β can be determined. For example, β can be deduced for an image by one of two ways. First, if the image weighting function for the R system is not known exactly, then consider a boundary between two large homogeneous areas (sky and trees, grass and dirt), where each area is larger than the collecting area of a sensor in the R image plane. Then at the boundary, $2\bar{\varrho}=\varrho_1+\varrho_2$. Substitution in (8) and (9) yields

$$\beta=\frac{2(r_1c_2-r_1c_2)}{(r_1-r_2)(c_1+c_2)}. \tag{11}$$

If the weighting function for an R sensor is known, then a second determination of β can be made from any region in the image by application of Eqs. (8) and (9) to all surfaces within the range of the weighting function.

The constant α, relating the two image transfer functions can also be deduced from Eq. (10), for at any point in the image:

$$\beta\bar{\varrho}=\frac{\varrho_i(\alpha r_i-c_i)}{c_i}=\frac{\varrho_k(\alpha r_k-c_k)}{c_k} \tag{12}$$

combining (13) with (8) we find that across any border

$$\alpha=\frac{c_i-c_k}{r_i-r_k} \tag{13}$$

providing a local determination of α. Thus, both α and β can be determined locally in the image.

It is now possible to solve the lightness scale normalization problem. Although without an explicit value for $\bar{\varrho}$ the reflectivity of any surface cannot be determined, we do have sufficient information to find a surface whose reflectivity matches $\bar{\varrho}$. Once such a patch has been found, then Eq. (8) can be used to determine the remaining reflectivities. The gray scale will thus always be normalized to the average reflectivity $\bar{\varrho}$. For such a patch $S_0, \varrho_0=\bar{\varrho}$ and from (10)

$$\frac{c_0}{r_0}=\frac{\alpha}{1+\beta}, \tag{14}$$

where α and β are determined from (11) and (13). Therefore, if a patch yields a C to R ratio of c_0/r_0, then that patch has an average reflectivity $\bar{\varrho}$.

To summarize, lightness or reflectance scales can be normalized using only image signals provided that

i) There are two imaging systems, one sensitive to scattered light, the other not.

ii) The form of the transfer function of the two image sensing systems is known

iii) These two transfer functions differ by a constant (which need not be known)

iv) The distribution of scattered light in the optical system is constant (but it need not be known).

If the transfer function is linear, then the average reflectance $\bar{\varrho}$ will be proportional to the average of the c_i' signals, and the normalization is trivial. If the transfer function is nonlinear, however, the computation becomes elaborate. Some non-linear forms of the transfer function, however, can yield simple computations. For example, an ideal transfer function would be one such that $c_0/r_0=1$ (i.e., $c_0=r_0$) when a patch had an average reflectance. By pairing two such imaging systems, then the transfer function need not be known. As will be seen shortly, this solution is apparently used by the brain.

5.3. Scattered Light Model

Within the human eye, the rods and cones constitute two characteristically different imaging systems. Because the directionally sensitive cones are oriented toward the pupil (Laties, 1968; Enoch and Laties, 1971) they must be relatively insensitive to scattered light. Rods, however, are not directionally sensitive and will respond to scattered light.

For each individual rod or cone, the signal, V, in any retinal region can be described by the Naka-Rushton equation (Naka and Rushton, 1966; Normann and Werblin, 1974):

$$V = \frac{V^* e}{e+\sigma}, \tag{15}$$

where e is the image intensity and σ and V^* are constants of the system. (V^* is the maximum voltage and σ is the intensity that elicits a signal one-half V^*). This relation can be recast to indicate the form of T^{-1}:

$$e = \frac{\sigma V}{(V^* - V)}. \tag{16}$$

Now replace the signal V by the rod and cone signals $V_r = r'$ and $V_c = c'$ [see Eq. (7)] and let σ_r and σ_c be the appropriate constants for each system. Then following Eq. (10), we find that the relation between the image intensities c_i and r_i and the signals c' and r' available to the brain will be

$$\frac{c_i}{r_i} = \frac{\sigma_c V_c (V_r^* - V_r)}{\sigma_r V_r (V_c^* - V_c)} = \frac{\varrho_i}{\varrho_i + \beta \bar{\varrho}} \tag{17}$$

with $\alpha = \sigma_r/\sigma_c$. An analysis similar to the preceding section can now be made to solve for β and α in terms of the available signal values $V_c = c'$ and $V_r = r'$. Such a computation to determine a region having an average reflectance, $\bar{\varrho}$, is clearly quite complex and ludicrous.

Instead, consider the much simpler result that would occur if the brain merely was required to determine whether the rod and cone signals were equal. Substituting V_c equals V_r in Eq. (17) we find that

$$\frac{\sigma_c}{\sigma_r} \frac{(V_r^* - V_r)}{(V_c^* - V_c)} = \frac{\varrho_0}{\varrho_0 + \beta \bar{\varrho}}, \tag{18}$$

where ϱ_0 is the patch of "standard" reflectance.

To solve Eq. (18) a decision must be made regarding how the two V-log$I$ curves for the rod and cone systems will be normalized with respect to each other. The simplest solution would be to scale the intensity values according to the equivalent background levels of each system. Much evidence supports the notion that each receptor system sets its own equivalent background level for any steady-state level of illumination (Barlow and Sparrock, 1964; Graham, 1965; Crawford, 1947). In effect, the constants σ_c and σ_r are describing the equivalent backgrounds of the two systems in the particular illuminant – be it achromatic or chromatic. Such a scale change in the effective intensity units is equivalent to a shift of the V-log$I$ curves along the horizontal axis in Fig. 1 for each intensity is merely changed by a factor proportional to σ. Now let the two R and C curves cross the C curve at a point where V_c equals $1/A$ of its maximum signal V_c^*. Then it can be shown that the rod response at this cross-over point will be

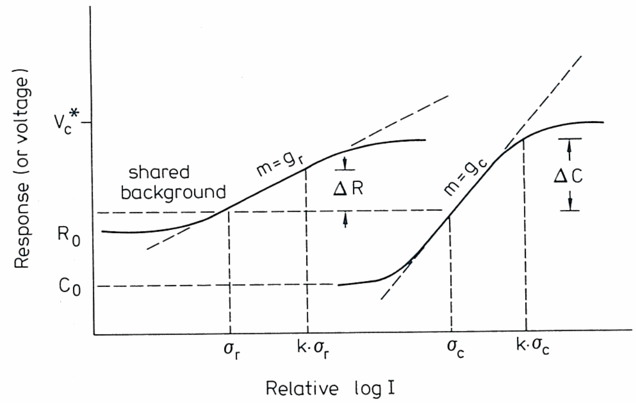

Fig. 1. Hypothetical response curves for rods and cones, adapted from Normann and Werblin (1974). Increments on equivalent backgrounds σ_c and σ_r lead to signals ΔC and ΔR, assuming that the response levels for $1/2\,\sigma_c$ and $1/2\,\sigma_r$ are equalized. Gains g_c and g_r are the slopes of the respective curves at their inflector points (see Appendix 1)

$$V_r = V_r^* [1 + \sigma_r (A-1)/\sigma_c]. \tag{19}$$

Substituting into Eq. (18) we obtain

$$\frac{V_r^*}{V_c^*} \cdot \frac{A}{[1 + \sigma_r (A-1)/\sigma_c]} = \frac{\varrho_0}{\varrho_0 + \beta \bar{\varrho}}. \tag{20}$$

But the normalization of the rod and cone intensity scales to these respective equivalent backgrounds[1] corresponds to setting $\sigma_r = \sigma_c$. Hence for all A's we can obtain the same solution for ϱ_0:

$$\varrho_0 = \frac{\beta \bar{\varrho}}{(V_c^*/V_r^*) - 1}. \tag{21}$$

Thus, as long as the saturating voltages V_c^*, V_r^* and the scatter coefficient β remain constant in the two imaging systems, a natural strategy is to assume $\bar{\varrho}$ is generally constant also, at least over a wide enough field of view. For the normalization of lightness scales, it is not necessary that a patch of average reflectance always be determined, but only that the reference is always a patch of the same reflectance, regardless of the scene. Thus a surface of reflectivity ϱ_0 has a unique interpretation, for it can serve as a reflectivity against which other surface reflectivities can be compared. I will call a surface reflectance satisfying this condition as having a unique gray reflectivity.

1 Note that another possibility for normalization of the rod and cone intensity scales instead of a horizontal shift along log$I$ is to raise the rod response function by a bias voltage. In this case, a similar solution can still be obtained

To summarize again the conditions for which Eq. (21) will yield a unique gray reflectivity:

i) The rod and cone transfer functions are approximated by the Naka-Rushton equation and each have different saturation constants (V_c^* and V_r^*).

ii) Each system independently normalizes its V-$\log I$ curve to its own equivalent background.

iii) The rods are much more sensitive to scattered light than the cones.

iv) The distribution of scattered light in the optical system is constant.

v) The reflectance of a scene on the average is constant.

Note that the first two (design) constraints are commonly accepted and are very general properties of visual system, the third is a specific but well documented fact, the fourth is derived from a physical limitation imposed upon all optical systems, and the final condition is an assumption about the nature of the external world.

6. Psychophysical Comparisons

6.1. Unique Gray Reflectance

To demonstrate that the human visual system may indeed use a lightness normalization scheme suggested by the scattered light model, we must obtain independent estimates of V_r^*, V_c^*, and β and $\bar{\varrho}$. Several estimates of the first two parameters are available. Of these perhaps the data of Normann and Werblin (1974) are the most direct, although they were obtained from *Necturus* retina. Appendix 1 shows that V_r^* and V_c^* correspond to the slopes of the receptor response curve plotted versus $\log I$ as shown in Fig. 1. Thus, the ratio V_c^*/V_r^* is the same as the ratios of rod-cone sensitivities when plotted on a $\mathrm{Log}\, I$ scale. For moderate levels of illumination (10 cd/m^2) the Normann and Werblin data suggest a ratio of about 2. Of course, this value is for isolated receptor potentials in a non-mammalian species.

In man, there are two methods that can be used to estimated the gain ratio V_c^*/V_r^*. Although each method has a different theoretical underpinning, it is of interest that they both yield ratios similar to that of Normann and Werblin.

6.1.1. Detective Quantum Efficiency. The detective quantum efficiency (DQE) of a visual process is a ratio of the actual detecting ability of the eye to the maximum conceivable detecting ability. It is equivalent to the square of the ratio of the signal-to-noise of the output to the input signal-to-noise ratio. But if the V_r and V_c values in Eq. (21) are taken as signal-to-noise ratios of the output signal, and if the output noise (or background) is shared by both systems, then the ratio of (DQE) for the rods and cones will equal the ratio of the sensitivities of the two systems if the operating characteristic is plotted against $\log I$ (see Appendix 4).

Several authors have derived or measured DQE for a variety of conditions, mostly for rod vision (DeVries, 1943; Rose, 1948; Jones, 1959; Barlow, 1962; Hallett, 1969; van Meeteren, 1978). Of these, only Jones (1969) and van Meeteren (1978) provide data that allow DQE to be estimated for both rods and cones under similar conditions. Jones' values range from about 1.2 to 2.0 for peak DQE ratios of cones to rods, yielding an average value of 1.3 for V_c^*/V_r^*. Van Meeteren's data, using an entirely different paradigm, yield a V_c^*/V_r^* value of about 1.5 for retinal illuminances greater than 1 troland.

6.1.2. Brightness Scaling. A second method for estimating the relative sensitivities of the rod and cone systems is obtained from data for brightness scaling using Stevens' (1966) magnitude estimation technique. Consider Fig. 1 once more. The sensitivity of the operating characteristic is such that if I is increased by a factor k, then the response will increase by a factor k^g from its reference level, where g is the slope as indicated. The exponent in Stevens' power function is thus a measure of the gain of the brightness system, provided that the shared noise (or backgrounds) is held constant (see Appendix). If the noise increases with luminance, then the sensitivites will be overestimated, each by a constant amount. Thus, the ratios of the lowest exponents should be the most appropriate.

Data of Stevens and Stevens (1963) show two sets of brightness estimation data obtained under similar conditions, and one for scotopic and the other for photopic levels. The exponents are 1/3 and 1/2 respectively, yielding a sensitivity ratio g_c/g_r of 1.5. Since $g_c/g_r = V_c^*/V_r^*$, this estimate agrees well with that derived from the DQE, and will be used for V_c^*/V_r^*.

The remaining parameters[2] in Eq. (21) are β, and $\bar{\varrho}$. Values for the scattering coefficient β range from a few percent to over 50% depending upon the nature of the source and the angular distance of the retinal region from the source (Fry and Alpern, 1953; LeGrand, 1969). An average value for extended sources suggested by Wysecki and Stiles (1970) is 10%. This corresponds to the product $\beta_0\bar{\varrho}$, with a value of 40% for maximum scatter computed uniformly over a 10′

2 One neglected parameter has been a constant that corrects for the fact that the cones doe not sample a single image point, but an area described by the Stiles-Crawford effect. This constant depends upon pupil size and has a value of about 0.9 for photopic levels. Considering the accuracy of the approximations, this factor has been neglected, but does become important at very low light levels

region and attenuated as θ^{-2} elsewhere. With $\beta\bar{\varrho}=0.1$ and $V_c^*/V_r^*=1.5$, the reference gray reflectance, ϱ_0 is 22%. According to the Munsell gray scale, mid-gray is a reflectance of 19%. Direct measurements that I have made an several observers of this unique gray sensation yield experimental values of 16–20% for central 2 deg fields.

6.2. Effect of Overall Light Level

In sunlight and snow, the rod system is saturated and its sensitivity goes to zero. Under these conditions, the fraction V_c^*/V_r^* becomes very large and the value of ϱ_0 in Eq. (21) goes to zero regardless of the amount of scatter. Thus, the unique gray reflectance is a surface of near zero reflectance. Under these conditions, all surfaces must appear "whitish" or light gray.

The opposite extreme is at night, when only rod vision is active. Now the cone sensitivity is lost, and their saturation voltage, V_c^*, becomes very small, causing the denominator of Eq. (21) to go to zero. (It can never be negative.) The unique gray reflectance then takes on a value of 1.0, regardless of β. Thus a surface such as snow that is 100% reflecting must appear grayish, and all other surfaces will appear darker still[3]. This is the realization of the old adage that "At night, all cats are gray" – even Persion whites!

6.3. Effect of Spot Size

As the area of a light source changes, the sensitivities of the rod and cone systems will also change. As the spot size becomes smaller, the larger integrating area of the rod fields causes the rod sensitivity, g_r, to fall faster. Thus for a point source, $g_c > g_r$. This relation causes ϱ_0 to go to zero as the spot size goes to zero (especially in the dark, where β also goes to zero). Hence the unique gray reflectance is near zero, and all point sources against dark (or dim) backgrounds should appear "white" or luminous[3].

6.4. Abrupt Changes in Illumination Level

If neutral density filters are placed over our eyes, then the world appears darker. The effect is quite pronounced even if the overall retinal illumination is reduced by only a factor of two or three. Have our receptor activities changed appreciably by this light reduction? It seems unlikely, considering that a factor of three is a small fraction of a 10^7 dynamic range.

However, the rod and cone systems do not have the same kinetics, with the rods being much more sluggish (Sakitt, 1976). When the retina is dimmed, therefore, the restabilization of rod activity will generally take longer than that of the cones, and hence the rod activity will be greater than cone. In effect, the value of β has been raised, not by scattered light, but by a high "dark light" level for rods. The reflectance seen as unique gray will thus again approach 1.0, causing all surfaces to appear "dark". Following removal of the filter and subsequent light adaptation, the reverse occurs until adaptation is complete.

6.5. Increasing Scattered Light

Equation (21) is based upon the assumption that the light striking a region of the retina comes through the same aperture of the optical system for both the R and C images. However, if the side of the eyeball is illuminated by a penlight, then the internal scattered light level is raised considerably. The rod (R) signals will then be greatly elevated as compared with the cone (C) signals. The model then predicts that the field should become a darker gray, on the average, for only very highly reflecting patches will produce strong enough cone signals to compensate for the increased rod response. (The situation is similar to the change in rod-cone balance below photopic levels.) This "gray-out" of the visual field with scleral illumination is noted whenever the technique is used to cast shadows of the vessels onto the retina (Purkinje, 1825). However, to my knowledge, on one has previously explained why the visual field should appear so much darker.

7. Conclusions

Although I have not been able to prove that the lightness normalization problem is solved by comparing rod and cone activities, Eq. (21) does illustrate how a solution can be reached by such comparison. Three

3 Luminous objects: Eq. (21) describes how a reference gray level may be computed by the visual system. The equation places no constraints upon the relation between this reference reflectance, ϱ_0, and the gray sensation. Consider the possibility, suggested by Land (1965), that an internal gray scale is created by comparing ratios of intensity sensations across boundaries. The lightness sensation value, V, then becomes a function of ϱ_i/ϱ_0:

$$V=f(\varrho_i/\varrho_0).$$

Because the maximum reflectance is 1.0, there is an upper bound, $\lambda=1/\varrho_0$, which imposes a limit on V_{max} for all reflecting surfaces. (This is the sensation "white".) If λ can be specified independently, then all values of V greater than V_{max} must be luminous surfaces. Several independent sources of information can aid the observer in estimating λ. First, if the area is a source, then the distribution of scattered light falls off in a distinctive manner for angles away from this source, especially if it is a point (LeGrand, 1969). Second, experience can provide us with reasonable estimates of λ or V_{max}. Preliminary experiments suggest that the criterion adopted by observers varies widely, with some choosing a V_{max} generally associated with reflectances of 0.8–0.9. Thus, under some conditions, reflecting surfaces will appear luminous although they are not sources. Highlights may be one such example physiological constraints were used: 1) cones are directionally selective, rods are not; 2) the sensitivity of cone systems is much greater than that of rods; and 3) the rod and cone systems adjust their operating characteristics to their own equivalent backgrounds. In addition, two simplifying assumptions were made: 1) The distribution of scattered light in the optical system is constant, and 2) the reflectance of a scene on the average is constant. Finally, a major assumption was that there is a sensation of "unique gray" that is accessible to the observer and which corresponds to the condition of equal rod and cone activity (or perhaps more properly, equal S/N). With these assumptions and constraints, a patch having a standard reflectance can be identified, thereby normalizing the lightness scale[4]. Parameter estimation suggests such a reference patch will have a reflectivity of 20%, corresponding to a Munsell 5 value. Although some of the parameters are difficult to assess for everydays scenes, estimates obtained for a variety of special conditions do yield quite reasonable unique gray reflectances. Furthermore, in a qualitative way, the behavior of the human visual system is remarkably consistent with a very simple hypothesis for normalizing lightness scales.

Acknowledgements. The comments and criticisms of Drs. D. Marr, B. Sakitt, and S. Ullman were most helpful, with special thanks to DM for heading me toward vision understanding as a computational problem. This work was supported under NSF grant No 77-07569-MCS.

8. Appendix

8.1. Response of Receptor System: Naka-Rushton Relation

Naka and Rushton (1966) have suggested the following hyperbolic relation between the retinal response V and flash intensity I:

$$V/V^* = I/(I+\sigma), \tag{1}$$

where V^* is the saturation response and σ is a constant. Williams and Gale (1978) have shown that a considerable range of biophysical and psychophysical data may be interpreted by this response relation.

On a V-$\log I$ plot (Fig. 1), the slope of the response function will be

$$\frac{dV}{d(\log I)} = V(1 - V/V^*). \tag{2}$$

At the inflection point of the V-$\log I$ curve, the second derivative is zero. Differentiating (2) and setting the result to zero shows that this inflection occurs when the response is $V^*/2 = \sigma$, with a slope equal to $V^*/4$. Thus, at the point of inflection, the ratio of the maximum sensitivities of the cone and rod systems g_c/g_r will be proportional to their saturation voltages V_c^*/V_r^*.

8.2. Response of Receptor Aggregates: Power and Probability Considerations

The text analysis is confined solely to the situation where the image responses are from isolated receptors. Although it should be clear that the equivalent background (σ) normalization scheme can be applied to an aggregate of receptors of the same type, a more detailed elaboration of the σ-normalization procedure seems in order.

Let τ be the summation time over which the active receptors will contribute to the activity of a neuron. Within time τ, each receptor can be triggered at most M times, which is equivalent to the maximum number of quantal units (EPSP's) it may contribute to a neuron. Let n_i be the actual number of excitatory units contributed by the L^{th} receptor in the interval τ, and let k be the total number of receptors that feed the neuron.

Principle of Neuronal Equivalence (Assumption): At any level in a sensory system, within any neighborhood, all neurons of a similar functional class (i.e., that have inputs and outputs of a similar functional nature), will have equal activity on the average.

The purpose of this principle is to prohibit the condition where some neurons in a network always have unusually high or low activity. Such a condition is not only inefficient, but distracting and will lead to a change in the signal code for an identical stimulus presented in different regions of the network. The neural equivalence principle can be effected by adaptive mechanisms such as inhibitory feedback, cell size regulation, synaptic counts and strength, and clearly represents a memory state[5]. Prolonged changes in input (such as by deprivation or sensory adaptation) must then yield compensatory changes that return the average firing rate of a neuron to its neighborhood mean. Returning to the response of an aggregate receptor system, the total number of active inputs in time τ will be

$$\sum^{k} g n_i, \tag{3}$$

4 Clearly a similar solution can also be proposed for color normalization, but it is not necessary for the activities of all cone types to be compared with rod activity. One comparison suffices provided that the activities of the remaining types of cone responses are compared with the rod-normalized cone system (which can be $\bar{y}$)

5 A second method of effecting the principle is to build each neural level so that if the summation area of a second level neuron increases (thereby increasing its total input), then the number of second level neurons must increase at the same rate (thereby diluting the contribution of each input)

where g is the individual receptor sensitivity and n_i is a function of log intensity (see Fig. 1). However, by the principle of neuronal equivalence, the effectiveness of each input must be reduced in proportion to k. The average input to a neuron will then be

$$\frac{1}{k}\sum^{k}\frac{n_i}{k}=g\frac{\bar{n}}{k}. \tag{4}$$

Similarly, the average total number of inactive inputs will be $g(M-\bar{n})/k$.

Now the activity of a neuron reflects the power of its input. However, the principle of neural equivalence requires that the average firing rate of every neuron on the same level be the same. Thus, the signal is carried by the power fluctuations about the mean activity level (i.e., the mean level is discounted at the next higher level.) The total signal power of a neuron is thus the normalized sum of the RMS changes in active and inactive input:

$$P=\sum^{k}\frac{(gn_i/k-g\bar{n}/k)^2}{g\bar{n}/k}+\sum^{k}\frac{[g(m-n_i)/k-g(m-\bar{n})k]^2}{g(M-\bar{n})/k}, \tag{5}$$

$$P=\frac{g}{k}\sum^{k}\frac{(n_i-\bar{n})^2}{\bar{n}[1-\bar{n}/M]}. \tag{6}$$

If $\bar{n}\ll M$, then the power is a random variable with a χ^2 distribution, with a mean value $\bar{P}$ of

$$\bar{P}=\frac{g(k-1)\lambda}{k}\simeq g\lambda, \tag{7}$$

where λ is the mean of the Poisson probability function characterizing the delivery rate of the receptors' EPSP's. (Note that λ is proportional to log intensity for any adaptation level.)

The power of a neuron is thus proportional to the number of its inputs, as well as the sensitivity of the neuron (receptor) providing these inputs. The relative power of the rod and cone inputs, therefore, will be proportional to their relative sensitivities on a V-log$I$ plot as in Fig. 1. Equation (21) will thus remain valid regardless of the number of receptors that comprise the aggregate (or receptive field).

8.3. Weber Fraction

If the activity of a neural system has a χ^2 distribution, then the mean-to-sigma ratio will remain constant at $[(k-1)/2]^{1/2}$. In the case where a receptor system must decide whether the signal level has changed in a region, its task will be to compare P_{sn} with $P_{n'}$ with k remaining constant. Hence detectability will be a function of λ_{sn}/λ_n only. For a given level of detectability, therefore, λ_{sn}/λ_n will be constant, and hence the Weber fraction will also remain constant[6].

However, if the principle of neural equivalence holds, then it is not necessary for k to remain constant also, for Eq. (7) is independent of k, for large k. The number of receptor inputs may change with retinal eccentricity, for example, and the Weber fraction will still remain constant.

8.4. Detective Quantum Efficiency (DQE)

DQE is the ratio of the actual detecting ability of a system to the detecting ability of an ideal observer that makes optimum use of all of the signal collected (Jones, 1959). In terms of signal-to-noise ratios, DQE may be defined as follows:

$$\mathrm{DQE}=\left(\frac{\text{observed } s/n}{\text{ideal } s/n}\right)^2. \tag{8}$$

We must now obtain estimates for the signal and noise level where detection is optimal. In the experiments of Blackwell (1946) used by Jones (1959) to estimate DQE, test flashes were superimposed upon a steady background field. If adaptation were complete, then the average power in the detecting network must be near zero, invoking the principle of neuronal equivalence. Deviations from this value will arise from two sources: (1) external shot noise of the incoming photons; or (2) spontaneous internal activity. Because DQE reaches a maximum well beyond the Poisson-limited range of luminances with detection optimal where Weber's law holds, the major source of noise must be spontaneous internal activity, rather than photon noise. Thus the noise level P_n will be the same for both the average rod and cone systems.

Combining Eqs. (7) and (8), we obtain at a given luminance level:

$$\frac{\mathrm{DQE(cone)}}{\mathrm{DQE(rod)}}=\frac{P_{sn}(\mathrm{cone})/P_n}{P_{sn}(\mathrm{rod})P_n}=\frac{g_c}{g_r}=\frac{V_c^*}{V_r^*}. \tag{9}$$

Unfortunately, the DQE for both systems is not available at the same luminance level, because the most sensitive system sets the threshold. However, from Jones' (1959) calculations, it can be seen that in the region where either the rod or either the cone system limits detection, then DQE is roughly constant. Thus these values were extrapolated to obtain the separate rod and cone system DQE's at intermediate photopic ranges (see Fig. 4 of Jones, 1959).

6 Note that detectability will follow a power-law ROC

8.5. Power Law Exponent

Let the subjective sensation of brightness, B, be some function f of the receptor response, R:

$$B=f(R), \tag{10}$$

where $R=g\log I$, with g equal to the system sensitivity and I is the test spot luminance. Now the power-law result for brightness states that

$$B=I^e, \tag{11}$$

where e is an exponent that is different for rod and cone systems (approximately 1/3 and 1/2 respectively). Then

$$\frac{dB}{dI}=eI^{e-1}=e\frac{B}{I}. \tag{12}$$

But from (10)

$$\frac{dB}{dI}=f'(R)\frac{dR}{dI}=\frac{gf'(R)}{I}. \tag{13}$$

Combining (12) and (13) we obtain two equations, one for the cone and the other for the rod system:

$$\begin{aligned} g_c f'(R)&=e_c B_c,\\ g_r f'(R)&=e_r B_r. \end{aligned} \tag{14}$$

To compare the relation between these two brightness functions, divide one equation by the other to obtain

$$\frac{g_c}{g_r}=\frac{e_c}{e_r}, \tag{15}$$

Thus, the ratio of exponents will equal the ratio of sensitivities of the two systems, where sensitivity is defined by Fig. 1. But Appendix 1 shows that $g_c/g_r = V_c^*/V_r^*$ and hence the ratios of the exponents will also equal the ratio of the two saturation constants in text Eq. (21).

References

Aguilar, M., Stiles, W.S.: Saturation of rod mechanism of the retina at high light levels of stimulation. Opt. Acta **1**, 59–65 (1954)

Barlow, H.B.: Increment thresholds at low intensities considered as signal/noise discriminations. J. Physiol. **136**, 469–488 (1957)

Barlow, H.B.: Measurements of the quantum efficiency of discrimination in human scotopic vision. J. Physiol. **160**, 169–188 (1962)

Barlow, H.B., Sparrock, J.M.B.: The role of afterimages in dark adaptation. Science **144**, 1309–1314 (1964)

Blackwell, H.R., Blackwell, O.M.: Rod and cone receptor mechanisms in typical can atypical congenital achromatopsia. Vis. Res. **1**, 62–107 (1961)

Blick, D.W., MacLeod, D.I.A.: Rod threshold: influence of neighboring cones. Vis. Res. (in press)

Crawford, B.H.: Visual adaptation in relation to brief conditioning stimuli. Proc. R. Soc. London, Ser. B **134**, 283–302 (1949)

DeValois, R.L.: Analysis of response patterns of LGN cells. J. Opt. Soc. Am. **56**, 966–677 (1966)

DeVries, H.: The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. Physica **10**, 553–564 (1943)

Enoch, J.M., Hope, G.M.: Directional sensitivity of the foveal and parafoveal retina. Invest. Ophthalmol. **12**, 497–503 (1973)

Enoch, J.M., Laties, A.M.: An analysis of retinal receptor orientation. Invest. Ophthalmol. **10**, 959 (1971)

Fry, G.A., Alpern, M.: The effect of a peripheral glare source upon the apparent brightness of an object. J. Opt. Soc. Am. **43**, 189–195 (1953)

Fry, G.A., King, V.M.: The pupillary response and discomfort glare. J. Illum. Eng. Soc. **4**, 307–324 (1975)

Gouras, P.: Color opponency from fovea to striate cortex. Invest. Ophthalmol. **11**, 427–433 (1972)

Graham, C.H.: Vision and visual perception. New York: Wiley 1965

Hallett, P.E.: Quantum efficiency and false positive rate. J. Physiol. **202**, 421–436 (1969)

Horn, B.K.P.: Obtaining shape from shading information. In: The psychology of computer vision. Winston, P.H. (ed.), p. 115. New York: McGraw-Hill 1975

Hurvich, L.M., Jameson, D.: Some quantitative aspects of an opponent colors theory. J. Opt. Soc. Am. **45**, 602–616 (1955)

Jones, R.C.: Quantum efficiency of human vision. J. Opt. Soc. Am. **49**, 645–653 (1959)

Land, E.H.: The retinex. Am. Sci. **52**, 247–264 (1964)

Land, E.H., McCann, J.J.: Lightness and retinex theory. J. Opt. Soc. Am. **61**, 1–11 (1971)

Naka, K.J., Rushton, W.A.H.: *S*-potentials from colors within the retina of fish (*Cyprinidae*). J. Physiol. **185**, 536–555 (1966)

Laties, A.M.: Histological techniques for the study of receptor orientation. Tissue Cell **1**, 63 (1968)

LeGrand, Y.: Light, colour, and vision. London: Chapman & Hall 1968

Makous, W., Boothe, R.: Cones block signals from rods. Vis. Res. **14**, 285–294 (1974)

McCann, J.J., Benton, J.L.: Interaction of long-wave cones and the rods to produce color sensations. J. Opt. Soc. Am. **59**, 103–107 (1969)

Normann, R.A., Werblin, F.S.: Control of retinal sensitivity. I light and dark adaptation of vertebrate rods and cone. J. Gen. Physiol. **63**, 37–61 (1974)

Purkinje, J.: Beobachtungen und Versuche zur Physiologie der Sinne, Bd. 2. Berlin: Reiner 1825

Rose, A.: The sensitivity performance of the human eye on an absolute scale. J. Opt. Soc. Am. **38**, 196–208 (1942)

Rushton, W.A.H.: The retinal organization of vision in vertebrates, p. 12. In: Biological receptor mechanisms. Symp. of Soc. Exp. Biol. XVI. Cambridge: University Press 1962

Sakitt, B.: Psychophysical correlates of photoreceptor activity. Vis. Res. **16**, 129–140 (1976)

Stevens, S.S.: Duration, luminance, and the brightness exponent. Percept. Psychophys. **1**, 96–100 (1966)

Stevens, J.C., Stevens, S.S.: Brightness function: effects of adaptation. J. Opt. Soc. Am. **53**, 375–385 (1963)

Stiles, W.S.: Color vision: the approach through increment threshold sensitivity. Proc. Nat. Acad. Sci. **45**, 100–114 (1959)

Stiles, W.S., Crawford, B.H.: The luminous efficiency of rays entering the eye pupil at different points. Proc. R. Soc. London. Ser. B **112**, 428 (1933)

Ullman, S.: On the visual detection of light sources. Biol. Cybernetics **21**, 205–212 (1976)

Meeteren, van, A.: On the detective quantum efficiency of the human eye. Vis. Res. **18**, 257–267 (1978)

Wald, G.: Human vision and the spectrum. Science **101**, 653–658 (1945)

Westheimer, G.: Dependence of the magnitude of the Stiles-Crawford effect on retinal location. J. Physiol. **192**, 309 (1967)

Wiesel, T., Hubel, D.H.: Spatial and chromatic interaction in the lateral geniculate body of the rhesus monkey. J. Neurophysiol. **29**, 115 (1966)

Williams, T.P., Gale, J.G.: Compression of retinal responsivity: V-$\log I$ functions and increment thresholds. Vis. Res. **18**, 587–590 (1978)

Woodham, R.J.: Reflectance map techniques for analyzing surface defects in metal castings. Massachusetts Institute of Technology report AI-TR-457, 1978 (available from the Artificial Intelligence Laboratory)

Wyszecki, G., Stiles, W.S.: Color science: concepts and methods, quantitative data, and formulas. New York: Wiley 1967

Received: February 15, 1979

W. Richards
Department of Psychology
Massachusetts Institute of Technology
Cambridge, MA 02139, USA