ESSAY CONCERNING COLOR CONSTANCY*

Dorothea Jameson and Leo M. Hurvich

Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6196

CONTENTS

INTRODUCTION .............................................................................................. 1
  Current Interest ...................................................................................... 1
  Historical Overview .............................................................................. 3

RELEVANT VARIABLES ...............................................................................
  Visual Sensitivity ................................................................................... 5
  Chromatic Sensitivity ............................................................................ 6
  Neutral Adaptation and White Light ....................................................... 9
  Contrast, Assimilation and Receptive Fields ........................................... 11
  Postreceptoral Adaptation or Biasing ..................................................... 14
  Visual Cortex and Double-Opponent Cells ........................................... 15

REMARKS ON COMPUTATIONAL APPROACHES ........................................ 16

INTRODUCTION

Current Interest

Issues of constancy have arisen in the study of perception whenever the senses have been examined as information systems that mediate knowledge of the characteristics of the physical world. The kinds of constancy are manifold. They include mappings on sensory surfaces that are somehow converted into external space location maps so integrated that they serve not only as efficient and precise indicators of distance and direction information, but as mediators of sensorimotor integration and skill control as well (Guthrie et al. 1983; Jay & Sparks 1984). Objects, moreover, preserve their objective identities despite a

*This is the tenth in a series of prefatory chapters written by eminent senior psychologists.
variety of changes in the energy maps projected on the sensory receiving surfaces whether these changes relate to the object (or event) sizes, shapes, intensities, or qualities. At the same time, objects are systematically dependent, in terms of their phenomenal appearances, on the different local contexts in which they are embedded in the objective environment. Emphasis on only one or the other of these factors—i.e., identity preservation or context-dependent perception—tends to ignore the richness of the total information that our sensory mechanisms contribute to the cognitive systems that fashion the so-called real world as we know it.

Some of the current resurgence of interest in the constancy problem can be attributed to the advent of the computer in its various degrees of technological sophistication and processing capacity; the computer as a tool for testing intricately detailed hypotheses, as a tool for developing simulations, and as a would-be substitute for a human perceiver—i.e., a perceptual robot or a “task” robot with “machine vision.” The last use, or better, goal, is one that seems to encourage an understandable tendency toward oversimplification of the problem. We would have no quarrel with such a tendency if it were not that the oversimplification that might be useful for the machine as tool to perform specific tasks in the robot context is somehow carried over into the analysis of perception, even though the oversimplification distorts the nature of the human perceptual problem.

In a 1986 issue of the Journal of the Optical Society, a Feature Section was devoted to computational approaches to color vision. In his introduction the feature editor explicitly recognized the focus on the machine task aspect of the approach. “If we want to address robots in higher-level languages that we understand about objects, we must make them see the way we do” (Krauskopf 1986). Eight of the twelve feature papers dealt with color constancy, and the primary aim was to find algorithms or computational approaches that would yield means for deriving constant surface reflectance properties of objects for different and initially unknown illuminants. Apart from its use in the computation of surface reflectance characteristics for object recognition, perceptual information about the different conditions of illumination as relevant in its own right was largely if not totally ignored in these papers. Our own judgment is that human visual systems (including both higher- and lower-order processes) are likely to have evolved a design that provides perceptual information about change as well as constancy—about light, weather, and time of day, as well as about the relatively constant physical properties of mainly opaque objects within a scene. To what extent current technology

---

1 The phrase “computational vision” is already gaining wide currency (e.g., Boynton 1988).
2 Stanford University has applied for a patent based on research directed toward this goal, a fact that emphasizes its obvious relation to machine design (Maloney & Wandell 1986).
might or might not find such information relevant for the tasks of present-day or near-future robots we are not prepared to say, but that luminosity information, as well as opaque surface information, is relevant for biological organisms of human or other species is an assumption that we are prepared to make.

**Historical Overview**

Early experimental studies of the perceptual constancies typically examined the degree of constancy manifested under different conditions. [For an integrative theoretical discussion of the various constancies, with an emphasis on Gestalt principles, see Koffka (1935).] With respect to color (including both the chromatic and achromatic brightness or lightness properties), the central issue was the same as it is today. How can the surface reflectance characteristics of the distal object be recovered to achieve an approximately constant surface percept despite the fact that the retinal image of the object depends on both its surface reflectance \( R \) and the incident illumination \( I \), \( R \times I \), when \( R \) is constant but \( I \) is both unknown and changes from one situation to the next? Helmholtz’s conjecture was both best known and most widely accepted. In his text on experimental psychology, Woodworth (1938) included Helmholtz’s own statement of his view, which we quote here.\(^3\)

> Colors are mainly important for us as properties of objects and as means of identifying objects. In visual observation we constantly aim to reach a judgment on the object colors and to eliminate differences of illumination. So, we clearly distinguish between a white sheet of paper in weak illumination and a gray sheet in strong illumination. We have abundant opportunity to examine the same object colors in full sunlight, in the blue light from the clear sky, and the reddish yellow light of the sinking sun or of candlelight—not to mention the colored reflections from surrounding objects. Seeing the same objects under these different illuminations, we learn to get a correct idea of the object colors in spite of difference of illumination. We learn to judge how such an object would look in white light, and since our interest lies entirely in the object color, we become unconscious of the sensations on which the judgment rests.

Woodworth also cites Hering’s views on color constancy. Hering, not surprisingly, disagreed with Helmholtz’s analysis. He called attention to the various peripheral factors (pupillary changes, retinal adaptation, and physiological contrast mechanisms) that actually must alter the sensory effects of visual stimulation under different conditions of illumination, and that, with continued visual experience, Hering thought would also alter the state of the central mechanisms involved in perception—the kinds of changes we would today refer to as visual plasticity. Hering’s view led directly to his concept of “memory color.”

\(^3\)The statement quoted here from Woodworth was abstracted by him from the first edition of Helmholtz’s *Physiological Optics* (1866, p. 408). In Southall’s (1924) English translation of the third edition, Helmholtz’s discussion appears in Volume 2, pp. 286–87.
The color in which we have most consistently seen an external object is impressed indelibly on our memory and becomes a fixed property of the memory image. What the layman calls the real color of an object is a color of the object that has become fixed, as it were, in his memory; I should like to call it the memory color of the object. . . . Moreover, the memory color of the object need not be rigorously fixed but can have a certain range of variation depending on its derivation. . . . All objects that are already known to us from experience, or that we regard as familiar by their color, we see through the spectacles of memory color.

(Hering 1920)

Woodworth’s summary chapter on the perception of color captures the flavor of the experimental work on approximate color constancy during the period between 1900 and the late 1930s. Typically, the experiments were designed to determine the degree of lightness constancy for various conditions, sometimes by sample matches made to a display of surfaces of different reflectances under different levels of illumination and/or shadow conditions, sometimes by matches made between rotating disks of various average reflectances. The term “albedo” came into common use as the relative reflectance index, and the measure in these experiments was the degree to which the albedo determined the visual matches for the different conditions. Arithmetic (Brunswik) or logarithmic (Thouless) ratios were developed to express the departures of the experimental matches from those predicted for perfect lightness constancy. Ordinarily the data fell somewhere between perfect retinal image light matches and perfect object constancy, although occasionally overcompensation for illumination differences was observed. Considerable effort was devoted to determining the efficacy of various cues for judging illumination, a requirement, in the Helmholtz context, for solving the reflectance problem; and in the same context, measures were compared for children of various ages. For the most part, children did not seem very different from adults, although the results differed for different experiments and were particularly susceptible to effects of instructions. Instructions have always been recognized as crucial in such experiments (MacLeod 1932; Katz 1935; Hurvich & Jameson 1966), and they continue to recur as an experimental variable (Arend & Reeves 1986; Arend & Goldstein 1987). The extremes can best be summarized by the difference between making an adjustment to make a particular part of a display look identical to the same area in a differently illuminated display, as contrasted with an adjustment to make a particular surface in a display seem identical in its surface characteristics to the same object in a differently illuminated display. Behavioral experiments on nonhumans used “identification” as indexed by a trained response, and these results, too, suggested that fish and primates are able to identify objects in different illuminations in terms of their surface reflectances.

Not included in Woodworth’s summary was the classical experiment of Hess & Pretori (1894). Although their aim was to measure the effects of
brightness contrast between two (adjacent) center/surround displays, their results can readily be analyzed in constancy terms. That is, for a center area of one reflectance and a surround of different reflectance, a uniform increase in illumination would produce a proportional increase in light reflected from each surface, and the ratio of reflected light of center-to-surround in the retinal images would remain unchanged despite the proportional increase in each. The measured contrast ratio for the matched area in the center/surround comparison display would also be constant if the observers were exhibiting perfect lightness constancy. Our own replot of the Hess & Pretori data (Jameson & Hurvich 1964, 1970) shows that their observations encompass a range of findings that depend systematically on the surround-to-center contrast ratio of each test display. When this ratio is low (equivalent to surround reflectance lower than center’s), the center appears to increase in perceived brightness as center and surround are both increased proportionally in illumination; as the contrast ratio is made higher (equivalent to surround reflectance higher than center’s), the center appearance approaches constancy; and as the contrast ratio is made still higher (equivalent to surround reflectance much higher than center’s), the dark center appears to become blacker with proportional increase in illumination of both center and surround. We have reported findings similar to these for a patterned array of different achromatic patches, and cite in our report concordant results from other laboratories (Jameson & Hurvich 1961a, 1964).

RELEVANT VARIABLES

Visual Sensitivity

Light sensitivity is so well known to be controlled by the level of illumination to which one is adapted that it hardly needs documentation here. Although it is most often illustrated by the dark-adaptation curve, for relevance to the important constancy issue, only the photopic segment of that threshold sensitivity function, the cone region, describes the course of sensitivity recovery of interest. Moreover, the reflection of this recovery function, which shows the increasing threshold energy requirement with increase in level of background light, makes clear the decrement in light sensitivity with increase in adaptation level. Because the visual response depends on the product of stimulus × sensitivity, a major part of the compensation for illumination changes (in addition to the small contribution of pupillary changes) obviously occurs at a very peripheral level, and largely in the retinal light receptors. For the range of adaptation levels within which Weber’s law holds, it is often assumed that contrast sensitivity (and by extension, suprathreshold contrast perception) will be constant, and thus account for perceived lightness constancy. It is essentially another statement of the ratio hypothesis proposed by Wallach (1948). Were this a perfectly compensatory mechanism, then there
would certainly be no need for experience with, or judgments of, different levels of illumination, because their effects, at least for uniform illuminations and diffuse object surfaces, would never be registered at all beyond the most peripheral level of the visual system. But the situation is not quite this simple. If contrast sensitivity is measured with sinewave stimuli as a function of spatial frequency to determine the human contrast-sensitivity function, both the level and the form of this function change with average level of illumination. This dependence on illumination has important implications for visual perception; it is one of the findings that make it most unlikely that form perception depends on a straightforward Fourier processing of a scene by the visual system (Kelly & Burbeck 1984). It also suggests that all sharply focused edges between surfaces of different reflectances will not appear equally sharp at different light levels. Kelly & Burbeck believe that at low spatial frequencies, contrast sensitivity is closely related to mechanisms of lateral inhibition, which are spatially more diffuse than the excitatory processes. The dependence of the effectiveness of such mechanisms on illumination level is consistent with our own long-held conviction that visual adaptation must involve postreceptoral changes as well as receptoral sensitivity adjustments (Hurvich & Jameson 1958, 1960, 1961, 1966; Jameson 1985; Jameson & Hurvich 1956, 1959, 1961b, 1964, 1970, 1972; Varner et al 1984).

**Chromatic Sensitivity**

In 1905, von Kries made an analysis of the way the visual system might compensate for changes in the spectral quality of illumination to make it possible to identify object colors, and proposed that the three different spectrally selective mechanisms of the retina (cone types) suffer relative decrements in overall light sensitivity in proportion to the relative strengths of their individual stimulation by the prevailing illumination. This analysis is qualitatively consistent with the way both the threshold and suprathreshold spectral luminosity functions vary in form with chromatic adaptation (Jameson & Hurvich 1953; Hurvich & Jameson 1954). Thus, for example, exposure to longwave light selectively reduces light sensitivity in the same region of the spectrum, as it should if the contribution of the longwave cone signal to light sensitivity were reduced in amplitude. However, von Kries’s postulated changes in the balance of sensitivities would not change the forms of the three individual wavelength vs receptor sensitivity functions, but only their amplitudes; hence additive color matches that depend on the selective absorptions of the three different cone pigments would be unaffected by the sensitivity adjustments. Within reasonable limits, such matches are so unaffected, but only if the state of adaptation is uniform throughout retinal image areas of both the test and matching fields.
If this is not the case and the matches are “asymmetric” (for example between test field in one eye for one state of chromatic adaptation and matching field in the other eye for a different adaptation), then differences in responsiveness between the two states of adaptation can be registered by changes in the proportions of the matching lights. Such asymmetric color matches make it clear that the von Kries rule of linear, proportional changes in amplitude of receptor sensitivities cannot account for all the data (Hurvich & Jameson 1958; Jameson & Hurvich 1972). Departures from this rule are systematic. That is, the measured changes in proportions of the matching lights vary systematically with the luminance level of the test field relative to the surround luminance to which the eye is adapted. The departures from the proportionality rule are, moreover, in the directions that would have been predicted from a nonconstancy phenomenon known as the Helson-Judd effect (Helson 1938; Judd 1940). Spectrally nonselective surfaces seen against a spectrally nonselective background all appear achromatic (white through grays to black) in white light. When illuminated by chromatic light, samples whose reflectances are near the background level continue to appear gray, those above the background level take on the hue of the illuminant, and those below take on a hue that is complementary to that of the illuminant. Hue shifts for chromatic samples tend to behave similarly—i.e. as if intermixed with the illuminant hue or with its complementary, depending on the relative reflectances of sample and background. Such departures from perfect color constancy with changes in spectral quality of illumination are reminiscent of those described above for lightness constancy, and both sets of phenomena imply that perceived contrast between objects of different surface reflectance varies with the level and kind of illumination in which they are seen and to which the visual system is adapted.

We should emphasize that the magnitudes of these perceptual changes are not so great as usually to prevent object identification by color, particularly for distinctly colored surfaces that, under most ordinary illuminants, undergo perceived hue, saturation, or brightness shifts that still do not move them out of one color category and into another, which would certainly be the case were there no compensatory changes in visual sensitivities (Jameson 1983). For particular kinds of arrays that contain strong colors, but with subtle color differences, however, the state of chromatic adaptation can make the difference between seeing a pattern and failing to perceive that the surface is anything but uniformly colored. This statement is based on our studies of wavelength discrimination for test lights viewed within surrounds to which the observer is adapted (Hurvich & Jameson 1961). The consequences of chromatic adaptation for such discriminations are not a priori obvious. It might be anticipated that exposure, for example, to longwave light, which reduces the sensitivity of the longwave receptor, would selectively impair
discriminability between just detectably different long wavelengths. Instead, the opposite occurs. In the longwave spectral region, the threshold wavelength difference is actually decreased; thus discriminability is improved, relative to what it is for neutral adaptation. And conversely, wavelength discrimination is relatively impaired in the midwave spectral region (where light sensitivity remains high). Qualitatively, what happens in this situation is that the perceived redness is somewhat depressed in the longer wavelengths, making slight differences in the yellowness of these same lights more obvious, whereas the perceived midspectrum greenness is enhanced, and tends to mask slight differences in the yellowness of these lights that can be detected reliably in a neutral state of adaptation. Changes analogous to these occur in other spectral regions for other kinds of chromatic adaptation.

Anyone who has had the opportunity to observe paintings hung in the same surroundings in both daylight and at night under incandescent illumination is likely to be aware of the disappearance or enhancement of such subtle hue differences. In most of these situations, the state of chromatic adaptation is probably determined primarily by the spectral quality of the illuminant. This assumes that the different surfaces in the field of view will be sufficiently varied so that the space average reflectance will not be far from neutral or spectrally nonselective. Some paintings, of course, are sufficiently large so that only the gamut of reflectances within the painting itself enter, with the illuminant, to affect the adaptation state. But here too the discriminability of similar hues and saturations will be dependent on the quality as well as level of the average light reflected from the surface area within the field of view as one inspects the painting, and it will differ for different illuminants.

In connection with paintings of the sort just mentioned, it should be pointed out that even for a very large painting that is very nearly monochromatic, such as Ad Reinhart’s canvas called “Red Painting” (red geometric figure against red background, 6.5 ft by 12.5 ft), which hangs in the Metropolitan Museum of Art in New York, continued inspection of the painting does not rob it of its redness and transform it into a gray painting. Fortunately for the artist, chromatic adaptation of the von Kries sort need not be complete; that is, the balance of sensitivities need not be completely compensatory so that the space average product of the reflectances × illumination yields a neutral or achromatic response. For highly selective reflectances or illuminants this is seldom the case; rather, the sensitivity balance only partially compensates for the effective adapting light rather than completely compensating for it. In brief, there are degrees of chromatic adaptation (Jameson & Hurvich 1956), as well as degrees of light and dark adaptation. Complete adaptation to strongly chromatic light is a special case; it does occur in a so-called Ganzfeld situation—that is, when the eye is exposed to completely uniform illumination throughout the entire surface of the retina (Hochberg et al 1951). With
prolonged exposure to a Ganzfeld, all visual effect of light fades away and we become, as it were, sightless.

**SURFACE METAMERS** Surface metamers constitute a special case of an illuminant-dependent departure from color constancy. This case comes about because the appearance of a given surface material under one illuminant can be precisely matched for the same illuminant by a variety of different dyes and paint mixtures used to color other material samples. The spectral reflectance distributions of the samples can differ markedly, but the samples are visually identical. Such surfaces are thus, by definition, all surface-color metamers for this one illuminant. If the illuminant is changed, the surface color matches no longer hold. The different samples take on different hues and saturations that deviate, one from the next, in directions and amounts that are governed by their particular spectral reflectances in relation to the spectral characteristics of the new illuminant. [For a detailed technical discussion of surface metamers, illuminants, and distortion transformations, see Wyszecki & Stiles (1967).] The color changes cannot be predicted without a priori knowledge of the spectral distributions that are involved, but, in general, they will be more significant the more irregular the spectral reflectance and illuminance distributions. With the increased use of fluorescent light sources that contain localized spectral energy peaks, the so-called "color rendering" properties of illuminants have required the increased attention of lighting engineers and illuminant manufacturers. Visual mechanisms of color adaptation do not, even in principle, solve this problem caused by illuminant energy peaks and high degrees of surface-color metamerism.

**Neutral Adaptation and White Light**

Chromatic adaptation is, by commonsense definition, measured as a departure from adaptation to white light. By common sense as well, white light is light that looks white or achromatic. But what looks white or achromatic is, quite obviously, any one of a variety of very different spectral distributions depending on other variables in the viewing situation. Consider only one series of such illuminants whose energy varies smoothly and systematically across the visible spectrum in a way that nearly parallels the energy output of a physicist’s ideal black body raised to increasing temperatures. Such illuminants are characterized by so-called color temperatures (kelvin, K); lights of high color temperatures (such as light from the north sky, about 10,000 K) have their energy output more heavily weighted in the short wavelengths, whereas artificial incandescent light of the sort used for indoor illumination (2400–2800 K) is relatively impoverished in shortwave energy but has comparatively high energy output in the longwave region of the visible spectrum. Illumina-
tion that is a mixture of skylight and noonday sunlight (color temperature of about 5500 K) has a relatively balanced energy distribution. This continuum of illuminants is approximately the one referred to in Helmholtz's statement that "we have abundant opportunity to examine the same object color in full sunlight, in the blue light from the clear sky, and the reddish-yellow light of the sinking sun or of candlelight . . .", although our incandescent lights are less "reddish-yellow" than either the sinking sun or Helmholtz's candlelight.

In controlled laboratory test situations, uniform light fields from this whole gamut of color temperatures can be perceived as white light, but the perception depends on a multiplicity of interacting variables that include level of total light energy, exposure duration, area of light field, and prior light exposure (Hurvich & Jameson 1951a,b; Jameson & Hurvich 1951b). The gamut of color temperatures perceived as white increases with energy level whatever the parametric value of each of the other variables. That is, at high levels the relatively desaturated blue or yellow hues seen at lower light levels are somehow veiled or weakened. Since the cone system adapts rapidly, chromatic adaptation might well be a contributing factor responsible for the neutral percept for all illuminants except the one that approximates an equal-energy distribution. The latter illuminant (with some individual variation probably due to differences in ocular media) in our experiments had no perceptible hue at any energy level for any of the exposure durations or field sizes we examined. Results of other experiments designed to test for chromatic adaptation effects were consistent with the conclusion that it is only a near equal-energy illuminant that leaves the visual system in a neutrally balanced equilibrium state (Jameson & Hurvich 1951a).

For opaque surfaces of spectrally nonselective reflectances, it is only for conditions that produce such a physiologically neutral equilibrium state of adaptation that all gray-scale levels of the nonselective surfaces can be expected to appear equally achromatic as whites through grays to blacks. Illuminants that produce other adaptation states will alter the perceived neutrality in accord with the Helson-Judd effect, tinting the lighter samples toward the illuminant hue and the darker ones toward its complementary. The extent of the perceived departures from strict neutrality of such surface colors will be minimal for illuminants very similar to the physiologically neutral one, and increasingly more noticeable for illuminants that are more heavily weighted toward one or another end of the spectrum. If the visual scene includes a variety of spectrally selective as well as nonselective surfaces, then the neutral or nonneutral appearances of the latter will further depend on the other surfaces in the array. In addition to illumination and reflectance characteristics, additional variables such as size and proximity become relevant for all the perceived surface colors.
Contrast, Assimilation, and Receptive Fields

The systematic departures from color constancy that carry information about illumination are essentially color contrast effects. They include both (a) brightness or lightness contrast that accentuates the perceived difference between the lightest and darkest objects or reflectances as mentioned above for both surfaces (Jameson & Hurvich 1961a, 1964) and sinewave gratings (Kelly & Burbeck 1984), and (b) color contrast that accentuates perceived differences in the complementary yellow-to-blue and red-to-green hue dimensions (Jameson & Hurvich 1961b). In retinal images of natural scenes that contain three-dimensional objects and surface reflectances made up of both specular and diffuse components, contrast accentuates the differences between highlight and shadow, and contributes to the three-dimensionality of the scene, even if the image is not of the scene itself but of a two-dimensional photographic display. Shadowing is so effective a cue for three-dimensional shape that even shadowing that is produced by border contrast, rather than a gradation in either illumination or reflectance, can result in perceived depth variations across a perfectly flat surface. A good example is the familiar Mach scallop or fluted effect that perceptually “curves” adjacent edges forward and back into the surface plane when one views contiguous rectangular samples of a gray scale that is regularly ordered from light to dark.

Lateral interactions are common to the anatomy and neurophysiology of visual systems. Although at least in some species there may be contact influences that spread across the retinal receptor layer itself, in primates and thus probably also in humans, the more significant lateral interactions seem to occur at postreceptoral levels. In the color processing system, the three-variable spectral analysis of retinal image light occurs, as it were, in three parallel classes of cone receptors, each with a characteristic spectral sensitivity determined by its particular cone photopigment. Light absorption is signalled by graded hyperpolarizing electrical responses in each cone class, and gives rise to synaptic changes that result, ultimately, in postreceptoral “neural images.”

A significant recombination in the color processing system involves a transformation from the three different light absorption maps of the receptor mosaic that yields another set of three maps essentially based on a set of three different sums and differences governed by the signal strengths in the different receptor types. In our model based on psychophysical evidence (see Hurvich 1981), one of the neural systems is activated in accord with a difference between the weighted signal strengths of the midwave-sensitive receptor and the summed short- and longwave-sensitive receptors, a second in accord with a difference between the weighted signal strengths of the shortwave-sensitive receptor and the summed mid- and longwave-sensitive recep-
tors, and a third in accord with the weighted signal strengths of the signals summed from all three receptor types. It should be noted here that opponent neural processing as fundamental to color vision is by now universally accepted, but the specific models proposed by different investigators differ in their detailed formulations. A recent computational proposal (not yet implemented by experiment) suggests use of sinewave spectral power distributions to most efficiently evaluate a subset of these formulations, including our own (Benzschawel et al 1986). All models require differencing mechanisms for hue processing, in accord with Hering's original hypothesis. The three overlapping spectral separations achieved by the selective photopigments are thus sharpened in the two differencing systems of the neural map, and essentially lost in the third. But since this spectral sharpening requires neural activation related to more than any single one of the adjacent cones, it comes at the expense of the spatial discreteness potentially available at the retinal receptor level. Thus the effective spatial grain in the neural maps is necessarily coarsened relative to that of the individual cones of the retinal mosaic.

Spatial, simultaneous color contrast has been a recognized characteristic of perception since at least the time of Leonardo da Vinci, and it has been exploited by artists who often exaggerate both hue and brightness contrast for pictorial effect (Jameson & Hurvich 1975). Because of contrast, any formal process expression for perceived color for a specified retinal light image array must include not only (a) the spectral sensitivities of the three classes of photopic light receptors, (b) coefficients to express the amplitude balance these receptors brought about by adaptation of the von Kries type, and (c) the interactions that give rise to the difference and sum functions that characterize spectral opponent processing in the neural image, but also (d) the mutual lateral neural interactions that occur within each class of the triplex of processing systems at this level (Jameson & Hurvich 1959). The effects of the latter are readily measured by perceptual scaling techniques and by color matches made to individual, uniform samples within an array compared with matches to the same samples in the presence of parts or all of the remaining array. Quantitative modeling of the effects by simultaneous equations that include spatial terms can describe them to a rough approximation (Jameson & Hurvich 1961b, 1964), but a physiologically more realistic model, and one that intrinsically subsumes more spatial variables, involves filtering by a difference of Gaussians (DOG) at the opponent neural level. Such functions are idealized representations of neural receptive fields of the circularly symmetric, spatially antagonistic, center/surround type. Psychophysically determined threshold interaction effects have been used to estimate the critical spatial dimensions within which only excitatory summative effects (receptive field center effects) occur within a small central foveal region of the visual field (Westheimer 1967). When such estimates are compared with those
derived from other kinds of psychophysical experiments, such as measures of
sinewave contrast sensitivity that typically involve larger retinal areas, there
are differences in calculated receptive field center diameters, although the
shapes of the derived sensitivity profiles are very similar (Kelly & Burbeck
1984). The nonhomogeneity of the receptor mosaic—that is, the decline in
numbers of cones per unit area from fovea to periphery of the retina (and
corresponding decline in numbers of related postreceptoral cells)—is accom-
panied by expansion of receptive field center diameters with increasing
distance from the foveal projection; but there is also considerable size vari-
ation within any particular projection area (Hubel & Wiesel 1960). Thus, the
spatial grain of the neural maps, although coarser, follows the grain of the
retinal receptor mosaic, but in a graded band, so to speak, rather than being
singularly determined by retinal location.

Spatial mixture and blending of hue and/or lightness are effects that are
opposite to border contrast since they reduce, rather than accentuate, differ-
ences in contiguous image areas. In our own analyses of these phenomena,
the variation in receptive field size within a particular locus referred to above
has seemed to provide the kind of physiological basis needed to account for
the fact that both sharp edges between adjacent image areas and apparent
spreading of different hues across the image boundaries can occur. Such
effects, variously called assimilation or spreading, are particularly striking in
repetitive patterns whether striped or curvilinear, and they can readily be
observed in decorative fabrics and other motifs as well as in the paintings of
some contemporary artists (Jameson & Hurvich 1975).

What is seen in such patterns depends on the sizes of the uniform elements
within the pattern imaged on the retina relative to the cone diameters, and to
the diameters of both the center and surround regions of the related neural
receptive fields. If the image elements are small relative to the cone diamet-
ers, then true spatial light mixture occurs; if they are small relative to the
receptive field centers, then some degree of spatial blending or assimilation
occurs; and if they are larger, then assimilation gives way to spatial contrast.
These changes can be observed most easily by decreasing or increasing
viewing distance from the pattern, thus controlling the relative sizes by
increasing or decreasing, respectively, the width (in the stripe example) of the
pattern elements in the retinal image. In this case, color constancy fails with
change in distance: For example, stripes that are seen close up as red
alternating with blue become increasingly reddish purple and bluish purple
stripes farther away. Complete light mixture with failure of spatial resolution
requires very distant viewing. Far enough away, a striped pattern can look
uniform. It is the intermediate range that is of most interest, because here
there is both good pattern resolution and partial hue mixture. Also, at just the
right distance within the intermediate range, it is possible to attend to the
striped field as a whole and see the stripes as reddish purple and bluish purple, or, alternatively, to concentrate on the adjacent stripes at the center of gaze (where the receptive fields are smallest in the foveal region) and see them as vividly red and blue with no trace of the purple mixture hue. To the casual viewer, the nonconstancy of adjacent stripe color that can occur when scanning such a pattern at the critical viewing distance is usually not noticed as such without deliberate attention, but what is noticed is a kind of visual liveliness that fabric designers sometimes strive for.

Since resolution and mixture depend on neural receptive field center sizes, the fact that, for some retinal image dimensions, both can occur simultaneously and at the same location suggests that the two effects result from processing in different neural systems with different receptive field dimensions; and indeed, process modeling using scaled receptive field (DOG) filtering gives a good qualitative match to the perceptual effects (Jameson 1985). Receptive fields of different scales are used commonly in computational models, and their dimensions have typically been based on analyses of psychophysical data indicating that sinewave contrast sensitivity requires a number of different spatial processing "channels" for different regions of the spatial frequency dimension. [A good critical summary and relevant references can be found in Kelly & Burbeck (1984).] It is also concluded from the dependence sinewave contrast sensitivity on luminance level that the effectiveness of the inhibitory surround region of receptive fields is decreased at low luminances and increased at higher ones. Thus, the relative effectiveness of the mutual lateral interactions that give rise to spatial contrast both at edges and across more extended retinal image areas (see von Békésy 1968) would be expected to vary with luminance in the same way and provide a physiological basis for the perceived increase in object color contrast in bright light.

Postreceptoral Adaptation or Biasing

It seems clear that change in the spectral quality and quantity of the adapting illuminant not only changes the balance of sensitivities at the receptor level, but that it also changes the balance of excitatory and inhibitory influences that are related to both spectral and spatial processing in the color related systems at the postreceptoral level. In addition to the evidence from our own studies of asymmetric color matches, perceptual scaling data, and discrimination functions discussed earlier in this essay, and the evidence from sinewave contrast functions mentioned above, additional evidence for the involvement of postreceptoral mechanisms comes from a very different experimental and analytical paradigm. This paradigm is the two-color increment threshold technique employed in the many exemplary experiments and analyses carried out by W. S. Stiles. Pugh & Kirk (1986) have published a comprehensive
historical review of this work, including references to others (among whom Pugh was an important contributor), that outlines the changes in Stiles's own interpretation of such discrimination thresholds and provides the basis for the current interpretation that the mechanism for adaptation to the background light in this paradigm cannot be restricted exclusively to the triplex of retinal light receptors but must also involve postreceptoral adaptation effects in the neural differencing mechanisms—i.e. at the spectrally opponent level of neural color processing. In their review, the authors emphasize that, although Stiles had started from the hypothesis that analysis of his psychophysical data would reveal activities and adaptation effects only in the cones, by 1967 he himself pointed out that difference signals may also make an important contribution to the discriminations in his experimental paradigm.4

We do not intend to imply here that the postreceptoral influences envisaged by all investigators concerned with this issue are necessarily identical with those that we have hypothesized to account for a variety of different psychophysical and perceptual findings. For example, D'Zmura & Lennie (1986) postulate variable weights that are adaptation-dependent applied to the adaptation-scaled cone signals at the differencing level. Whether their specific formulation would yield effects at the cortical level equivalent to our postulated postreceptoral, incremental or decremental, equilibrium level or set-point shifts that depend on lateral opponent interactions is not directly evident. Their discussion of physiological mechanisms leaves uncertain the level (or levels) of neural processing at which the postreceptoral adaptation effects occur (as does our own model of these effects), and even includes an expression of uncertainty about whether the kinds of adjustments to scaled cone signals that they postulate for their second stage are actually made by the visual system. Clearly, independent evidence on this issue from visual neurophysiology is both lacking and needed. Some of our own psychophysical experiments that compare adaptation to steady light fields with adaptation to the same lights for an equivalent duration but with interpolated dark intervals that permit partial recovery of cone sensitivity have led us to the conclusion that postreceptoral mechanisms (at some level) recover from chromatic adaptation shifts very slowly before the neutral equilibrium level is restored (Jameson et al 1979). Such relatively long-term biasing suggests a potential contribution to the adaptation effects at processing levels as far removed from the retinal receptors as the visual projection areas of the cortex.

Visual Cortex and Double-Opponent Cells

Cells that show opponent spectral characteristics are known to exist in the primate all the way from the retina, through the lateral geniculate nucleus

4For an earlier suggestion that this might be so, see Hurvich 1963.
(LGN), to various cortical projection areas. Although cortical cells in area 17 and beyond usually have receptive fields that are organized in such a way that the cells are preferentially sensitive to lines and edges with particular orientations, some of which have been reported also to be spectrally selective and opponent, there are also cortical cells with circularly symmetric receptive fields that are characterized by spectral opponency both in the centers and in the antagonistic surrounds (De Valois et al 1982; Michael 1978a,b; Jameson 1985 for additional references). Recent work by Livingstone & Hubel (1984; Hubel & Livingstone 1987) has localized such cells, thought to be related to the parvocellular system of the LGN, in cluster-like formations, *blobs*, in area 17, and has suggested that these double-opponent blob cells feed into *thin stripe* formations in area 18, from which there are also anatomical connections back to area 17 as well as with other visual projection areas. Such double-opponent cells conveniently display characteristics similar to the difference-of-Gaussians receptive fields combined with spectral differencing for two hue systems and broadband spectral sensitivities for an achromatic system, which are consistent with our interpretations of psychophysical and perceptual data. Despite this convenient convergence, we do not intend to imply either that these are *the* relevant physiological findings for neural color processing or that our own analyses are anything but oversimplified and incomplete. It is with this caveat, and the further caveat that these are certainly not the only collections of cells or brain areas involved, that they are included in the digest shown in Table 1. The suggestion in this digest that the connections to area 17 from area 18 as well as from 17 to 18 might be related to changes in state related to the establishment of “memory color” is our own speculation, and it is no more than that. Interconnections with other subdivisions and other brain areas would certainly be required for colors of particular hue categories to be regularly associated with objects of particular forms and particular contexts.

From the point of view of understanding visual perception, or even a circumscribed aspect of the mechanism such as color processing, in terms of visual neurophysiology, we are barely at the starting line ready for the first halting step. From a perspective of 20 or more years back, progress in visual neurophysiology has been rapid and impressive. But examined from today’s perspective, the missing details and the nearly totally unexplored functional specializations of the different relevant brain areas, as well as of their mutual interrelations, loom even more impressively large.

**REMARKS ON COMPUTATIONAL APPROACHES**

We mentioned in the introductory paragraphs of this essay that issues related to object color constancy are a common focus of computational approaches; in
Table 1  Some relevant aspects of color processing

<table>
<thead>
<tr>
<th>Retinal light stimulus</th>
<th>Space (and time) average of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>von Kries adaptation (proportionality rule weighted for degree of adaptation)</td>
<td>Direct light $\times$ surface reflectances (specular and diffuse components)</td>
</tr>
<tr>
<td>Additional postreceptoral activation</td>
<td>Influence on:</td>
</tr>
<tr>
<td>Receptive field effects (control by spatial sums and differences)</td>
<td>Amplitudes of three photopic sensitivity functions</td>
</tr>
<tr>
<td>Activation of cortical sensory area 17</td>
<td>Control of magnitudes of input signals to postreceptoral spectral differencing mechanisms and summative luminosity mechanism</td>
</tr>
<tr>
<td>Area 17 local cortical connections between blobs</td>
<td>Locally weighted space (and time) average of:</td>
</tr>
<tr>
<td>Activation of cortical area 18</td>
<td>Difference and sum effects within adjacent postreceptoral neural elements</td>
</tr>
<tr>
<td>Reciprocal cortical connections between area 17 and area 18</td>
<td>Influence on:</td>
</tr>
</tbody>
</table>

Hurlbert’s (1986) words, computations that will “extract the invariant spectral-reflectance properties of an object’s surface from the varying light that it reflects.” Part of the problem considered by some computational studies is the separate extraction of the illuminant properties from specular highlights in a three-dimensional scene or representation thereof (D’Zmura & Lennie 1986; Lee 1986), and another part is the separation of shadows from material changes (Gershon et al 1986). Many of these approaches are concerned to some extent with one or another version of the retinex algorithm proposed by Land (1983, 1986; Land & McCann 1971) to specify lightness and color in constant terms related to constant reflectances and independently of illumination (Arend & Reeves 1986; Brainard & Wandell 1986; D’Zmura & Lennie 1986; Hurlbert 1986; Worthey & Brill 1986). Land’s computational procedures for describing perceived colors have undergone a number of mod-
ifications since he was first surprised by his own observation that the wide gamut of hues he was able to recognize in a photographic slide projection did not require wavelengths in the projected image that he associated with those hues, nor did they require mixtures of wavelengths from three different parts of the spectrum as he would have anticipated from the technology of colorimetry (Land 1959). Although others saw his demonstrations as instances of simultaneous color contrast, Land was not interested in contrast explanations, whether cognitive or physiological. As a physicist looking for another account from the physics of light, he proposed that the different colors seen in the natural image could be attributed to (and computed by) the ratios of almost any pair of longer and shorter wavelengths or wavelength distributions used to form the projected image or to illuminate the original scene. The first significant change in this anti-trichromatic, or at least nontrichromatic, idea was in the direction of traditional color theory. The two-record account was modified to a three-layer, three-light-record account in which lightness ratios were computed for each record separately, with the maximum lightness in each assigned a value of 1.0. Such a procedure yields a three-variable chromaticity and photometric lightness space normalized with respect to the maximum lightness, taken to represent “white,” with hue designations assigned to various regions in the space in accord with the hue names assigned to the three different light records. We would describe this procedure as akin to the application of a von Kries adaptation rule for the normalization, and a Young-Helmholtz type of theory for the color coding. Further modifications of the specifics of the retinex procedure include the computation of each lightness ratio record across reflectance boundaries, akin to Wallach’s (1948) account of achromatic lightness constancy; a reset correction to retain a maximum of 1.0; a logarithmic transformation; and the introduction of a ratio threshold. The latter serves to discount gradual lightness changes within reflectance boundaries of the sort that would be produced by an illumination gradient, thus eliminating the gradient from the computation and presumably from the perception as well. In his 1983 paper, Land includes a transformation from what we described above as a chromaticity and lightness space, which he calls the color three-space, to a red-green, yellow-blue, white-black opponent color three-space. This is another step in the direction of currently accepted color theory. In a still more recent report (Land 1986), an alternative algorithm is presented that involves photometric measurements of the surface pattern with a small and a large photometer aperture (the latter having a diminishing sensitivity profile), a log transform of the record at each of the two very different scales, and then a differencing operation. This alternative algorithm for the first time in retinex computations relaxes the strict coupling between computed lightness at a point on a surface and surface reflectance at that location. The procedure, although described differently, is implicitly akin
COLOR CONSTANCY

to the mechanism proposed by von Békésy (1968) to account for simultaneous contrast. This most recent change in retinex formulation thus brings the computational approach closer to the center/surround receptive field based modeling that we, and many others, have been engaged in for some time. The retinex operations do not yet, however, include the receptive field dependences required to subsume the systematic departures from lightness and color constancy that occur with change in level and quality of illumination. Nor do they yet include in the photometric procedures provision for change in retinal image size with change in viewer-to-surface distance, and thus the distance-dependent departures from perceived color constancy that can vary from assimilation to contrast effects for the same reflectance pattern which we discussed above (see the section on Contrast, Assimilation, and Receptive Fields).

It seems predictable that computational approaches to the old issue of color constancy will not for long continue to seek direct and precise perceptual correlates of constant surface reflectances, but will increasingly embody the more realistic approach of object identification through approximate invariance of color category. As we have pointed out elsewhere, there are some colors (e.g. the colors of haystacks, concrete and other masonry) that are difficult to categorize under any illuminant and that change quite noticeably with change in viewing conditions. For objects of this sort, color identification, rather than contributing to object identification, is more likely a result of it. It also seems predictable that approaches that include computations to extract illumination information as well as surface color will probably begin to incorporate shadow, as well as highlight effects, and to recognize the biological significance of such information as such for purposes other than being discounted. We have already cited attempts to separate shadow from material changes across surfaces, but we should add here that with no change in shadow, illumination, or reflectance, perceived differences can also result from apparent differences in object shape and orientation. Thus, a surface seen as a trapezoid under glancing illumination can appear less light than it does when the observer's set is manipulated so that the same surface is seen as a normally illuminated square lying flat on a receding plane (Hochberg 1978). Effects of this sort, when they occur, are clearly not under the control of any variables in the light stimulus, but rather point to mutual influences between different specialized processing systems tempered by well-practiced adaptive behavioral responses of the individual.

In the long run, the kind of widely encompassing computational approach that seems to us to offer the most promise for modeling of perceptual effects is exemplified by Edelman's neuronal group selection theory (Reeke & Edelman 1988). The theory is based on biological considerations, with both variability and selection emphasized not only as evolutionary but also as
developmental principles. In development, selection for neuronal connectivity is elaborated by selective mechanisms for differential cell growth and survival, and followed during early experience by selection, through modification of synaptic strengths, among diverse preexisting groups of cells to shape and adapt the behavior of the organism. An appealing feature of the computational model based on this theory is the processing in parallel of unique responses to individual stimuli (the automaton sampling system called *Darwin*), and of generic responses to stimulus class (the automaton sampling system called *Wallace*). There is high-level reciprocal connectivity between these systems, and a natural emergence of similarity-based categories that are relevant to the adaptive needs of organisms.

It seems to be agreed that surface color recognition is a useful component of object identification, and it is our judgment that such recognition is adequately accomplished by category matching and does not require precise matching-to-sample by the three color variables of hue, brightness, and saturation. It seems also to be agreed that context and instructions can modify actual experimental matching between extremes that approximate reflectance matches, on the one hand, and on the other hand, an illumination-dependent range of perceived hues, saturations, and brightnesses that include, but are not restricted to, a set of approximate reflectance matches. Both the systematic changes and the categorical constancies are perceptually available for recording in experiments, and more importantly, for adaptive responses to objects recognized in the environment and to the illumination conditions of that environment. Recognition and identification require some degree of perceived constancy, but we could cite too many examples of identification and recognition, whether of persons, objects, buildings, or landscapes, despite aging, fading, season, and illumination, to assume that the systematic changes related to such different conditions are not also perceptually informative in important ways.

**Literature Cited**

Guthrie, B. L., Porter, J. D., Sparks, D. L. 1983. Corollary discharge provides accurate eye position information to the oculomotor system. *Science* 221:1193–95
Helson, H. 1938. Fundamental problems in color vision. I. The principles governing changes in hue, saturation, and lightness of nonselective samples in chromatic illumination. J. Exp. Psychol. 23:439–76
MacLeod, R. B. 1932. An experimental investigation of brightness constancy. Arch. Psychol. 23(135):1–102