Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 5163–5169, August 1983
Physics

Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image*†

EDWIN H. LAND

The Rowland Institute for Science, 100 Cambridge Parkway, Cambridge, Massachusetts 02142

Contributed by Edwin H. Land, May 31, 1983

In the Color Vision Symposium at the Academy in April 1958, we showed a series of experiments which demonstrated that "whereas in color-mixing theory the wavelengths of the stimuli and the energy content at each wavelength are significant in determining the sense of color . . . in images neither the wavelength of the stimulus nor the energy at each wavelength determines the color. This departure from what we expect on the basis of colorimetry is not a small effect, but is complete . . ." (1, 2). The initial and most engaging experiment comprised taking two black-and-white photographs of the same scene, one through a red filter and one through a green filter, and projecting these two black-and-white pictures in superposition on the screen to yield a single black-and-white panchromatic image of the scene. When a red filter was placed in the path of the light from the projector that contained the picture taken through a red filter, the whole scene became dramatically colored as if in many respects it were a standard full-color photograph. The first paradox was that the radiation coming to the eye of the observer consisted only of various ratios of red light to white light which should have yielded only a variety of pinks. The second paradox was that the *overall* ratio of light from the one projector to light from the other projector could be changed markedly without changing the color names of the objects in the colored picture: the colors of the individual objects must be determined by the ratio of red light to white light, but a change in the overall ratio of red light to white light did not change the colors.

In light of the understanding which we now have, this simple experiment, which was a shock to the intuitive understanding of all of us, turns out to be the most sophisticated experiment we could have undertaken.

For the flavor of the many experiments described at the Symposium, I refer you to the two papers (1, 2) at that time. Here, I want to turn to the quantitative procedures which we now use. We prepared a laboratory display which we dubbed a "Mondrian" (although it actually is closer to a van Doesburg), utilizing about 100 colored papers. A paper of a given color would appear many times in different parts of the display, each time having a different size and shape and each time being surrounded by a different set of other colored papers. One reason for the design was to prohibit the superposition of afterimages of areas onto other areas (3), and another reason for the design was to obviate explanations of results in terms of the size or shape or surrounding of any given paper.

The Mondrian is illuminated by using three 35-mm slide projectors with no slides in the slide holder. The output of each projector/illuminator is controlled independently. An interference filter passing long waves is placed in the path of one projector, a middle wave filter, in the path of the second, and a short-wave filter, in the path of the third (Fig. 1). One may think of these as relating roughly to the three visual pigments. A telescopic photometer (Spectra Pritchard photometer, model 1980A), placed roughly where the observers will be, receives and measures radiation from about 1/16th of a square inch on each chosen area of the Mondrian when it is pointed at that area. The instrument is calibrated so that at any wavelength it reports directly in watts per steradian per square meter.

Let me call your attention to these four papers: yellow, white, green, and blue. The telescope is pointed at a yellow paper. The short-wave and middle-wave illuminators are turned off, and the whole Mondrian is illuminated with the long-wave illuminator. The output of this projector is then changed until the meter reads exactly "one" (0.1 W per Sr^2 per m^2). The long-wave illuminator is turned off and the middle-wave illuminator is turned on. Its output is adjusted until the meter reads one. This ensures that the amount of middle-wave energy now reaching the meter from that small patch is equal to the amount of long-wave energy. Finally, after the middle-wave illuminator is turned off, the short-wave illuminator is turned on, and its output is set so that the meter (which we must remember is reading the radiation to our eyes) reads one. All three illuminators are now turned on. While looking at the Mondrian as a whole, we note that the yellow paper looks yellow. We now turn our attention to the white paper, pointing the telephotometer at it. We go through the same procedure of illuminating with one illuminator at a time and of setting each illuminator so that the light coming this time from the white paper to the meter, and hence to our eyes, measures one for the long wave and one for the middle wave and one for the short wave. Thus, we have arranged to have coming to our eye from the piece of white paper exactly the same flux—the same wavelength composition, the same energy composition—which a moment earlier we had arranged to have coming to our eye from the piece of yellow paper. The somewhat indigestible question is "what color will the piece of paper be which was white in the Mondrian previously?" Keep in mind that the information now coming to our eye from that piece of paper dictates classically that, if one, one, and one coming to our eye gave yellow, then one, one, and one must again be yellow. This conviction dates back to Newton's proposition V (4):

> Corporum naturalium colores e genere radiorum derivantur, quos maxime reflectunt,

which we translate as "The colors of natural bodies are derived from the type of rays which they reflect to the greatest degree." In the English edition of Newton's "Optics" (5), he states that "Every Body reflects the Rays of its own Colour more copiously than the rest, *and from their excess and predominance in the reflected Light has its Colour.*"

When we turn on all three illuminators and send to our eye from the originally white paper exactly the radiation which came to our eye from the yellow paper, we can properly expect the

* Presented at the annual meeting of the National Academy of Sciences, Apr. 25, 1983, Washington, DC.

† This is paper no. 3 of a series. Paper no. 2 is ref. 2.

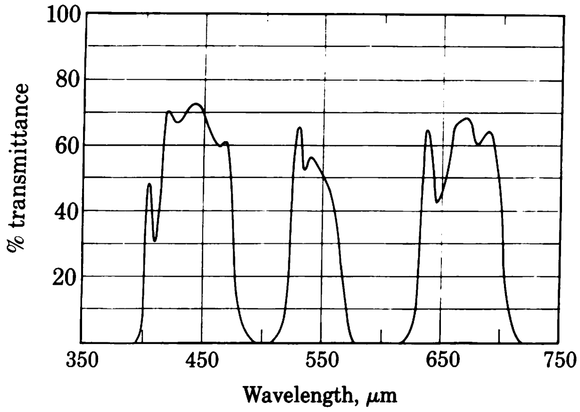


FIG. 1. Transmission bands of filters used to illuminate the Mondrian.

paper which initially looked white now to look yellow. On the contrary, however, it is white!

We now repeat the "one-one-one" experiment with the green paper and once again find that it is still green seen with the same package of radiation coming to our eye as came first from the yellow and second from the white. Similar results are obtained from a blue piece of paper. In spite of the large changes in the relative outputs of the illuminators that were necessary to cause the yellow, the white, the green, and the blue papers all to reflect one, one, one, the colors of the Mondrian changed remarkably little.

It is vital to note that, when synchronized camera shutters are placed on the three illuminants, limiting the time of illumination to 0.1 sec, there is an even closer approximation to independence of color from the composition of the radiation from an area to the eye. Thus, the high degree of independence cannot be accounted for on the basis of visual phenomena such as adaptation, pigment bleaching, wandering of the eye, and so on. Rather, one must search for a precise computational process extending over large areas of the field of view and possibly over the whole field of view.

As a first step in such a search, it seems wise to distinguish between domains of independence and domains of dependence. The first lesson taught by the early pictures seen with red and white light was that the results in terms of color were only weakly dependent on the relative amounts of energy used for showing the long-wave black-and-white picture and the short-wave black-and-white picture. The exigencies of this kind of color photography, and indeed of the established kinds of color photography as proposed by James Clerk Maxwell, lead to three separate black-and-white records of the world being photographed, one for long wavelengths, one for middle, and one for short. With the primitive photographic materials available to him, Maxwell was not stimulated to treat the three records as independent; indeed, in continuing the thought pattern of Newton and Young, he was deeply immersed in the concept of *mixing* colors—witness the Maxwell color-mixing top. Maxwell did not realize that the three black-and-white records are independent and that they can be projected without relating the flux from one projector to the flux from the others over a remarkably wide range.

If you contemplate the Mondrian while I make these sweeping changes in relative illumination on the separate wave bands, you cannot possibly understand why there are not concomitant changes in color—as long as you think in terms of color-mixing at each point. If, however, you can imagine the Mondrian as being a composite of three independent images, one carried by long waves, one carried by middle waves, and one carried by short waves, I can then introduce the proposition that, first, each of these images is unaltered by a change in the flux with which it is carried and second, that it is the comparing of these three complete images, rather than the merging of their fluxes, which produces the array of colors.

Let us first address the aspect of the proposition which says that each of the images is unaltered by a change in the flux with which it is carried. A corollary observation is that when I illuminate the Mondrian with one narrow band of wavelength, the image which you see *is* unaltered by an overall change in illumination (nor would it be altered by oblique illumination such as would be provided by a point source positioned near a corner of the display). We are led to realize that what we mean by an image, at least as the basis of color vision, is an array of areas with different lightnesses relative to each other. The fact that this array of lightnesses is not altered as a function of illumination leads us to the proposition that the image is not dependent on the quantity of illumination. Similarly, the imagined image of the real Mondrian on any one of the three wave bands will be independent of the flux on that wave band. You probably noticed when we were changing from one wave band to another in illuminating the Mondrian that an area which was light on one wave band might be very dark on a second wave band. If you were very observant, you would have seen that each colored area was characterized by three lightnesses and that, conversely, if you know the three lightnesses of any area, you will know what color it will be when all three illuminators are turned on. Since the lightnesses are independent of the quantity of illumination, it follows that the color of any area should be independent of the relative quantities of illumination. There is left, then, the great final question: What is the computational basis for each of the three independent images-in-terms-of-lightness?

For the sake of analysis, let us assume an imaginary retina with a single kind of cone containing one broadband pigment peaking somewhere in the middle of the visible spectrum. We cannot see what we know as an image prior to the operation of the computational system that generates lightnesses. We cannot look at a photograph without processing by means of the lightness computation. What we know as reality is the experience at the terminal end of this computation. Since we all use the same computational mechanism, we share the terminal experiences. We name them, talk about them, train ourselves to relate to them and to handle them. The role of the computational mechanism is to generate a common denominator of permanences in an environment where the illumination is changing in intensity and wavelength composition, in which shadows merge with objects while objects fade into shadows. In this agonizingly complex environment, a visual system which transferred the actual external display to a corresponding internal presentation would have little or no practical value. A visual system did evolve in which internal structures and processes are so related to external events as to generate an alphabet and a vocabulary of useful visual "realities," synthesized permanences which, though to a significant extent a product of internal action, can be treated as if they are external. The very success of this evolutionary program convinces us that the products of this synthetic partnership *are* external realities. I think the best way to learn about this partnership is to go at once to the study of a proposed biological mechanism for making color. Thus, we turn back to our imaginary retina containing only one set of cones, with peak sensitivity at the center of the spectrum of visible rays. The pattern formed by the lens of the eye whenever it is imaging the outside world on the retina would be an unimaginable hodgepodge (unimaginable because we can never see it). What we seek is an operation to perform on the output of this simplified retina that will yield an image. Since the pigment uses only one band of wavelengths, the image should resemble what we see in a black-and-white photograph. The formulation discussed below promises to satisfy that need.

Although we know that a large part of the visual process is carried on beyond the retina, let us assume for a moment that all of the processes are carried on in the retina. This will assist in our topographical visualization of the mechanism of the transducing process which must convert the hodge-podge arriving at the retina to a useful, synthesized reality.

On the basis of experiments such as the one-one-one, we accept the first postulate of retinex theory (6, 7) that there are three identical independent lightness-making mechanisms, one for long waves, one for middle waves, one for short waves, each served by its own retinal pigment. The basic task of color vision theory becomes the determination of the nature of these mechanisms. I would like to discuss several examples of mechanisms. The first two are essentially similar, one being dynamic, the other being static. My own taste is for the dynamic one if it is neurally feasible.

Version 1: In the dynamic version, at brief intervals, one retinal cell or another in the family of cells containing the same kind of pigment will emit a signal which will proceed radially outward. This signal will be proportional to the logarithm of the intensity of the radiation falling on that retinal element. From it will be subtracted the logarithm of the intensity of the radiation falling on an adjacent retinal element. To this result will be added the output of the next adjacent cell minus the output of its neighbor, and so on. (This procedure is akin to relating the initial point to later points by the product of successive ratios.) The signal will proceed in this way, with freedom to branch within the rule of implied directionality inherent in the concept of a signal radiating from a single source. From time to time, one cell or another may be the initiator of a signal. At any given time, not more than a few hundred cells need be "radiating."

A basic property of this process as so far described is that it will give a constant relationship of any given point to a starting point, irrespective of changes in relative illumination—providing that there is a threshold value below which the difference in the logarithmic outputs of adjacent cells is called zero.

Now imagine that we are looking down on the retinal plane and giving our attention to a point on it. At any instant, hundreds of signals will be arriving at the point. Each signal will report the value of the unique relationship between this point of intersection and one of the photosensitive cells at which the signal originated. The average of these values will characterize the point. Let me name this average the "lightness number."

You will recall that for the sake of analysis we assumed a retina with a single kind of cone containing a pigment peaking in the middle of the spectrum. Let us now assume that these cones use the middle-wave pigment. The lightness number at which we have just arrived is based on the information from the whole family of cones containing that pigment. We must now imagine two more families, one with pigment peaking in the short-wave part of the spectrum and another with pigment peaking in the long-wave part of the spectrum. Each of these families will operate independently of the other two, and each of them will generate its own lightness number at any single point of intersection. Any point in the field of view will be characterized by three lightness numbers. They are independently derived and must be kept independent.

The best way to represent them is as a point in a three-space. Each axis of the three-space will be identified with the range of lightness numbers for one of the three families of cones (Fig. 2). Thus, the three lightness numbers which fully characterize any point of intersection will together designate a point in the three-space and each point in the three-space will be a different color from all the rest.

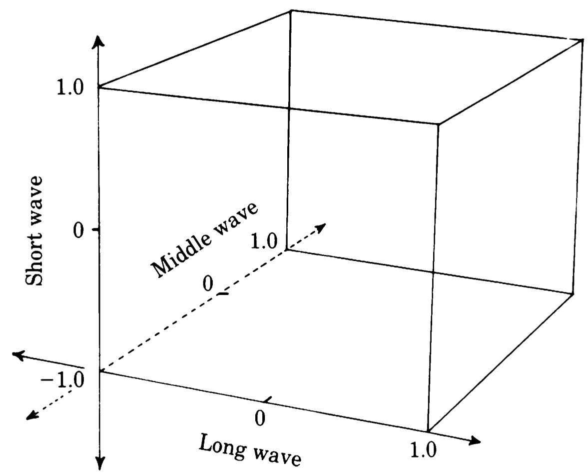


FIG. 2. Coordinate system for color three-space. The cube is sketched in as an aid in spatial visualization. Only the axes are significant.

When an image of a colored Mondrian is formed on the retina and the computational technique proposed for the retina is pursued, it is found that the three-space is populated in an orderly way. The points on one of the internal diagonals turn out to look black at one end, run through gray, and are white at the other end of the diagonal. There is a domain in which the greens reside, another for the reds, still another for the blues, and yet another for the yellows. It is a triumph of this computational technique that the overall variation in the composition of the illumination in terms of flux at a given wavelength or in terms of relative flux between wavelengths does not disturb the reliability with which a paper which looks red, no matter where it resides in the Mondrian, will have the same three lightness numbers as the other papers which look red. It will therefore be part of a family of reds which appears in one domain of the three-space. Similarly, all the blues or greens or yellows, *wherever resident in the Mondrian* and however haplessly illuminated, will appear in their appropriate domains in the three-space (Fig. 3A). It is the computation that leads each paper to have its position in the three-space; the proof of the pudding is that all things that appear in the same region of the three-space are the same color as one another, whatever their history in terms of geography and illumination on the Mondrian may have been.

Version 2: The static version of the mechanism is best described as a conceptual entity including symbolism for wave bands, even though some of the description is redundant with respect to version 1. To predict the perceived color of an area in any visual scene, the computation shown in Fig. 4 is carried out. The visual field is broken into unit areas. The relative reflectance, R, of the target area, i, is computed with respect to some other area, j, along a path drawn between the two areas by using the formula shown, where Λ designates the particular wave band (long, middle, or short) and I is the intensity. The threshold operation on the ratios along the path is included to remove the effects of nonuniform illumination over the scene: variations gradual enough to be below threshold are dropped out; all others are considered significant and contribute to the computation. The average of many such computed relative reflectances is taken to determine the value we define as average relative reflectance at area i.

Conceivably, this average of relative reflectances (not fluxes) could be taken over every area in the visual field, but as few

 Proc. Natl. Acad. Sci. USA 80 (1983)

as one or two hundred is usually sufficiently accurate. The average is taken over areas from the entire visual field and not just those nearby; experiments indicate there may be nearly as much contribution from distant areas as nearby ones.

Since the above computation is carried out three times, once for each of the three wave bands, three numbers become associated with each unit area. These three numbers designate a point in a three-dimensional color space, a point which proves

A

B

C

D

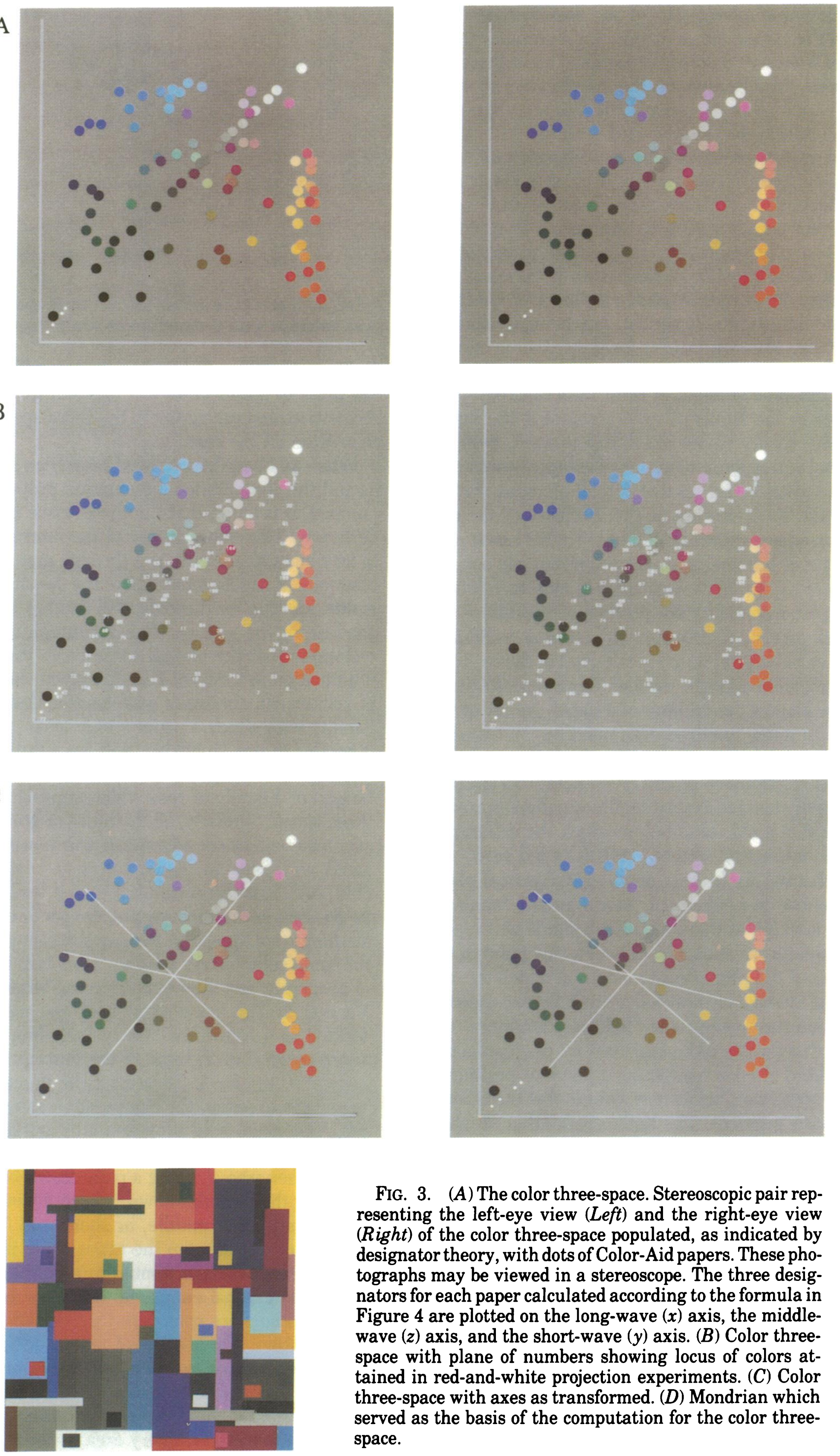
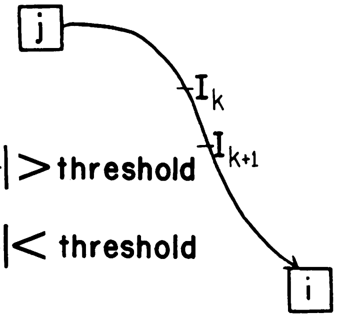


FIG. 3. (*A*) The color three-space. Stereoscopic pair representing the left-eye view (*Left*) and the right-eye view (*Right*) of the color three-space populated, as indicated by designator theory, with dots of Color-Aid papers. These photographs may be viewed in a stereoscope. The three designators for each paper calculated according to the formula in Figure 4 are plotted on the long-wave (x) axis, the middle-wave (z) axis, and the short-wave (y) axis. (*B*) Color three-space with plane of numbers showing locus of colors attained in red-and-white projection experiments. (*C*) Color three-space with axes as transformed. (*D*) Mondrian which served as the basis of the computation for the color three-space.

The relative reflectance of i to j:

$$R^{\Lambda}(i;j) = \sum_k \delta \log \frac{I_{k+1}}{I_k}$$



$$\delta \log \frac{I_{k+1}}{I_k} = \begin{cases} \log \frac{I_{k+1}}{I_k} & \text{if } \left|\log \frac{I_{k+1}}{I_k}\right| > \text{threshold} \\ 0 & \text{if } \left|\log \frac{I_{k+1}}{I_k}\right| < \text{threshold} \end{cases}$$

Average Relative Reflectance at area i:

$$\bar{R}^{\Lambda}(i) = \frac{\sum_{j=1}^{N} R^{\Lambda}(i;j)}{N}$$

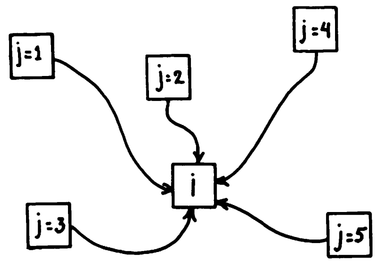


FIG. 4. Computation of average relative reflectance (designator).

to be invariant with large changes in quantity and composition of illumination of the field of view. Experimentally it is found that if two points in any visual scene are the same color (even though the wavelength composition reflected from them may be very different), they will be represented by this computation at the same point or closely adjacent points in this color three-space.

Version 3: Horn (8) proposed one scheme for computing lightnesses for the three retinexes. It has many of the required basic properties demonstrated earlier. The algorithm involves computing the local ratio of fluxes in a particular isotropic fashion, doing a threshold operation, and then carrying out the mathematical inverse of the original ratio operation. His implementation of the critical inverse operation postulates a network of cells connected locally in particular feedback loops whose mutual interaction is intended to generate the necessary long-distance dependencies.

There are questions raised by this implementation—questions of just how long-range the interactions are, the importance and handling of the boundaries, and an overall scaling problem. Nevertheless, it illustrates another approach to implementing retinex calculations.

Versions 1 and 2 of the computational scheme for the individual retinexes will yield designators for blacks, for whites, for the range of whites, for light sources in daylight and light sources in the dark—e.g., signal lights. The values will all fall in the color three-space.

The computation does not depend on using an initial standard such as a white or a black area. Although the computation is self-calibrating, the average that is used as a standard is not an average of fluxes but an average of ratios that are derivable in a field of unknown, changeable, and unknowable flux.

This computation will handle tasks which our earlier proposals could not and there is an ultimate conceptual simplicity in the present approach. Nevertheless, the results of extensive quantitative tests of retinex theory using the earlier computational technique are valid and valuable (9, 10).

It is too soon to say what computational mechanisms will finally be most appropriate for the cellular anatomy. Both mechanisms 1 and 2 achieve all the operational objectives, and we can turn to a three-space based on them for a predictive and descriptive tool. Perhaps the most satisfying test of its predictive ability is to find where the colors in the red-and-white picture, and some variants of them, are located, after first predicting where they ought to be located. As you will remember from our two-color projection experiments, we project two records, a middle-wave record (made on black-and-white transparency positive film with a Wratten 58 green filter) and a long-wave record (made on black-and-white transparency positive film with a Wratten 25 red filter). Let us consider the red-and-white pictures in terms of lightness numbers which, for convenience, let me call designators. The middle-wave and short-wave designators are equal to each other because when we project the middle-wave record with white light, we are, in effect, placing on the screen three identical images in superposition:

(*i*) the image carried by the portion of the white light to which the short-wave photosensitive cells respond,
(*ii*) the image carried by the portion of the white light to which the middle-wave photosensitive cells respond,
(*iii*) and the image carried by the portion of the white light to which the long-wave photosensitive cells respond.

When we add a second projector equipped with a red filter and project the long-wave record, the energies transmitted by this red projector are combined geographically with the energy from the long-wave portion of the white light that previously had been placed on the screen. From this summation of long-wave energies, a final image, in terms of lightness, is then computed for the long-wave retinex.

Thus, the short-wave and middle-wave retinexes will have equal designators for every place in a projected red-and-white image, and the long-wave retinex will have designators computed from the combined long-wave energies. When we photograph a variegated colored scene and compute the trio of designators for the colors that will appear in a red-and-white picture, all the points in three-space that represent the original scene fall on a plane that cuts through the whole array of colors in the three-space (Fig. 3*B*). The numbers in the plane are the "addresses" of points in the original scene. It is obvious that, although all the colors are confined to a plane, that plane passes through enough important color domains in the color space to yield a very colorful picture. There are many other interesting experiments that can be carried out to study this plane. For example, when we move the red filter from the long-wave record to the middle-wave record, all the colors change to quite different new ones falling on the same plane in predictable new positions.

At the beginning of this paper, we saw that if we kept the same designators, which we did automatically in the Mondrian when we changed the flux on each wave band, the color stayed whatever it had been—the yellow stayed yellow, the green stayed green—because the change in the relative flux did not change the trio of designators for the yellow area or for the green area. Both of them had their proper fixed place in the three-space. The color in the Mondrian stayed constant because changing the flux did not change the designators. Now let us seek an experiment in which we keep the flux from a paper constant but change the designators. We have placed a circular piece of paper onto the Mondrian and, in the slide holder of the illuminator (which, of course, has no image in the slide since you are illuminating and looking at the actual Mondrian), we have placed a slide that has an individually computed neutral-density filter with a hole in it. We have done that for each of the three illuminators—the long, middle, and short—so that the projected holes are in registration with the circular paper. Thus we have obtained, by computing the designators for the circle, a value of 1.5 for the long wave, 1.5 for the middle wave, and 1.5 for the short wave. These designators denote a white far out in the color three-space. By pressing a button on the slide changer, we remove from all three illuminators the three neutral filters with the holes in them so that the illuminators will light the entire Mondrian uniformly for each of the three wave bands. There will be no change in the flux coming from the circular paper because the flux illuminating it was coming through the *holes* in the three neutral filters. We have done nothing to change that flux. By removing the neutrals from the rest of the field of the illuminators, we will cause the generation of three new designators for the circular paper on the Mondrian, one for each wave band, having the computed values 0.5, −0.2, and 0.4. Let us turn to the color three-space to see what our designators predict for the new color of the circle. Looking at the color three-space stereoscopically, estimate the positions of 0.5, −0.2, and 0.4. Now you can see that that position in the color three-space is occupied by a purply magenta. If you now turn to the Mondrian while we remove the neutral filters, thus causing the generation of these designators, you will see that the white paper has gone to a deep purple. Thus we have carried out the experiment of keeping the flux the same and changing the designators.

This experiment with the circle which we change from white to purple by changing the designators is of particular value for studying the question of the significance of the immediate surround—that is, of the influence on a given piece of paper of the papers that are its neighbors. The very design of the Mondrian is supposed to demonstrate that the immediate neighborhood has no special significance because a paper of a given type is the same color whatever its shape or size or position in the Mondrian. Nevertheless, the question arises so frequently that it is good to have another strong demonstration. In the same slide holders that hold the neutral filters with the holes in them, we can place opaque annuli also having holes of the same size in register with the holes in the various neutral filters. We can then, at will, observe the Mondrian and the circular piece of paper on it with a wide ring that looks black between the circle and the rest of the Mondrian. If we then change the designators deriving from the whole Mondrian, we find that the circle still responds accurately to the instruction of the designators, even though there is now being included in the computation the very wide black ring between the circular paper and the rest of the Mondrian.

While the experiment with the black annulus shows that the computation for the lightness of a point is carried out over extensive areas of the field of view, there is nothing in the experiment to tell us to what extent the computation is carried on in the retina as opposed to the cortex. The following experiment, carried out in collaboration with David Hubel and Margaret Livingstone, shows that the cortex is essential to the computation (11).

A Mondrian was modified by covering half of it with velvet. In the middle of and just to the velvet side of the vertical line where the velvet met the rest of the Mondrian, we placed a target area having a fixation point within it. The flux from the target area to the observer's eye was kept constant while the illumination of the Mondrian was altered to create a new set of designators for the target area. For the normal observer, this procedure altered the color of the target area from purple to white. An observer who was normal, *except that for medical purposes his corpus callosum had been cut through,* sat in a chair several feet in front of the target area. For all of us, the two left-hand sides of the retinas are connected to the left sides of our brains, and the two right-hand sides of our retinas are connected to the right sides of our brains. Consequently, one half of the field of view is served by one half of the brain and the other half of the field of view, by the other half of the brain. If the computation for the Mondrian were carried out entirely in the retina, the results of the computation for any point in the field of view would be independent of which half of the brain was used to read the results. The subject with the split brain would report "white" when the rest of us saw white and "purple" when the rest of us saw purple. This was not the case.

If the computations are carried out *not* in the retina, but in the cortex, there should be a striking difference between the results for a normal observer and results for a subject with a severed corpus callosum. If we assume that normally the computation extends over the whole field of view, it follows that the computation would extend across both halves of the brain by way of the corpus callosum. When the normal observer regards a Mondrian, half of which is velvet and half of which is the standard Mondrian, the color he sees in the target area depends on the whole computation across the whole field of view. When our subject with the severed corpus callosum regards the same field of view, one half of his brain will be computing for Mondrian plus target and the other half of his brain, for black velvet plus target. When the verbal half of his brain is dealing with *Mondrian* plus target, he reports "purple" or "white" at the same time that we do. When the verbal half of his brain is dealing with *velvet* plus target, then he reports only "white" while we are seeing the changes from white to purple. (We determined this choice of presentations by having him look at the display directly or with a mirror.) Thus, the experiment establishes that the computations (*i*) are long-range and (*ii*) require the cortex.

A summary rule for retinex theory is that color is always a consequence, never a cause. To understand opponent phenomena in terms of retinex theory, one should endeavor to restate any experiment in terms of the computation on each of the three independent retinexes. Recent and current findings with double opponent cells (D. Hubel and M. Livingstone, personal communication) (described, for example, as "red on (+) center, red off (−) surround, green off (−) center, green on (+) surround") open the question of whether there is a transformation from the long-, middle-, and short-wave axes we have shown for the color three-space to a set of axes lying within the color space and consistent with the double opponent cell's postulated red-green, yellow-blue, and white-black response. Such a transformation does exist: it is a simple linear "almost rotation": one axis is not quite orthogonal to the other two, and the scale factors are not quite unity (Fig. 3*C*). Versions 1 and 2 of the computation depend on nonsymmetry to establish directionality of paths, so that double opponent cells would have to project to at least a mathematical stratum which introduced nonsymmetry.

Semir Zeki has discovered color-reading cells in the V4 region of the prestriate visual cortex of the rhesus monkey (12). The image of a Mondrian is formed on the retina of the anesthetized monkey, and Mondrian experiments analogous to those we carried out on ourselves are carried out with the animal. The results for the monkey, as reported by his cortical cells, are gratifyingly similar to our own.

What is left to be done in the further exploration of retinex theory?

(*i*) For the mechanisms that determine lightness by using thresholds, it would be good to know more about quantification of these thresholds. For example, a color photograph can be put surprisingly far out of focus, when projected, without losing the retinex interactions that lead to color. One should avoid both neural and computational proposals that depend on sharp edges.

(*ii*) The anatomical continuum must be filled in for the member of each family of cones (i.e., those with the same spectral sensitivity) which unites them into a computing entity for the individual retinexes.

In conclusion, retinex theory is a powerful tool that explains many of the basic phenomena of color vision. It predicts many others. We are finding it most rewarding to use as a tool for all visual investigations ranging from the nature of visual pigments, to the source of color in Mach bands, to time constants in Benham's top phenomena, to the study of the relationship between the wavelengths of three illuminants for taking a picture and the wavelengths of the three illuminants used for viewing a picture and indeed to the spectrum. It is particularly useful to be able to understand many of a large variety of "effects" in the simple terms of lightness, designator theory, and the color three-space.

1. Land, E. H. (1959) *Proc. Natl. Acad. Sci. USA* **45,** 115–129.
2. Land, E. H. (1959) *Proc. Natl. Acad. Sci. USA* **45,** 636–644.
3. Daw, N. (1962) *Nature (London)* **196,** 1143–1145.
4. Newton, I. (1729) *Lectiones Opticae (London)* Prop. V, p. 215.
5. Newton, I. (1704) *Opticks (London)* Prop. X Prob. V, p. 135.
6. Land, E. H. (1964) *Am. Sci.* **52,** 247–264.
7. Land, E. H. (1981) in *Encyclopedia of Physics, Vision and Color*, eds. Lerner, R. & Trigg, G. (Addison-Wesley, Reading, MA), pp. 1090–1097.
8. Horn, B. K. P. (1973) *On Lightness*, M.I.T. Artificial Intelligence Laboratory Memo No. 295.
9. Land, E. H. & McCann, J. J. (1971) *J. Opt. Soc. Am.* **61,** 1–11.
10. McCann, J., McKee, S. & Taylor, T. (1976) *Vision Res.* **16,** 445–458.
11. Land, E., Hubel, D., Livingstone, M., Perry, S. & Burns, M. (1983) *Nature (London)* **303,** 616–618.
12. Zeki, S. (1980) *Nature (London)* **284,** 412–418.