MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

Color Vision and Image Intensities:
When are Changes Material?

John M. Rubin & W. A. Richards

ABSTRACT

Marr has emphasized the difficulty in understanding a biological system or its components without some idea of its goals. In this paper, a preliminary goal for color vision is proposed and analyzed. That goal is to determine where changes of material occur in a scene (using only spectral information). This goal is challenging for two reasons. First, the effects of many processes (shadowing, shading from surface orientation changes, highlights, variations in pigment density) are confounded with the effects of material changes in the available image intensities. Second, material changes are essentially arbitrary. We are consequently led to a strategy of *rejecting* the presence of such confounding processes. We show there is a unique condition, the *spectral crosspoint*, that allows rejection of the hypothesis that measured image intensities arise from one of the confounding processes. (If plots are made of image intensity versus wavelength from two image regions, and the plots intersect, we say that there is a spectral crosspoint.) We restrict our attention to image intensities measured from regions on opposite sides of an edge because material changes almost always cause edges. Also, by restricting our attention to luminance discontinuities, we can avoid peculiar conspiracies of confounding processes that might mimic a material chage. Our *crosspoint conjecture* is that biological visual systems interpret spectral crosspoints across edges as material changes. A circularly symmetric operator is designed to detect crosspoints; it turns out to resemble the double-opponent cell which is commonplace in biological color vision systems.
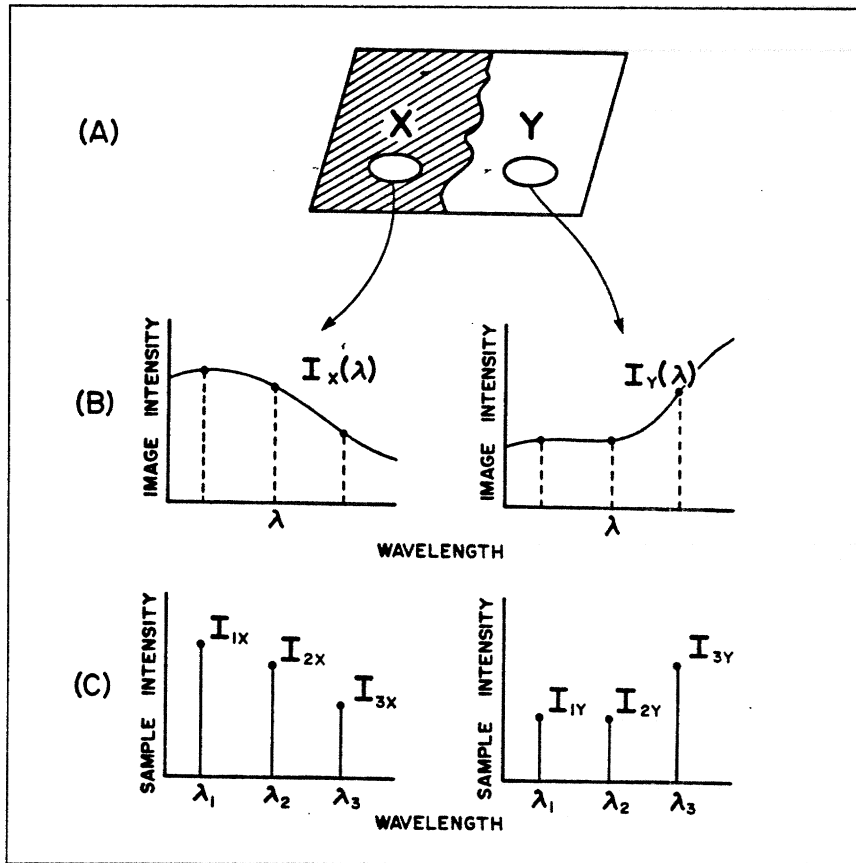
Figure 1. Notation. A) X and Y are two regions in an image. B) The available information is continuous functions of image intensity versus wavelength, $I_X(\lambda)$ and $I_Y(\lambda)$, from regions X and Y respectively. C) Discrete representation of this information is possible by taking a number (here, three) of spectral samples (narrowband in this illustration, at wavelengths λ_1, λ_2, and λ_3).

Our goal is thus similar to the one first proposed by Land (1977), but more modest. Land was concerned, as we are, with computing information about surface properties of materials in a scene. His objective was to determine the reflectance of regions using only the available intensity information. Land reached his objective by assuming a greatly simplified Mondrian world—a flat, shadowless world composed of regions of uniform reflectance. We prefer to deal with the natural world in its entirety without unnecessary simplification, and instead limit our objectives. In this way we can exploit the regularities in the world as assumptions or constraints in the solution of our problem. The simple goal we will analyze, to reiterate, is just to determine when changes in an image arise from one material's bordering (or occluding) another in a scene. We shall see below that the problem to be solved in achieving our modest goal is still formidable.

GLOSSARY OF SYMBOLS

I. Processes

$\hat{M}$—material changes

$\hat{S}$—shadow (point source)

$\hat{H}$—highlight or gloss

$\hat{O}$—change in surface orientation

$\hat{P}$—change in pigment density

$\hat{S}_{pen}$—shadow-with penumbra (extended source)

$\hat{O}_{simple}$—monotonic change in surface orientation

II. Quantities measured in Image

$I_X(\lambda)$—the continuous function of image intensity versus wavelength measured in a region X in an image.

$I_{i,X}$—a discrete spectral sample at wavelength λ_i

III. Continuous Scene Variables

$\rho(\lambda)$—the albedo or spectral reflectance function of a material.

$E_D(\lambda)$—the diffuse component of the illumination.

$E_S(\lambda)$—the synthetic component of the illumination.

$E(\lambda)$—the illuminant taken as a whole (when components are irrelevant).

α—the penumbra constant.

γ_X—the thickness of pigmentation in region X.

IV. Discrete Scene Variables

If $Q_X(\lambda)$ is some function of wavelength in region X, then $Q_{i,X}$ is a discrete sample of that function, at wavelength λ_i, from region X.

Table 1. A glossary of symbols.

Table 1. A glossary of symbols.

2   The Scope of the Problem

The problem in achieving the proposed goal of finding changes of material (hereafter denoted by the symbol $\hat{M}$[1]) is that any given image intensity can arise in many different ways, depending on the particular processes in effect. Thus given a single intensity value in the image, it is generally not possible to decide which of the many possible events in the world produced it. It can be said that the act of imaging a scene, or projecting it into two dimensions, *confounds* the effects of material changes, our interest here, with the effects of other processes. These confounding events include shadows, surface orientation changes, highlights, and variations in pigment density. Furthermore, the quest for recovering $\hat{M}$ (material) changes from image intensities must succeed under a range of spectrally different illuminants. A system that only worked properly given an illuminant having the spectral composition of the mid-day sun would be very limited. At dawn and dusk, when sunlight reddens, the system might fail.

[1] The circumflex accent is used in this paper to denote abstract *processes*; it is intended to keep discussion about processes distinct from talk about simple variables or functions.

We can now summarize the problem that early color vision faces, if its goal is to discover material changes:

The problem in determining where changes of material occur in a scene is that in the available image intensity values, the effects of many processes may be confounded.

In the remainder of section 2, a notational scheme will be developed. In section 3, a general outline of the solution will be presented.

Notation. A notational scheme is shown in figure 1. Regions in the static *image* will be denoted by X and Y. (The time variable will be ignored.) These letters will also be used to refer to regions in a *scene* which are the inverse projections of the image regions. The context will make clear the correct referent. The image intensity measurable in region X (or Y), say, as a function of wavelength, will be denoted by the function $I_X(\lambda)$ (or $I_Y(\lambda)$). Note that $I_X(\lambda)$ (or $I_Y(\lambda)$) represents *all* the information available from region X: it is the continuous spectral distribution of image intensity from X (or Y), as shown in the middle graph in figure 1. In section 5, we will refer to discrete approximations of $I_X(\lambda)$, which can be generated by sampling the image intensity from X at spectral points $(\lambda_1, \ldots, \lambda_n)$, as shown in the bottom graphs in figure 1. In discrete computations, I_{jX} will denote the image intensity measured from region X at λ_j.

Since many symbols are used in this discussion, a glossary is presented in table 1.

3   The Theme of the Solution: A Negative View

3.1   The Strategy

Given two spectral energy distributions $I_X(\lambda)$ and $I_Y(\lambda)$, how might we determine whether they arise from an $\hat{M}$ (material) change? Little can be said about the spectral nature of $\hat{M}$ changes; they are essentially arbitrary. No simple equations can relate image intensities from two regions composed of different materials. The major confounding processes—shadows (hereafter denoted $\hat{S}$), highlights ($\hat{H}$), surface orientation changes ($\hat{O}$), and changes in pigment density ($\hat{P}$)—however, produce lawful changes in the image. Simple equations can capture this lawfulness, as will be seen in section 4. Suppose that by examining image intensities from two regions we could eliminate the possibility that they arose from either an $\hat{S}, \hat{H}, \hat{O},$ or $\hat{P}$ change. Since we are assuming the illuminant to be spectrally invariant in a neighborhood, we would like to conclude that the change was due to a material change.

A preliminary conjecture is thus proposed:

When a difference between the image intensities $I_X(\lambda)$ and $I_Y(\lambda)$ (taken from two image regions X and Y) does not arise from one of the confounding processes $\hat{S}$, $\hat{H}$, $\hat{O}$, or $\hat{P}$, then this difference between X and Y is due to a difference in materials, or $\hat{M}$.

The conjecture above suggests a computational strategy of attempting to *reject* measurements of image intensities as arising solely from shadow ($\hat{S}$), highlight ($\hat{H}$), surface orientation change ($\hat{O}$), or pigment density change ($\hat{P}$), the lawful confounding processes in a scene. Our rejection strategy will be correct only if there are no confounding processes other than the ones mentioned. But, as will be discussed later, if in the course of rejecting the presence of $\hat{S}$, $\hat{H}$, $\hat{O}$, and $\hat{P}$, we also reject the presence of any of a large class of other possible confounding processes, the strategy will be a powerful one and the conjecture will be useful for practical purposes.

It is important to note that *rejecting* the presence of a lawful process is often much easier than *accepting* it. As mentioned above, lawful processes are associated with equations. These equations relate quantities measurable in the image (constants in the equations) to scene properties (variables in the equations) which are not directly measurable. Typically, these scene property variables (reflectances, for example) are constrained to take values within a certain interval. As will be seen below, rather than attempting to solve a system of equations (subject to constraints on the values of variables), it is often computationally simpler to determine *whether* the system (with constraints) can be solved. In this sense, rejecting a solution (process) is easier than accepting one. By analogy, disproving a conjecture about number theory (some universally quantified equation or inequality) with a single counterexample is simpler than demonstrating the theoremhood of the conjecture.[2]

Here's a simple example. Consider the equations $I_1 = JK$ and $I_2 = J^K$, where J and K are variables, and I_1 and I_2 are constants. If the variables J and K are constrained to have values greater than unity, then a simple test can be made that might determine that the system has no solution. Specifically, if $\min(I_1, I_2) < 1$, then there is no solution that obeys the constraints[3] . Intuitively,

---

[2] Of course, rejecting the occurrence of a process or event is logically equivalent to accepting the occurrence of its negation. But usually either a process or its complement, and not both, can be characterized mathematically. So the logical equivalence breaks down in favor of practical considerations, such as describability. Intuitively, rejecting the presence of a shadow is child's play compared to accepting the presence of a nonshadow. How could the class of all (visually interesting) events-that-are-not-shadows possibly be characterized? That class is certainly a peculiar collection of odds and ends, including such diverse members as paths of fireflies and holes in the ground. Visual systems "don't care" if they're accepting processes or rejecting negations of processes; it only matters to us when we *characterize* what we think the system might be doing.

[3] Clearly the product of two numbers each greater than unity is greater than unity. And a number greater than unity raised to a power greater than unity is also greater than unity. The rejection condition $\min(I_1, I_2) < 1$ is just the contrapositive of the two statements above.

this decision about unsolvability is computationally easier than actually finding a solution.

3.2 The Strategy Applied to some Confounding Processes

Figure 2 illustrates intuitively how the rejection strategy will be applied. Each of the first three panels ($\hat{O}$, $\hat{H}$, $\hat{P}$) show the effects (in a graph of image intensity versus wavelength) of a surface orientation change, a highlight, or a change in pigment density in a planar patch of a single material. (Shadow changes $\hat{S}$ are similar to highlights.) As will be shown below, the effect of a surface orientation change is to cut down the direct illumination by a constant fraction. Highlight or gloss is a situation of increased image intensity at all wavelength. (A shadow is the opposite.) Finally, pigment density changes, such as the variations seen in the grain of wood, can be characterized as several light-absorbing filters in sequence. The effect of a stack of filters is to reduce the available light according to a power relation. (The concentration of a dye dispersed in liquid affects transmitted light according to a power law more precisely.) What all these natural processes have in common is that they act to increase or decrease image intensities across wavelength. Violation of those sorts of displacements can be used to reject the presence of shadow, orientation change, gloss or highlight, or change in pigment density. Therefore, if two spectral functions of image intensities are not related by one always lying above the other, then the functions come from two regions that might be composed of different materials. The lower right portion of figure 2 illustrates such a situation, which is a candidate for a material change.

In the next section we examine simple models of the confounding processes. In section 5, we address the problems: How many channels does a visual system require in order to reject the presence of particular processes in a scene? What sort of computations with available image intensities allow rejection of confounding processes?

4 The Physics Behind the Scenes

4.1 The Image Intensity Equation

When light is reflected from a surface into the eye, the image intensity $I(\lambda)$ depends on several factors. The surface properties of the object interact with the geometry of the viewing situation and
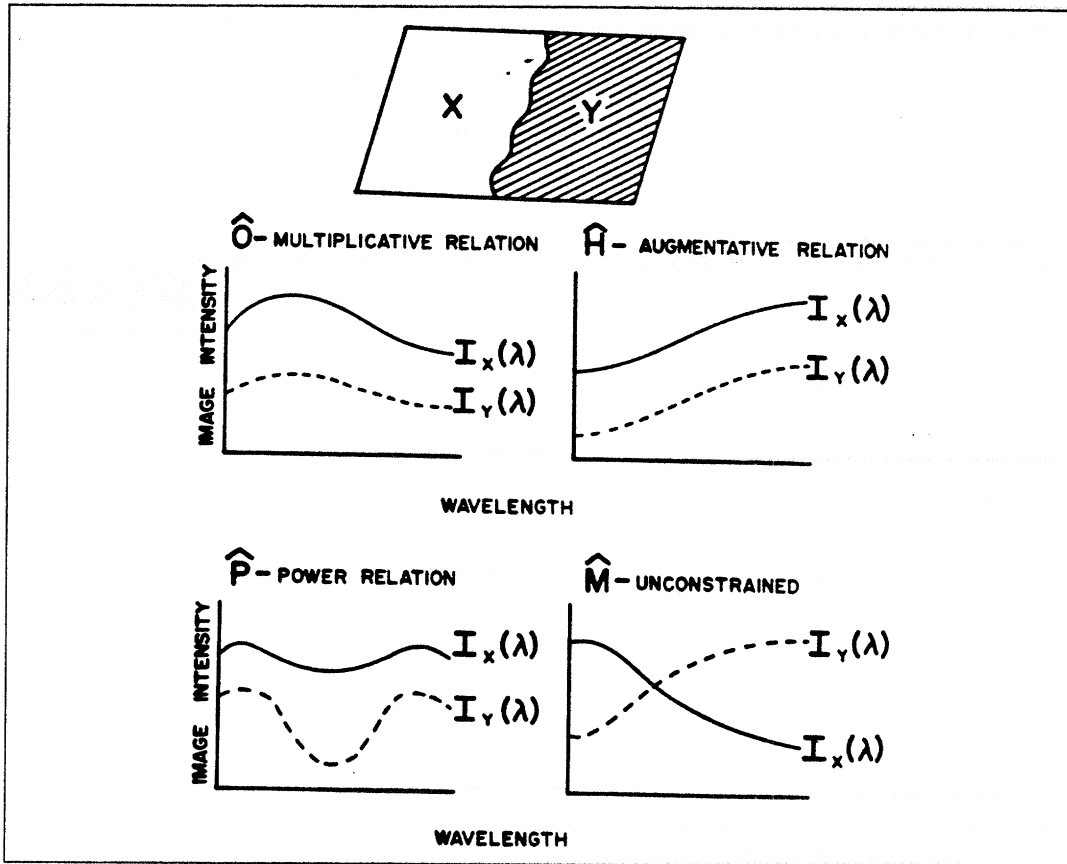
Figure 2. Three different relations are illustrated that can hold between two spectral functions of images intensities from regions X and Y. At top (left), a multiplicative relation (surface orientation change) is illustrated. At right is depicted an augmentative relation (highlight, for example). Bottom (left) is a power relation, typical of pigment density changes. At right is a change that is *not* an increase or decrease of intensity across wavelength. Such a situation is a possible material change.

the spectral nature of the light source to produce image intensity.[4] In the case of a matte surface[5], these effects combine multiplicatively to yield the following image intensity equation:

$$I(\lambda) = \rho(\lambda)E(\lambda)[\mathbf{N} \cdot \mathbf{L}]R(\theta, \phi) \tag{1}$$

where $\rho(\lambda)$ is the reflectance or albedo of the surface, $R(\theta, \phi)$ is the bidirectional reflectance distribution function (see Horn & Sjoberg, 1979) which describes properties of the surface dependent on its orientation with respect to viewer (ϕ) and light source (θ), and $[\mathbf{N} \cdot \mathbf{L}]$ is the angular relation between the surface normal, $\mathbf{N}$, and the illuminant direction $\mathbf{L}$. Any collection of multiple light sources is

---

[4]Technically, $I(\lambda)$ for a point x,y on the surface is called image *irradiance*. Image *intensity* takes into account other factors such as pupil size and luminance constants. These details are not important to the argument here, and will be omitted.

[5]For specular surfaces, a different image intensity equation holds. See Appendix I.
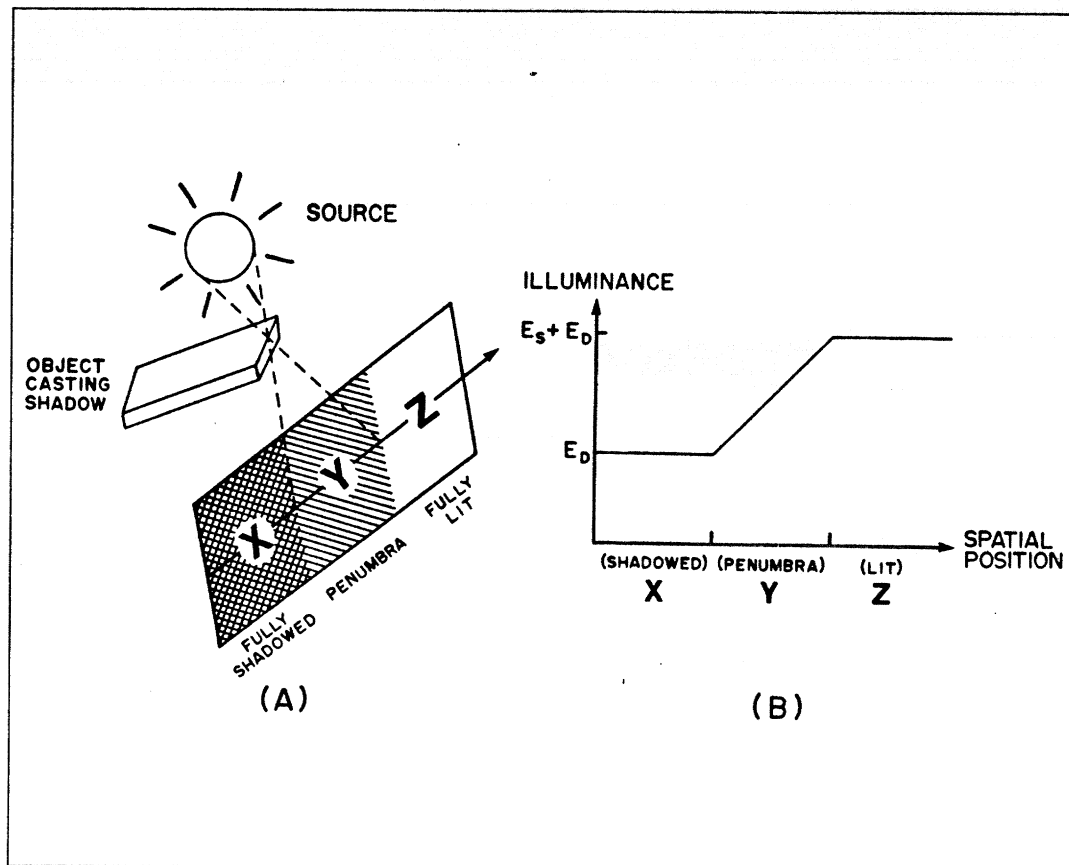
Figure 3. How shadow-with-penumbra is cast. A) Between the full lit and fully shadowed regions, some proportion of the source can reach the object. B) An (achromatic) image intensity profile is shown for the surface.

equivalent to a single source (called here the "synthetic source") on unshadowed portions of surfaces (Silver, 1980), so regardless of the complexity of the illumination, a single function $E(\lambda)$, together with directions θ and ϕ, will characterize the direction and spectral nature of the illuminant for unshadowed surfaces.

The effect of many natural processes that affect images can usually be characterized simply as acting on one or more of the multiplicative factors in equation (1). For example, a shadow corresponds principally to a reduction of the illuminant $E(\lambda)$. Surface orientation changes correspond to to changes in the $[N \cdot L]$ term, and pigment density changes affect only the $\rho(\lambda)$ term. (Highlight or gloss requires that a specular term be added to equation (1).) All of the processes discussed above can occur in a region of a single material.

We now proceed to examine how common natural processes other than material changes will affect the image intensity equation, and what these confounding processes have in common that allows their rejection *en masse*.

4.2 Shadows

Suppose a shadow ($\hat{S}$) falls on a surface composed of a single material, and furthermore, suppose that the only changes across the surface are these due to shadow. (Specifically, there will be no changes in surface orientation or variations in pigment density on the surface.) Then the shadow can be described by the following equations:

$$I_{lit}(\lambda) = [E_S(\lambda) + E_D(\lambda)]\rho(\lambda)$$
$$I_{shade}(\lambda) = E_D(\lambda)\rho(\lambda)$$

$(2a)$

where $\rho(\lambda)$ is the albedo of the material. The image intensities, as functions of wavelength, from the lit and shadowed regions are $I_{lit}(\lambda)$ and $I_{shade}(\lambda)$, respectively. $E_D(\lambda)$ represents a diffuse component of illumination striking both lit and shadowed regions, and $E_S(\lambda)$ is some additional synthetic source striking only the lit region.[6]

If the illuminant is an extended source, like the sun, there will be penumbrae (see figure 3), and a third equation must be included:

$$I_{pen}(\lambda) = [\alpha E_S(\lambda) + E_D(\lambda)]\rho(\lambda)$$

$(2b)$

where α is a constant between 0 and 1, and $I_{pen}(\lambda)$ is the image intensity (as a function of wavelength) in some region in the penumbra. (The symbol $\hat{S}_{pen}$ will denote shadows with penumbrae.)

4.3 Surface Orientation Changes

Suppose that two image regions X and Y differ only in their surface orientations. That is, assume that X and Y receive the same illumination (no self-shading) and are composed of the same material. This process as been dentoed by the symbol $\hat{O}$. The patches X and Y that differ only in surface orientation have respective surface normals N_X and N_Y which form angles θ_X and θ_Y with the

[6]If light is available from the shadowed region, it arises from some collection of sources. According to Silver (1980), a collection of sources is equivalent to a single source provided that they all illuminate the same surface patch. We call this synthetic source $E_D(\lambda)$. Furthermore, this diffuse synthetic source acts as if it lies in the direction of the viewer's eye (Horn, 1975). Next consider the lit region. Again, by Silver's proof, there is a single source that is equivalent to the illumination reaching the lit region. Let's call this source $E_{S^*}(\lambda)$, and its direction L_{S^*}. Since we can see the lit region, we know that there is some component of L_{S^*} in the direction of the viewer. So we can perform a simple vector decomposition of the synthetic illuminant for the lit region into a component in the viewers direction (so this component is identical to the synthetic diffuse illuminant $E_D(\lambda)$ above), and some other component $E_S(\lambda)$ with a new direction L_S. So in equations (2a) $E_D(\lambda)$ abbreviates a product of a term representing the spectral flux of the illuminant, a bidirectional reflectance term, and a surface orientation term, with the illuminant direction L in the viewer's direction. Likewise, $E_S(\lambda)$ stands for the product of of the same three terms above, but in some synthetic direction L_S.

(synthetic) light source direction **L**, and angles ϕ_X and ϕ_Y with the viewer direction **V**. Using the image intensity equation (1), the expected image intensities from regions X and Y differing only in surface orientation are:

$$I_X(\lambda) = \rho(\lambda)E(\lambda)[\mathbf{N}_X \cdot \mathbf{L}]R(\theta_X, \phi_X)$$
$$I_Y(\lambda) = \rho(\lambda)E(\lambda)[\mathbf{N}_Y \cdot \mathbf{L}]R(\theta_Y, \phi_Y)$$

(3a)

where again $I_X(\lambda)$ and $I_Y(\lambda)$ are the image intensities from regions X and Y as functions of wavelength.

Both equations of (3a) have a pair of multiplicative terms not involving wavelength. Let $\beta_X = [\mathbf{N}_X \cdot \mathbf{L}]R(\theta_X, \phi_X)$, and let β_Y be similarly defined. Equations (3a) can now be simplified with this consolidation of constant factors:

$$I_X(\lambda) = \beta_X \rho(\lambda)E(\lambda)$$
$$I_Y(\lambda) = \beta_Y \rho(\lambda)E(\lambda)$$

(3b)

4.4 Changes in Pigment Density

The reflectance of materials is determined primarily by the density of a pigment in some embedding layer of the material, and the thickness of the pigment layer. The embedding layer in leaves, for example, is cellulose, and the usual pigment is chlorophyll. Some surfaces may be unevenly pigmented. The grain of a wooden table provides a good example of a change in pigment density across a surface. This sort of change will be labeled $\hat{P}$.[7]

Several different laws relate changes in pigment density to changes in reflectance. (See Judd & Wyszecki, 1963.) Kubelka and Munk (1931) formulated a law that deals with the thickness and density of pigment on a solid background. Beer's law describes the effect of concentration of a dye dispersed in a liquid on the transmittance of the liquid, or equivalently, the thickness of a series of transmitting filters. The laws all differ in their details, yet there are some themes common to both. The pigmentation laws describe a *smooth transition* from an unpigmented state to a state of saturated pigmentation. (Often, the unpigmented materials are white or grey, and darken as pigment density increases. Caucasians are usually pale in the winter, and darken gradually during the summer if allowed to frequent the beach.) More can be said about the smooth transition. Loosely speaking,

---

[7]The difference between pigment density changes and changes of material is important. $\hat{P}$ changes are changes in the density of a single pigment in a single sort of embedding material; $\hat{M}$ changes are changes between two different pigments and (or) two different embedding materials.

the *shape of the albedo function is preserved* by the pigment density change. Two properties of the change of the albedo function make the notion above more precise. First, for a material characterized by Kubelka-Munk analysis or Beer's law, a change in pigment density affects the albedo function by either strictly increasing or decreasing its value at all wavelengths. Second, two albedo functions of the same material (but having different pigment densities) are related monotonically[8]. The conditions above will be called the *normal pigmentation conditions,* and pigmentation processes which obey them will be called *normal.*

Beer's law (Wyszecki & Stiles, 1973) will serve as an example of a pigmentation law. Suppose X and Y are two regions that transmit light through stacks of identical filters. Suppose further, that X and Y have the same surface orientation, and differ only in the number of filters in the stack. Then the image intensities $I_X(\lambda)$ and $I_Y(\lambda)$ that we could expect to measure are:

$$I_X(\lambda) = [\rho(\lambda)]^{\gamma_X} E^*(\lambda)$$
$$I_Y(\lambda) = [\rho(\lambda)]^{\gamma_Y} E^*(\lambda) \qquad (4)$$

where $\rho(\lambda)$ is the transmittance of a thin filter, and γ_X and γ_Y are the number of such filters in regions X and Y, respectively. Since the viewer–object–illuminant geometry is identical in X and Y, the two multiplicative terms of the image intensity equation (1) that deal with scene geometry and do not appear in (4) can be considered "absorbed" as a constant factor in $E^*(\lambda)$.

Two facts about equations (4) warrant attention. First, a power relation is described between the reflectances of two regions of different pigment thicknesses. This is the simplest such pigment relation, and can be considered as a base from which more complex laws develop. Second, note that regardless of the function $\rho(\lambda)$, the first *normal pigmentation condition* holds. That is, $[\rho(\lambda)]^{\gamma_X} > [\rho(\lambda)]^{\gamma_Y}$, for all λ, or vice versa. Since the illuminant and scene geometry are identical in regions X and Y, the *normal pigmentation condition* above is preserved in the available image intensities. Similar relationships between pairs of albedoes will hold for more complicated pigmentation laws (Kubelka & Munk, 1931), because such laws require monotonicity.

4.5 Highlights

The image intensity equation (1) only applies to matte surfaces. Highlight or specularity, a condition when a surface acts as a partial mirror, is a common confounding process. In a highlighted region.

[8]Two single-valued functions f and g are *related monotonicaly* if $f(x) > f(y)$ implies $g(x) > g(y)$. In particular, two monotonically related functions have their local minima and maxima at the same values in their domain.

image intensities are due to both a reflection of the light source, and a matte component due to the albedo of the underlying material. Appendix I shows that if most of the illumination of a region is due to the source that is reflected in the highlight, then the highlight $\hat{H}$ can be described as a process that strictly augments image intensities at all wavelengths. That is, if the *normal highlight* condition holds, the image intensity measurable in a highlighted region is the same as that of a neighboring matte region *plus* some (always positive-valued) function of wavelength[9] . Note this is identical to the first of the *normal pigmentation conditions*. Highlight differs from pigment density change in that it is always a positive change, and it does not necessarily preserve the shape of the image intensity function of wavelength from a neighboring matte regions.

5 Details of Process Rejection: Spatial and Spectral Samples

Our computational strategy is to examine image regions and to attempt to reject them as arising from the action of a confounding process on a single material. How might rejections be made? If a process is lawful, as the major confounding processes are, then some system of equations characterizes the effects of the process. Given measurements of image intensities, if there is always a *plausible solution*[10] to the equations describing a process, then we can always mathematically *interpret* the image intensities as arising from that process. When we can always interpret image intensities as being due to a particular process, we can never *reject* the occurrence of that process. Whether or not there always exist plausible solutions depends on the number of samples being taken. In this section we examine the minimum information needed to be sometimes able to reject the presence of the major confounding processes. We will answer the questions: What is the minimum number of spatial and spectral samples needed to reject image intensities as arising from each of the confounding processes? How would increasing the number of samples increase the number of correct rejections?

(Spectral samples will at first be assumed to be taken at a single wavelength. In Appendix II, it is shown that essentially any sort of spectral sample preserves the results of the following sections, which are derived using narrowband samples.)

---

[9]Since the word "highlight" denotes a region of greater brightness than surrounding regions, the normal highlight condition is reasonable and necessary to avoid mathematical anomalies such as "highlights" which are darker than surrounding regions.

[10]A *plausible solution* to a set of equations is one which assigns meaningful values to physical variables. For example, a solution which assigns a number greater than one to a variable representing a reflectance is not physically meaningful and must be discarded

5.1 Rejecting the $\hat{S}$ Hypothesis

Impossibility of Rejecting $\hat{S}$ with One Spectral Sample. Suppose one narrowband spectral sample is taken at some wavelength λ_1. Two image intensities can then be measured: I_{1X} and I_{1Y}, the intensities at λ_1 in image regions X and Y (see fig. 2). Can we reject the possibility that the *only* difference between regions X and Y is due to a shadow? Since either X or Y can be lit, and the other in shade, there are two equations to examine. In more formal terms, we need to know if we can rule out the existence of a plausible solution to either of two discrete versions of equations (2a):

$$I_{1X} = (E_{1D} + E_{1S})\rho_1$$
$$I_{1Y} = E_{1D}\rho_1 \tag{5a}$$

corresponding to X lit and Y shadowed, or

$$I_{1X} = E_{1D}\rho_1$$
$$I_{1Y} = (E_{1D} + E_{1S})\rho_1 \tag{5b}$$

corresponding to X shadowed and Y lit. Note that ρ_1 is the albedo of the material at λ_1, and E_{1S} and E_{1D} are the strengths of the synthetic and diffuse components of the illuminant.

A *shadow interpretation* is some solution to equations (5a) or (5b) that assigns physically meaningful values to variables. If the measured values of I_{1X} and I_{1Y} can be shown to be inconsistent with both equations (5a) and (5b), then there canot be a shadow interpretation. On the other hand, if *any* values of I_{1X} and I_{1Y} yield a plausible solution to the equations, then there will *always* be a shadow interpretation. If there is always an $\hat{S}$ interpretation, rejection is impossible, and we must resort to a greater number of spectral samples.

Intuitively, we are given two measurements of image intensity at only one wavelength. We are faced with the question: Is there any pair of image intensities that *cannot* be construed as arising from lit and shadowed portions of the same material? Given two measurements, one will be darker than the other. The darker region can be interpreted as being a shadowed continuation of the lighter region. So any pair of measurements can be construed as a shadow. Therefore, there can be no rejection of the shadow interpretation with only a single spectral sample.

The casual argument above can be made rigorous. Is there ever a possibility that both equations (5a) *and* (5b) will have no solution? First we'll consider solutions to equations (5a). One solution is given as follows:

$$E_{1D} = \delta$$
$$\rho_1 = \frac{I_{1Y}}{\delta}$$
$$E_{1S} = \delta(\frac{I_{1X}}{I_{1Y}} - 1)$$

(6)

where δ is the parameter of the one-dimensional solution space of equations (5a). (Two equations in three variables almost always have a one-dimensional solution space.) This solution is acceptable or plausible as long as all the variables take positive values, and $\rho_1 \in (0, 1)$. The latter constraint is satisfied by restricting δ to the interval (I_{1Y}, ∞). Note that E_{1D} and ρ_1 are always positive, since δ is restricted as above, and the measurement I_{1Y} is positive. Therefore all variables will take on positive values if E_{1S} is positive, or, equivalently, if $I_{1X} > I_{1Y}$. So whenever $I_{1X} > I_{1Y}$, there is a plausible solution to (5a). By the symmetry of equations (5b), it is clear that whenever $I_{1Y} > I_{1X}$, equations (5b) will have a plausible solution. Therefore, there will *always* be a solution to either (5a) or (5b), regardless of the values of measured image intensities. So there is always a shadow interpretation for monochromatic measurements; the darker region can be construed as the shadowed one. Hence, monochromacy is inadequate to reject measured image intensities as arising from shadows[11] .

Two Spectral Samples.

Perhaps adding a second spectral sample will help. Will samples at two wavelengths from each of regions X and Y *sometimes* allow the rejection of the shadow hypothesis about the X-Y difference? Or will any sets of measurements *always* have a shadow interpretation, as in the monochromatic case?

Intuitively, dichromacy should allow some rejections of the shadow hypothesis. Suppose that in the first spectral sample, region X is lighter than region Y. Then the same ought to hold for the second spectral sample if Y is to be construed as a shadowed continuation of the same material that composes X. So if one region is lighter in the first spectral sample, but darker in the second spectral sample, we probably aren't looking at a shadow. This follows from noting that in the case of shadow, both regions reflect a diffuse component of illumination, while the lit region has a synthetic component in addition.

A more formal demonstration can be made, by rewriting the shadow equations (2a) in discrete

---

[11]The claim that monochromacy does not allow shadow rejections only applies to the type of computations described here. Of course there might be some type of *achromatic* computation, involving, say, texture, that would allow the rejection of shadow.

form for two spectral samples:

$$I_{1X} = (E_{1D} + E_{1S})\rho_1$$
$$I_{2X} = (E_{2D} + E_{2S})\rho_2$$
$$I_{1Y} = E_{1D}\rho_1$$
$$I_{2Y} = E_{2D}\rho_2$$

$$(7a)$$

corresponding to X lit and Y shadowed, or

$$I_{1X} = E_{1D}\rho_1$$
$$I_{2X} = E_{2D}\rho_2$$
$$I_{1Y} = (E_{1D} + E_{1S})\rho_1$$
$$I_{2Y} = (E_{2D} + E_{2S})\rho_2$$

$$(7b)$$

corresponding to X shadowed and Y lit.

Once again, the left-hand side terms are the measured image intensities, and the right-hand side contains the variables for which we attempt to solve the equations. Given any four measurements I_{1X}, I_{1Y}, I_{2X}, and I_{2Y}, will there always be a plausible solution to either equations (7a) or (7b)?

Consider first solutions to (7a). We have four equations in six unknowns, suggesting a two-parameter family of solutions.

Below is a solution to equations (7a), parameterized by δ and ϵ:

$$E_{1D} = \delta$$
$$E_{2D} = \epsilon$$
$$\rho_1 = \frac{I_{1Y}}{\delta}$$
$$\rho_2 = \frac{I_{2Y}}{\epsilon}$$
$$E_{1S} = \delta(\frac{I_{1X}}{I_{1Y}} - 1)$$
$$E_{2S} = \epsilon(\frac{I_{2X}}{I_{2Y}} - 1)$$

$$(8)$$

The constraints of the values of the variables are once again that ρ_1 and $\rho_2 \in (0, 1)$, and that all the variables take on positive values. So extending the results of the previous section on a single spectral sample, it is clear that whenever both $I_{1X} > I_{1Y}$ and $I_{2X} > I_{2Y}$, there is a solution to equations (7a). Symmetrically, whenever both $I_{1Y} > I_{1X}$ and $I_{2Y} > I_{2X}$, there is a solution to equations (7b). So there is a dichromatic rejection condition for shadows. If $(I_{1X} > I_{1Y}$ and $I_{2Y} > I_{2X})$ or $(I_{1Y} > I_{1X}$ and $I_{2X} > I_{2Y})$, there will be no plausible solution to equations (7a), nor to (7b), and the presence of shadow can be rejected.
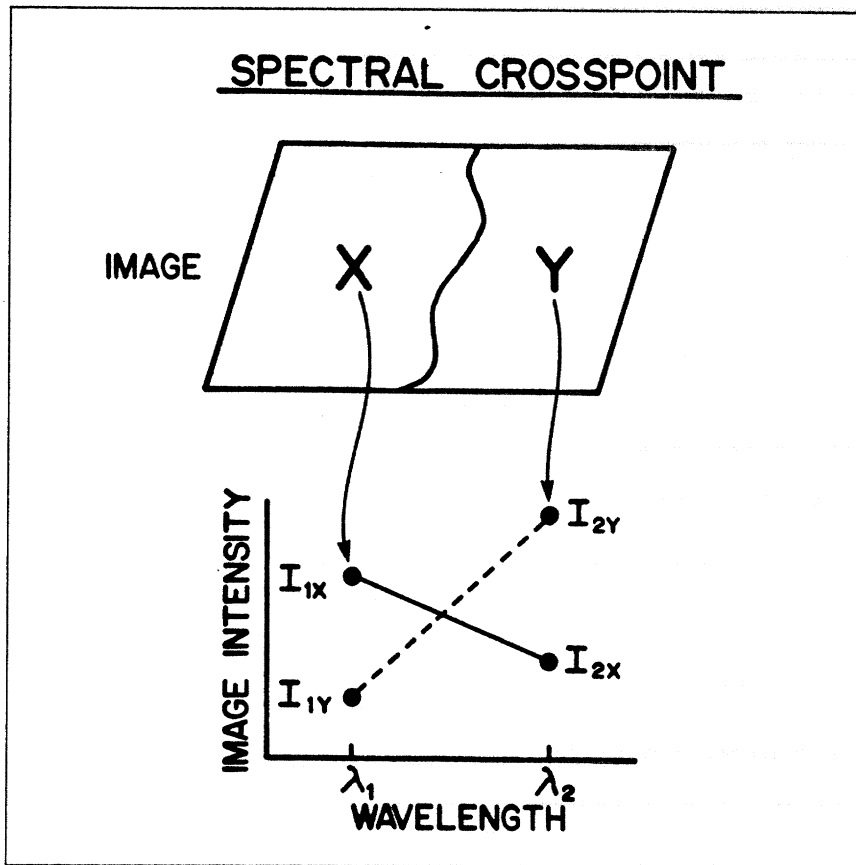
Figure 4. Rejection of $\hat{S}$ by crosspoint in plots from regions X and Y of image intensity versus wavelength.

There is a simple geometric interpretation of the dichromatic rejection condition: $\hat{S}$ can be rejected if the graphs of image intensity versus wavelength for regions X and Y intersect, or have a crosspoint. Furthermore, it has just been proved that *the crosspoint condition is unique and minimal.* It is minimal in the sense that it involves two spatial and two spectral samples; no smaller number of samples would do. It is unique in that, given two spatial and spectral samples, only the crosspoint condition holding among the four measurements can accurately lead to a rejection of $\hat{S}$. Whenever there is no crosspoint, a shadow interpretation, possibly wrong[12] , can be found. A final point. Given two line segments (two spectral samples at two spatial locations yields two line segments in our system of representation) there is only one topological property that they have: intersection (or non-intersection). Fortunately, intersection, or crosspoint, has been shown to be a physically interesting condition in this problem, as well as a topologically interesting one. Figure 4 illustrates the dichromatic rejection condition for $\hat{S}$.

[12]A *wrong interpretation* is one that assigns values to physical variables that do not correspond to the actual values. Interpretations of sensory data are discussed in more detail in Richards *et al.*, 1981.

n-chromacy. It is easy to extend the dichromacy result above to the general case of n spectral samples $(\lambda_1, \ldots, \lambda_n)$. Imagine taking pairs of these spectral samples and plotting a line segment from region X and another segment from region Y. If any pair of spectral samples yields a crosspoint (the line segments intersect), then $\hat{S}$ can be rejected for the X–Y change. More formally, we can reject $\hat{S}$ n-chromatically if there exist $j \neq k$ such that the line segment from I_{jX} to I_{kX} intersects the line segment from I_{jY} to I_{kY}. Equivalently, we can reject $\hat{S}$ n-chromatically if there exist $j \neq k$ such that $(I_{jX} - I_{jY})(I_{kX} - I_{kY}) < 0$.

5.3   Rejecting the $\hat{O}$ Hypothesis

One Spectral Sample. The surface orientation change equations (3b) can be re-expressed as a single equation:

$$\frac{I_X(\lambda)}{I_Y(\lambda)} = \frac{\beta_X}{\beta_Y} \tag{9}$$

Equation (9) describes a simple proportionality holding over wavelength for image intensities. Clearly, if only a single spectral sample is taken at each of X and Y, a trivial proportionality will hold between the measurements I_{1X} and I_{1Y}, regardless of their values. (That is, for all measurements of image intensities, there are appropriate values of β_X and β_Y so that a discrete version of equation (9) holds.) Hence a surface orientation change cannot be rejected monochromatically.

Two Spectral Samples. Will a second spectral sample sometimes allow the rejection of image intensities as arising from a surface orientation change? And if so, under what conditions? Equation (9) can be rewritten in discrete form for two spectral samples:

$$\begin{aligned}\frac{I_{1X}}{I_{1Y}} &= \frac{\beta_X}{\beta_Y} \\ \frac{I_{2X}}{I_{2Y}} &= \frac{\beta_X}{\beta_Y}\end{aligned} \tag{10}$$

It is clear (by combining equations (10)) that we can reject the $\hat{O}$ hypothesis whenever

$$\frac{I_{1X}}{I_{1Y}} \neq \frac{I_{2X}}{I_{2Y}} \tag{11}$$

Note that the nonproportionality condition above is narrower than the crosspoint condition that we derived for shadows; a crosspoint allows the rejection of both shadow and surface orientation change interpretations of image intensities.

5.4 Rejecting the $\hat{P}$ Hypothesis

In the preceding two sections, relationships among image intensities (at two spectral samples) were derived such that the discovery of the relationship could be taken as evidence that certain confounding processes were not occurring in the image. The spectral crosspoint was the less strict of the two criteria derived. In this section, a different tack is taken. First, it is taken as obvious that monochromacy is inadequate to reject the presence of $\hat{P}$. Next, it will be pointed out that a spectral crosspoint is good evidence that the image intensities under consideration do not arise from a change in pigment density[13]. It suffices to show (normal) pigmentation does not produce crosspoints. But this follows immediately from the first of the *normal pigmentation conditions*. Specifically, equations (4) can be rewritten for two spectral samples as

$$\frac{I_{1X}}{I_{1Y}} = \frac{\rho_1^{\gamma X}}{\rho_1^{\gamma Y}}$$
$$\frac{I_{2X}}{I_{2Y}} = \frac{\rho_2^{\gamma X}}{\rho_2^{\gamma Y}} \tag{12}$$

which implies (by the first *normal pigmentation condition*) that if $\frac{I_{1X}}{I_{1Y}} > 1$, then $\frac{I_{2X}}{I_{2Y}} > 1$, or vice versa. Therefore, if $\frac{I_{1X}}{I_{1Y}} > 1$ and $\frac{I_{2X}}{I_{2Y}} < 1$ (one of two possible spectral crosspoint conditions), then the presence of a normal pigmentation process can be rejected.

5.5 Rejecting the $\hat{H}$ Hypothesis

The presence of a spectral crosspoint can also be taken as evidence that no normal highlight is present. Since a (normal) highlighted region has greater image intensity at all wavelengths than a neighboring matte region, no spectral crosspoints can arise from such a pair of regions. (Also note that shadow changes are formally equivalent to highlights, in that lit and higlighted regions correspond to striclty augmentative changes to image intensities in neighboring shadowed and matte regions, respectively.) Hence the spectral crosspoint allows the rejection of normal highlights.

---

[13]More precisely, it will be shown a crosspoint implies either (*not* $\hat{P}$) or ($\hat{P}$, but not a *normal* pigmentation process).

5.6 Rejecting Special Cases of $\hat{S}$, $\hat{O}$, and $\hat{H}$

Shadows with Penumbrae. In the event of an extended source, shadows with penumbrae ($\hat{S}_{pen}$) can be rejected monochromatically (see Appendix II), but it should be noted that monochromatic rejection of the penumbra case requires three independent spatial samples, in contrast to the two spatial samples needed for dichromatic rejection of $\hat{S}$ (point source). That is, for sharp (point–source–generated) shadows, two spatial and two spectral samples are the minimum needed for rejection. For shadows with penumbrae, three spatial samples at a single wavelength suffice for rejection. *Nonmonotonicity* of the discrete plot of image intensity versus postion in the image at any wavelength is sufficient for rejecting $\hat{S}_{pen}$. (Nonmonotonicity is when three collinear spatial samples are neither strictly increasing or decreasing. This is identical to the condition that the center spatial sample of the collinear three has the greatest intensity, which cannot occur when the line passes through lit, penumbra, and shaded regions.)

"Simple" surface orientation changes. A special class of $\hat{O}$ changes can be rejected with a single spectral sample taken over three spatial regions. Let $\hat{O}_{simple}$ be the subset of $\hat{O}$ changes involving no inflections. $\hat{O}_{simple}$ is similar to $\hat{S}_{pen}$ in that both describe spatially monotonic processes. Clearly then, $\hat{O}_{simple}$ can also be rejected monochromatically using three spatial regions with the *nonmonotonicity* test. (See Appendix II, which treats the shadow–with–penumbra case, $\hat{S}_{pen}$.)

Highlights and Monchromacy. Consider three collinear samples passing through a normal highlight such that the center sample is taken in the highlight, and the end samples are taken in nearby matte regions of the same material. A luminance profile (at a single spectral sample) shaped like an upside-down vee can be expected for our piecewise linear representation. Therefore, a monotonic luminance profile, or a vee-shaped profile, can allow the rejection of the possibility that the three samples correspond to matte–highlight–matte regions.

5.7 Summary of Rejecting Individual Confounding Processes

In the sections above, it was shown that a single spectral sample was not sufficient (except in special cases involving three spatial samples) to reject image intensities as the effect of a single major confounding process. Two spectral samples, however, were proven adequate to sometimes allow rejection of image intensities as the effect of a single process. The strictest rejection criterion occurred with surface orientation changes; nonproportionality was shown sufficient for rejection of $\hat{O}$. The

broadest criterion was the spectral crosspoint, derived from the shadow equations. Since shadow and (normal) highlight are similar in being processes that allow any strictly augmentative change over wavelength, the spectral crosspoint also allows the rejection of highlight. Furthermore, the crosspoint was shown to be the unique and minimal rejection criterion for $\hat{S}$. The crosspoint is a special case of nonproportionality, and hence allows the rejection of surface orientation changes as well. (And the crosspoint is more secure than the nonproportionality condition in situations of noisy measurements of image intensity.) Changes in pigment density are strictly augmentative (or subtractive) changes over wavelength, and therefore cannot cause crosspoints. So the crossoint is sufficient to reject $\hat{P}$. Since each of the major confounding processes $\hat{S}, \hat{H}, \hat{O}$, and $\hat{P}$ can be rejected dichromatically by the presence of a spectral crosspoint, the crosspoint criterion for rejection of single confounding processes is adopted.

5.8 Rejecting Combinations of Processes

The discussion above focused on what happens when a *single* confounding process occurs. But by inspecting any photograph, it is obvious that arbitrarily chosen neighboring regions of the image may depict elements of the scene which differ in pigment density, surface orientation, and shadowing. (In particular, joint occurrences of $\hat{O}$ and $\hat{S}$ are common. Architecture provides examples in which one face of a polyhedral structure shades another.)

How can the analysis be extended to cover instances of *combinations of processes*? Perhaps we can restrict our attention to some subset of neighboring image regions in which material–change–mimicking conspiracies among the confounding processes are unlikely.

Which Neighboring Image Regions should be Examined?. It would be computationally exhausting to examine all possible neighboring image regions for material changes. And unnecessary as well. Material changes invariably produce with luminance discontinuities or edges in the image. If finding material changes is the goal being pursued, there's no point in looking for them where they're not going to be. That is, for the purpose of discovering material changes, "neighboring regions" should be taken to mean regions separated by edges.

Can several confounding processes coincide at a single edge? As pointed out above, shadows and surface orientation changes can occur at a single edge. But aside from this case of self-shading, it is generally believed that at an edge, a single process predominates (Marr & Hildreth, 1980; Marr, 1982).

(In the case of self-shading, no crosspoint can occur. By examining the shadow and surface orientation change equations, it can be seen that a joint occurrence of $\hat{S}$ and $\hat{O}$ is mathematically equivalent to an occurrence of $\hat{S}$ alone. And $\hat{S}$ changes cannot induce crosspoints. So self-shading cannot cause crosspoints.)

5.9 Accepting $\hat{M}$ by Default

When a crosspoint occurs across an edge, we know it does not arise from one of our confounding processes. If the reasoning above is correct, it is highly unlikely that such a crosspoint is caused by a peculiar conspiracy among $\hat{S}$, $\hat{H}$, $\hat{O}$, and $\hat{P}$ across an edge. It is worth asking if we have fully captured the range of possible confounding processes by our models of $\hat{S}, \hat{H}, \hat{O}$, and $\hat{P}$. All that could be said was that a highlighted region had greater image intensity at all wavelengths than a nearby matte region of the same material. The same is true for lit versus shadowed regions. And a slightly more stringent condition held for pigment density changes. It is clear from the discussion above that *any process that acts on a scene variable(s) (that is a function of wavelength) by strictly increasing or decreasing its value (at all wavelengths)* can be rejected with the spectral crosspoint. Thus, when a spectral crosspoint occurs across an edge, it is almost always due to a material change.

5.10 False Targets: Effectiveness of the Spectral Crosspoint in $\hat{M}$ Detection

A word should be said about the theory of $\hat{M}$ detection presented here. If our characterization of confounding processes is accurate, most spectral crosspoints across edges will be $\hat{M}$ changes. Exceptions include instances of highlight and pigment density changes that aren't *normal*, and rare (measure zero) coincidences of confounding processes at edges. So in signal detection language, the false target rate of crosspoints in $\hat{M}$ detection is low; most crosspoints will be material changes. As for the hit rate, it need not be the case that most $\hat{M}$ changes cause crosspoints in the image. Intuitively, only about half of the material changes in a scene will cause crosspoints. This should not be disturbing. The crosspoint computation is extremely easy, and provides immediate strong assertions. Of course a full theory of $\hat{M}$ detection is likely to involve other spectral computations, as well as nonspectral (textural) computations.

6 Relation to Psychophysics and Neurophysiology

The theory we have presented begins with the physics underlying several scene processes affecting image intensities. Simple analysis of equations describing the processes has shown the crosspoint is the unique criterion for rejecting each of the confounding processes, and thus a good tool for spectrally identifying material changes. But what has this to do with color vision? The implication of the results above is our *crosspoint conjecture* that biological visual systems interpret spectral crosspoints as material changes.

Psychophysical Evidence. An interesting fact about human color vision is that it becomes unstable under isoluminance conditions (Evans, 1948). Boundaries defined only by a chromatic change are weak; colors from one side of such a boundary are likely to invade the other. Color vision seems to require luminance discontinuities for stability. It was argued (section 5.8) that the only place in an image it makes sense to seek material changes is at edges. If color vision were concerned with the detection of material changes, then it would be an efficient system only if its computational resources were employed at edges. Besides, spectral computations carried out in arbitrary neighboring regions would be difficult, if not impossible, since many confounding processes could be acting together between the two regions. The number of hypotheses to be entertained would become combinatorically nasty (that is, all possible subsets of $\hat{S}$, $\hat{H}$, $\hat{O}$, and $\hat{P}$), and the biological value of enduring such tedium is dubious.

Number of Spectral Channels. Biological color vision systems are usually di- or trichromatic (Walls, 1942). Why are there no pentachromatic systems, say? The theory suggests an answer. Dichromacy was shown sufficient to enable a system to make strong $\hat{M}$ assertions across edges. Adding a third spectral sample increases the chances of finding a crosspoint. That is, trichromacy seems to provide higher hit and lower miss rates in $\hat{M}$ detection than dichromacy. Why not add spectral channels *ad infinitum*? The spectral reflectances of natural objects are almost always functions that change slowly over wavelength (Krinov, 1971). (There appear to be rarely more than three extremal points in reflectance function of wavelength.) It seems unlikely that a tetrachromatic system would detect enough additional crosspoints to be evolutionarily advantageous.

Comment on Physiology and Neurophysiology.

Trying to design an operator to detect spectral crosspoints provides some insights into neurophysiology. Although biological photopigments are broadband (Dartnall,1962), Appendix III
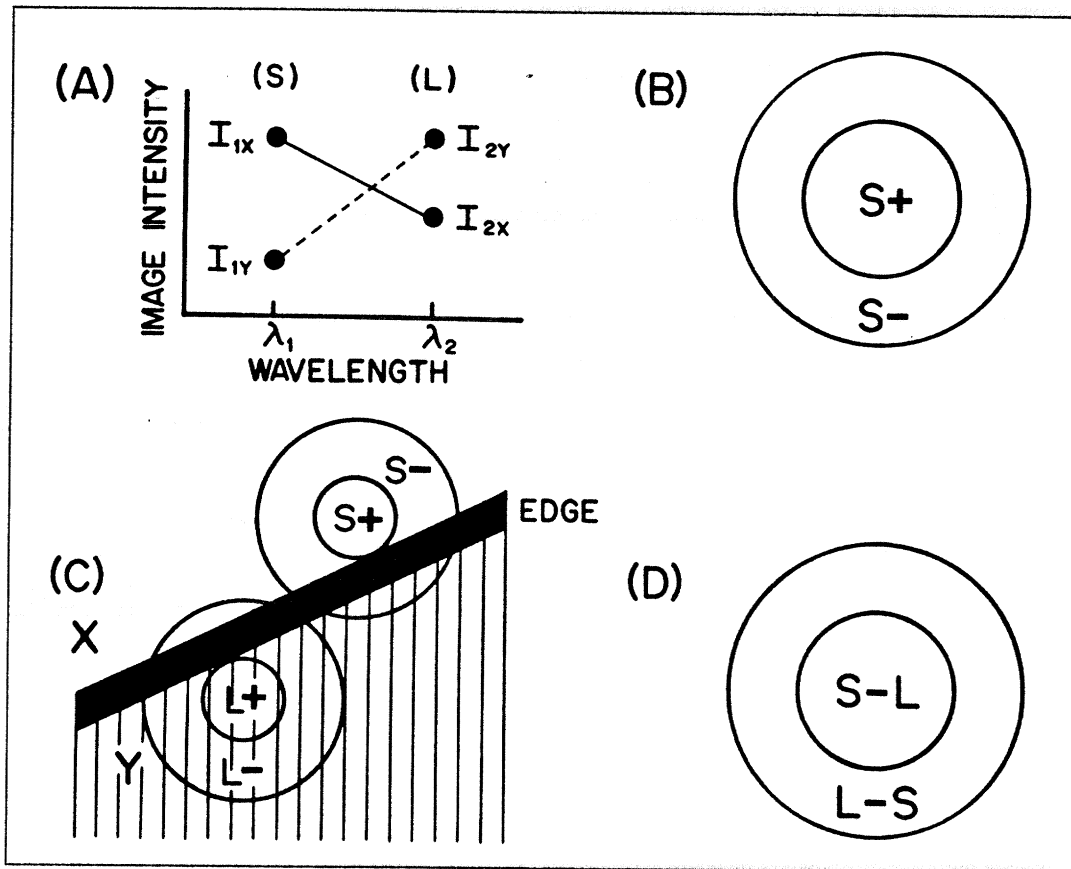
Figure 5. Designing an operator for detecting crosspoints. A) A crosspoint. B) A "Land unit," a circularly symmetric operator that compares image intensities at a single spectral sample (here short wavelength, or S), over two spatial regions. C) The S^+EL^+ crosspoint detector. Two different Land units, S and L, are combined with a logical **and** across an edge, E. A nonzero response indicates a crosspoint. D) The S^+EL^+ unit resembles the double opponent unit. See text.

shows that the particular shape of the photopigments used for taking spectral samples, and their degree of overlap, have no effect on the basic crosspoint finding.

The crosspoint (fig. 5a) is equivalent to the following inequality among image intensities:

$$(I_{1X} - I_{1Y})(I_{2X} - I_{2Y}) < 0 \qquad (13a)$$

which is identical to

$$(I_{1X} - I_{1Y})(-I_{2X} + I_{2Y}) > 0 \qquad (13b)$$

Suppose we desire a crosspoint operator that is circularly symmetric (for tessellation efficiency, say). Let the photopigments be S and L, having their respective λ_{max}'s at short and long wavelengths (see

fig. 5a). Equation (13b) can be rewritten:

$$(S_X - S_Y)(-L_X + L_Y) > 0 \tag{13c}$$

where S_X denotes image intensity in region X viewed through photopigment S, and so on.

The operator depicted in fig. 5b is ideal for computing the $(S_X - S_Y)$ factor in equation (13c). This sort of unit we call a "Land unit" since it was suggested by Land's (1977) retinex theory of color vision. The Land unit has zero output in a homogeneous field, and maximum output when the center receives more energy at short wavelengths than the surround.

Obviously, the firing of a single Land unit, even a maximal firing, does not provide information about the presence or absence of a crosspoint. A single unit can only inform about events at single spectral sample. Consider, however, the nonzero firing of a pair of Land units across an edge, as shown in fig. 5c. The S Land unit's nonzero firing implies $S_X > S_Y$; the L Land unit's firing implies $L_Y > L_X$. The conjunction, or logical anding, of the two conditions over and edge provides a test for crosspoint. We propose, then, a crosspoint detector composed of the conjunction of two (spectrally different) Land units on opposite sides of an edge. Note that without the requirement of an edge, this unit might respond to gradations of image intensities due to the effects of confounding processes. We call this operator an S^+EL^+ unit[14] .

Two further points should be raised about crosspoint detection with the SEL operator. First, the unit shown in fig. 5c only detects the type of crosspoint shown in fig. 5a. If Land units are constrained to have only positive response, as neurons are, a second type of unit, L^+ES^+, would be required to detect the other type of crosspoint ($I_{1X} < I_{1Y}$).

Second, a positive response of an on-center off-surround Land unit on one side of an edge implies a positive response of an off-center on-surround unit on the other side of the edge. In fig. 5c, imagine replacing the L^+ unit on the Y side of the edge with an L^- unit (off-center) on the X side. Suppose the L^- unit and an S^+ unit, both on the X side of an edge, were logically anded in a single unit. Such a unit might be sketched as in fig. 5d; it resembles a double-opponent cell, found in the retina of goldfish and in the primary visual cortex of primates (Daw, 1972)[15] . A final point should be

---

[14]This unit can be compared to Marr and Ullman's (1981) S^+TS^- unit for directional selectivity. The two units are similar in that they have three components, the outer two of which are "polar" forms of the same computation (S^+ and S^- differ in sign of contrast; S^+ and L^+ are different spectral samples. Also, the center components of the two units perform more sophisticated computations than the outer components.

[15]The unit depicted in fig. 5d may be misleading in suggesting an *arithmetic* sum of two Land units, rather than a *logical* anding. Only detailed quantitative neurophysiological study of such double-opponent units can reveal the computation they perform.

made about the double–opponent operator. It does not produce maximum output when it precisely straddles a edge with spectral crosspoint. The operator's largest response occurs when it is slightly offset from the material change boundary. This is reminiscent of properites of the $\nabla^2 G$ convolution operator[16] for luminance discontinuities (Marr & Hildreth, 1980).

7 Conclusion

Color vision systems evolved to solve certain problems in making sense of natural images. Natural images are complicated things, however, caused by a myriad of processes. While first intuition might be to study color vision over a simplified or restricted domain, we feel much can be understood by considering the more complicated natural domain. Certain regularities emerge from the confusion of the world, and these lawful relationships can be exploited as assumptions in the solution of sensory problems. Even given assumptions about the physics underlying scenes, natural images are still complicated. While ambitious goals can be successfully pursued in limited domains (Land, 1977), we feel that generally, more modest goals are appropriate to complicated domains. Our simple goal is to detect changes of material in an image using spectral information.

One unusual characteristic of material changes is that they are unconstrained. Since an unconstrained process cannot be sought directly (for what equations could we seek solutions?), we resort to inferring material changes by rejecting a class of processes that confound spectral image intensities. Rejecting the presence of a lawful process, it was pointed out, is often computationally much simpler then solving explicitly for the variables of the process.

First, a computation based on a single spectral sample was sought that would allow, at least on some occasions, the rejection of the presence of one of the major confounding processes. But it was shown that no such computation exists (of the type discussed here); *all* neighboring pairs of regions look like lit and shaded portions of the same surface when viewed monochromatically. (It was shown, however, that special classes of confounding processes $\hat{S}_{pen}$ and $\hat{O}_{simple}$, as well as $\hat{H}$, could be rejected monochromatically using three spatial regions.) Next, it was found that adding a second spectral sample (for two spatial regions) did allow us to sometimes be able to reject the presence of a confounding processes. The spectral crosspoint was the broadest criterion derived, and it was shown that its discovery allowed the rejection of each of the confounding processes $\hat{S}$, $\hat{H}$, $\hat{O}$, and $\hat{P}$. (Furthermore, the spectral crosspoint was proved the *unique and minimal* criterion for the rejection

---

[16]The laplacian of a two-dimensional gaussian.

of $\hat{S}$.) It was argued that by restricting the search for spectral crosspoints to edges, that complications arising from the joint occurrence of several confounding processes were unlikely. (And if several confounding processes did occur coincidentally, a crosspoint could only be generated under very unusual circumstances of illumination.) The spectral crosspoint allows us to say with great confidence that image intensities do not arise from $\hat{S}$, $\hat{H}$, $\hat{O}$, $\hat{P}$, or any of a large class of possible confounding processes. The simultaneous rejection of this large class of processes allows us to infer with great confidence that a material change is taking place. Our *crosspoint conjecture* is that biological visual sytems interpret crosspoints as material changes.

When no crosspoint occurs, more sophisticated computations are still possible (Richards, Rubin & Hoffman, 1981). The crosspoint strategy is not foolproof. While material assertions will be almost always correct, there will be many missed material changes. Not all material changes produce spectral crosspoints. A good strategy for visual systems would be to locate the maximum absorption frequencies of their photopigments in order that a maximum number of crosspoints be detected[17]. A study of the reflectances of natural objects could perhaps reveal if biological photopigments are located in wavelength in such a manner as to maximize the detection of crosspoints in an organism's environment.

Finally, it was shown that an operator for the detection of crosspoints, constrained to use only simple arithmetic functions, to output only positive values, and to be circularly symmetrical, is the S^+EL^+ unit, which performs a logical **and** of two spectrally different units across and edge. The SEL unit resembles the double-opponent cell commonly found in biological color vision systems.

APPENDIX I
Details about Highlights

The image intensity equation (1) does not apply to highlight. Highlight occurs on certain surfaces when two conditions hold. First, the viewer direction, surface normal, and illuminant direction must be approximately coplanar. Second, $\phi \approx \theta$; that is, the angle of incident illumination must be nearly equal to the angle of reflection to the viewer. When the two conditions above hold, certain materials will display mirrorlike qualities, and the illumination source will be imaged on the surface. Let $E_{source}(\lambda)$ be the source that is imaged in the highlight. We can expect the following image intensities

---

[17]Also, the photopigments should be located to minimize "doublecrossings," or situations in which a pair of complementary and cancelling crosspoints occur between the spectral sample points. Such situations would cause material changes to be missed.

from two neighboring regions of the same surface orientation that differ only in that one region has a highlight:

$$I_{matte}(\lambda) = (E_{source}(\lambda) + E_{other}(\lambda))\rho(\lambda)$$
$$I_{highlight}(\lambda) = \delta E_{source}(\lambda) + (1 - \delta)(E_{source}(\lambda) + E_{other}(\lambda))\rho(\lambda) \tag{15a}$$

where $I_{matte}(\lambda)$ and $I_{highlight}(\lambda)$ are the image intensities expected from the matte and highlighted regions, and $E_{other}(\lambda)$ represents all illumination other than the source $E_{source}(\lambda)$ reflected in the highlight. Note in the equation for the highlighted region that there are matte and specular components (Evans, 1948; Horn, 1977) in some linear combination determined by fraction δ. The specular component is just some coefficient multiplied by the source; the albedo plays no role. The matte component involves all illumination, not just the direct source, as well as the albedo $\rho(\lambda)$ of the material.

Next, note the highlight equation can be rewritten as follows:

$$I_{highlight}(\lambda) = \delta L(\lambda) + I_{matte}(\lambda) \tag{15b}$$

where $L(\lambda) = E_{source}(\lambda) - [E_{source}(\lambda) + E_{other}(\lambda)]\rho(\lambda)$. Therefore, the highlight becomes a purely additive process whenever $L(\lambda) > 0$. We shall call this condition the *normal highlight* condition. It is equivalent to the following:

$$\frac{E_{source}(\lambda)}{E_{source}(\lambda) + E_{other}(\lambda)} > \rho(\lambda) \tag{15c}$$

for all λ in the visually useful range.

Since reflectances of surfaces usually have maxima around .7 (Krinov, 1971), the *normal highlight* criterion above is, loosely speaking, that the source provide a little more than twice the illumination than the diffuse light. Normal sunlit scenes will obey this criterion.

APPENDIX II

Shadows with Penumbrae ($\hat{S}_{pen}$): Rejection with One Spectral Sample

Here we seek to rule out the existence of a solution to either of the following discrete equations:

$$I_{1,X} = (E_{1S} + E_{1D})\rho_1$$
$$I_{1,Y} = (\alpha E_{1S} + E_{1D})\rho_1 \tag{16a}$$
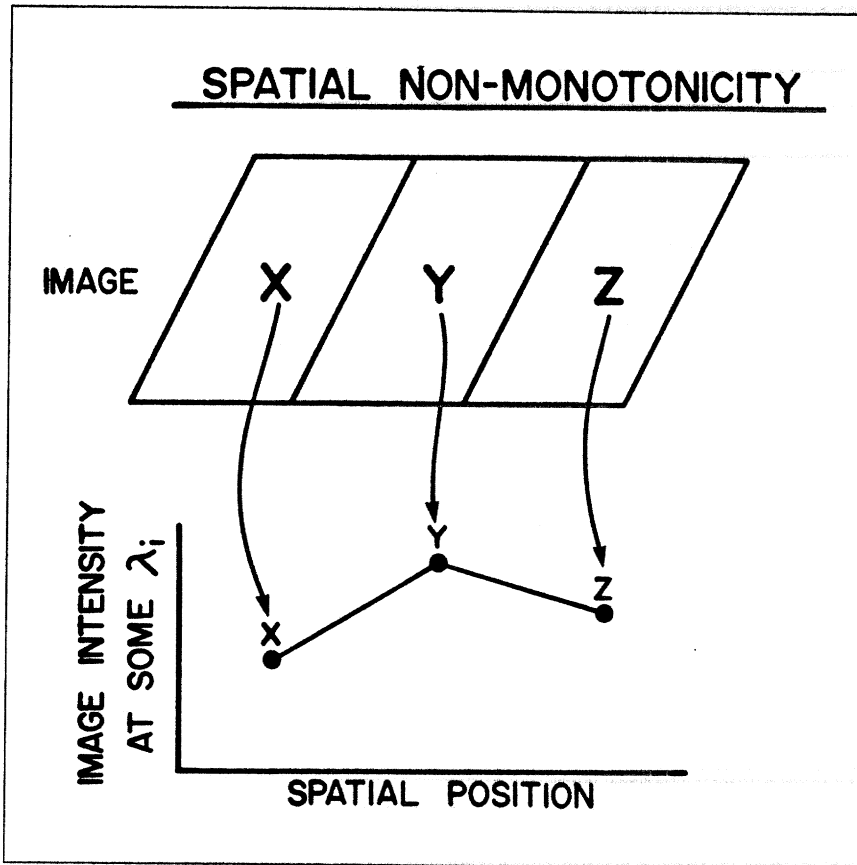$$I_{1,Z} = E_{1D}\rho_1$$

SPATIAL NON-MONOTONICITY

Figure 6. Rejecting $\hat{S}_{pen}$ by the nonmonotonicity of the discrete plot of image intensity versus spatial position at some wavelength.

where X, Y and Z denote three collinear spatial regions being tested for their correspondence to lit, penumbra, and shaded regions, respectively, of a shadow with penumbra, or

$$
\begin{aligned}
I_{1X} &= E_{1D}\rho_1 \\
I_{1Y} &= (aE_{1S} + E_{1D})\rho_1 \\
I_{1Z} &= (E_{1S} + E_{1D})\rho_1
\end{aligned}
\tag{16b}
$$

corresponding to the other case that Z is the lit region and X is shaded.

A solution to equations (16a), parameterized by δ is given below:

$$
\begin{aligned}
E_{1D} &= \delta \\
\rho_1 &= \frac{I_{1,Z}}{\delta} \\
E_{1S} &= \delta\left(\frac{I_{1,X}}{I_{1,Z}} - 1\right) \\
a &= \frac{I_{1,Y} - I_{1,Z}}{I_{1,X} - I_{1,Z}}
\end{aligned}
\tag{17}
$$

The constraints are that all the variables be positive, and $\alpha \in (0, 1)$. E_{1S} is positive if $\frac{I_{1,X}}{I_{1,Z}} - 1 > 0$ Now the only interesting restriction is that $\alpha \in (\mathbf{0}, \mathbf{1})$.

$$\alpha > 0 \quad \text{implies} \quad I_{1,Y} - I_{1,Z} > 0 \quad \text{implies} \quad I_{1,Y} > I_{1,Z}$$
$$\alpha < 1 \quad \text{implies} \quad I_{1,X} - I_{1,Z} > I_{1,Y} - I_{1,Z} \quad \text{implies} \quad I_{1,X} > I_{1,Y}$$

Therefore, any violation of the restriction $I_{1,X} > I_{1,Y} > I_{1,Z}$ implies the unsolvability of equations (16a). By the symmetry of the alternate equations (16b), it is clear they provide the restriction $I_{1,Z} > I_{1,Y} > I_{1X}$. The pair of restrictions from (16a) and (16b) together imply that the three collinear image intensities must be either strictly increasing or strictly decreasing; that is, they must be monotonic. This rejection condition of *nonmonotonicity* is illustrated in figure 6, and can be summarized as follows:

If the discrete plot of image intensity versus position across the $\hat{S}_{pen}$ candidate is nonmonotonic, we can reject $\hat{S}_{pen}$. Equivalently, we can reject $\hat{S}_{pen}$ if $(I_{1,X} - I_{1,Y})(I_{1,Y} - I_{1,Z}) < 0$.

It is clear that with additional spectral samples, the nonmonotonicity of the discrete plot of image intensity versus wavelength *at any wavelength* is sufficient to reject $\hat{S}_{pen}$

However, in practice, there is a problem with the type of operator suggested by the theory: at what scale should triplets of spatial regions be examined for nonmonotonicity? Shadows–with–penumbrae can occur with large variations in spatial extent in images. Should a whole range of triplets be examined? The computational complexity of such a task should make one wary of it. We take this scale problem to imply that, for practical use, the crosspoint condition is preferred.

APPENDIX III
Broadband spectral samples cannot induce crosspoints.

It will be shown that a crosspoint cannot arise in $\hat{P}$, $\hat{O}$, or $\hat{S}$ situations from an overlapping of samples.

1] Assume the change from X to Y is due to one of the confounding processes. Therefore, the continuous spectral energy distributions $I_X(\lambda)$ and $I_Y(\lambda)$ do not intersect. (Otherwise the proof in section 5.1 could be shown wrong by taking samples straddling the intersection.) Assume, without loss of generality, that $I_X(\lambda)$ is always greater than $I_Y(\lambda)$.

2] Assume that our two spectral samples will be measured by "photopigments" P_1 and P_2 where each P_i is simply a function mapping some wavelength interval into the unit interval $(0, 1)$.

3] The sample measured at λ_i in region X, S_{iX}, and in region Y, S_{iY}, are defined as follows.

$$S_{iX} = \int_{\lambda_{i1}}^{\lambda_{i2}} I_X(\lambda) P_i(\lambda)\, d\lambda$$

$$S_{iY} = \int_{\lambda_{i1}}^{\lambda_{i2}} I_Y(\lambda) P_i(\lambda)\, d\lambda \tag{18}$$

where $\lambda_{i2} > \lambda_{i1}$ and the interval $(\lambda_{i1}, \lambda_{i2})$ is the range of wavelengths over which the "photopigment" P_i is sensitive.

4] $S_{iX} > S_{iY}$, for all i, follows directly from the fact that $I_X(\lambda) > I_Y(\lambda)$, for all λ.

5] Therefore, spectral crosspoints cannot be induced by overlapping spectral samples.

REFERENCES

Dartnall, H.J.A., "The identity and distribution of visual pigments in the animal kingdom," Ch. 18 in *The Eye*, 2, H. Davson, ed., Academic Press, New York, 1962.

Daw, N.W., "Color-coded cells in goldfish, cat and rhesus monkey," *Investigative Ophthalmology* 11 (1972), 411-417.

Evans, R.M., *An Introduction to Color*, Wiley, New York, 1948.

Horn, B.K.P., "Image intensity understanding," *M.I.T. Artificial Intelligence Laboratory Memo 335* (1975).

Horn, B.K.P., "Understanding image intensities," *Artificial Intell.* 8 (1977), 201-231.

Horn, B.K.P. & Sjoberg, R.W., "Calculating the reflectance map," *Applied Optics* 18 (1979), 1770-1779.

Judd, D.B. & Wyszecki, G., *Color in Business, Science, and Industry*, John Wiley and Sons, Inc., New York, 1963.

Krinov, E.I.., *Spectral reflectance properties of natural formations*, National Research Council of Canada Technical Translation 439, 1971.

Kubelka, P. & Munk F., "Ein Beitrag zur Optik der Farbanstriche," *Z. tech. Physik.* **12** (1931), 593.

Land, E.H., "The retinex theory of color vision," *Scientific American* **218** (1977), 108-128.

Marr, D., *Vision: A computational investigation into the human representation and processing of visual information,* Freeman and Company, San Francisco, 1982.

Marr, D. & Hildreth, E., "Theory of edge detection," *Proc. R. Soc. Lond. B* **207** (1980), 187–217.

Marr, D. & Ullman, S., "Directional selectivity and its use in early visual processing," *Proc. R. Soc. Lond. B* **211** (1981), 151-180.

Richards, W.A., Rubin, J.M. & Hoffman, D.D., "Equation counting and the interpretation of sensory data," *M.I.T. Artificial Intelligence Laboratory Memo 614* (1981).

Silver, W., "Determining shape and reflectance using multiple images," (June, 1980).

Walls, G., *The Vertebrate Eye and its Adaptive Radiation,* Hafner Publishing Co., New York, 1942.

Wyszecki, G. & Stiles, W.S., *Color Science: Concepts and Methods, Quantitative Data and Formulas,* John Wiley & Sons, Inc., New York, 1967.