XXVIII. NEUROPHYSIOLOGY

Academic and Research Staff

Prof. J.Y. Lettvin
Dr. M.H. Brill
Dr. E.R. Gruberg
Dr. E.A. Newman
Dr. G.M. Plotkin
Dr. S.A. Raymond

Graduate Students

L.R. Carley
D.A. Cohen
C.L. Epstein
B. Howland
L.L. Linden
K.J. McLeod
L.L. Odette

1. TECTAL STUDIES OF AMBYSTOMA

National Institutes of Health (Training Grant 5 T01 EY00090)
Bell Laboratories (Grant)

Edward R. Gruberg, William Harris
[William Harris is with the Department of Neurobiology, Harvard Medical School, Boston, Mass.]

In an earlier study Gruberg and Solish showed that salamander somatosensory spino-tectal fibers took the same course and ended in the same layer of the intermediate tectal neuropil as 5-hydroxytryptamine (5-HT or serotonin) fibers as seen by Falck-Hillarp fluorescence. In order to determine whether the spino-tectal fibers are themselves serotonergic, we have carried out anatomical, biochemical, genetic and physiological tests. By using a specific 5-HT fiber poison, 5,7-dihydroxytryptamine (5,7-DHT), we have been able to abolish electrically recorded contralateral somatosensory units while the overlying visual units appear normal. 5,7-DHT poisoning also abolished 5-HT fluorescence in the intermediate tectal neuropil.

Genetically eyeless salamanders, or normal salamanders bilaterally enucleated early in development, have a shift in their somatosensory input to the superficial tectum. There is the same shift in Falck-Hillarp serotonin fluorescence. This shift is not due to a sprouting of 5-HT fibers. With the use of high-performance liquid chromatography, we have found the same amount of tectal serotonin in normal and eyeless salamanders.

We have traced the origins of the spino-tectal cells by using HRP tectal

injections. A majority of these cells lie in a thin lamina on the ventral border of the spinal gray. This is the same layer where serotonin cell bodies are also found. In addition, we have verified the results obtained with Falck-Hillarp fluorescence by injecting ^{3}H-serotonin into the tectum and subsequently carrying out radioautography.

Collectively, these results strongly suggest that at least some somatosensory fibers in the salamander tectum are serotonergic.

2. THE BASAL OPTIC SYSTEM

National Institutes of Health (Training Grant 5 T01 EY00090)
Bell Laboratories (Grant)

Edward R. Gruberg, Keith Grasse
[Keith Grasse is with the Department of Psychology, Dalhousie University, Halifax, Nova Scotia.]

We have begun experiments to investigate the anatomy and physiology of the basal optic area (BOA) in the frog. The BOA in the ventral tegmentum receives a direct input from the retina, consisting of the largest diameter fibers found in the optic nerve.

With the animal on its back and using a ventral penetration through the upper mouth, we have been able to record routinely single units in the BOA. Most units are directionally sensitive to moving stimuli. So far we have found two classes of units: those responsive preferentially to vertical moving stimuli and those to horizontal moving stimuli. A wide variety of sizes and shapes of stimuli elicit responses. Vertical units increase firing to either slow upward movement or slow downward movement; the opposite direction gives a null response. Horizontal units increase firing either to slow naso-temporal movement or to slow temporo-nasal movements. They, too, give null response in the opposite direction. All these units yield broad tuning curves of response vs direction of movement. The best response is obtained for most units when stimuli are moved with angular velocity in the range of 0.1° to 1°/sec.

We have also begun an HRP study of the connections of the BOA. Since the target

region is very small, we have developed a method for delivering volumes of 1/2 nanoliters or less of HRP solution. We have found ipsilateral projections to the BOA arising from three principal areas: the posterolateral tegmental field, the posterior nucleus of the thalamus, and a wide extent of the ventral thalamus, particularly in the anterior part. So far, we have found no clear contralateral input.

Publications

Gruberg, E.R., E.A. Newman, and P.H. Hartline, "The Python and the Rattlesnake: A Comparison of Infrared Trigemino-Tectal Techniques," Abstracts for Society for Neuroscience Meeting, Atlanta, Georgia, 1979, p. 708.

Gruberg, E.R., E. Kicliter, E.A. Newman, L. Kass, and P.H. Hartline, "Connections of the Tectum of the Rattlesnake Crotalus viridis: An HRP Study," J. Comp. Neurol. 188, 31-42 (1979).

Gruberg, E.R. and J.Y. Lettvin, "Anatomy and Physiology of the Binocular System in the Frog Rana pipiens," Brain Res., accepted for publication.

Newman, E.A., E.R. Gruberg, and P.H. Hartline, "The Infrared Trigemino-tectal Pathway in the Rattlesnake and in the Python," J. Comp. Neurol., accepted for publication.

3. RETINAL OPERATORS THAT NULL OUT RIGID 3-SPACE TRANSLATIONS

Bell Laboratories (Grant)

Michael H. Brill

Suppose an artificial planar retina of photosensors receives (via optical elements) an image of light reflected from a matte coplanar array of reflectances. Suppose the reflecting surface is parallel to the retina, is uniformly illuminated, and undergoes rigid translation in 3-space — a motion that is arbitrary but constrained so the edge of the surface never passes over the retina. Such a situation might be encountered by an airborne camera looking directly down at flat terrain — neglecting the earth's curvature. Without prior knowledge (such as ground speed in the above application), how can the retina detect that no motion other than a rigid translation is going on?

One solution to this kind of problem, proposed by Pitts and McCulloch,[1] is to generate all possible translated versions of the initial image internally, and match

them to the actual image as it evolves in time. If the instantaneous image is tolerably close to any translated version of the initial image, then it is construed as that translated version. One problem with this method is that parts of a translating image newly emerging into view of the retina are not matchable to a translated version of the initial image. This is rather easily solved by arranging that each instantaneous image be compared to the previous image rather than to the first image, and ignoring the edge of each image in the comparison. A more difficult problem is to do all the shifting and comparing in real time.

We have developed another method, not so general but easier computationally: The retina performs a nonlinear differential operation on the point-of-view-corrected light-energy image $h(x,y,t)$* such that the result of the operation is zero if the image is due to light reflected from a matte plane parallel to the retina undergoing 3-space translation. The differential operator is of order three, and is evaluated far enough away from the retinal edge so as not to incur artifacts. The operator is derived below.

First, the form of $h(x,y,t)$ is derived for the motion in question.

In setting up the geometry, the following specifications are made without loss of generality: In the 3-space (Cartesian coordinates η,ξ,ζ), the retina is at $\eta = 1$, and a 3-space point is represented on the retina by conical projection through $(0,0,0)$. Retinal rectangular coordinates (x,y) are chosen so that the x axis is parallel to the ξ axis in the 3-space, and the y axis is parallel to the ζ axis.

Then the projected image of a point (η,ξ,ζ) on the retina is the point

$$x = \frac{\xi}{\eta}, \qquad y = \frac{\zeta}{\eta}. \tag{1}$$

Suppose these coordinates represent a point on an object at time t that had at $t = 0$ the coordinates $(\hat{\eta},\hat{\xi},\hat{\zeta}) = (\eta + \alpha, \xi + \beta, \zeta + \gamma)$. Then the image of the object point at $t = 0$ was at the retinal point

$$\hat{x} = \frac{\xi + \beta}{\eta + \alpha}, \qquad \hat{y} = \frac{\zeta + \gamma}{\eta + \alpha}. \tag{2}$$

Here α, β, and γ are unspecified functions of time t.

The ensemble of viewed points (η,ξ,ζ) lies on a plane whose equation at t = 0 is $\eta = s$ (where s is a constant). All these points are translated by the same vector (α,β,γ) at time t, whence

$$(\hat{x},\hat{y}) = \left(\frac{sx + \beta}{s + \alpha}, \frac{sy + \gamma}{s + \alpha} \right). \tag{3}$$

Suppose at time t the viewed-plane point (η,ξ,ζ) reflects light power $I(\xi,\zeta,t)$ uniformly through solid angle 2π. The retina is presumed to receive all the light from this point at $(x,y) = (\xi/\eta, \zeta/\eta)$, and the power received is

$$H(x,y,t) = \frac{I(\xi,\zeta,t)}{2\pi r^2} \cos\theta,$$

where r is the distance from (η,ξ,ζ) to $(1,x,y)$, and θ is the angle the light path makes with the normal to the retinal plane. If the retinal image is real, then

$$\cos\theta = \frac{\eta}{\sqrt{\eta^2 + \xi^2 + \zeta^2}} = \frac{1}{\sqrt{1 + x^2 + y^2}}$$

$$r^2 = (1+\eta)^2 + (x+\xi)^2 + (y+\zeta)^2 = (1+\eta)^2 (1+x^2+y^2).$$

Thus

$$H(x,y,t) = \frac{I(\xi,\zeta,t)}{2\pi(1+\eta)^2 (1+x^2+y^2)^{3/2}}. \tag{4}$$

Translation of the viewed plane corresponds to the identification

$$I(\xi,\zeta,t) = I(\xi+\beta,\zeta+\gamma,0) = I(\hat{\xi},\hat{\zeta},0).$$

From Eq. 4,

$$\begin{aligned} I(\xi,\zeta,t) &= 2\pi H(x,y,t)(1+\eta)^2 (1+x^2+y^2)^{3/2} \\ I(\hat{\xi},\hat{\zeta},0) &= 2\pi H(\hat{x},\hat{y},0)(1+\hat{\eta})^2 (1+\hat{x}^2+\hat{y}^2)^{3/2}. \end{aligned} \tag{5}$$

If one defines at each t the point-of-view-corrected image

$$h(x,y,t) = H(x,y,t)(1+x^2+y^2)^{3/2},$$

then

$$h(x,y,t) = \frac{(1+\hat{\eta})^2}{(1+\eta)^2}\, h(\hat{x},\hat{y},0) \equiv \frac{(1+\hat{\eta})^2}{(1+\eta)^2}\, f(\hat{x},\hat{y}). \tag{6}$$

From the initial condition $\eta = s$, Eq. 3, and the condition $\hat{\eta} = \eta + \alpha$, it follows that

$$h(x,y,t) = \left[1 + \frac{\alpha}{1+s}\right]^2 f\left(\frac{sx+\beta}{s+\alpha}, \frac{sy+\gamma}{s+\alpha}\right). \tag{7}$$

Now define $a \equiv s/(s+\alpha)$, $b \equiv \beta/(s+\alpha)$, $c = \gamma/(s+\alpha)$, and rewrite Eq. 7 as follows:

$$h(x,y,t) = \left[1 + \frac{s}{1+s}\left(\frac{1}{a} - 1\right)\right]^2 f(ax+b,\ ay+c). \tag{8}$$

This is the general form of h for the motion in question.

There are two important limits in which h takes particularly simple forms: When $s \ll 1$ (the <u>microscopic</u> limit), $h(x,y,t) = f(ax+b,\ ay+c)$; and when $s \gg 1$ (the <u>telescopic</u> limit), $h(x,y,t) = (1/a^2)\, f(ax+b,\ ay+c)$. For both of these forms, we now derive specific relations between the partial derivatives of h and a,b,c. Henceforth $\dot{a}$ will be the time derivative of a, and so on; h_x will be the partial derivative of h with respect to x, and so on; and f_1 will be the partial derivative of f with respect to its first argument, and so on.

In each limiting case, the strategy will be to solve for pure functions of time in terms of partial derivatives of h (readily computed by the retina), and then take spatial derivatives of these functions of time. These derivatives will be zero when h has the desired form. However, the same functions of the partial derivatives of h will <u>not</u> have this property when h does not have the desired form; their nonzero spatial derivatives will cue targets moving on a rigid background.

a. Microscopic Limit [$h(x,y,t) = f(ax+b, ay+c)$]

We begin by computing all first partial derivatives of h to get expressions that explicitly involve a, b, c, and their first time derivatives:

$$h_x = af_1; \qquad h_y = af_2$$
$$h_t = (\dot{a}x+\dot{b})f_1 + (\dot{a}y+\dot{c})f_2. \tag{9}$$

By removing f_1 and f_2 from these equations, and defining the pure-time functions $A = \dot{a}/a$, $B = \dot{b}/a$, $C = \dot{c}/a$, one obtains

$$h_t = A(xh_x + yh_y) + Bh_x + Ch_y. \tag{10}$$

As desired, this equation relates the partial derivatives of h — computable from the image — to pure-time functions characteristic of the motion. To generate enough equations at each point (x,y) to solve for A,B,C, we differentiate Eq. 10 with respect to x and y, thereby getting three linear equations in A,B,C with coefficients being functions of the partial derivatives of h:

$$h_t = A(xh_x + yh_y) + Bh_x + Ch_y$$
$$h_{tx} = A(xh_{xx} + yh_{xy} + h_x) + Bh_{xx} + Ch_{xy} \tag{11}$$
$$h_{ty} = A(xh_{xy} + yh_{yy} + h_y) + Bh_{xy} + Ch_{yy}.$$

This set of equations can be solved for A, B, and C (e.g., by Cramer's rule). A, B, and C are then functions of the partial derivatives of h whose gradients can be evaluated for any h. A nonzero gradient is a diagnostic for image changes that are not rigid translations in 3-space.

Do translating planes that are parallel to the retina produce the only visual inputs h that are nulled out by the above operators (which we henceforth call $\vec{Z}$)? We have not been able to determine yet whether these geometrical constraints are the only ones generating $h(x,y,t) = f(ax+b, ay+c)$. One can show, however, that this form for h is the only one satisfying $\vec{Z}h = 0$.

[Proof: First, note that Eqs. 11 represent the general first integration of $\vec{Z}h = 0$, but A, B, and C are now arbitrary functions of time, not necessarily deriving from a, b, and c. Second, note that a solution of Eq. 10 (the first of Eqs. 11) automatically solves the other two equations. Thus the only PDE to be solved is Eq. 10. To show the uniqueness of the form of Eq. 8 as a solution to Eq. 10, first construct from the first-integrated time functions A,B,C the functions a_1, b_1, c_1 such that

$$A = \frac{\dot{a}_1}{a_1}, \qquad B = \frac{\dot{b}_1}{a_1}, \qquad C = \frac{\dot{c}_1}{a_1}.$$

Functions with this property are

$$a_1 = e^{\int A\, dt}, \qquad b_1 = \int a_1 B\, dt, \qquad C_1 = \int a_1 C\, dt.$$

Then Eq. 10 becomes

$$a_1 h_t = \dot{a}_1(xh_x + yh_y) + \dot{b}_1 h_x + \dot{c}_1 h_y. \tag{12}$$

Finally, change variables in the above differential equation to the new variables $T = t$, $X = a_1 x + b_1$, $Y = a_1 y + c_1$. This transforms the partial derivatives to

$$\begin{aligned} h_t &= h_T + h_X(\dot{a}_1 x + \dot{b}_1) + h_Y(\dot{a}_1 y + \dot{c}_1) \\ h_x &= h_X a_1; \qquad h_y = h_Y a_1. \end{aligned} \tag{13}$$

Substitution of Eqs. 13 into Eqs. 12 yields the condition $h_T = 0$, whence h depends only on $X = a_1 x + b_1$, and $Y = a_1 y + c_1$.]

Equation 10 is a familiar differential form: $x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ is the Lie generator of isotropic dilation; $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ are the Lie generators of translation in x and y, respectively. Newman and Demus[2] (in a program of research that paralleled our own) also parametrized motion with the Lie group formalism, including a generator $x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$ for rotations as well as those above. One might note that, since they were dealing with a missile-tracking application, their dilation operator on h

might more appropriately have been replaced by $xh_x + yh_y - 2h$ as in (b) below.

Instead of differentiating their analogue of Eq. 10 to provide the necessary equations to solve for the pure-time variables, Newman and Demus generated nine equations for each retinal point by evaluating their analogue of Eq. 10 at the point itself and at the eight points surrounding it.

This illustrates an important freedom in our method, to be optimized in pattern-recognition applications: The differentiation operations generating the second and third of Eqs. 11 can be replaced by any two independent nonidentity spatial linear operations (such as convolutions). Optimizing these operators may reduce the noise attendant in digital computations approximating the derivative. Another strategy for noise suppression is to take the gradient of nonlinear functions $\phi(A,B,C)$ rather than of the pure-time functions A,B,C themselves. This freedom is also allowed in our formalism.

A final word should be said about the cases in which the inversion of Eq. 11 is impossible (e.g., when $h_x = h_y = 0$, which is the condition of a local spatial uniformity). If a determinant test shows that Eq. 11 is not invertible at a point, the operator at that point must be immediately zeroed and the inversion avoided. A,B,C can then be interpolated in the singular region. This circumvents the singularity, and still permits the advancing (or retreating) edge of a target moving on a uniform background to yield nonzero values for $\vec{Z}h$. Thus the operator can function as a target spotter even when $\vec{Z}h$ is zeroed where Eq. 11 is not invertible.

b. Telescopic Limit $\left[h(x,y,t) = \frac{1}{a^2} f(ax+b, ay+c)\right]$

In a manner analogous to that of the microscopic limit, one can show that

$$h_t = A(xh_x + yh_y - 2h) + Bh_x + Ch_y, \tag{14}$$

where $A = \dot{a}/a$, $B = \dot{b}/a$, and $C = \dot{c}/a$. Operating on this equation with two independent spatial linear operators allows one to solve for A,B,C for each picture element, as before. The proof of the uniqueness of the form $\frac{1}{a^2} f(ax+b, ay+c)$ as a solution to Eq. 14 remains to be completed.

We thank Drs. H. Resnikoff (National Science Foundation), P. Burt and A. Rosenfeld (University of Maryland Computer Vision Laboratory), and E. Barrett (JAYCOR, Alexandria, VA) for helpful discussions.

Footnotes and References

*If the retinal plane is one unit of distance from the projection center, h is the light-energy distribution on the retina, multiplied by $(x^2 + y^2 + 1)^{3/2}$ to correct for the anticipated distance and cosine errors if the viewed scene were a matte plane parallel to the retina.

1. W. Pitts and W.S. McCulloch, "How We Know Universals: The Perception of Auditory and Visual Forms," in W.S. McCulloch, Embodiments of Mind (The M.I.T. Press, Cambridge, Mass., 1965), pp. 46-66.

2. T.G. Newman and D.A. Demus, "Lie Theoretic Methods in Video Tracking," Workshop on Image Trackers and Autonomous Acquisition Applications for Missile Guidance, 19-20 November 1979, U.S. Army Missile Command, Redstone Arsenal, AL.

4. APPARENT REFERENCE-FRAME PARADOX IN GENERAL RELATIVITY

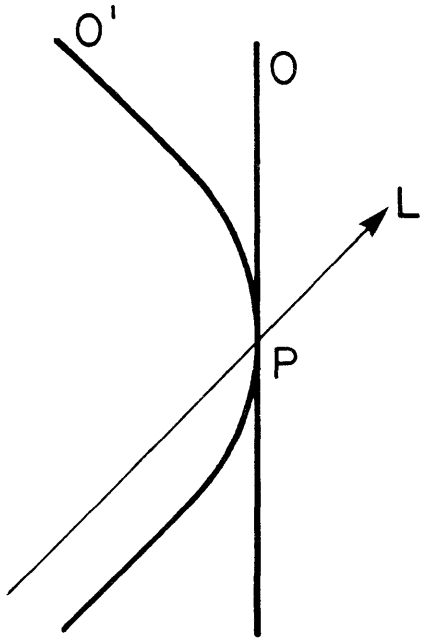
Bell Laboratories (Grant)

Michael H. Brill

Computations connecting General Relativity with its three fundamental tests — perihelion advance of Mercury, light bending, and red shift, all in the sun's gravity field — make no note of the noninertial status of the earthbound reference frame.[1-3] How would the measurements change if the observer were stationary and freely falling at the surface of the earth? When we examined the problem of transforming between measurements of inertial and noninertial observers in a uniform gravity field (a case much simpler than the true gravitational field of the earth), the following apparent paradox arose:

The strong Principle of Equivalence in General Relativity states that observer accelerations are locally indistinguishable from gravity fields. (This is to be contrasted with a less stringent interpretation — the weak Principle of Equivalence — which asserts the equivalence of inertial and gravitational mass.) If we adopt the interpretation of Ohanian[4] (and others) that global departures from strong equivalence are due to gravity gradient effects (tidal forces), then uniform gravity fields should display a global strong equivalence, since a uniform field has zero gradient.

On the other hand, reference frames in General Relativity are commonly interpreted as local Lorentz coordinate systems.[5] Suppose in flat space-time the world lines of inertial observer O and accelerating observer O' are tangent to each other at point P, and that both observers see a light at P (see diagram). If O and O' adopt the same local Lorentz frame at P, they should record the same measurements of L at P. But the Principle of Equivalence says O' should find his acceleration the same as a uniform gravity field, through which light L should appear frequency-shifted relative to the measurement made by O. In view of these contradictory arguments, does O make the same frequency measurement as O' on light L?



It would seem that if O and O' make identical measurements at P, then this would forbid interpretations of the Principle of Equivalence stronger than the equivalence of inertial and gravitational mass. If O and O' make different measurements at P, then it would seem that reference-frame transformations are not the same as coordinate transformations, no matter how locally these transformations are constrained.

One aspect of the problem not heretofore discussed is the motion of the light source as seen by O and O'. When O' sees L at P, he also construes that the light source was moving with respect to him at the time of emission of L from the source. It is possible that the Doppler shift of the light as seen by O' exactly cancels the gravitational frequency shift demanded by the strong Principle of Equivalence, and thus renders identical the measurements of L by O and O'. We are now addressing this possible resolution of the paradox, and are discovering that an important unresolved consideration is the coordinate system in which the "equivalent" gravity field is expressed.

References

1. M.H. Brill, "Perception and the Observer in Relativity Theory," M.I.T. RLE Progress Report No. 121, January 1979, pp. 150-153.

2. M.H. Brill, "Transformations between Observer Reference Frames in General Relativity," 1979, unpublished.

3. M.H. Brill, "Observer Reference Frames in General Relativity," AAAS General Meeting (poster session), San Francisco, January 3-8, 1980, Abstract volume, p. 135.

4. H.C. Ohanian, "What Is the Principle of Equivalence," Am. J. Phys. 45, 903-909 (1977).

5. C. Misner, K. Thorne, and J.A. Wheeler, Gravitation and Cosmology (Wiley, New York, 1972), Chap. 6.

5. COMPUTER-SIMULATED OBJECT-COLOR RECOGNIZER

Bell Laboratories (Grant)

Michael H. Brill

If a spatially diffuse illuminant with spectral power distribution $I(\lambda)$ is incident on a set of coplanar matte reflectors with spectral reflectances $r_i(\lambda)$, a trichromatic visual system (artificial or natural) will represent the reflected lights as tristimulus values

$$Q_{ij} = \int I(\lambda)\, r_i(\lambda)\, q_j(\lambda)\, d\lambda, \qquad (1)$$

where $q_j(\lambda)$ $(j=1,2,3)$ are the spectral sensitivities of the eye.

If the visual system is an object-color recognizer, it must compute functions of Q_{ij} that are invariant with respect to spectral changes in $I(\lambda)$. What is computed will be related to object spectral reflectances and not to the illuminant spectral power distribution (SPD).

More particularly, it is desired to find a function of Q_{ij} that is invariant with respect to interchange of illuminants in a spectral equivalence class (characteristic of naturally occurring lights). Implicit in this treatment is a restriction on the allowed reflectance spectra, and possibly also on the spectral sensitivities of the photosensor.

In previous work,[1] we noted that tristimulus volume ratios are illuminant-invariant if the spectral reflectances are linear combinations of three particular functions of wavelength. In this case the equivalence class of illuminant SPDs is almost unrestricted. It is required only that noncoplanar object-color tristimulus vectors remain noncoplanar when the light is changed. We note that the three reflectance-spectrum basis functions need not be "known" by the recognizer, because tristimulus volume ratios will in any case be illuminant-invariant.

On the basis of this principle, we have designed (in computer simulation) an object-color recognizer,[2] and compared its performance with that of a recognizer computing scaled-integrated reflectances as in the work of McCann et al.[3]

The general recognizer framework is shown schematically in Fig. XXVIII-1. The sensor turns reflected light spectra $I(\lambda)\ r_i(\lambda)$ into tristimulus vectors $\underline{Q}_i$ via sensors $q_j(\lambda)$. (Spectra for lights $I(\lambda)$ and pigments $r_i(\lambda)$ are tabulated and multiplied in our computer simulation.) By using the reflected-light tristimulus vectors from a set of known reflecting objects (known in the sense of being color-named ahead of time), a transformation is derived for removing contingency on the illuminant spectrum $I(\lambda)$. The three-vectors $\underline{\alpha}_i$ emerging from this transformation we call object-color three-vectors. They are compared (via a Euclidean distance measure) with a set of stored template three-vectors $\underline{\alpha}_i^{(0)}$ corresponding to the same reflectors under a calibration light $I_0(\lambda)$. All the $\underline{\alpha}_i^{(0)}$ are associated with distinct color names. The template closest to any given object-color vector is assigned as the putatively identified color, and is given the color name of that template. Identification scores are computed as the number of correct pairings of $\underline{\alpha}_k$ with $\underline{\alpha}_i^{(0)}$, divided by the total number of pairings.

The transformation used for comparison with the volumetric method was a streamlined version of the Retinex theory[3]: In a particular tristimulus basis, the illuminant-invariant quantities are taken to be "scaled integrated reflectances"

$$\alpha_{ij} = \frac{Q_{ij}}{Q_{j\,max}} = \frac{\int I(\lambda)\ r_i(\lambda)\ q_j(\lambda)\ d\lambda}{\int I(\lambda)\ r_{max}(\lambda)\ q_j(\lambda)\ d\lambda}\ , \qquad (2)$$

where $r_{max}(\lambda)$ is the reflectance spectrum for a white reflector. The α_{ij} are the components of the object-color vectors. The templates are similarly constructed with

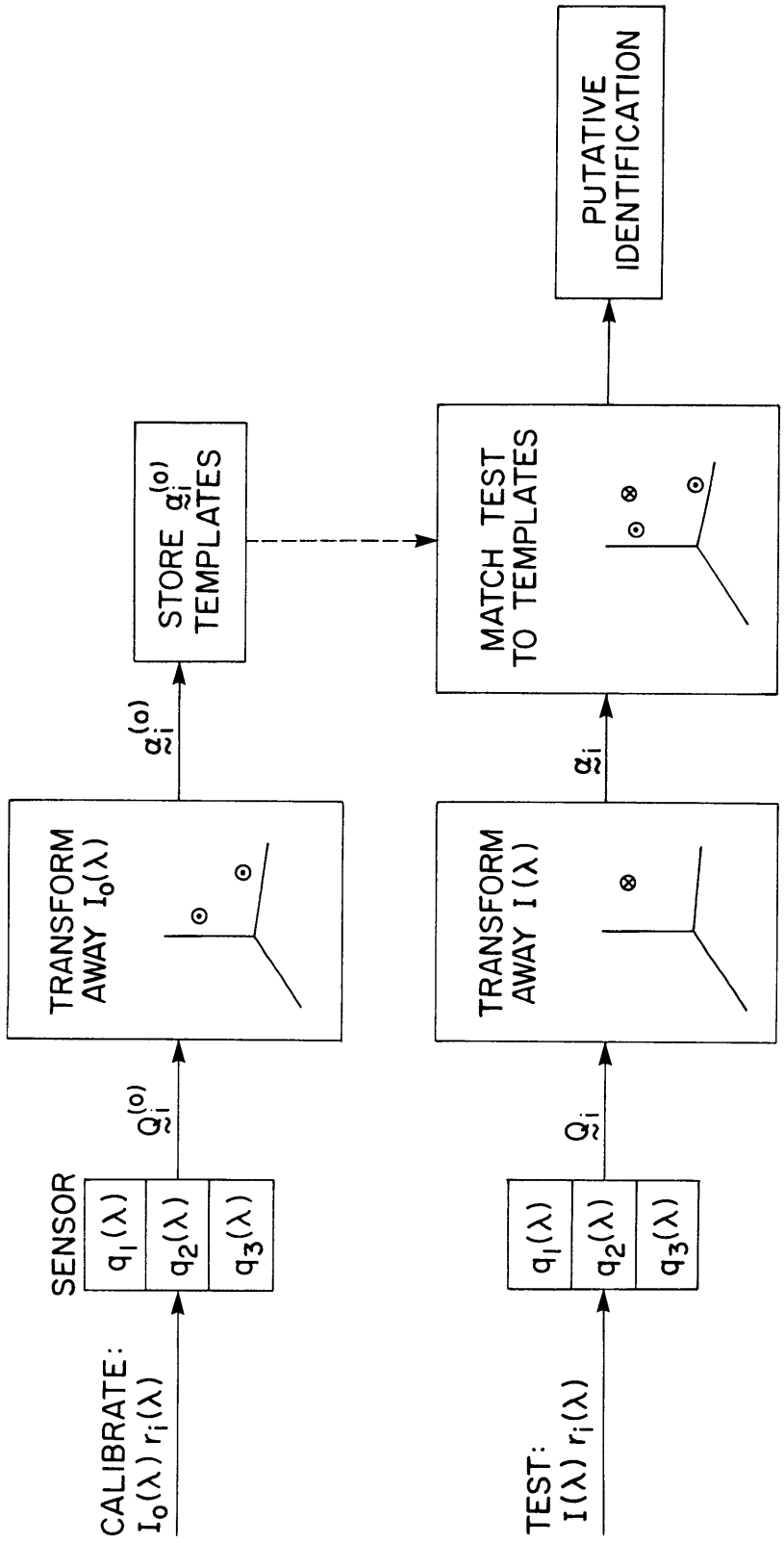


Fig. XXVIII-1. General recognizer.

$$\alpha_{ij}^{(0)} = Q_{ij}^{(0)}/Q_{j\,max}^{(0)}. \tag{3}$$

The assurance that the reference white is in the visual field and its assumed prior identification are what render our version of the Retinex theory "streamlined." Further processing to determine the identity of this reflector or a suitable substitute if it is absent from the visual field can only compromise color constancy; the "streamlined" theory has a maximum of color constancy, even though it may not agree as well as the complete theory with a more general interpretation that involves psychophysical data. Also, we note the following procedural differences from McCann et al.[3]: McCann et al. did actual spectrophotometry, whereas we obtained tabular illuminant and reflectance spectra (at 16 wavelengths) and multiplied them together to give reflected-light spectra; also, McCann et al. used Brown and Wald's[4] spectrophotometric experiments on the human eye to generate $q_j(\lambda)$, but we used Pearson and Yule's[5] choice of basis for C.I.E. tristimulus curves that are nonnegative and have minimum overlap. We did this in order to preserve metamers dictated by the C.I.E. system. (As will be seen, choice of linear combination of C.I.E. curves is immaterial to the volumetric algorithm, since volume ratios are linear-group invariants.)

Whereas in the above method a single reflector (white) was used as a reference, the volumetric method requires <u>three</u> reference reflectances. The algorithm based on the volumetric theory can be derived from the assumption that

$$r_i(\lambda) = \sum_{k=1}^{3} \alpha_{ik} u_k(\lambda), \tag{4}$$

where $u_k(\lambda)$ are the reflectances of the three reference objects. (This assumption agrees with Cohen's conclusion[6] based on a principal-components analysis on over 400 Munsell pigments.) Then

$$\int I(\lambda)\, r_i(\lambda)\, q_j(\lambda)\, d\lambda = \sum_{k=1}^{3} \alpha_{ik} \int I(\lambda)\, u_k(\lambda)\, q_j(\lambda)\, d\lambda \tag{5}$$

whence

$$Q_{ij} = \sum_{k=1}^{3} \alpha_{ik} P_{kj} \tag{6}$$

(where P_{kj} is the j^{th} tristimulus value of the k^{th} reference color). Given one tristimulus row-vector $\underline{Q}_i$ and three reference tristimulus row-vectors making square matrix [P], it follows that

$$\underline{\alpha}_i = \underline{Q}_i[P]^{-1} \tag{7}$$

is illuminant-invariant.

Each component α_{ik} is a tristimulus volume ratio. To see this, use Cramer's rule to solve $\underline{Q}_i = \underline{\alpha}_i[P]$ for $\underline{\alpha}_i$, and note that α_{ik} is a ratio of determinants (volumes):

$$\alpha_{i1} = \frac{V_{0i23}}{V_{0123}}, \qquad \alpha_{i2} = \frac{V_{01i3}}{V_{0123}}, \qquad \alpha_{i3} = \frac{V_{012i}}{V_{0123}}, \tag{8}$$

where, e.g., V_{0123} is the directed volume of the parallelepiped generated by the origin of tristimulus space and the reflected-light tristimulus points generated by the reference objects 1,2,3.

In the general schematic of Fig. XXVIII-1, the i^{th} object-color three-vector is given by $\underline{\alpha}_i = \underline{Q}_i[P]^{-1}$ and the i^{th} template is given by $\underline{\alpha}_i^{(0)} = \underline{Q}_i^{(0)}[P^{(0)}]^{-1}$. The performance of the volumetric recognizer depends on how illuminant-invariant the inferred $\underline{\alpha}_i$ actually are for real spectra.

To reiterate the principle of the recognizer, it is presumed that, ahead of time, all the objects in the universe of discourse are examined (with prior knowledge of their color names) under a light $I_0(\lambda)$. The true names of several reference colors guaranteed to appear in the visual field (three in the case of volumetric theory and one in the case of the comparison theory) are given together with their tristimulus values under a test light. Then the other object colors under this new light are putatively named by matching their object-color vectors with the templates.

The data base for the present recognizer consisted of the following.

There were 37 illuminant spectra:

a. 10 black-body radiators with T = 2000-10,000°K in 1000°K increments, and also T = ∞°K.

b. 5 Abbot-Gibson daylights, with correlated-color temperatures 4800, 5500, 6500, 7500, and 10,000°K.

c. C and A illuminants, and high-pressure sodium and mercury lights.

d. 18 fluorescent lights (spectra furnished by Dr. W. Thornton of Westinghouse Corporation).

The calibration-light spectrum was equal-energy at all wavelengths.

We used 45 artists' pigment spectra from the spectrophotometric data of Barnes.[7] All of Barnes' spectra were used except the blacks, greys, and whites, which we replaced with a spectrally nonselective 3-percent black and a 98-percent white. In all, there were 47 pigment spectra. The white was the reference color in the scaled-integrated reflectance algorithm, and for the volumetric theory we used French ultramarine blue, emerald green, and cadmium red. (These were observed to be qualitatively different from each other, but were not chosen by any quantitative means.)

The results of the simulation are shown in Table XXVIII-1. The fractions are the total number of recognition errors (incorrect template matches) divided by the

Table XXVIII-1. Error-rate results.

	Volumetric Recognizer	Scaled-Integrated Reflectance Recognizer
All lights	209/1628 = 12.2%	353/1702 = 20.7%
Nonfluorescent	57/836 = 6.8%	103/874 = 11.8%
Fluorescent	152/792 = 17.9%	249/828 = 30.1%
Sodium	14 errors	26 errors
Gold fluorescent	42 errors	36 errors

total number of recognitions. For the volumetric algorithm, the three reference object-color templates are preserved but the corresponding object-color vectors are removed from the score computation (since they are recognized perfectly by

definition). Thus the total number of putative recognitions is given by 44 object colors times 37 illuminants. Similarly, in the comparison algorithm, the reference white is eliminated as a test vector but not as a template for possible confusion. The total number of recognitions in that case is given by 46 pigments times 37 illuminants. Absolute numbers of errors are listed for worst-case illuminants in the fluorescent (gold fluorescent, used in photo-resist manufacturing) and nonfluorescent (high-pressure sodium) classes. This is done to indicate that a large number of errors came from a few illuminants that are clearly not conducive to color constancy via any theory. As expected, fluorescent lights generally gave less color constancy than nonfluorescents. Although scaled-integrated reflectance theory gives substantial color constancy as revealed by the object-color recognizer, we concluded from the results of Table XXVIII-1 that volumetric theory is more effective in producing color constancy (illuminant-invariant assessment of chromatic relations).

After presenting this work to the Optical Society of America and at SRI International, we were prompted by a question from Dr. D.H. Kelly (SRI) to consider the conditions under which the volumetric algorithm would recognize object colors seen in a television reproduction. It turns out that so long as the system has a gamma of one and the taking sensitivities of the TV camera are linear combinations of the spectral sensitivities of the recognizer (Maxwell-Ives criterion), the object-color recognizer automatically corrects for the transformation of the image between real life and TV. Equivalently (subject to the above restrictions), the algorithm corrects for transformations of the image due to photographing it through three other filters. This can be seen as follows:

Let the system taking sensitivities be

$$P_j(\lambda) = \sum_{k=1}^{3} b_{jk} q_k(\lambda). \tag{9}$$

Corresponding to the i^{th} reflector, the color separations have transmittances

$$T_{ij} = \sum_{k=1}^{3} b_{jk} \int I(\lambda)\, r_i(\lambda)\, q_k(\lambda)\, d\lambda = \sum_{k=1}^{3} b_{jk} Q_{ik}, \tag{10}$$

where Q_{ik} are the tristimulus values of the i^{th} object when it is seen directly

by the recognizer. The transmittances T_{ij} of the color separations modulate (in a spectrally nonselective way) the spectral power distributions of three light sources (or phosphors) $S_j(\lambda)$. The reprojected lights $T_{ij}S_j(\lambda)$ add together on a white screen, and are reflected back to the "eye". The tristimulus values due to the i^{th} object represented in the processed image are

$$Q_{i\ell} = \sum_{j=1}^{3} T_{ij} \int S_j(\lambda)\, q_\ell(\lambda)\, d\lambda = \sum_{j=1}^{3} \sum_{k=1}^{3} b_{jk} Q_{ik} \int S_j(\lambda)\, q_\ell(\lambda)\, d\lambda$$

$$= \sum_{k=1}^{3} \left[\sum_{j=1}^{3} b_{jk} \int S_j(\lambda)\, q_\ell(\lambda)\, d\lambda \right] Q_{ik} \equiv \sum_{k=1}^{3} C_{\ell k} Q_{ik}. \qquad (11)$$

This is an invertible linear transformation on the original tristimulus vectors, which — as has been demonstrated — is automatically compensated by the volumetric algorithm.

These findings will be the subject of two forthcoming papers.

References

1. M.H. Brill, "Further Features of the Illuminant-Invariant Trichromatic Photosensor," J. Theor. Biol. 78, 305-308 (1979).
2. M.H. Brill, "Computer Simulation of Object-Color Recognizers," J. Opt. Soc. Am. 69, 1405 (1979).
3. J.J. McCann, S.P. McKee, and T.H. Taylor, "Quantitative Studies in Retinex Theory," Vision Res. 16, 445-458 (1976).
4. P.K. Brown and G. Wald, "Visual Pigments in Single Rods and Cones in the Human Retina," Science 144, 45-52 (1964).
5. M.L. Pearson and J.A.C. Yule, "Transformations of Color Mixture Functions without Negative Portions," J. Color and Appearance 2, 30-36 (1973).
6. J. Cohen, "Dependency of the Spectral Reflectance Curves of the Munsell Color Chips," Psychon. Sci. 1, 369-370 (1964).
7. N.F. Barnes, "Color Characteristics of Artists' Pigments," J. Opt. Soc. Am. 29, 208-214 (1939).